The Economic Value of Social and Emotional Learning

Technical Report *in* Journal of Benefit-Cost Analysis · March 2015

CITATIONS
38

READS
7,325

6 authors, including:

Clive Belfield
City University of New York - Queens College
150 PUBLICATIONS **4,705** CITATIONS
SEE PROFILE

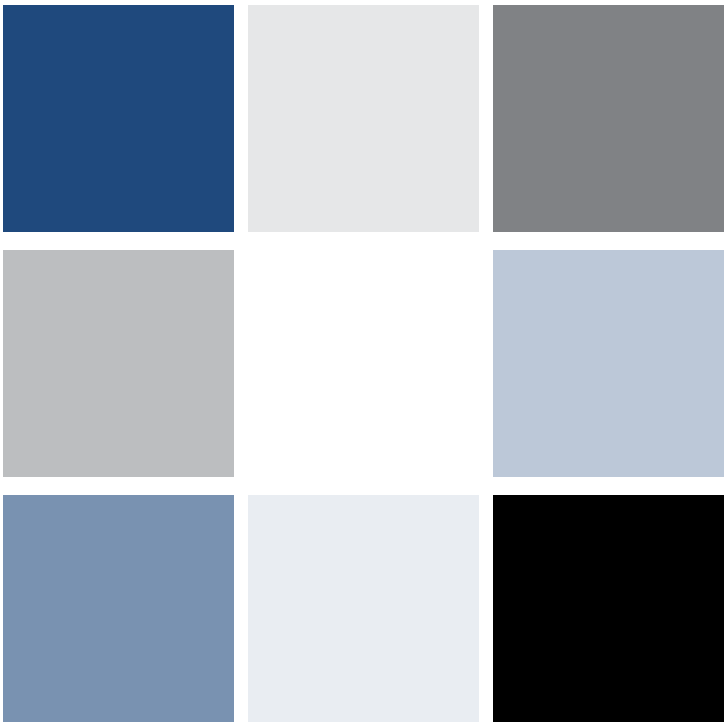
Alli Klapp
University of Gothenburg
26 PUBLICATIONS **440** CITATIONS
SEE PROFILE

Henry M. Levin
Columbia University
283 PUBLICATIONS **8,496** CITATIONS
SEE PROFILE

Sabine Zander
Columbia University
4 PUBLICATIONS **230** CITATIONS
SEE PROFILE

THE ECONOMIC VALUE OF SOCIAL AND EMOTIONAL LEARNING

February 2015 (Revised)



Clive Belfield

Brooks Bowden

Alli Klapp

Henry Levin

Robert Shand

Sabine Zander

Center for Benefit-Cost Studies in Education

Teachers College, Columbia University

www.cbcse.org

Funding for this research was provided by the NoVo Foundation (novofoundation.org).

The Economic Value of Social and Emotional Learning

Clive Belfield, Brooks Bowden, Alli Klapp, Henry Levin, Robert Shand and Sabine Zander

Center for Benefit-Cost Studies in Education
Teachers College, Columbia University
www.cbcse.org

February 2015 (REVISED VERSION)

Funding for this research was provided by the NoVo Foundation (novofoundation.org).

Acknowledgements: The authors greatly appreciate comments and support from Steve Aos, John Bridgeland, Mark Greenberg, Damon Jones, Stephanie Lee, Annie Pennucci, and Roger Weissberg.

SUMMARY

Benefit-cost analysis is a tool for evaluating the economic profitability of an investment. It has been used in education since the 1960's to determine the rate of return on both individual and social investments in education. Essentially, benefit-cost analysis compares the monetary cost of an investment with the monetary value of its outcomes. For example, by reducing high school dropouts there are costs to the student in foregone income by staying out of the labor market. But there are also gains to the student in terms of higher income, better health, and lower likelihood of involvement in the criminal justice system, all which can be measured, in terms of a monetary return on investment. The taxpayer also makes an investment in education through paying a considerable portion of its direct cost and gets a return through higher tax revenues and lower costs of public services for health, public assistance, and criminal justice. And, society obtains returns by using its resources in its most productive ways, at least partially reflected in economic returns.

It is not necessarily the case that the benefits of an educational investment always justify the costs. For some interventions, the benefits will exceed the costs. But in others the investment is not found to be compensated by its returns: even if an intervention is effective, it may be too expensive to implement. Thus, the purpose of a benefit-cost (BC) study is to provide an accurate measure of costs and benefits to determine not only if the benefits are greater than the costs of the investment, but by how much.

In the educational setting, we seek investments that have the highest return to the taxpayer and to society. In the past, BC studies have been limited largely to increases in educational attainment and to improvements in cognitive test scores. But it is now becoming widely recognized that social and emotional learning in schools can be as important as or even more important than cognitive gains in explaining important developmental and life outcomes (Durlak et al. 2011; Heckman and Kautz 2012; Levin 2012). Social and emotional skills are less commonly considered in educational evaluations, in part because they are more challenging to measure than attainment and test scores. As such skills have gained prominence, it is important to integrate them into BC studies for consideration in educational policy and decision-making.

Because of the long history of the Center for Benefit-Cost Studies in Education (CBCSE) at Teachers College and its predecessor organizations in publishing BC studies in education (e.g. Levin 1975, Levin and McEwan 2001, and other publications at www.cbcse.org), it was requested by the NoVo Foundation and the Collaborative for Academic, Social, and Emotional Learning (CASEL) to apply BC analyses to studies of social and emotional learning. The purpose was to both demonstrate the BC method in this domain as well as to gain early perspectives on the potential economic returns to investments in social and emotional learning.

In this Report, we review the available evidence on the economic value of Social and Emotional Learning (SEL). We utilize a formal method to perform economic valuations with respect to changes

in Social and Emotional (SE) skills to guide future evaluations of reforms that target SEL. Our main contribution is to demonstrate BC analysis using recent impact evaluations of six prominent SE intervention:

- 4Rs;
- Positive Action;
- Life Skills Training;
- Second Step;
- Responsive Classroom; and
- Social and Emotional Training (Sweden).

Table S1

Social and Emotional Learning Interventions

Intervention	Grades and Student Groups
4Rs Learning and literacy program to combat aggression/violence	Grades K–5 Disadvantaged
Positive Action School curriculum/activities to promote positive thinking, actions, and self-concept	Grades 3–8 All
Life Skills Training Classroom intervention to reduce substance abuse/violence	Grades 6–12 At-risk students
Second Step Social skills curriculum to improve problem-solving/emotional management	Grades PK–10 Disadvantaged
Responsive Classroom Improve teacher efficacy to influence SE skills and school community	Grades 3–5 All
Social and Emotional Training (Sweden) Classroom intervention to support cognitive and SE competencies	Grades 1–9 All

Table S1 provides the descriptions of the six interventions.

These interventions were chosen because they are prominent in the literature and provide diversity in terms of their goals, measures of outcomes, and student populations. For each intervention we have constructed tables of ingredients and their costs; alongside, we have created benefit maps to summarize the possible benefits each intervention might confer and calculated the monetary value of the portion of benefits that could be identified and quantified based on the results reported in the impact evaluation. We then computed appropriate economic metrics – benefit-cost ratios and net present values – and performed sensitivity testing to see if the results are robust to alternative specifications.

Our central question is whether a range of different SEL interventions, both individually and in the aggregate, show benefits that exceed their costs. We recognize the fact that we have captured only a portion of their benefits because not all of their effects are fully measured and can be readily converted into monetary measures of their benefits. Further, the portion of benefits that we have been

able to capture may differ considerably among the different interventions' depending upon the goals of each and measurability of the associated benefits. W*e caution that it is inappropriate to compare individual benefit-cost ratios among the six interventions in the absence of a more complete accounting of benefits for each*. Thus, we focus primarily on the basic question of whether the available measures of benefits are equal to or exceed costs for each and summarize our results with an overall average benefit-cost ratio among the six interventions.

The most important empirical finding is that each of the six interventions for improving SEL shows measurable benefits that exceed its costs, often by considerable amounts. There is a positive return on investments for all of these educational reforms on social and emotional learning. And the aggregate result also shows considerable benefits relative to costs, with an average benefit-cost ratio of about 11 to 1 among the six interventions. This means that, on average, for every dollar invested equally across the six SEL interventions, there is a return of eleven dollars, a substantial economic return. These findings are robust to the imposition of different assumptions on the sources and construction of benefits and costs, and a full accounting for benefits, as shown in the benefit maps, would provide an even larger return.

However, we emphasize that in addition to the benefit-cost test that we have imposed on these interventions, an important contribution of this report is its demonstration of methods and applications to estimating benefits and costs of SEL interventions. Only if a formal method is applied in a transparent way will economic evaluations of SEL interventions help improve resource allocation within the school setting.

This methodological demonstration can advance understanding of SEL in the future. Because our calculations have not captured all of the benefits of each of the interventions and the proportions that we have captured probably vary considerably among them, *we caution the reader that it is inappropriate to compare directly the results for specific interventions relative to the others*. The important findings are that each of the SEL approaches shows benefits that exceed costs and that the average return is very high, but differences in the quality and availability of comparable data preclude precise comparisons among the interventions for the following reasons:

- **Evaluation Quality**—We drew upon the best evaluations for each intervention that were available. However, even among relatively high quality evaluations there was variation. Had the evaluation designs and measurements been of consistent quality, there might have been some differences in results. Both data quality and research methods account for some potential differences in our results.
- **Cost Estimation**—Our cost method is the well-established ingredients method, which requires identification of all of the resources that were required to produce the results that were found in the evaluations. Since all evaluations had been done in the past, it was necessary to reconstruct the interventions as they had been implemented at the time of their reported effectiveness results. Precise details of the interventions were rarely available, and attempts to contact informed observers were not always successful. Even when successful, relying on retrospection of details can often provide incomplete or misleading results. To as great an extent as possible we have attempted to use multiple sources to construct the cost estimates, but differences in the information base attainable for each of the interventions varied considerably in comprehensiveness and quality.
- **Multiple Benefits**—Most SEL interventions have multiple goals and benefits, and this contrasts with interventions to improve cognitive test results in a particular subject. For example, an attempt to reduce aggression may also improve impulse control and later reduce juvenile crime and/or may

raise academic achievement. We have tried to capture some of this heterogeneity in our benefit maps, but even these are dependent upon an incomplete knowledge-base. And as the maps reveal, only a portion of the possible benefits are measured in the specific evaluation studies. Consequently, our measures of benefits are based upon a limited set of dimensions, a source of downward bias in our benefits measure. The actual benefits may be considerably higher if we were able to identify all effects and convert them into monetary benefits.

- **Benefits Measures**—There is a major challenge in taking outcomes such as reductions in substance abuse or aggression or other improvements in behavior and attitudes and converting them into monetary values to society. We have used shadow prices, which reflect the amount that society is willing to pay for such improvements. But shadow prices depend upon specific assumptions, constructs, and data availability, and there is likely to be some underlying variance, largely unmeasured, in the magnitude of potential shadow price estimates. We believe that the shadow prices used to estimate benefits in this report are both plausible and defensible, but they are still subject to variability.

Taken in conjunction, the two contributions of this Report suggest a promising agenda for future research on the economics of social and emotional learning. From one perspective, it is feasible to apply benefit-cost analysis to SEL interventions, and these interventions do offer high economic returns as educational investments. Overall, SEL interventions are likely to pass a benefit-cost test. From another perspective, there is considerable additional research to be performed to establish the full extent and magnitude of the benefits of SEL. The full economic value of SEL is not yet established. Our hope is, therefore, that this investment stimulates the generation of a refined knowledge base and a greater focus on the development of benefit-cost applications to SEL initiatives.

CONTENTS

Summary 3

1. Introduction 8

2. Evidence on the Economics of SEL 10
- 2.1 Program-Based Evidence 10
- 2.2 Earnings and SE Skills 11
- 2.3 Educational Achievement and SE Skills 13

3. Foundations for Economic Evaluation of SEL 15
- 3.1 Economic Framework 15
- 3.2 Ingredients Method for Costs 16
- 3.3 Shadow Price Methods for Benefits 18

4. SEL Interventions 21
- 4.1 4Rs 21
- 4.2 Positive Action 25
- 4.3 Life Skills Training 30
- 4.4 Second Step 34
- 4.5 Responsive Classroom 37
- 4.6 Social and Emotional Training 42

5. Developing Methods and Evidence for Economic Evaluations of SEL 46

References 49

Appendices

Appendix I Benefits Maps 56

Appendix II 62

1. INTRODUCTION

Social and emotional skills are increasingly being recognized as important for child development (Weissberg et al., 2003; Zins et al., 2004). These skills, which include competencies such as self-awareness, self-management, social awareness, relationship skills, and responsible decision-making, may help students progress further in their education and may also enhance personal, economic and social well-being in youth and adulthood. In recent reviews of the evidence on interventions to develop social and emotional skills, Durlak et al. (2011, Table 2) and Sklad et al. (2012) identify durable and substantively important gains in social and emotional skills, attitudes, positive social behavior, conduct problems, emotional distress and academic performance. These gains should lead to substantial increases in personal and social wellbeing.

According to Durlak et al. (2011), social and emotional learning combines youth development with the promotion of particular competencies, with the aim of enabling students to respond appropriately to environmental demands and fully take advantage of opportunities. Ultimately, social and emotional competencies encourage a shift to an internal locus of control, allowing individuals' choices and actions to better accord with their own values. These competencies relate to "soft skills" and personality traits that, according to Heckman and Kautz (2012), predict success in school, the labor market, and in life. Notably, social and emotional competencies do not just raise academic achievement and educational attainment. They also foster personal satisfaction and growth, help individuals become better citizens, and reduce risky behaviors like violence and drug use (Durlak, Weissberg and Pachan, 2010; Collaborative for Academic, Social, and Emotional Learning, 2013).

However, although the positive development of social and emotional skills through education has been established, the economic value of gains in these skills has received little attention. But, within a resource-constrained education system, instructional activities in social and emotional skills must compete with instruction directed toward test score improvements and increases in educational attainment. To date, most education research has focused on these last two objectives, with the presumption that this will have important economic consequences. For attainment, this focus makes sense: the evidence on the high economic value of graduating from high school or completing college is substantial. For achievement, however, the justification is much less certain: the impact of cognitive gains on economic performance is extremely modest; and most interventions fail to generate long-term boosts in cognition (Levin, 2012). Moreover, whatever influences are driving the high returns to attainment, they are only partially mediated through improvements in cognitive functioning (Heckman and Kautz, 2012). A suggestive and plausible alternative is that much of the effect of education on economic outcomes and personal well-being is based on non-cognitive or social and emotional development such as effort, motivation, curiosity, empathy, caring, and the many other dimensions that can be encompassed by social and emotional skills. These skills may drive attainment, which then becomes the focus for educational investments. On balance, it is likely that both social and emotional learning and cognitive learning are efficient investments. But the former domain has received very little attention from economists, and the policy debate instead emphasizes cognitive gains and test score accountability. In this environment, it is critical to identify and enumerate the value of social and emotional skills.

In this Report, we demonstrate an approach to conducting economic evaluations of six school-based social and emotional learning interventions and place these evaluations in the broader context of the economics of social and emotional learning. These six interventions are selected because each

has been evaluated using a generally, acceptable, research method in terms of impacts on a range of social and emotional (SE hereinafter) skills. We estimate each intervention's costs based on the ingredients employed during the implementation previously evaluated. We utilize the effects estimated in the evaluations to estimate economic benefits of the interventions to society. We then calculate the benefit-cost ratios and net present values to determine if the benefits generated by each program outweigh the costs of implementation. The intention here is not to rank these interventions in terms of their efficiency for reasons of differential completeness of the underlying studies in identifying effects and measurement challenges in identifying benefits. Instead, it is to provide a demonstration of how benefit-cost analysis of social and emotional interventions should be undertaken and to see if measured benefits exceed costs, understanding that a full assessment of benefits of each intervention would yield even higher benefits. Thus, we pay considerable attention to methodological challenges in estimating costs and calculating benefits; in turn, these challenges highlight broader research issues in evaluating social and emotional interventions.

Our Report is structured as follows. We begin with a review the evidence on the development of SE skills and in particular how these skills influence future earnings and are mediated through changes in education levels. Next, we set out a framework that includes both the economic principles and the methodological approaches for estimating the value of SE skills. We then apply this framework to six SE interventions to calculate their costs, benefits, and net benefits. Finally, in light of these six applications, we illustrate important conceptual, empirical, and methodological issues for research on the economic value of SEL. Overall, by providing the foundations for economic evaluations of social and emotional skills, this research should enhance their presence in future research and policy discussions.

2. EVIDENCE ON THE ECONOMICS OF SEL

2.1 Program-Based Evidence

There is substantial evidence on positive impacts of SEL interventions (SCDRC, 2010; Durlak et al., 2011; and Sklad et al., 2012). These impacts include large gains in SE skills, attitudes, positive social behavior, and academic performance, as well as reduced conduct problems of student disruption and emotional distress. These gains, measured across hundreds of interventions, are substantively large (with effect size gains of 0.2-0.6); and, when based on teacher reports, are consistently statistically significant (Durlak et al., 2011, Table 2). In a subsequent review, Sklad et al. (2012) emphasize gains in social-emotional skills and positive self-image. Concomitant with these gains were other effects in terms of social behavior, substance abuse, mental health disorders, and achievement.

SEL interventions vary significantly in application and mission: some are integrated into regular classroom instruction, others are directed at a specific adverse behavior (e.g. substance abuse), and others are intended to have a general effect (e.g. on school climate) rather than individual students. SEL interventions can be targeted to specific grades or be applied across the full span of grades. Variations in the practices for each intervention with respect to lessons, classroom strategies, and professional development are given in detail in SCDRC (2010, Table 1.1, Panels 1-3) and Sklad et al. (2012, Tables 3 and 4). Hence, although we refer to SEL interventions as a group, we recognize there is considerably diversity.

Some benefit-cost analyses of SEL interventions have been performed.[1] Applied to the context in Washington state, Lee et al. (2012) have calculated costs, benefits, and net benefits for three well-known SEL interventions: Promoting Alternative Thinking Strategies program (PATHS); the Seattle Social Development Project (SSDP); and Life Skills Training (LST).[2] For the SSDP, the participant costs are $3,030 and the benefits are $5,800; the net benefits are strongly positive at $2,770. For LST, the costs are $30 and the benefits are $1,290; the net benefits are therefore $1,260. At the national level, Jones et al. (2008) estimate net benefits of $810 per student in LST. Looking at how SEL interventions ameliorate substance abuse by youth, Miller and Hendrie (2008) perform benefit-cost analysis for three interventions. For, LST, they estimate costs of $290, benefits of $5,960, and therefore net benefits of $5,670. For SSDP, they estimate costs of $3,200, benefits of $19,000, and therefore net benefits of $15,800. Finally, for the Social Competence Program, Miller and Hendrie (2008) estimate costs of $350, benefits of $2,500, and therefore net benefits of $2,150. Given the approaches used, these benefit-cost analyses show that SEL interventions can yield positive, sometimes substantial, economic returns. However, since the studies were done independently, they use different approaches to both cost and benefit measurement. Therefore, they are not directly comparable. Across the broader literature on youth behavior, there are many benefit-cost analyses of programs to combat delinquency.[3] Many of these programs are intended to modify some SE skills (such as student conduct, educational progress, and externalizing behaviors). But their main focus is on a narrow class of behaviors, such as teenage pregnancy or juvenile crime; few are intended to affect a broad array of outcomes simultaneously.

