



# Welcome to CNG

## The Life-Cycle of Exam Items from Conception to Retirement

- Program will begin at 9:30 a.m.
- Feel free to sit anywhere you prefer
- Enjoy light refreshments and networking until then

# Thank you to our Full Sponsors!

PROMETRIC



**interrel**  
ADVANCING INTERESTS



Castle

SCANTRON.  
SMART STARTS HERE

CASTLE IS A SCANTRON ASSESSMENT SOLUTIONS BUSINESS

 **Pearson**  
VUE

**btlSurpass**   
*Powering Assessment*

# Thank you to our Supporting Sponsors!



Dental Assisting National Board, Inc.  
*Measuring Dental Assisting Excellence®*



Today's Sponsor



Thank you!



Marya Ryan, PhD, PMP, CAE  
Owner  
OneEarth Consulting, Inc.

# Today's Speakers

## *Phase 1: Job Analysis*

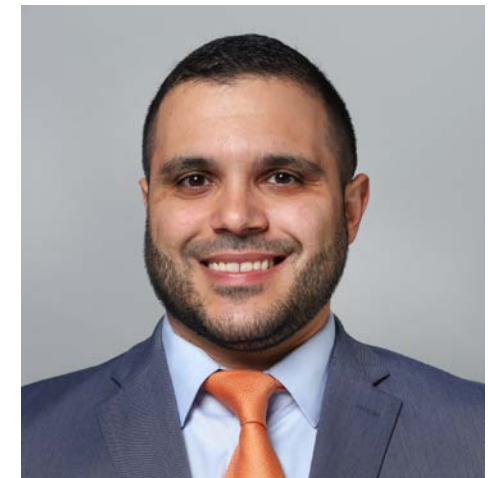
Carissa Homme

Manager of Test Development and Certification  
Competency & Credentialing Institute



Frank Perna

Sr. Manager, Test Development  
Association for Supply Chain Management



# Today's Speakers

## *Phase 2: Item Development*

Jolene Riordan  
Assessment Specialist  
American Dental Association



Kay Manger-Hague  
Director – Credentialing Operations  
Academy of Nutrition and Dietetics



# Today's Speakers

## *Phase 3: Item Analysis*

Tara McNaughton, Ph.D.

Director of Psychometrics and Assessment  
American Osteopathic Association





# Job Analysis at CCI

May 15, 2019

CNG Chicago

Carissa Homme, PhD





Questions welcome

Simplified examples

# Outline

- About CCI
- Job Analysis Overview
- Key Players
- The Job Analysis Process

# About CCI



# Job Analysis Overview

- Also known as practice analysis, role and function study, or role delineation study
- Goal: identify the tasks and knowledge that are important to the competent performance of a professional
- Required by accreditation bodies
- How: creation, administration, and analysis of an online survey *or* panel based study

# Job Analysis Overview

- Occurs every 5 years for CNOR and every 3 years for CSSM
  - Timing reflects pace of change in the field
  - Also done at the start of a new certification build
- SMEs: 10-15
- Commitment: 2 in-person meetings, completion of online surveys, and 3 remote calls across 5-6 months

# Key Players

## Volunteers

- Volunteers
  - Subject Matter Experts (SMEs)
  - Must hold the credential
  - Receive points towards their next recertification
  - Group should reflect diversity in the profession



# Key Players

## Volunteers

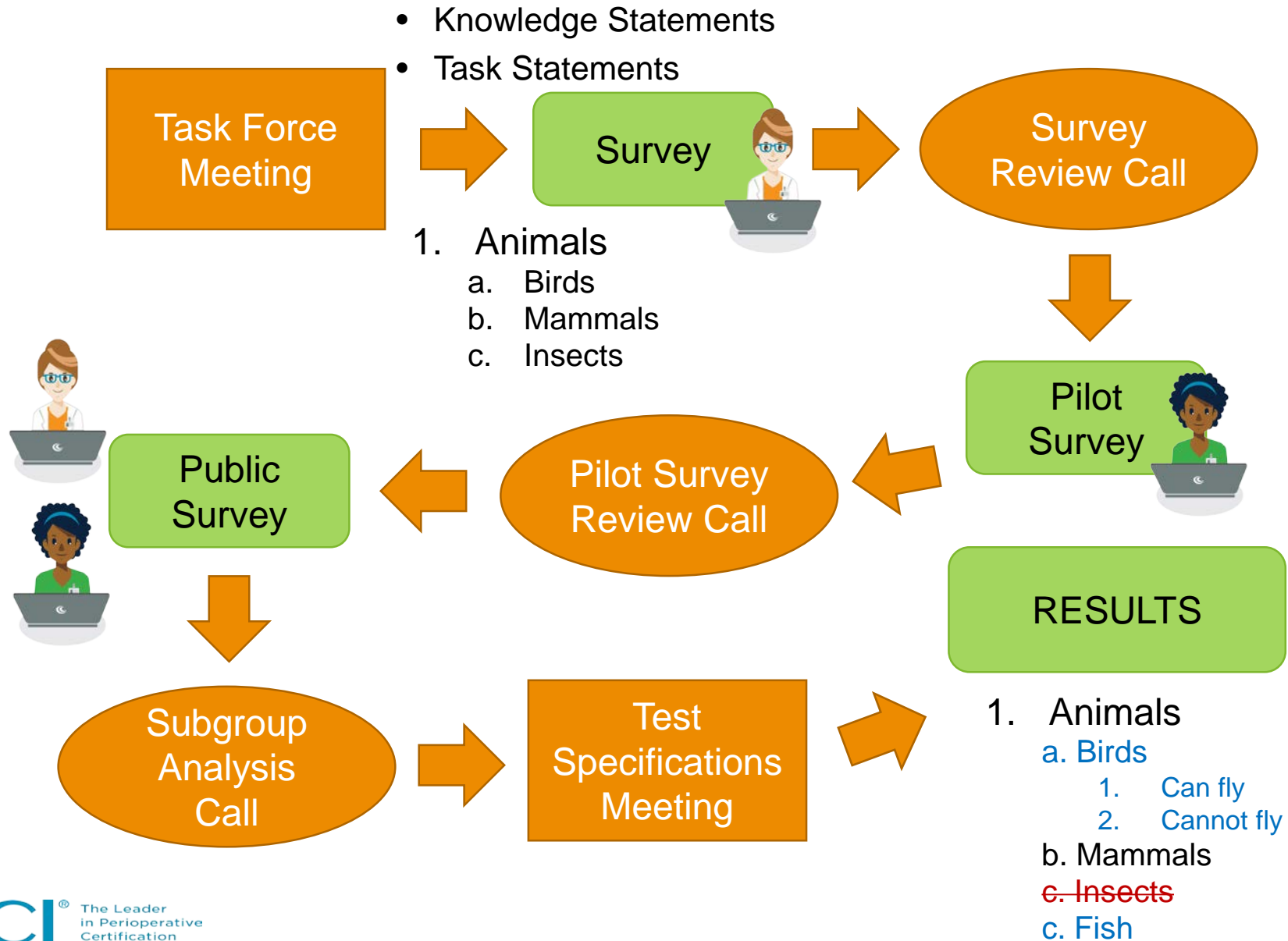
- Volunteers
  - Subject Matter Experts (SMEs)
  - Must hold the credential
  - Receive points towards their next recertification
  - Group should reflect diversity in the profession



## CCI and Prometric

- CCI Staff
  - Facilitate meetings
  - Manage test development process
  - Recruit and coordinate volunteers
- Prometric Partners
  - Present the material at meetings
  - Perform psychometric analyses
  - Publish and administer exams
- CCI Certification Council
  - Determine cut scores
  - Approve test development volunteer selections
  - Certification program oversight

# Job Analysis Process





# Crosswalk

- Follows the Job Analysis
- Compares old blueprint to the new
- Goal: place items into the new blueprint

## 1. Animals

- a. Birds
- b. Mammals
- c. Insects

## 1. Animals

- a. Birds
  - 1. Can fly
  - 2. Cannot fly
- b. Mammals
- ~~c. Insects~~
- c. Fish



- Typically Remote
- Requires a minimum of 2 SMEs
- Similar activity: Pool Review / Item Bank Review

# Summary

- About CCI
- Job Analysis Overview
- Key Players
- The Job Analysis Process

Questions?



