



THE TRUTH ABOUT CLASS SIZE REDUCTION: PROBLEMS, EXAMPLES, AND ALTERNATIVES

IB-2014-A • April 2014 | by Ross Izard, Research Associate

SUMMARY

Throughout the United States, class size reduction (CSR) is heralded as an effective way to improve academic outcomes. However, the research surrounding the effects of CSR is inconclusive at best. It is often difficult to determine the reliability of the studies conducted on the topic, and much of the research wavers between small, temporary academic gains in certain student groups and no gains at all.

Previous attempts to implement CSR on a large scale have proven exorbitantly expensive and unwieldy. Two states in particular illustrate the dangers of CSR:

- California's CSR program, which aimed to reduce class sizes in grades K-3 to 20 students, ultimately wound up costing roughly \$1.7 billion per year and produced more problems than it solved.¹ The program's goals have been largely abandoned in the face of growing financial problems within the state.
- In Florida, a program was adopted which attempted to reduce class sizes in both primary and high schools across the state. This program has cost roughly \$20 billion since 2002 and is expected to see its annual costs rise to nearly \$4 billion in the coming years.² It has produced few clearly correlated gains in academic achievement.

The high-cost, low-return educational solution offered by large-scale CSR is not an effective use of taxpayer dollars. However, targeted small-scale application of educational techniques that reimagine what schools and classrooms look like could be beneficial. For this reason, Colorado would be better served by pursuing two complementary courses of action:

- **Encourage exploration into new, promising educational strategies such as blended learning.** By doing so, the state could help provide educators with the information they need to make well-informed, effective decisions about how to improve their unique students' outcomes.
- **Promote autonomy in decision-making among schools and school districts.** Educators thus would be empowered to utilize the best strategy for their own unique students and within their specific contexts.

QUESTIONABLE BENEFITS

A variety of research exists on the topic of class size reduction (CSR), much of which comes to different conclusions about its efficacy. Many contemporary studies find no significant benefits in CSR programs. However, the positive results of the seminal, late-1980s Tennessee STAR experiment—the foundation of nearly all modern CSR programs—led some scholars to conclude that class size reductions can have beneficial effects for low-achieving students in primary grades.³ Some studies even go so far as to predict specific long-term wage increases (around 4.4 percent) as a result of class size reductions at an early age.⁴

While these studies initially seem promising, it is important to note that the majority of purported CSR benefits are generally small and have the largest impacts on minority, low-income, or low-performance students. Benefits also seem to be largely concentrated in kindergarten and first grade, with few demonstrable effects found after third grade. Moreover, many of CSR's small but measurable improvements are not cumulative; they tend to level out as students progress through the education system.⁵

Contemporary research indicates that many of the previously reported benefits of CSR are only correlational, or merely based on measuring two factors—student achievement and class size, in this case. As such, some CSR supporters assume a causal relationship that may or may not exist between the two.⁶ Many variables exist in the educational world that could potentially affect student outcomes. Class compositions, parental involvement levels, teacher strategy, socio-economic conditions, and school facilities themselves, along with school leadership, expectations, and culture all can impact academic performance.

Therefore, it is extremely difficult to accurately separate and gauge the effects of CSR programs. Supposed benefits based only on loose correlation are at best only partially helpful and at worst entirely inaccurate.

Many studies that tout the benefits of CSR on education are actually meta-studies; they simply review, aggregate, and draw conclusions from samples of previous research. While some meta-studies can be helpful, many can reach misleading conclusions. A meta-study is only useful if one can be certain that the studies underlying it are unbiased, selected based on unbiased criteria, and free of major flaws. Such certainty is often impossible in the realm of CSR research for four reasons:

1. **Many of the studies used to support various CSR meta-study conclusions do not use consistent definitions of “class size.”** One study may define the term as the size of an individual class, but another may define the concept as a school's overall student-teacher ratio. Thus, one cannot easily be certain that the study accurately measures the correct variables. Unreliable conclusions often follow.
2. **Studies based on student-teacher ratios can be skewed.** By including non-academic (physical education, music, etc.) teachers and teachers who work in more than one classroom in the calculations, schools can artificially decrease the reported number of students per teacher. As a result, the conclusions of any ratio-based CSR analyses can be skewed. This disparity helps explain why many schools report relatively low student-teacher ratios despite the prevalence of larger classes within those school themselves. In addition, the use of student-teacher ratios does not account for the nature of the physical facilities in which students are taught. Some schools may have small classes in individual, fully enclosed classrooms. Others may attempt to capitalize on available space by combining multiple classrooms in gymnasiums or other large, open spaces, thereby introducing a variable that could significantly affect academic achievement calculations.⁷
3. **Meta-studies often attempt to aggregate data collected from studies measuring different outcomes for different populations.**⁸ Measuring outputs equally across areas and populations can lead to false conclusions. The approach ignores a variety of factors outside of CSR that could impact student performance, such as socio-economic conditions, teaching strategies, school culture, and school facilities. Output measurements also could be out of sync; not all schools or school districts use the

same tests or standards to measure student achievement. For these reasons, CSR's effect could be very different in one area than in another, making aggregate analyses that draw overarching conclusions inherently unreliable.

4. **Meta-studies can fall victim to selection bias.** Researchers may “cherry pick” only studies that best support their conclusion instead of taking a broad cross section of the existing research, or they may weigh various studies differently, potentially affecting conclusions. As an example, a series of significant meta-studies (here called literature reviews) based on 277 estimates from 59 different studies concluded that CSR programs had little discernible effect.⁹ However, when the same estimates were examined using a different set of criteria, an entirely different conclusion was reached.¹⁰ Additionally, the studies that underlie meta-study results can sometimes be poorly controlled or executed, making their data of questionable value when drawing overarching conclusions in the policy space.

Finally, other studies have found little to no empirical evidence that CSR positively impacts students' academic performance, even in cases of massive reductions in class size (40 percent).¹¹ Instead, they assert that support for CSR programs is driven more by the self-interest of teachers, administrators, and parents than by rational cost-benefit analyses.¹²

Indeed, it is frequently argued that CSR programs serve the teachers' unions' interests, not those of the children. After all, fewer students per class translates into more teachers per school, which serves to inflate union membership—and thus the political and monetary leverage of the union itself—while simultaneously pleasing the teachers they serve.¹³ Viewed in this way, it is easy to see why some of the largest proponents of CSR programs are teachers' unions, as well as why CSR programs are generally popular politically despite their questionable efficacy.

THE LAW OF UNINTENDED CONSEQUENCES

Many CSR supporters have used Tennessee's Project STAR—arguably the most widely studied experiment in the realm of class size effects—both as an example of successful CSR implementation and as a template for future such programs. STAR sought to measure the impact of class size on academic performance in grades K-3. While a great deal of debate has followed about whether the program was rigorously enough controlled to provide compelling evidence of positive CSR effects, experts still generally accept the study as a relatively effective foray into CSR implementation.

However, STAR does not, as many CSR supporters would have the public believe, justify large-scale educational system overhaul. Instead, it serves as an example of successful small-scale, targeted implementation of class size reduction. The most commonly cited phase of the STAR program was implemented in Tennessee's 17 poorest school districts, and only in grades K-3. Its resultant benefits did not warrant a push for large-scale CSR; they simply established that such programs are best utilized in early education and among minority and underprivileged students. The narrow findings did not stop other states from attempting to implement similar programs on a massive scale, however, resulting in outcomes ranging from disappointing to counterproductive.

In California, wide-spread CSR implementation did not proceed smoothly. Beginning in 1996, the state heavily incentivized schools to limit their primary school class sizes to 20 students or less. Eager to obtain the additional state funding that CSR could bring, schools rushed to comply. Unfortunately, the policy's serious problems only became evident after it went into effect.

When California required that every small class be given a separate classroom of equal square footage, it resulted in the loss of nearly 1,400 computer labs, music rooms, and childcare facilities.¹⁴ Schools in every state generally operate within fixed facilities that are not easily expanded. Thus, most schools have

a finite number of classrooms to utilize. Forcing smaller class sizes creates more classes in fixed facilities. Unsurprisingly, such changes can necessitate the appropriation of a school building's limited extra space, potentially depriving students of valuable resources that would otherwise have been available.

Some temporary solutions to this problem exist, such as the use of external, mobile classrooms called modular buildings. Even so, many schools may be unable to use such tools due to issues of land availability, regulatory restrictions, or funding priorities. In these cases, schools can actually wind up *losing* assets despite the increased levels of funding that often accompany CSR programs. This unfortunate result was clearly seen in California following its roll out of large-scale CSR.

