Lunch with EDNA

Multiple Choice Questions for Applying, Analyzing, and Evaluating Student Learning

Hosted by **Anita Parker**, Lead Educational Developer, Centre for Teaching and Learning, University of Alberta (aparker@ualberta.ca)
September 20th, 12:05 to 12:50 pm

Join us for Lunch with EDNA where Anita will set the table with the merits and limitations of traditional multiple-choice questions (MCQ), then dish out examples of MCQ requiring students to go beyond remembering to apply, analyze, and evaluate. If writing MCQ is on your Whine List, come join us for Lunch!

Assessment Matters

Assessment of student learning in higher education is intrinsic to the teaching and learning process. It is a way for students to engage meaningfully with course content. Effective assessments align with learning goals and provide direct evidence of student achievement. Assessments can evoke an array of negative emotions, such as fear, dread, anxiety, and intimidation, or even positive emotions, such as confidence and optimism (Parker et al., 2021).

Traditional Assessments

Traditional exams (closed-book, in-person, and invigilated) have deep roots in higher education. These exams can include written response questions, but they primarily consist of multiple choice questions (MCQs).

Traditional exams with MCQs do offer a number of advantages.

- A large number of concepts can be covered in one sitting.
- Having only one correct answer promotes objectivity.
- Ease of administration.
- Ease of marking.
- Invigilation can discourage cheating.

However, traditional exams with MCQs have drawbacks.

- Students prepare by memorizing information; they opt to cram, focus only on surface learning, and data-dump on test day.
- Students' test anxiety can affect their performance.
- Cheating still occurs.
- Authenticity is lacking. "As is often pointed out, few students end up with jobs where they

get paid to fill out multiple-choice test bubble sheets" (Frey et al., 2012).

Alternative Assessments

Alternative (authentic) assessments can take the form of oral presentations, essays, performances, group projects, peer-assessment, journals, self-assessment, field experiences, oral exams, and portfolios, wherein students construct their own responses rather than select from ones presented.

An **open-book exam**, although similar to a traditional exam in terms of structure and purpose, may also be considered an alternative assessment. In an open-book format, students are permitted access to resources during an exam. This format (in-person or online, invigilated or not) offers many advantages for cultivating and measuring student learning.

1. The format is arguably more authentic because students prepare, access information, think and problem solve in ways that are similar to their real-world, future careers.
2. It can still be summative and rigorous; it can still be invigilated to discourage cheating.
3. It can positively impact student emotions by reducing anxiety and increasing confidence.
4. Students must adequately prepare because memorization will not provide them with the required level of conceptual understanding.
5. If students collaborate during the exam (whether permitted or not), they are still learning.

Unless instructors are primarily interested in assessing memory of knowledge and skills, closed book exams are often misaligned with higher-level and critical thinking objectives. Open-book multiple choice questions with formative feedback can be:

- aligned with higher-level thinking objectives;
- rigorous in terms of item statistics; and,
- objectively scored with meaningful feedback for remediating misconceptions in students.