Thank You



# Item Life Cycle- Job Task Analysis Process

Frank Perna

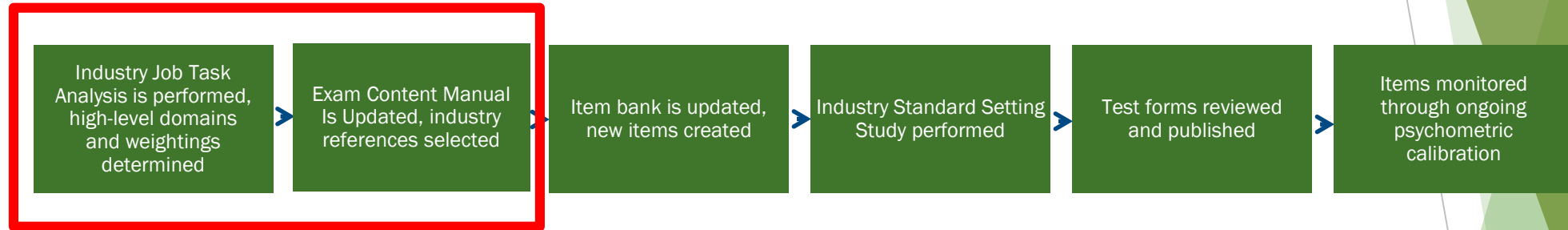
May 15, 2019

# Item Life Cycle- Job Task Analysis Process

Frank Perna

May 15, 2019

# Overview of The ASCM Exam Development Process

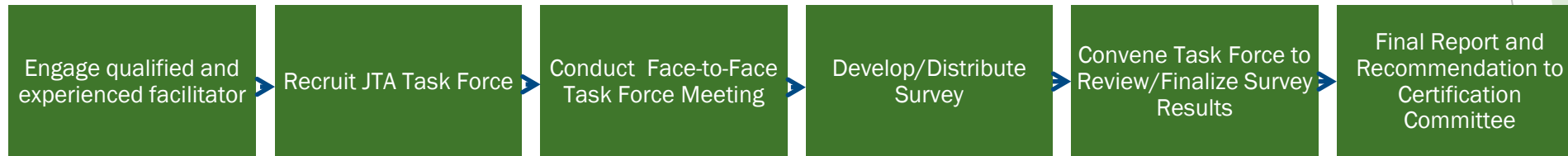


# The JTA- What and Why?

- Worldwide (if applicable) research survey to validate the knowledge and skills required of individuals working in content-specific job roles for today's work environment
- Accrediting bodies require a JTA-like process to take place when a new Certification is being created, and periodically thereafter to validate and update content
- The results of this project influences certification content domains and weightings
- Frequency of process can vary depending on the nature of the industry for which the JTA is performed



# ASCM's JTA process



# Selecting a Facilitator

- ASCM contracts a 3<sup>rd</sup> party Psychometrician to facilitate JTAs
  - Can be an internal employee, if applicable
  - Many Test Development vendors offer JTA services
- Tips for selecting a facilitator
  - Consider budget
  - Review resumes holistically
  - Rely on your network
  - Check References
  - Ensure mutual goals are established (contract and beyond)
  - Ability to keep a group engaged is AS IMPORTANT (if not more) than knowledge of the process!!

# Recruiting a Diverse JTA Task Force

We utilize ASCMs database, social media, and our network of volunteers and members to recruit a diverse, holistic group

## ▪ Considerations used for selecting participants

- Industry
- Stage in career
- Country/Region
- Certified/Not-Certified
- ASCM Member/non-member

## ▪ Tips for recruiting a Task Force

- Again, consider budget
- Don't invite the "same old volunteers"
- Set expectations/time commitment early on
- Be as diverse as your industry and your budget allow
- Conduct interviews- having a lot of experience doesn't necessarily make you a good volunteer
- Look for team players

# Face-to-Face Task Force Meeting

ASCM holds JTA Task Force Meetings at HQ in Chicago, bringing the group of SMEs together for two consecutive days

During the meeting, the following typically takes place:

- Brief overview of ASCM done by staff (especially important with non-members in the room)
- Orientation to the JTA process by facilitator
- High-level discussion of what the job roles for individuals certified will entail
- Deep dive into the knowledge, skills, and tasks required to complete roles
- Once knowledge, skills, and tasks are established, they are grouped together into overarching content domains
- Based on the groupings, the group will then discuss and determine the percentages by which each domain represents the body of knowledge as a whole
- Finally, the group discusses the overview of what the survey might look like, including scales, demographic info requested, etc.

# Survey Development and Distribution

Along with our facilitator and Certification Staff, our Research Department will compile the survey and distribute it using a strategic approach with input from our Marketing Team.

## Survey Considerations:

- It should go to as diverse a group as possible- Not just members or non-members engaged with the organization
- It should go to people who are qualified to pursue the Certification- Consider job titles, eligibility requirements, etc.
- Incentives for participation- What can be offered to encourage responses? (free periodical, CEU points, raffles, etc.)
- Determine response goals- How many minimum responses, what margin for error is acceptable, etc.
- Survey length- Keep it as short as possible while still covering the necessary topics. This can be challenging!
- Work with Marketing to determine the best distribution plan, including email cadence, other media channels , etc.

# JTA Survey Sample

**B. DESIGN THE SUPPLY CHAIN** - Please assess the frequency and importance of each task relative to this topic.

	How often do you (or does someone you supervise) perform each task?					How important is each task?			
	Daily/Weekly	Monthly	Quarterly	Annually	Never	Very important	Important	Somewhat important	Not important
Identify customer and business requirements	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Identify future state	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Perform gap analysis between current and future state	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**B. DESIGN THE SUPPLY CHAIN** - How important are the following knowledge areas to this topic?

	Very Important	Important	Somewhat important	Not important	I don't know
Advanced business acumen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Product lifecycle stages	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Market research	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

# Results Review and Final Report

After the survey is closed and results are compiled and organized, the JTA Task Force comes together again in a virtual meeting to review the results and finalize the content domains along with the knowledge, skills and tasks that fit within each one.

## During this review:

- The Task Force will consider the importance and frequency with which respondents rated each input
- Inputs with extremely low ratings can be considered for removal
- Based on the finalization of knowledge, skills, and tasks within each high-level domain, the group will reconsider their weightings for each domain and finalize the percentages

Once this meeting is complete, The facilitator will create an Executive Report of the findings of the JTA, which is provided to the Exam Committee to aid them in updating the Exam Content Outline

At this point, the JTA is complete!

Questions?



Thank You

# The Life Cycle of an Item

Jolene Riordan, MLS, Assessment Specialist  
Department of Testing Services, American Dental  
Association

# ADA Department of Testing Services

# ADA Department of Testing Services (DTS)

- The ADA Department of Testing Services (DTS) is a shared service that provides professional psychometric and test development services for high stakes testing programs in dentistry and healthcare.
- Test Development Team Purpose:  
*Develop, organize, and maintain high quality examination content that is relevant, current, accurate, free from bias, and in accordance with test specifications, thereby facilitating the valid, reliable, and fair evaluation of candidate knowledge and skills in alignment with the purpose of each examination.*

# ADA Department of Testing Services (DTS)

Joint Commission on National Dental Examinations (JCNDE)	Council on Dental Education and Licensure (CDEL)
<ul style="list-style-type: none"> <li>• National Board Dental Examinations (NBDE)                             <ul style="list-style-type: none"> <li>• Part I</li> <li>• Part II</li> <li>• INBDE (under development)</li> </ul> </li> <li>• NBDHE</li> </ul>	<ul style="list-style-type: none"> <li>• Dental Admission Test (DAT)</li> <li>• Advanced Dental Admission Test (ADAT)</li> </ul>
Outside clients	ADA Board of Trustees
<ul style="list-style-type: none"> <li>• Optometry Admission Test (OAT)</li> <li>• Canadian Dental Admission Test (CDAT)</li> <li>• Additional clients</li> </ul>	<ul style="list-style-type: none"> <li>• Dental Licensure Objective Structured Clinical Examination (DLOSCE)</li> </ul>

# ADA Department of Testing Services (DTS)

Examination Content, Scoring, Analysis, New Development		
<p><b>Test Development</b></p> <p>Conducts Test Construction Team (TCT) meetings for seven examination programs (80+ meetings annually)</p>	<p><b>Research &amp; Development - Psychometrics</b></p> <p>Oversees analysis and scoring of examinations (45,000+), professional investigations, and technical publications in support of examination programs</p>	<p><b>New Psychometric Development &amp; Innovations</b></p> <p>Provides psychometric support in the development of new testing programs.</p>
DTS Operations		
<p><b>Test Administration</b></p> <p>Oversees application processing and test vendor administrations (40,000+ examinations)</p> <p>Responds to phone calls, live chats, emails, faxes (nearly 70,000 annually)</p> <p>Resolves testing day problems</p>	<p><b>Test Security and Fraud Prevention</b></p> <p>Test security policies, procedures, and candidate appeals; risk assessment</p>	<p><b>Communications</b></p> <p>Provides communications for stakeholders and communities of interest</p>
<p><b>Client Services/Special Projects</b></p> <p>Project management and services to outside clients</p>	<p><b>Volunteer and Meeting Coordination</b></p> <p>Oversees volunteer activities and meeting logistics for TCT and governance meetings.</p>	

# VOLUNTEER SELECTION

# Recruitment of Subject Matter Experts

- Traditionally SMEs were appointed to a set Test Construction Team (TCT) for a term of five years
- We are moving to a pool system for volunteer SME selection beginning in 2020
  - Volunteers complete an online application to be entered into the test constructor pool
  - Applications can be submitted throughout the year
  - The governance committee approves applicants for inclusion in the pool (*in the past volunteers were selected only for open slots in TCTs*)
  - Provides flexibility in recruitment for examination development activities



# Volunteer Pool Selection

- Volunteers are asked to select the examination programs they are interested in contributing to
- Volunteer recruitment begins about six months prior to each meeting
- Volunteers are selected from the pool for exam development opportunities based on:
  - Area of expertise
  - Representation of the field (educators, clinicians, private practitioners)
  - Geographic location
  - Availability (most meetings are scheduled for 2.5 days)
  - Past participation (to avoid recruiting the same dentist too often!)