1 Throughout, we report all money values in 2013 dollars and, where appropriate, express amounts in present values at the start of each SEL intervention. To avoid spurious precision, amounts are rounded to the nearest $10 or $100 where appropriate.

2 See http://www.wsipp.wa.gov/BenefitCost. Evidence on PATHS is currently under review.

3 See Weimer and Vining (2009) and the compendia of evidence at: www.colorado.edu/cspv/blueprints/; youthinfo.gov; and wsipp.wa.gov.

Nonetheless, these programs are typically found to have benefits that exceed their costs. Miller and Hendrie (2008, Table 12) estimate net benefits for 17 general substance abuse programs and find almost all have positive net benefits. In their review for Washington State of programs for children and adolescents, Lee et al. (2012) find positive net benefits for seven.

Although these analyses contribute to the debate on the economic value of SEL, the literature and evidence base are limited. More importantly, significant empirical and methodological challenges remain. First, few studies report costs in detail. Even fewer use the ingredients method, relying instead on more casual information such as budgetary data or estimates by the program developers. Second, it is not clear if these programs can be compared either to each other or to the broader class of delinquency prevention interventions. There are significant programmatic differences with respect to the students targeted, the size of the program, and the outcomes intended for improvement (SCDRC, 2010, Table 1.1, Panel 6). There are also significant differences in how the programs are evaluated. These differences include: the number of impacts that are converted into monetary benefits; the time horizon for analysis (with some studies including only immediate benefits and others only long-term benefits); and the method by which benefits are calculated (i.e. how impacts are translated into dollar amounts).[4] Unless the method is consistently applied, comparing results across existing evaluations is highly problematic.

The economic evidence on the value of SE skills therefore needs to be expanded to include more interventions and particularly those that affect SE skills directly (rather than youth conduct). The research also needs to be undertaken in a standardized way, such that findings from separate benefit-cost analyses are comparable and so can be generalized. We discuss these issues in more detail in our conclusion.

2.2 Earnings and SE Skills

One approach to establishing the value of SE skills is to look at how they directly influence labor market outcomes. There is now an expanding literature on the association between SE skills and subsequent earnings (Murnane et al., 2001; Waddell, 2006; Drago, 2011). This evidence indicates strong positive impacts on earnings when SE skills are enhanced.

However, a critical aspect of this research is how SE skills are defined and measured. These skills can be defined in terms of psychological traits, behaviors, attitudes towards others, and performance on specific tasks. In turn, these constructs can be measured in different contexts (e.g. school versus home) and using different instruments. Yet, labor market research relies almost exclusively on the Rosenberg scale of self-esteem and Rotter scale for locus of control. These scales are rarely used in evaluations of particular interventions. Thus, it is not possible to directly translate the impact of an intervention into gains in earnings.

In addition, there are methodological challenges to identifying earnings gains from social and emotional (SE) skills. One challenge is that the construct chosen to measure SE skills is typically the only one representing an individual's behavioral traits; it may be that other traits – correlated with SE

4 According to Durlak et al. (2011, 413), the average follow-up period for SEL interventions is less than two years. Thus, in most cases it is not possible to see whether there have been long-run impacts. Yet, it seems likely that some SEL effects will persist, even if the immediate effects are stronger than those over a longer-term (Sklad et al., 2012). It is possible to extrapolate from immediate impacts to future ones, although this extrapolation adds imprecision.

skills – are driving the association. Another challenge is that education itself may be positively correlated with SE measures; and education is of course a significant determinant of earnings. Both these challenges may offset any underlying association. De Araujo and Lagos (2013) examine how earnings are influenced by self-esteem (their construct for SE skill); they find that self-esteem has no influence after controlling for locus of control attributes and for education. Nevertheless, if self-esteem affects educational attainment and thence earnings, this still represents an economic benefit from SE skills. A final challenge arises when these SEL constructs are measured simultaneously with earnings: it seems plausible that someone with high earnings might report (or have acquired) high SE skills. On the last of these, de Araujo and Lagos (2013) find that high self-esteem is indeed jointly determined with wages.

Notwithstanding these challenges, research evidence consistently finds strong associations between SE skills – broadly defined – and earnings. Using NLSY79, de Araujo and Lagos (2013, p.1987) estimate that a "one standard deviation increase in [self-reported] self-esteem leads to a 30.46% increase in real wages". This increase is mostly mediated through attainment, which is estimated to be 1.5 years greater for those with higher self-esteem. If we assume baseline returns to attainment at 10% per year, the estimated gain in earnings from a one standard deviation increase in self-esteem is therefore 15%. This is still a very large gain in earnings. However, Drago (2011), also using NLSY79, finds smaller estimates of self-esteem (at 4%). Using NELS88, Segal (2013, p.767) reports a one standard deviation increase in "[teacher-reported] misbehavior is associated with a 4% decrease in earnings" at age 28. Also using NELS88, Eren and Ozbeklik (2013) employ the Rosenberg and Rotter scales in 10th grade and estimate that a one-standard deviation increase in non-cognitive ability is associated with an increase of 9% in earnings.[5]

Although promising, a precise association between SE skills and earnings remains to be determined. The economic value of these impacts of enhanced SE skills can only be generalized under two restrictive assumptions. It is necessary to assume that these constructs (self-esteem, misbehavior, self-control, etc.) accurately and fully reflect SE skills. As well, it is necessary to assume that SE skill differences are stable over childhood (such that SE skills in one grade reflect SE skills throughout the school years). These assumptions – both of which are highly debatable – represent important areas for future research.

Yet, given these two assumptions, it is possible to approximate the present value of lifetime earnings gain from a one standard deviation increase in SE skills. (We use one standard deviation for illustrative purposes; most interventions cannot yield effects this large). Based on current lifetime earnings profiles from the American Community Survey and the Current Population Survey, the average present value lifetime earnings of a child who is currently in 3rd grade would be $575,000.[6] These amounts represent how much a lifetime's earnings are worth at 3rd grade (adjusting for inflation). Given the above evidence, if SEL goes up by one standard deviation, earnings go up by 4-15% (net of additional schooling). Using the lower bound of these numbers, the average expected earnings gain from a one standard deviation increase in SE skill in 3rd grade is $46,000. A conservative estimate – using the weakest correlation between SE skills and earnings – would yield earnings gains of $23,000. In other words, if an intervention raised a child's SE skills at that age, this would be worth $23,000 in terms of gains in future productivity alone. Although not precise and subject to the many caveats listed above, these figures indicate very large long-term labor market effects from enhanced SE skills.

5 Results from Murnane et al. (2001) and Waddell (2006) are not easily translated into effect size gains.

6 Details of this calculation are available from the authors. Under these two assumptions, the present value amount can be calculated for any year group.

2.3 Educational Achievement and SE Skills

An alternative approach to estimating the economic value of SE skills is to examine how they affect education levels. From this association it should then be possible to draw on the vast amount of research on the economic returns to education to establish the benefits of SE skills. Overall, research evidence typically shows a positive association between SE skills and education. In their review of 35 SEL interventions, Durlak et al. (2011, Table 2) report an average effect size gain in achievement of 0.27.

From an economic perspective the association between SE skills and achievement is complex. Where SE skills enhance achievement, it should be possible to value SE skills by valuing the labor market gains from higher achievement. However, this approach is likely to undervalue SE skills as these almost certainly have much more diffuse and long-term behavioral and attitudinal implications than achievement does. Academic achievement gains often fade-out within a short period (see Kinsler, 2012; Jacob et al., 2008). Moreover, for some groups academic achievement *per se* may not be especially important for future lifetime outcomes and may even be inversely related to SE skills (as established in studies of high school dropouts who pass the GED, see Heckman and LaFontaine, 2006). Recent research using longitudinal data by Heckman and Kautz (2012) describes a weak association between achievement and future life outcomes; and Castex and Dechter (2013) find that cognitive ability has actually declined as a determinant of wages over the last two decades. Heckman and Kautz (2012) also argue that SE skills and achievement as measured by standardized test scores are strongly confounded: because SE skills are expressible as behavior on tasks and taking an achievement test is a task, then higher achievement scores must also reflect higher SE skills. From an economic perspective, distinguishing the two constructs is problematic and may lead to double-counting in benefit-cost analysis because of their overlap. Finally, it might be noted that a focus on the cognitive gains of SEL programs serves to reinforce the idea that schools teach only 'knowledge' and downplays the idea that schools ought to develop broader social and emotional skills.

A number of studies have – despite these caveats – evaluated SE programs in terms of how they increase achievement, and these achievement gains can be translated into present value money benefits in terms of higher earnings. (As with the earnings analyses referred to above, an implicit assumption is that achievement gains are stable and persist over grade levels). To estimate the economic gains from achievement, we adapt estimates generated by Belfield and Levin (2009, Table 4).[7] Specifically, we adjust their estimates of achievement gains – mediated through changes in attainment – for inflation and population weights. These adjustments yield a present value gain in earnings from a one standard deviation increase in 3rd [8th] grade math test scores of $34,300 [$40,700]. Applied to the specific results reported by Durlak et al. (2011), a one standard deviation increase in SE skills in 3rd [8th] grade is therefore valued at $9,000 [$11,000]. Although still substantial, these education-mediated values are considerably below the direct estimates reported above.

Thus, the association between SE skills and education should be evaluated more broadly than through individual achievement gains. Although Zins et al. (2007) make a strong case for evaluating SE skills in terms of academic outcomes, these outcomes should include school behaviors and school attitudes, as well as test scores (Zins et al., 2007, Figure 1.4). One direct way that has clear economic

7 To avoid the methodological challenges noted above, Belfield and Levin (2009) express achievement gains in terms of attainment gains (using the association between test scores and high school graduation).

consequences is to look at rates of grade retention and placement in special education. So far, this approach – even as it only refers to substantive behaviors and not attitudes – has not been extensively investigated. Moreover, it seems unlikely that the benefits of SE are confined solely to the individual student. The benefits of SE skills should be looked at using general measures at both classroom-level (such as peer effects, teacher efficacy, or teacher retention) and school-level (such as school quality or school climate). That is, in an economic sense there are externalities such that a healthier emotional and social climate induced directly by SEL interventions at the level of individual students should improve classroom harmony and school functioning. Unfortunately, to our knowledge there is no straightforward way to estimate the economic value of general school quality/climate differences.[8] It seems plausible that in the absence of SEL, schools would incur greater costs in terms of heightened security, student support, or facilities upkeep. Related evidence on 'compensating wage differentials' (in this case, the need to pay teachers more to work in more disruptive school environments) has not found large costs when students exhibit worse behavior (Goldhaber et al., 2010). At this time it is therefore possible to calculate the economic value of SE skills – as mediated through education – only in terms of individual test score gains or more productive social behaviors and not in terms of other school-wide or student-level effects.

8 Again, this task is made more difficult by the idiosyncratic instruments used by evaluators to measure school climate.

3. FOUNDATIONS FOR ECONOMIC EVALUATION OF SEL

3.1 Economic Framework

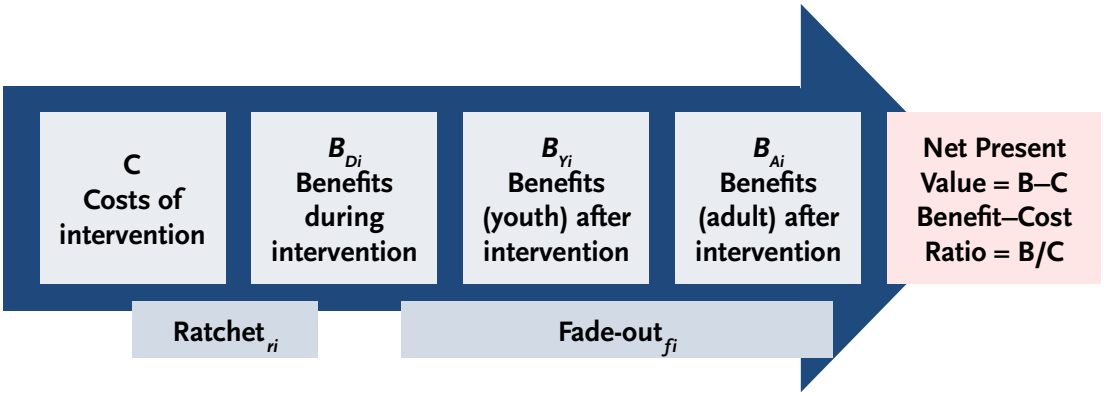
A framework for economic evaluation of SEL interventions – linking costs and benefits together in a standardized approach – is given in Figure 1.

To begin, the costs of each intervention must be estimated. These costs should be based on the opportunity cost of the resources required (Levin, 1975; Nas, 1996, p.61). They should be expressed in present values (discounted) from the start-date of the intervention to account for the fact that money expenditures that happen earlier as reflected in opportunity cost are given greater weight by society (see McEwan, 2002, p.39).

Next, the benefits of the intervention must be estimated, ideally based on causal impact evaluations. These too should be expressed as present values at the start date of the intervention: present investments in SEL engender future benefits and the further these benefits are in the future, the less valuable they are. By convention, a societal perspective is adopted whereby all benefits are calculated regardless of who reaps them. These benefits can be separated into three components. First, there are immediate benefits B_D – those occurring as the intervention is being delivered. Second, there are post-intervention benefits during youth B_Y – those benefits that occur after the intervention has been delivered but while the participants are still in school. Third, there are post-intervention benefits during adulthood B_A – those benefits that occur after the participants have left school. There may be many separate benefits in each time period and each benefit must be modeled over time. For the immediate benefits, the issue is how rapidly the benefits accrue once the intervention has begun (specified as the 'ratchet function' r_i). Some interventions may work immediately (perhaps through control of behaviors); others may work more slowly (perhaps through changing attitudes). Holding all other factors constant, immediate benefits are preferred to later ones. For the post-intervention benefits, the critical issue is their durability ('fade-out function' f_i). Some interventions may yield benefits only during implementation; others may change behaviors beyond the time frame of intervention. Critically, all benefits from an intervention should be counted: otherwise – tautologically – analyses that include more benefits will appear more beneficial.

Costs and benefits need to be reported in an equivalent way. For these analyses, all dollar amounts are reported as present values of future benefits from the year in which the intervention began with a discount

Figure 1
Cost-Benefit Framework for Social-Emotional Learning



rate of 3.5% applied to future costs and benefits. Dollar values are adjusted for inflation and reported in 2013 prices. Unless otherwise stated, the general ratchet function is assumed to be zero (i.e. the impacts of the program only occur in the year in which they are measured). Unless otherwise stated, the fade-out function is zero: the benefits are assumed to persist through school and adulthood. (Although this may seem to be a strict assumption, it is unlikely that interventions are delivered under the assumption that they will only have temporary effects. In our sensitivity testing, we vary the rate of fade-out).

Two efficiency metrics can be derived from comparing costs and benefits: the benefit–cost ratio *(=B/C)* and the net benefits *(=B-C)*.[9] Interventions with benefit-cost ratios that exceed one or for which benefits exceed the costs are considered efficient. Interventions with the highest net benefits or benefit-cost ratios are most preferable from an efficiency standpoint. Given that SEL interventions are often at substantially different scales of implementation, the net benefit measure may be more informative. The net benefit measure indicates the total amount of resource saved: interventions that generate the largest total dollar saving are preferred (even if they have a lower benefit-cost ratio). Calculation of these metrics is subject to sensitivity testing. In this analysis, a general form of sensitivity analysis is performed: salient benefits and costs that cannot be monetized are mapped, along with analysis as to how their inclusion would influence the results.

It is important to specify this framework in detail because interventions can differ in important respects. Most obviously, they can differ in how much resource is allocated to them: how much they cost and who bears that cost. Next, they can differ in the number and types of separate domains of behavior they impact. Those interventions that reduce juvenile crime and substance abuse will be more beneficial than those that impact in just one of these domains. Also, interventions can differ with respect to the durability of their impacts. Those that yield short-term gains in academic achievement are unlikely to be as beneficial as those that ensure students graduate from high school, for example. In conjunction, there are a series of context assumptions (year of evaluation, inflation adjustments) that can create variation in analyses. These differences have to be ascertained when comparing interventions.

Application of this framework – consistent measurement of costs and benefits, reporting of standardized metrics, and consideration of differences – should yield meaningful estimates of the economic value of interventions to boost SE skills.

Moreover, this framework gains even greater salience in light of current practice in SEL evaluations. For most interventions there are important uncertainties and knowledge gaps. Most evaluations give a very incomplete sense of how valuable – and hence how attractive to policymakers – a particular intervention is. This economic framework serves to illustrate these uncertainties and gaps. As such cost-benefit analysis yields information on what we do know and illustrates the possible importance of what we do not know.

3.2 Ingredients Method for Costs

The ingredients method is a cost-accounting approach that is compatible with the economic concept of opportunity cost and has been used successfully to undertake cost and cost-effectiveness studies in education (Levin and McEwan, 2001; IOM/NRC, 2014). With this method all the ingredients or

9 A third metric is the internal rate of return, defined as the discount rate that makes the present value of benefits equal to the present value of costs. This metric conveys less information than the other two metrics but may be useful as a shorthand way to compare programs of similar scale and duration.

inputs used to implement an intervention – net of the costs of the alternative or 'business as usual' program – are identified and specified. Inputs should be counted as incremental beyond what is usually required in the absence of the intervention. All inputs should be included and evaluated for their costs, regardless of who pays for them (with their financing source identified).

Program ingredients are grouped into personnel, materials/equipment, facilities and other inputs. Ideally, information on ingredients is collected from semi-structured instruments administered to key personnel responsible for implementing the intervention. Each ingredient is then priced out using prices from independent sources based on actual market outcomes (or shadow prices if market prices do not reflect opportunity costs or are unavailable).[10] The cost of the intervention is then reported as the total sum of all ingredients multiplied by their unit prices. Also, some interventions induce additional resource use if the intervention succeeds in improving earlier educational outcomes (e.g., if more students are induced to attend college). When monetary benefits are calculated, additional costs engendered by success are typically counted as 'negative benefits' and put on the benefits side of the ledger.

