



# NC Public Schools – READY for Success

## Designing NC's Measures of Student Learning (Common Exams)

January 5, 2011  
State Board of Education

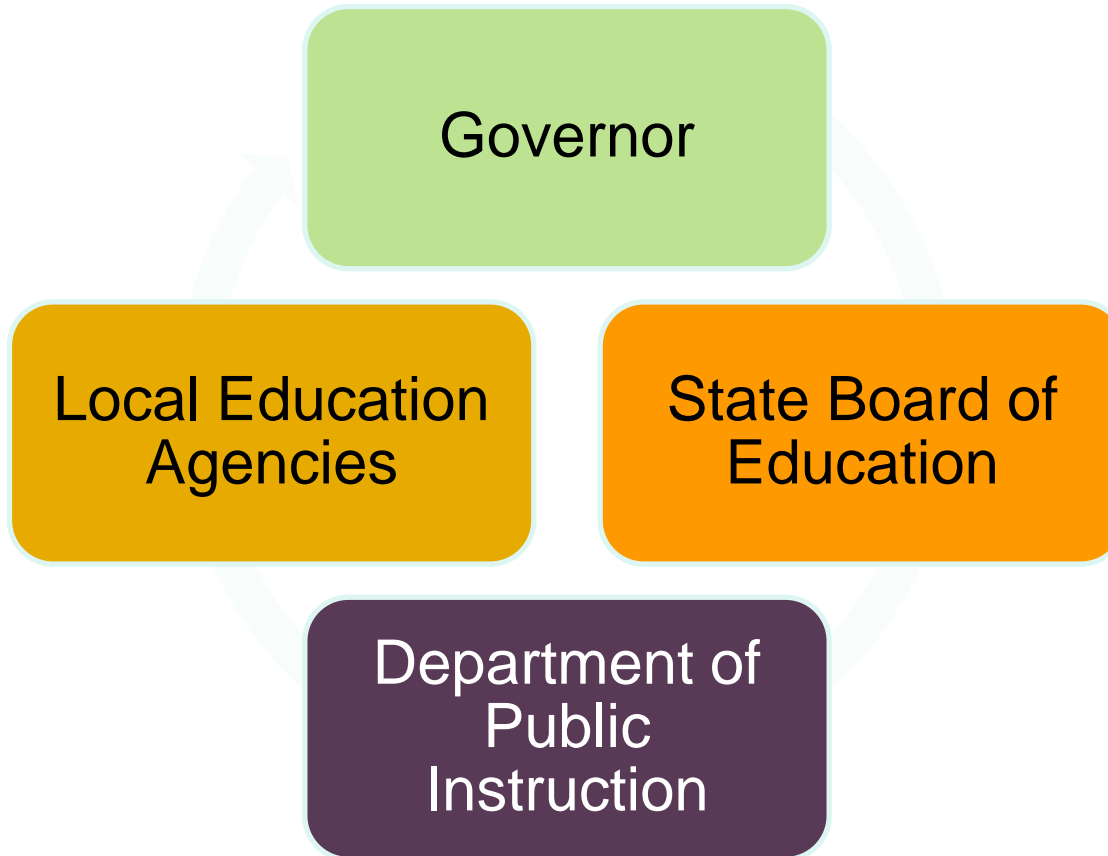
Dr. Rebecca Garland  
Dr. Lynne Johnson



# Overview

1. Context for Educator Effectiveness Work
  - What We Want
  - What We Have
  - What We Still Need
2. Measures of Student Learning (MSL) Quick Facts
3. MSL Design Process
4. Enthusiasm, Concerns, and Other Teacher Feedback
5. Challenges Ahead
6. Next Steps for 2011-12

# Who are “We?”



# Context for Educator Effectiveness Work



# What We Want:

Develop effective teachers and leaders in preparation programs

An effective teacher in every classroom and leader in every school

Achievement and growth for all students

Use professional development to increase effectiveness of teachers and leaders

# What We Have

## Statewide NC Educator Evaluation System

- All 115 school districts
- Many charter schools

## Standard, Statewide Measures of Student Growth and Achievement

- English Language Arts
- Mathematics
- Some years of Science
- Career and Technical Education



