



Assessment System Design Topics and Issues

North Dakota Assessment Task Force

Barry Topol
Ed Roeber
John Olson

November 5, 2015



Presentation Overview

I. Potential North Dakota Assessment Vendors

- Vendor Overview
- Factors to Consider in Selecting an Assessment Vendor
 - Nature of the Assessment System
 - Size of the State Program and Number of Students to be tested
 - Technical Quality
 - Experience Working with States
 - Customer Support Requirements

II. Build Vs. Buy Vs. Hybrid Considerations

- Descriptions
- Advantages and Disadvantages
- Criteria
- Key Criteria Descriptions



Presentation Overview

III. Key Decision Points and Components of a State Assessment RFP

- Major Decision Points
- Timing
- Cost-related Issues

IV. Q&A

3



Potential Vendor Review

November 5, 2015 www.assessmentgroup.org 4

State Assessment Vendors



Vendor	Size	Strengths	Capabilities	Other
Pearson	Very Large	Operations PARCC vendor	Full (dev., ops, scoring, IT), build & buy	Largest vendor; 3 major losses in last 3 years (CA, TX, FL)
ETS	Very Large	Development, Psychometrics	Full, build and buy?	Big wins in 2015 (CA, TX)
DRC	Large	Customer Service	Full, build and buy	Just acquired CTB
AIR	Large	CBT system for SBAC**	Full, build and buy?	Lots of SBAC state wins
Measured Progress	Medium	Customer Service	Full, build and buy	Uses TAO and eMetric for CBT
Questar	Small	Technical Capability	Full, build and buy	New Management Team

** Proprietary system only. Open source system has issues.

5

State Assessment Vendors



Vendor	Size	Strengths	Capabilities	Other
Measurement, Inc.	Medium	Scoring (Human and AI)	Full – Might partner for buy or hybrid bids	Scoring subcontracting
ACT	Large	ACT; big district player; Psychometrics	Aspire is OTS program	Would likely partner with Pearson on hybrid opportunities
NWEA	Medium	CAT; MAP product	Large district player	Limited federal approval of MAP for use in state assessment

Many other specialized vendors including Renaissance Learning, Scantron, Curriculum Associates, College Board, etc.

ASG expects three to four vendors would bid on a North Dakota assessment program

6



The Consortium Option

- Two assessment consortia (SBAC, PARCC) formed in 2010 to develop the next generation of assessments
- Each consortium received roughly \$180M in development funding from USED
- 28 states and DC participated in the first year of testing which went relatively well
 - Assessments have more rigor than older tests
- Each consortium has its own approach to assessment and how its assessment is implemented in member states

7



SBAC & PARCC – Fast Facts

SBAC and PARCC Assessment Consortia

Consortium	Member States		Assessment Model	Testing Time	Testing Windows	CBT Test Price		Other Products	Implementation Methodology
	2015	2016				Base Product**	Other Products+		
SBAC Grades 3-8, 11	18	15	Computer Adaptive	Math 3-4 Hrs ELA 4-4.5 Hrs	2*, both parts given last 12 weeks of the year	~\$22.30	~\$5.20	Interim Assessment, Digital Library	SBAC responsible for content, Individual or group of states hire vendor for implementation
PARCC Grades 3-11; 2 HS Math Tracks	10 + DC	7 + DC	Fixed Form	Math 4-4.5 Hrs ELA 4+-5+ Hrs	1, at 90% of School Year	~\$25.00	TBD	Diagnostic Assessment, Speaking and Listening, Formative Tools	PARCC/Pearson responsible for content; Pearson for all implementation

* One window for multiple choice questions, one window for essay type questions
 ** Includes membership fee of \$6.20 per student for SBAC and \$1 per student for PARCC
 + Paper based tests are an additional \$10

8

Factors to Consider in Vendor Selection



- Nature of the Assessment System
 - Online (CBT) or Paper and Pencil Test (PPT)
 - Custom Developed or Off-the-Shelf (OTS) or Hybrid
- State Size/Student Count
- Technical Quality
 - Alignment to Standards (state, CCR)
 - Validity – Measures what it is supposed to
 - Comparability – Scores can be compared over years/programs
 - Psychometric Analysis

