Getting It Right: Implementing Innovative Item Types

It takes time and attention to design and implement innovative items.

A six-step design process (Parshall & Harms, 2008) that improves the quality of innovative item types within any exam program is presented in CER, Fall 2008.

The tasks of Subject Matter Experts (SMEs) should be approached differently when developing innovative item types instead of traditional item types.

In this session we will:

- Define “Innovative Items” and show a few examples
- Briefly describe the 6 steps of the model
- Focus on Subject Matter Expert (SME) involvement in Steps 1, 4, 6
6-Step Design & Development Process


Innovative Item Types

Think of a "continuum" of innovation –
- from discrete items with modest changes
- to elaborate, performance-based simulations

Almost any departure from Multiple Choice questions has been termed "innovative".

Any item type that is new to your exam program could be considered innovative for you.

Multiple Response

1) You want your students to categorize information in a form that can be sorted by topics. Which of the following programs should they use? (Select all that apply.)
   - database
   - email
   - graphics
   - presentation
   - spreadsheet
   - web browser
   - word processing
Development of Innovative Items

Certain item development processes
- improve the quality of the innovative items, and
- provide a framework for including the item types as part of normal test development

Step 1 – Analyze Construct Needs

- **Purpose**: Identify deficiencies in construct coverage and potential value of new item types
- **Activity**: Test developers and SMEs consider: program goals, areas currently being measured well and those that are not, identify missing pieces in construct coverage
- **Result**: List of a few specific areas of content or cognition to target

Step 2 – Select Innovations

- **Purpose**: Consider specific innovations and their potential match to construct needs
- **Activity**:  
  - Consider aspects of innovation  
    - Assessment structure, response action, media inclusion, interactivity, complexity, fidelity, and scoring methods  
  - Evaluate match of innovations to construct needs
- **Result**: Selection of a few specific innovative item types for prototype development
Steps 1 and 2

- Steps 1 and 2 are the “construct” phases – and Subject Matter Expert (SME) involvement may vary by agency (some agencies using SMEs in 1, some agencies using SMEs in 2)

Focus the SMEs on Step 1?

- What construct are you measuring?
- Could it be measured without innovation?
  - Use the innovation only when the answer is “no”
  - Save your efforts for the constructs that can only be measured with the innovation

Example -

- Which test topic requires innovation in measurement?
- VOTE BY RAISING YOUR HAND:
  (X = your field)
  A. History of X
  B. Theory of X
  C. Ethics of X
C – Ethics of X

- MCQs should do a good job of measuring the history of X field
- MCQs should do a good job of measuring the theory of X field
- Many agencies are DISSATISFIED with their ability to test ETHICS via MCQs, so that topic should be the focus for our innovation.

Sound?

- Some fields require sound as an innovation.
- Examples are heart sounds, ECG monitors, engine sounds
- Other examples?

Ordered Response/Sequence

- Many professions have a sequence of steps that must be followed
- In the traditional MCQ formats they ask what should be done FIRST
- An innovation might be required to test knowledge of the full set of steps
Shopaholic Example

• Let’s say we’re working on an ordered response item about Shoe Shopping for a Shopaholic examination

What is the correct sequence of events in Shoe Shopping?

• A. See new shoes, discover price, remove old shoes, try on shoes, attempt to walk in shoes, purchase shoes
• B. Remove old shoes, see new shoes, try on shoes, attempt to walk in shoes, purchase shoes, discover price
• C. Discover price, remove old shoes, try on shoes, attempt to walk in shoes, purchase shoes, see new shoes

What are you measuring?

• If candidates won’t attempt long-list items, you’re not measuring the sequence of Shoe Shopping (guessing, skipping)
• Is there only one best sequence for certified shopaholics, or might some do A and some do B?
• Is each step in the sequence clear? (See new shoes = “I spy a new pair of shoes” or “I see a new pair of shoes on my feet”
What construct are you measuring?

