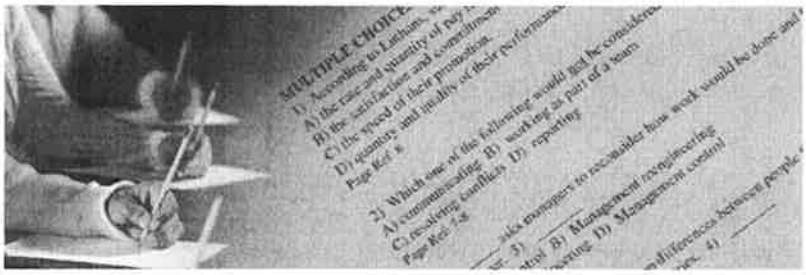
Creating Effective Multiple Choice Questions

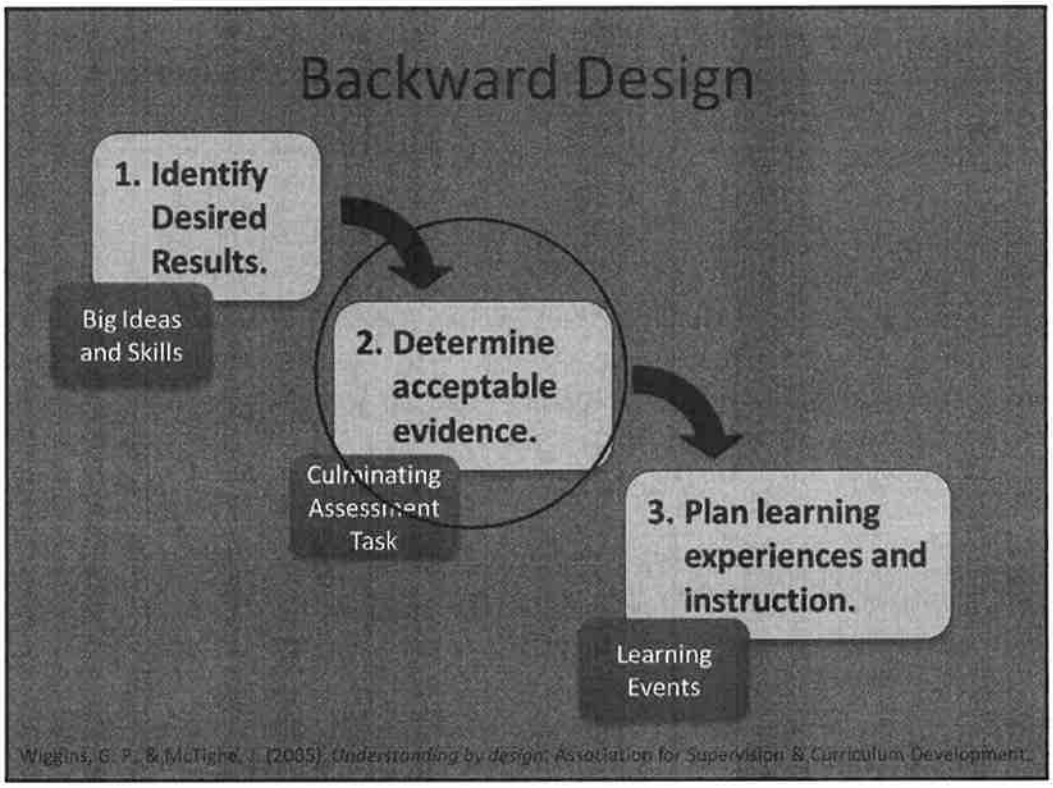


Dan Hubert
Associate Program Director
Learning Outcomes Assessment

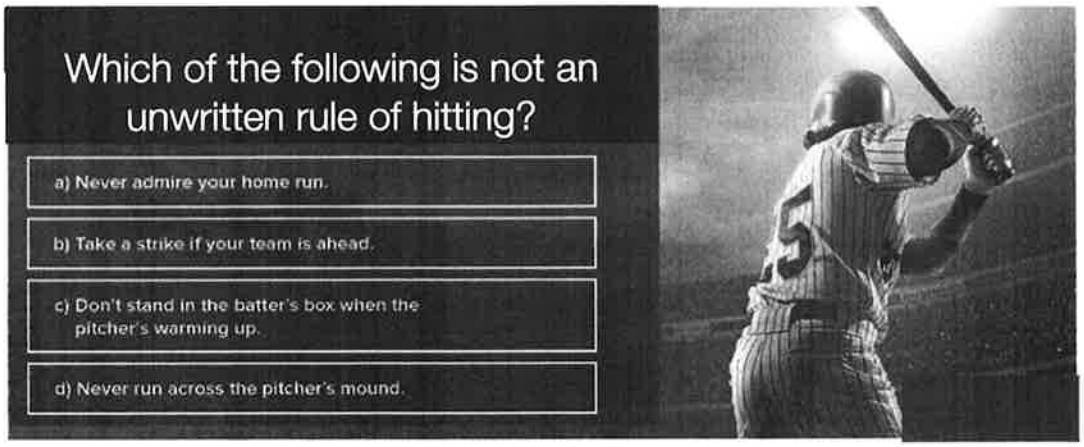
Today's goals

By the end of this workshop you will be better able to…

1. Describe components as well as advantages/disadvantages of using MCQs
2. Understand properties of well- vs poorly-designed MCQs
3. Recognize cues used by test-wise students
4. Draft various cognitive orders of learning MCQs relevant to your discipline



What makes you uncomfortable about this question and answer choices?