# Benefits of Volunteering as a SME

- Awarded CE credit for each meeting
- Opportunity to network and collaborate with peers around the country
- Contribute to the continued excellence of the profession by sharing knowledge and expertise
- Enjoy a stay in Chicago within a block or two of the Chicago Water Tower
- Receive instruction on writing quality items that they can use in their own classrooms

# Item Development

- Item writing assignments are sent to the SMEs several months prior to the meeting start date
  - SMEs are instructed to upload items to a secure site provided by the ADA
  - Depending on the needs of the examination program, item writing activities may also be scheduled on-site
- Item review and examination review meetings are held on-site

# Expectations of Volunteer SMEs

- Ability to be flexible and adaptable
- Ability to work with different groups
- Ability to attend for the duration of scheduled meetings
- Intellectual curiosity and openness
- Strong writing and communication skills
- Expressed interest and enthusiasm for the role
- Receptive to feedback from staff

# ITEM DEVELOPMENT

# Role of DTS Assessment Specialists

## The Assessment Specialists

- ❖ Facilitate item development
- ❖ Review, edit, and revise items according to grammar and style guidelines
- ❖ Maintain item disposition (status throughout an item's life cycle, cognitive level, etc.)
- ❖ Ensure the blueprint is met for each examination form
- ❖ Work with R&D to insure exam specs are met
- ❖ Deliver the examination to the vendor

# Role of SMEs

## The Subject Matter Experts (SMEs)

- ❖ Include item writers, item reviewers, examination content reviewers
- ❖ Write the items
- ❖ Provide images and case materials
- ❖ Ensure the currency of the items
- ❖ Ensure the accuracy of the items
- ❖ Ensure the items are fair
- ❖ Ensure the language is clear
- ❖ Verify item content is important to test on this examination
- ❖ Verify items are at an appropriate level of difficulty for this examination
- ❖ Ensure items fit to the examination specifications
- ❖ Review items before they are placed on an examination
- ❖ Conduct sensitivity checks

# Item Development

**The SMEs are the experts.**

We rely on them to write items that are

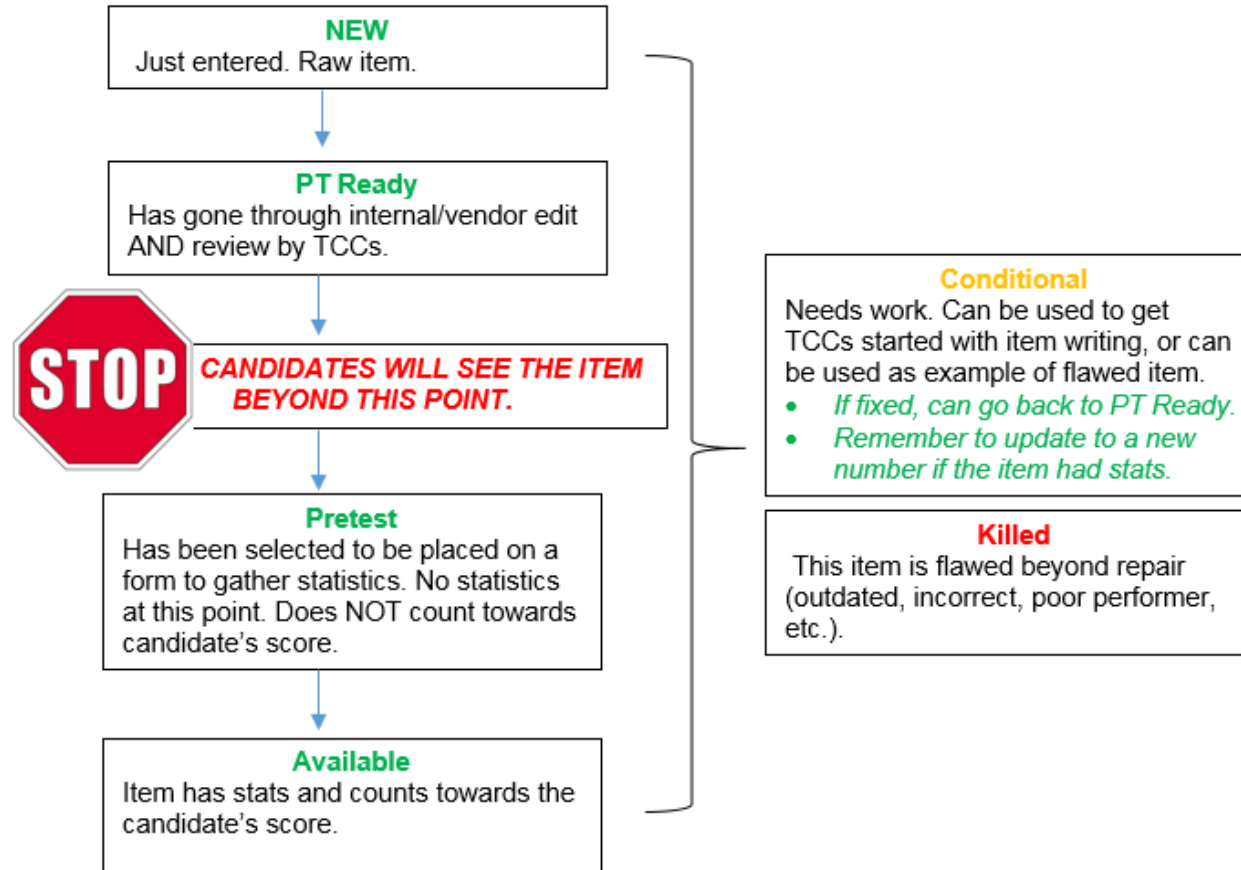
- ✓ Current
- ✓ Valid
- ✓ Important
- ✓ Relevant
- ✓ Fair



# How is Exam Content Determined?

- Practice Analysis
  - To provide this necessary content related validity evidence, it is necessary to periodically review and define the domain of knowledge, skills, and abilities that are relevant to the safe practice of dentistry.
  - The Practice Analysis informs the examination blueprint.

# Item Status Throughout the Item's Lifetime



# Item Review

- Items are reviewed at several stages of item development:
  - Newly written items are edited to style by staff as they are entered in the database
    - *The content and core concept must be maintained. If a suggested edit may affect the content, but is unacceptable as written, the item is sent back to the item writer or reviewed by the appropriate TCT.*
  - A second editorial review is conducted by an external vendor/editor
    - *These edits are themselves reviewed internally prior to being entered in the database*
  - Newly written items are reviewed by the TCT prior to being given a status of PT Ready (ready to be pretested)
  - Items are reviewed and given a status of Pretest when selected for pretesting on an examination form by the TCT

## Item Review (*continued*)

- Examination forms are reviewed in their entirety by a TCT prior to submission to the vendor
- Items exhibiting differential item functioning are identified by the R&D team for review by the appropriate TCT
- Items are also reviewed as needed
  - When guidelines change
  - If changes are made to blueprint after a practice analysis
  - To ensure currency

# ORIENTATION TO ITEM DEVELOPMENT:

*ITEM WRITING AND ITEM REVIEW*

# Item Development Orientation

- New SMEs are provided with a webinar orientation prior to their first meeting to introduce them to the item development process and to answer any questions they may have.
  - Required to sign the DTS Contributor Agreement Form prior to attending
    - Includes:
      - Confidentiality Agreement
      - Copyright Agreement
      - Conflict of Interest Statement
  - Instructed to attend the webinar in a private location where they won't be disturbed and others will not be able to listen in

# Item Development Orientation

- SMEs are introduced to internal ADA style guidelines
  - Internal editorial style
  - Accepted item formats
- SMEs are instructed to avoid common item writing flaws
  - Negatively stated stems
  - Unfocused stems
  - True/False
  - “All of the above....” / “None of the above...”
  - Overlap among distractors
  - Cued items (word or part of a word echoed in stem and key)
  - Items testing more than one concept

# Item Development Orientation

- Common item writing flaws (continued)
  - Stereotypes, assumptions, slang, inflammatory language
  - Excess verbiage
    - Window dressing
    - Red herrings
    - Teaching
  - Branching items
  - Double jeopardy
  - Opposite distractors
  - Made-up distractors
  - Esoteric knowledge, trivia, “factoids”



# SME Item Review

As SMEs review items, they are instructed to:

- ✓ Read each item and all response options completely.
- ✓ Verify the key is correct.
- ✓ Verify the content is important to this examination.
- ✓ Ensure each item tests only one single concept.
- ✓ Ensure the item and any associated images are original (exam content is copyrighted by the ADA and must not appear anywhere else).
- ✓ Verify the content is important to test on the examination.
  - ❖ *No esoteric knowledge, factoids, or trivia.*

# Item Fairness

SMEs are asked also to verify the ***fairness*** of the items by asking themselves:

- ✓ Is the item stated clearly?
- ✓ Is the item at an appropriate level of difficulty for this examination?
- ✓ Is the item free from unnecessary or “tricky” language, regionalisms, inflammatory content, and “slang”?
- ✓ Is the item free from stereotypes or assumptions?
- ✓ Does it fit the examination specifications?

