

Modeling Instruction has a Large Effect Size (Cohen's D)

Modeling Instruction meets the standard for moderate evidence established by the What Works Clearinghouse. The effect size is large: 0.91 for more than 1600 students in regular first-year physics courses in public high schools. (Analysis by Eric Brewe, Ph.D. in 2010.)

Quantitative evidence supporting the impacts of Modeling Summer Institutes has been established through a well-designed, well-implemented study undertaken as part of the Modeling Workshop Project evaluation (Hestenes 2000). Findings of this study have been cited many times and used in national-level, influential reports (Clewell 2004) and by the U.S. Department of Education to identify Modeling Instruction as one of two Exemplary K-12 Science Programs (Expert Panel Review 2001).

Study Design. The study was a quasi-experimental, repeated measures design. The repeated measures design maintains a high level of validity by reducing variability. In this design, teachers who applied to participate in the Modeling Workshops were asked to administer the Force Concept Inventory (FCI) (Hestenes et al. 1992) to their students near the end of the school year (i.e., post-instruction) prior to participating in the Modeling Workshop during the subsequent summer. The FCI measurement was then repeated with different students in the same course (which minimized practice effects). Teachers who then participated in a second Modeling Workshop in the next summer repeated the measurement. Teachers, teaching in the same school with similar students, matched the data, and the approximate equivalence of the students pre-instruction was established. This was done because pre-instruction FCI scores consistently show no difference on the average (25.38% - 26.80%). Together, the measured equivalence criteria and the repeated measures design established this quasi-experimental study as meeting the standard for moderate evidence established by the What Works Clearinghouse (WWC 2008).

Instruments and Measures. The study utilized the Force Concept Inventory, a 30-question conceptual test that has been the standard instrument for evaluating conceptual understanding of introductory mechanics since publication (Hake 1998). Estimates of reliability employing Cronbach's alpha measured on the FCI posttest range from the mid 0.80s to the mid 0.90s and average higher than $\alpha = 0.85$, which provide evidence for the reliability of the FCI (Osborn Popp 2000, Hake 2002).

Data Collection. Data for this study were collected at two points: first, baseline FCI posttest data were collected from teachers who applied to attend a Modeling Workshop – these are the comparison data. Then data were collected again, after teachers had completed a Modeling Workshop – these are the treatment data. Baseline data (i.e., data from the comparison group) were collected during the spring semester of 1996-1997 from 1,653 students, and data from the treatment group were collected from 2,018 students during fall and spring of 1998-1999. The matching criteria applied were: data were collected only from students of the same 26 teachers who were in the same school and teaching the same first-year physics course in the baseline and treatment group, in order to maximize the likelihood of equivalence of groups prior to instruction. (The course was regular physics for 22 teachers, honors physics for two teachers, 9th grade physical science for one teacher, and principles of technology for one teacher.)

Data Analyses. In order to establish the effects of a Modeling Workshop on student

understanding, the study used the post-instruction FCI score achieved by students as the dependent variable. The independent variable was whether the students' teacher had previously completed two four-week summer Modeling Workshops. These data were analyzed using a two sample t-test for equivalence of means, to compare the FCI scores of students at the end of the school year but prior to their teacher attending a four-week summer Modeling Workshop, with the FCI scores of students at the end of the school year after their teacher attended their second four-week summer Modeling Workshop. **The before-workshop average of 42% (N=1,653) was significantly different than the after-workshop average of 53% (N=2,018), $p < 0.001$.** Furthermore, these differences between before- and after-workshop averages were all large as indicated by the effect size, $d = 0.91$. The 95% confidence interval on effect size was 0.84 – 0.98.

Summary of Evidence. The data supporting the impact of Modeling Workshops on high school physics teachers clearly meet the evidence standards with reservation as established by the What Works Clearinghouse. This is the maximum attainable from a quasi-experiment. (The study design did not include random assignment to treatment groups, but did include checks of group equivalence. It is difficult to conduct an experimental design in physics, for the vast majority of high schools have only one physics teacher. Rural schools typically have one section of physics.)