The rapid implementation of California's CSR program also created a nearly instant demand for roughly 25,000 new teachers that the market could not readily supply. As a result, many positions were filled by applicants without certification or prior teaching experience.¹⁵ While the notion that teacher certification is inextricably tied to teaching quality is highly suspect, the fact remains that California's class reduction policy initially resulted in a general statewide decrease in teacher quality. This decline in quality was particularly prevalent in low-income districts, largely negating any beneficial effects of CSR in areas where it might have achieved the best results. It took half a decade for the market to adjust, resulting in wasted resources and lost opportunities.

Finally, the immense pressure placed upon principals and administrators to meet certain class size targets can lead to outcomes contrary to CSR program goals. In California, such self-defeating results were seen in the creation of a large number of ineffective combination classes, which contain students from more than one grade level. Because multiple single-grade classes were seen as a practical impossibility under the CSR requirements, many mixed classes were created to save space and teacher time. These combination classes had a clearly negative impact on student performance in California that tended to grow worse in higher grade levels.¹⁶

The pressure caused by CSR implementation also may be compounded by the unavoidable intersection with performance targets and performance-based school rankings. Principals attempting to achieve performance targets may overburden the best teachers with large class sizes while attempting to minimize the number of students within poor-performing teachers' classes. Such adjustments could eventually lead to a "burn-out" effect in which skilled teachers leave the education field or find other schools with lighter burdens. Equally problematic is the practice of providing high-quality teachers with smaller classes as a reward. Doing so will likely reduce the overall outcomes of a school because the majority of students work in larger classes with less effective teachers.¹⁷

California stands as a warning against hasty, broad-based implementation of class size reduction. The program cost an estimated \$1.7 billion per year at its inception, an amount that has only risen in the subsequent years. The strategy has been all but abandoned as the state has moved from a time of budget surplus to one of economic strain.

THE MONEY SINK

California is not the only state to have attempted large-scale CSR implementation. In 2002, Florida instituted an even broader program that aimed to reduce class sizes across all grades in both primary *and* secondary schools. The program has cost an estimated \$20 billion since it began. Its annual costs are expected to reach \$4 billion in the coming years.

For all the money spent, Florida's CSR program has shown little in the way of verifiable positive results. CSR supporters point to Florida's increased overall academic achievement during the past decade as evidence

of the program's success. However, Harvard University scholar Matthew Chingos finds such assertions to be misleading. Florida's student achievement was already improving in the years prior to the CSR program. Further, policymakers enacted a variety of other changes to the state's education system around the same time. Changes included the following:

- **The A-Plus Accountability and School Choice Program**, which created a system of grade-based school evaluations, took effect in 1999, and was heavily revised in 2002.
- **A variety of school choice programs** were launched, including the Opportunity Scholarships Program, the McKay Scholarships for Students with Disabilities Program, and the Corporate Tax Credit Scholarship Program.
- **Just Read, Florida!** sought to improve reading and comprehension skills in schools across the state, in part by training teachers extensively in scientifically-based reading instruction.
- **Substantial funding changes occurred across Florida's education system due to CSR implementation.** Per pupil funding rose roughly evenly across all districts during the implementation of Florida's CSR program. However, districts that had already met the average class size requirements of the CSR program were free to use additional funds as they saw fit, potentially instituting programs or adding services that could have impacted student achievement.

It is extremely difficult to isolate which, if any, of these programs and policies had a statewide effect on student achievement, and virtually impossible to accurately attribute the state's gains to CSR alone.

In fact, Chingos' research seems to suggest that Florida's sweeping CSR mandate has had little effect on much of its target population. Four years after CSR's adoption in Florida, fourth through eighth grade math scores had, statistically speaking, not been affected. Scores rose only as much as they would have had the schools simply been given the additional resources without the requirement to reduce class size. By the same token, students in grades seven and eight—the grades that experienced the largest relative class size reduction—actually experienced a statistically significant decrease in reading achievement by 2006. Grades four through eight as a whole experienced no discernible improvement. Moreover, the touted non-academic benefits of CSR—reduced absenteeism, suspension rates, or violent incidents—did not decrease in any statistically significant way as a result of Florida's program.¹⁸

All told, Florida's CSR program has had little to no effect on key outcomes among students in grades four through eight. No reliable research currently exists to substantiate claims of benefits in Florida's K-3 or high school student populations, either. It remains possible that the program could have some small benefits among some students in the earliest grades—this would be consistent with the findings from STAR—but those benefits have yet to be credibly measured or evaluated. Florida has clearly spent billions of dollars on a program with highly questionable effects. This approach ties up valuable resources that could be used in other, more effective ways.