Edward Jones advertisement on MLB.com on 10/1/18

Quality assessment items…

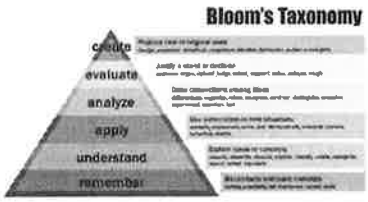
…are clearly stated and unambiguous to the learner

…go beyond rote learning, definitions, & terminology

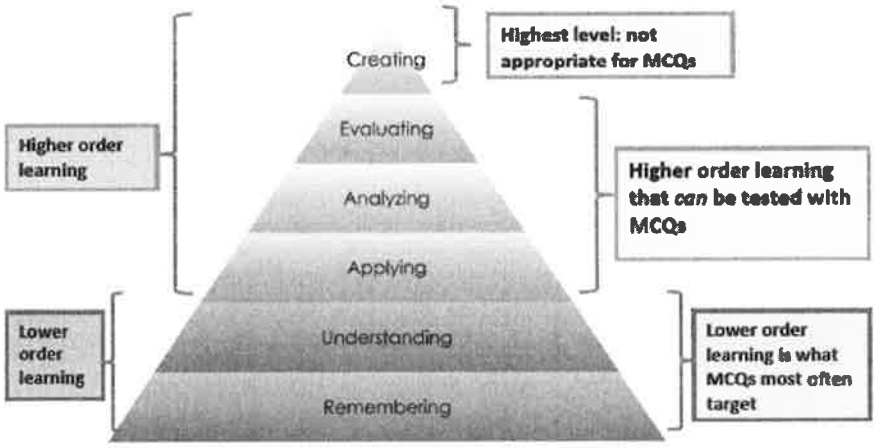
…evaluate broad understanding of material

…can be difficult for those who don't know material, but straightforward for those who do

…assess a variety of cognitive levels

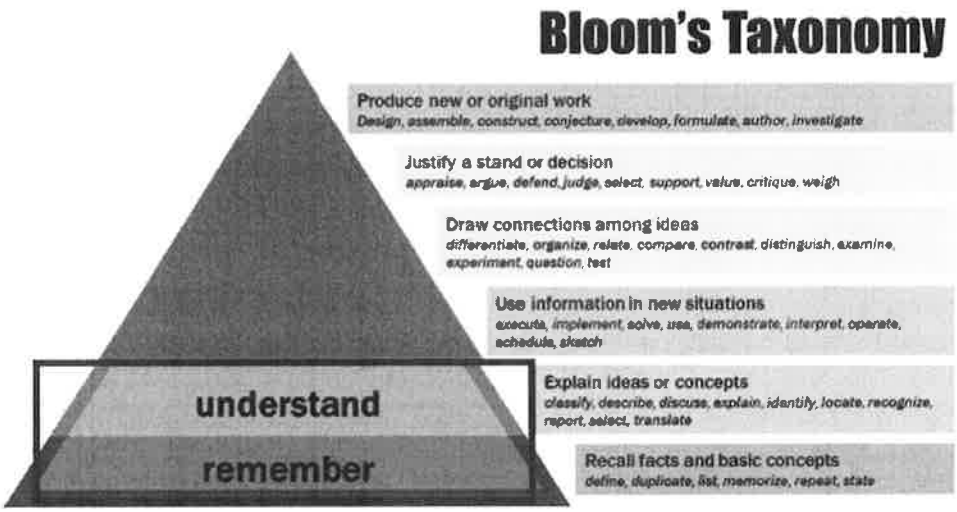


Cognitive Orders of Learning/Thinking



https://mcgill.ca/skillsets/files/skillsets/mcq_handout3.pdf
A Taxonomy for Learning, Teaching & Assessing: A Revision of Bloom's Taxonomy of Education Objectives. Anderson & Krathwohl (Eds.)

Evidences of Learning



Remembering and Understanding

(Recall is typically lowest cognitive order of learning—usually fact-based)

[STEM→] The mean of a set of exam scores is also called the:

OPTIONS

A. most frequently occurring score.
B. 50th percentile.
C. measure of the range of scores.
D. arithmetic average.*

DISTRACTERS

MC Item Development Process

1) Draft the stem and correct answer

2) Create options/distracters

3) Review, revise, confirm

Drafting a Stem and Answer

- "Question" or "Completion" format
- Positively-worded
- Use a full, complete sentence
- Include as much of the option statement as possible
- Limit irrelevant material
- Use *clear*, *succinct* and *familiar* language
- Avoid negative wording unless absolute necessary.