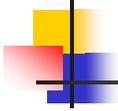
9

Factors to Consider in Vendor Selection



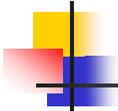
- Experience Working With States
- State Customer Service Requirements
 - Will the vendor give the state sufficient attention?
 - How big is the state's assessment department?
 - What level of project management and support will be required of the vendor?
- Build Vs. Buy
 - Capabilities and Product Offering(s)
 - Cost
 - Time to Market
 - State Ownership of Test Items/Test Security

10




Build versus Buy versus Hybrid

November 5, 2015
www.assessmentgroup.org
11




Definitions/Factors

- Build – Vendor develops a custom assessment specifically aligned to the states standards and requirements

Pros	Cons
<ul style="list-style-type: none"> Get what you want Alignment to standards High quality High amount of control <ul style="list-style-type: none"> Design – structure, content Timing of administration Delivery method (CBT/PPT) 	<ul style="list-style-type: none"> Cost – Likely most expensive option Time to Market – 2 years to implement due to development and field testing requirements Can not compare student results to those of other states

November 5, 2015
www.assessmentgroup.org
12



Definitions/Factors

- Buy – State purchases an “off the shelf” assessment from a vendor

Pros	Cons
<ul style="list-style-type: none"> Cost – Generally the low cost option Time to Market – Fastest (< 1 Yr.) Alignment to CCR Standards – Possible but state needs to verify Comparability – Perhaps in the future depending on the product 	<ul style="list-style-type: none"> Control – State has no control over content, design Quality – Items testing critical thinking skills is an issue Need to augment test for alignment? Detailed reporting information and sub-scores may not be available Tests in areas such as Science or Social Studies may not be available

November 5, 2015
www.assessmentgroup.org
13



Definitions/Factors

- Hybrid – A vendor combines an existing “off the shelf” test with custom-generated items. Custom items are field tested in Yr. 1 and available for use in Yr. 2.

Pros	Cons
<ul style="list-style-type: none"> Time to Market – Can launch quickly in year 1, using OTS test component Cost – less expensive than a full custom assessment Better alignment to state and/or CCR standards than OTS product Generally speaking, there is better quality and control than with an OTS product 	<ul style="list-style-type: none"> Quality – Varying quality of products by vendor. State needs to ensure the quality and alignment of the OTS component meets its needs. Comparability – Can only compare OTS items and only if enough other states use that component of the assessment

4



Build vs. Buy vs. Hybrid

Build Vs. Buy Vs. Hybrid - Summary										
Assessment Option	Cost*	Control			Online Requirement	Comparability W/other states	Quality+	Other Tools and Support	Internal Requirements	Time To Market**
		Structure	Content	Timing						
Custom Developed Assessment	Highest	High	High	High	Up to State	Low	High	Depends	Med- High	1 year to Dev. 1 year to FT
Buy Off the Shelf Assessment	Lowest	Low	Low	High	Most Likely Up to State	TBD	?	High	Lowest	6 months
Develop Hybrid Assessment	Middle	Med	Med	High	Most Likely Up to State	TBD	Med/High	Depends	Med- High	6-9 months

* Consider technology upgrade costs; Costs to replicate a CAT is prohibitive for an individual state unless items are shared
 + Quality and depth of alignment to College and Career Ready standards; Depth of Knowledge (DOK) of items; Reporting cabability
 ** From RFP award to test administration. Allow 2 to 5 months for RFP development and approval depending on state requirements/regulations

November 5, 2015
www.assessmentgroup.org
15



Key Criteria Description

- Time to Market – The amount of time it takes to have an assessment ready for administration
 - New items must be field tested on a prior year assessment before they can be used on an operational assessment so a custom test takes 2 years to fully implement
 - The assessment cycle is a full year, so ideally a contract for a custom test should be issued a year prior to actual test administration. A contract for an OTS or hybrid test should be issued 6-9 months prior to testing; additional time is needed to issue the RFP, analyze responses, and award the contract.

16




Key Criteria Description

- Quality – To some degree, quality is in the eye of the beholder but in the assessment world, quality is generally measured in 4 ways:
 - Alignment – How well does the test align to state and College and Career Ready (CCR) standards?
 - Validity – Is there evidence that the test measures what it says it measures? Do test items adequately test the intended construct and only the intended construct?

