

Making Money Matter: How Districts Can Use the School Improvement Fund to Drive Achievement

Through a series of research reports, The Chalkboard Project has pointed teachers, principals, superintendents, and school boards to interventions with the highest likelihood to improve K-12 student achievement. This issue paper highlights the limited, but growing, number of programs proven effective through experimental trials. Targeted expansions of these rigorously tested programs, with some of the newly available School Improvement Fund resources, would spark progress and rapidly move Oregon students into the academic top tier.

Introduction

Oregon's investment in K-12 education slipped during the 2001–2003 recession and, consequently, Oregon fell from an above- to below-average K-12 spending state. Since then, a recovering economy has left the state in a relatively strong, short-term fiscal position. The Oregon Legislature is poised to appropriate \$260 million to the School Improvement Fund (SIF). As currently designed, the Fund targets spending to a variety of early childhood programs, including pre-Kindergarten programs, K-3 class size reductions, full day kindergarten, and literacy programs. Districts could also invest in a relatively broad array of other programming, including teacher mentoring, extended instructional time for at-risk students, new curricula and instructional materials, English as a Second Language services, and vocational education programs.¹

The Chalkboard Project supports the increased spending that would accompany the SIF and is convinced that—if invested well—the higher spending could translate into improved student achievement. But across the United States, the K-12 system's track record of turning more resources into better outcomes is mixed at best. As the non-partisan Coalition for Evidence Based Policy puts it:

“...over the past 30 years the United States has made almost no progress in raising the achievement of elementary and secondary school students, according to the National Assessment of Educational Progress, despite a 90 percent increase in real public spending per student.”²

¹ Senate Bill 318-B as ordered by the House on May 25, 2007.

² Coalition for Evidence-Based Policy. November 2002. Bringing Evidence-Driven Progress to Education: A Recommended Strategy for the US Department of Education. page i.

The Coalition points to a shortage of evidence-based education policies as a key reason the country has not seen stronger progress on achievement. The problem is two-fold. First, districts and states have lacked a willingness to rigorously evaluate educational programs, so the knowledge base on what works in a K-12 context is relatively thin. Second, districts and states have underinvested in the few areas where solid research exists.

The balance of this issue paper reviews a number of policy areas that are candidates for funding through the SIF and highlights programs that have met the research “gold standard”: achievement gains proven through an experimental trial³ that included a carefully designed treatment and control group. It is not practical to expect, nor does Chalkboard advocate, that districts restrict their SIF spending to this limited number of scientifically tested interventions. However, Chalkboard strongly encourages districts to review these programs, identify those that complement their on-going strategies, and strategically invest a meaningful share of their SIF allocations in one or two interventions backed by rigorous evidence. By doing so, districts would increase the likelihood of short-term achievement gains, which, in turn, would improve the chances for the reauthorization of the SIF.

One-on-One Tutoring

One-on-one tutoring of students deemed at risk of reading failure has proven effective through a variety of delivery models. The proven programs share the common feature of identifying readers at risk of reading failure and pairing them, one-on-one, with a tutor for an hour or more a week.

³Through experimental trials, researchers randomly assign students to treatment and control groups. With appropriate sample sizes and a careful design, the socio-economic and academic characteristics of the students are very similar. Consequently, researcher can attribute any subsequent difference in achievement measured between the groups to the treatment.

The most comprehensive, proven tutoring approach is embedded in the Success for All (SFA) program, which is a comprehensive, whole school reform. Through SFA’s one-to-one tutoring, certified teachers worked with K-3 students, concentrating about 40 percent of their efforts on first graders. Initial informal reading inventories identified which students were candidates for assistance, and then regular assessments were performed at eight-week intervals. Tutors supported these students for additional 20-minute reading sessions focusing on skills practice and addressing specific learning deficits. Tutoring activities are aligned with and supplemented the regular classroom curriculum.⁴