Work time: Create Stems + Answer

- Create 3 lower-cognitive order test item stems + correct answers in either "question" and "completion" formats
- If you brought a test/quiz/exam, locate items/item stems that are lowest cognitive order/fact-based and
  - Create new or revise item stems based on those initial items
  - Confirm the cognitive learning level for which you are assessing

3.3 The Six Categories of the Cognitive Process Dimension and Related Cognitive Processes*

Process Categories	Cognitive Processes and Examples
1. REMEMBER—Retrieve relevant knowledge from long-term memory.	
1.1 Recognizing	(e.g., Recognize the dates of important events in U.S. history)
1.2 Recalling	(e.g., Recall the dates of important events in U.S. history)
2. UNDERSTAND—Construct meaning from instructional messages, including oral, written, and graphic communication.	
2.1 Interpreting	(e.g., Paraphrase important speeches and documents)
2.2 Exemplifying	(e.g., Give examples of various artistic painting styles)
2.3 Classifying	(e.g., Classify observed or described cases of mental disorders)
2.4 Summarizing	(e.g., Write a short summary of the events portrayed on videotapes)
2.5 Inferring	(e.g., In learning a foreign language, infer grammatical principles from examples)
2.6 Comparing	(e.g., Compare historical events to contemporary situations)
2.7 Explaining	(e.g., Explain the causes of important eighteenth-century events in France)
3. APPLY—Carry out or use a procedure in a given situation.	
3.1 Executing	(e.g., Divide one whole number by another whole number, both with multiple digits)
3.2 Implementing	(e.g., Determine in which situations Newton's second law is appropriate)
4. ANALYZE—Break material into constituent parts and determine how parts relate to one another and to an overall structure or purpose.	
4.1 Differentiating	(e.g., Distinguish between relevant and irrelevant numbers in a mathematical word problem)
4.2 Organizing	(e.g., Structure evidence in a historical description into evidence for and against a particular historical explanation)
4.3 Attributing	(e.g., Determine the point of view of the author of an essay in terms of his or her political perspective)
5. EVALUATE—Make judgments based on criteria and standards.	
5.1 Checking	(e.g., Determine whether a scientist's conclusions follow from observed data)
5.2 Critiquing	(e.g., Judge which of two methods is the best way to solve a given problem)
6. CREATE—Put elements together to form a coherent or functional whole; reorganize elements into a new pattern or structure.	
6.1 Generating	(e.g., Generate hypotheses to account for an observed phenomenon)
6.2 Planning	(e.g., Plan a research paper on a given historical topic)
6.3 Producing	(e.g., Build habitats for certain species for certain purposes)

Revised Bloom's Taxonomy Process Verbs, Assessments, and Questioning Strategies

Level of Taxonomy	Definition	Process Verbs		Assessments		Question Stems
Creating	Generating new ideas, products, or ways of viewing things Designing, constructing, planning, producing, inventing	Act Arrange Assemble Combine Compose Construct Create Design Develop Devise Formulate	Generate Improve Infer Invent Imagine Plan Predict Prepare Revise Show Write	Advertisement Poem Blueprint Cartoon Collage Film Formula Invention New game	Newspaper Painting Plan Play Song Story Video	-Can you design a…to…? -Can you see a possible solution to…? -How would you devise your own way to…? -What would happen if…? -How many ways can you…? -Can you create new and unusual uses for…?
Evaluating	Justifying a decision or course of action Checking, hypothesizing, critiquing, experimenting, judging	Argue Assess Choose Compare Conclude Criticize Debate Decide Defend	Determine Evaluate Justify Prioritize Rate Recommend Support Tell why Value	Conclusion Debate Editorial Investigation Judgment Opinion	Recommendation Report Survey Verdict	-Is there a better solution to…? -What do you think about…? -Do you think…is a good or bad thing? -How would you feel if…? -How effective are…? -What are the pros and cons of …?
Analyzing	Breaking information into parts to explore understandings and relationships Comparing, organizing, deconstructing, interrogating, finding	Calculate Categorize Classify Compare Contrast Diagram Differentiate Discover Distinguish Examine Experiment	Group Interpret Investigate Order Organize Question Relate Research Sequence Solve Survey	Chart Checklist Database Diagram Graph Illustration Investigation	List Outline Plan Questionnaire Report Spreadsheet Summary	-Which events could not have happened? -How is …similar to …? -What are some other outcomes? - Why did …occur? -What was the problem with…?

