Considering our workshop title...

Hey you are *here*, thinking about assessment ahead of the term start - kudos to you!!

Assessment is hard to do *well*....

Why are multiple choice exams challenging when they *seem so easy*!
The Main Points for Today

1. Designing multiple choice exams takes time and expertise.
2. The stakes involved in multiple choice exams are high.
3. Instructors can develop the skills necessary for designing, creating, and revising multiple choice exams and items.
4. Both students and instructors benefit from the use of high quality multiple choice exams.

Discuss: What can your own experiences tell you?

- Turn to someone beside, in front, or behind you and....
  - Introduce yourself (your name and faculty)
  - Briefly describe your experiences with multiple choice exams
    - as a learner (if applicable)
    - as an instructor (if applicable)
  - If time allows ask them:
    - Why are you attending this workshop?
    - What do you hope to gain from this workshop?
Group Debrief: What did we learn?

What faculties are represented in this group?

Which of the following best describes our experiences as learners with multiple choice exams?

A. Positive  
B. Negative  
C. Neutral  
  ▶ WHY?

Group Debrief: What did we learn?

Which of the following best describes our experiences as instructors with multiple choice exams?

A. Positive  
B. Negative  
C. Neutral  
  ▶ WHY?
Did you know? There are many common issues related to multiple choice exams

At the exam-level
- Nature of content Emphasized
- Adequacy of time allocated
- Balance of difficulty
- Others?

At the item-level
- Use of unfamiliar vocabulary
- Unfamiliar cultural references
- Linking to previous knowledge
- Over-emphasis on reading
- Fit of format to content
- Others?

Did you know? One of the strongest predictors of how you will assess as an instructor is how you were assessed as a learner (Wiliam, 2017)

Now reflect: How your own learning experiences might affect how you offer multiple choice exams?
Theory to practice link: This workshop reflects both ‘best’ practices and my own classroom experiences

Consider what my students might really saying with these:

Your exams are cover what we talked about in class.

I feel your exams are fair.

Your exams make me think harder than I have before.

Today’s workshop is focused on 7 practices related to the lifecycle of a multiple choice exam

- **Design**
  1. Focus on the big ideas taught
  2. Assess fit of content to multiple choice exams
  3. Create exams to assess both higher and lower cognitive skill levels

- **Create**
  4. Develop equitable items
  5. Build technically sound items

- **Use/Revise**
  6. Provide a distraction-free environment
  7. Seek information to improve the exam
1. Design your exam to assess the ‘big ideas’ that have been taught.

If you were going to a surgeon regarding issues with your gall bladder which of these questions would you hope your surgeon addressed in medical school?

- What medical evidence should be considered when deciding if a gall bladder needs to be removed?
- Where is the gall bladder located?

2. Consider what content taught can be assessed by a multiple choice exam.

- What distinguishes multiple choice exams from other possibilities?
  - E.g., ranking, matching, completion, short answer, long answer, performance assessment
- What would help you assess fit of content to item type?
  - What are some examples from the big idea activity that would be better assessed using another possibility?
3. Balance different levels of cognition in the multiple choice exam.

Chose one of the following: What level of Blooms is required for these instructional objectives?

Students will:

1. Recall the basic purposes for commercial advertising
2. Describe the basic techniques advertisers use to sell products to consumers
3. Observe a series of television commercials and identify in each one the selling technique(s) employed
4. Differentiate the observed television commercials on the basis of their effectiveness in promoting a product/service
5. Script and perform a commercial designed to sell a product of their choice
6. Judge the effectiveness of the commercials created by their peers using a class-generated set of criteria

A. Creating
B. Evaluating
C. Analyzing
D. Applying
E. Understanding
F. Remembering
### Bloom's Revised Taxonomy Verbs: Some Ideas