While SFA suggested promise through numerous quasi-experimental evaluations, findings from a large-scale experimental trial validated the approach. The trial, which consisted of 35 schools serving more than 20,000 students across the United States, found SFA participants outperformed their non-SFA counterparts on three literacy outcomes: word identification, word attack, and passage comprehension. SFA participants showed the biggest gains in word attack and outperformed 64 percent of their counterparts.⁵

Rigorous evidence also backs one-on-one tutoring with the Lindamood Phonemic Sequencing curriculum. An experimental trial placed 180 kindergarten students with poor phonological-processing skills into categories: tutoring aligned with the Lindamood curriculum, tutoring

⁴Slavin, R., Madden, N., Karweit, L., & Wasik, B.A. (1990). *Success for All: Effects of variations in duration and resources of a school-wide elementary restructuring program* (No. 2). Baltimore, MD: Center for Research on Effective Schooling for Disadvantaged Students. Retrieved November 6, 2004, from <http://www.ed.gov/offices/OERI/At-Risk/cdspubs.html>

⁵See Borman, Geoffrey D., Robert E. Slavin, Alan Cheung, Anne Chamberlain, Nancy Madden and Bette Chambers. 2006 “Final Reading Outcomes of the National Randomized Field Trial of Success for All.” Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA

supported by two alternative curricula, and no tutoring. Students placed in the Lindamood exhibited greater improvements in word attack skills, word identification, and reading comprehension than students enrolled in the alternative tutoring programs or who received no tutoring at all.⁶

The Reading Recovery program has demonstrated achievement gains in four separate experiments. The program supplements regular classroom instruction with one-on-one tutoring for at-risk students that generally run 30 minutes daily for 12–20 weeks. Participating districts must establish a Reading Recovery site, which involves a one-year training program for a teacher leader. Gleaning findings across the four trials, the US Department of Education concluded Reading Recovery had positive effects on general reading achievement and phonemic awareness and potentially positive effects on fluency and comprehension.⁷

Finally, Oregon's Start Making a Reader Today (SMART) is a low-cost program that has measured sustained reading gains among participants. SMART pairs community volunteers with K-2 students, who are deemed at risk of failing reading standards. The program currently serves about 11,000 students in 260 elementary schools. Community volunteers receive minimal training and work with students in 30-minute sessions twice a week during school hours. District costs, which run about \$300 per child per year, primarily cover the cost of a SMART coordinator. An experimental trial showed students tutored through the SMART program outperformed non-tutored peers on national tests of word

identification, word comprehension, and passage comprehension.

Significant expansions of one-on-one tutoring for early, at-risk readers may be the best evidence-based option districts could pursue with the SIF resources. Not only has the approach proven effective across multiple trials and conditions, but it also lacks the implementation challenges that face class size reduction policies.

Class Size

Class size reduction strategies are arguably the most analyzed and debated K-12 educational reform of the past several decades. Class size reduction policies are very expensive and require additional teachers and classroom space. Moreover, large-scale class size reduction policies can abruptly increase the demand for teachers and, consequently, lower teacher quality.⁸ Researchers have advanced hundreds of articles that support, condemn, or stand silent on the efficacy of the approach.

Standing above all these analyses are the findings of a single scientific experiment—the Tennessee Student/Teacher Achievement Ratio project, or Project STAR—which randomly assigned students and teachers to large and small classes. The only wide-scale, randomized field study of its kind, the Tennessee STAR project reduced class sizes in Kindergarten through Grade 3. All Tennessee elementary schools were invited to participate in the project and ultimately 79 schools did, which resulted in more than 6,000 students per grade level. Students and teachers were randomly assigned to one of three class types: regular classes (student-teacher ratios of 22–26); regular classes with full-time aide (student-teacher ratios of 22–26 and a full-time teacher's aide); and small classes (student-teacher ratios between 13 and 17).