The ingredients method is strongly preferred over reliance on budgets (and very strongly preferred over reliance on statements from program deliverers). Agency budgets are rarely comprehensive, covering all relevant ingredients, and are almost never itemized in a way that clearly identifies how much is spent on a given intervention. Using budgets it is very difficult to disentangle the incremental costs of an intervention from the regular school operations. As well, budget statements reflect local prices and not what an intervention would cost if another agency decided to implement it in their local context. If budgets were used, interventions in areas with high prices would be automatically disadvantaged in any economic comparisons.

The ingredients costing method is generally applicable across all types of educational interventions. However, for SEL interventions the critical issue is how to specify incremental costs relative to the business as usual provision. Where an SEL intervention is delivered as an after-school program, then incremental costs are straightforwardly measured as the cost of that program. However, if the within-school curriculum is altered to emphasize SEL, it may be difficult to identify what incremental resources are involved. More importantly in this case, curricular changes that emphasize SE skills ought to enhance SE skills – that is their purpose. Logically when time and resources are reallocated to a new goal or focus such as SEL, other outcomes obtain less time and resources and so are impaired. At issue is what skills are being displaced with any reallocation of resources and effort from one focus to a different one.

Critically, the costs of SEL interventions must be measured with reference to the overall goals of the education system as well as the specific SE skills that are being targeted. Those who argue that there is no cost to adding an SE intervention to the regular teaching schedule, and therefore that there are no additional instructional costs or displacement of productive activities, should provide evidence that this can be done. Similarly, the burden of proof is on those who argue that time spent on SE skills will not compromise achievement. Certainly, there is a limit to the number of goals and activities that can be addressed within a given time frame. It might be possible to meet this burden of proof: there may be slack in teacher schedules so that they can do more; or it may be that enhanced SE skills complement cognitive learning such that achievement scores and SE skills move mutually upward; or it may be that SEL can indeed be seamlessly integrated into the existing curriculum. However, these possibilities should be substantiated with evidence. As a general rule, instruction in one area should be assumed to

10 These prices are taken from appropriate datasets of statistical agencies (e.g. the Bureau of Labor Statistics or the National Center for Educational Statistics); an extensive catalog of prices of education inputs is available from cbcse.org.

displace instruction in another area: the working assumption is that there is no 'free lunch' in terms of resource use to boost SE skills (or achievement).

3.3 Shadow Price Methods for Benefits

The benefits of interventions are derived from shadow pricing – placing monetary values on – the interventions' impacts. A shadow price is based upon the societal willingness to pay for a specific impact (Levin and McEwan, 2001: Chapter 4, p.60-61). Conventionally, market prices indicate willingness to pay. But where market prices are missing or are distorted, then shadow prices for willingness to pay must be derived independently. For example, it does not make sense to conceptualize student aggression as a 'market' (with buyers and sellers of aggression) and so there is no easy way to observe a 'price' for student aggression. Instead, we have to calculate how much we think society is willing to pay to reduce student aggression by a certain amount. That amount will depend on how important society feels about combating aggression; as a lower bound, the value can be represented by how much is currently spent to combat aggression.

In theory, shadow prices can be calculated for most behaviors that are the target of SEL interventions. The fundamental approach of shadow pricing is to measure these behaviors on a scale and then calculate shadow prices for each notch up or down the scale. There are several different techniques for deriving shadow prices (e.g. the burden method, the hedonic method) and the validity of each technique depends on the particular impact being priced out (Boardman et al., 2011).

In practice, deriving shadow prices for impacts from SEL interventions is complex. For some impacts, approximate shadow prices already exist. Published studies have estimated the shadow prices of an array of social 'ills', such as smoking, teenage pregnancy, child abuse, asthma, and being a career criminal (Maynard and Hoffman, 2008; Cohen and Piquero; 2009, Table 12; and Cohen et al., 2010, Table 8); for shadow prices of ADHD during the intervention years, see Jones and Foster (2009). However, as we describe below, these shadow prices have to be adapted to specific population groups or recalibrated to specific impacts before they can be applied to SEL interventions.

For other impacts of SEL, new shadow prices must be derived. Given the existing evidence base and methodological research, it is not as yet possible to estimate shadow prices for all possible impacts. Durlak et al. (2011 Table 2) catalog six categories of impacts arising from SEL interventions: (1) academic performance; (2) SE skills; (3) positive social behavior; (4) conduct problems; (5) emotional distress; and (6) attitudes. For academic performance, it is possible to derive or adapt existing shadow prices (see above). For SE skills, it is possible to calculate or approximate shadow prices in terms of increased earnings (see above). For behaviors and conduct, it is possible to derive shadow prices where these actions have clear economic consequences (e.g. a specific juvenile crime). For emotional distress, it is also possible to derive shadow prices, but typically these impacts are less likely to be expressed in ways that do have obvious economic consequences. For the last category – attitudes – it is very difficult to derive shadow prices (Heckman and Kautz, 2012; Borghans et al., 2008). Fundamentally, shadow pricing can only be applied to behaviors (not attitudes) that can be measured with validity.

One short-cut would be to find a 'unit shadow price' for specific SE skills, i.e. an overall willingness to pay for increments in these skills (or a measure that encompasses most of this willingness to pay). Two approaches – the mediation of general SE skills through earnings and education – have been discussed above. Two other possible techniques for deriving a unit shadow price are to look at aggression or the economic burden per 'high-risk' youth.

Child aggression is typically measured using the aggregated Behavioral Assessment System for Children (BASC). This system has been used extensively in research and is validated (see Reynolds and Kamphaus, 2002, Chapter 3). Helpfully for shadow pricing, BASC T-scores can be scaled to behaviors as: clinically significant (>70), at-risk (60-70), average (40-60) and below average (<40). Also, there is evidence on how aggression persists over time: studies have shown how aggressive and disruptive behavior in primary school predicts aggressive behavior through middle and high school and whether different trajectories of aggressive and disruptive behavior are associated with a range of negative life outcomes in youth and early adulthood.[11] Thus, in theory BASC shadow price notches could be calculated. However, we are not aware of studies that have priced out changes in BASC T-scores either in terms of contemporaneous or lagged benefits. Also, studies typically disaggregate the BASC according to the teacher or child reporter; and it is less clear how the different scores of these reporters independently translate into behaviors and, so, into shadow prices (Reynolds and Kamphaus, 2002). For future research we believe this approach is promising, although it would require a dedicated research endeavor to establish accurate shadow prices for the BASC scale (and so is beyond the scope of our current endeavor).

Another option is to derive a unit value for general youth delinquency that subsequently leads to adult delinquency. For example, Cohen and Piquero (2009, Table 12) have calculated the lifetime social burden per career criminal, one-time offender, high school dropout and drug user over the lifetime. However, there are two issues in applying these values to enhancements in SE skills. One concern is that it may not be valid to translate impacts across any of the six SE categories described above into a general youth delinquency status. This translation requires assumptions about how aggressive behavior relates to criminal activity or externalizing behaviors relate to drug use, for instance. A second issue is that this approach focuses only on the long-term consequences of SE skills and not on the immediate impacts, and the latter may be the most salient (particularly if there is uncertainty about fade-out of SE skills). Overall, this approach requires significant extrapolation and inference to produce shadow prices. However, it can serve as an approximation when no other evidence is available and can be applied in sensitivity analysis.

In summary, the use of a single SEL index – in terms of earnings, or mediated through education, or via the BASC, or a general delinquency index – poses a number of methodological challenges. Therefore, our approach is to derive separate shadow prices for each intervention based directly on the impacts reported for that intervention. This approach, which is more research-intensive, ensures that each intervention is evaluated with respect to its intended objectives.[12] As a sensitivity check, however, we apply these more general approaches to trace their implications for SEL.

Currently, it is not possible to determine a shadow price for every impact from each intervention: there is insufficient data and research evidence to do this. This creates a risk that interventions where data on benefits do exist will appear more beneficial than interventions where the evidence base is more sparse. It also carries the related risk that interventions where more outcomes are evaluated will appear more efficient than those where only a few outcomes are investigated. To illustrate the extent to which impacts are captured by shadow pricing, we create benefit 'maps' for each intervention. These maps show the outcomes evaluated, the instruments used to measure these outcomes, and whether

11 These outcomes include early sexual activity, teenage pregnancy, school dropout and drug abuse (e.g., Montague et al. (2011) uses BASC scores to link behavior in grades 1-5 to outcomes at age 19/20).

12 A similar approach is undertaken by (Lee et al., 2012).

these outcomes can be expressed as shadow prices. The maps show whether extant interventions can – under current research circumstances – be evaluated appropriately using benefit-cost analysis.

Our benefit-cost analyses of six interventions are reported below. These interventions were selected using several evaluation criteria. Primarily, interventions were selected where there was existing evidence of effectiveness based on a methodologically rigorous evaluation. In addition, interventions were selected with the expectation that costs data could be obtained and that benefits could be assigned shadow prices. Interventions operating across a range of dimensions were also selected, so as to illustrate the variation in efforts to enhance social and emotional learning.

4. SEL INTERVENTIONS

4.1 4Rs

Program

The 4Rs Program (Reading, Writing, Respect, & Resolution) focuses on social and emotional learning and literacy development in grades K-5, with an overall goal of ameliorating aggression and violence at an early age. The 4Rs curriculum, specific to each grade, aims to increase pro-social behavior and help students develop cooperative problem-solving skills. At each grade level, there are seven units, each based on one literary work, that highlight themes such as conflict, diversity, and relationships, and the curriculum reinforces those themes through skills practice.

4Rs has been found to be effective at reducing aggression. In a recent longitudinal evaluation of relatively disadvantaged third-grade children, Jones et al. (2011) employed a school-randomized, experimental design across 1,184 children and 146 teachers in 9 treatment and 9 control group public elementary schools in New York City. The evaluation found improvements of 4Rs compared to 'business-as-usual' elementary schooling on: student self-reports on hostile attributional bias, aggressive interpersonal negotiation strategies and depression; teacher reports of attention skills, aggressive and socially competent behavior; and on math/reading achievement. We apply the results from this study to calculate the costs and benefits of 4Rs.

Costs

We estimate the costs of replicating this implementation of the 4Rs program. We consider only the costs of the program above and beyond the resources students already receive as part of their regular instruction in school, i.e., we identify the incremental costs of introducing the programs into existing school activities. All cost estimates exclude any costs associated with the conduct of research activities and reflect only program implementation costs throughout the course of the first two years of the evaluation study. Costs are adjusted for inflation using the CPI-U into 2013 prices and national prices are applied. A discount rate of 3.5% is used.

The cost estimates for 4Rs are adapted from Long et al. (2014, Tables 3 and 4).[13] The analysis by Long et al. (2014) used the ingredients method and high quality data to calculate a whole school costs analysis of 4Rs across grades.[14] Long et al. (2014) calculated costs for all three years during which the longitudinal evaluation of the 4Rs program took place (2004-2007). As the impact evaluation only refers to two years of 4Rs, only two years of costs are reported here and national rather than local prices are applied to provide comparable costs of replication.

Personnel accounted for the most significant portion of program costs. This category includes teachers, center administrators, 4Rs staff developers, and consultants. Incremental school principal time and parental time were not included as these were assumed to be trivial per student. Teacher input was counted for instructional delivery of the curriculum, initial training, ongoing training and workshops. Hours of teacher input were collected from the Center's weekly accounting logs and

13 We appreciate communications from Long et al. (2014) to help us adapt their estimates.

14 Long et al. (2014) compiled the ingredients to implement the 4Rs program through review of documents from the evaluation study (e.g. teacher logs on what they did each day in terms of program implementation etc.), extensive research into the budgets and meetings with the program's director and accountant. The researchers did not interview teachers or principals directly.

teacher time logs. Overall, the program did not require a particular experience level for these teachers. Teachers devoted time to the delivery of the 4Rs curriculum in each year of program participation. Per year, this totaled to a 7-unit, 21–35 lesson (number depending on grade level), literacy-based curriculum devoted to SEL. Students received one 4Rs lesson per week, one hour per lesson. Before beginning 4Rs, classroom teachers received initial training of approximately 25 hours. Teachers also received on-going training, working with 4Rs staff developers or administrators on-site at their school in individual or small group sessions (approximately 12 contacts each year). Teachers also attended workshops throughout the school year.

Teacher cost was estimated using a national average price for an elementary school teacher in 2013. The same rate was applied for teaching time, initial training, ongoing training and workshops. Prices were adjusted to include fringe benefits valued at 29.5% of salary. Annual salaries were divided by an average academic year of 1,260 hours to convert to hourly wages.

Data on other personnel was also taken from Center's accounting logs and measured in days. Center administrators included the Executive Director and two administrators (one full-time, one part-time). There were seven 4Rs staff developers working on hourly schedules. (The time of consultants for translation services and curriculum development was not included as this was assumed to be a developmental and not an operating cost). The hourly wage rates (including fringe benefits) for administrators and developers were taken from the database catalog of prices at www.cbcse.org.

Facilities space included workshop space and the Center office space. Two facilities were rented out to host 4Rs workshops. For the workshop space, this was measured as a single unit. The cost per unit was represented as a set fee that the facilities charged for rent. This data was collected using the Center's accounting records. Additionally, school classroom space for instruction of the 4Rs student cohort was costed out. Space estimates were based on class sizes of 26 students receiving one 4Rs lesson per week of one hour per lesson. Classroom space costs are expressed in national average costs.[15]

The 4Rs program included kits and basic instructional materials (but no computer materials).[16] These were costed according to invoice prices. Travel costs for 4Rs personnel were included based on invoiced amounts. Other subsistence ingredients (such as basic foodstuffs) were also included.

As noted above, a critical feature of costs analysis is that the costs must be measured in relation to outcomes. For 4Rs, this means that we have to decide whether or not to include teaching time as part of the program's incremental costs. The program was intended to fit into the standard literacy block of the school curriculum in an equivalent classroom space and serve as a partial substitute for regular literacy instruction. The teaching time – and classroom space – is therefore not strictly incremental to the implementation of 4Rs. The only difference was in how these students were taught and so, given the same amount of instructional time, the difference in costs between the methods might therefore be very small. However, the decision as to whether to include instructional time depends on how the impacts of 4Rs are measured. If the impacts that are measured are only those that 4Rs is intended to influence (and not those intended or inherent to the provision received by the control group), then it

15 For educational facilities rental rates are not generally available as national prices. We apply school building construction costs (adjusted for cost of land, development, furnishings and equipment) amortized over 30 years. Costs associated with initial training for 4Rs program are amortized over 3 years. We do not amortize ongoing coaching and support to teachers. For classroom space we use new elementary school national median prices amortized over 30 years at a 3% interest rate. Class size estimates are from http://schools.nyc.gov/AboutUs/data/classsize/classsize_2013_11_15.htm. The school year is assumed to be 36 weeks.

16 We amortize initial training and initial materials (4Rs Kits) over 3 years at an interest rate of 3.5%.

Table 1

Costs for 4Rs Program (Single Cohort for Two Years)

Categories/Ingredients	Year 1	Year 2	Total
Participants	630	630	630
Personnel:	91%	93%	92%
Training and support	50%	66%	0%
Teachers - 4Rs delivery	41%	27%	33%
Facilities:	8%	6%	7%
Rental Morningside office	1%	1%	1%
Rental Workshop Space	0%	0%	0%
Classrooms	7%	5%	6%
Materials and equipment:	1%	1%	1%
4Rs Kits	1%	0%	1%
Supplemental materials	0%	1%	1%
Total	100%	100%	100%

would not be surprising to see a big impact from the program. Therefore, we report resource costs both with and without instructional time (and the related instructional facilities costs). Including instructional time is conservative, suggesting that 4Rs only influences SE skills and not other skills. Excluding instructional time is possibly too restrictive in that it implies that 4Rs manages to enhance SE skills whilst also maintaining all other skills at the same standard as the business-as-usual approach.

For the two-year program providing 4Rs to a cohort of 630 participants, the present value total cost is $262,300 if instructional costs are excluded or $426,600 if instructional costs are included. Hence, the average cost without instructional time/facilities included is $420; and with instructional time/facilities, the average cost is $680.

Benefits

There are many impacts from 4Rs that could potentially be translated into money benefits. Based on the available evaluations, we have generated a map of these benefits (see Appendix I, Map 1). There are impacts in: understanding and handling feelings; achievement; academic skills; and other domains. These impacts have specific outcomes that are measurable, but only a subset of these outcomes can be monetized at present.[17] In addition, 4Rs has impacts that are not easily measured (e.g. cooperation or dealing with diversity). Although these impacts are valuable in their own right, if they cannot be measured then it is not possible to assign shadow prices to them.

In this benefit-cost analysis we are able to consider three behaviors. To correspond with the costs analysis, we measure these behaviors at the end of two years of the intervention, i.e. in spring of fourth grade. Our focus is on teacher reports of ADHD symptoms, social competence, and aggressive

17 In particular, we are not able to consider the general class of impacts measured using the BASC such as child attitudes and thoughts, even as these may predict subsequent behavior.

behavior. The results based on Jones et al. (2011, Table 3) yield statistically significant effect size impacts after two years for these behaviors.[18] These effect size gains are: 0.12, for reduced ADHD symptoms; 0.14, for social competence; 0.13, for reduced aggression (average from 0.05-0.21). These effect size gains are equivalent to moving from the median position to the 44th-45th percentile of outcomes.

Shadow prices for ADHD, social competence, and aggression are derived using the cost-of-illness method. That is, the shadow prices are based on what society currently spends on these conditions through the health care system. For the baseline estimates, the ratchet effect is assumed at zero and decay rate is assumed to be infinity. That is, there are no effects in the first year of the intervention, and there are no effects that exist beyond the implementation of the intervention; there are only effects in the second year of the intervention. For ADHD, spending is from Jones et al. (2009). For social competence and aggression, spending on oppositional defiant disorder (ODD) and conduct disorder (CD) respectively are applied; these are from Foster et al. (2005).[19] These spending figures are for a similarly disadvantaged population to the 4Rs group. All these estimates are conservative in that the 'cost-of-illness' method typically excludes some important costs (e.g. family expenditures).

The annual present value benefits of moving from the median burden to the 44th-45th percentile burden are: $2,490 for ADHD; $1,360 for Social Competence; and $4,470 for aggression. Therefore, the present value sum of benefits, i.e. the total immediate benefits of 4Rs expressed as a present value back to the first year of program delivery, is $8,320.

