

Randomized Controlled Trials Commissioned by the Institute of Education Sciences Since 2002: How Many Found Positive Versus Weak or No Effects

July 2013

Since the establishment of the Institute for Education Sciences (IES) within the U.S. Department of Education in 2002, IES has commissioned a sizable number of well-conducted randomized controlled trials (RCTs) evaluating the effectiveness of diverse educational programs, practices, and strategies (“interventions”). These interventions have included, for example, various educational curricula, teacher professional development programs, school choice programs, educational software, and data-driven school reform initiatives. Largely as a result of these IES studies, there now exists – for the first time in U.S. education – a sizable body of credible knowledge about what works and what doesn’t work to improve key educational outcomes of American students.

A clear pattern of findings in these IES studies is that the large majority of interventions evaluated produced weak or no positive effects compared to usual school practices. This pattern is consistent with findings in other fields where RCTs are frequently carried out, such as medicine and business,¹ and underscores the need to test many different interventions so as to build the number shown to work.

Based on our careful count, what follows is the breakdown of positive versus weak or null findings in the IES studies:

- A total of 90 interventions have been evaluated in IES-commissioned RCTs. Of these:
 - 11 interventions (12%) were found to produce positive effects;
 - 79 interventions (88%) were found to produce weak or no positive effects.
- Focusing on the subset of 77 interventions evaluated in RCTs that our review identified as having no major study limitations (e.g., differential attrition, inadequate statistical power)²:
 - 7 interventions (9%) were found to produce positive effects;
 - 70 interventions (91%) were found to produce weak or no positive effects.

The methods we used in developing this count are described in endnote 3, and a complete list of the IES-commissioned RCTs is attached.

References

¹ In medicine: reviews have found that 50-80% of positive results in initial (“phase II”) clinical studies are overturned in subsequent, more definitive RCTs (“phase III”). John P. A. Ioannidis, “Contradicted and Initially Stronger Effects in Highly Cited Clinical Research,” *Journal of the American Medical Association*, vol. 294, no. 2, July 13, 2005, pp. 218-228. Mohammad I. Zia, Lillian L. Siu, Greg R. Pond, and Eric X. Chen, “Comparison of Outcomes of Phase II Studies and Subsequent Randomized Control Studies Using Identical Chemotherapeutic Regimens,” *Journal of Clinical Oncology*, vol. 23, no. 28, October 1, 2005, pp. 6982-6991. John K. Chan et. al., “Analysis of Phase II Studies on Targeted Agents and Subsequent Phase III Trials: What Are the Predictors for Success,” *Journal of Clinical Oncology*, vol. 26, no. 9, March 20, 2008.

In business: of 13,000 RCTs of new products/strategies conducted by Google and Microsoft, 80-90% have reportedly found no significant effects. Jim Manzi, *Uncontrolled: The Surprising Payoff of Trial-and-Error for Business, Politics, and Society*, Perseus Books Group, New York, 2012, pp. 128 and 142. Jim Manzi, *Science, Knowledge, and Freedom*, presentation at Harvard University’s Program on Constitutional Government, December 2012, linked [here](#).

² The criteria we used to assess whether the RCT was well-conducted are summarized in the [RCT Checklist](#), and include such items as: (i) the program and control groups were similar in their pre-program characteristics; (ii) the study had low sample attrition, and similar attrition rates for the program versus control group; (iii) the study sample was large enough to detect a meaningful effect of the intervention; (iv) the study measured outcomes for all individuals assigned to the program group, regardless of whether or how long they participated in the program; and (v) study outcomes were assessed with valid measures.

