

Program Evaluation using Virtual Comparison Group Analyses



SC Formative Assessment Conference

Nov. 19 and 20, 2009

Michael P. Dahlin and Jeff C. Marshall



Agenda

- Introduction
- VCG Uses
- VCG Construction
- Results and Interpretation
- Strengths and Weaknesses



Introduction

- Randomized, Controlled Experimental Design
 - Gold standard for evaluating new programs or curricula
 - Often not realistic in educational settings
 - Random assignments not possible
 - Teachers aware of group assignments

Introduction, continued



- Virtual Comparison Groups
 - Alternative to experimental design
 - Compares student growth to a group of similar students within NWEA's Growth Research Database (GRD)
 - Captures roughly 10% of the U.S. Student population
 - Spans roughly 12 years
 - key demographic data plus academic growth data on math, reading, language usage and the sciences

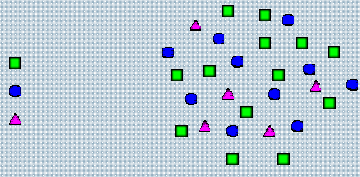
VCG Uses

Appropriate uses:



- Tailored Normative Growth Targets
 - School population differs greatly from published norms
- Program Evaluation Examples –
 - Did students who received a new math curriculum demonstrate greater than “typical” math growth?
 - Did a PD program lead to accelerated growth for students of participating teachers?

VCG Construction Process:



- Identify treatment group
- Identify all matching students from GRD
- Randomly select control group

Possible School Selection Criteria



- School Type (NCES)
 - Regular, Special Ed., Vocational, etc.
- School Status (NCES)
 - Charter, Magnet, Title I
- Urban-Rural Classification (NCES)
- % students receiving free or reduced price lunches (NCES)
- % students who are classified as below the poverty line (SAIPE)
- State Location



Possible Student Selection Criteria



- Grade
- Subject
- Test Administration Dates (pre- and post)
- Initial Test Score
- Match on a prior year's test score
- Gender
- Ethnicity



Filter level grouping



Level	RIT Difference	SES %Difference	Date Range	Use Prior Year
1	+/- 1	+/- 5%	+/- 7 days	N
2				Y
3	+/- 2			
4			+/- 10 days	
5		+/- 10%		
6		+/- 15%		
7	+/- 3			
8			+/- 12 days	
9	+/- 4			
10	+/- 5			

Results and Interpretation



- [Performance by School](#)
- [Performance by RIT Band](#)
- [Performance by Ethnicity](#)
- [Link to Pivot Table Results](#)

Strengths and Weaknesses



- Strengths
 - Additional demographic filters are possible than with traditional NWEA norms
 - Can more easily account for the amount of instructional time received
- Weaknesses
 - Certain combinations of demographics not well represented within GRD, resulting in few or no appropriate VCG matches
 - Validation process still on-going, though preliminary findings are encouraging

For More Information, Contact:



Kingsbury Center at NWEA:

<http://www.nwea.org/our-research/-kingsbury-center-nwea>

Michael Dahlin: michael.dahlin@NWEA.org

Jeff Marshall: marsha9@clemson.edu