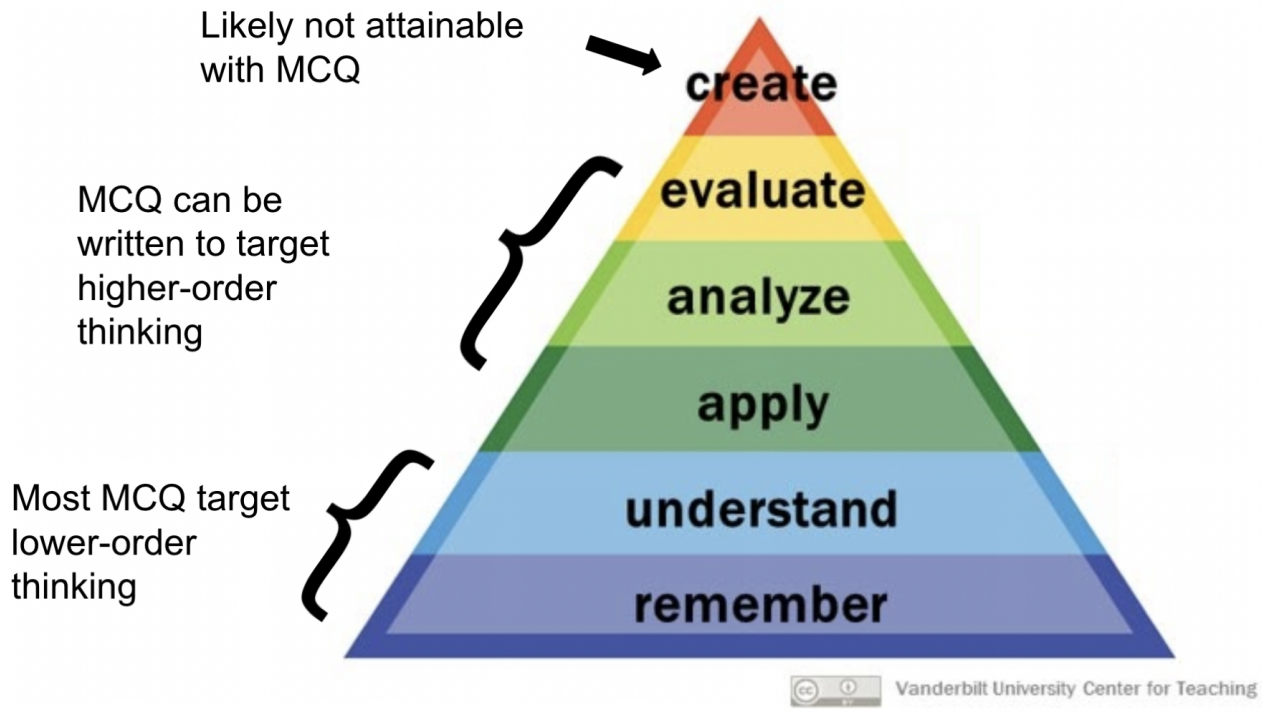
Myth: Open book exams are easy for students because all they do is look up the answers in the book.
Fact: if you test for memory, then yes this is what happens. But if you test for higher-level thinking such as analysis and evaluation, then no.

Myth: Multiple choice questions cannot test higher-level thinking.
Fact: It all depends on how you construct the question. Avoid test-wiseness and simple memory recall.

Myth: Students will cheat.
Fact: Students can, in principle, cheat on any assessment. More importantly, is the assessment one that will challenge them and test the objectives? If so, then, in principle, even if students worked together, they would be learning. This is the point of the objectives.

Source:
Open-Book Higher-Level Thinking Multiple Choice Assessments (20:23)
2021 Festival of Teaching and Learning, University of Alberta
Jacqueline P. Leighton (Faculty of Education)

Bloom's Taxonomy: Apply, Analyze, Evaluate



Traditional MCQs usually reside at the lower two levels—remember and understand. These questions are limited to basic knowledge and common problems and can be easily located in a textbook or searched for online. Such questions are not suitable for an open-book format.

Alternative MCQs that are suitable for open-book are carefully constructed to target higher Bloom's levels of apply, analyze, and evaluate.

Illustrative Example Moving up Bloom's Taxonomy

Remember

At the Remember level of Bloom's taxonomy, students recognize or recall concepts. Consider this sample MCQ at the Remember level (lowest level).

Which of the following lists the stages of the creative process in the correct order?

A. orientation, preparation, incubation, illumination, verification*
B. orientation, incubation, preparation, illumination, verification
C. preparation, orientation, incubation, illumination, verification
D. preparation, orientation, incubation, verification, illumination

Learning outcome: Students will be able to **list** the stages of the creative process.
In this question, students themselves aren't listing the stages, but they are choosing from a few list options. All that is required to answer this question is to recall the order of certain pieces of related information. This concept is easily searched online or in a paper or electronic document, and it could likely be answered without studying or thinking too hard.
This question is not good for an open-book format.