From AACSB Assurance of Learning (AoL) 10.16.16 Notebook

Revised Bloom's Taxonomy Process Verbs, Assessments, and Questioning Strategies

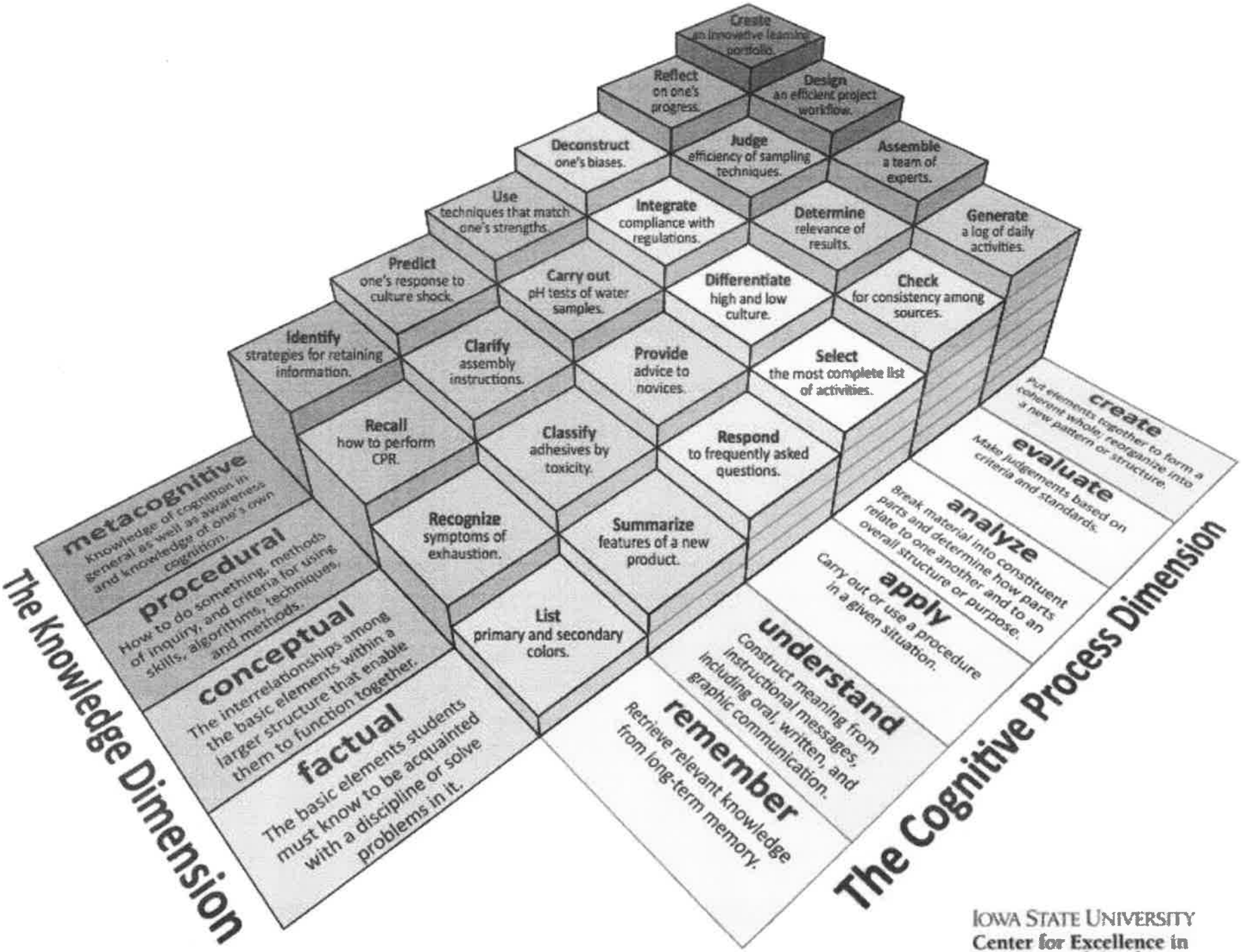
Level of Taxonomy	Definition	Process Verbs		Assessments		Question Stems
Applying	Using information in another familiar situation Implementing, carrying out, using, executing	Adapt Apply Calculate Change Compute Demonstrate Dramatize Draw Experiment Illustrate	List Make Manipulate Practice Produce Sequence Show Solve Teach Use	Demonstration Diagram Experiment Illustration Journal Lesson Map Model	Performance Poster Prediction Presentation Report Scrapbook Simulation	-Do you know of another instance where…? -Can you group…? -Which factors would you change…? -What questions would you ask of…? -From the information given, can you develop a set of instructions about…?
Understanding	Explaining ideas or concepts Interpreting, summarizing, paraphrasing, classifying, explaining	Ask Calculate Convert Describe Discuss Explain Give examples Identify Locate	Observe Recognize Report Research Retell Review Summarize Tell	Debate Definition Dramatization Example Explanation Label List	Outline Quiz Recitation Reproduction Story Problems Summary Test	-Can you write in your own words? -How would you explain…? -What could happen next? -Who do you think…? -What was the main idea…?
Remembering	Recalling information Recognizing, listing, describing, retrieving, naming, finding	Choose Cite Define Describe Give example Group Know Label List Listen Locate	Match Memorize Name Quote Recall Recite Record Repeat Select Underline	Definition Fact Label List Quiz	Reproduction Test Workbook Worksheet	-What happened after…? -How many…? -What is…? -Who …? -Can you name…? -Which is true or false?

From AACSB Assurance of Learning (AoL) 10.16.16 Notebook

Fink's Taxonomy Verbs for Learning Outcomes

Foundational Knowledge	Associate, Compare, Contrast, Describe, Define, Explain, Give example, Identify, Illustrate, Indicate, List, Name, Paraphrase, Recite, Recognize, Remember, Repeat, Restate, Tell
Application	Analyze, Assess, Critique, Calculate, Create, Coordinate, Demonstrate, Draw, Employ, Estimate, Give example, Illustrate, Imagine, Interpret, Judge, Locate, Make decisions, Manage, Measure, Operate, Perform, Prescribe, Record, Solve, Use
Integration	Associate, Blend, Combine, Compare, Connect, Contrast, Correlate, Differentiate, Integrate, Intermix, Join, Link, Relate, Synthesize, Unite
Human Dimensions	Acquire, Advise, Advocate, Behave, Communicate, Collaborate, Cooperate, Empathize, Express, Feel, Help, Influence, Initiate, Inspire, Interact, Involve, Lead, Mediate, Motivate, Negotiate, Nurture, Promote, Protect, Reconcile, Resolve, Reflect, Respect, Respond, Share, Support, Unite
Caring	Commit to, Decide to, Demonstrate, Develop, Discover, Explore, Express, Identify, Interpret, Pledge, Recognize, Value, Reflect, Renew, Revitalize, Share, State, Value
Learning to Learn	Analyze, Construct knowledge, Critique, Create a plan, Describe how to, Develop a learning plan, Identify resources, Identify your learning style, Identify needs, Inquire, Formulate Frame questions, Generalize knowledge, Predict performance, Reflect, Research, Self-assess. Self-regulate, Self-monitor, Set goals, Take responsibility, Transfer knowledge