Students in STAR's small classes

⁶Torgesen, Joseph, Richard Wagner, Carol Rashotte, Elaine Rose, Patricia Lindamood, Tim Conway, & Cyndi Garvan (1999). "Preventing reading failure in young children with phonological processing disabilities: Group and individual responses to instruction." *Journal of Educational Psychology*, 91, 579-593.

⁷US Department of Education. 2007. *WWC Intervention Report: Reading Recovery*. Accessed at <http://www.whatworks.ed.gov>.

⁸See Bohrnstedt, George and Brian Stecher. September 2002. *What We Have Learned About Class Size Reduction in California*. CSR Research Consortium.

performed better than students in regular and regular/aide classes in all locations and at every grade level. STAR's kindergarten students in small classes showed measurable achievement gains over students in regular classes and regular classes with aides. The STAR-achievement advantage persisted through first grade with small class students scoring in the 64th percentile and 59th percentile in reading and mathematics, respectively, on the Stanford Achievement Test (SAT). Meanwhile, students in regular classes scored in the 53rd percentile in reading (11 percentage points lower) and the 47th percentile (12 percentage points lower) in mathematics.⁹

Researchers concluded that the achievement gains were concentrated in Kindergarten and 1st Grade. By reviewing cohorts of new students who entered the program in Kindergarten, they found the achievement gains were established in the first two years (K-1) and then declined slightly in Grade 2 and 3. Researchers speculated that the class size effect may be pronounced in K-1 because children are not well socialized or familiar with classroom routines. They further argue that teachers can more easily manage the “better socialized” students in Grades 2 and 3.¹⁰

After 1st grade, the evidence on class size reductions weakens considerably.¹¹ At best, according to Princeton University's Alan Krueger, broad class size reduction policies may produce positive outcomes, but they are just as likely to yield no achievement improvement at all. At worst, Stanford University's Eric Hanushek argues class size reduction policies have a very weak,

⁹ See Boyd-Zaharias, Jayne. *Project STAR The Tennessee Student/Teacher Achievement Ratio Study: Background and 1999 Update*. HEROS, Inc. Lebanon, TN.

¹⁰ See Word, Elizabeth et. al. *The State of Tennessee's Student/Teacher Achievement Ratio (STAR) Project: Final Summary Report 1985-1990*. Tennessee Department of Education. Nashville, TN.

¹¹ See Mishel, Lawrence and Richard Rothstein editors. 2002. *The Class Size Debate*. Economic Policy Institute. Washington, DC. Table 1-2, Page 14.

virtually non-existent link to student achievement. In short, the literature suggests that if school districts proceed in this area, they should do so with a full understanding that empirical evidence offers no guarantee of achievement gains for class size reductions beyond the very early grades.

The cost-effectiveness of class size reductions in the early grades may improve if they start with students from low-income and minority families and are coupled with complementary policies (e.g., evidence-based professional development for teachers).¹²

High-quality Preschool

In the last forty years, much research has focused on the role of high-quality early childhood education in closing the achievement gap between advantaged and disadvantaged children. Rigorous research suggests high quality education for low-income preschool-aged children can boost elementary school achievement, lower enrollment in special education, and reduce grade retention. The strength of the preschool argument is built on the findings of two model programs: High/Scope Perry Preschool and the Carolina Abecedarian Project.

The High/Scope Perry Preschool Project, which operated during 1962–1967, is most frequently cited demonstration of the effects of early childhood development programs on disadvantaged children. The high-quality program consisted of certified teachers with bachelor's degrees with each teacher serving no more than 8 students. Daily classes ran 2.5 hours in length, and teachers conducted home visits weekly. Participants were tracked until age 40, which allowed researchers to measure long-term educational gains but also other important social outcomes (e.g., criminal behavior and welfare assistance costs). Perry attendees had higher high school completion rates than their non-program

¹² Mishel (2002) pp 92-93.

counterparts (65 percent vs. 45 percent) and outperformed the non-program group on school achievement tests at ages 9, 10, and 14; and on literacy tests at ages 19 and 27. At age 40, 76 percent of Perry students were employed compared to 62 percent of the non-program peers.¹³