³ The following summarizes the methods we used in developing the count of positive versus weak or null findings:

- (i) In cases where the study measured intermediate outcomes (e.g., teacher content knowledge) and more ultimate, policy-relevant outcomes (e.g., student achievement), we counted the effect on the ultimate outcomes.
- (ii) In cases where the study measured both interim and longer-term outcomes – e.g., achievement at the end of one school year versus two school years – we counted the effect on the longer-term outcomes (as the best gauge of whether the intervention’s effects were sustained or dissipated).
- (iii) If the study found that the intervention produced no statistically-significant effect on the outcomes described in (i) and (ii), we counted the finding as one of “weak or no positive effects.”
- (iv) In the two cases where the study found an effect that was statistically significant but small in magnitude (0.05 and 0.06 standard deviations, respectively), we counted these as findings of “weak or no positive effects.”
- (v) One RCT compared the effectiveness of different curricula but did not have a services-as-usual control group and therefore is not included in our count (which seeks to address how often evaluated interventions produce better outcomes than what schools are doing anyway).
- (vi) Some of the RCTs evaluated multiple interventions; thus the number of interventions evaluated is higher than the number of studies.

List of IES-Commissioned RCTs

Abe, Y., Thomas, V., Sinicrope, C., & Gee, K. A. (2012). *Effects of the Pacific CHILD Professional Development Program*. (NCEE 2013–4002). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

http://ies.ed.gov/ncee/edlabs/regions/pacific/pdf/REL_20134002.pdf

Arens, S. A., Stoker, G., Barker, J., Shebby, S., Wang, X., Cicchinelli, L. F., & Williams, J. M. (2012). *Effects of curriculum and teacher professional development on the language proficiency of elementary English language learner students in the Central Region*. (NCEE 2012-4013). Denver, CO: Mid-continent Research for Education and Learning.

http://ies.ed.gov/ncee/edlabs/regions/central/pdf/REL_20124013.pdf

Black, A. R., Somers, M.-A., Doolittle, F., Unterman, R., and Grossman, J. B. (2009). *The Evaluation of Enhanced Academic Instruction in After-School Programs: Final Report* (NCEE 2009-4077).

Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

<http://ies.ed.gov/ncee/pubs/20094077/pdf/20094077.pdf>

Bos, J., Sanchez, R., Tseng, F., Rayyes, N., Ortiz, L., and Sinicrope, C. (2012). *Evaluation of Quality Teaching for English Learners (QTEL) Professional Development*. (NCEE 2012-4005). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

http://ies.ed.gov/ncee/edlabs/regions/west/pdf/REL_20124005.pdf

Campuzano, L., Dynarski, M., Agodini, R., and Rall, K. (2009). *Effectiveness of Reading and Mathematics Software Products: Findings From Two Student Cohorts* (NCEE 2009-4041). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

<http://ies.ed.gov/ncee/pubs/20094041/pdf/20094041.pdf>

Dynarski, Mark, Roberto Agodini, Sheila Heaviside, Timothy Novak, Nancy Carey, Larissa Campuzano, Barbara Means, Robert Murphy, William Penuel, Hal Javitz, Deborah Emery, and Willow Sussex.

Effectiveness of Reading and Mathematics Software Products: Findings from the First Student Cohort, Washington, D.C.: U.S. Department of Education, Institute of Education Sciences, 2007.

<http://ies.ed.gov/ncee/pdf/20074005.pdf>

Cavalluzzo, L., Lowther, D., Mokher, C., and Fan, X. (2012). *Effects of the Kentucky Virtual Schools' hybrid program for algebra I on grade 9 student math achievement*. (NCEE 2012-4020). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

<http://ies.ed.gov/ncee/edlabs/regions/appalachia/pdf/20124020.pdf>

Coe, Michael, Hanita, Makoto, Nishioka, Vicki, and Smiley, Richard. (2011). *An investigation of the impact of the 6+1 Trait Writing model on grade 5 student writing achievement* (NCEE 2012–4010).

Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

http://ies.ed.gov/ncee/edlabs/regions/northwest/pdf/REL_20124010.pdf

Condelli, L., Cronen, S., Bos, J., Tseng, F., and Altuna, J. (2010). *The Impact of a Reading Intervention for Low-Literate Adult ESL Learners* (NCEE 2011-4003). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. <http://ies.ed.gov/ncee/pubs/20114003/pdf/20114003.pdf>

Constantine, J., Player D., Silva, T., Hallgren, K., Grider, M., and Deke, J. (2009). *An Evaluation of Teachers Trained Through Different Routes to Certification, Final Report* (NCEE 2009-4043). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. <http://ies.ed.gov/ncee/pubs/20094043/pdf/20094044.pdf>

Cordray, D., Pion, G., Brandt, C., Molefe, A, & Toby, M. (2012). *The Impact of the Measures of Academic Progress (MAP) Program on Student Reading Achievement*. (NCEE 2013-4000). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. http://ies.ed.gov/ncee/edlabs/regions/midwest/pdf/REL_20134000.pdf

Corrin, W., Lindsay, J. J., Somers, M-A., Myers, N. E., Myers, C. V., Condon, C. A., & Smith, J. K. (2012). *Evaluation of the Content Literacy Continuum: Report on Program Impacts, Program Fidelity, and Contrast*. (NCEE2013-4001). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. http://ies.ed.gov/ncee/edlabs/regions/midwest/pdf/REL_20134001.pdf

Drummond, K., Chinen, M., Duncan, T.G., Miller, H.R., Fryer, L., Zmach, C., & Culp, K. (2011). *Impact of the Thinking Reader® software program on grade 6 reading vocabulary, comprehension, strategies, and motivation* (NCEE 2010-4035). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. http://ies.ed.gov/ncee/edlabs/regions/northeast/pdf/REL_20104035.pdf

Finkelstein, N., Hanson, T., Huang, C.-W., Hirschman, B., and Huang, M. (2011). *Effects of Problem Based Economics on high school economics instruction*. (NCEE 2010-4002rev). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. http://ies.ed.gov/ncee/edlabs/regions/west/pdf/REL_20104022.pdf

Garet, M., Wayne, A., Stancavage, F., Taylor, J., Eaton, M., Walters, K., Song, M., Brown, S., Hurlburt, S., Zhu, P., Sepanik, S., and Doolittle, F. (2011). *Middle School Mathematics Professional Development Impact Study: Findings After the Second Year of Implementation* (NCEE 2011-4024). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. <http://ies.ed.gov/ncee/pubs/20114024/pdf/20114024.pdf>

Garet, Michael S., Stephanie Cronen, Marian Eaton, Anja Kurki, Meredith Ludwig, Wehmah Jones, Kazuaki Uekawa, Audrey Falk, Howard Bloom, Fred Doolittle, Pei Zhu, and Laura Szejnberg. *The Impact of Two Professional Development Interventions on Early Reading Instruction and Achievement* (NCEE 2008-4030). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. <http://ies.ed.gov/ncee/pdf/20084030.pdf>

Glazerman, Steven, Eric Isenberg, Sarah Dolfin, Martha Bleeker, Amy Johnson, Mary Grider, and Matthew Jacobus. (2010). *Impacts of Comprehensive Teacher Induction: Final Results from a Randomized Controlled Study* (NCEE 2010-4027). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. <http://ies.ed.gov/ncee/pubs/20104027/pdf/20104027.pdf>

Gleason, P., Clark, M., Tuttle, C. C., and Dwoyer, E. (2010). *The Evaluation of Charter School Impacts: Final Report* (NCEE 2010-4029). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. <http://ies.ed.gov/ncee/pubs/20104029/pdf/20104029.pdf>

Goodson, B., Wolf, A., Bell, S., Turner, H., & Finney, P. B. (2011). *Effectiveness of a Program to Accelerate Vocabulary Development in Kindergarten (VOCAB): First Grade Follow-up Impact Report and Exploratory Analyses of Kindergarten Impacts* (NCEE 2012-4009). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. http://ies.ed.gov/ncee/edlabs/regions/southeast/pdf/REL_20124009.pdf