From: http://thepeakperformancecenter.com/educational-learning/thinking/blooms-taxonomy/verbs-learning-objectives/finks-taxonomy-verbs/



IOWA STATE UNIVERSITY
Center for Excellence in Learning and Teaching

Basic Rules for Drafting Options

Follow the grammatical structure of the stem

List vertically, separate lines

3-5 total options, ~ same length

Arrange in logical order (e.g. low - high, A - Z)

Randomize correct answer

Avoid weak distracters

Characteristics of Distractors

Plausible enough to be chosen and possibly be argued as correct

True statements that don't answer the question

Often incorporate common student errors or misconceptions

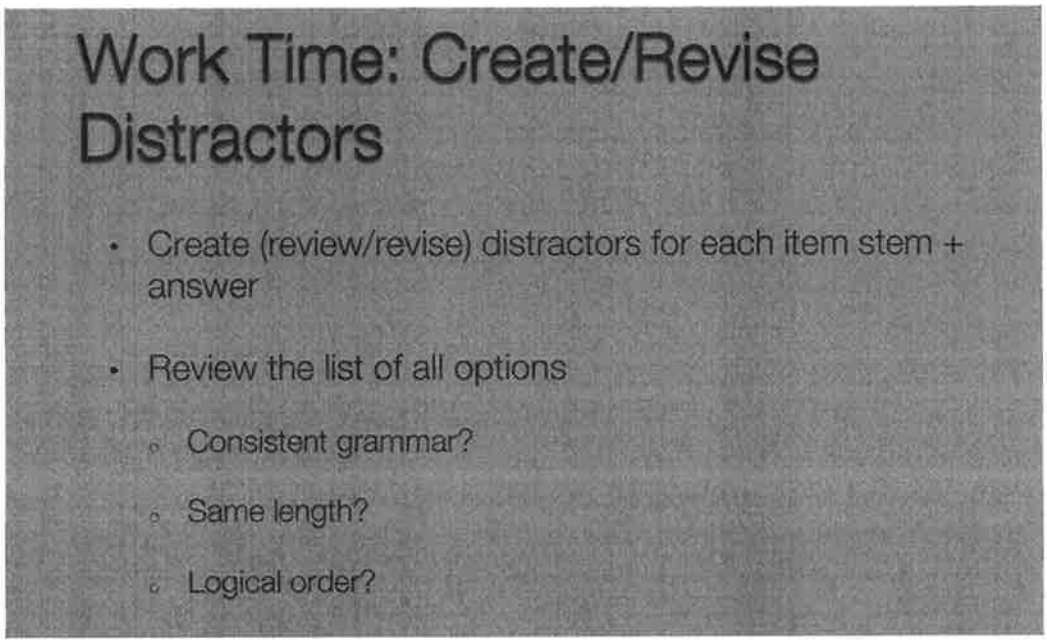
Familiar, yet incorrect, words or phrases

Each one stands a reasonable chance of selection by those who doesn't fully know the material

Example

Of the following campus buildings, which one opened most recently for use by the ND community?