Understand

At the Understand level of Bloom's taxonomy, students explain concepts.
Consider this sample MCQ at the Understand level.

Which of the following describes what takes place in the PREPARATION stage of the creative process, as applied to the solution of a particular problem?

A. The problem is identified and defined
B. All available information about the problem is collected*
C. An attempt is made to see if the proposed solution to the problem is acceptable
D. The person sets the problem aside and gets involved with an unrelated activity

Learning outcome: Students will be able to **describe** the stages of the creative process.
In this question, students themselves aren't describing the preparation stage, but they are choosing from a list of descriptions. Students are required to understand the meaning of preparation; the options are not exact definitions. This concept is easily searched online or in a paper or electronic document, and it could likely be answered without studying or thinking too hard.
This question is not good for an open-book format.

Apply / Analyze

At the Apply and Analyze levels of Bloom's taxonomy, students use information in new situations and draw connections between ideas.
Consider this sample MCQ at the Apply and Analyze levels.

"The story is told of the famous German organic chemist Auguste Kékulé who was struggling with the problem of how the six carbon atoms of benzene were linked together. He was getting nowhere with the problem, and one day fell asleep in front of the fireplace while he was pondering on it. He dreamt of molecules twisting and turning around like snakes. Suddenly, one of the snakes swallowed its own tail and rolled around like a hoop. Kékulé woke up with a start, and realized that his problem could be solved if the six carbon atoms of benzene were attached to each other to form a ring. Further work showed that this was entirely correct."

The above passage illustrates a particular phase of the creative process. Which one is it?

A. Preparation
B. Incubation
C. Orientation
D. Illumination*
E. Verification

Learning outcome: Students will be able to **analyze** and **categorize** a situation in the context of the stages of the creative process.
In this question, students read and analyze a story and apply what they know about the concept (the creative process) to categorize this story into a certain stage (illumination). The story does not explicitly state that the scientist is experiencing the illumination stage. The question does not ask students to define or describe the illumination stage.
This IS a good question for an open-book format.

This question could be made even better with "because" statements.
The above passage illustrates the ___(1)___ stage because ___(2)___. Which row identifies (1) and (2) to correctly complete the statement?

a. *(1) incubation; (2) [correct reason]*
b. *(1) illumination; (2) [correct reason]**
c. *(1) incubation; (2) [incorrect reason]*
d. *(1) illumination; (2) [incorrect reason]*

Evaluate

At the Evaluate level of Bloom's taxonomy, students justify a stand or position.
Consider this sample MCQ at the Evaluate level.

A student was asked the following question: Describe the particular stage of the creative process illustrated by the reading above. As an answer, the student wrote the following:

"The story of the German organic chemist Auguste Kékulé is an illustration of the illumination stage of the creative process. The illumination phase follows a period of incubation and is a time when one one experiences an "aha" moment; seemingly out of nowhere, the solution to a problem presents itself. Kékulé had previously been struggling with his problem and was now relaxed (sleeping), which indicates that the preparation stage had ended (or was on hold), and he was taking his focus off the problem (i.e. letting it incubate). Kékulé spontaneously realized a potential solution to his problem, and he could now move into the verification stage to determine if his idea was correct."

How would you judge this student's answer?

