HOW STATES ARE IMPLEMENTING THE ASSESSMENT AND ACCOUNTABILITY REQUIREMENTS OF NCLB

Stanley Rabinowitz, Ph.D.

Presentation at the 2002 Edward F. Reidy, Jr., Interactive Lecture Series
October 10-11, 2002, Nashua, NH
1. What must states/local districts do to be in compliance with the assessment and accountability provisions of the Act?

2. What provisions of the Act are still in need of clarification/regulation?

3. Is there a difference between the *letter* and *spirit* of the Act?

Still many unresolved questions

RILS 2002
PRESENTATION SET OF QUESTIONS

1. What are states doing?
2. What are states thinking about doing?
3. What should states be thinking about doing… (but may not have gotten to yet)?

Presentation not inclusive but addresses issues interesting and/or important (to me)

Interactive Session: what other issues fall within categories 1 - 3?
ASSESSMENT ISSUES VS ACCOUNTABILITY ISSUES

Set up as separate provisions (legislation, negotiated rule making, draft regulations) but...

- Where are they consistent?
- Where do they interfere with each other?
- On the one hand (*flexibility*—*assessment*)...
- On the other hand (*rigidity*—*accountability*)
- Interaction of decisions
- Informal vs. Formal Sanctions
ASSESSMENT ISSUES VS ACCOUNTABILITY ISSUES (continued)

- Assessment System relationship to Accountability System
  - Transferability of terms and concepts
  - Evaluation
  - Coherence (Alignment)
  - Tradeoffs (reliability vs. validity)
ASSESSMENT:
FILLING THE GAPS

Already compliant states vs. Gap states

Strategies

– Full CRT
– Augmented NRT
– CRT Light
– CRT + NRT (different grades)
– CRT + NRT (same grades)
ASSESSMENT: FILLING THE GAPS (continued)

Issues
- Alignment: different for each option and NAEP
- When to add additional tests (grades, content)
- How to add additional tests
- What about science, social studies, etc.?
- High School:
  Exit Exams: first time vs. cumulative passing rates
  End-of-Course Exams:
   multiple options vs. core
   minority student participation rates
- Local assessments
- Spring vs. Fall assessment
- How to use 2006 date?
ASSESSMENT: SETTING PERFORMANCE LEVELS

- Most direct interaction with accountability
- Existing programs vs. new/evolving programs
- World-class standards vs. Title I
- State vs. Federal definitions of Proficiency
- How many levels to set?
  - Overall
  - Above Proficient
ASSESSMENT:
STANDARD SETTING

What is the expected relationship across grades and content areas?

Are there known problem grades/content areas or successful initiatives?

Can you explain empirical differences?

Approaches:

• Grade by grade, content area by content area
• G x G, CA x CA, smooth
• Specific grades, extrapolate, smooth (3 – 8 vs. 3 – 5 – 8)

RILS 2002
ASSESSMENT:
STANDARD SETTING  (continued)

Multiple Methods – Pros and Cons
- Convergent Validity
- How to combine multiple, “unalike” data?

Conclusions
- Method less important than how well you do it
- Multiple methods primary advantage: provide degrees of freedom for adjustments
- Need coherence (interpretability, plausibility) across grades and content area
- Use a priori information (political and statistical) followed by impact data
ASSESSMENT: GRADE-BY-GRADE COMPARISONS

Comparative year-to-year student achievement data

Vertical Scale

- Full CRT Model
- Mixed CRT-NRT Model

Percent Proficient

- Weaknesses of Vertical Scales
  - Construct
  - Statistical
  - Practical

- Alignment requirements

Adjustments as new grades/content areas come on line

RILS 2002
ASSESSMENT:
SPECIAL EDUCATION

- Reasonable adaptations and accommodations
- Must provide one or more alternate assessments for students who cannot participate in regular assessment system
- Interaction with degree of disability?
- Inconsistent with accountability provisions
ASSESSMENT:
ENGLISH-LANGUAGE LEARNERS

• Assess in a valid and reliable manner:
  – Reasonable accommodations
  – Use of language and form “most likely to yield accurate and reliable information”

• Make every effort to develop linguistically accessible academic assessment measures
ACCOUNTABILITY: SETTING THE STAGE

• Principles of Reform
  – Most proposed models are defensible (though not necessarily NCLB-approvable)...if you accept their assumptions
  – Status vs. Improvement vs. Longitudinal vs. Mixed Models
  – Social Studies “test”
  – If you don’t know: Use NCLB
  – If you do know: Crosswalk to NCLB
ACCOUNTABILITY: SETTING THE STAGE  
(continued)

• Historical Context
  – Prior experience with accountability models
  – Process used to develop accountability models (prior and new)

• Practical Considerations
  – Identify right schools
  – Identify right number of schools
  – Interaction with minimum N considerations
ACCOUNTABILITY: SINGLE STATEWIDE SYSTEM

• Pre-NCLB accountability system coexistence
• All schools treated identically
• Title I schools vs. all schools
  – Procedures vs. Sanctions
  – Elementary vs. High School
• What about small schools?
• What about Alternative Schools?
• Need to translate local systems into NCLB language and intent
ACCOUNTABILITY: USE OF INDEX

• Cross Content Areas vs. Within Content Areas
• Number of levels (below and above Proficient)
• Values of levels (relative to Proficient): NCLB intent
• Determine point on index equivalent to % Proficient
  – Distributional simulations
  – 100 vs. 100+
• Time frame (...2006...2008......2014)

RILS 2002
ACCOUNTABILITY: AYP

- 2002 – 2014
  - Linear growth
  - Jagged growth
  - Curvilinear growth
- 95% vs. every child
- Status standard (% Proficient or Index)
- Use of “Safe Harbor” (% Proficient or Index)
- Multiple starting points
  - Set ceiling at school level
  - Perceptions of fairness
- Aggregation
  - Across years
  - Across sub-groups
- Adjustments as different grades and content areas added
- NAEP Audit

RILS 2002
ACCOUNTABILITY: MINIMUM N

• Reliability of Assessment System vs. Reliability of Accountability System
• Sufficient Reliability
• 20 – 50 (41)
• Alternate Questions: Credible results
  – What N identifies the “correct” schools?
  – What N does not exempt low performing schools and subgroups?
• 20…30…40.................................................................1000
• Multiple N problem

RILS 2002
CONCLUSIONS

• Still numerous unanswered questions
• Landscape continues to change
So…
• Develop and submit Assessment and Accountability Systems…
• Consistent with state Reform Principles and History…
• Crosswalked to NCLB Provisions and Intent
You can either…
• Submit plan you don’t want to do (and have it approved!)
• Submit Plan consistent with above principles
  – Approved: OK
  – Disapproved: begin the negotiations

RILS 2002