Furthermore, the study meets the criteria for drawing causal conclusions. *Temporal precedence* is clear because student scores rise only after the teacher attends the Modeling Workshop. The repeated measures design of the study ensures *covariance of cause and effect*, as student scores rise only after the teacher attends the Workshop. Further, because the sample size is substantial, the study is resilient to threats of internal validity such as selection bias or single sample threats.

References cited:

- Clewell., B., et. al. (2004). Review of Evaluation Studies of Mathematics and Science Curricula and Professional Development Models. (Urban Institute study commissioned by the GE Foundation). <http://www.urban.org/publications/411149.html>
- Expert Panel Review (2001): Modeling Instruction in High School Physics. (Office of Educational Research and Improvement, U.S. Department of Education, Washington, DC). See <http://modeling.asu.edu/ModlInstr-ExemplarySci-2001.pdf>
- Hake, R (1998). Interactive-engagement vs. traditional methods: A six thousand-student survey of mechanics test data for introductory physics courses. *Am. J. Phys.* **66**, 64-74.
- Hake, R. (2002), Lessons from the physics education reform effort, *Ecology and Society* **5**(2): 28; online at <http://www.ecologyandsociety.org/vol5/iss2/art28/>. *Ecology and Society* is a free online "peer-reviewed journal of integrative science and fundamental policy research" with tens of thousands of subscribers in many nations.
- Hestenes, D., Wells, M., and Swackhamer, G. (1992). Force Concept Inventory, *The Physics Teacher* **30**, 141-158. <http://modeling.asu.edu/R&E/Research.html>
- Hestenes, D. (2000). Findings of the Modeling Workshop Project (1994-2000) (from Final Report submitted to the National Science Foundation, Arlington, VA). <http://modeling.asu.edu/R&E/ModelingWorkshopFindings.pdf>
- Osborn-Popp, S. (2000). Personal Communication via e-mail on October 16, 2000 with Modeling Workshop Project Statistical Consultant.
- What Works Clearinghouse (2008). WWC Procedures and Standards Handbook v 2.0.

End note:

This document is a slightly modified excerpt by Sharon Osborn Popp and Jane Jackson, of the *Investing In Innovation (I3)* proposal to the U.S. Department of Education by the American Association of Physics Teachers. Submitted May 12, 2010. Unfunded. CFDA # 84.396B PR/Award # U396B100210. Title: Teacher Enhancement for STEM Education Reform: A National Network of Modeling Instruction Sites. Philip W. Hammer, Principal Investigator. The excerpt is from the narrative section on pages 14 to 16, entitled “Strength of Research, Significance of Effect, and Magnitude of Effect: Moderate Evidence in Support of a Validation Grant”, **written by Eric Brewster, Ph.D., science education faculty at Florida International University**. Data were given him by Sharon Osborn Popp, Ph.D., internal evaluator of the Modeling Workshop Project at Arizona State University (1995-2000). Data were Phase II teachers of Leadership Modeling Workshops (summers 1997 & 1998 at University of Wisconsin at River Falls, University of Akron - Ohio, and Arizona State University).

Demographics (contributed by Jane Jackson): Of the 26 teachers who contributed matched data,

- 22 taught regular first-year physics, two taught honors first-year physics, one taught 9th grade physical science, and one taught principles of technology.
- 25 taught at public schools, and one at an independent school.
- 18 taught at suburban and/or middle class schools, three at small rural schools, and three at urban high-poverty schools.
- 18 teachers were male.
- They taught in 16 states across the nation (eight taught in Wisconsin, three in Arizona).
- 18 teachers did not have a degree in physics or physics education. Content areas of degrees were: four chemistry, four physical/general science, three life science, two math, two engineering, and one each of elementary education, dietetics, and meteorology.

[Posted at <http://modeling.asu.edu> in December 2014. Clarified and re-titled in August 2018 by Jane Jackson.]