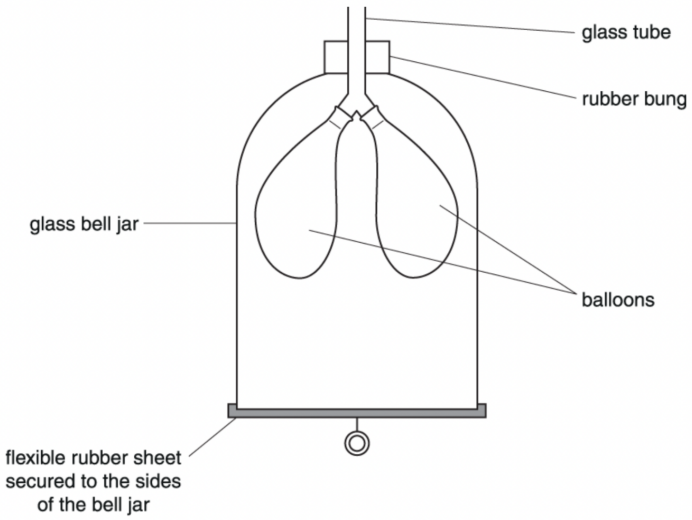
A. Excellent: Student provides a correct answer and demonstrates a thorough understanding of the concept*
B. Good: Student provides a correct answer, but the response contains an error that may indicate a minor misunderstanding of the concept
C. Mediocre: Student provides a correct answer, but the response contains error(s) that may indicate a misunderstanding of the concept
D. Unacceptable: Student provides an incorrect answer, and the response contains error(s) that indicate a significant misunderstanding of the concept

Learning outcome: This question could align with several learning outcomes with students expressing and justifying their thoughts.
The student is required to first read the story. Then the student reads a sample response. Finally, the student evaluates the quality of the sample response. The actual question is, "How would you judge this student's answer? Excellent, good, mediocre, or unacceptable?" Essentially, this is a peer review situation, and the options are a rubric. This particular question provides an "excellent" sample of student work, but it could be altered to other rubric levels. This IS a good question for an open-book format.

Notice that higher-order MCQs are longer, requiring more writing by the instructor and more reading by the student. To encourage students to take more time with a longer reading, consider linking one reading to several questions (and therefore to a larger percentage of the exam mark). Questions with longer readings, such as cases, can be challenging to write because they may be open-ended and of a choose-the-best-answer nature. Consider writing these questions collaboratively with a colleague or have a trusted peer read over your work and offer feedback.

More Sample Higher-Order MCQs

Understand, Apply, Analyze
Context: Biology 20



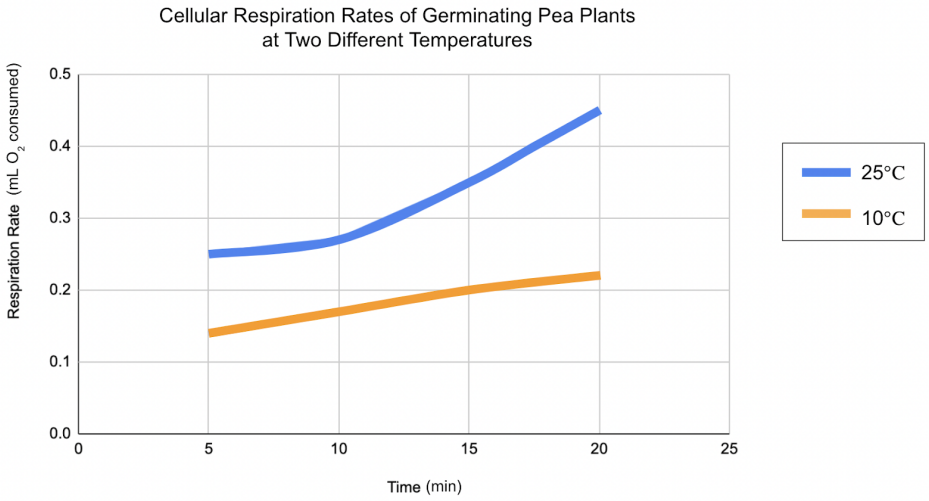
The mechanism of human breathing can be demonstrated using a glass bell jar that has an open top (air can move in through the glass tube) and a flexible rubber sheet for a bottom. (It does not have a glass bottom.) Two deflated balloons represent the right and left lungs. This model is commonly found in biology classrooms.