# What We Need

A required, standard, objective measurement of the effect each teacher has on each student's learning

Measures of student growth in all content areas

A common statewide process for identifying “effective” teachers and leaders



# What We Need

A required, standard, objective measurement of the effect each teacher has on each student's learning

Measures of student growth in all content areas

A common statewide process for identifying "effective" teachers and leaders

## **Completed (Summer 2011)**

Added Sixth and Eighth Standards

## **In Progress (2011-12)**

Measures of Student Learning Design Process  
(Phase One of Three Completed)

## **Targeted for Completion (February 2012)**

Second Set of State Board Policies on Educator Effectiveness

# Measures of Student Learning (MSL) Quick Facts

# Guiding Principles

- North Carolina's experienced teachers know their students and their content
- They are the best-qualified to provide input on meaningful assessment of currently non-tested grades and subjects
- Valid measures of what students know and are able to do will likely exceed traditional multiple-choice assessment



# What MS Ls Are

- Measures of what students know and are able to do after completing a course or grade
- Tightly linked to the instruction that a teacher delivers
- **One** part of how North Carolina will evaluate the effectiveness of its teachers
- Similar to the common summative assessments that many districts already have in place

# What MS Ls Are Not

- Multiple-choice standardized exams for all areas of the Standard Course of Study
- Assessments that need to be delivered with the same level of security as EOCs and EOGs
- Designed without teacher input
- The only source of data used to make decisions about a teacher's effectiveness
- Part of the school accountability model

# The Balancing Act

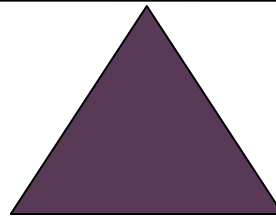
Given limited resources to dedicate to this work:

Freedom from Bias in  
Results

High Levels of Reliability

Ability to Feed Results into  
EVAAS or Another  
Sophisticated Growth  
Model

High Levels of Content  
Validity, with Performance-  
based Tasks



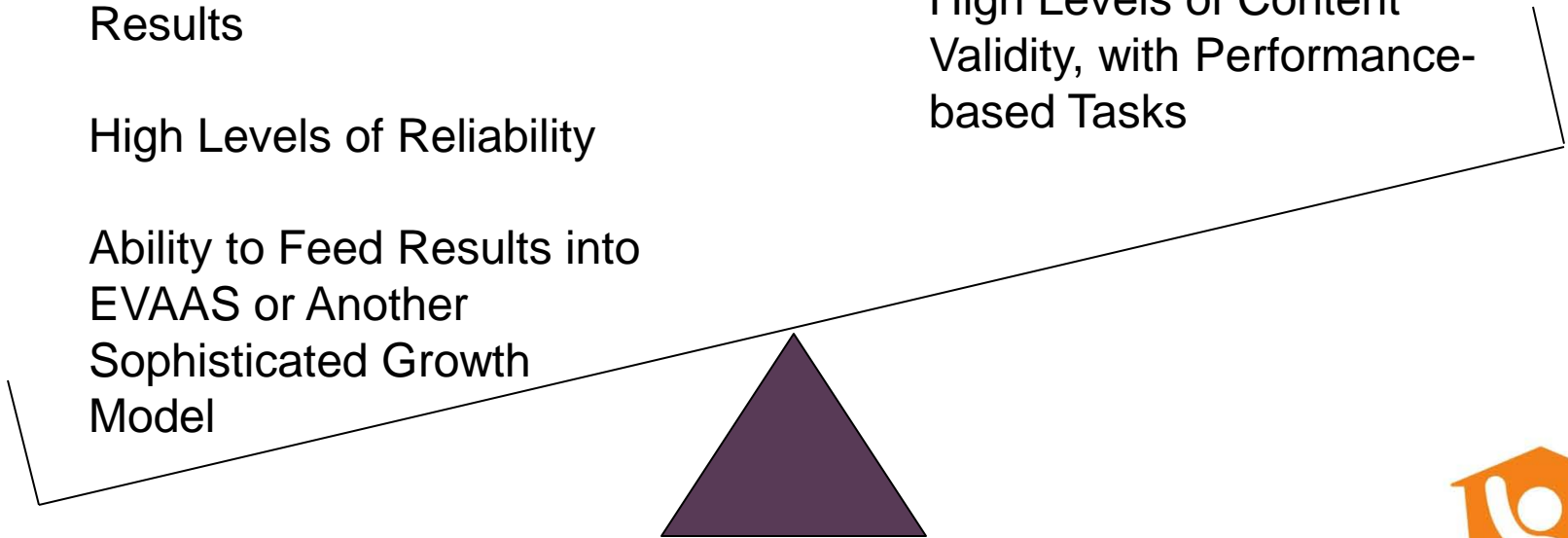
# The Balancing Act

Freedom from Bias in Results

High Levels of Reliability

Ability to Feed Results into EVAAS or Another Sophisticated Growth Model

High Levels of Content Validity, with Performance-based Tasks



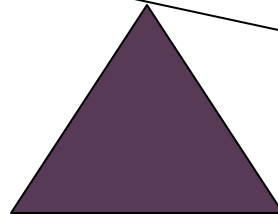
# The Balancing Act

Freedom from Bias in Results

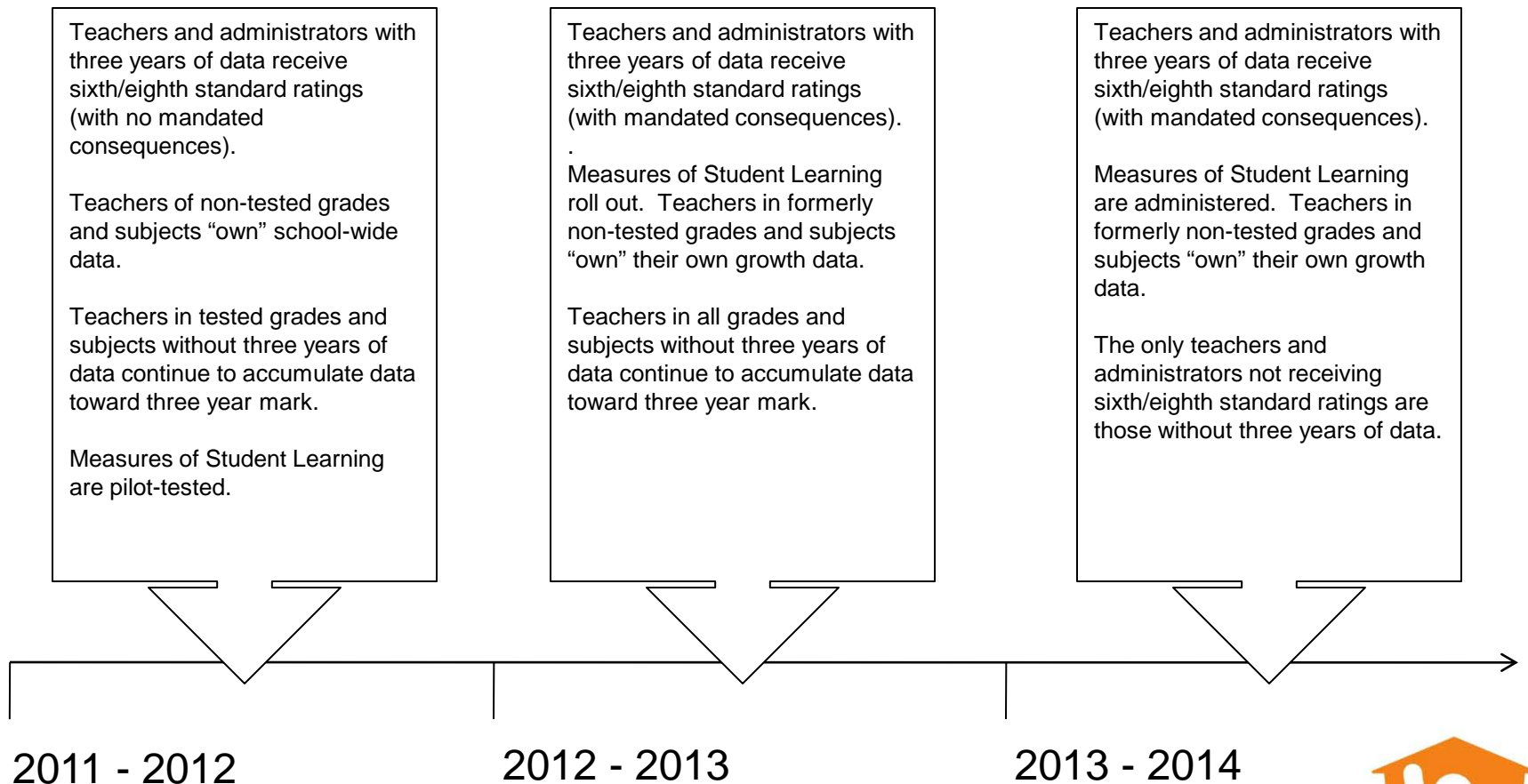
High Levels of Reliability

Ability to Feed Results into EVAAS or Another Sophisticated Growth Model

High Levels of Content Validity, with Performance-based Tasks