**QUESTIONS?**

**Thank You**

# Life Cycle of an Item: Item Development

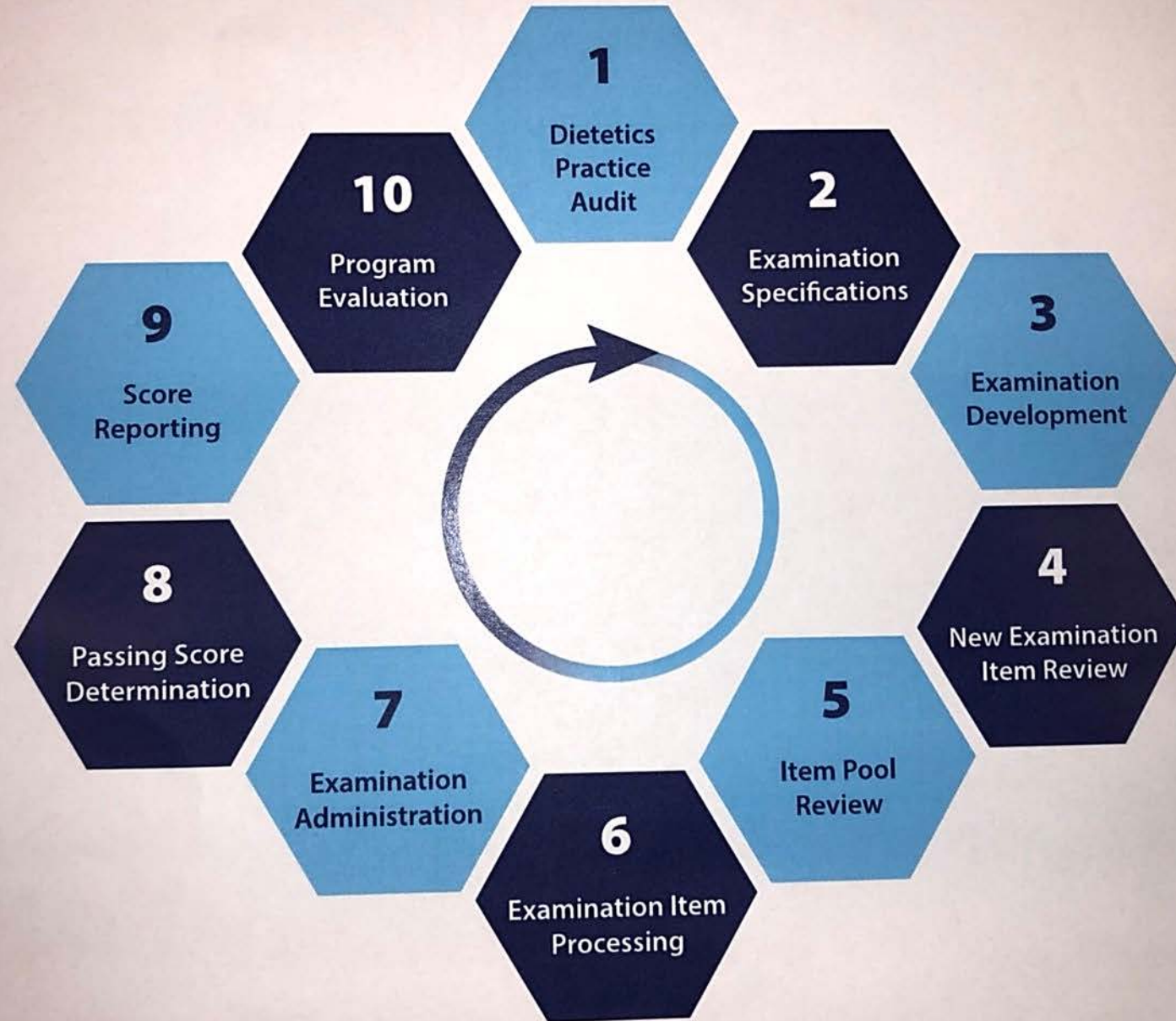
---

May 15, 2019

CNG Chicago

Kay Manger-Hague, RD

# Examination Development & Testing Program



## PLANNING AND DEVELOPMENT

### 1. Dietetics Practice Audit

A Dietetics Practice Audit describes the knowledge and skills necessary to perform competently at an identified level of practice. It serves as the basis for test specification development. Using a role delineation study is among the most desirable methods for specification development, because it assists in ensuring that the certification test is job related, representative of practice, and geared to the appropriate responsibility level. CDR conducts a Dietetics Practice Audit every five years.

### 2. Test Specifications

Test specifications are a detailed blueprint for constructing a test. They include a description of the content to be tested, the proportion of the test to be devoted to the different areas of content within domains, and the characteristics of acceptable test items. Test specifications derived from the practice audit verified by actual practice provide evidence in support of test content validity and establish its defensibility and credibility.

### 3. Test Item Development

New items (questions) are prepared by individuals selected from diverse practice areas and population subgroups who are trained in the specifics of good test construction principles.

Criteria applied to writing test items are:

- (1) *relevance and criticality to entry-level practice;*
- (2) *accuracy, currency and clarity;*
- (3) *regional and institutional differences; and*
- (4) *conformity with test specifications.*

### 4. New Test Item Review

Test items are reviewed by professional test editors to eliminate technical flaws, ambiguities, and potential bias. All test items are reviewed by experienced item writers to verify appropriate classification and conformance with item writing criteria.

Editorially and technically sound items are pretested as unscored items on a test. This ensures that the scoreable portion of the test includes good performing items.

### 5. Test Item Pool Review

Annually, experienced test reviewers appointed by the Commission on Dietetic Registration review items for content accuracy, currency, relevance to entry-level practice and one best answer.

### 6. Test Item Pretesting

Only test items that have survived content, measurement, and editorial review are suitable for inclusion in the computer-based testing item pool.

## ADMINISTRATION AND SCORING

### 7. Test Administration

Registration eligibility requirements are established by the Commission on Dietetic Registration. The Commission contracts with ACT, Inc. to administer the tests on computer at over 230 test sites. Special testing needs, such as those for religious observance and physical handicaps, are accommodated under standardized secure conditions.

### 8. Passing Score Determination

A passing score study is periodically conducted by experienced dietetics professionals representing diverse practice

areas and population subgroups. The use of systematic judgment of content experts in these studies establishes the minimum level of acceptable professional performance expected on a certification test. CDR uses a criterion-referenced approach for determining the passing score. This criterion-referenced passing score becomes the basis for equating future examinations, thus ensuring that all test versions are of equal difficulty level.

## REPORTS AND EVALUATION

### 9. Score Reporting

A score report announces the examinee's performance on the certification test. The report includes a total scaled score as well as two subscores; Food and Nutrition and Foodservice Systems Management.

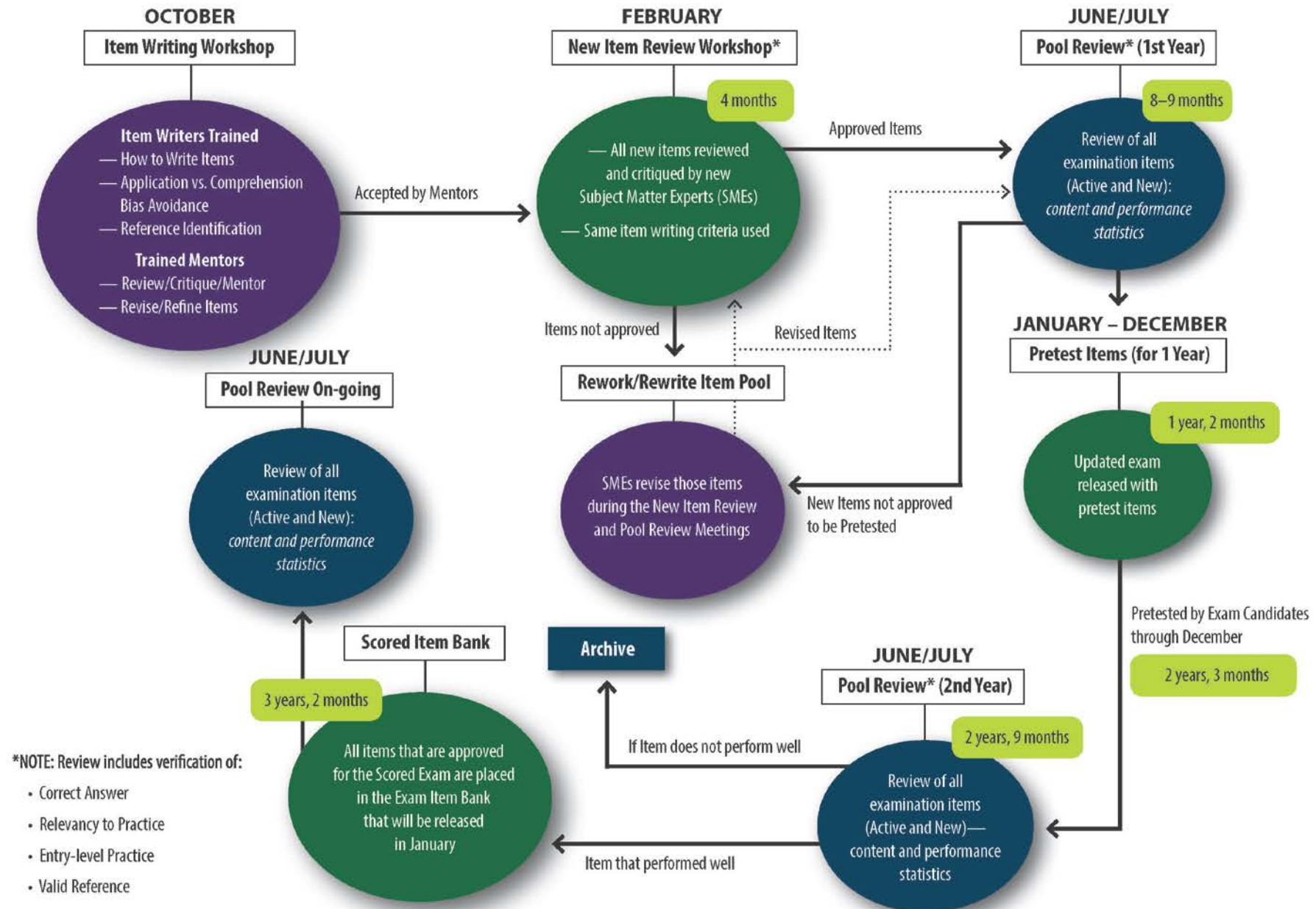
Twice annually, the Commission provides dietetic education programs with both institutional test summary reports of examinee group performance and individual scores by name when authorized by the examinee.