<table>
<thead>
<tr>
<th>Remembering: Remembering previously learned material</th>
<th>Understanding: Explaining ideas or concepts</th>
<th>Applying: Using information in another familiar situation</th>
<th>Analyzing: Breaking information into parts to explore understandings and relationships</th>
<th>Evaluating: Justifying a decision or course of action</th>
<th>Creating: Generating new ideas, products, or ways of viewing things</th>
</tr>
</thead>
<tbody>
<tr>
<td>choose, count, describe</td>
<td>discuss, explain, give examples, illustrate</td>
<td>act, administer, arrange, calculate, change</td>
<td>categorize, characterize, classify</td>
<td>appraise, argue, assess</td>
<td>adapt, anticipate, collaborate, combine</td>
</tr>
<tr>
<td>outline, identify, locate</td>
<td>generalize, paraphrase</td>
<td>observe, dramatize, draw, establish, interview</td>
<td>contrast, diagram, discriminate</td>
<td>conclude, critique</td>
<td>compile, compose, devise, facilitate, create, design</td>
</tr>
<tr>
<td>outline, identify, locate</td>
<td>restate, retell, review</td>
<td>imitate, implement, inform, instruct, interview</td>
<td>distinguish, estimate, examine</td>
<td>judge, justify, prioritize</td>
<td>incorporate, integrate, intervene, make up, invent, plan, pretend</td>
</tr>
<tr>
<td>outline, identify, locate</td>
<td>summarize, trace</td>
<td>modify, prepare, paint, participate, perform</td>
<td>dissect, focus, separate, limit, subdivide</td>
<td>prove, recommend, validate</td>
<td>model, negotiate, organize, propose, rearrange, reconstruct</td>
</tr>
<tr>
<td>outline, identify, locate</td>
<td>record, relate, reproduce, select</td>
<td>transfer, use, transfer</td>
<td>research</td>
<td>reinforce, reorganize, review, structure</td>
<td>reinforce, reorganize, review, structure</td>
</tr>
</tbody>
</table>

### Consider: How well does each exam type measure a cognitive level?

<table>
<thead>
<tr>
<th>Cognitive Level</th>
<th>Multiple choice</th>
<th>Short/Long Answer</th>
<th>Performance Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Remembering</td>
<td>★★★★★</td>
<td>★★★★</td>
<td>★</td>
</tr>
<tr>
<td>Understanding</td>
<td>★★★★★</td>
<td>★★★★</td>
<td>★</td>
</tr>
<tr>
<td>Applying</td>
<td>★★★★★</td>
<td>★★★★</td>
<td>★</td>
</tr>
<tr>
<td>Analyzing</td>
<td>★★★★★</td>
<td>★★★★</td>
<td>★</td>
</tr>
<tr>
<td>Evaluating</td>
<td>★</td>
<td>★</td>
<td>★★★★★</td>
</tr>
<tr>
<td>Creating</td>
<td>★</td>
<td>★</td>
<td>★★★★★</td>
</tr>
</tbody>
</table>

**How can this verb list help?**

**How might this information influence your design of a multiple choice exam?**
How can a table of test specificity/blueprint help?

EDPY 303 Midterm Exam Blueprint - Fall 2019

<table>
<thead>
<tr>
<th>Topics</th>
<th>Remembering/Understanding</th>
<th>Applying and Above</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assessment Audiences</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Curriculum, Instruction, and Assessment Alignment</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Bloom’s Taxonomy</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>Fair Assessment</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Reliability and Validity</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Selected-Response Items</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Completion Items</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Gathering Evidence of Learning</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>22</strong></td>
<td><strong>38</strong></td>
</tr>
</tbody>
</table>

Content areas tested

EDPY 303 Midterm Exam Blueprint - Fall 2019

Course Content

<table>
<thead>
<tr>
<th>Level of Thinking – Blooms’ Taxonomy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topic #</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>6</td>
</tr>
</tbody>
</table>

When/how in the lifecycle can this activity be helpful?

Short or long response might be better?

True/False, Matching, Completion, Multiple-choice?
Why Balance Items Assessing Higher and Lower Cognitive Skills?

Design practices for high quality multiple choice exams

- Link the big ideas you are assessing with what and how you teach.
- Ensure what you are choosing to assess using a multiple choice exam is a good fit with the format.
- Design exams so that there is a good balance of cognitive skill levels:
  - Blooms taxonomy is helpful for thinking about what you cognitive level you want students to be able to do and know.
  - Exam blueprint/Table of test specifications is a useful tool to make explicit the emphasis of content and level of cognition.
Discuss: What are some common reasons for selecting correct or incorrect multiple-choice answers?

Turn to someone beside, in front, or behind you and....

Correct choice:
- Possesses required information or skill
- Uses partial information that favours answer
- Uses clues in the item and/or test
- Makes a blind, lucky guess

Incorrect choice:
- Lacks required information or skill
- Uses partial information or misinformation that favours a distracter
- Makes a blind, unlucky guess

Introducing some language for multiple choice items....

Calculus was independently developed by Newton and
A. Barrow
B. Kepler
C. Leibniz
D. Pascal

These are the “alternatives”, “choices”, or “options”.
This is the stem.

This is the “keyed response” or correct choice.

These are the “distracters”, (alternatives that are incorrect).
4. Create items that are equitable for all learners

1. The purpose of the cluss in furmpaling is to remove
   A. cluss-prags
   B. tremalis
   C. cloughs
   D. plumots

2. Trassig is true when
   A. lusp trasses the vom
   B. the viskal flans, if the viskal is donwil or zortil
   C. the belgo frulls
   D. dissles lisk easily

3. The sigla frequently overskes the trelsum because
   A. all siglas are mellious
   B. siglas are always votial
   C. the trelsum is usually tarious
   D. no trelsa are feskable

What makes these items less equitable?