- Shopaholics must know the sequence events in shoe shopping
- Can it be measured without innovation?
- If not, Step 2 is Select the Innovation

Innovative Item: Less cumbersome

- Build a list in the correct sequence for shoe shopping:
  - Try on shoes
  - See new shoes
  - Purchase shoes
  - Remove old shoes
  - Attempt to walk in shoes
  - Discover price

Step 3 – Design Initial Item Prototypes

- **Purpose:** Develop and review draft prototypes
- **Activity:**
  - Develop Prototypes
  - Internal Review of Prototypes
  - Determine Revisions (and prototypes to eliminate)
- **Result:** Set of prototypes that have been internally evaluated and revised
Example 1 – Initial Prototype

Example 1 – First Revision

Example: 2 Stages of Prototype Revision
Example: 2 Stages of Prototype Revision

Example: Final Item Design

Agency Specific?
- Some agencies might use SMEs in step 3
- Some agencies might use staff or vendor in step 3
- We’ll jump to 4
**Step 4 – Iteratively Refine Item Types Designs**

- **Purpose:** Refine item type designs
- **Activity:** Iteratively complete the following steps:
  - **Step 4a:** Develop item writing materials and sample items
  - **Step 4b:** Conduct usability testing
  - **Step 4c:** Evaluate and revise item type designs
- **Result:** Approval of specific item types for pilot testing

---

**Design Stage**

**Step 4**
Iteratively Refine Item Type Designs

- Develop Item Writing Materials & Sample Items
- Evaluate & Revise Item Type Designs
- Conduct Usability Testing

---

**Step 4a: Develop Item Writing Materials and Sample Items**

- For each item type, iteratively develop:
  - Draft item *templates* for each item type
  - Initial item writing *guidelines*
  - Item writer *training*
- Write a few sample items, ensuring representation of all item features
What Are Templates?

- Templates are a structured means of collecting/storing item data.
- Possible aspects to template design:
  1. Database fields
  2. Screen layout elements
  3. More specificity in either of the above (i.e., sub-template)

Sample Template

**Word Processing Task Description:**
For this task you will use a word processor to

**Task Progress:**

**Simulated word processing program**
Changes to standard menu options:
Text to appear initially on screen:

**Navigation Options:**
Exit, Help, Status indicator

Sample Prototype

*Description*: For this task you will use a word processor to create and edit a document.
**What Can Templates Do?**

- Templates are used in the item writing process. They provide:
  - **Structure** – they help guide and constrain item writing
  - **Efficiency** – they can save time & money in programming efforts and media development
  - **Security** – they can help replace an “exposed” item with a close variant

**Templates & Prototypes**

- A prototype is an instantiation of a template;
  - a completed item or task that is built using the framework of the template
- Templates and Prototypes can be used
  - in the planning/design stage
  - in the item writing stage

**Item Writing Guidelines**

- Standard item writing guidelines have a proven effectiveness in contributing to high quality traditional items.
- Most innovative items include new elements not addressed by existing guidelines, e.g.:
  - Multimedia
  - New response actions
Examples of New Guidelines

• An audio item should use no more than two people speaking.
• The video in an item stimulus should be no more than 30 seconds in length.
• Each hot spot item should include a single correct area and four incorrect areas.
• No more than half the responses in a multiple-response item should be keys (correct).

Item Writer Training

• Item writer training prepares SMEs to write high quality items.
• Even skilled item writers will need to be given examples and training in each new item type.
• Examples of useful materials include
  – Visual prototypes
  – Samples of completed templates

“How To” versus “Do Not…”

• Traditional item formats are so well-established that many agencies begin the training with a long list of Do Not....
• Innovative items will require a structured How To list
• The time/expense of developing innovative items suggests a high level of structure
Change in Level of Structure

- Writers of traditional item types are typically given training then set free to create
- That model is unlikely to succeed for innovative item types

Level of Structure

- If you lose the structure, you are likely to lose the measurement
  - What are you measuring?
  - Hot spot example

Click on the place to begin the incision
Start at the End - Scoring

• Indicate the place to begin the incision
  – What are you measuring?
  – What is the correct response?
  – How close must examinees be to the correct response?
  – Where does “incorrect” begin?
  – Will the scoring be right/wrong?
  – Is partial credit appropriate?

Start at the End

• Knowing what you’re measuring is tied to knowing how to score it
• Be sure all stakeholders are in agreement with scoring plan EARLY in the process

Enforcing structure

• For traditional MCQs, many agencies have documented guidelines for:
  – Focused stems
  – Positive context
  – One best answer
  – Clarity of statements
  – Blueprint classification
But allow and enter into the bank items with:

- Unfocused stems
- Negative context
- More than one correct answer
- Unclear phrases or terminology
- Missing or incorrect blueprint classification

Why Maintain Sloppy Item Banks?