Hanson, T., Dietsch, B., and Zheng, H. (2012). *Lessons in Character Impact Evaluation*. (NCEE 2012-4004). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. <http://ies.ed.gov/ncee/edlabs/regions/west/pdf/20124004.pdf>

Heller, J.I. (2012). *Effects of Making Sense of SCIENCE™ professional development on the achievement of middle school students, including English language learners*. (NCEE 2012-4002). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. http://ies.ed.gov/ncee/edlabs/regions/west/pdf/REL_20124002.pdf

Heppen, J.B., Walters, K., Clements, M., Faria, A., Tobey, C., Sorensen, N., and Culp, K. (2012). *Access to Algebra I: The Effects of Online Mathematics for Grade 8 Students*. (NCEE 2012-4021). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. http://ies.ed.gov/ncee/edlabs/regions/northeast/pdf/REL_20124021.pdf

Hitchcock, J., Dimino, J., Kurki, A., Wilkins, C., and Gersten, R. (2010). *The Impact of Collaborative Strategic Reading on the Reading Comprehension of Grade 5 Students in Linguistically Diverse Schools*. (NCEE 2011-4001). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. http://ies.ed.gov/ncee/edlabs/regions/southwest/pdf/REL_20114001.pdf

Institute of Education Sciences. *Final Report on the National Assessment of Title I: Summary of Key Findings*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, 2007. http://ies.ed.gov/ncee/pdf/20084014_rev.pdf

James-Burdumy, S., Deke, J., Lugo-Gil, J., Carey, N., Hershey, A., Gersten, R., Newman-Gonchar, R., Dimino, J., Haymond, K., and Faddis, B. (2010). *Effectiveness of Selected Supplemental Reading Comprehension Interventions: Findings From Two Student Cohorts* (NCEE 2010-4015). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. <http://ies.ed.gov/ncee/pubs/20104015/pdf/20104015.pdf>

James-Burdumy, Susanne, Brian Goesling, John Deke, and Eric Einspruch (2010). *The Effectiveness of Mandatory-Random Student Drug Testing* (NCEE 2010-4025). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. <http://ies.ed.gov/ncee/pubs/20104025/pdf/20104025.pdf>

Judkins, David, Robert St.Pierre, Babette Gutmann, Barbara Goodson, Adrienne von Glatz, Jennifer Hamilton, Ann Webber, Patricia Troppe, and Tracy Rimdzius. *A Study of Classroom Literacy Interventions and Outcomes in Even Start* (NCEE 2008-4028). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. <http://ies.ed.gov/ncee/pubs/20084028/pdf/20084028.pdf>

Kushman, Jim, Makoto Hanita, and Jacqueline Raphael. *An Experimental Study of the Project CRISS Reading Program on Grade 9 Reading Achievement in Rural High Schools: Final Report*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, 2011. http://ies.ed.gov/ncee/edlabs/regions/northwest/pdf/REL_20114007.pdf

Martin, Taylor, Sarah J. Brasiel, Herb Turner, and John C. Wise. (2012). *Effects of the Connected Mathematics Project 2 (CMP2) on the Mathematics Achievement of Grade 6 Students in the Mid-Atlantic Region* (NCEE 2012-4017). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. http://ies.ed.gov/ncee/edlabs/regions/midatlantic/pdf/REL_20124017.pdf

Newman, D., Finney, P.B., Bell, S., Turner, H., Jaciw, A.P., Zacamy, J.L., & Feagans Gould, L. (2012). *Evaluation of the Effectiveness of the Alabama Math, Science, and Technology Initiative (AMSTI)*. (NCEE 2012-4008). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. http://ies.ed.gov/ncee/edlabs/regions/southeast/pdf/REL_20124008.pdf