# Implementation Timeline



# Four “Buckets” of Assessment

1

EOCs, EOGs  
and VoCATS

2

Category One  
of MSLS

With appropriate  
resources and time,  
these MSLS can be  
validated  
psychometrically

3

Category Two of  
MSLS

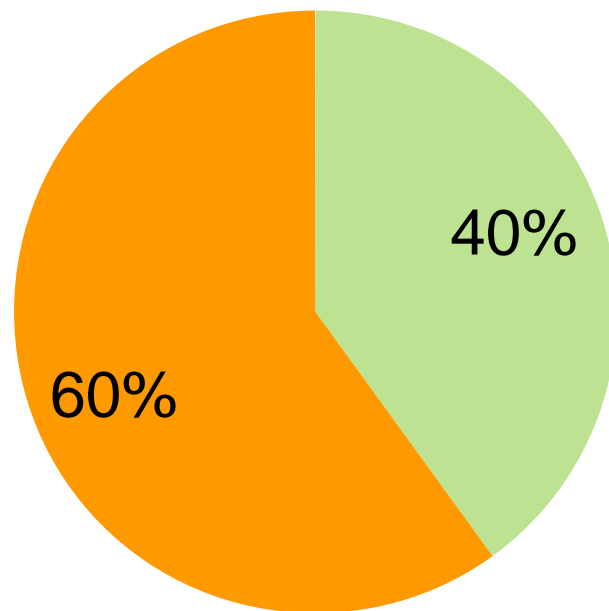
With the heavy  
emphasis on  
performance, these  
MSLS cannot be  
validated  
psychometrically

4

Locally  
Developed  
Courses

# A Picture of Assessment

## North Carolina Teacher Workforce



- EOCs, EOGs, or VoCATS
- Currently Non-Tested Grades and Subjects



# Four “Buckets” of Assessment

1

EOCs, EOGs  
and VoCATS

2

Category One  
of MSLS

With appropriate  
resources and time,  
these MSLS can be  
validated  
psychometrically

