Multiple Methods for Examining Outcome Data with Implementation Data: Alternatives for Determining Key Factors of Effectiveness

Wendy Wolfersteig,
Arturo Valdivia, Albert Kopak
Southwest Interdisciplinary Research Center
Arizona State University

Wendy.Wolfersteig@asu.edu

ASU
Arizona State University

Outcome Measures

- Standardized test scores are typical
- But --- how do these relate to other measures of the project effectiveness?
- Can we measure impact of other variables on the test score outcomes?

Example 1

- Exploratory - To determine the relationships between the Arizona Reading First (K-3 reading) program implementation factors from teacher surveys and the reading performance for students who received the program during three consecutive years
- What key factors of implementation had significantly impacted the outcome of high test scores on the DIBELS assessment.

Example 1

Two steps were undertaken to conduct this study on key implementation factors to relate to DIBELS scores.

1. Exploratory factor analysis was used on the 2008 teacher survey in order to find a number of latent factors.
   - seven factors emerged: Coach (10 items), Meetings (5 items), Buy-in (11 items), Assessment (6 items), Fidelity-Modify Core (11 items), Observations (4 items) and Instruction (6 items).
2. Regression analysis obtained using the least squares method was used in a linear model.
   - the dependent variable was the last result of the ORF DIBELS test (ORF end-final for second or third grade, depending on the group),
   - the regressors (independent variables) were the DIBELS test from previous periods (in each group) and the teacher survey factors found during the first step.

Example 1

- Assessment was the only implementation factor of 7 that was statistically significant for both the second and third grade results. These questions focused on having an assessment system in place. Sig 2nd – Coach & Observations; Sig 3rd – Fidelity/Modify Core Curr
- Using end of year ORF for second grade as the outcome variable, the 12 independent variables had an adjusted R^2 = .764 ; the variables account for 76% of the variance of the DIBELS test score. The ANOVA showed the F-test was significant (F = 679.264, p = .000), therefore the hypothesis that all the regressor parameters were zero was rejected.
- Using end of year ORF for third grade as the outcome variable, the 12 independent variables had an adjusted R^2 = .804; the variables account for 80% of the variance of the DIBELS test score. The ANOVA showed the F-test is significant (F = 993.361, p = .000), therefore the hypothesis that all the regressor parameters were zero was rejected.

Example 2

- Additional analyses were needed to provide confirmation and validation for those results: 2009 Teacher Survey, same 7 factors

- *Multiple Regression Results:* Ordinary least squares (OLS) regression models with teacher factors predicting students' end of year 'raw' test scores - estimated separately for each grade level.

- *Logistic Regression* was conducted to identify which teacher factors were associated with students' odds of being consistently 'categorized' as *Benchmark* readers compared those students whose reading proficiency decreased from one assessment (fall 2008) to the next (spring 2009).

Example 2

Summary of Significant OLS and Logistic Regression Results with Teacher Factors Predicting Students' End of Year DIBELS Scores and Categories by Grade Level

	Scores: OLS Positive	Scores: OLS Negative	Category: Regression Positive	Category: Regression Negative
Assessment	K	-2, -3	+K, +1	--
Observations	--	-K	+3	
Instruction	+2	--	--	--
Meetings	--	-1, -3	--	-1, -2, -3
Professional development	+K, +3	--	--	--
Buy-in	+K, +1, +2	-3	+1, +2,	-3
Coach	--	-K	+3	--

Example 2

Summary of study 2

- It appears that the effect of teacher factors on DIBELS reading test scores and or benchmark score category status is grade-specific.
- Some factors were positively associated with test score outcomes in certain grades while the same factors were negatively associated with test scores for other grades
- These findings reiterate the notion that certain teacher factors may have a greater influence on students' test scores according to grade level.
- Indeed, the inconsistency in these results could also be cautiously viewed as preliminary evidence that suggests that students, based on their grade level, respond differently to teacher factors compared to students in other grades.

Example 3

Same question of what impacts students' test scores

Different data sets:

- AIMS test data: 3rd – 8th grade raw scores for reading, writing, math
- T4S: teacher observations in 7 areas; 4 composite areas used for this study
  - Communicates selected standards or objectives to all students
  - Key vocabulary emphasized
  - Ensure student engagement throughout the learning
  - Use formative assessment to determine the instructional needs of all students

Example 3

- Mean T4S score for each grade in a given school (all 3rd graders by school got the same average T4S sum score (aggregated by both grade and school; condensed in order to merge with test score data)
- STATA v 10
- AIMS raw scores by grade by school match given T4S mean score by grade by school
- Follows exactly what we did in Example 2, so there is a sense of consistency
- Ran a bi-variate correlation and OLS regression to assess the relationship between the T4S score and the raw writing, math, and the reading scores.

Example 3

Regression coefficients

- T4S with reading b=0.9362275, t(4601)=4.14, p<.01 – for every one point increase in T4S mean scores there is a 0.936 increase in reading test scores

- T4S with writing b=0 .1552708, t(4604)=1.45, p=.14 – for every one point increase in T4S mean scores there is a 0.1552708 increase in writing test scores (not sig)

- T4S with math b=1.224186, t(4602)=4.33, p<.01 – for every one point increase in T4S mean scores there is a 1.224186 increase in math test scores

Conclusion

- Possible to examine the data relationships from various statistical perspectives to determine effectiveness of implementation strategies on test score outcomes.
- Lesson: Manipulation of data elements needs much time and consideration
  - Development of factors from survey or instruments
  - Matching of different types of scores with different teachers, students, grade categories, etc.