



ITEM WRITING

Applied Measurement
Professionals, Inc.

Robert C. Shaw, Jr., PhD

Program Director

Introduction

- This presentation is designed to introduce the process of writing items for a professionally produced certification examination





A test item

- measures one unit of content
- contains a stimulus (the question)
- prescribes a particular response form





Those who write items with little measurement error possess

- complex technical job skills
- sophisticated thinking
- deep understanding of candidates' likely responses to items
- awareness of candidates' perceptions about job content
- the ability to write with a high degree of precision

Preparing to Write

- You must link your items to the prescribed
 - task in the detailed content outline
 - cognitive level
- You will write multiple-choice items with four-options



Why multiple choice?

- Dichotomous (right/wrong) scoring encourages measurement precision
- Valid inferences are strongly supported because each item measures one specific bit of content





Why multiple choice?

- The flexible format allows efficient measurement of a variety of objectives
- Examinees cannot bluff their way to receiving credit
 - although they can correctly guess
 - we will talk more about minimizing effective guessing among candidates as a source of measurement error



Item Components

- Stem
 - The statement or question to which candidates respond
 - The stem can also include a chart, table, or graphic
 - The stem should clearly present one problem or idea



Example Stems

- Direct question
 - Which of the following are included among risks of solid-organ transplantation?
- Incomplete statement
 - Solid-organ transplantation includes risks, which include



Among the four options will be the

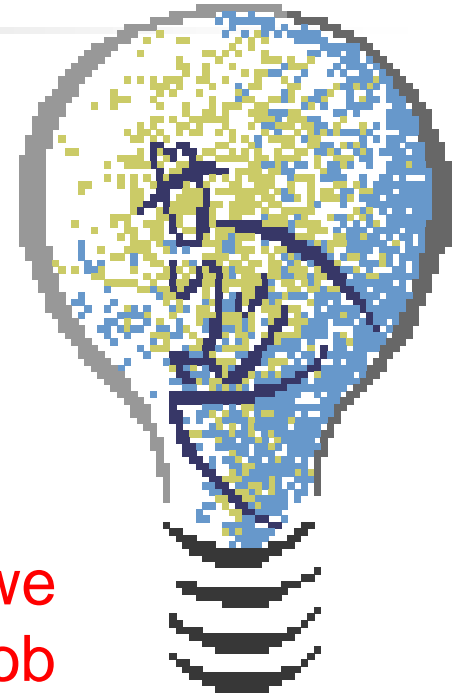
- Key, the *best* or *most appropriate* of available responses
- Distractors are *plausible* yet incorrect responses to the stem



Cognitive levels

- Recall
- Application
- Analysis

Cognitive levels are designated because we recognize that varying dimensions of the job require varying levels of cognition. By linking items to cognitive levels, test scores better represent job-related competencies.





Cognitive levels

- Recall items

- require an effort of rote memorization, so the correct response was previously learned
- are NEVER situationally dependent



Recall item

Which of the following is the most reliable indication of graft rejection in a heart transplant recipient?

- A. fever
- B. ✓ hypotension
- C. angina
- D. weight gain



Recall item

Which of the following solid organs is most commonly transplanted?

- A. heart
- B. lung
- C. ✓ kidney
- D. liver



Cognitive levels

- Application items
 - use interpretation, classification, translation, or recognition of elements and relationships
 - Any item involving manipulations of formulas, no matter how simple, are application level
 - Items using graphics or data tables will be at least at the application level
 - have keys that depend on the situation presented in the stem



Application item

A 34-year-old female with a diagnosis of primary pulmonary hypertension has been referred for lung transplantation. Initial laboratory studies are

T. Bili	1.9	WBC	6.4
Sgot	39	Hgb	17.4
Sgpt	50	Hct	50.1
Alk Phos	91	Plt	180K
Hbsab	+	PA	114/48 mean 74
Hbsag	-	C.O.	3.2



Application item

With which of the following are these results most consistent?