When a biology student pulled downward on the flexible rubber sheet, the two balloons filled with air. Which of the following statements **best** describes why the balloons filled with air?

a. *Pulling on the rubber sheet increased the volume of the jar and caused air to move passively along a concentration gradient.*
b. Pulling on the rubber sheet disrupted the oxygen molecules in the jar and in the balloons, increasing their kinetic energy.
c. The student was likely blowing carbon dioxide into the glass tube at the same time she was pulling on the rubber sheet.
d. There must have been a hole in the rubber sheet, because the balloons should not have filled with air.

Notes:
Students could be asked to predict what will happen when the rubber sheet is pulled downward. What if the balloons didn't inflate? What could have gone wrong (such as a collapsed lung)?

Understand, Apply, Analyze
Context: Biology 20



In a lab experiment, students measured the respiration rate of germinating pea plants at two different temperatures, 10℃ (cold) and 25℃ (warm). Considering the above graph of their experimental data, which of the following statements is the **most accurate** interpretation?

a. *Respiration rate is positively affected by temperature.*
b. Temperature only affects oxygen consumption between 15 and 20 minutes
c. Germinating plants thrive in cooler temperatures.
d. Warmer temperatures are required for energy production.

Notes
Laboratory science is the ideal scenario for higher order questions. Students can be provided with a short case about what a student did in a lab and the data they obtained. The question can be based on analyzing and interpreting the data.

Understand
Context: Biology 30

The human brain receives a non-stop flow of information from the body's sensory receptors 24 hours a day. For example, sensory receptors in your skin are continually detecting the presence of the clothes you are wearing.

What concept **best explains** why you don't notice your clothing touching your skin most of the time?

a. *Sensory adaptation, because the brain is able to ignore redundant, unimportant messages*
b. Sensory adaptation, because receptors have the ability to turn on and off as needed
c. Accommodation, because the brain is able to focus on important messages only
d. Accommodation, because receptors can ignore unimportant stimuli

Notes
This question is lower level; however, it showcases a template from which higher order questions can be built:

Stem: A story, case, or data table.
Question: Which of the following statements best explains [the concept(s)]
Options:
a. [Correct answer], because [correct reason].
b. [Correct answer], because [incorrect reason].
c. [Incorrect answer], because [correct reason].
d. [Incorrect answer], because [incorrect reason].

Understand, Apply, Analyze
Context: Microbiology 2XX - excerpt from a paper discussed in class

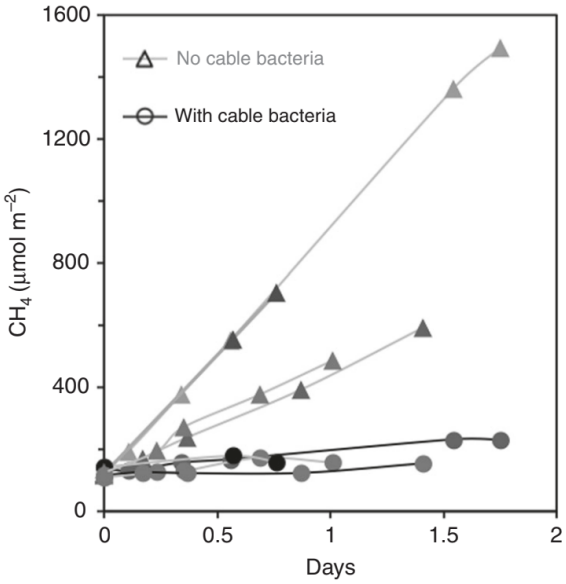


Fig. 3 CH_4 emissions. The emitted amount of CH_4 from replicate rice pots with cable bacteria (circles) and without cable bacteria (triangles) normalized to the surface area of the pots as a function of time.

Which of the following statements **best** explains Figure 3 and why?

a. Correct statement and correct explanation
b. Mostly correct statement but incorrect/plausible explanation
c. Incorrect statement but correct explanation
d. Incorrect statement and incorrect/plausible explanation

Source: Scholz, V. V., Meckenstock, R. U., Nielsen, L. P., & Risgaard-Petersen, N. (2020). Cable bacteria reduce methane emissions from rice-vegetated soils. *Nature Communications*, *11*(1), 1878.

