



Dr. Michael Herrick

HerrickResearch

The Critical Role of Assessment in Experimental and Quasi-Experimental Designs

Teaching American History
Project Directors Meeting

New York City

January 6-8, 2009

TEACHING AMERICAN HISTORY GRANTS

- LEARNING BY DOING (LBD)
- MAKING AMERICANS, MAKING AMERICA (MAMA)

WHY QUASI-EXPERIMENTAL?

- Federal government requires experimental designs
- Random assignment to groups not feasible
- Matched control groups was possible

GPRA Performance Goal

Students in experimental and quasi-experimental studies of educational effectiveness of Teaching American History projects will demonstrate higher achievement on course content measures than students in control and comparison groups

INSTRUMENTATION

LBD

- Downloaded NAEP items
- Assembled test

MAMA

- Project participants wrote test items
- Assembled draft of test
- Piloted tested items
- Conducted item analysis to eliminate/
revise bad items
- Finalized test

METHODOLOGY

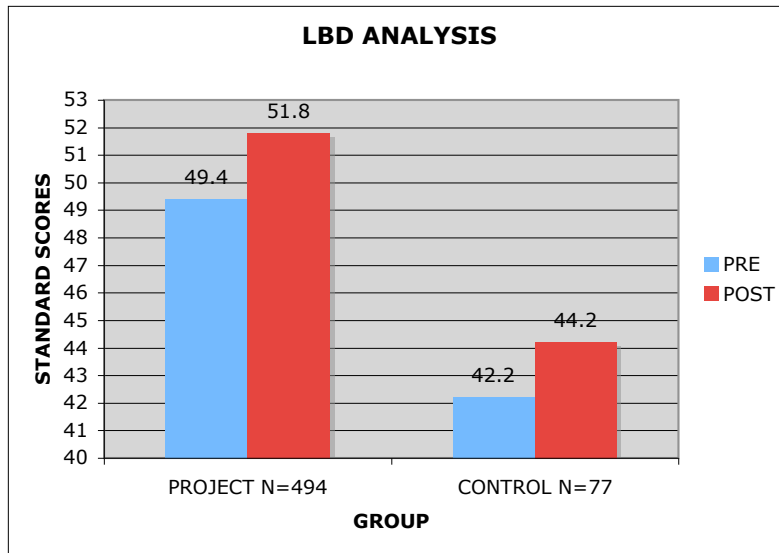
LBD

- Downloaded NAEP items
- Assembled test
- Matched demographics of project schools with non project schools
- Obtained control group schools
- Assigned school IDs for pre/post matching
- Administered tests to both groups
- Scored tests
- Matched school and student IDs
- Compared standard scores from pre to post for LBD and control groups
- Computed statistical difference between groups using SPSS

MAMA

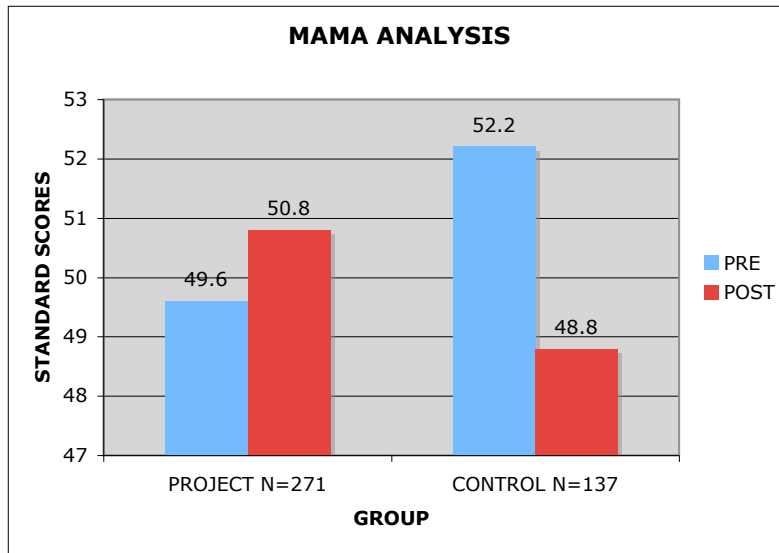
- Project participants wrote test items
- Assembled draft of test
- Pilot tested items
- Conducted item analysis to eliminate/revise bad items
- Finalized test
- Matched demographics of project schools with non project schools
- Obtained control group schools
- Assigned school IDs for pre/post matching
- Administered tests to both groups
- Scored tests
- Matched school and student IDs
- Compared standard scores from pre to post for MAMA and control groups
- Computed statistical difference between groups using SPSS

LBD ANALYSIS



- Post test scores increased in both Project and Control groups
- Difference in mean “gains” between groups were not statistically significant at .05 level.

MAMA ANALYSIS



- Project group: post test scores increased
- Control group: post test scores declined
- Difference between groups of mean “gains” from pre to post test were statistically significant at .01 level.

CONCLUSIONS AND IMPLICATIONS

- Assessment choice in evaluation makes a difference in results
- Alignment of assessment content to project content is key
- Standardization of assessment is not as important to evaluation quality as is content validity of assessment
- Program participants are in best position to write valid assessments