### 10. Program Evaluation

A comprehensive technical report, which includes statistical data, is provided by the test contractor to the Commission on Dietetic Registration. This report and feedback from dietetic programs and examinees are used by the CDR Examination Panel in evaluating the certification testing program. The Standards for Educational and Psychological Testing and established Examination Panel procedures are used as the basis for the evaluation process.



# Item Development Activities and Timeline





The Commission on Dietetic Registration is soliciting the services of registered dietitians and dietetic technicians, registered to assist in the annual item writing workshop for the Registration Examination for Dietitians and/or Dietetic Technicians. During this workshop, volunteers are trained by Pearson VUE consultants to construct items for the examinations. Participants include RDN, RD, NDTR or DTR employers, practitioners and educators who possess knowledge of entry-level practice.

CDR ID Number \_\_\_\_\_ Credential (RD or DTR) \_\_\_\_\_

Name \_\_\_\_\_  
(Last) (First) (Middle)

Address \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

Daytime Phone \_\_\_\_\_ Cell Phone \_\_\_\_\_

E-Mail Address \_\_\_\_\_

**PLEASE CHECK AS APPROPRIATE:**

I would like to write items for the Registration Examination for:

Dietitians                       Dietetic Technicians                       Both

Area of Expertise:

Clinical Nutrition                       Community Nutrition                       Foodservice Systems Management

Other \_\_\_\_\_

I am currently working in dietetics as an: (Please select one from the list below.)

- Educator of dietetics students
- Educator of dietetic technician students
- Employer of entry-level RDs
- Employer of entry-level DTRs
- Entry-Level Registered Dietitian (practicing 3 or less years)
- Entry-Level Dietetic Technician (practicing 3 or less years)

Please return this form to:  
**KAY MANGER-HAGUE, DIRECTOR – CREDENTIALING OPERATIONS**  
Email: [khague@eatright.org](mailto:khague@eatright.org) Fax: 312/899-4772

# Questions?

---

May 15, 2019

CNG Chicago

Kay Manger-Hague, RD



A M E R I C A N  
O S T E O P A T H I C  
A S S O C I A T I O N

# ONGOING TEST ITEM ANALYSIS

Tara McNaughton

*May 15, 2019*

# The Item Lifecycle

---



# Develop

- 
- Determine content area
- Identify teaching point
- Develop item

# Use

## Initial Use

- Content review
  - Exam Construction
- Baseline statistics/Key validation
- Operational statistics

## Subsequent Use

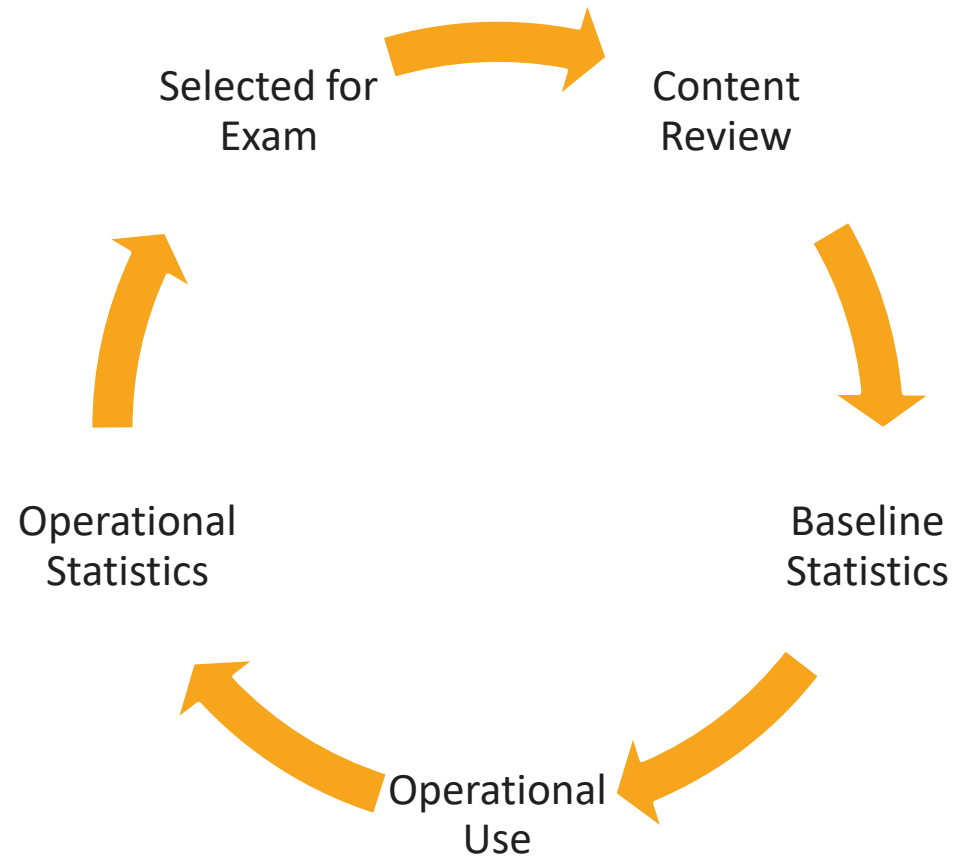
- Content review
  - Exam Construction
- Baseline statistics/Key validation
- Operational statistics
- Measurement consistency

# Retire

- 
- Content relevancy
- Statistical issues
- Indications of item compromise
- Duplicate item
- Better items tap same teaching point
- Other?

# Ongoing Item Use

---





# Use

## Content Reviews

- SMEs groups
  - Diverse individuals
    - Gender
    - Age
    - Region
    - Specialization
    - Setting
  - Diverse groups
    - Item writers
    - Item reviewers/editors
    - Exam construction
    - Exam review
    - Key validation

## Statistics for Initial Use

- Key validation **or**
- Pilot/Pre-Testing/Screen
  - Baseline statistics
    - P-Value
    - Pt. Biserial/Pt. Measure
- Initial use/Prior to CAT use
  - Baseline statistics
  - Operational statistics
    - Item difficulty
    - Fit

## Statistics for Subsequent Use

- Baseline statistics
  - Each use
  - Periodically for CAT exams
- Operational statistics
  - Displacement
    - Measurement consistency

# Baseline Statistics

Traditional Item Statistics	Rasch Equivalent	Indicates
<ul style="list-style-type: none"><li>■ P-Value<ul style="list-style-type: none"><li>— Proportion of examinees getting the item correct</li></ul></li></ul>	<ul style="list-style-type: none"><li>■ Pt-Measure<ul style="list-style-type: none"><li>— Correlation between score response and ability measures</li></ul></li></ul>	<ul style="list-style-type: none"><li>■ How well targeted the item was to examinee ability levels</li><li>■ Item's association with overall test results</li></ul>
<ul style="list-style-type: none"><li>■ Pt-Biserial<ul style="list-style-type: none"><li>— Correlation between score response (right/wrong) and total test score</li></ul></li></ul>		