17




Key Criteria Description

- Comparability – Can results of this years assessment be compared to previous and future years? Can they be compared to other state's results that use the same instrument?
- Reporting – Can the results be validly reported at various levels (e.g., standards, sub-standards and items) to policymakers, administrators, teachers, parents and students?
- *Note: USED Peer Review also evaluates the quality of a state's assessment*

November 5, 2015 www.assessmentgroup.org 18



Key Criteria Description

- Control – Does the state or the vendor have control over:
 - The structure of the assessment?
 - Type of items
 - Administration mode (paper vs. online)
 - Content of the assessment?
 - Alignment to state and CCR standards
 - Nature/type of the items
 - Administration dates/windows?
 - Reporting content?

19



Key Decision Points and Components of a State Assessment RFP

November 5, 2015 www.assessmentgroup.org 20



State Assessment RFP

- A RFP is the best way to procure a state assessment regardless of the nature of the test (build, buy, hybrid)
 - Defines the scope of the project
 - Describes the products and services needed
 - Outlines legal and other requirements
 - Permits selection of the highest value solution
 - (An RFI is sometimes issued to gather information on products and services)
- **Several Key Decisions must be made and reflected in the RFP**

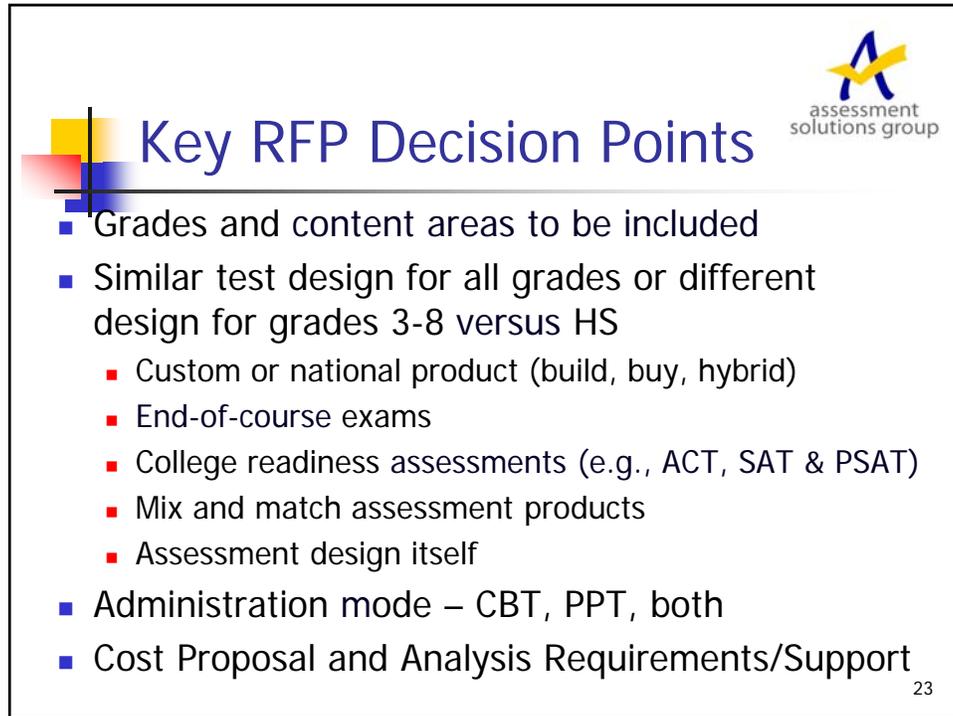
21



Key Decision - General or Specific RFP

Nature of RFP	Advantages	Disadvantages
Specific – State defines: <ul style="list-style-type: none"> • Test Design • Number of items and forms • Administration steps • Reporting requirements 	<ul style="list-style-type: none"> • State gets exactly what it wants • All responders bid the same program; responses are easily comparable 	<ul style="list-style-type: none"> • May preclude vendors from offering innovative alternatives • May preclude vendors from offering improvements to a program
General – State defines: <ul style="list-style-type: none"> • General outcomes desired • General rather than specific work statement 	<ul style="list-style-type: none"> • Encourages respondent creativity • Can generate several different designs from respondents 	<ul style="list-style-type: none"> • State may not get what it wants • Difficult to compare costs/value of different designs