ALTERNATIVE APPROACHES

The lackluster results of previous CSR research and implementation culminate in an important question: Are CSR programs the most effective use of taxpayer dollars in Colorado's education system? In light of the sparse and often inapplicable evidence of CSR program benefits, the myriad of peripheral problems that such programs can create, and the enormous costs and risks associated with their implementation, the answer is clearly no. In a time when national tests show only about 40 percent of Colorado's students reach proficient status or better in reading and math, taxpayers deserve to see their money spent on the pursuit of real, effective, financially prudent educational reforms.¹⁹ Resources should not be squandered on programs that offer only meager or isolated benefits at great cost.

When traditional solutions to complex problems fail to produce the desired results, it is often more helpful to simply start over than to attempt to tailor, modify, and force fit the same ineffective solutions again and again. Indeed, the lack of progress in Colorado's educational system seems to indicate the need for more than retrofitted, by-the-book reform. The system needs to be fundamentally reconsidered.

Such reconsideration includes rethinking what a modern classroom should look like. Perhaps rather than envisioning the ideal classroom as one in which a small group of students receive lectures from a single teacher, Colorado's educators, legislators, and parents should consider other, potentially more effective strategies capable of improving educational outcomes for more students at a lower cost. Models aimed at accomplishing such results already exist in the form of blended learning, a type of schooling that combines online, computer-based learning and specially tailored face-to-face teaching. The blended approach to education has shown promising results despite upending many traditionally accepted ideas about education, including the idea that smaller classes are inherently more beneficial to students than larger ones.

Early forays into the realm of blended learning have shown tremendous promise. For instance, Carpe Diem Collegiate High School and Middle School in Yuma, Arizona, has utilized a blended learning strategy to achieve math and reading proficiency rates of 92 percent among its students. An average of 40 percent of Carpe Diem's students also attained advanced performance levels in those subjects.²⁰ Notably, many Carpe Diem students come from minority or low-income households, making the school's results particularly impressive in light of the fact that only 48 percent of Hispanic students and 44 percent of black students passed the AIMS mathematical assessment between 2005 and 2011. Similarly, just 68 percent of Hispanic students and 67 percent of black students passed the AIMS reading assessment during the same seven-year period.²¹

Carpe Diem has achieved its impressive results at costs much lower than one might expect; the school boasts that its program costs only 60 to 70 percent as much as traditional schooling on a per-pupil basis.²² The school's blended-learning strategy also opens the door to more cost-effective facility building, long considered one of the most prohibitive costs in education. Because blended learning approaches to education often do not require large numbers of individual classrooms, Carpe Diem's academic facility only cost 23 percent as much as a nearby traditional school to build.²³

Finally—and perhaps most interestingly—Carpe Diem's 280 students achieved their impressive math and reading scores despite the school having only six full-time teachers on its staff. Indeed, students at Carpe Diem often spend much of their time learning on computers in a large, open room that more closely resembles a business office than a school. Every 55 minutes, students rotate between the large, computer-based learning center and more traditional classrooms for specially focused face-to-face instruction.²⁴ The apparent success of this educational model flies in the face of commonly held notions about what education should look like, and seems to directly contradict the idea that class size alone is a major determining factor behind academic success. Further exploration will be necessary to firmly establish this finding, but initial results seem to indicate that the deciding factor in academic outcomes may not be the size of the group a student sits in, but the way in which information is conveyed in that setting.