We perform several sensitivity tests based on different assumptions. First, we assume that the ratchet effect is linear. That is the benefits after one year are half the size of the benefits after two years. This assumption is as plausible as assuming there is no effect until after two years of the intervention (but it is less conservative). With a linear ratchet effect the total benefits are $12,630. Second, we apply a decay rate from Washburn et al. (2011). This decay rate is one-third within two years after the cessation of the intervention (and one-third more for the subsequent year). With this decay rate, the total benefits are $16,370. Third, we assume that the decline in ADHD will have an impact on labor market outcomes. Fletcher (2013) estimates ADHD to be associated with a fall in labor market attachment of 5%. Using the CPS data from 2009-2013, we calculate average lifetime earnings of $383,100 in present value terms in 4th grade. With an effect size change of 0.12, the value of the change in labor market attachment is $580-$780. With labor market effects included, the total benefits are $8,930. We consider these to be only a portion of the total benefits, that portion that could be identified and measured, and refer to these as the minimum benefits of the intervention.

Benefit-Cost Results for 4Rs

The benefit-cost comparisons for 4Rs are given in Table 2. The estimates of costs per participant are $1,410 and $2,590. But to be conservative, the latter cost estimate is used as the baseline. The minimum commensurate benefits are $8,320. Our main concern is whether the minimum estimated benefits exceed the costs per participant, the gain in net present value when costs are deducted from

18 Unadjusted descriptive statistics for these behaviors show no difference between intervention and control groups for 'Child Social Competence'; adverse effects for 'Child Aggression'; and positive effects for 'Child ADHD symptoms' (Jones et al., 2011, Table S1).

19 The annual incremental costs per ADHD=($4100-$1800) per year (Jones et al., 2008, Figure 1, estimates for youngest year). For Social Competence: annual incremental costs per ODD=($2750-$1490) per year (Foster et al., 2005, Table 1). For Aggression: annual incremental costs per CD=($5630-$1490) per year (Foster et al., 2005, Table 1). These figures are in 2000 dollars rounded to nearest ten. Moving down from the median to the 45th percentile is associated with costs that are 3% lower per percentile. The discount rate is 3.5%.

the benefits. The net present value gain at baseline is $5,730 per participant for 4Rs a large gain in benefits relative to costs. We remind the reader that our focus is on whether the minimum estimated benefits exceed the costs, not comparisons with the other interventions, given that the proportion of overall benefits for each intervention captured in the estimates can differ substantially.

The estimates for the net present value for delivering 4Rs to 100 students are given in Table 2. To be conservative, the cost estimate including instructional time is used as the baseline. Therefore, the net present value gain per 100 students is $764,000.

Table 2

Benefit-Cost Results for 4Rs for 100 Students

	Costs	Benefits	Net Present Value
Baseline Estimate (AC2 incl. instruction time)	$68,000	$832,000	$764,000
Sensitivity tests:			
AC2; B + labor market gains	$68,000	$89,300	$8,250
AC1 (excl. instruction)	$42,000	$832,000	$790,000
AC1; B + labor market gains	$42,000	$893,000	$851,000
AC2; ratchet=0.5	$68,000	$1,263,000	$1,195,000
AC1; ratchet=0.5	$42,000	$1,263,000	$1,221,000
AC2; decay rate=0.33	$68,000	$1,637,000	$1,569,000
AC1; decay rate=0.33	$42,000	$1,637,000	$1,595,000
Average across sensitivity tests	*$55,000*	*$1,203,000*	*$1,148,000*

Sources: Table 1 above. *Notes:* Present values at third grade (d=3.5%). Baseline model assumes AC2 and baseline benefits. Sensitivity tests sorted by NPV.

To ascertain how robust are the gains in minimum benefits relative to costs, we test economic assumptions that might alter the net returns in Table 2. These alternatives apply the lower cost estimate (excluding instructional time), apply different assumptions about the fade-out and decay rate of effects, and include labor market benefits. Given the conservative assumptions for the baseline estimate, all the sensitivity tests yield higher net present values than the baseline. Across the seven tests, the average net present value is estimated at $1,148,000 per 100 participants. (Moreover, even these results do not include many of the possible impacts from 4Rs itemized in the benefits map). Finally, given the baseline costs of $680, a threshold benefit level can be identified: that is, the economic value of 4Rs is positive even if only the benefits of reduced aggression are accounted for ($4,470); and the value is positive even if only the benefits for ADHD are accounted for ($2,490). Thus, programs such as 4Rs may be economically justified based on very partial measures of benefits in relation to full costs.

4.2 Positive Action

Program

Positive Action is a school-based curriculum and supplemental set of school cultural and family activities designed to promote students' positive thinking, actions, and self-concept. This intervention is

intended to increase intrinsic motivation to learn and reinforce positive behavioral choices. It consists of a series of short lessons at each grade level organized into six units: self-concept; positive actions for your body and mind; managing yourself responsibly; treating others the way you would like to be treated; telling yourself the truth; and improving yourself continually. Supplementary materials can be used by school leaders, teachers, counselors, or families to help students apply the main lessons of the program to target student development in specific areas, such as prevention of bullying or drug abuse. The program is based on the theory that positive actions make us feel good about ourselves, and therefore become self-reinforcing over time.[20]

Positive Action has been extensively studied, notably in two large-scale, multi-year randomized controlled trials.[21] One study took place in 20 elementary schools in Hawaii and evaluated outcomes across a range of academic, behavioral, school quality, sexual activity, substance abuse, and violence outcomes (Snyder et al., 2010, 2012, 2013; Washburn et al., 2011; and Beets et al., 2009).[22] The other experimental evaluation of Positive Action was conducted in Chicago. With random assignment at the school-level, outcomes from Positive Action were evaluated at seven treatment schools over six years of implementation across grades 3-8. A series of studies examined outcomes such as academic achievement, student behavior, bullying, drug use, violence, and mental health (Lewis et al., 2012, 2013ab; Li et al., 2011; and Bavarian et al., 2013).

We perform our benefit-cost analysis using the outcomes from the trial in Chicago. We chose this site for several reasons. It is the most recent trial of Positive Action. Also, the outcomes measured in that trial most closely align with social and emotional learning outcomes discussed above. Specifically, assignment to Positive Action yielded impacts on: student self-reported normative beliefs about aggression, bullying and other violent or disruptive behaviors; parental reports of bullying or conduct problems; and on disciplinary actions based on administrative data (Lewis et al., 2013). There was a 0.38 effect size reduction in self-reported bullying, as well as there were significant impacts on parent-reported bullying, disruptive behaviors, and disciplinary actions.

Costs

We estimated costs using the ingredients method based on interviews from two sites that were part of the Chicago program evaluated by Lewis et al. (2013). The interviews were conducted with a teacher and a counselor who served as site coordinators at two elementary schools.[23] We estimate costs using the ingredients method and calculate only those costs that were incremental to the delivery of 'business-as-usual' schooling. All costs are estimated in 2013 dollars and expressed in present values with a discount rate of 3.5%. Although randomization occurred at the school level and all students in schools assigned to treatment were eligible to receive Positive Action services, only the initial third grade cohort was

20 Program developers illustrate that dynamic with the Thoughts-Actions-Feelings Circle (TAF), in which thoughts lead to actions which lead to feelings about oneselves. Those feelings, in turn, lead to new thoughts, repeating the cycle in either a positive or negative manner. A positive cycle reinforces intrinsic motivation to learn, whereas a negative cycle undermines intrinsic motivation. The program therefore systematically emphasizes thoughts and actions that lead to a positive cycle.

21 In addition, other studies have examined alternate versions of Positive Action (as a pre-kindergarten or family program (see Flay, 20012, 2010 respectively). Other studies have evaluated Positive Action using quasi-experimental designs (Flay, Allred and Ordway, 2001; Flay and Allred, 2003).

22 This research was included in the meta-analysis by Durlak et al. (2011) and has met What Works Clearinghouse evidence standards (as a progress report by Flay et al., 2006).

23 Despite repeated requests to a sample of schools, we were unable to obtain contact with other personnel involved in the delivery of Positive Action in Chicago.

followed intensively in terms of measurement and implementation support. Yet, while that cohort was intended to receive Positive Action for six years, student attrition was very high such that average length of treatment was only 3.1 years. Therefore, we estimate the costs of 3 years of treatment, expressed as a present value in 3rd grade.

By far the largest program ingredient is personnel time. This includes the reallocation of classroom teacher time to deliver the program's approximately 140 15-minute lessons. Classroom teacher time was also incurred for lesson preparation, initial and ongoing training, and for implementation support. Other personnel were also involved in the program. A coordinator at each site, often a teacher or counselor, organized the program and provided delivery support to teachers as well as organized regular parent meetings and school-wide assemblies. School principals spent time supporting the program, and parents contributed time attending meetings and occasionally volunteering in classrooms. Also, at one site a school aide served as an interpreter at parent meetings. Positive Action staff and trainee staff at the University of Illinois-Chicago provided training and implementation support.

Estimates for the amount of time each of these groups contributed to the program were made based on interviews with site coordinators. National average prices for each ingredient in terms of wage time were then obtained from relevant databases.[24] Fringe benefit rates of 29.5% for personnel in K-12 education and 28.5% for personnel in higher education were applied. Annual salaries were divided by an average academic year of 1,260 hours to convert to hourly wages.

Other ingredients include facilities for training, meetings, and most importantly classroom space for program delivery.[25] Facilities costs were based on rental rates for classroom space.[26] Also, program materials were costed, including curriculum kits with workbooks and worksheets for teachers and supplementary kits for counselors, school administrators, and families, and stickers as positive reinforcement for students. Prices for program materials were obtained from the Positive Action catalog and amortized over 5 years, based on average teacher turnover, at a 3.5% interest rate.

Finally, the extent to which teacher time devoted to the program should be considered incremental is unclear. Although there are no additional financial outlays – since teachers would be teaching even in the absence of the program – there is an opportunity cost of lost instruction in what would have happened in the absence of the program. To be conservative, we consider all teacher time to be an incremental cost.

The costs of Positive Action at the two sites are given in Table 3.[27] In total, there were 140 participants in the program over this three year period. The total costs for this cohort was $70,710. The average cost per participant was therefore $510.

We undertake two sensitivity tests on costs. For the baseline estimate all students in a school were assumed to receive Positive Action. All school-wide costs, such as coordinator, principal, and trainer

24 These included: the Database of Educational Resource Prices at cbcse.org; and data from the Department of Labor, the Bureau of Labor Statistics, the National Center for Education Statistics, the National Education Association, and the Integrated Postsecondary Education Data System.

25 We assumed an average elementary school classroom size to be 900 square feet, and conservatively assumed that meetings and trainings would take place in a similarly sized regular classroom.

26 In the absence of a significant rental market for educational spaces, we estimated the hourly opportunity cost of using educational facilities by amortizing new construction prices over 30 years at a 3.5% interest rate. Elementary school construction prices were obtained from the 16th Annual School Construction Report at http://webspm.com/~/media/457BAA6BC0164B86A1D04E80DD696569.pdf.

27 From the interviews we found notable variation in program implementation between the two sites. Program treatment was notably more intensive in one site (and so variable costs were higher). But the site also served a larger student population (so fixed costs per student were lower).

Table 3
Costs for Positive Action (3-Year Program)

Ingredient	Share of Costs (Site A)	Share of Costs (Site B)	Share of Costs (Sum)
Participants (single cohort for 3 years)	*80*	*60*	*140*
Personnel:			
Classroom teacher	74%	59%	69%
School staff (coordinator, principal, aide)	11%	8%	10%
Management (implementation support)	4%	10%	6%
Parent time	2%	8%	4%
Training (incl. travel/facilities)	1%	1%	1%
Facilities:			
Classroom space	6%	10%	7%
Materials and equipment:			
Program materials	3%	5%	3%
Total Resource Cost	**100%**	**100%**	**100%**

time as well as school culture materials and activities, are divided among six cohorts. However, coordinators at both sites indicated that treatment intensity in non-evaluated cohorts was lower. Therefore, school-wide costs were instead allocated across students in the evaluated cohorts who received intensive treatment. Under this more stringent assumption, the total program cost was $142,200 or $1,020 per student.

A second sensitivity test was based on length of program receipt. The baseline estimates assume an average of 3 years of program receipt. But students in later years of the study could have benefited from increased program effectiveness due to teacher experience or improved overall school climate (as more of their peers had received the program). As a check, we estimate the program costs spread across six years of program delivery. Under this assumption, program costs using the baseline estimates dividing school-wide costs across 6 cohorts would be $131,800, or $940 per student.

Benefits

Positive Action is likely to have many impacts on social and emotional outcomes. These include improvements in personal behavior, mental health, achievement and academic behaviors, and school climate (see the Benefits Map in the Appendix I). However, not all these impacts can be measured in such a way as to allow them to be monetized. In addition, some outcomes, such as school quality and academic achievement, overlap with one another, and so to count both of them would be likely to double-count the consequences of Positive Action. Again, we warn against comparing benefit-cost outcomes among interventions when the portions of the benefits that are captured by the evaluations and the ability to convert them to benefits are unknown among the alternatives.

For this benefits analysis, we estimate the shadow price of bullying reduction. Estimates of the shadow price for bullying are only approximate. The economic consequences of bullying include: days of missed school (both for the victim and for the suspended or expelled perpetrator); school personnel time to respond to bullying cases; school practices and training programs to mitigate bullying; parental

time to resolve bullying; resources of social services; resources for alternative placement of perpetrators; and justice system expenditures for cases that involve the legal system.

To our knowledge, no study has estimated the benefits of these reductions using a rigorous shadow pricing technique.[28] A very lower bound estimate is to apply the number of school absences that are associated with bullying and price these out using society's willingness to pay for a day of schooling.[29] Under this assumption, an effect size reduction of 0.38 in bullying for three years yields 18 fewer days of bullying. We assume these days build cumulatively over the three year period (equivalent to the ratchet effect for other SEL programs).

Benefit-Cost Results for Positive Action

The benefit-cost results for Positive Action are reported in Table 4. Measurable benefits exceed costs by a considerable amount.

Table 4

Benefit-Cost Results for Positive Action Per 100 Participants

	Costs	Benefits	Net Present Value
Baseline	$51,000	$258,000	$207,000
Sensitivity tests:			
S1. Zero ratchet effect	$51,000	$810,000	$759,000
S2. Slower fade-out	$51,000	$392,000	$341,000
S3. Education benefits	$51,000	$337,000	$286,000
S4. Delinquent youth benefits	$51,000	$1,007,000	$956,000
S5. Intensive program	$60,000	$130,000	$70,000
Average across S1-S5			$443,600

Sources: Table 3 above. *Notes:* Present values at third grade (d=3.5%). Sensitivity tests assume costs or benefits from baseline model unless other specified. For delinquent youth benefits, see Appendix Table 1.

We undertake five sensitivity tests. For the first two tests we vary the ratchet and fade-out effects respectively. For our baseline estimates, we assume that the ratchet effect is cumulative: that is, the benefits accumulate as the intervention is delivered. This assumption is as plausible as linearity (but it is less conservative). Also, for the baseline estimates the fade-out rate is immediate after the final year of evaluation (that is, there are no effects on bullying beyond the length of the program). For our first sensitivity check, we assume the ratchet effect is immediate: that is, the effects of the intervention are immediately equal to the final year impact. Obviously this increases the benefit-cost ratio. As a third

28 For discussions that at least attempts to estimate this issue, see www.principals.org/Content.aspx?topic=The_Financial_Costs_of_Bullying_Violence_and_Vandalism; www.casafeschools.org/FactSheet5rev2.pdf; and www.highmarkfoundation.org/pdf/publications/HMK_Bullying%20Report_final.pdf

29 Societal willingness to pay for a school day is estimated using average daily expenditures for a day of public school nationally in 2014 prices. This amount is $70 per day ($13,400/180, see nces.ed.gov/programs/digest/d12/tables/dt12_215.asp). There is also a value of increased achievement with fewer days of absenteeism (Gottfried, 2010, 2011ab).

test, we assume that the impacts decline linearly to zero six years after the initiation of the intervention. This too increases the benefit-cost ratio.

The third sensitivity test draws on recent evidence on the long-term effects of bullying reported in Wolke et al. (2013). Although this study is not representative of the U.S. population, it tracks longitudinally students who are bullies, victims, both, and neither. With long-term follow-up of the students, Wolke et al. (2013) identify strong deleterious effects of bullying on all those involved across a range of domains. Specifically, compared to those who have no experience of bullying, victims are 1.1 times as likely to have dropped out of high school, bullies are 1.6 times as likely, and those who are victim/bully are 3.1 times as likely, to drop out of high school. As an approximation, we can use these proportions, along with estimates of the economic burden of high school failure, to derive the shadow price of bullying. Accounting for the impact of Positive Action on bullying, we calculate that the present value benefit per participant from lower rates of bullying is (conservatively) $3,370. These benefits are easily greater than the costs.[30]

As a fourth sensitivity check we calculate benefits based on impacts reported in Beets et al. (2009) for Positive Action implemented in school in Hawai`i. At this site, the main statistically significant impacts were on substance abuse and violent behavior. For substance abuse, the treatment group rate was 4.0% compared to 7.6% in the control group; for violent behavior the treatment group rate was 6.1% compared to 2.2% (Beets et al., 2009). We use the shadow price of a substance abuser and career criminal from Cohen and Piquero (2009), adjusted to correspond to other shadow prices applied here. This shadow price is then multiplied by the number of reductions in substance abusers or career criminals yielded from a program serving 140 participants. The advantage of this method is that it avoids direct modeling of the lifetime impacts and the fade out of Positive Action. The disadvantage is that it requires extrapolation from behaviors in 5th grade to youth delinquency. Therefore, to be conservative, we assume 20% fade-out each year.[31] As well, these shadow prices do not count the burden of delinquency before youth. To avoid double-counting benefits we assume only one of these scenarios applies, i.e. only the substance abuse rate is reduced or the career criminal rate but not both. The calculations of benefits are given in Table 1 of Appendix II. The minimal estimates of benefits exceed costs by considerable amounts.

Finally, we assume that the program is delivered as intensively in the second year as it is in the first year. Under this assumption, the average cost of the program rises to $600. Again, the benefits easily exceed these costs.

Overall, the sensitivity analyses indicate that even a partial estimate of benefits for Positive Action exceeds costs across all conditions. Although this analysis yields results that are favorable to Positive Action, the sensitivity tests in Table 4 make clear that the benefit-cost ratio varies widely depending on which model assumptions are made.

4.3 Life Skills Training

Program

Life Skills Training (LST) is a school-based classroom intervention to reduce substance abuse and violence (www.lifeskillstraining.com). The intervention is generally delivered to at-risk students in middle and/or high school. Life Skills Training teaches social and emotional skills to build confidence and

30 Details of this calculation are available from the authors.

31 This fade-out function approximates to the correlation between 8th grade and adult behaviors of 0.38 reported by Hawkins et al. (1998).

self-esteem, equip youth with the skills needed to resist peer pressure, and generally improve social and emotional competence to reduce anxiety and improve a range of health outcomes. The curriculum is divided into three key components: knowledge and skills needed to resist use of alcohol, tobacco, and other drugs; personal management skills; and general social skills to build assertiveness. The program duration involves sessions over three years, with most of these in the first year of the middle school program followed by booster sessions.