November 5, 2015
www.assessmentgroup.org
22

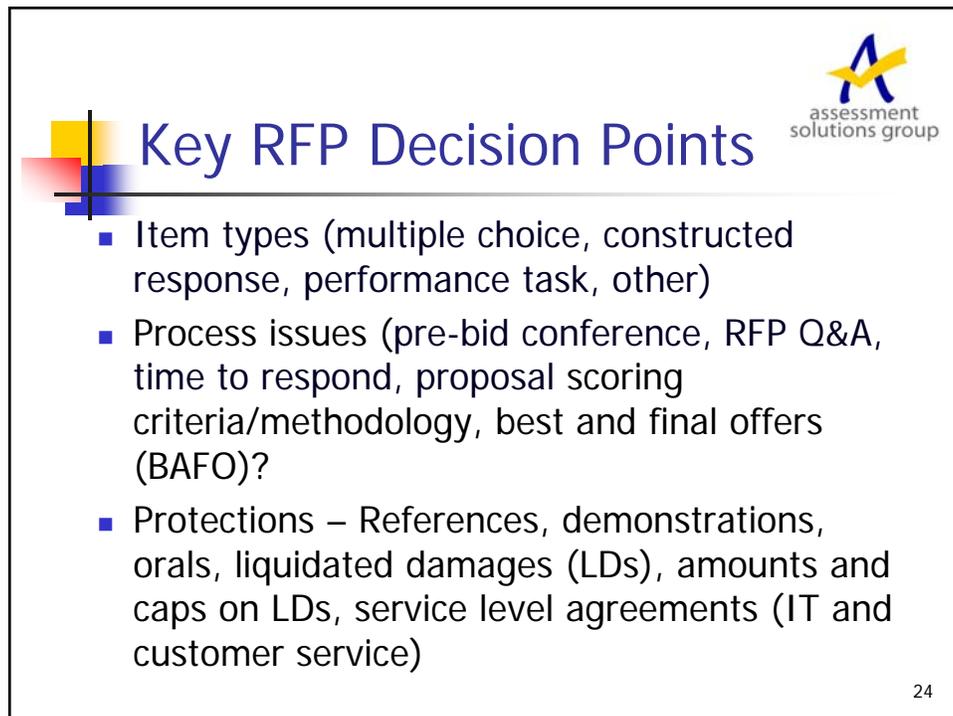


assessment solutions group

Key RFP Decision Points

- Grades and content areas to be included
- Similar test design for all grades or different design for grades 3-8 versus HS
 - Custom or national product (build, buy, hybrid)
 - End-of-course exams
 - College readiness assessments (e.g., ACT, SAT & PSAT)
 - Mix and match assessment products
 - Assessment design itself
- Administration mode – CBT, PPT, both
- Cost Proposal and Analysis Requirements/Support

23



assessment solutions group

Key RFP Decision Points

- Item types (multiple choice, constructed response, performance task, other)
- Process issues (pre-bid conference, RFP Q&A, time to respond, proposal scoring criteria/methodology, best and final offers (BAFO)?
- Protections – References, demonstrations, orals, liquidated damages (LDs), amounts and caps on LDs, service level agreements (IT and customer service)

24



ASG Recommends a 2-3 Month Process for Developing an RFP*

Week	Activity
Week 1	Collect relevant information and materials from state (incl. old RFPs); Meet with state staff to review key requirements
Weeks 2-3	Begin drafting RFP
Week 4	State review of 1 st draft (primarily scope of work)
Week 5	Revise
Weeks 6-7	State review 2 nd draft (including boilerplate and evaluation rubrics)
Week 8	Revise
Week 9	State review 3 rd and final draft with all parts included (cost sheets); Finalize RFP and deliver to state for release
Week 10	State RFP finalization activities; release RFP

*Will be longer if state review, approval and legal processes are extensive

25



Questions?

November 5, 2015 www.assessmentgroup.org 26