Preschool Curriculum Evaluation Research Consortium (2008). *Effects of Preschool Curriculum Programs on School Readiness* (NCER 2008-2009). Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education. Washington, DC: U.S. Government Printing Office. http://ies.ed.gov/ncer/pubs/20082009/pdf/20082009_rev.pdf

Randel, B., Beesley, A. D., Aphorp, H., Clark, T.F., Wang, X., Cicchinelli, L. F., & Williams, J. M. (2011). *Classroom Assessment for Student Learning: The impact on elementary school mathematics in the Central Region*. (NCEE 2011-4005). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. http://ies.ed.gov/ncee/edlabs/regions/central/pdf/REL_20114005.pdf

Rolfhus, E., Gersten, R., Clarke, B., Decker, L., Wilkins, C., and Dimino, J. (2012). *An evaluation of Number Rockets: A Tier 2 intervention for grade 1 students at risk for difficulties in mathematics*. (NCEE 2012-4007). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. http://ies.ed.gov/ncee/edlabs/regions/southwest/pdf/REL_20124007.pdf

Silvia, S., Blitstein, J., Williams, J., Ringwalt, C., Dusenbury, L., and Hansen, W. (2010). *Impacts of a Violence Prevention Program for Middle Schools: Findings From the First Year of Implementation* (NCEE 2010-4007). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

<http://ies.ed.gov/ncee/pubs/20104007/pdf/20104007.pdf>

Social and Character Development Research Consortium (2010). *Efficacy of Schoolwide Programs to Promote Social and Character Development and Reduce Problem Behavior in Elementary School Children* (NCER 2011-2001). Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.

<http://ies.ed.gov/ncer/pubs/20112001/pdf/20112001.pdf>

Somers, M.-A., Corrin, W., Sepanik, S., Salinger T., Levin, J., and Zmach, C. (2010). *The Enhanced Reading Opportunities Study Final Report: The Impact of Supplemental Literacy Courses for Struggling Ninth-Grade Readers* (NCEE 2010-4021). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

<http://ies.ed.gov/ncee/pubs/20104021/pdf/20104021.pdf>

Weinstock, P., Bos, J., Tseng, F., Rosenthal, E., Ortiz, L., Dowsett, C., et al. (2012). *Evaluation of Program for Infant/Toddler Care (PITC): An On-site Training of Caregivers* (NCEE 2012-4003). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

http://ies.ed.gov/ncee/edlabs/regions/west/pdf/REL_20124003.pdf

Wijekumar, K., Hitchcock, J., Turner, H., Lei, P.W., and Peck, K. (2009). *A Multisite Cluster Randomized Trial of the Effects of CompassLearning Odyssey® Math on the Math Achievement of Selected Grade 4 Students in the Mid-Atlantic Region* (NCEE 2009-4068). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

http://ies.ed.gov/ncee/edlabs/regions/midatlantic/pdf/REL_20094068.pdf

Wilkerson, S. B., Shannon, L. C., Styers, M. K., and Grant, B. (2012). *A study of the effectiveness of a school improvement intervention (Success in Sight)*. (NCEE 2012-4014). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S.

Department of Education. http://ies.ed.gov/ncee/edlabs/regions/central/pdf/REL_20124014.pdf

Wilkins, C., Gersten, R., Decker, L., Grunden, L., Brasiel, S., Brunnert, K., and Jayanthi, M. (2012). *Does a summer reading program based on Lexiles affect reading comprehension?* (NCEE 2012-4006). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

http://ies.ed.gov/ncee/edlabs/regions/southwest/pdf/REL_20124006.pdf

Wolf, Patrick, Babette Gutmann, Michael Puma, Brian Kisida, Lou Rizzo, Nada Eissa, and Matthew Carr. *Evaluation of the DC Opportunity Scholarship Program: Final Report* (NCEE 2010-4018). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. <http://ies.ed.gov/ncee/pubs/20104018/pdf/20104018.pdf>