A. Washington Hall
B. Notre Dame Field House
C. Walsh Family Hall of Architecture
D. Duncan Student Center



Verbal Cues

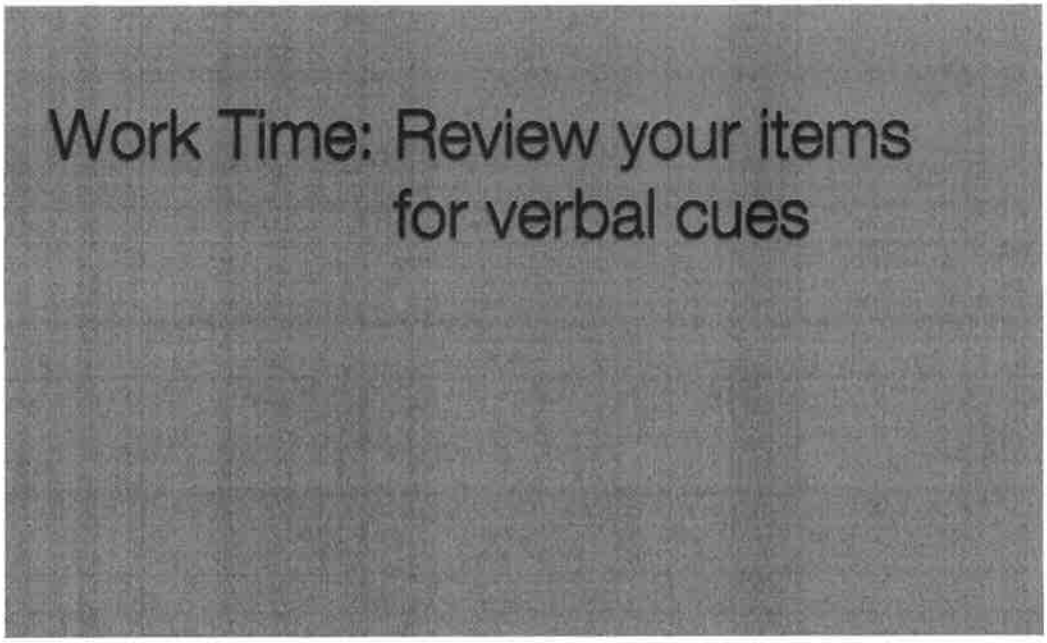
<u>Correct Answers</u>

- One option is longer, more detailed, or more complex
- Key word appears in the stem and in only one option
- One option with a vague word or phrase like "usually," "typically" or "may be"

Verbal Cues

Distracters

- Two options with the same meaning
- Specific determiners: absolute or extreme terms like "always," "never," or "all"
- An option that is clearly implausible or humorous
- An error in grammar or spelling
- **One option in textbook/lecture language (correct), others in everyday language



Rewriting a Wordy Stem

Suppose you are a mathematics professor who wants to determine whether or not your teaching of a unit on probability has had a significant effect on your students. You decide to analyze their scores from a test they took before the instruction and their scores from another exam taken after the instruction. Which of the following t-tests is appropriate to use in this situation?

a. Dependent samples.

b. Heterogeneous samples.

c. Homogeneous samples.

d. Independent samples

Rewritten, it may look something like this...

When analyzing your students' pretest and posttest scores to determine if your teaching has had a significant effect, an appropriate statistic to use is the t-test for:

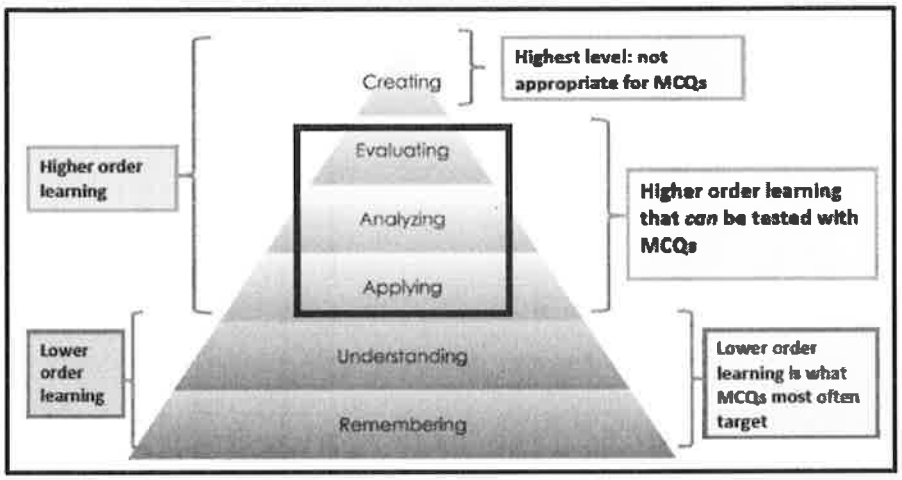
a. Dependent samples.

b. Heterogenous samples.

c. Homogenous samples.

d. Independent samples

Higher Cognitive Orders of Learning



Higher Cognitive Order Items

Some examples include…

- Comparison-Application
- Incomplete scenario
- Premise-Consequence
- Analogy
- Evaluation
- Case study

Applying, Analyzing, and Evaluating

The mean of a set of test scores is called the:

A. most frequently occurring score.
B. 50th percentile.
C. arithmetic average.*
D. measure of the score range.

Recall/Remember

A university-developed aptitude test used in Honors Program admissions decisions was given to a group of seven applicants. Their scores were: 70, 89, 94, 72, 72, 80 and 98. The mean score for these test results is:

A. 72
B. 76
C. 82*
D. 90

Application

Table 5. Item Flipping

Original Items:	Flipped Items:
Which of the following best describes what is meant by 'formative assessment'? A. is based on the student's attitudes, interests and values B. is designed primarily to evaluate learning C. is usually high-stakes D. provides information to modify teaching and learning*	A teacher uses a strategy called *Thumbs Up, Thumbs Down* with her students. This illustrates the use of: A. affective assessment B. formative assessment* C. diagnostic assessment D. summative assessment

According to Piaget's theory of cognitive development, what is 'accommodation'?

A. the ability to think logically
B. the diminishing of a response to a frequently repeated stimulus
C. altering one's existing schemas as a result of new information*
D. an inability to understand perspectives besides one's own

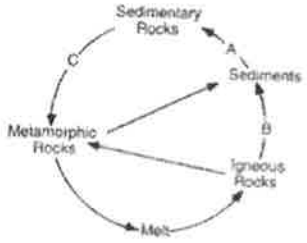
After Sarah learned that penguins can't fly, she had to modify her existing concept of birds. This best illustrates the process of:

A. Accommodation*
B. Conservation
C. Habituation
D. Egocentrism

(Adapted from: ProProfs, n.d.)