Like Perry, the Carolina Abecedarian Project is a long-running experimental demonstration that compares education and economic outcomes between an intensive preschool treatment and control groups. The treatment consisted of a center-based education program for 0-5 year olds that focused on language development. Researchers measured achievement gains in reading and math at ages 8, 12, and 15. Moreover, at age 21 project attendees were almost three times as likely to have attended a four-year college than their control counterparts.¹⁴

The Perry and Abecedarian models are relatively high cost at \$7,415 and \$8,574 per child per year, respectively. And both models are multi-year interventions: Perry runs two years, Abecedarian for five. To date, neither Perry nor Abecedarian has been replicated on a broad scale in typical classroom settings. While federal Head Start and the Oregon PreKindergarten programs borrow features of the Perry model and operate at a comparable cost per child, neither has accrued Perry's body of evidence. Head Start research—both experimental and non-experimental—suggests the program may generate academic and economic benefits that are comparable to the model programs, but the evidence is still building.¹⁵

¹³ Schweinhart, Lawrence. 2004. *The High/Scope Perry Preschool Study Through Age 40: Summary, Conclusions, and Frequently Asked Questions*. High/Scope Educational Research Foundation. Ypsilanti, MI.

¹⁴ Masse, Leonard and W. Steven Barnett. *A Benefit Cost Analysis of the Abecedarian Early Childhood Intervention*. National Institute for Early Education Research. New Brunswick, NJ.

¹⁵ See Isaacs, Julia B. 2007. *Cost-Effective Investments in Children*. The Brookings Institution. Washington DC. Page 9.

Oregon school districts interested in expanding preKindergarten programs with SIF resources should note research supports only high-quality models. Program designs tied to the proven Perry and Abecedarian models have a higher likelihood of driving academic gains than those that are not. No such evidence yet exists for broad-coverage, low-cost programs.

Full Day Kindergarten

Full-day kindergarten has gained support as a scheduling option in public schools and has simulated discussion among parents, teachers, administrators and school boards who are looking to the empirical research to determine the potential implications of the extended day. The evidence on the academic effects of full-day kindergarten is ambiguous at best.

Unlike preschool or class size interventions, full day kindergarten has never been evaluated through a randomized experimental trial. Lacking that rigorous evidence, researchers have recently turned to longitudinal databases and—after the fact—tease out the effects of full day kindergarten from other factors that drive achievement (e.g., socioeconomic profile of students, teacher quality). These quasi-experimental studies have concluded that participation in full day kindergarten may generate short-term achievement gains, but those gains fade over time and are largely eliminated between first and third grades.¹⁶ A recent RAND study summarizes near consensus of the research to date:¹⁷

¹⁶ See Votruba-Drazal, E. 2005. Full-day vs. part-day kindergarten: Children's academic trajectories through first grade. Paper presented at the Society for Research on Child Development. Atlanta, GA. See also Rathburn, A. and J. West. 2004. *From Kindergarten through Third Grade: Children's beginning school experiences*. U.S. Government Printing Office: National Center for Educational Statistics.

¹⁷ See Le, Vi-Nhuan et. al. 2006. *School Readiness, Full-Day Kindergarten, and Student Achievement: An Empirical Investigation*. RAND. Santa Monica, CA.

“There appears to be growing evidence that attendance in full-day kindergarten programs is not associated with long-term academic benefits. However, it may be possible to redesign early education programs so that nonacademic skills are improved, which may translate into improved academic achievement”

Chalkboard continues to support strategic expansions of full-day kindergarten in the longer run and anticipates that, with time, research will prove the intervention effective for some students under some conditions. School districts pursuing an immediate expansion of full-day kindergarten should proceed cautiously, avoid universal implementation, and give the program its best chance for success. Students with the most disadvantaged backgrounds—low-income or non-English speaking families—could potentially benefit from full-day kindergarten but, even there, the evidence is mixed. Districts should also examine the rigor of the second-half of the kindergarten day and ensure curriculum is aligned with those in the upper grades.