A series of impact evaluations of LST have been conducted (see Botvin et al., 2006; Spoth et al., 2008, Griffin et al., 2011). These studies have examined a range of effects, including reductions in crime, smoking initiation, substance abuse, and risky behaviors (Griffin et al., 2006; Spoth et al., 2008). For each impact, LST is statistically significantly associated with lower delinquency.

Based on these impacts, two benefit-cost analyses of LST have already been conducted (Lee et al., 2012; Miller and Hendrie, 2008).[32] Adapted to be consistent with our other analyses, the results for these two studies are given in Table 2 of Appendix II. The studies differ with respect to estimated costs and how benefits are calculated. In Lee et al. (2012), the benefits are mediated through reductions in crime. In Miller and Hendrie (2008), the benefits are mediated primarily through changes in 'quality of life'. However, both analyses estimate that the costs of Life Skills Training are very low ($40 and $290 respectively) and that the benefit-cost ratios are extremely large although significantly different (at 46:1 and 21:1 respectively).

Our analysis follows these studies. We describe costs in detail, adjusting for displaced instructional time as appropriate, and itemize the array of possible benefits from LST. We also update the results into 2013 dollars.

Costs

Based on descriptive reports of LST, we calculate the costs by adapting the ingredients identified in the estimates from Miller and Hendrie (2008).[33] Personnel ingredients are based on time spent teaching LST, as well as time on training. Average national teacher salaries are applied, including fringe benefits.[34] Other ingredients include materials and subsistence for teachers. These are priced from several sources.[35] We include costs for personnel time for program delivery.

These costs are given in Table 5. We estimate the cost of three years of LST at $130 per student. These costs are very low because LST is a very light-touch program. However, LST has changed. Earlier evaluations of LST were based on implementation where instruction was within regular class time. In current versions of LST the students are pulled out of class to participate in the program. Thus, we note that these costs may not necessarily represent the same implementation in terms of intensity, training, oversight and support as was used to obtain the published effects used to estimate the benefits.

32 There has also been a cost-effectiveness analysis of LST compared with other programs that may reduce prescription drug abuse (Crowley et al., 2014).

33 Ideally, the full ingredients approach should be applied to a version of LST that had been evaluated. However, the identities of recently-evaluated sites were unavailable, and after contacting one of the earlier evaluation sites, we discovered that current program delivery in the school district differs substantially from descriptions of the program as evaluated. Therefore, our cost estimates are based on published descriptions of the program in evaluations. In fact, these cost estimates are roughly in line with those estimated by Miller and Hendrie (2008); they are higher than those of Lee et al. and Aos et al. (2004, 2012) as their estimates of the latter omit personnel costs for program delivery.

34 http://www.bls.gov/news.release/archives/ecec_12082010.pdf

35 These include Western CAPT, Maryland Blueprints, and Aos et al. (2004), as well as the LST website.

Table 5
Costs for Life Skills Training (3-Year Program)

Ingredient	Share of Costs
Participants	25
Teacher time	54%
Training	9%
Implementation fidelity	11%
Facilities:	
Classroom space	6%
Materials:	
Manual and Student Guides	21%
Total Cost	**100%**

Benefits

Our estimate of benefits is derived from impacts reported in Botvin, Griffin, and Nichols (2006). Although LST is primarily a substance abuse prevention program, Botvin, Griffin, and Nichols (2006) estimated the effects of the program on other risky behaviors with the hypothesis that the program could reduce behaviors besides substance abuse through increasing general social and emotional competencies. This study identified a statistically significant reduction in "Delinquency in past year" (odds ratio of 0.684), "Frequent fighting in past year" (odds ratio of 0.742), and "Frequent delinquency in past year" (odds ratio of 0.643) after one year of LST for sixth grade students. Thus, assuming a given population of 100 at-risk students where 50 are delinquent, if those students had received one year of LST, conservatively, only 37 would be delinquent.

Our estimates of benefits follow a similar approach to those for Positive Action. We shadow price the immediate monetary consequences (as the program is only one year, there is no ratchet effect). We then estimate lifetime impacts assuming that these impacts persist. This is likely a lower-bound estimate of the benefits of LST, as delinquency is an ancillary program outcome and because we may not be able to fully capture the economic burden of delinquency.[36]

At baseline, 53.2% of the combined intervention and control groups displayed any of the above delinquent behaviors, with no statistically significant differences between the groups. Therefore, we can estimate that 1,263 of the 2,374 students in the intervention group had previously exhibited delinquent behaviors, whereas only 937 students in the intervention group reported delinquent behaviors in the follow-up survey three months after the intervention, for a reduction of 326 delinquent students.

As with Positive Action, the immediate economic burden of acts of delinquency can include the value of lost instruction due to absenteeism and truancy, lost instructional time and costs of school staff time devoted to disciplinary actions such as suspensions and expulsions, and the burden experienced

36 As described by Botvin, Griffin, and Nichols (2006), this burden includes "destroying others property, throwing objects at people or cars, shoplifting, stealing from others, taking something from someone by force, or intentionally vandalizing a school or other building."

by crime victims due to loss or destruction of property. Conservatively, we estimate the value of the reduction in delinquency as a reduction in a single act of the least burdensome delinquent behavior per year for the 326 students whose behavior was positively impacted by the program. This approach is used because the data cannot be disaggregated to show exactly which delinquent acts were avoided and to what extent.

If a large portion of the reduction in delinquent acts is in the most severe acts and for the most serious offenders, the lifetime benefits of delinquency reduction could be significantly larger, ranging up to Cohen and Piquero's (2009) estimate of $877,030 per career criminal. It is highly unlikely, however, that all of the 53.2% of students identified as exhibiting any delinquent behaviors at baseline could be characterized as "moderate juvenile offenders" or "career criminals," so we proceed with a more cautious approach of estimating the economic burden of an individual delinquent act. The National Association of Secondary School Principals estimates the loss due to vandalism at $400 per incident and the direct costs of school suspensions at $170 per incident, which translate to $430 and $180 in 2013 dollars, respectively.[37] If 26.8% (1-0.742) of the cases of vandalism, all resulting in suspensions, were averted, LST could save $80 per participant per case per year; if we assume those effects persist without fadeout from the intervention in grade 6 through the end of grade 12, the present value of the benefits are $450 per participant.

Benefit-Cost Results for LST

Using our estimates of costs and benefits we can calculate the net benefits for LST, that is the benefits minus costs. This calculation is made within our framework (see above). As noted, the costs of the program are $13,000 for 100 students. With a conservative estimate, the net benefits are $32,000 per 100 participants. Thus, consistent with other studies, we find a strongly positive return to Life Skills Training.

We apply several sensitivity tests to the benefits of Life Skills Training. First, we assume that the only benefit derives from the case of vandalism that is averted per offender: this would reduce the benefits to $80 and yield a negative net present value. For other sensitivity tests, we apply alternative estimates of lifetime costs per at-risk youth: one test applies the benefits per career criminal averted from Cohen and Piquero (2009); another test relies on the association between LST and substance abuse and the benefit of averting such abuse; and a third test applies shadow price estimates from a recent health intervention delivered to 8th grade students to reduce delinquency.[38] Each of these is based upon riskier assumptions for which the evidence on effectiveness of LST is less direct. These sensitivity tests are reported in Table 6.

Overall, Life Skills Training yields a very large present value. Under a range of sensitivity tests, that value may be considerably larger, although we note that these tests are not very precise. Of course, given the very low cost of LST and its targeting of at-risk students, the program does not have to be especially effective in order to generate economic value.

---

37 www.principals.org/Content.aspx?topic=The_Financial_Costs_of_Bullying_Violence_and_Vandalism

38 For the first two tests, we adapt the estimates for career criminals and substance abusers reported in Table 1 of Appendix II. We adjust these amounts for delivery to 10th grade students and for the impacts of LST on being in this status (accounting for fade-out). Conservatively, we report this shadow price assuming the program averts one delinquent per 100 students. Finally, we adapt the shadow prices that Kuklinski et al. (2012) apply in their evaluation of a health intervention to reduce youth deviant behaviors. Adapting their benefits and adjusting them to 2013 prices, the lifetime benefits of $3,100.

Table 6

Benefit-Cost Results for Life Skills Training Per 100 Participants

	Costs	Benefits	Net Present Value
Baseline	$13,000	$45,000	$32,000
Sensitivity tests:			
S1. 1 case vandalism averted	$13,000	$8,000	($5,000)
S2. Career criminal probability	$13,000	$201,000	$188,000
S3. Substance abuse reduction	$13,000	$650,000	$637,000
S4. Delinquency reduction	$13,000	$310,000	$297,000
Average across S1-S4			$279,000

Sources: Table 1 above. *Notes:* Present values at third grade (d=3.5%). Baseline model assumes baseline costs and baseline benefits. Sensitivity tests assume costs or benefits from baseline model unless other specified.

4.4 Second Step

Description

Second Step is a classroom-based social skills curriculum for pre-school through the junior year of high school, with a distinct curriculum for each grade. The program builds on cognitive behavioral intervention models and consists of interactive lessons that relate to problem-solving and emotional management. The *Second Step* program is designed for school-wide implementation and is implemented by classroom teachers. The curriculum is intended to help students develop empathic behaviors and improve their skills in communication, social problem-solving, and critical thinking. The overall goal is for children to identify and understand their emotional state and to manage and communicate these emotions appropriately and so increase social competence and reduce aggressive and delinquent behaviors.

Second Step has been extensively studied. A series of impact evaluations have been conducted in the United States and other countries, and many report positive impacts with students from a variety of age groups, socioeconomic and racial backgrounds, and geographic settings. Between 1989 and 2012 there were 35 studies of Second Step that focused on children with or at-risk of emotional disturbance. However, none of these met What Works Clearinghouse evaluative standards (NCES, 2013).[39]

To calculate the costs and benefits of Second Step, we use the impacts from a recent study by Espelage et al. (2013). This study presents first-year results from a 3-year school-randomized controlled trial in 36 Midwestern schools (Chicago/Illinois and Wichita/Kansas). This study utilizes a sample of 6th-grade students who received the 15-lesson sixth grade curriculum. A large proportion of these students were eligible for free or reduced-price lunch (treatment 72.7%, control 75.6%). Thus, the sample is more disadvantaged relative to the national population. Espelage et al. (2013) evaluated the impact of Second Step on reducing youth physical and sexual violence: after one year, the researchers found a substantial reduction in self-reported physical aggression among students.

39 Studies have examined a range of effects, including: social competence, school performance, and satisfaction with life (Holsen et al., 2008); positive approach, caring, suppression of aggression, and consideration of others (Cooke et al., 2007); knowledge about empathy, anger management, impulse control, and bully-proofing (Edwards et al., 2005); decline in anxious and depressed behavior (Schick et al., 2005); and aggression and cooperation (Frey et al. 2005).

Costs

We estimate the costs of replicating this implementation of the Second Step program based primarily on descriptions of the program.[40] We consider only the costs of the program above and beyond the resources students already receive as part of their regular instruction in school, i.e., we identify the incremental costs of introducing the programs into existing school activities. All cost estimates exclude any costs associated with the conduct of research activities and reflect only program implementation costs throughout the course of the first year of implementation.

Without access to our usual sources for compiling and clarifying data on ingredients, we had to derive the costs of Second Step using a top-down approach. See Table 7. As described in Espelage et al. (2013), Second Step requires four hours of training, and teachers also completed implementation logs to see what parts of the curriculum had been implemented. Given hourly rates for teacher salaries as well as trainer costs, we estimate the cost of Second Step implementation at $50 per year per student when instructional time is not accounted for.[41]

Table 7
Costs for Second Step (1-Year Program)

Ingredient	Share of Costs
Participants	1,940
Teacher resource for training/logs (including facilities, materials)	10%
Trainer resource (including facilities, materials)	1%
Instructional time (incl. classroom space)	89%
Total Cost (excl. instructional time)	**$92,700**
Total Cost (incl. instructional time)	**$849,300**
Average Cost per student (TC1/n)	**$50**
Average Cost per student (TC2/n)	**$440**

Notes: Present value (*d*=3.5%) in 2013 dollars.

In addition, Second Step requires instructional time. Again as described in Espelage et al. (2013), Second Step involves up to 25 instructional hours over the course of an academic year (6th grade). A lower-bound estimate of total instructional hours available per school year is 800, so Second Step absorbs about 3% of total instructional time (or 2% of the entire time in school). Given school expenditures of $12,500 per student, this amounts to $390 for delivery of Second Step on the assumption that it is an addition to the regular school instruction. Overall, therefore, the total resource requirement for Second Step, including instructional costs, is $440. Thus, the critical issue for resource use for Second

40 Normally we use informed observers to assist in identifying the required ingredients and their purpose and implementation. Despite repeated attempts within the Chicago Public School system, we were unable to find personnel who had implemented Second Step and had direct information on ingredients usage. A representative of the Office of Social and Emotional Learning at CPS was helpful in confirming some details on the general implementation of the program and how it has evolved over time.

41 With hourly annual teacher pay at $55,190 (+29.5% benefits and multiplied by 70% for overheads, facilities, and management). Contact hours are 1,200 per year. Class size is assumed to be 20-24 students. Training is needed annually. Completion of implementation logs requires 2.5 hours annually. Trainer costs are based on salaries of $90,000 (with the same benefits, overheads rates, and class sizes; but contact hours of 600 per year). Materials costs are estimated at $20 per student.

Step is the extent to which it displaces or substitutes for the business-as-usual curriculum. If Second Step displaces current practice it will have a much lower cost than if it is a supplement to the regular curriculum.

Benefits

The array of outcomes and benefits of Second Step are shown in the benefits map (Appendix I). These outcomes encompass substance abuse, delinquency, and changes in sexual health status.

Potentially, where these outcomes translate into behaviors, they can also be assessed in monetary values. However, the extent to which these outcomes overlap or confound each other is unknown. As well, it is not possible to tell whether these outcomes are at the expense of achievement gains. Therefore, we apply two approaches to estimating benefits. One is to focus on a single, specific behavior for which plausible shadow prices exist. (This excludes drug knowledge and attitudes but includes sexual risk behaviors which may either induce sexually-transmitted diseases or teenage pregnancy). The other is to use an overall 'delinquent youth' shadow price that is intended to capture all the outcomes in aggregate.

The first approach is to calculate a shadow price for aggression. The outcomes of Second Step are from Espelage et al. (2013) who identified a 42% reduction in self-reported physical aggression. Using estimates from Foster et al. (2005) for the medical resources required to address cases of aggression, this outcome yields benefits of $4,320.[42]

As an alternative, we apply the shadow price of an at-risk youth calculated by Cohen and Piquero (2009). As reported in Appendix II, this shadow price is (conservatively) $877,030. From a baseline rate of 5% at-risk youth, Second Step would yield 2.1 percentage points fewer substance abusers to a new rate of 2.9% (corresponding to a 42% reduction in the number of at-risk youth). If the program does not exhibit fade-out, Second Step would yield benefits of $18,420 per participant. This is a very large benefit compared to the costs. More plausibly, we assume fade-out of 10% per annum until adulthood. Nevertheless, this still yields present value benefits of $7,550 per Second Step participant.

Benefit-Cost Results for Second Step

The costs and minimum net benefits for Second Step are reported in Table 8. Assuming that the costs of instruction are counted, the program costs for one year are $440. The benefits in terms of reduced aggression are $4,320. The baseline estimate therefore yields a net benefit of $388,000 per 100 participants.

The sensitivity tests show that Second Step is likely to yield positive net benefits under a range of scenarios. For example, if we apply the benefits from reducing the number of at-risk youth, then the net present value per 100 students increases from the baseline to $711,000. If we assume fade-out is not immediate (within one year) but is only complete after three years, then the net present value increases even further to $796,000. Finally, if we assume that achievement and other school outcomes are unchanged – even as the level of aggression is much lower – then the appropriate cost measure should exclude instructional costs. To the degree that a calmer environment reduces the need for discipline and disruption and makes the school safer, it is even conceivable that academic

42 Calculations assume costs per aggressor of $1,490 in 2000 dollars with 5% of youth population in this category. Fade-out is immediate after the program is completed.

Table 8

Benefit-Cost for Second Step Per 100 Participants

	Costs	Benefits	Net Present Value
Baseline	$44,000	$432,000	$388,000
Sensitivity tests:			
S1. At-risk youth	$44,000	$755,000	$711,000
S2. Slower fade-out	$44,000	$840,000	$796,000
S3. Unchanged achievement	$5,000	$432,000	$427,000
Average across S1-S3			$581,000

Sources: Table 7 above. *Notes:* Present values at third grade (d=3.5%). Sensitivity tests assume costs or benefits from baseline model unless other specified. For delinquent youth benefits, see Appendix Table 1.

achievement could increase. In this case, the net resource requirement for Second Step is very small. Yet, it seems unlikely either that it is costless to displace instructional time or that instructional efforts in one direction do not jeopardize efforts in another direction. The overall result suggests a good return on an SEL investment under a variety of assumptions.

4.5 Responsive Classroom

Description

Responsive Classroom is a pedagogic approach that focuses on how teachers both teach and interact with elementary school students. The approach is designed to provide teachers with strategies, structures, practices, and techniques to improve their self-efficacy, to impact student social and emotional, academic, and other non-academic outcomes, and to build a strong school community.[43] Through two one week sessions, coaching, and published materials, teachers are trained how to incorporate ten key practices into their teaching philosophy and pedagogy. Teachers use the Responsive Classroom approach to provide emotional support and to proactively manage the class. These practices lead to increased student motivation and engagement, which in turn increases academic skill acquisition.

Responsive Classroom has been identified as an effective program based on CASEL criteria.[44] It is integrated with classroom instruction and provides opportunities for students to practice SEL skills. Also, the program provides tools to assess implementation and evaluate student behavior (CASEL, 2013, Table 3). Early research found that the program improved social skills and reduced problem

43 Responsive Classroom was developed and is provided by the North East Foundation for Children. The program is composed of ten key practices: morning meeting; rule creation; interactive modeling; positive teacher language; logical consequences; guided discovery; academic choice; classroom organization; working with families; and collaborative problem-solving (see https://www.responsiveclassroom.org/).

44 CASEL (Collaborative for Academic, Social, and Emotional Learning) reviews SEL programs annually to identify effective SEL programs that promote student self-awareness, self-control, social awareness, relationship building, and problem solving. The criteria for the review are that the programs: are well-designed and classroom based; must provide repeated opportunities for students to practice new skills and behaviors; must offer training and implementation support; must be evidence-based (CASEL, 2013).

behaviors (Elliot, 1999; 1995; 1993). More recently, a quasi-experimental evaluation of Responsive Classroom found that the program positively impacted academic performance, teacher self-efficacy, children's social skills, and children's perception of school (Rimm-Kaufman et al., 2007; Rimm-Kaufman and Chiu, 2007).