- How many of you refuse to bank items exhibiting the flaws just mentioned?
- You should be saying "Return to sender!"

It’s worse with Innovative Formats

- The problem is greatly exacerbated with innovative item formats
- Don’t accept items that don’t meet your specifications
  - Don’t store them
  - Don’t spend SME time reviewing them
  - Impose STRUCTURE upon the process
Getting it Right, Now!

You should have a high level of structure for SMEs, plus keeping a focus on measurement. Those who went before you had to do the trial & error, on their dimes!

Step 4b: Conduct Usability Testing

- Conduct multiple rounds of usability testing
- Usability refers to the degree to which a software program:
  - Is easy to learn
  - Contains the necessary functions to let the user complete the tasks

Usability and Validity

- Validity concerns related to usability:
  - Make sure we’re not testing examinee ability to use the testing software
  - Examinee time should be spent on item content, not interface interpretation
- Innovative items make these implications especially important
**Effects of Usability Testing**

- Yields software that:
  - Can be learned more quickly
  - Is more efficient to use
  - Results in fewer user errors

- Financial benefits
  - Reduces programming
  - Prioritizes development
  - Reduces maintenance and support

**Usability Method – Think Aloud**

- “Tell me what you’re thinking as you go…”
- 5-6 participants
- 3-4 rounds

**Isolate extraneous input**

- In traditional MCQs item writers sometimes include extraneous information that is detected and removed during the SME review
- Extraneous information in innovative formats might be detected during usability reviews
- Does anyone remember the patient with jaundice from a pilot video clip for a nursing examination?
Step 4c: Evaluate & Revise Item Type Designs

- Evaluate each item type
- Include all stakeholder groups
- Determine necessary changes
  - Based on usability testing and stakeholder reviews
- Decide which item types to eliminate
  - Prioritize implementation if necessary
- Revise remaining item prototypes

SMEs plus

- Evaluating and revising the item types might involve SMEs, staff, vendors, board members, consultants, etc.

SME Review of Traditional MCQs

<table>
<thead>
<tr>
<th>Committee of SMEs</th>
<th>Committee of SMEs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Military model, with Commander in Chief</td>
<td>Democracy</td>
</tr>
<tr>
<td>Leader determines timing for review of each item</td>
<td>&quot;Majority rules&quot; determines timing for review of each item</td>
</tr>
<tr>
<td>Leader decides to accept, revise or reject item</td>
<td>&quot;Majority rules&quot; decides to accept, revise or reject item</td>
</tr>
</tbody>
</table>
Innovative Items

• Suggest that you reject the two previous models
• (While it’s still new) Ensure that a test developer is working with the SMEs
• If 1 SME is holding up group decision on an item – endure with him/her
• Endure MUCH longer than you would in traditional MCQ

Differences of Opinion Among SMEs

• While the committee is wrapped up in the details of the innovative presentation, the objecting SME may be focusing on Step 1 – What is being measured?
• Or, the objecting SME may be Starting at the End – how will we score this item? What is the correct response? How far off is a wrong response?

• Or, the objecting SME may realize that the construct being tested doesn’t require innovation (back to MCQ)
Step 5 – Pilot Test Innovative Item Types

- **Purpose**: Pilot test item type implementation and outcomes
- **Activity**:
  - Implement a full system test of each item type
  - Conduct thorough item analysis of examinee response data
- **Result**: For each item type-
  - a (small) list of remaining changes needed, or
  - approval for operational use

Step 6 – Produce Final Materials

- **Purpose**: Produce final documentation for each new item type
- **Activity**:
  - Produce final versions of:
    - Item templates
    - Item writing guidelines
    - Item writer training
  - Produce other documents related to exam program processes and procedures
- **Result**: Full preparation for implementing the new item types

For more information on the model:

Speaker Contact Information

Lynn C. Webb, Ed. D.  Cynthia G. Parshall, Ph. D.
Testing Consultant    Measurement Consultant

testing@lwebb.com    cparshall@CBTmeasurement.com

Please send an email request for the full presentation