Understand, Apply
Context: Microbiology 2XX lab report question

During this lab, you processed two x 1g meat samples: one from sample A and one from sample B for total counts.
For all meat samples, duplicate 500uL from the 10-4 were plated and incubated. The resulting counts were: 60 and 70 CFU for meat B and too few to count for meat A.

Calculate the average CFU per gram for meat sample B, assuming 1 g is equivalent to 1mL and there are 1,000,000 cells/mL.

1.3×10^{12} CFU/g
1.3×10^{5} CFU/g
6.5×10^{12} CFU/g
6.5×10^{5} CFU/g

Notes:
Calculation questions are, by nature, higher order. Students must understand the underlying concepts, laboratory procedures, and formulas and be able to use the formulas for new situations.
The distractors represent common calculation errors. (For example, 1,000,000 cells/mL is not needed for this question.)
Idea for an Evaluate level question: Students read steps for completing a lab, and then evaluate these steps as to whether or not they would lead to the desired outcome.

Context: Dr. Leighton's EDPY 3XX course

During a class discussion, we highlighted Bronfenbrenner's (1979) **bioecological model** as showing the transactional nature of multiple levels of influence on human development. The model conceptualizes adolescents as nested within four levels. The influence of the family is situated within the **microsystem** level, along with peers, school, religious institutions, and health services.

Does this mean that the influence of the family on the adolescent is equivalent to the influence of peers, school, religious institutions, and health services on the adolescent?

a. Yes. All the variables in the microsystem would have the same influence on the adolescent. This is a strength of the bioecological model.
b. No. All the variables in the microsystem would not have the same influence because the family is the most important variable to adolescent development, as compared to peers, school, religious institutions, and health services. This is a weakness of the bioecological model because it does not distinguish among variables in the microsystem.
c. Yes. All the variables in the microsystem would have the same influence on the adolescent. The fact that these variables are not distinguished indicates that, in principle, they all have the same impact on the teen.
d. No. All the variables in the microsystem would not have the same influence because the family is the most proximal variable to the adolescent. The adolescent is embedded immediately within the family. This is a weakness of the bioecological model because it does not distinguish among variables in the microsystem.*

Notes:
From Dr. Leighton: *This is the best answer. It shows a critical understanding of the model in light of the data presented in class about the proximity of the family influencing adolescent responses and behaviour via identity and autonomy.

Source:
Open-Book Higher-Level Thinking Multiple Choice Assessments (20:23)
2021 Festival of Teaching and Learning, University of Alberta
Jacqueline P. Leighton (Faculty of Education)

Using excerpts from readings and class discussions is an excellent way to make sure questions are not easily Googled and students have been engaged with the course regularly.

Other higher-order tasks that can be built into MCQs:

- Interpret the meaning of an article excerpt.
- Apply previously acquired knowledge to a new situation.
- Analyze text, such as poetry, and identify patterns and relationships.
- Predict the effect of something known on a new situation; explain your prediction.
- Evaluate the steps of a plan toward accomplishing a goal.
- Evaluate the validity of an argument.

Question for Discussion and Sharing

As Educational Developers and Instructors, we may find ourselves creating MCQs for exams. From what we have covered in this session, and from your own experience, what are some tips-and-tricks, steps, considerations, helpful hints, concerns, etc. on this endeavour?