For our benefit-cost evaluation we use evidence from a recent randomized control trial to evaluate the impact of Responsive Classroom on academic achievement in reading and math (Rimm-Kaufman et al., 2014). The sample included 24 schools in a large, diverse mid-Atlantic school district. The treatment group consisted of 13 schools and 1,467 students. Students entering third grade in 2008-2009, fourth grade in 2009-2010, and fifth grade in 2010-2011 were included as participants. Overall, the evaluation was not able to detect a statistically significant impact of the program directly on math or reading achievement at the end of fifth grade (Rimm-Kaufman et al., 2014: Table 4). However, there were statistically significant achievement gains when the program was implemented with fidelity. Structural equation modeling analyses were utilized to estimate the mediated effect of the program through fidelity of implementation (and mediated effect by student's initial ability). Fidelity was related to effect size gains of 0.26 *SD* in math and 0.30 *SD* in reading (Rimm-Kaufman et al., 2014: Table 5). For students with low initial math achievement, implementation of Responsive Classroom in treatment schools was related to an increase of 0.89 *SD* (*p<.01)* in math and 0.52 *SD* (*p<.05)* in reading (Rimm-Kaufman et al., 2014: Table 6). For initially high achieving math students, implementing Responsive Classroom with high fidelity in treatment schools was related to an increase of 0.49 *SD* (*p<.01)* in reading (Rimm-Kaufman et al., 2014: Table 6). Therefore, we perform benefit-cost analysis only for the schools in which the program was faithfully implemented, *with the express caution that our findings should not be viewed as representative of typical or average schools.*

Costs

No prior information on the costs of Responsive Classroom was available. Our approach was to estimate the ingredients needed to replicate the implementation from the evaluation by Rimm-Kaufman et al. (2014). We were able to obtain ingredients data from the evaluation team based on implementation observation data and other records related to the experiment.[45] In addition, supplementary information was collected from the developers regarding the training provided and how the program has evolved over time.

As per our framework, all cost estimates in our main analyses are incremental and exclude any costs associated with the conduct of research activities and reflect only program implementation costs corresponding to the years of the evaluation study. All costs are estimated in 2013 dollars and expressed in present values using a discount rate of 3.5%. Our data are representative of the whole sample of 13 treatment schools. However, we were unable to interview school-level staff. Because this information was not collected in full accordance with the ingredients method, we perform a series of sensitivity tests on costs.[46]

In Rimm-Kaufman et al. (2014), randomization occurred at the school level because the program targets the entire school including principals and teachers. The intent is to create an environment

45 We are grateful to Sara Rimm-Kaufman and Julia Thomas at the University of Virginia, and Philip Pohlmeyer and the staff at Responsive Classroom for their assistance.

46 In the ideal case, ingredients information should be collected directly from those personnel who are directly delivering the program or from skilled observers.

where teachers can grow together and to provide the support needed to take risks to advance their skills. The program was rolled-out over three years to longitudinally follow one cohort of students from 3rd through 5th grades.

Personnel ingredients represent the largest input category. Over the course of the three years, two one-week summer training sessions were provided to 3rd-5th grade teachers. These training sessions were supplemented with coaching, online support, and workshops provided by the developer. Principals also participated in the summer training sessions and also received ongoing support and coaching from the program developer. The schools that adopted Responsive Classroom also participated in the annual developer's conference, which provided additional professional development and networking opportunities.

Personnel costs were estimated using national prices available from the U.S. Department of Labor. To estimate a daily rate, the annual median salary of a principal or elementary teacher was divided by 180 days. The hourly rate assumes a seven hour day for teachers and an eight hour day for principals. If teachers on average spend five years teaching, we assume that the training will last for five years and amortize the cost of the time accordingly using a 3.5% rate. NEFC reported that teachers may remain in the profession longer as a result of participating in the program and working in a school that implements the program well. Thus, we include a sensitivity test that assumes teachers spend 10 years teaching.

Facilities ingredients include the space used for training sessions. All of the training sessions were held during the summer in district space, such as classrooms. We utilize the hourly rate for classroom space for all trainings.[47] We also include 9.19% of the average student expenditure in 2013, divided by 180 days, as overhead for the use of facilities to cover additional expenses that would not have otherwise occurred.[48] The daily overhead rate was amortized for five years using a rate of 3.5%.

Materials included books published by the developer that complement the training provided. Teachers and principals received books and a manual during training. Each school received a set of Responsive Classroom books for their library. The costs of these materials were based on their 2013 sticker prices, with a 20% discount for bulk purchase and amortized with a 3.5% rate assuming the life of each book would be five years.[49]

We were unable to ascertain the degree of displacement of other activities during the school day due to implementing Responsive Classroom. The program is implemented during the regular school day as a set of embedded practices. The evaluators and NEFC reported that the program provides guidance to teachers to improve classroom organization and schedule management. Through improved transitions and time management, the program does not displace any other class time. Because we were not able to obtain access to interview school-level staff to confirm this, we include a sensitivity analysis that incorporates additional costs for implementation time.

Cost information on Responsive Classroom is given in Table 9. The total cost of the program relative to business-as-usual is $1.32 million over three years. Across the 1,467 participants, this yields an average program cost of $900.

47 Please refer to the Positive Action section above for details.
48 Source: http://asumag.com/maintenance/36th-annual-maintenance-operations-cost-study-schools?page=2
49 Information from individual communication with NEFC.

Table 9

Costs for Responsive Classroom (3-Year Program)

Ingredient	Share of Costs
Participants (single cohort for 3 years)	*1467*
Personnel:	
Principals	13%
Teachers	70%
NEFC Training	14%
Facilities:	
Training space	1%
Materials and equipment:	
Books/manuals	1%
Total Resource Cost	**$1,316,250**
Average Cost per student (=TRC/n)	**$900**

Notes: Present Values at third grade (*d*=3.5%). National prices. 2013 dollars. Amounts rounded to $10.

We conducted four tests of the sensitivity of the costs per student estimate. First, we included costs for the evaluation personnel who observed the program during the study.[50] Two researchers (either study coordinators or graduate students) observed each teacher five times in class for 60 minutes. Including the time for these personnel increases the cost for the program by $30-$50 per student. Second, we modified our assumptions about the duration of the effects of the training. In the baseline, we assumed that training would be needed every five years as teachers exit the program. If Responsive Classroom reduces teacher turnover because of positive effects on teacher satisfaction, training may be needed less frequently. If we assume that the average time a teacher remains in the classroom was increased to 10 years, the cost per student would drop by $100. Third, we re-estimated the costs of the training using a market price for training instead of the prices paid to the developer. In fact, the costs of training by the developer at the time of the evaluation correspond closely to current market prices for general training. The cost estimate is therefore unchanged.

Finally, we included classroom space and the time that teachers spent providing a morning meeting for students during class time.[51] Although these meetings are only one aspect of the program, their implementation may impact the school day in a costly way. The morning meeting is a 20 minute class meeting designed to build community, teach social skills, and reinforce academic skills and is intended to be held daily. If the cost of teacher time and classroom space are included on the basis of this daily meeting during the school year, the cost per student for Responsive Classroom increases by $1,060 (to $1,960).

50 These costs were not included in our main analysis under the assumption that the observations did not result in any feedback for the teachers regarding their performance. The presence of the evaluators could have changed the instructional environment, perhaps through accountability pressures. However, the findings were not shared with the teachers or principals, so this sensitivity test is conservative.

51 These costs were not included in our main analyses because the teacher's time and the classroom space used during the class day are included in the base cost of schooling provided to students in both the treatment and control groups.

Benefits

For benefits, we estimate the economic value of achievement gains from Responsive Classroom. Although there was no overall program impact, when results are limited to strong implementation of Responsive Classroom, there was an effect size gain by 5th grade of 0.26 in math and 0.3 in reading (Rimm-Kaufman, 2014, Table 5). These fidelity-determined results are applied here under the justification that comparable effects and costs are being applied. That is, we have estimated costs for a faithfully implemented program that can be combined with the corresponding fidelity-adjusted outcomes. There are limitations with this approach. There could be many reasons why a school that implements Responsive Classroom with high fidelity would have greater gains in student achievement. Therefore, we are unable to know if the results are caused by the program or if they are due to some other difference between the schools. We recommend that these results be interpreted with caution, and they may not be generalizable.

The benefits map (Appendix I) shows that, based on evidence from other studies, there are other important outcomes besides achievement. These include understanding and handling of feelings (measured as social skills, social competence, and school perception) and delinquency, as well as corollary outcomes for teachers (self-efficacy and relationship with the child). However, it is not possible to monetize these impacts either because they are either outside the evaluation by Rimm-Kaufman (2014) or because no shadow prices are currently available.

Based on analysis of the association between achievement and earnings, we calculate the present value of earnings gains from a one standard deviation achievement gain of $34,300 in third grade (see Section 2). Therefore, with an effect size change in overall achievement of 0.26 (the lower bound of gains in math and reading), the value of the change in labor market attachment is $8,920. We note that this value assumes that the gains from Responsive Classroom are permanent (until age 18), even though many interventions experience test score fade-out (e.g. Heckman and Kautz, 2012).

Benefit-Cost Results for Responsive Classroom

Table 10 shows the benefit-cost results for Responsive Classroom on a per participant basis. With costs of $900 and benefits of $8,920, the net present value per 100 students is $802,000.

Table 10

Benefit-Cost Results for Responsive Classroom Per 100 Participants (high fidelity implementation only)

	Costs	Benefits	Net Present Value
Baseline	$90,000	$892,000	$802,000
Sensitivity tests:			
S1. Costs include evaluator observations	$95,000	$892,000	$797,000
S2. Costs with implementation	$196,000	$892,000	$696,000
S3. Fade-out of 10% pa	$90,000	$384,000	$294,000
S4. Fade-out of 25% pa	$90,000	$89,000	($1,000)
Average across S1-S5			$446,500

Sources: Table 9 above. *Notes:* Present values at third grade (d=3.5%). Baseline model assumes baseline costs and baseline benefits. Sensitivity tests assume costs or benefits from baseline model unless other specified.

Given the very large baseline benefits, we perform sensitivity tests that show lower bounds to Responsive Classroom. Applying the higher cost estimates described above, the net benefits are reduced. Nevertheless, even including the costs of morning meetings, the benefits exceed the costs by $696,000 per 100 participants. The key sensitivity parameter is the fade-out of benefits from achievement gains. If the achievement gains fade-out by 10% per year until age 18, the benefits are reduced to $3,840. When the fade-out function reaches 25% per year, then the benefits are equal to the costs. Thus, for programs of this scale, the break-even fade-out rate is approximately one-quarter. Looking across the four (conservative) sensitivity tests, the net present value per 100 students is $446,500. However, if we evaluate the program based only on the overall average effects on achievement, the costs would outweigh the benefits. Future studies of the effect of SEL programs on achievement outcomes would benefit from measuring SEL outcomes, such as those in Appendix I, rather than limiting them to standardized test performance.

4.6 Social and Emotional Training

Program

For contrast with the previous programs from the U.S., we decided to apply our methods to an extensive program that was highly documented and evaluated in Sweden. Our final benefit-cost analysis is of Social and Emotional Training (SET) for a program implemented and evaluated for Swedish school children. The SET program is a classroom-based intervention for grades 1-9 designed to support students' cognitive and social and emotional competencies, learning and development. The program has a similar curriculum to the U.S. Providing Alternative Thinking Strategies (PATHS) program (cite).

The SET program has been evaluated using a longitudinal quasi-experimental design at several different time-points, and the findings show positive effects for the program on a number of social and emotional outcomes. The evaluation addressed 41 classes and 52 teachers at two schools implementing SET and 14 classes at one control school located in the same neighborhood. The impacts of the program are primarily focused on social and emotional competencies linked to mental health issues, although reduced drug use was also considered for students in the older grades. Positive impacts from SET were identified both after two and five years of the program (Kimber, Sandell and Bremberg, 2008; Kimber and Sandell, 2009).

To calculate the costs and benefits of the SET program, the impacts from a five-year follow up (Kimber and Sandell, 2009) are used. In a five-year follow up (Kimber, Sandell and Bremberg, 2008) the results showed positive and significant effects on five of seven variables for the SET-students: internalizing problems (*d*=. 56), externalizing problems (*d*= .42), mastery as reflecting self-efficacy or hopelessness (*d*= .36), self-image and self-esteem (*d*= .54), contentment in school (*d*= 60) and bullying (*d*= 35). Effect sizes were medium and no relationship was found between the treatment group and the promotion of social skills (*d*= .07).[52] The cohort size for the intervention was 1,028 youth, although 837 students participated in the evaluation (663 treatment, 174 control).

52 In the two-year follow-up, there were impacts in terms of: in grade 3, psychological well-being (d= .95) and ability (d= .56); and in grades 4-9, body image (d= .48), psychological well-being (d= .33), aggressiveness (d= .32), relation with others (d= 33), attention seeking (d= .32) and bullying (d= .39). In addition, there were significant positive effects for the program on alcohol abuse (d= 0.26) and drug use (d= 0.23).

Costs

No prior information on the costs of the SET program is available. The costs were estimated by using the ingredients method, and information about the ingredients was collected by interviewing the personnel who implemented the program and from information materials on the program (Kimber, 2001). U.S. prices were used.

This evaluation has estimated the total costs for the entire program over the five years, reported in Table 3 of Appendix II. The cost of the program is calculated for the 1,028 students who participated at the outset of the program.[53] The costs are in 2013 prices and a discount rate of 3.5% is used. Both the operating costs such as teacher salaries and administrative costs and capital costs such as rent costs for classrooms, facilities and materials have been estimated.

Personnel categories include teachers, principals, assistant principals, counselors and program developers. Teacher input was counted for curriculum delivery, initial (two full days) training and ongoing training. Hours of teacher training were collected from the program developer who implemented the program. Teachers in grades 1-5 delivered two 45 minute sessions of SET instruction (the program) every week, each year of the program, a total of 40 sessions per year. In grades 6-9, teachers delivered 45 minutes sessions of SET instruction every week, each year of the program, a total of 40 sessions per year. Teachers received initial and ongoing training for 33.5 hours the first year. In the second year some teachers received additional training and, due to teacher turnover, new teachers received training separately. Training was provided by the program counselor. Data on the national average salaries for teachers in compulsory school, special education teachers, counselors, assistant principals and principals has been taken from Statistics Sweden (Statistics Sweden, 2013).

Facilities for the SET program included school space and training space. The two-day initial training was held in an auditorium at one of the schools. The ongoing training was also held in the facilities at the schools by using classrooms or other smaller meeting-rooms.[54] The costs for facilities are based on the yearly rent in 2013 prices that the two schools pay to the municipality.

The material used in the SET program was a manual used by the teachers which was copied at the schools. Students created their own work-books using paper and other materials provided. The price for the paper was collected from several companies selling paper, and an average was calculated. No travel costs were incurred for training.

In Table 11 the costs are presented. For the five-year program, the present value of total cost is $555,260 and the average cost per student is $540. However, if we exclude the cost for teacher instruction time, the total cost is $143,000 and the average cost falls to $140.

We test this cost estimate for sensitivity to the cohort size. Specifically, as the intervention was delivered to a cohort of only 837 (out of the 1,028 intended participants), then the costs may be apportioned across this number of youth. If the cohort size falls to 837, the average cost rises to either $660 or $170.

Benefits

To calculate the benefits we use the impacts on drug use reported for students in grades 7-9 in Kimber et al. (2009). For drug use, students were assessed at three time points and divided into zero/light

53 Turnover affected the number of students who participated in the program for the full the five years, although the sample of students in the classes remained stable.

54 The costs for classroom were calculated from the 2013 rent costs to a square meter cost per hour. On average, a classroom is 56 square meters and a class size of 25 students.

Table 11

Costs for SET (5-Year Program)

Ingredient	Share of Costs
Participants	*1,028*
Personnel	
Teachers (Training/ongoing coaching)	16%
Teachers (SET instruction)	74%
Administration (Training/support/meetings)	3%
Facilities	
Auditorium	<1%
Workshop space	<1%
Classrooms	5%
Materials/equipment	
SET Manual	<1%
Supplemental materials	1%
Total Resource Cost 1 (with teacher SET instruction)	**$555,260**
Total Resource Cost 2 (without teacher SET instruction)	**$143,000**
Average Cost 1 per student (=TRC2/1,028)	**$540**
Average Cost 2 per student (=TRC1/1,028)	**$140**

Notes: Discounted by 3.5% to year 1. Prices in U.S. dollars (2013) $1=6SEK. Amounts rounded to $10.

users and heavier users. For the zero/light group, students receiving the SET treatment reported a decrease in drug use of 5 percentage points (6% to 1%) from year 1 to 5; zero/light students in the control school reported an increase of 6 percentage points (2% to 8%) from year 1 to 5. This difference equates to an effect size gain of d=0.64. For the heavier users, there was an increase for SET treatment group from 1% to 5% and for the control group from 1% to 15% (an effect size, d=0.32).[55]

Our benefits map for SET shows many other possible benefits from the program (Appendix I). However, these benefits could not be included in this analysis. Some benefits, such as alcohol, smoking and volatile substances abuse, are likely to be directly confounded with our selected measure of substance abuse. Inclusion of these other benefits would therefore lead to double-counting. Other impacts, such as delinquency and behavior measures of mental health, might also be indirectly confounded with drug use even if these impacts could be accurately shadow priced. Finally, none of the array of social competencies that SET promotes (relations with others, etc.) have shadow prices available and therefore cannot be monetized.

For drug use, we base our shadow prices on what society currently spends on these behaviors through the health care, criminal and judicial systems. For this analysis we use the calculations made

55 Nationally, among males [females] age 15-16, 2% [1%] are drug abusers (illicit drug use during the past 30 days). The proportion raises with age, for 17-18 year-olds the rates are 6% and 2% respectively (The National Board of Health and Welfare, 2013).

by Nilsson and Wadeskog (2013); other spending estimates are given by Nilsson and Wadeskog (2008) and The National Board of Health and Welfare (2013). These estimates are conservative in that they do not include certain costs such as individual losses of income.

Across Sweden, the estimated annual burden of drug use was $3.9 billion ($26 billion kronor) in 2011; this amounts to $450 per capita nationally (SOU, 2011). Of this aggregate amount, 42% was indirect losses of production due to sick leave and premature death, 27% was for spending on the criminal justice system, 26% was for health and social care treatments, and the remaining 5% was for insurance and private health care. Per drug user, the present value of the social burden is estimated at $102,920 in 2013 dollars. Given the respective proportions of youth who are drug users in the intervention versus comparison group, there is a net reduction of 0.0982 drug users.[56] This translates into a benefit of $7,510.

