

Creating quality multiple-choice exams: The good, the bad and the ugly

CTL Teaching Institute: From Surviving to Thriving
August 2019



Cheryl Poth, Ph.D.
Associate Director (Assessment), CTL
cpoth@ualberta.ca

Considering our workshop title...

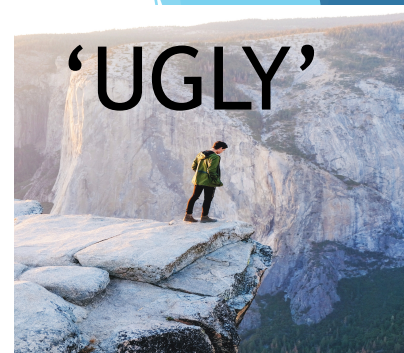


- ▶ Hey you are *here*, thinking about assessment ahead of the term start - kudos to you!!

C. Poth - Creating quality multiple-choice items

'Bad'

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



- ▶ Why are multiple choice exams challenging when they *seem so easy*!

3

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.

C. Poth - Creating quality multiple-choice items

4

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?

C. Poth - Creating quality multiple-choice items

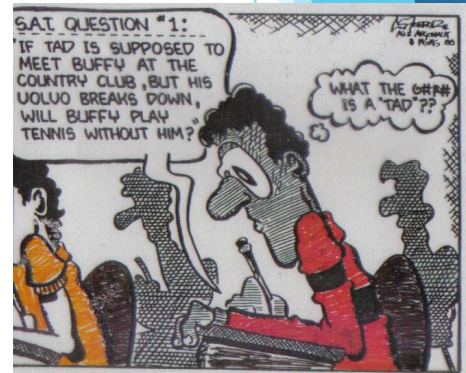
5

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?

C. Poth - Creating quality multiple-choice items

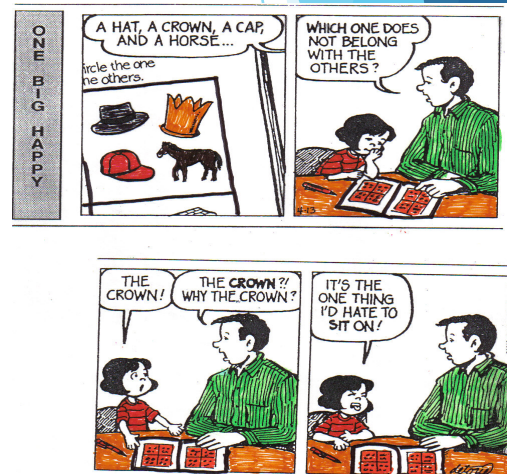


Group Debrief: What did we learn?

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

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

C. Poth - Creating quality multiple-choice items



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?

C. Poth - Creating quality multiple-choice items

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?

8



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

Now reflect: How your own learning experiences might affect how you offer multiple choice exams?

C. Poth - Creating quality multiple-choice items

9



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.

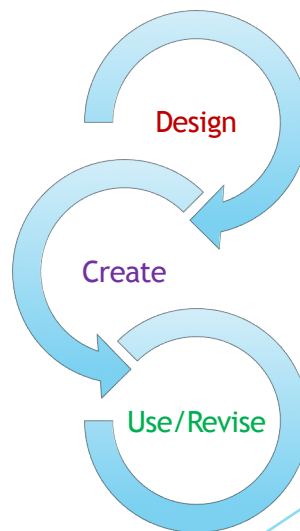
Your exams make me think harder than I have before.

C. Poth - Creating quality multiple-choice items

10

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



C. Poth - Creating quality multiple-choice items

11

1. Design your exam to assess the ‘big ideas’ that have been taught.

Consider the importance of the following...

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?



What can this activity tell us about the content you should assess in your multiple choice exam?