- A. viral hepatitis
- B. portal hypertension
- C. secondary biliary cirrhosis
- D. ✓ passive hepatic congestion



Application item

The following pressures were obtained from a right heart catheterization for a 56-year-old man under evaluation as a heart transplantation candidate:

PA Systolic	56
PA Diastolic	30
PA Mean	39
PCWP	12
Cardiac Output	5.1
Cardiac Index	2.3



Application item

What is the pulmonary vascular resistance in wood units?

- A. 2.3
- B. $\sqrt{5.2}$
- C. 12
- D. 17

When an item requires a calculation, it is best to generate wrong responses using data from the stem.



Cognitive levels

- Analysis items
 - use information synthesis, problem solving, and evaluation of the best response
 - require candidates to find the problem from clues and act toward resolution



Analysis item

A 40-year-old male with alcoholic liver disease and diuretic resistant ascites is listed for liver transplantation. He shows no evidence of encephalopathy, has an ammonia level of 20 $\mu\text{mol/L}$, and a sodium level of 128 mg/dL . Which of the following dietary recommendations from the transplant nurse should best optimize his nutritional status?

- | | <u>sodium</u> | <u>protein</u> |
|------|---------------|----------------|
| A. ✓ | low | unrestricted |
| B. | low | low |
| C. | unrestricted | low |
| D. | Unrestricted | unrestricted |



Analysis item

A transplant nurse takes a call regarding a 25-year-old for whom consent for donation was given. It is 24 hours after admission and the following information is reported:

HR 115 BP 95/60 Temp 39 WBC 18.5

neutrophils-segs	30%	bands	25%
lymphocytes	32%	monocytes	8%
eosinophils	4%	basophils	1%



Analysis item

Which of the following should the transplant nurse recommend before the donation procedure begins?

- A. Repeat laboratory tests.
- B. ✓ Initiate broad-spectrum antibiotics.
- C. Send blood samples for culture.
- D. Call the transplant center medical director.



Tips for writing items to test higher cognitive levels

- Incorporate job-related situations that require analysis of *multiple issues* to arrive at a solution
- *Avoid explicitly identifying the problem* when prompting trouble-shooting



Tips for writing items to test higher cognitive levels

- Give *clues* to the problem and ask for the best course of *action*
 - Which of the following should the transplant nurse do first?
 - Which of the following should the transplant nurse recommend?



Other item formats

- Complex multiple choice (CMC)
 - are best for situations with multiple correct responses
 - may incorporate a direct question or incomplete statement stem format
- will typically present four elements



CMC items

When reviewing pretransplant and posttransplantation needs of the diabetic patient in an educational program for health care professionals, which of the following issues should be included?

Elements

1. the need for recipient pretransplant cardiac evaluation
2. the possible effect of immunosuppression on glucose control
3. requirements for Rh match between the donor and recipient
4. current patient and graft survival statistics

Options

- A. 1, 2, and 3 only
- B. 1, 2, and 4 only
- C. 1, 3, and 4 only
- D. 2, 3, and 4 only



Use an efficient and clear option format

- Write options with similar lengths
 - Novice item writers tend to produce keys that are longer and more detailed than distractors
 - Test wise candidates will be drawn to the longest response



Put as many words as possible into the stem

The psychometrician should recommend

- A. **that the panel write** longer more difficult to read stems.
- B. **that the panel write** distractors of length similar to the key.

The psychometrician should recommend **that the panel write**

- A. longer more difficult to read stems.
- B. distractors of length similar to the key.



Seek balance among options

Undesirable

- A. high blood pressure
- B. low blood pressure
- C. high temperature
- D. low heart rate

Desirable

- A. high blood pressure
- B. low blood pressure
- C. high heart rate
- D. low heart rate



Write in third-person style

- “A transplant nurse is reviewing”
- “A physician is evaluating”
- “A laboratory received”
- “A social worker met with”
 - Specifically avoid pronouns like “you” and “your”



Cause each option to flow from the stem

- If you write an incomplete statement at the end of the stem, then evaluate the grammar of each option when linked to the stem

Native nephrectomy may be required prior to kidney transplantation when a potential recipient has

- A. lupus erythematosus.
- B. anuria.
- C. ✓ ureteral reflux.
- D. Type II diabetes.