Scully, Darina (2017). Constructing Multiple-Choice Items to Measure Higher-Order Thinking. Practical Assessment, Research & Evaluation, 22(4). Available online: http://pareonline.net/getvn.asp?v=22&n=4

Bloom's Taxonomy



In the above diagram, path A is indicative of _____.

1. burial and lithification
2. cooling and uplifting
3. heating and crystallization
4. weathering and deposition

1. Quantity A
$x^2 + 1$

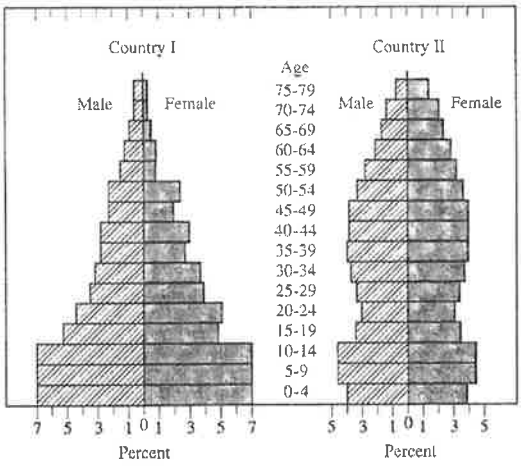
Quantity B
$2x - 1$

A. Quantity A is greater.

B. Quantity B is greater.

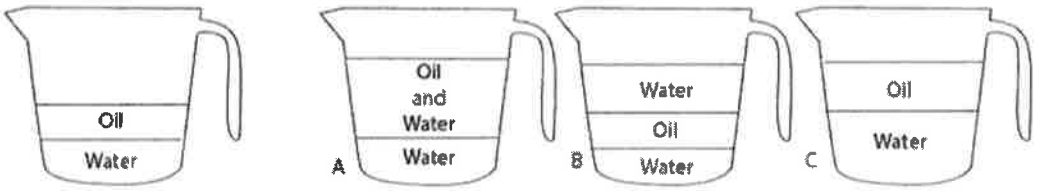
C. The two quantities are equal.

D. The relationship cannot be determined from the information given.

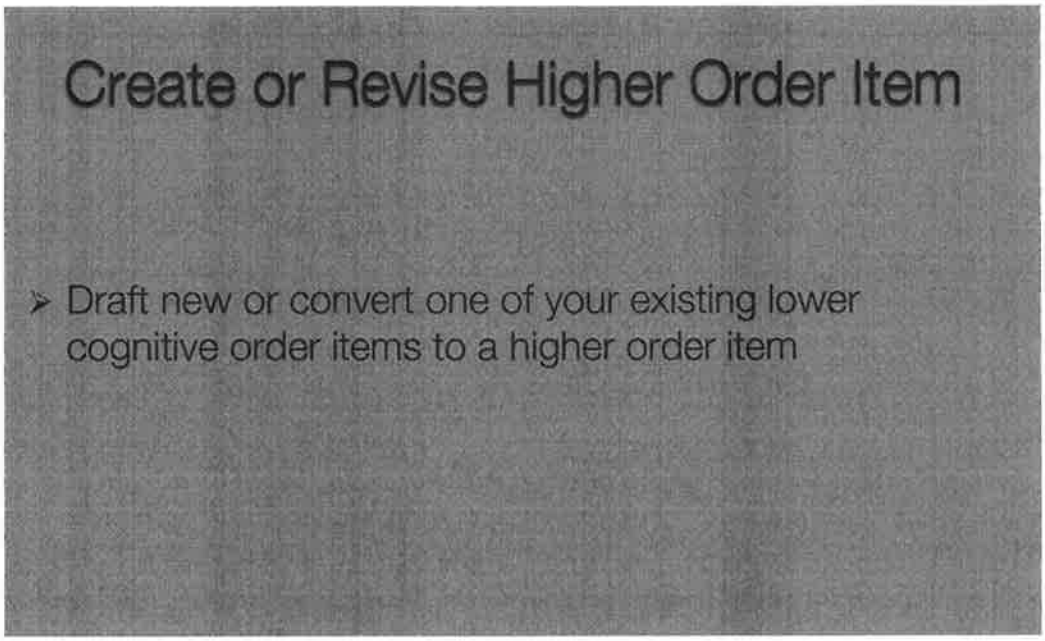


Which best approximates the ratio of males to females among individuals below 15 years of age?

	C 1	C 2
A.	1 : 1	1 : 1
B.	0.75 : 1	0.75 : 1
C.	0.5 : 1	0 .5 : 1
D.	1 : 1	0.8 : 1
E.	0.75 : 1	1 : 1



A glass container holds water and oil. If you add more water, which of the following choices--A, B, or C--best represents the resulting mix?



Avoid these common mistakes/issues

Stems with "except" (*All are signs of XYZ <u>except</u> ...*)

- Negative wording of stem (*Which of the following is NOT...*)
- Inconsistencies in grammar, length, or style
- Microscopically-fine distinctions between options (unless absolutely necessary).
- All / None of the above, All, Always, or Never
  "Both A & B"

Trying to intentionally "trick" students

Test Design, Length, Balance

Is your goal to…

→ Fill the **entire** class time?

→ Keep the students **busy**?

→ Use more questions than fewer?