Although the success of Carpe Diem is admittedly anecdotal, similar results have been noted at other schools using blended learning strategies. For instance, KIPP LA in California, a nonprofit charter school that utilizes a blended learning approach to education, achieved a college acceptance rate of 93 percent in a geographic area where only 10 percent of students pursue higher education. Eighty-six percent of KIPP LA's students from 2011 are currently pursuing college degrees.²⁵

Rocketship, a charter management organization that makes use of blended learning in its California schools,

has seen significantly positive results grow out of its alternative approach to education. In 2010, two Rocketship schools were ranked among the top 15 California schools with predominantly low-income student populations. Rocketship's success may serve as unproven but promising evidence that alternate approaches to education can help close the learning gap observed between low-income or minority students and their peers.²⁶ Much like Carpe Diem, the technological focus of blended learning has allowed Rocketship schools to save a great deal of money. Computer-based learning labs result in per-school savings of roughly half a million dollars—money that is subsequently used to support higher pay for its comparatively small number of teachers, after-school activities, and leadership programs.²⁷

Unfortunately, very little hard research has been conducted on the possible benefits of blended learning in the United States. Some Colorado school districts have already begun experimenting with modified blended learning programs, but no concerted effort has yet been made to fully examine and understand the potential benefits and limitations of the approach. Promoting further exploration of blended learning and other divergent approaches to education can only serve to empower schools, school districts, and parents across Colorado and the United States to make well-informed, carefully weighed education policy decisions with potentially real benefits instead of relying on exorbitantly expensive placebo programs.

Most importantly, the relative failings of large-scale CSR programs such as those in California and Florida illustrate the consequences of high-level decision making in situations that merit more carefully targeted, case-by-case solutions. Every student body is different, and some approaches—including CSR and certain forms of blended learning—may prove to be highly effective in some contexts and utterly ineffective in others.

For this reason, the decision of which approach to use in any given situation should be left up to those who best understand the unique situation being faced: the schools themselves. By both further developing an understanding of new educational techniques and increasing the level of autonomous decision-making among schools and school districts, Colorado can empower its educators to do what is best for their unique students instead of forcing them to utilize large-scale templates that could—and often do—prove ineffective.

ENDNOTES

- ¹ California Department of Education, <http://www.cde.ca.gov/ls/cs/k3/facts.asp>.
- ² "2009–10 Florida Education Finance Program," DOE Information Database Workshop, Summer 2009, available at <http://www.fldoe.org/eias/databaseworkshop/ppt/fejp.ppt>.
- ³ H.M. Cooper, "Does Reducing Student-to-Teacher Ratios Affect Achievement?" *Educational Psychologist* 27 no. 1 (1989): 16; Alan B. Krueger, "Experimental Estimates of Education Production Functions," *Quarterly Journal of Economics* 114 no. 2 (1999): 497-532; B. Nye, L.V. Hedges, S. Konstantopoulos, "Are Effects of Small Classes Cumulative? Evidence from a Tennessee Experiment," *Journal of Educational Research* 94 no. 3 (2001): 336–345; Alan B. Krueger and Diane M. Whitmore, "The Effect of Attending a Small Class in the Early Grades on College-Test Taking and Middle School Test Results: Evidence from Project STAR," *The Economic Journal* 111 no. 468 (2001): 1-28.
- ⁴ Peter Fredricksson, Bjorn Ockert, and Hessel Oosterbeek, "Long-Term Effects of Class Size," *The Quarterly Journal of Economics* 128 no. 1 (2012): 249-85.
- ⁵ Eric A. Hanushek, "Some Findings from an Independent Investigation of the Tennessee STAR Experiment and from Other Investigations of Class Size Effects," *Educational Evaluation and Policy Analysis* 21 no. 2 (1999): 143-63.
- ⁶ Robert E. Slavin, "Class Size and Student Achievement: Small Effects of Small Classes," *Educational Psychologist* 27 no. 1 (1989): 99-110.

This information stems from the following: Eugene M. Lewit and Linda Schuurmann Baker, "Class Size," *The Future of Children* 7 no. 3 (1997): 112-21; Melvin Borland, Roy M. Howsen, & Michelle W. Trawick, "An Investigation of the Effect of Class Size on Student Academic Achievement," *Education Economics* 13 no. 1 (2005): 73-83.