Tips and tricks

- Avoid negatives in the stem; or, if needed, then bold and italicize.
- Avoid using, what we used to call, multiple multiples.
- Remove extraneous information/context as it can be confusing or give answers to other questions.
- Keep the length of answers option similar.
- Don't have more than 4 answers.
- Avoid a 3 - 1 split among the alternatives - (A 3-1 split is when the distractor contains a significant word or phrase that does not appear in the keyed answer).
- Avoid items with a keyed answer that is much less or more wordy than the distractors.
- I use case studies scenarios and then am able to ask 4-5 questions related to the scenario. The case study may be lengthy, but the questions are usually “shorter.”
- When teaching math, creating answers which hold possible calculation results along the way to the final answer will test higher level understanding.
- Would be great to have a quality control system...test out questions prior with other colleagues or students in another class/level in the program.
- The questions being asked need to be connected to the course outcomes; the importance of Learning outcomes to measure using the right domain.
- The number one suggestion I would give instructors is to design the item to measure the learning outcome or objective (match the verb to the assessment).
- Do an item analysis.
- Make sure the distractors are plausible.
- I struggle with constructing MC questions. I have found ChatGPT very helpful.
- Another tip I would give is to write an MCQ that asks students to answer "how" or "why" type questions. This asks students to truly understand the concept.

Considerations

- Don't provide blatant obvious correct answers. The students should be able analyze the answers.
- Keep the language as accessible as possible for ELL and neurodivergent learners.
- I liked the point of if you make a student read a lot of text, to then ask multiple questions about that text. Never thought about that before.
- Try to get students to obtain higher levels of learning through MCQs, not basic regurgitative info.
- Do not use ambiguous and idiomatic expressions that are an obstacle to understanding, particularly for English Language Learners.
- We want to challenge students, but finding the appropriate level of application could be difficult.
- Students need knowledge to work with to use with higher-level thinking processes. The first step is teaching the content, the second step is teaching students how to think.
- Reduce the number of questions on your exams as you move up Bloom's.
- It likely takes more time to create the questions but fewer questions will be required.
- Not exactly multiple choice, but having students write MCQ can access the "creative" realm while sometimes giving them an appreciation of how difficult they can be to write!
- Someone else said it, but creating multiple choice questions in a collaborative environment can lead to better peer review in your test writing process.
- I find that some students are not used to higher order types of MCQs. We can model how to approach these questions during "lecture" time.

Questions, concerns

- My students are mature adults and have test anxiety. Any tips?
- Is there a staff member at U of A CTL who would provide feedback on multiple choice questions once we develop them? *Yes, if you are a U of A instructor, you can request a consultation at ctl.ualberta.ca. If you are an instructor at another EDNA institution, contact your teaching centre.*
- When you had the students evaluate the exam student's answer and applied the rubric, could their answer not be subjective?
- MC can depend on strong reading skills. How to avoid disadvantagering learners who are not as skilled here, particularly when it's not germane to the outcomes?
- I cover a lot of material. Any tips to cone it down a bit for MCQ's?
- Would it make sense to make the questions less text-heavy, as these text intensive questions could create an unnecessary barrier for students for whom English is not the home language?
- Do you have any suggestions on how to iterate term after term?
- Do you know of any digital auto-item writing tool if you input the content, context, and verb?
- Thinking about timing for these higher-order questions, do you have a formula for calculating the time it would take to answer one?
- I love this concept!!!.....now how can I apply this to an Introductory Financial Accounting course??!!

References

Clegg, V. L., & Cashin, W. E. (1986). Improving multiple-choice tests. *Center for Faculty Evaluation and Development, Kansas State University*.

Frey, B. B., Schmitt, V. L., & Allen, J. P. (2012). Defining authentic classroom assessment. *Practical Assessment, Research, and Evaluation*, *17*(1), 2.

Mueller, J. (2018). Authentic assessment toolbox. https://jonfmueller.com/toolbox/

Parker, A. M., Watson, E., Dyck, N., & Carey, J. P. (2021). Traditional versus open-book exams in remote course delivery: a narrative review of the literature. *Proceedings of the Canadian Engineering Education Association (CEEA)*. https://ojs.library.queensu.ca/index.php/PCEEA/article/view/14845

Sambell, K., McDowell, L., & Montgomery, C. (2013). *Assessment for learning in higher education.* New York: Routledge.