CMC items



- You should use elements equally among options
 - Novice writers tend to use correct elements more often
 - Test wise candidates then need only discard the least frequently used elements to find the key



CMC items

- When there are 3 correct elements, use 3 elements in every option

- A. 1, 2, and 3 only
- B. 1, 2, and 4 only
- C. 1, 3, and 4 only
- D. 2, 3, and 4 only



CMC items

- When 2 elements are correct, use each element twice and use 2 elements in every option

A. 1 and 3 only
B. 1 and 4 only
C. 2 and 3 only
D. 2 and 4 only

A. 1 and 2 only
B. 1 and 4 only
C. 2 and 3 only
D. 3 and 4 only

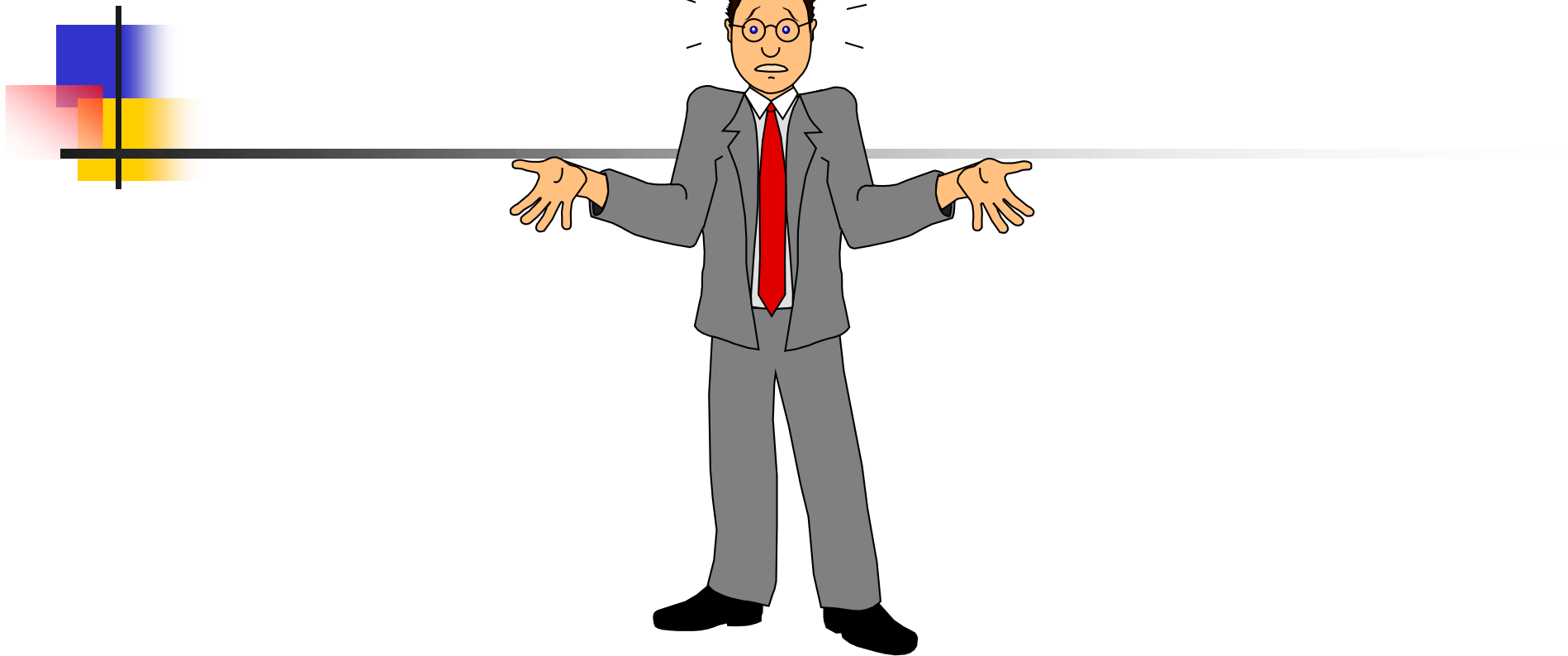
- CMC items with 4 elements, but only two correct responses may be more challenging for candidates