An item can be considered more equitable when it is...

- free from racial, ethnic and sexual bias
- free of irrelevant material
- stated in appropriate and clear language
- free from pop cultural references that would not be familiar to all students
5. **Build items that are technically sound**

4. The fribbled breg will minter best with an
   A. Derst
   B. Morst
   C. Sorter
   D. Ignu

5. Among the reasons for tristal doss are
   A. The sabs foped and the foths tinzed
   B. The kredges roted with orots
   C. Few rakobs were accepted in sluth
   D. Most of the polats were thonced

6. Which of the following (is, are) always present when trossels are being gruven?
   A. Sot
   B. Vost
   C. Runt and vost
   D. Shum and vost

7. The mintering function of the ignu is most effectively carried out in connection with
   A. Raxma tol
   B. The groshing stantol
   C. The fribbled breg
   D. A frally sush

---

An item is considered more **technically sound when it is**..

- the item is free of verbal clues to the answer
- the stem is focused on a single, meaningful problem
- key words in the stem are emphasized as needed
- distracters are all plausible
- alternatives are homogenous and parallel in structure
- alternatives are in some order that is logical and easily understood
- legitimate use of ‘all of the above’ and ‘none of the above’
- the item has correct spelling and grammar
Guidelines for a High-Quality Multiple Choice Distractors

- Use common student errors
- Use language appropriate to the students
- Alternatives should all be homogeneous (length and complexity)
- Alternatives should be arranged in a logical order (e.g. pyramid or reverse-pyramid, alphabetical order, numerical order, etc.)

Try to Avoid:

- The 3:1 Split...One of the alternatives stands out “like a sore thumb”
- A multiple-choice item that is actually a “true-false” item because nearly all students can eliminate two distracters
- Unjustified use of “all/none of the above”

What If I have existing items?

If you have access to a test bank or old exams - consider adapting them in the following ways:

- Increase the level of cognition
- Change item type
- Reword to focus it on the “big idea”
- Improve its technical soundness
- Revise to be more equitable
Creating practices for high quality multiple choice exams

- Design items that are equitable for all.
  - Is there anything that could ‘inadvertently get in the way’ of a valid correct answer.
- Ensure items are technically sound.
  - Can the distractors and stems be considered high quality?
  - Are there any apparent clues?

6. Provide a distraction-free environment for writing the exam

▶ Discuss: What are some common distractions that can occur?

Distractions:
- Individuals making noise
- Fire alarm
- Poor items
- Others?

Prevention:
- Consider alternative ways of administering (LAC)
- Reviewing procedures ahead of time
- Others?
7. Seek information for improving your exam.

- How many have written your own exams prior to administration?
  - What information did you glean?
- How many have revised your exams after they have been administered?
  - What information did you use?

Did you know? Information is available to improve exams and items

- How long did students need to complete the exam?
  - When did the first person, ½, and the last person leave?
- What was the exam reliability?
  - Make sure all the items could be judged unequivocally correct.
- What was the individual item difficulty?
  - Range from Easy >0.9, Moderate 0.5 to 0.9, Difficult <0.5
- To what extent did the individual item discriminate among higher and lower achieving students?
  - Helps decide to keep, change or toss out.
Revising practices for high quality multiple choice exams

- Prevent as many distractions as possible.
- Seek and use information for improving your exam
  - Was enough time provided?
  - Was the exam reliable?
  - Can the exam results be considered valid?

The lifecycle of high quality assessment

- **Now**: Invest time thinking about your assessment strategies
- **Throughout the term**: Revise items, ensure alignment of content to assessment
- **End of term**: Seek information for improvement
How did we do relative to the main points?

Have you increased your understanding of

✓ The time and expertise it takes to design multiple choice exams?
✓ The high stakes involved in multiple choice exams?
✓ How instructors can develop the skills necessary for high quality multiple choice exams and items?
✓ Why both students and instructors can benefit from the use of high quality multiple choice exams?

Just before we wrap up: Now consider what you heard about at the beginning about:

► Why attend this workshop?
► What did you hope to gain from this workshop?
► To what extent did we reach your goals?
► What questions remain?
Thank you for coming....
Don’t forget about how the CTL can help....

▸ Resources available on our website
  ▸ Literature and more podcasts are forthcoming!
▸ Request an individual consultation
  ▸ The link is on our homepage!
▸ Register for an assessment-related workshop or seminar
  ▸ Check our calendar of events or subscribe to our listserv for more information.
▸ Don’t see what you want?
  ▸ Tell us about something you want to see offered