C. Poth - Creating quality multiple-choice items

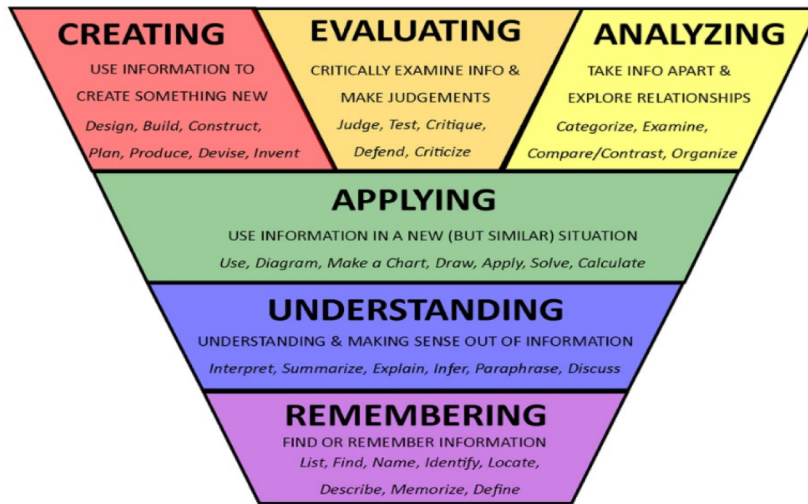
2. Consider what content taught **can be** assessed by a multiple choice exam.

- ▶ 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?

C. Poth - Creating quality multiple-choice items

13

3. Balance different levels of cognition in the multiple choice exam.



<https://carleton.ca/viceprovost/blooms-taxonomy/>

Bloom's Taxonomy



Dr. Benjamin Bloom
Creator of a
Taxonomy of
Educational
Objectives

14

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

C. Poth - Creating quality multiple-choice items

15

Bloom's Revised Taxonomy Verbs: Some Ideas

Remembering: Remembering previously learned material

choose	find	locate	outline	record	sequence	tell
count	identify	match	quote	relate	show	underline
define	label	memorize	recall	reproduce	state	view
describe	list	name	recite	select	sort	write

Understanding: Explaining ideas or concepts

convert	generalize	paraphrase	restate	summarize		
discuss	give examples	predict	retell	trace		
explain	illustrate	report	review			

Applying: Using information in another familiar situation

act	chart	discover	imitate	modify	prepare	transfer
administer	collect	dramatize	implement	operate	provide	use
arrange	compute	draw	inform	paint	solve	utilize
calculate	contribute	establish	instruct	participate		
change	determine	extend	interview	perform		

Analysing: Breaking information into parts to explore understandings and relationships

break down	compare	deduce	dissect	experiment	research	
categorize	contrast	diagram	distinguish	focus	separate	
characterize	correlate	differentiate	estimate	limit	subdivide	
classify	debate	discriminate	examine	point out		

Evaluating: Justifying a decision or course of action

appraise	conclude	defend	judge	prove	support	
argue	criticize	evaluate	justify	recommend	validate	
assess	critique	interpret	prioritize	reframe		

Creating: Generating new ideas, products, or ways of viewing things

adapt	compile	develop	incorporate	model	produce	reinforce
anticipate	compose	devise	integrate	negotiate	progress	reorganize
collaborate	construct	facilitate	intervene	organize	propose	revise
combine	create	formulate	invent	plan	rearrange	structure
	design	generate	make up	pretend	reconstruct	

How can this verb list help?

Consider:

How well does each exam type measure a cognitive level?

	Multiple choice	Short/Long Answer	Performance Assessment
Remembering	★★★★★	★★★☆☆	★☆☆☆☆
Understanding	★★★☆☆	★★★★★	★★☆☆☆
Applying	★★★☆☆	★★★★★	★★★★☆
Analyzing	★★★☆☆	★★★★★	★★★★☆
Evaluating	★☆☆☆☆	★☆☆☆☆	★★★★★
Creating	★☆☆☆☆	★★★★☆	★★★★★

How might this information influence your design of a multiple choice exam?

C. Poth - Creating quality multiple-choice items 17

How can a table of test specificity/blueprint help?

EDPY 303 Midterm Exam Blueprint - Fall 2019

C. Poth - Creating quality multiple-choice items

Content areas tested	Topics	Taxonomy levels tested		
		Remembering/ Understanding	Applying and Above	
	Assessment Audiences	1	1	2
	Curriculum, Instruction, and Assessment Alignment	3	7	10
	Bloom's Taxonomy	3	8	7
	Fair Assessment	3	4	7
	Reliability and Validity	3	3	6
	Selected-Response Items	5	10	15
	Completion Items	0		
	Gathering Evidence of Learning	4		
	Total	22	38	60