# Baseline Statistics – Targets

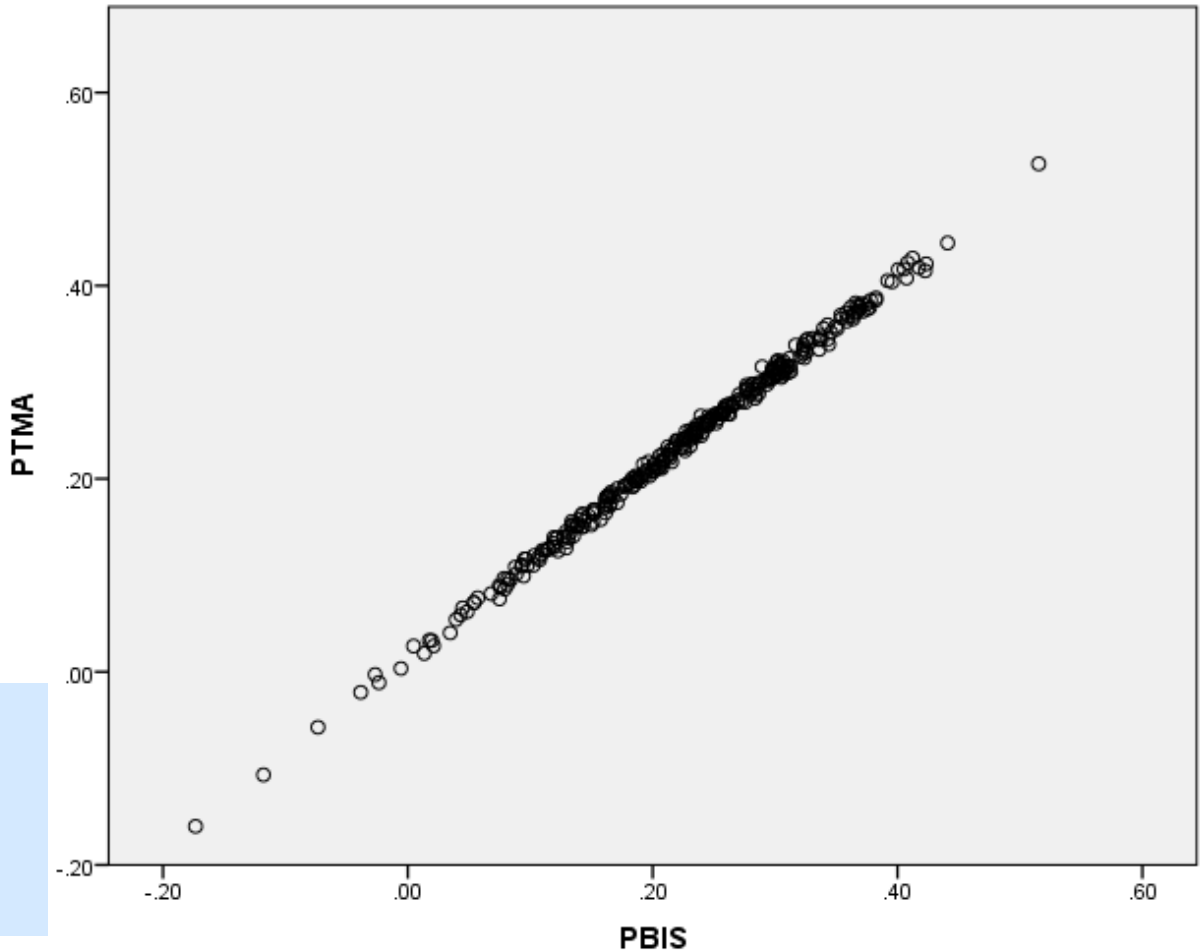
Statistics	Looks Like	Certification Targets
<ul style="list-style-type: none"><li>■ P-Value<ul style="list-style-type: none"><li>– Proportion of examinees getting the item correct</li></ul></li></ul>	<ul style="list-style-type: none"><li>■ 0.00 to 1.00<ul style="list-style-type: none"><li>– No one to everyone</li></ul></li></ul>	<ul style="list-style-type: none"><li>■ .40-.80</li><li>■ .30-.90</li></ul>
<ul style="list-style-type: none"><li>■ Pt-Biserial/Pt. Measure<ul style="list-style-type: none"><li>– Item total correlations</li></ul></li></ul>	<ul style="list-style-type: none"><li>■ -1.00 to 1.00<ul style="list-style-type: none"><li>– Extreme negative to extreme positive association</li></ul></li></ul>	<ul style="list-style-type: none"><li>■ .20 or higher</li><li>■ .10 – .19</li><li>■ .01 – .09</li></ul>

# Pt Biserial (PBIS) vs Pt Measure (PTMA): One example

- 320 Item exam
- ~400 examinees
- Difference between Pt. Measure and Pt. Biserial

	Min	Max	Mean	Std. Deviation
Difference (PTMA-PBIS)	-0.01	0.03	0.011	0.006

For key and distractor Pt. Measures add:  
DISFILE=yourfile.txt  
To your control file



# Operational Statistics

---

## Initial Use

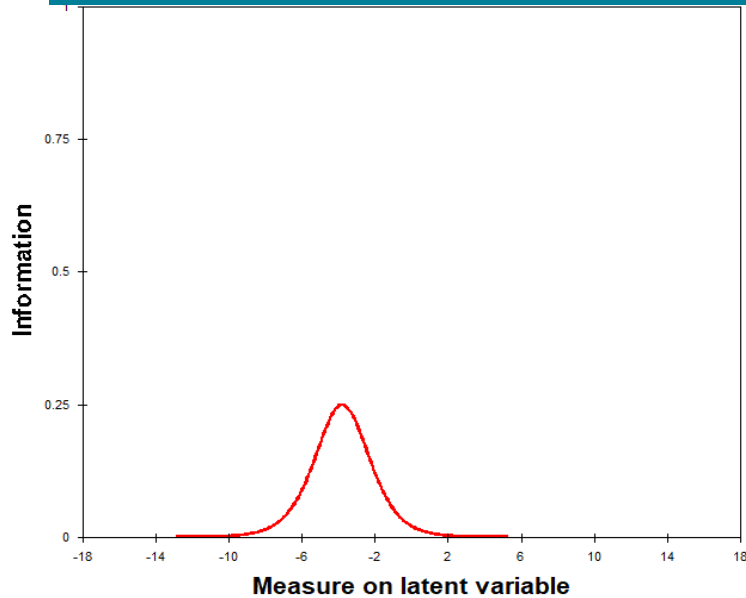
- Item difficulty
  - Values approx. -3.00 to 3.00 logits
- Fit
  - Values possible 0.00 –  $\infty$
  - Expected value is 1.00
  - Acceptable range 0.50 – 1.50
- Items with good initial statistics
  - Linking or anchor items
    - Common item equating

## Subsequent Use

- Item difficulty
- Fit
- Displacement
  - Consistency of difficulty level with anchor value

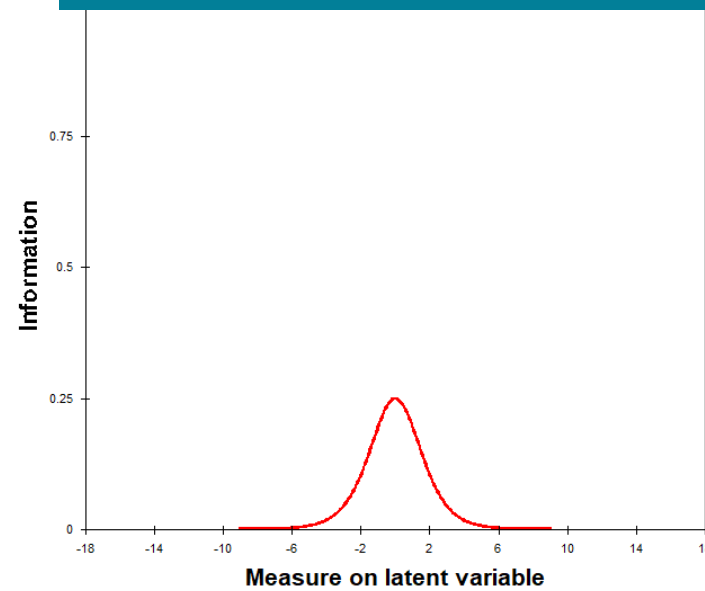
# P-Values and Item Information

P-Value= .98, PTMA .16



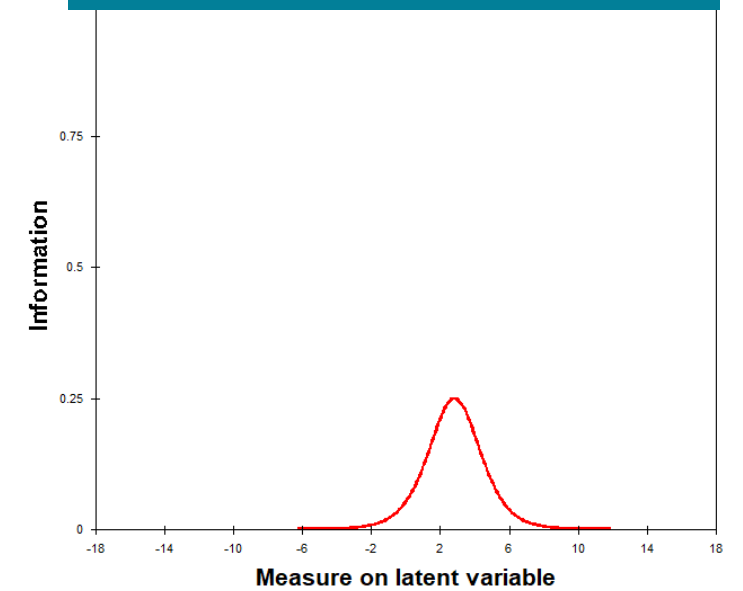
Max information at = -3.80

P-Value= .62, PTMA .14



Max information at = 0.00

P-Value= .10, PTMA = .10



Max information at = 2.82

Pass Point = 0.10

# Displacement

---

- Estimates difference
    - Anchored value
    - Item difficulty if not anchored
  - Indicates the appropriateness of the anchor value
- Displacement Criteria to recalibrate
    - 0.30/Robust z
    - 0.50
    - 0.60
    - 1.00
    - $t = \text{displacement} / \text{standard error}$ 
      - $df = \text{number of responses}$
  - Over 1.00 may be consider item parameter drift (IPD)