Benefit-Cost Results for SET

The net benefits for the SET program are presented in Table 12. The baseline estimates per participant are of costs at $540 and benefits at $7,510, yielding a net present value per 100 participants of $697,000. The SET program easily passes a benefit-cost test: the program is relatively inexpensive per participant; it is highly effective for the population of substance abusers; and the economic burden per substance abuser is very large.

As shown in the bottom panel of Table 12, the SET program is unlikely to have a benefit-cost ratio that is less than one. Under the assumption that the cohort is smaller than expected (837 participants instead of 1,028), and therefore program costs are higher, the net present value per 100 participants is $685,000. Counting only the benefits from heavy users, the benefits are only reduced slightly and the net present value per 100 participants is $322,000. Finally, even if the fade-out rate is 60% within the first year, the net present value per 100 participants is still over $200,000. Moreover, the net present value is likely to be even greater than reported here, given that only the public burden of substance abuse is included (and not the private burden).

Table 12

Benefit-Cost Results for SET Per 100 Participants

	Costs	Benefits	Net Present Value
Baseline	$54,000	$751,000	$697,000
Sensitivity tests:			
S1. Smaller cohort of students	$66,000	$751,000	$685,000
S2. Heavy users only	$54,000	$376,000	$322,000
S3. 60% fade-out in year 1	$54,000	$300,000	$246,000

Sources: Table 11 above. *Notes:* Present values (d=3.5%) in 2013 dollars.

56 Calculation based on 90% light drug usage (1% treatment, 8% control) and 10% heavy drug usage (5% treatment, 15% control). Of the population, 4% are assumed to be drug users at age 17-18 (The National Board of Health and Welfare, 2013)

5. DEVELOPING METHODS AND EVIDENCE FOR ECONOMIC EVALUATIONS OF SEL

The above analyses indicate that SEL interventions can easily pass a benefit-cost test. In fact, the weighted average benefit-cost ratio across all six interventions with prior evidence of effectiveness indicates that identified benefits outweigh the costs by a factor of 11:1, with an average net present value per 100 participants of $618,380. This finding suggests that had we been able to obtain measures and monetary values of the full range of benefits identified in each of the benefit maps in Appendix I, the benefits would exceed costs by even larger magnitudes. However, a broader conclusion from these analyses is that – if economic evaluations are to have an important influence on the development of SE provision – then significant changes in current evaluation practice are needed (see also Belfield and Levin, 2014). These changes are summarized in Box 5.1.

Box 5.1
Recommendations for Evaluation Practice

Costs Analysis
1. Report all intervention inputs
2. Specify how intervention differs from 'business-as-usual'

Benefits Analysis
1. Draw benefit maps
2. Identify which impacts can and cannot be monetized
3. Identify how impacts overlap

Methodology
1. Calculate shadow prices for behaviors
2. Examine resource consequences for schools with students with low SE skills

A second concern is with the measurement of impacts and hence benefits. Across SE interventions, outcomes are rarely measured in a consistent way. As our case studies illustrate, evaluations vary in how many domains are measured and what scales are used for measurement. This can lead to benefit-cost analyses where simply the number of benefits varies (regardless of their value). It also means that, instead of determining the appropriate way to measure benefits and to apply shadow pricing techniques, the analyst must build each economic case in idiosyncratic fashion. With more classification and clarification of impacts, many more shadow prices could be estimated and more benefit-cost analysis performed. In addition, it is not clear how persistent the outcomes are. Some programs may affect outcomes in a very 'fast-acting' way. Some programs may have effects that endure over a long time period. Both types of programs will be much more beneficial than where effects are delayed or transitory. At present, these ratchet and fade-out attributes cannot be identified from the available evidence and it would therefore be preferable to have outcome data measured over a protracted period of time. All these attributes matter in economic evaluations, and without them it is very difficult to make comparisons across programs.

Applying a direct economic perspective, the main concern is that the methods by which impacts can translate into money benefits is underdeveloped. Even across the six broad categories of outcomes applied in reviews there is lack of clarity as to how outcomes overlap or confound. For example, high achievement may cause high social and emotional skills, e.g. by raising self-respect and a sense of control; the opposite association may hold (as argued by Heckman and Kautz, 2012); or they may be mutually determined. Evaluating an intervention in terms of both gains in achievement and social and emotional skills may therefore be double-counting. Also, it is not clear if SEL outcomes can be expressed monotonically as benefits. That is, when two interventions enhance SEL outcomes by 5% and 10% respectively, it cannot necessarily be assumed that the benefits from the latter are twice those of the former intervention. Similarly, the benefits of SEL may vary across the ability spectrum: improving SEL skills for at-risk students may have very different economic consequences from improving SEL skills across the general population.

There are two clear areas where more shadow pricing would help. First, at this time, there is no mechanism for converting the Teacher Rating Scale (TRS) of the BASC into economic metrics. The BASC-TRS is a widely-used, validated scale for child behavior, and yet its economic implications have not been investigated at all to our knowledge. Second, there is limited evidence on the resource implications within schools from low SE skills (especially negative social behavior or student conduct problems). If SE skills are low and create disruptions and learning problems, both schools and public welfare agencies will require more resources: to hire more counselors or social services personnel; to compensate current personnel who are working with more delinquent students; and to cover other related expenses such as injuries, sick pay, insurance payments, and recruitment costs due to increased quits (Pouliakis and Theodossiou, 2013).[57] Itemized thus, it is seems likely that improving SE skills could yield sizeable cost savings. For example, serious behavior problems can result in expensive special education services and retention in grade or referral to other social agencies. But as yet the dollar amounts of these behaviors are unknown.

In addition, existing shadow prices need to be validated. Even in two areas where some progress has been made – on the association between SE skills and earnings or education – the need for substantial further investigation remains. Alternative scales need to be applied to identify the labor market gains from social and emotional skills more precisely and to model how these gains and skills are mutually determined (Heckman et al., 2006). Also, more emphasis should be put on the impact of SE skills on attainment (such as graduation from high school and college) rather than achievement; the economic value of attainment is much more robust than of achievement, not least because of 'fade-out' in cognitive gains.

To address these issues, a range of approaches need to be applied. On the costs side, the ingredients method of calculation is straightforward (although we have noted that it is rarely applied to SE interventions). Moreover, there are two important analytical issues. First, few evaluations give descriptions in comparative terms, i.e. how the intervention differs from business-as-usual. If treatment contrast is not well-documented, it is very difficult to calculate incremental costs. The primary example where this difficulty arises is with in-class treatments and the extent to which these are a substitute for, or addition to, regular classroom instruction. Second, retrospective calculation of cost data is very

57 Evidence from Goldhaber et al. (2010) shows that these compensating wage differentials may be difficult to estimate in practice using wage equations. An alternative technique using contingent valuation methods has not yet been applied, nor has analysis of school spending patterns according to student behavior.

difficult. Persons with knowledge of how resources are allocated are often unavailable or are unable to describe program ingredients in detail. During our investigation we struggled to find subject interviewees who could describe the interventions which had been evaluated. This might not matter if SE programs were stable such that current versions of the program resemble past versions. Also, it might not matter if SE programs were always improved with fidelity to program design. Unfortunately, neither condition is likely to hold. This makes cost estimation less precise.

On the benefits side, we recommend the use of benefit maps for each evaluation. These maps show which impacts are measured; whether these impacts are measured consistently; and if there is overlap or confounding across impacts. In turn, we can identify which impacts can be assigned shadow prices, which might be assigned shadow prices, and which cannot. Indeed, the benefit maps we have produced thus far highlight differences across evaluations, and these differences are sufficiently large that we caution, strongly, against any comparison of benefit-cost ratios across the field of social and emotional learning when the ability to capture the benefits among interventions is so variable. Also, we recommend more attention to the wider benefits of SEL. Unavoidably given the current evidence, studies focus on the individual-level benefits of SEL. But there is a strong case that SEL benefits are dispersed through a school or community and so economic value should be measured at that level. A better understanding of the economic value of school climate or classroom behavior would help bring this idea to the forefront.[58]

Overall, there are many caveats in comparing interventions: student groups differ; and a different set of outcomes are measured using different metrics over different time horizons. However, we emphasize that these caveats are not specific to benefit-cost analysis. They apply to all evaluation research that attempts to compare interventions. In fact, benefit-cost analysis serves to demonstrate the full heterogeneity of the research evidence on social and emotional learning, as well as highlight important gaps in the knowledge base. Equally importantly, this inquiry into the costs and benefits of specific SE interventions suggests that many such interventions might easily pass an economic test of a positive return to investment if sufficient and adequate data were available. Given constraints on education funding, passing this test is an important justification for making further commitments to enhance social and emotional learning.

58 We appreciate this suggestion from Damon Jones.

REFERENCES

Abramson, P. 2011. *16th Annual School Construction Report*. http://webspm.com/research/2011/02/school-construction-report/asset.aspx.

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

Bavarian, N., Lewis, K., DuBois, D., Acock, A., Vuchinich, S., Silverthorn, N., Snyder, F., Day, J., Ji, P., Flay, B.R. (2013) Using social-emotional and character development to improve academic outcomes: A matched-pair, cluster-randomized controlled trial in low-income, urban schools. *Journal of School Health,* 83(11), 771-779.

Beets, M.W., Flay, B.R., Vuchinich, S., Snyder, F.J., Acock, A., Li K-K, Burns, K., Washburn, I.J., and J. Durlak. 2009. Use of a social and character development program to prevent substance use, violent behaviors, and sexual activity among elementary-school students in Hawai`i. *American Journal of Public Health*, 99, 1438-1445.

Belfield, C.R. and H.M. Levin. 2009. Some economic consequences of improving mathematics reform. Monograph, Center for Benefit-Cost Studies in Education, www.cbcse.org.

Belfield, C.R. and H.M. Levin. 2014. Methodological issues in cost-effectiveness analysis. *Journal of Research in Educational Effectiveness*, forthcoming. Boardman, A. E., Greenberg, D. H., Vining, A. R., and Weimer, D. L. 2011. *Cost-Benefit Analysis: Concepts And Practice* (4th ed). Upper Saddle River, N.J.: Prentice Hall.

Boardman, A. E., Greenberg, D. H., Vining, A. R., & Weimer, D. L. 2011. *Cost-benefit analysis: Concepts and practice* (4th ed.). Upper Saddle River, New Jersey: Pearson Prentice Hall.

Borghans, L., Duckworth, A.L., Heckman, J.J. and B. ter Weel. 2008. The economics and psychology of personality traits. *Journal of Human Resources*, 43, 972-1059.

Botvin, G. J., Griffin, K. W., and Nichols, T. D. 2006. Preventing youth violence and delinquency through a universal school-based prevention approach. *Prevention Science,* 7(4), 403-408.

Castex, G. and E. Dechter. 2013. The changing role of education and ability in wage determination. *Journal of Labor Economics*, forthcoming.

Cohen, M.A. and A.R. Piquero. 2009. New evidence on the monetary value of saving a high risk youth. *Journal of Quantitative Criminology, 25*, 25-49.

Cohen, M.A., Piquero, A.R., and W.G. Jennings. 2010. Estimating the costs of bad outcomes for at-risk youth and the benefits of early childhood interventions to reduce them. *Criminal Justice Policy Review*, *21*, 391-434.

Collaborative for Academic, Social, and Emotional Learning (CASEL). 2013. *CASEL Guide: Effective Social and Emotional Learning Programs*. http://static.squarespace.com/static/513f79f9e4b05ce7b70e9673/t/526a220de4b00a92c90436ba/1382687245993/2013-casel-guide.pdf

Cooke, M. B., Ford, J., Levine, J., Bourke, C., Newell, L., & Lapidus, G. 2007. The effects of city-wide implementation of Second Step on elementary school students' prosocial and aggressive behaviors. *The Journal of Primary Prevention, 28*(2), 93–115.

Crowley, D. M., Jones, D.E., Coffman, D.L. and Greenberg, M.T. 2014. Can we build an efficient response to the prescription drug abuse epidemic? Assessing the cost effectiveness of universal prevention in the PROSPER trial. *Preventive Medicine, 62*: 71-77.

de Araujo, P. & S. Lagos. 2013. Self-esteem, education, and wages revisited. *Journal of Economic Psychology*, 34, 120-132

Drago, F. 2011. Self-esteem and earnings. *Journal of Economic Psychology*, 32(3), 480–488.

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

Durlak, J.A., Weissberg, R.P., and M. Pachan. 2010. A meta-analysis of after-school programs that seek to promote personal and social skills in children and adolescents. *American Journal of Community Psychology 45*:294-309.

Edwards, D., Hunt, M. H., Meyers, J., Grogg, K. R., & Jarrett, O. 2005. Acceptability and student outcomes of a violence prevention curriculum. *Journal of Primary Prevention, 26*(5), 401–418.

Elliot, S.N. 1993. *Caring to Learn: A Report on the Positive Impact of a Social Curriculum*. http://www.responsiveclassroom.org/past-research

Elliot, S.N. 1995. *The Responsive Classroom Approach: Its Effectiveness and Acceptability*. http://www.responsiveclassroom.org/past-research

Elliot, S.N. 1999. *A Multi-year Evaluation of the Responsive Classroom Approach: Its Effectiveness and Acceptability in Promoting Social and Academic Competence*. http://www.responsiveclassroom.org/past-research.

Eren, O. and S. Oxbeklik. 2013. The effect of noncognitive ability on the earnings of young men: A distributional analysis with measurement error correction. *Labour Economics,* 24, 293-304.

Espelage, D., Low, S., Polanin, J., and E. Brown. 2013. The impact of a middle school program to reduce aggression, victimization, and sexual violence. *Journal of Adolescent Health* 53, 180-186.

Flay, B.R. 2012. Strong quasi-experimental Evaluation of the Positive Action Pre-K program. Unpublished report, Oregon State University.

Flay, B.R. 2010. The Positive Action family program: a pilot randomized trial and replication. Unpublished report, Oregon State University.

Flay, B.R., Allred, C.G., and Ordway, N. 2001. Effects of the Positive Action program on achievement and discipline: Two matched –control comparisons. *Prevention Science, 2*(2), 71-90.

Flay, B.R. & Allred, C.G. 2003. Long-term effects of the Positive Action program. *American Journal of Health Behavior, 27*(Supplement 1): S6-S21.

Flay, B., Acock, A., Vuchinich, S., & Beets, M. 2006. Progress report of the randomized trial of Positive Action in Hawaii: End of third year of intervention. Available from Positive Action, Inc. 264 4th Avenue South, Twin Falls, ID 83301.

Fletcher, J. 2013. Adolescent depression and adult labor market outcomes. *Southern Economic Journal,* 80, 26-49.

Foster, E.M., Jones, D., & the Conduct Problems Prevention Research Group. 2005. The high costs of agression: Public expenditures resulting from conduct disorder. *American Journal of Public Health,* 95, 1767-1772.

Frey, K., Nolen, S., Van Schoiack Edstrom, L., & Hirschstein, M. 2005. Effects of a school-based social–emotional competence program: Linking children's goals, attributions, and behavior. *Journal of Applied Developmental Psychology, 26*(2), 171–200.

Goldhaber, D. Destler, K. and D. Player. 2010. Teacher labor markets and the perils of using hedonics to estimate compensating wage differentials. *Economics of Education Review*, 29, 1-17.

Gottfried, M.A. 2010. Evaluating the Relationship between Student Attendance and Achievement in Urban Elementary and Middle Schools: An Instrumental Variables Approach. *American Educational Research Journal, 47,* 434-465.

Gottfried, M.A. 2011a. The Detrimental effects of missing school: Evidence from urban siblings. *American Journal of Education, 117,* 147-182.

Gottfried, M.A. 2011b. Absent peers in elementary years: The negative classroom effects of unexcused absences on standardized testing outcomes. *Teachers College Record, 113.*

Griffin, K.W., Botvin, G.J., and Nichols, T.R. 2006. Effects of a school-based drug abuse prevention program for adolescents on HIV risk behavior in young adulthood. *Prevention Science,* 7(1), 103-112.

Griffin, K.W., Samuolis, J., and Wiliams, C. 2011. *Journal of Child and Family Studies, 20*: 319-325.

Hawkins, J.D., Herrenkohl T., Farrington D.P., Brewer D., Catalano R.F., and T.W. Harachi. 1998. A review of predictors of youth violence. In Loeber R., and D. P. Farrington (Eds). *Serious and Violent Juvenile Offenders: Risk Factors and Successful Interventions.* Thousand Oaks, CA: Sage Publications.

Hawkins, J.D., Kosterman, R., Catalano, R.F., Hill, K.G. and R.D. Abbott. 2005. Promoting positive adult functioning through social development intervention in childhood: Long-term effects from the Seattle Social Development Project. *Archives of Pediatrics & Adolescent Medicine, 159,* 25-31.

Heckman, J.J. and P. LaFontaine. 2006. Bias-corrected estimates of GED returns. *Journal of Labor Economics, 24,* 661-700.

Heckman, J. J., Stixrud, J., & S. Urzua. 2006. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics, 24*(3), 411-82.

Heckman, J.J. and T. Kautz. 2012. Hard evidence on soft skills. *Labour Economics,* 19, 451-464.

Holsen, I., Smith, B., & Frey, K. S. 2008. Outcomes of the social competence program *Second Step* in Norwegian elementary schools. *School Psychology International, 29*(1), 71–88.

IOM (Institute of Medicine) and NRC (National Research Council). 2014. Considerations in applying benefit-cost analysis to preventive interventions for children, youth, and families: Workshop summary. Washington, DC: The National Academies Press.

Jacob, B. A., Lefgren, L., and D.P. Sims. 2008. The persistence of teacher-induced learning. *Journal of Human Resources, 45*(4), 915-943.

Jones, D., Bumbarger, B.K., Greenberg, M.T., Greenwood, P. and S. Kyler. 2008. PCCD's Investment in Research-based Programs: A Cost-Benefit Assessment of Delinquency Prevention in Pennsylvania. Monograph, The Prevention Research Center for the Promotion of Human Development, Pennsylvania State University.

Jones, D., Foster, E.M., & the Conduct Problems Prevention Research Group. 2009. Service use patterns for adolescents with ADHD and comorbid conduct disorder. *Journal of Behavioral Health Services and Research*, PMID#: 18618263.

Jones, S.M., Brown, J.L. & J.L. Aber. 2011. Two-year impacts of a universal school-based social-emotional and literacy intervention: an experiment in translational developmental research. *Child Development, 82(2),* 533-554.

Kimber, B. (2001a). Livsviktigt, SET, Student Book. Stockholm: Ekelunds förlag.