Number of items at each taxonomy level for each content area

Course Content	Level of Thinking – Blooms' Taxonomy				Total %
	Remembering	Understanding	Applying	Analyzing	
Topic # 1	1-5, 6,7,10,13,14				20
2		8,9,11,12, 15-28, 31,32			40
3			29,30, 33-38		16
4			39,41-46		14
6				40,47-50	10
Total			30%	10%	100%

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

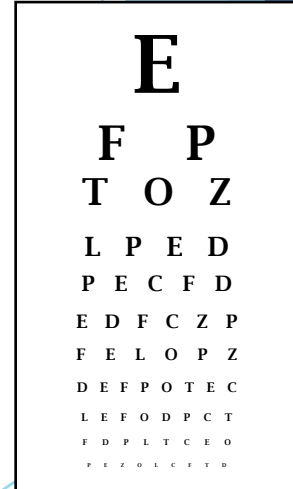
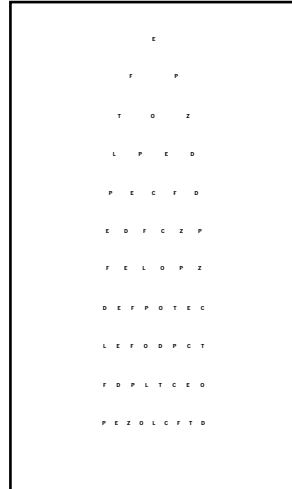
True/False, Matching, Completion, Multiple-choice?

Short or long response might be better?

Why Balance Items Assessing Higher and Lower Cognitive Skills?



C. Poth - Creating quality multiple-choice items



20

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.

C. Poth - Creating quality multiple-choice items

21

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

C. Poth - Creating quality multiple-choice items

22

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).

C. Poth - Creating quality multiple-choice items

23

4. Create items that are equitable for all learners

What makes these items less equitable?

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 overfesks the tresum because
 - A. all siglas are mellious
 - B. siglas are always votial
 - C. the tresum is usually tarious
 - D. no tresla are fesksable

C. Poth - Creating quality multiple-choice items

24

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

C. Poth - Creating quality multiple-choice items

25

5. Build items that are technically sound

What makes these items less technically sound?

4. The fribbled breg will minter best with **an**
- Derst
 - Morst
 - Sorter
 - Ignu**

6. Which of the following (is,are) always present when trossels are being gruvén?
- Sot
 - Vost**
 - Runt and **vost**
 - Shum and **vost**

5. Among the reasons for trystal doss are **are**
- The sabs foped and the foths tinzed**
 - The kredges roted with orots
 - Few rakobs were accepted in sluth
 - Most of the polats were thonced

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

C. Poth - Creating quality multiple-choice items

26

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

C. Poth - Creating quality multiple-choice items

27

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”**

C. Poth - Creating quality multiple-choice items

28

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

C. Poth - Creating quality multiple-choice items

29

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?

C. Poth - Creating quality multiple-choice items

30

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?

C. Poth - Creating quality multiple-choice items

31

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?

C. Poth - Creating quality multiple-choice items

32

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.

```

****TEST STATISTICS****
NUMBER OF EXAMINEES      118
TOTAL SCORE:
      MEAN      71.21
      VARIANCE  72.68
      STANDARD DEVIATION  8.53
      KR-20 RELIABILITY  0.8393
      S.E. OF MEASUREMENT  3.4173
  
```

$$p = \frac{\text{\# of correct answers}}{\text{\# of people taking the test}}$$

$$D = \frac{\text{\# correct (up)} - \text{\# correct (low)}}{\text{\# of students in either group}}$$

C. Poth - Creating quality multiple-choice items

33

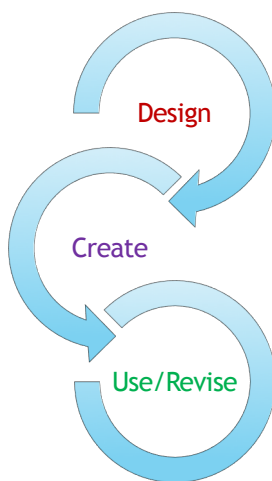
Revising practices for high quality multiple choice exams

- ✓ 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?

C. Poth - Creating quality multiple-choice items

34

The lifecycle of high quality assessment



C. Poth - Creating quality multiple-choice items

- ▶ **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

35

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?

C. Poth - Creating quality multiple-choice items

36

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?

C. Poth - Creating quality multiple-choice items

37

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

C. Poth - Creating quality multiple-choice items

38