3

Category Two of  
MSLS

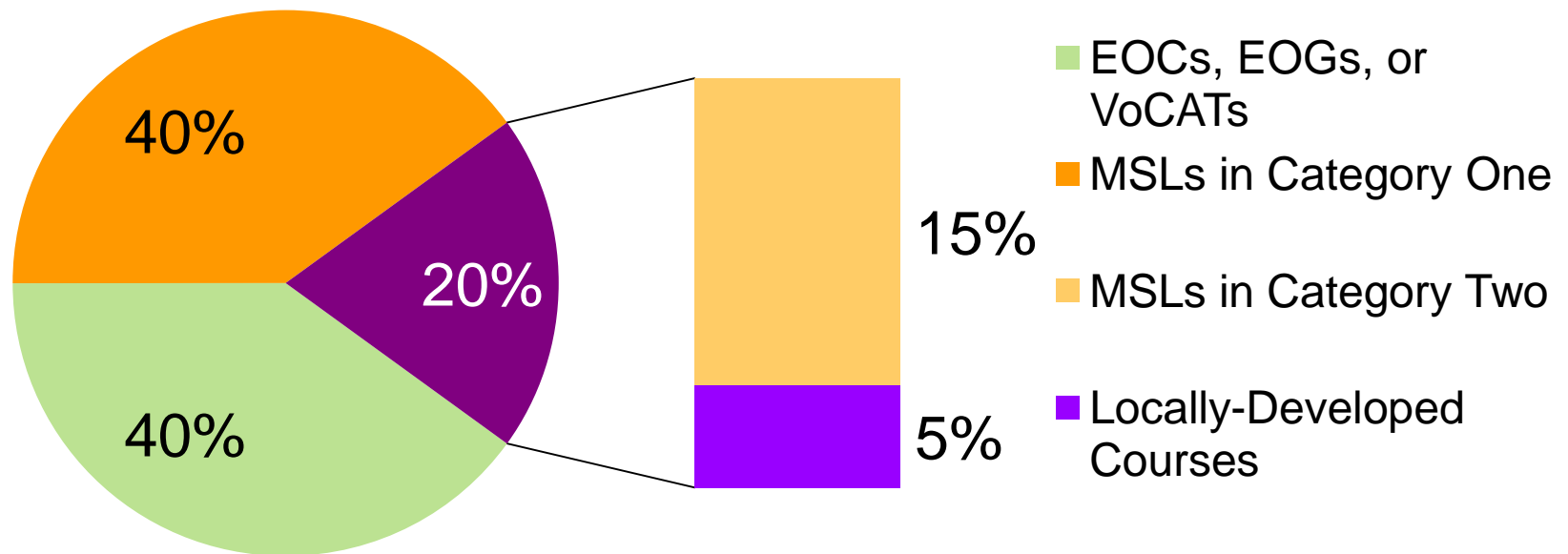
With their heavy  
emphasis on  
performance, these  
MSLS cannot be  
validated  
psychometrically

4

Locally  
Developed  
Courses

# A Picture of Assessment

## North Carolina Teacher Workforce



# MSL Design Process

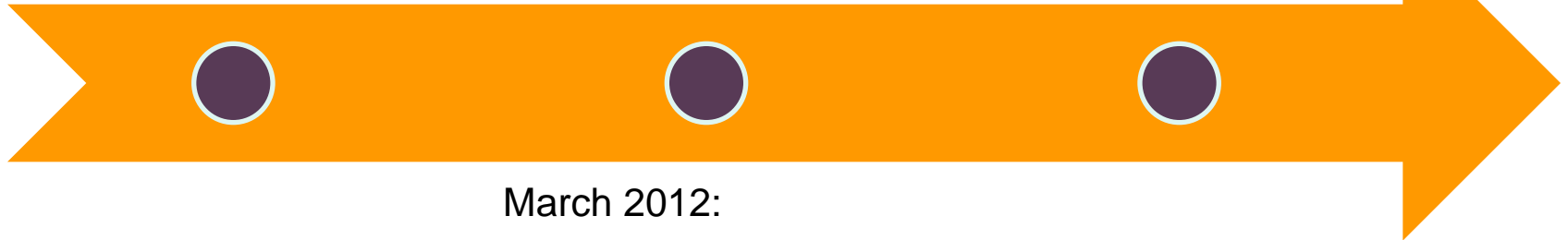
# Three Phase Process

October 2011:

Teachers design item specifications for all currently non-tested grades and subjects

April – May 2012:

Teachers create rubrics and guidance for administering and scoring Measures of Student Learning



March 2012:

Teachers vet open-source items and items generated by external vendor(s)

# Three Phase Process

## Phase I (October 2011)

Teachers design item specifications for all currently non-tested grades and subjects.