Write distractors with care

- When writing item *stems*, you should do all you can to help candidates clearly understand the situation and the question
- *Distractors* should be written with a more ruthless attitude

Things to avoid





Do not always place item keys in B and C positions

- Empirical studies show
 - writers tend to “hide” keys in one of the inner positions
 - candidates tend to look for correct responses in these same positions



Eliminate irrelevant sources of difficulty

- Stems should be plainly and simply written
- Use a few short sentences rather than longer complex sentences



Eliminate irrelevant sources of difficulty

- Use several small words rather than fewer polysyllabic words to describe the problem and ask the question
- Require relatively simple rather than complex computations with several steps



Avoid stereotypical or prejudicial language

- keep in mind the following potential sources of bias when writing
 - gender
 - ethnicity
 - age
 - religion
 - disabilities



Data ranges should not overlap

A. 36 - 38

B. 34 - 37

C. 32 - 35

D. 30 - 33

Neither of these option sets are acceptably formatted

A. 36 - 38

B. 34 - 36

C. 32 - 34

D. 30 - 32



Do not mix units of measure

- A. 7 days
- B. 15 days
- C. 1 month
- D. 1 year

wrong

- A. 7 days
- B. 15 days
- C. 30 days
- D. 365 days

right



Avoid mixing concepts in options

A. apples

B. oranges

C. bananas

D. hammers

A. heart rate

B. blood pressure

C. temperature

D. antimetabolite



Avoid creating subsets of one option in other options

- A. heart rate
- B. blood pressure
- C. temperature
- D. vital signs



Novice writers often provide clues to the key by

- placing similar phrases in the stem and key, even including identical words
- writing the key in more technical, detailed language



Novice writers often provide clues to the key by using more

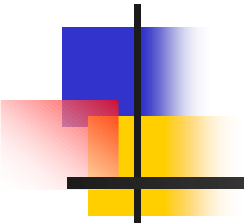
- ambiguous, more easily proven language (e.g., may, sometimes, usually) for keys
- absolute, less easily proven language (e.g., never, none, all, only, always) for distractors



“All of the above” or “None of the above” are lousy options

- An “All of the above” key allows candidates to correctly respond based on partial information
- “None of the above” is a poor distractor because candidates could almost always argue that it was correct
- Candidates’ perceive these options as filler written by someone who could or would not write a quality item

Are you test wise?



You are test wise if you can select the following item keys based on clues given in the item without knowing the content.

There are several *nonsense* items that follow to which test wise candidates could correctly respond if they are skilled at looking for clues from item construction flaws.



Item problem diagnoses

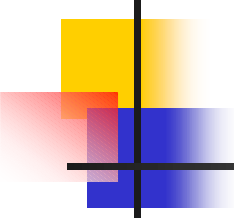
- The following are sample items received by ABTC
- These items illustrate some common construction flaws it would be best for you to avoid



A potential donor from Saudi Arabia who follows the Muslim faith is referred to an OPO. Which of the following Arabic speaking translators should communicate with the non-English speaking family.

- A. male social worker
- B. female nurse
- C. female physician
- D. male physician

■ Testwise candidates will be able to converge on the key by combining the most frequently used elements “male” and “physician.”



A transplant candidate receives an anergy panel for mumps, candida, ATG, and PPD. PPD and mumps results are positive. Which of the following is the best recommendation?