Note: You should complete your own exams in about ¼ of time it takes your students

Test Question Matrix – *Example 1*

40 questions– *Lower cognitive order heavy*

	Topic A	Topic B	Topic C	Topic D	
Remember	2	2	2	2	8 (20%)
Understand	5	5	4	2	16 (40%)
Apply	1	1	2	2	6 (15%)
Analyze	2	2	2	2	8 (20%)
Evaluate	--	1	--	1	2 (5%)
Create	--	--	--	--	--
Total	10 (25%)	11 (27.5%)	10 (25%)	9 (22.5%)	

Lower Order Learning

Higher Order Learning

Test Question Matrix –*Example 2*

30 questions – *Leaning to higher cognitive order*

	Topic A	Topic B	Topic C	Topic D	
Remember	1	1	1	1	4 (13.3%)
Understand	1	1	1	1	4 (13.3%)
Apply	3	2	3	2	10 (33.3%)
Analyze	2	3	2	3	10 (33.3%)
Evaluate	--	1	--	1	2 (6.7%)
Create	--	--	--	--	--
Total	7 (23.3%)	8 (26.7%)	7 (23.3%)	8 (26.7%)	

Lower Order Learning

Higher Order Learning

Practical suggestions

Do not write the test in one day

Write 1or 2 test items after class on note cards, or dictate a memo on your phone

General rule: each item should stand on its own; it should not depend on the answer to another question

Information in one item should not provide clues to another

Easiest test items written with definitive right and wrong answers

Suggestions (cont'd)

Have students write items as an assignment.

Grouping items under headings will improve student performance.

Ask a colleague to read over your test items to help ensure validity.

Publisher-provided test items – read each one before using it.

Listen to students' critiques of your questions.

After the exam…

Test Analyses

- Discrimination index
- **Distractor Evaluation**
- Frequency response
- **Point-Biserial correlation (PBS)**

Thank you for attending and participating!

Questions or follow up

Email -- dhubert@nd.edu

Office -- 350 DeBartolo Hall (in Faculty Lounge)
Phone -- 631-9148

Additional MC item examples for various cognitive thinking levels

Table 1. Item shells for different cognitive processes. Adapted from Haladyna (1997, 2004) to correspond with Crowe et al. (2008).

Cognitive process	Item shell	Example
Lower order cognitive skills		
Remember	• What is the best definition of __? • What is the best word to use for __?	What is the best definition of net primary productivity?
Understand/Comprehend	• What is an example of __? • What is the difference between __ and __?	What is an example of density-dependent regulation of population growth?
Higher order cognitive skills		
Apply	• Predict what happens when __. • What is a consequence of __? • Given __, what is the cause of __? • Solve a word problem or apply a formula in a new situation.	Predict what happens to production efficiency of polar bears when a lack of sea ice increases energy expenditure
Analyze	• To test the hypothesis __, a scientist would need to __. • What is the best conclusion for what is shown in this graph of __?	What would be the best way to test the effect of nitrogen on a phyto-plankton species in a local pond?
Evaluate	• What is the most effective/ appropriate for __? • What is the best [solution] for __?	An experiment finds that metapopulation dynamics of a butterfly is absent when habitat loss has occurred. What is the best explanation for this result?

Rauschert, E. S. J., S. Yang, and R. M. Pigg. 2019. Which of the Following Is True: We Can Write Better Multiple Choice Questions. Bull Ecol Soc Am 100(1):e01468. https://doi.org/10.1002/bes2.1468

Analysis	Breaking material down into its component parts to see inter relationships/hierarchy of ideas	*It is said that when organic chemist Auguste Kikuli was struggling with how the six carbon atoms of benzene linked together, he dreamt of molecules twirling and turning around like snakes. In his dream, one of the snakes swallowed its own tail and rolled around like a hoop. When Kikuli woke up, he realized that the six carbon atoms of benzene were attached to each other to form a ring. Further work showed that this was correct.* **Which phase of the creative process is illustrated by this example?** a) Preparation b) Incubation c) Orientation d) Illumination* **A male patient presenting to his doctor's office with polyuria, polydipsia, and polyphagia would best be treated with:** a) metformin b) aspirin c) sulfonylurea d) atorvastatin (Student needs to realize what disease state the patient is suffering from, then make a treatment recommendation based on that)
Synthesis	Producing something new or original from component parts	**Note:** There is debate over whether MCQ's can be used to test the synthesis level. That is because by its nature the synthesis level requires students to produce new or original material to answer the question. As MCQ's provide answer options there is no opportunity for something 'new' or 'original' to be developed as the answer will be within the provided list. It is therefore recommended that testing of the synthesis level is best achieved through short answer or essay type questions as they require the student to 'produce something new'.
Evaluation	Making a judgment based on a pre-established set of criteria	**Disregarding the relative feasibility of the following procedures, which of these lines of research is likely to provide us with the most valid and direct evidence as to evolutionary relations among different species?** a) Analysis of the chemistry of stored food in female gametes b) Analysis of the form of the Krebs cycle c) Observation of the form and arrangement of the endoplasmic reticulum d) Comparison of details of the molecular structure of DNA*

http://uhhcopfacultyresource.weebly.com/uploads/2/1/9/8/2198211/multiple_choice_and_true_false_exam_question_design_booklet.pdf Retrieved 9/23/2019