## Phase III

(April – May 2012)

Teachers create rubrics and guidance for administering and scoring Measures of Student Learning.

## Phase II (March 2012)

Teachers evaluate open-source items and items generated by external vendor(s).

# Phase I: Create Groups

Developed list of all non-tested courses and grades in the  
Common Core State Standards and NC Essential Standards

Grouped courses and grades together into like-content groups

Groups range from Extended Content Standards to Chemistry to  
Elementary Theatre Arts to Social Studies Electives

# Phase I: Select Members

Designed an online application system for interested educators to apply to join one of the design groups

Advertised for the design groups

Selected educators from over 1,500 applications

Responded to LEA concerns and notified teachers of final selection decisions

# Phase I: Ensure Representation

101 Local Education Agencies

10 Charter Schools

Department of Juvenile Justice and Delinquency  
Prevention

University of North Carolina

North Carolina Virtual Public School

# Phase I: Provide Training

1. The Measures of Student Learning design process and how the Measures fit into the State's educator effectiveness work (DPI Leadership and RttT Project Management)
2. Assessment design, including potential item types, reliability, and validity (Assessment Design and Development)
3. Overview of the Common Core and NC Essential Standards (Curriculum & Instruction)



# Phase I: Gather Feedback

Through three feedback protocols, teachers provided answers to the following critical question:

What does meaningful  
assessment in your  
content area look like?

# Preliminary Feedback:

To date, staff have consolidated input from K-2 English Language Arts and the Arts:

Subject Area	Teacher Input
K-2 ELA	<ul style="list-style-type: none"><li>• Student growth needs to be measured throughout the year, as opposed to a “one-time” opportunity</li><li>• Student growth is measured through student writing, speaking, listening, and reading</li><li>• Kindergarten ELA will need to have some type of pre-assessment</li><li>• In grades 1 – 2, the previous year’s score could count as the starting point for measuring growth</li><li>• Reading 3-D is too time-intensive and used for formative purposes</li></ul>



# Preliminary Feedback:

Subject Area	Teacher Input
The Arts	<ul style="list-style-type: none"><li>•Growth in the arts can be measured through performance-based tasks and selected-response items</li><li>•Performance-based tasks can assess the “creation” elements of the Essential Standards</li><li>•Selected-response items can assess the “culture and history” elements of the Essential Standards</li><li>•Performance-based tasks need to be broad enough to cover curricular options. For example, a scale can be sung, played on a string instrument, or played on a wind instrument</li></ul>



# Teacher Thoughts and Concerns:

Appreciation, pride, and even joy that their content areas are now being valued and that the State is recognizing that they impact the learning of their students

Worry about finding a valid way to measure student learning in an art or PE class that meets once a month and one that meets every day

Value the input of teachers into the process

Doubt over the sustainability of these Measures of Student Learning after Race to the Top ends

Worry about a “test-heavy” environment for students, especially young children

# Challenges Moving Forward



# Resources :

Originally, North Carolina budgeted \$120,000 for this effort

With a Race to the Top amendment, North Carolina has \$1.6 million to pay for a vendor to generate items and \$720,000 to involve more teachers more frequently

\$1.6 million may not be enough funding to generate the robust item set described by the teacher design groups

# Bias and Reliability in Grading:

Given the variety of items desired by the teacher design groups, educators will need to play a role in assessing student performance on Measures of Student Learning according to standardized rubrics

Their involvement introduces the risk for bias, even if teachers are not grading their own students' work

Teachers are concerned about when they will have time to grade the performance tasks in the Measures of Student Learning



# Equality Among Content Areas :

Balance of instructional time between content areas

Ratings generated by a mathematical model and those selected by principals based on data

# Next Steps for 2011 - 2012

# Next Steps: 2011-12

- Late December 2011/Early January 2012: Engage vendor to generate items
- March 2012: Design groups begin to return to vet items created by vendor

# Next Steps : 2011-12

- **April/May 2012:** Design groups return to design rubrics, administration instructions, and guidance on grading the Measures of Student Learning
- **Late Spring/Early Fall:** Pilot-test Measures of Student Learning