-
- ⁸ Alan B Kreuger, "Economic Considerations and Class Size," *The Economic Journal* 113 no. 485 (2003): F34-F63.
- ⁹ Eric A. Hanushek, "Assessing the effects of school resources on student performance: An update," *Educational Evaluation and Policy Analysis* 19 no. 2 (1997a): 141-164.
- ¹⁰ Kreuger, "Economic Considerations and Class Size."
- ¹¹ Slavin, "Class Size and Student Achievement: Small Effects of Small Classes"; Hanushek, "Some Findings from an Independent Investigation of the Tennessee STAR Experiment and from Other Investigations of Class Size Effects"; Caroline M. Hoxby, "The Effects of Class Size on Student Achievement: New Evidence from Population Variation," *The Quarterly Journal of Economics* 115 no. 4 (2000): 1239-85.
- ¹² Hoxby, "The Effects of Class Size on Student Achievement: New Evidence from Population Variation."
- ¹³ Terry M. Moe, "Collective Bargaining and the Performance of the Public Schools," *American Journal of Political Science* 53 no. 1 (2009): 156-174.
- ¹⁴ Eugene M. Lewit and Linda Schuurmann Baker, "Class Size," *The Future of Children* 7 no. 3 (1997): 112-21.
- ¹⁵ The preceding information comes from Christopher Jepsen & Steven Rivkin, "Class Size Reduction and Student Achievement: The Potential Tradeoff between Teacher Quality and Class Size," *Journal of Human Resources* 44 no. 1 (2009): 223-50.
- ¹⁶ David Sims, "A Strategic Response to Class Size Reduction: Combination Classes and Student Achievement in California," *Journal of Policy Analysis and Management* 27 no. 3 (2008): 457-78.
- ¹⁷ Nathan Barrett and Eugenia F. Toma, "Reward or punishment? Class size and Teacher Quality," *Economics of Education Review* 35 (2013): 41-52.
- ¹⁸ The information in this section was taken from Matthew M. Chingos, "The Impact of a Universal Class-Size Reduction Policy: Evidence from Florida's Statewide Mandate," Program on Education Policy and Governance, John F. Kennedy School of Government, Harvard University, 2011.
- ¹⁹ Forty percent represents the number of students attaining proficient or advanced NAEP scores in reading. 42% have attained proficient or advanced NAEP scores in math. See <http://nces.ed.gov/nationsreportcard/subject/publications/stt2013/pdf/2014465CO8.pdf> and <http://nces.ed.gov/nationsreportcard/subject/publications/stt2013/pdf/2014464CO4.pdf>.
- ²⁰ This information is based on statewide, standardized assessments conducted in Arizona. See <http://www.carpediemschools.com/results/>.
- ²¹ These percentages come from the 2013 Arizona Minority Student Progress Report. See <http://www.azhighered.gov/AMEPAC/AMEPAC%20MSPR%202013%20-%20Final.pdf>.
- ²² See <http://www.carpediemschools.com/results/>.
- ²³ Michael B. Horn and Heather Staker, "The Rise of K-12 Blended Learning," *The Innosight Institute*, 2011.
- ²⁴ Ibid.
- ²⁵ Brad Bernatek, Jeffrey Cohen, John Hanlon, and Matthew Wilka, "Blended Learning in Practice: Case Studies From Leading Schools Featuring KIPP Empower Academy," Michael and Susan Dell Foundation, 2012.
- ²⁶ Horn and Staker, "The Rise of K-12 Blended Learning."
- ²⁷ Rocketship Education informational page based on 2010-11 SY. See: <http://www.christenseninstitute.org/rocketship-education/>.

Copyright ©2014, Independence Institute

INDEPENDENCE INSTITUTE is a non-profit, non-partisan Colorado think tank. It is governed by a statewide board of trustees and holds a 501(c)(3) tax exemption from the IRS. Its public policy research focuses on economic growth, education reform, local government effectiveness, and constitutional rights.

JON CALDARA is President of the Independence Institute.

DAVID KOPEL is Research Director of the Independence Institute.

PAMELA BENIGNO is the Director of the Education Policy Center.

ROSS IZARD is Research Associate for the Education Policy Center. He is a graduate student in the Accelerated Master of Public Administration program at the University of Colorado at Denver. He holds a B.A. in Political Science from the University of Colorado at Boulder and maintains a strong interest in education policy and reform.

ADDITIONAL RESOURCES on this subject can be found at: <http://education.i2i.org/>.

NOTHING WRITTEN here is to be construed as necessarily representing the views of the Independence Institute or as an attempt to influence any election or legislative action.

PERMISSION TO REPRINT this paper in whole or in part is hereby granted provided full credit is given to the Independence Institute.