# Item Parameter Drift (IPD)

---

## Causes

- Educational/Curriculum changes
- Advances in the field
  - Technological
  - Established methods
  - Published guidelines
  - Other
- Cultural changes
- Item harvesting
  - Cheating

## Influenced by

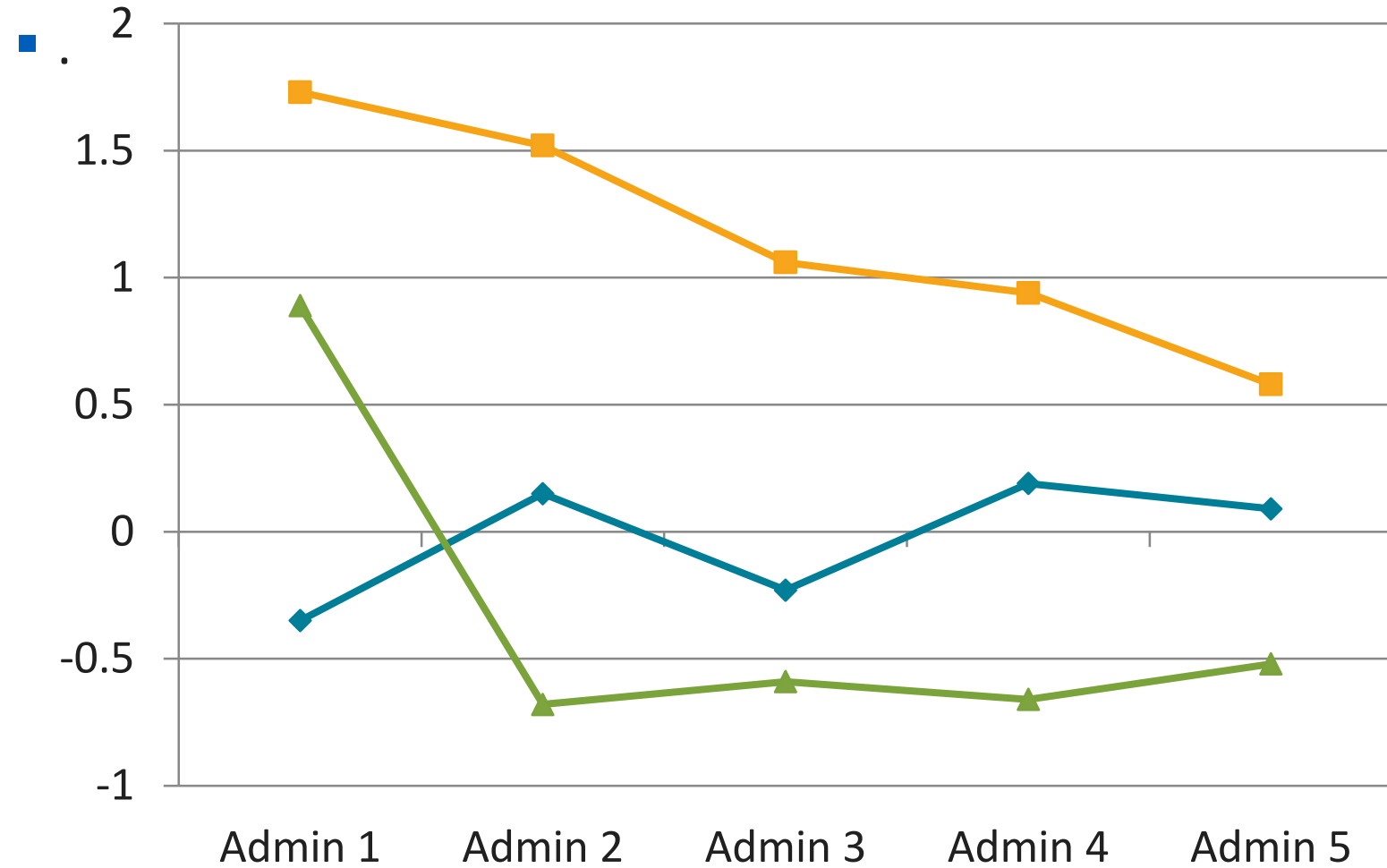
- Number of items in the bank
- Item exposure rates
- Number of Examinees
  - Too few
  - Too many

## Effects

- Ability estimates
- Pass rates
- Classification accuracy (Pass/fail)
- Exam validity
- Equating results



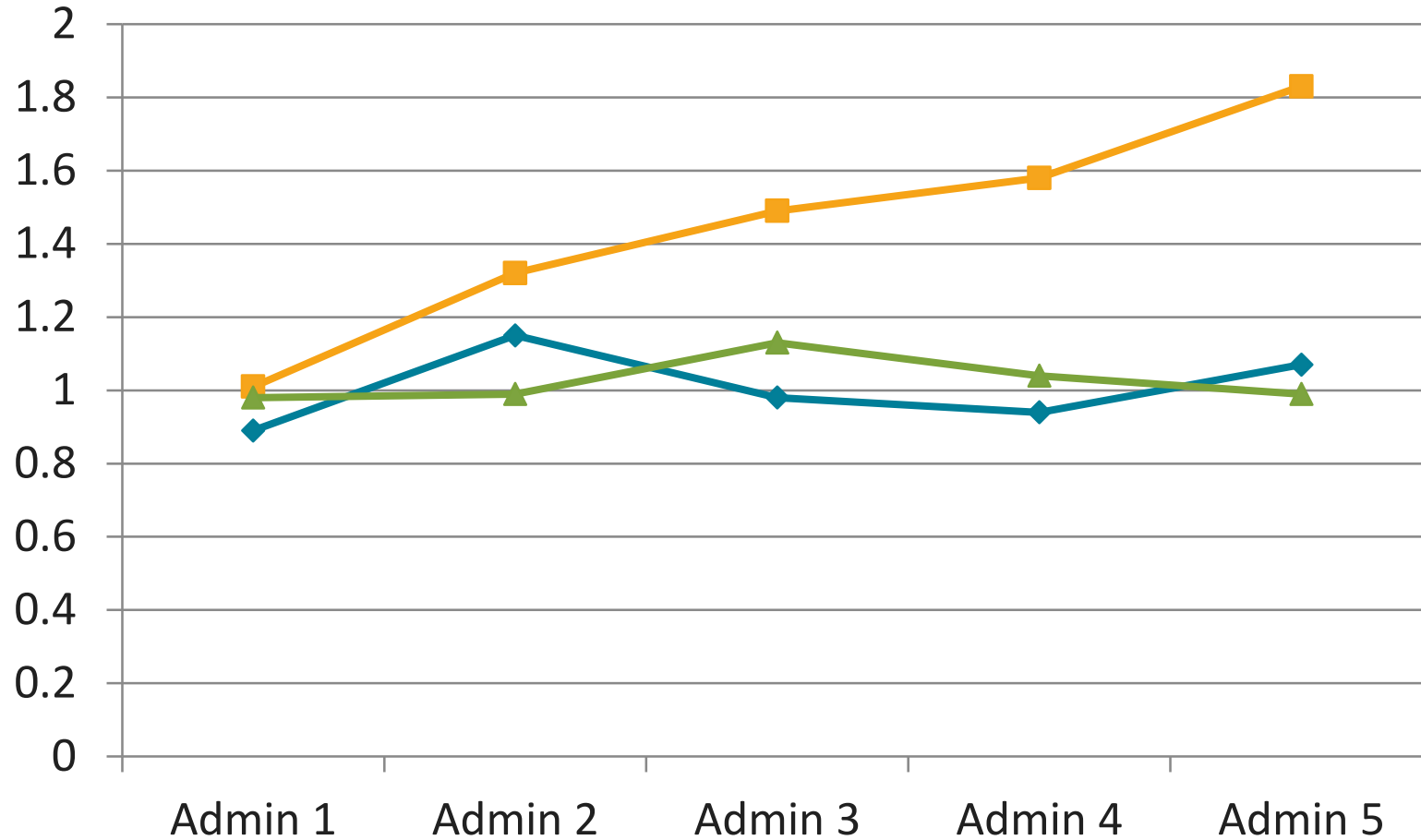
# IPD: Graphing Difficulty



◆ Item 1  
■ Item 2  
▲ Item 3

	Item 1	Item 2	Item 3
Min	-0.35	0.58	-0.68
Max	0.19	1.73	0.89
Absolute Drift	0.54	1.15	1.57

# IPD: Graphing Fit



- ◆ Item 1
- Item 2
- ▲ Item 3

	Item 1	Item 2	Item 3
Min	0.86	1.01	0.97
Max	1.18	1.83	1.15
Difference	0.32	0.82	0.18

# IPD: How Much Concern?

---

## IRT Models are Robust

- Rasch model robust with up to 25% of items demonstrating IPD
- 1PL, 2PL, and 3PL models also demonstrate robustness against IPD

## Concern Remains

- Larger proportions of test items with IPD
- Larger item drift increases impact of IPD
  - Even with fewer IPD items
  - 1.00 logit differences and greater
- Drift in anchor items
  - Item harvesting
  - Biases equated scores