Services to At-Risk Youth

Educators, policymakers, and researchers have recently redoubled their efforts to understand which students fail to graduate from high school on time and why. Research shows that student behaviors and academic performance in the middle grades can serve as strong, early predictors of on-time graduation.¹⁸ But while school districts are getting much better at predicting who is at risk of dropping out of school, considerably less is known about how to intervene and boost graduation rates or post-graduation outcomes. In this area, experimental trials point to at least two interventions worthy of consideration.

¹⁸ See Celio, Mary Beth and Lois Leveen. 2007. *The Fourth R: New Research Shows Which Academic Indicators Are the Best Predictors of High School Graduation*. Northwest Decision Research. Seattle WA.

The Minneapolis-based Check and Connect program focused attention on middle and high school aged students with learning, emotional, or behavioral disabilities. Schools would assign at-risk students a “monitor” who would periodically “check” academic (e.g., test scores, homework completion) and behavioral indicators (e.g., suspensions) associated with a high likelihood of dropping out of school. The monitors—who could be university graduate students or community members with social service backgrounds—worked with 25 to 35 students at a time and spent about an average one hour weekly with each student and his/her teachers and family.¹⁹ A randomized trial showed Check and Connect increased the likelihood of completing high school by 18 percentage points—61 percent of Check and Connect participants completed high school or a GED compared with 43 percent of control group students.²⁰ The intervention, as evaluated, is relatively expensive (\$1,600 per student per year) and the program’s developers are exploring lower-cost variations, which could include monitors who meet with small groups of students (three or four) rather than individually.²¹

Career Academies, which combine college preparatory work with technical and occupational courses in small learning environments, have boosted the post-school earnings of program participants. In a multi-site, experimental evaluation, the Manpower Demonstration Research Corporation (MDRC) reported the Academies significantly reduced dropout rates among students at high risk of school failure—32 percent of the non-Academy high-risk students dropped out of high school compared with 21 percent of their

¹⁹ See <http://ici.umn.edu/checkandconnect> for a complete description of the role of the monitor.

²⁰ See Sinclair, Mary F and others. 2005. “Promoting School Completion of Urban Secondary Youth With Emotional or Behavioral Disabilities.” *Exceptional Children*. Vol 71, No. 4. pp. 465-482.

²¹ Telephone conversation with Dr. Sandra Christensen on June 5, 2007.

high-risk Academy peers.²² Moreover, a subsequent report indicated young men involved with the program earned \$10,000 more than non-Academy men during a four-year follow-up period. The program had no measurable effect on earnings of young women. Career Academies are widely implemented across the United States, and in Oregon, fourteen Academies are formally associated with the Career Academy Support Network.

Conclusions

In light of K-12 education's well-documented link to economic productivity, the number of practices proven through rigorous research is surprisingly limited. The academic literature generally supports the theory that the earlier the interventions take place in a child's life, the higher is the likelihood that they will have a positive impact on student achievement. Rigorously tested interventions are concentrated in the early years and include one-on-one tutoring for young students falling behind in reading; targeted class size reductions in kindergarten and first grade; and high quality preschool programs. Recently, experimental trials have also validated two successful approaches to reducing high school dropout rates: well-documented Career Academies and the less well-known Check and Connect intervention.

Investing SIF resources exclusively on the limited list of evidence-based interventions is not practical for most districts. However, research suggests that districts that focus a meaningful share of their SIF allocation to one or more of these evidence-based policies stand a much better chance of measurably impacting achievement and graduation rates. In Chalkboard's view, quick and measurable success on student achievement would, in turn, strengthen the chances of reauthorization of the SIF two years hence.

²² Kemple, James J and Jason C. Snipes. 2000. *Career Academies: Impacts on Students' Engagement and Performance in High School*. MDRC.