- A. the patient is not a transplant candidate
- B. TB treatment and activate the candidate
- C. TB treatment and reevaluation in six months**
- D. no treatment and activate the candidate.

■ Test wise candidates will


- narrow their choices to parallel options B & C.
- eliminate A & D because of the negative wording they contain.



Heparin, which is the commonly used anticoagulant, acts by

- A. increasing blood platelets.
- B. depressing fibrin formation.**
- C. manufacturing prothrombin.
- D. increasing thrombin production.

- The unnecessary detail given in the stem could cue another item.
- B should begin with the verb *decreasing* to maintain parallel options.
- C should begin with *decreasing* to establish double-parallel options.



A blood group B patient has been on the kidney transplant cadaver list for 6 months. The dialysis physician recently stated the patient's only remaining access site is failing. A blood group D cadaver kidney becomes locally available. The coordinator should understand that

- A. recipient selection is not based on medical urgency.
- B. recipient selection guidelines are flexible to allow for locally defined medical urgency.**
- C. recipient selection prevents group B recipients from receiving group D organs.
- D. recipient selection only follows waiting time guidelines.


- The key contains the cue word 'locally' from the stem.
- D contains the word "only", which is often associated with false statements.
- The phrase 'recipient selection' should be added to the stem to enhance efficiency.



Coordinating the activity and interaction of the recovery teams is the responsibility of

- A. the Organ Recovery Coordinator.
- B. the most senior surgeon.
- C. the host hospital.
- D. no one person.

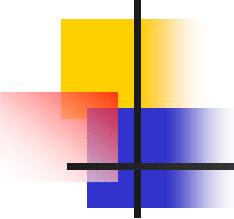
- The key contains two clue words from the stem “coordinate” and “recovery.”
- Distractor D is negatively worded and thus implausible to test wise candidates.
- Distractor C isn’t parallel to the other options since it does not describe a person.



Which of the following will assure adequate vital signs during donor transport to the OR?

- A. Do not worry about vital signs because the trip only takes a few minutes.
- B. Use a portable monitor with a heart rate and BP display.
- C. Stop every five minutes to check vital signs.
- D. None of the above.

- Candidates could argue that D is correct if they can think of a better option than the key.
- Distractor A is incongruent with the stem.



A donor's temperature drops to 34 deg C while in the ICU. The procurement coordinator should apply a Bear Hugger blanket and increase the temperature setting to achieve a normal core body temperature of

- A. 35.4 deg C.
- B. 35.8 deg C.
- C. 36.0 deg C.
- D. 37.0 deg C.

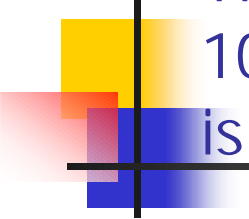
- It is best to offer constant, or nearly constant intervals between values in options.
- The stem unnecessarily increases difficulty with extraneous information.
- The real question is, “*What is normal core body temperature?*”



A normal SaO₂ falls within which of the following ranges?

- A. 95 - 99
- B. 91 - 96
- C. 87 - 91
- D. 80 - 86

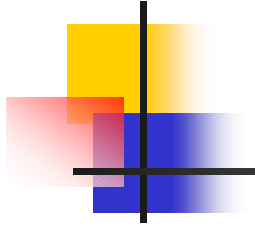
■ Some data ranges overlap and display nonconstant intervals.



The cardiac output of a donor with a heart rate of 100/min and a stroke volume of 50 mL per cardiac cycle is

- A. 5.0 L/min
- B. 0.5 L/min
- C. 2.0 L/min
- D. 50 mL/min

- All options should offer values using the same unit of measure.
- Values should be sorted in ascending order.



Summary





Examination Development & Review

- An examination committee must *unanimously* agree that a newly written item should be retained and subsequently used
- Items should be as polished as the writer can make them



Item submissions

- Submissions may be sent electronically as long as all task and cognitive level information from the top of each submission form is included
 - soneal@goAMP.com