## What To Do

- Monitor for IPD
- Review anchoring protocols
- Monitor difficulty and fit frequently in CAT exams
- IPD in candidate subgroups
  - More difficult to identify
  - Investigate anchor items with larger fit

# Investigate and Evade IPD

---

## Monitor for IPD

- Differential Item Functioning (DIF) with time as grouping variable
  - Mantel-Haenszel
- ANCOVA
- ICC/TCCs
- Lord's Chi-Square
- Cumulative sum (CUSUM method)
- Lagrange multiplier (LM)

## Larger Item Bank

- Improves classification accuracy and measurement precision when drift is present

## Keep Refreshing Content

- New material protects item bank
  - Lower exposure rates
  - Fewer re-used items lessens many impacts of IPD
    - Ability estimates
    - Classification accuracy (Pass/fail)
    - Pass rates
    - Exam validity

# IPD and Anchor Items

---

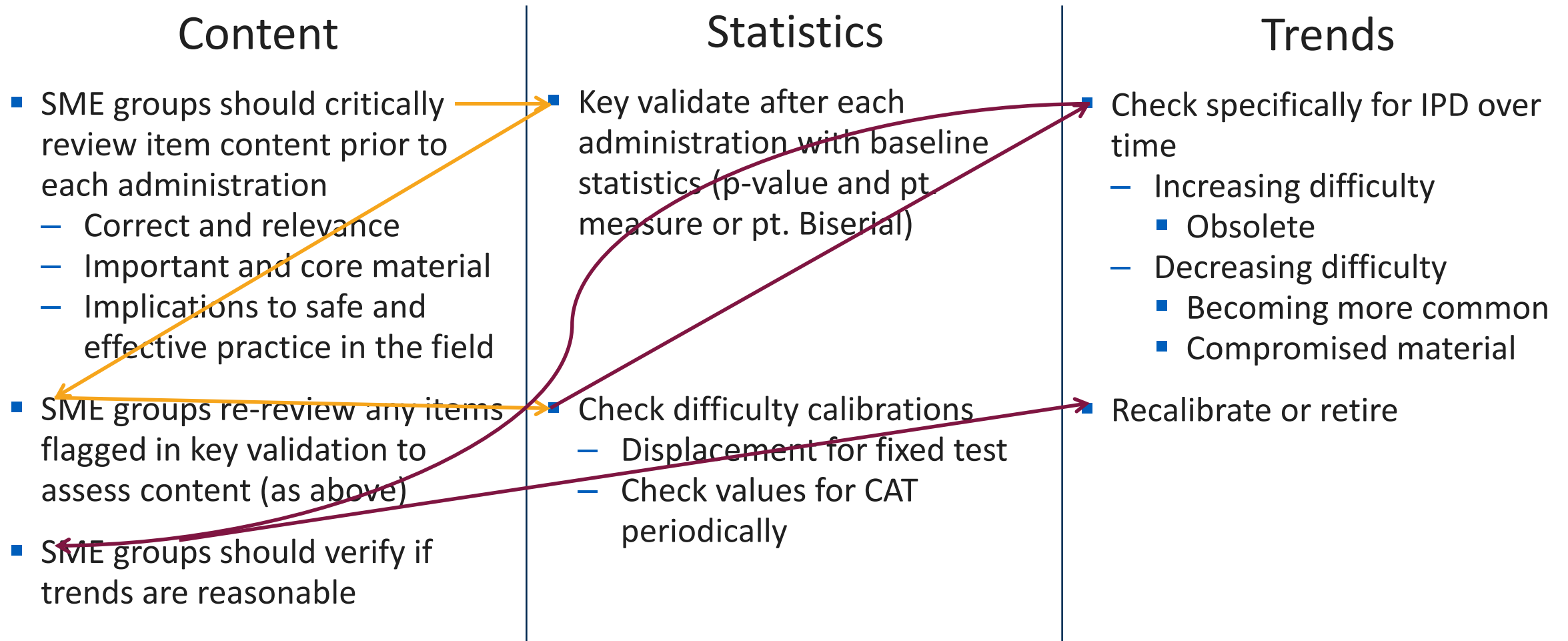
- Symmetrical
  - 50% of drift items easier
  - 50% of drift items harder
- Drift does not impact displacement statistics for non-drift anchor items
- Asymmetrical
  - 70% of drift items easier
  - 30% of drift items harder
- Displacement values for non-drifting items are pulled counter to the majority
  - i.e. non-drift items displace as more difficult
- Unilateral
  - 100% of drift items easier
- Displacement values for non-drift items may be flagged and unanchored as drifting more difficult

# Content and Statistics

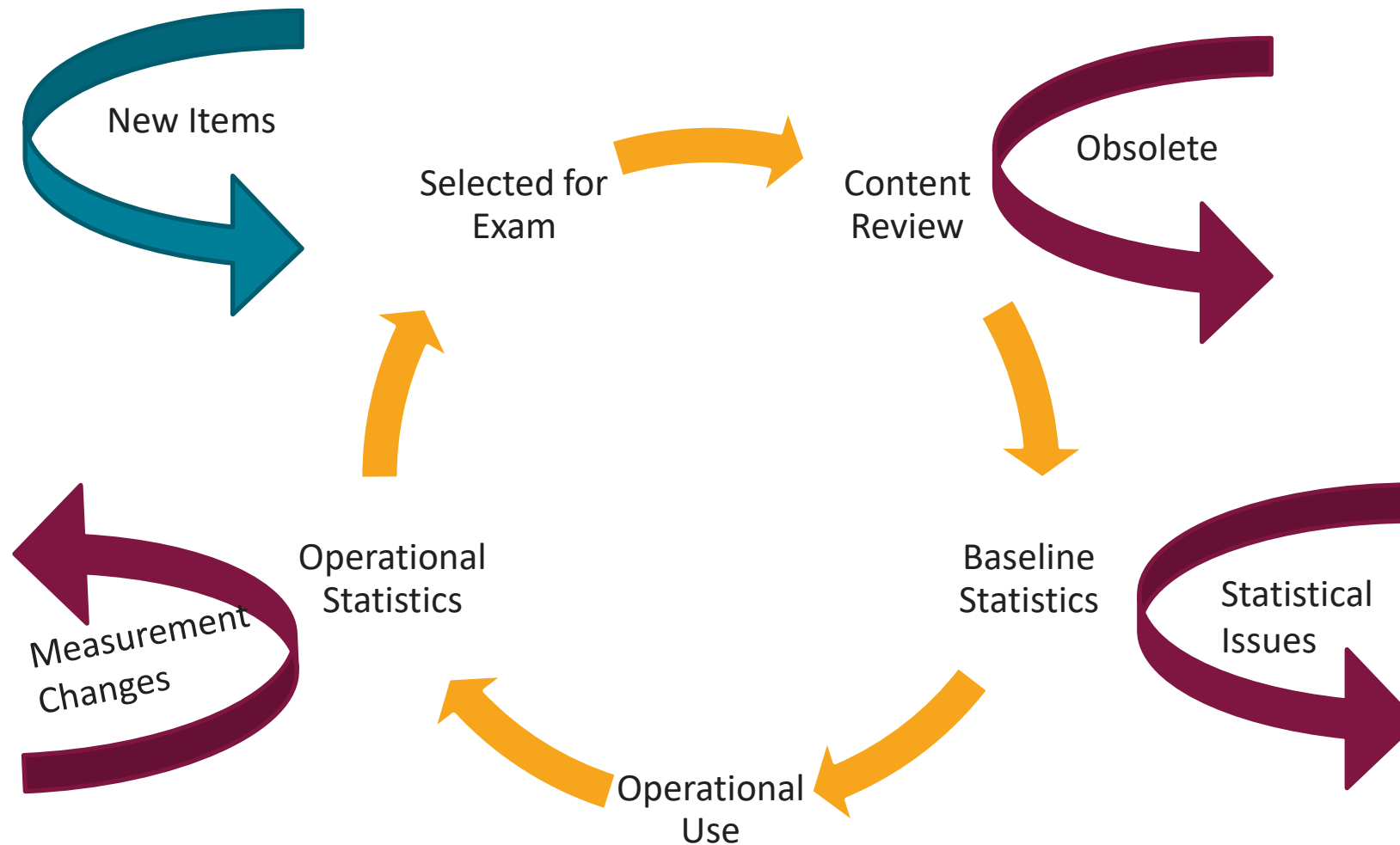
- .



# Content and Statistics



# Ongoing Item Use





# Summary

---

- Use appropriate statistics to flag items for SME groups at different points in the item lifecycle
- Refresh content with new material
- Key validate or pre-test items,
  - Baseline stats each time item is used
  - Periodically for CAT exams
- Check difficulty and fit to evaluate changes
  - Policy on when to allow items to recalibrate due to displacement
    - Fit statistics may also be useful
- Check trends across time to help identify when items should be flagged for SMEs to evaluate changes

# Discussion

---



# Job Openings?

# Next CNG Meeting Details

- **When:** Wednesday, August 7, 2019
- **Where:** TBD
- **Topic:** We welcome your suggestions! We will also review past surveys and calls for topics.



Thank you for attending!