Kimber, B. (2001b). Livsviktigt, SET, Teacher Manual. Stockholm: Ekelunds förlag.

Kimber, B., Sandell, R., & S. Bremberg. 2008. Social and emotional training in Swedish classrooms for the promotion of mental health: results from an effectiveness study in Sweden. *Health Promotion International, 23*(2), 134-143.

Kimber, B. and R. Sandell. 2009. Prevention of substance use among adolescents through social and emotional training in school: A latent class analysis of a five-year intervention in Sweden. *Journal of Adolescence, 32,* 1403-1413.

Kinsler, J. 2012. Beyond levels and growth: Estimating teacher value-added and its persistence. *Journal of Human Resources, 47*(3), 722-763.

Kuklinski, M.R., Briney, J.S., Hawkins, J.D., & R.F. Catalano. 2012. Cost-benefit analysis of *Communities That Care* outcomes at eighth grade. *Prevention Science, 12(2),* 150-161.

Lee, S., Aos, S, Drake, E., Pennucci, A, Miller, M., Anderson, L., and M. Burley. 2012. Return on investment: Evidence-based options to improve state-wide outcomes. Technical Appendix Methods and User Manual, www.wsipp.wa.gov.

Levin, H.M. 1975. Cost-effectiveness analysis in evaluation research. In M. Guttentag, & E. L. Struening (Eds.), *Handbook of evaluation research (Volume 2)*. Beverly Hills, CA: Sage.

Levin, H.M. 2012. More Than Just Test Scores. *Prospects*, UNESCO. 43(2), 269-284.

Levin, H.M. & P.J. McEwan. 2001. *Cost-effectiveness Analysis: Methods and Applications* (2nd ed.). Thousand Oaks: Sage Publications.

Lewis, K.M., Bavarian, N., Snyder, F., Acock, A.C., Washburn, I.J., and Flay, B.R. 2012. Direct and mediated effects of a social-emotional and character development program on adolescent substance use. *The International Journal of Emotional Education, 4*(1), 56-78.

Lewis, K., Schure, M., Bavarian, N., DuBois, D., Day, J., Ji, P., Silverthorn, N., Acock, A., Vucinich, S., & B. Flay. 2013a. Problem behavior and urban, low-income youth: a randomized controlled trial of Positive Action in Chicago. *American Journal of Preventative Medicine* 44(6), 622-630.

Lewis, K., Schure, M., Bavarian, N., DuBois, D., Day, J., Ji, P., Silverthorn, N., Acock, A., Vucinich, S., & B. Flay. 2013b. Effects of the Chicago trial of Positive Action on youth problem behaviors. *American Journal of Preventative Medicine, 44,* 622-630.

Li, K.K., Washburn, I., DuBois, D.L., Vuchinich, S., Ji, P., Brechling, V., Day, J., Beets, M.W., Acock, A.C., Berbaum, M., Snyder, F., and Flay, B.R. 2011. Effects of the Positive Action programme on problem behaviors in elementary school students: A matched-pair randomized control trial in Chicago. *Psychology & Health, 26*(2), 187-204.

Long, K., Brown, J.L., Jones, S.M., J.L. Aber L., & B.T. Yates. 2014. Cost Analysis of a School-Based Social and Emotional Learning and Literacy Intervention. Working Paper, Fordham University.

McEwan, P.J. 2002. Are cost-effectiveness methods used correctly? In Levin, H.M. & McEwan, P.J. (Eds.), *Cost-Effectiveness and Educational Policy: 2002 Yearbook of the American Education Finance Association* (p.37-53). Eye on Education: Larchmont, NY.

Maynard R.A. & S. Hoffman (Eds.), *Kids having kids*. Washington, DC: The Urban Institute Press.

Miller, T. and D. Hendrie. 2008. *Substance Abuse Prevention Dollars and Cents: A Cost-Benefit Analysis*. DHHS Pub. No. (SMA) 07-4298. Rockville, MD: Center for Substance Abuse Prevention, Substance Abuse and Mental Health Services Administration.

Missbruksutredningen. 2011. *Missbruket, Kunskapen, Vården: Missbruksutredningens forskningsbilaga*. Stockholm: Frizes, 2011.

Montague, M. Enders, C. Cavendish, W., and M. Castro. 2011. Academic and Behavioral Trajectories for At-Risk Adolescents in Urban Schools. *Behavioral Disorders* , 36, 2 , 141-156.

Murnane, R.J., Willett, J.B., Braatz, M.J. & Y. Duhaldeborde. 2001. Do different dimensions of male high school students' skills predict labor market success a decade later? Evidence from the NLSY. *Economics of Education Review*, 20, 311-320.

Nas, T.F. 1996. *Cost-Benefit Analysis: Theory and Application*. Sage Publications: Thousand Oaks, CA.

NCES. 2013. //ies.ed.gov/ncee/wwc/pdf/intervention_reports/wwc_secondstep_030513.pdf.

Nilsson, I., & A. Wadeskog. 2008. Varje drogfri dag en framgång – Socioekonomiskt Bokslut för Metadonprogrammet i Stockholm. Stockholm: SEE.

Nilsson, I., & Wadeskog, A. (2013). Trygghetens värde – sociala risker ur ett ekonomiskt perspektiv [The value of security – social risks from an economic perspective]. Retrieved May 5, 2014, from: http://www.seeab.se/Dalarna/Dalarna_2013.pdf.

Ozkan E. & S. Ozbeklik. 2013. The effect of non-cognitive ability on the earnings of young men: A distributional analysis with measurement error correction. *Labour Economics*, 24, 293-304.

Pouliakis, K. and I. Theodossiou. 2013. The economics of health and safety at work: An interdisciplinary review of the theory and policy. *Journal of Economic Surveys*, 27, 167-208.

Reynolds, C.R., and R.W. Kamphaus. 2002. *A Clinician's Guide to the Behavior Assessment System for Children (BASC)*. New York: Guilford Publications.

Rimm-Kaufman, S.E., Larsen, R.A.A, Baroody, A.E., Curby, T.W., Ko, M., Thomas, J.B., Merritt, E.G., Abry, T., DeCoster, J. 2014. Efficacy of the Responsive Classroom approach: Results from a 3-year, longitudinal randomized controlled trial. *American Educational Research Journal, 51*(3), 567-603.

Rimm-Kaufman, S.E., & Chiu, Y.I. 2007. Promoting social and academic competence in the classroom: An intervention study examining the contribution of the *Responsive Classroom* approach. *Psychology in the Schools, 44*(4), 397-413.

Rimm-Kaufman, S.E., Fan, X., Chiu, Y.I., & You, W. 2007. The contribution of the *Responsive Classroom* approach on children's academic achievement: Results from a three year longitudinal study. *Journal of School Psychology, 45,* 401-421.

Schick, A., & Cierpka, M. 2005. Faustlos: Evaluation of a curriculum to prevent violence in elementary schools.*Applied and Preventive Psychology, 11,* 157–165.

Segal, C. 2013. Misbehavior, education, and labor market outcomes. *Journal of European Economics Association*, 11, 743-779.

Sklad, M., Diekstra, R., De Ritter, M., Ben, J., and C. Gravestein. 2012. Effectiveness of school-based universal social, emotional, and behavioral programs: Do they enhance students' development in the are of skill, behavior, and adjustment? *Psychology in the Schools,* 49, 892-910.

Social and Character Development Research Consortium [SCDRC]. 2010. *Efficacy of Schoolwide Programs to Promote Social and Character Development and Reduce Problem Behavior in Elementary School Children.* (NCER 2011-2011). Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.

SOU (2011:6). *Missbruket, kunskapen, vården: Missbruksutredningens forskningsbilaga* [Drug Abuse, knowledge and medicin: The drug abuse investigation´s appendix]. Stockholm: Frizes.

Spoth, R. L., Randall, G. K., Trudeau, L., Shin, C., and Redmond, C. 2008. Substance use outcomes 5 1/2 years past baseline for partnership-based, family-school preventive interventions. *Drug and Alcohol Dependence,* 96, 57-68.

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

Snyder, F., Flay, B., Vuchinich, S., Acock, A., Washburn, I., and B.R. Flay. 2012. Improving elementary school quality through the use of a social-emotiional and character development program: a matched pair, cluster randomized, controlled trial in Hawai`i. *Journal of School Health*, 82, 11-20.

Snyder, F.J., Acock, A.C., Vuchinich, S., Beets, M.W., Washburn, I.J, and Flay, B.R. 2013. Preventing negative behaviors among elementary-school students through enhancing students' social-emotional and character development. *American Journal of Health Promotion, 28*(1).

Statistics Sweden (2013). Beskrivande data 2013 [Descriptive statistics 2013]. Retrieved February 6, 2014, from http://www.scb.se/sv_/Hitta-statistik/Statistik-efter-amne/Arbetsmarknad/Loner-och-arbetskostnader/Lonestrukturstatistik-primarkommunal-sektor/7627/7634/2012/28336/

The Swedish National Board of Health and Welfare, (2012). *The Public health report*. Stockholm: The Swedish Board of Health and Welfare.

Waddell, G. R. 2006. Labor-market consequences of poor attitude and low self-esteem in youth. *Economic Inquiry*, 44, 69–97.

Washburn, I.J., Acock, A., Vuchinich, S., Snyder, F.J., Li, K.K., Ji, P., et al. 2011. Effects of a social-emotional and character development program on the trajectory of behaviors associated with social-emotional and character development: Findings from three randomized trials. *Prevention Science, 12,* 314-323.

Weimer, D.L. and A. R. Vining (Eds). 2009. *Investing in the Disadvantaged. Assessing the Benefits and Costs of Social Policies*. Georgetown University Press: Georgetown, Washington DC.

Weissberg, R.P., Kumpfer, K., & Seligman, M. E. P. (Eds.). 2003. Prevention that works for children and youth: An introduction. *American Psychologist, 58*, 425-432.

Wolke, D. Copeland, W.E., Angold, A. & E.J. Costello. 2013. Impact of bullying in childhood on adult health, wealth, crime, and social outcomes. *Psychological Science,* 24, 1958-1970.

Zins, J.E., Weissberg, R.P., Wang, M. C., and H.J. Walberg. (Eds.). 2004. *Building Academic Success on Social and Emotional Learning: What Does the Research Say?* New York: Teachers College Press.

Zins, J.E., Bloodworth, M.R., Weissberg, R.P., and H.J. Wahlberg. 2007. The scientific base linking social and emotional learning to school success. *Journal of Educational and Psychological Consultation, 17,* 191-210.

APPENDIX I: BENEFITS MAPS

Benefits Map: 4Rs

Outcome Categories	Specific Outcomes	Measures	Monetizable
Understanding and handling feelings	Child aggression	SR - teacher	Y
	Child social competence	SR - teacher	Y
	Depressive symptoms	SR - student	
	Hostile attributional bias	SR - student	
	Normative beliefs about aggression	SR - student	
	Aggressive and prosocial fantasies	SR - student	
	Aggressive interpersonal negotiation strategies	SR - student	
Health-related	Attention skills - Child ADHD symptoms	SR – teacher	Y
Achievement	Reading	STest	P
	Math	STest	P
	Writing	—	
Academic skills	Academic skills	SR – teacher	P
Other	Attendance	Records	P
	Quality of classroom processes	Observations	
	Social– emotional functioning by teacher	SR – teacher	
Cooperation	—	—	
Building community	—	—	
Assertiveness	—	—	
Discussion skills	—	—	
Problem-solving	—	—	
Dealing with diversity	—	—	
Listening	—	—	

Notes: SR: Self-report; STest: Standardized test; Records: administrative/school records. P: potentially monetizable

Benefits Map: Positive Action

Outcome Categories	Specific Outcomes	Measures	Monetizable
Personal Behaviors	Substance abuse	SR – student SR – teacher SR – parent	Y
	Violence	SR – student SR – teacher SR – parent Records	Y
	Beliefs about aggression and violence	SR – student	Y
Mental Health	Depression	BASC	Y
	Anxiety	BASC	Y
	Sexual activity	SR – student	
	Affect	SR – student	
	Life satisfaction	SR – student	
Academic	Achievement	STest	P
	Grades	Records	P
	Retention	Records	P
	Academic ability/potential	SR – teacher	P
Academic Behaviors	Student motivation	SR – student SR – teacher	
	Attendance	Records	P
	Work habits and organizational skills	SR – student SR – teacher	
School Climate	School Quality	Surveys	

Notes: SR: Self-report; STest: Standardized test; Records: administrative/school records; Surveys: devised survey instruments; BASC: Behavioral Assessment System for Children. P: potentially monetizable.

Benefits Map: Second Step

Outcome Categories	Specific Outcomes	Measures	Monetizable
Substance abuse	Drug use	SR – student	Y
	Prescription drug abuse	SR – student	Y
	Alcohol use (as minor)	SR – student	Y
	Drug knowledge and attitudes	SR – student	
Delinquency	Aggression	SR – student	Y
	Delinquency	SR – student	Y
	Risky driving	SR – student	Y
Health	Sexual risk behaviors	SR – student	Y

Notes: SR: Self-report; STest: Standardized test; Records: administrative/school records; Surveys: devised survey instruments; BASC: Behavioral Assessment System for Children. P: potentially monetizable.

Benefits Map: LST

Outcome Categories	Specific Outcomes	Measures	Monetizable
Substance abuse	Drug use	SR – student	Y
	Prescription drug abuse	SR – student	Y
	Initiation of other illegal drug use	SR – student	Y
	Frequency of other illegal drug use	SR – student	Y
	Alcohol use (as minor)	SR – student	Y
	Smoking	SR – student	Y
	Drug knowledge and attitudes	SR – student	
Delinquency	Aggression	SR – student	Y
	Delinquency	SR – student	Y
	Risky driving	Records	Y
Health	Sexual risk behaviors	SR – student	Y
Social and Emotional Skills	Self-efficacy	SR – student	
	Decision-making skills	SR – student	

Notes: SR: Self-report; STest: Standardized test; Records: administrative/school records; Surveys: devised survey instruments; BASC: Behavioral Assessment System for Children. P: potentially monetizable.

Benefits Map: Responsive Classroom

Outcome Categories	Specific Outcomes	Measures	Monetizable
Achievement	Reading	STest	Y
	Math	STest	Y
Understanding and handling feelings	Social skills	SR – Teacher	
	Social Competence	SR – Teacher	
	Perception of school	SR – Student	
Delinquency	Reduction of Problem behavior	SR – Teacher, Parent, Student	P
Teacher Outcomes	Teacher Self-Efficacy	SR – Teacher	
	Teacher-Child Relationship	SR – Teacher	

Notes: SR: Self-report; STest: Standardized test; P: potentially monetizable.

Benefits Map: SET (Swedish)

Outcome Categories	Specific Outcomes	Measures	Monetizable
Substance abuse	Drugs	SR – student	Y
	Alcohol	SR – student	P
	Smoking	SR – student	P
	Volatile substances	SR – student	P
Delinquency	Aggressiveness	SR – student	P
Mental health	Psychological well-being	SR – student	P
	Body image	SR – student	
	Bullying	SR – student	P
	Attention seeking	SR – student	
Social competence	Talent/ability	SR – student	
	Relation with others	SR – student	
	Internalizing	SR – student	
	Externalizing	SR – student	
	Mastery	SR – student	
	ITIA	SR – student	
	Social skills	SR – student	

Notes: SR: Self-report; P: potentially monetizable.

APPENDIX II

Appendix Table 1
Adult Lifetime Benefits from Positive Action

	Control Group	Treatment Group	Reduction from PA	PV Benefit per avoided life-time profile[1]	Total PV Benefit per Participant
Substance abuser probability:					
Estimate 1	0.076	0.040	0.036	$877,030	$31,570
Estimate 2	0.188	0.101	0.087	$877,030	$76,300
Estimate 3	0.053	0.016	0.037	$877,030	$32,450
Estimate 4	0.041	0.011	0.030	$877,030	$26,310
Estimate 5	0.035	0.007	0.028	$877,030	$24,560
Career criminal probability:					
Estimate 1	0.061	0.022	0.039	$2,833,490	$110,510
Estimate 2	0.074	0.028	0.046	$2,833,490	$130,340
Estimate 3	0.038	0.011	0.027	$2,833,490	$76,500
Estimate 4	0.107	0.045	0.062	$2,833,490	$175,680
Estimate 5	0.054	0.013	0.041	$2,833,490	$116,170

Sources: 1 Cohen and Piquero (2009, Table 12). Substance abuser: Original value $650,000 adjusted for inflation using CPI to 2013 dollars and PV at age 6 using 3% discount rate. Career criminal: Original value $2.1m adjusted for inflation using CPI to 2013 dollars and PV at age 6 using 3% discount rate. Control/treatment group impacts from Beets et al. (2009). Estimates assume zero fade-out.

Appendix Table 2

Prior Cost-Benefit Analyses of Life Skills Training

	Lee et al. (2012)	Miller and Hendrie (2008)
Costs per participant	***$40***	***$290***
Benefits:		
Medical, other resource, work loss, quality of life		$5,960
Crime	$1,660	
Effects of HS graduation and smoking reduction on earnings and health care spending	$190	
Total Benefits per participant	***$1,850***	***$5,960***
B-C ratio	**46.3**	**20.6**
NPV per participant	**$1,810**	**$5,670**

Notes: Dollar values in 2013 prices, rounded to $10.

Appendix Table 3
Costs for SET (5-Year Program)

Ingredient	Year 1	Year 2	Year 3	Year 4	Year 5	Total Cost
Participants	*1,028*	*1,028*	*1,028*	*1,028*	*1,028*	*1,028*
Personnel:						
Teachers (Training/ongoing coaching)	$57,000	$17,870	$4,710	$4,710	$4,710	$89,000
Teachers (SET instruction)	$82,450	$82,450	$82,450	$82,450	$82,450	$412,260
Administration (Training/ support/meetings)	$7,270	$4,480	$1,450	$1,450	$1,450	$16,110
Facilities:						
Auditorium	$340	$0	$0	$0	$0	$340
Workshop space	$170	$100	$100	$100	$100	$570
Classrooms	$6,070	$6,070	$6,070	$6,070	$6,070	$30,350
Materials/equipment:						
SET Manual	$60	$60	$60	$60	$60	$310
Supplemental materials	$1,260	$1,260	$1,260	$1,260	$1,260	$6,320
Total Resource Cost 1 (with teacher SET instruction)	$154,630	$112,290	$96,110	$96,110	$96,110	$555,260
Total Resource Cost 2 (without teacher SET instruction)	$72,180	$29,840	$13,660	$13,660	$13,660	$143,000
Average Cost 1 (=TRC2/1,028)						$540
Average Cost 2 (=TRC1/1,028)						$140

Notes: Discounted by 3.5% to year 1. Prices in U.S. dollars (2013). Amounts rounded to $10.