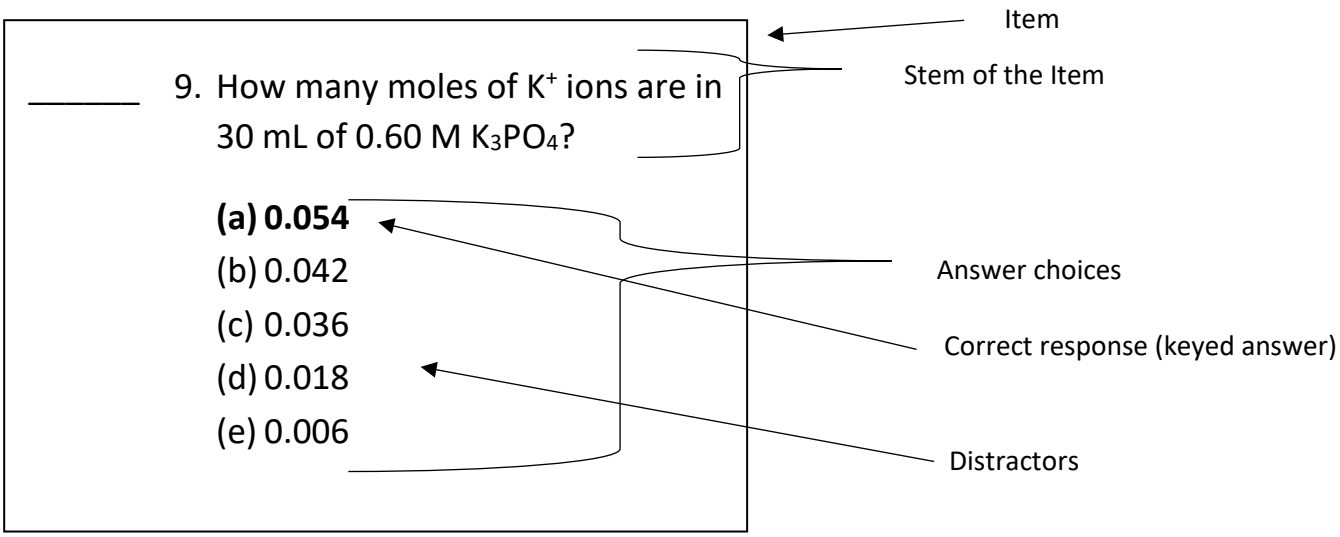
Electronic Supplementary Material (ESI) for Chemistry Education Research and Practice.


# APPENDIX 1
# ITEM WRITING FLAWS EVALUATION INSTRUMENT (IWFEI)
# USEAGE GUIDE

Jared Breakall, Dr. Chris Randles, and Dr. Roy Tasker

# Table of Contents

**Introduction to Multiple Choice Item Format**



**Is the test item clear and succinct?**

- The stem can only be interpreted as having one meaning.
- The stem doesn't include any extra information or wording (**Needed context is appropriate**).
- The answer choices don't include any extra information or wording
- There is clearly only one correct answer choice

Good Example

_____ 1. The atomic weight of silicon is 28.0855. Round this number to 4 significant figures.

(a) 28.0
(b) 28.08
(c) 28.09
(d) 28.086
(e) 28.1

The stem has only one interpretation with no extra information.

The answer choices don't include any extra information and only one answer choice can be interpreted as correct.

Poor Example

_____ 6. It takes 19 days for a particular nuclide to decay 30% of its original activity. What is the half-life of this nuclide?

(a) It would take 0.44 days
(b) It would take 11 days
(c) It would take 16 days
(d) It would take 27 days
(e) It would take 37 days

The stem could be interpreted as decaying from 100% to 70% or as decaying from 100% to 30%. This makes the question unclear.

The answer choices are not as succinct as possible. They include extra information/wording. "It would take" could be removed.

Poor Example

_____ 3. Aspirin is a pain killer that has a density of 1.40 g/cm$^3$. What is the amount (in moles) of aspirin, $C_9H_8O_4$, in a 325 mg tablet that is 100% aspirin?

(a) 0.00180 mol
(b) 0.00325 mol
(c) 0.467 mol
(d) 1.80 mol
(e) 2.80 mol

The density of Aspirin is extra information that is not needed to solve the problem. This introduces student ability to determine needed information as a variable in student performance. The question is no longer just testing the intended chemistry content.

**If the item uses negative phrasing such as "not" or "except", is the negative phrase bolded or capitalized?**

- The words "not" or "except" should be bolded or capitalized if included in the item.
- Avoiding "not" or "except" is ideal in most cases.

Good Example

_____ 1. Which of the following contains a triple bond?

a. ethylene
b. ethane
c. propene
d. benzene
e. propyne

The stem of this question doesn't contain negative phrasing. This is ideal for most items.

Good Example

_____ 1. Which of the following does NOT contain a triple bond?

a. Butyne
b. Pentyne
c. Hexyne
d. Benzene
e. Propyne

The negative phrase is capitalized.

Poor Example

_____ 1. All of the following processes are exothermic except:

a. Combustion of propane
b. Rusting of iron
c. Freezing of water
d. Melting of ice

The word 'except' is not bolded or capitalized.

**If the answer choices are numerical:**

**Are they listed in ascending or descending order?**

- Numerical answer choices should be listed in ascending or descending order. For example: 1,2,3 vs. 2,1,3

Good Example

_____ 12. What is the molality of a solution prepared by mixing 12.0 g benzene ($C_6H_6$) with 38.0 g $CCl_4$?

a. 0.240 *m*
b. 0.316 *m*
c. 0.508 *m*
d. 0.622 *m*
e. 4.05 *m*

The answer choices are written in ascending numerical order.

Poor Example

_____ 12. What is the molality of a solution prepared by mixing 12.0 g benzene ($C_6H_6$) with 38.0 g $CCl_4$?

a. 4.05 *m*
b. 0.240 *m*
c. 0.622 *m*
d. 0.316 *m*
e. 0.508 *m*

The answer choices are not written in ascending or descending numerical order.

This is *Not Applicable* if an item is K-type

Symbolic answer choices, such as electron configurations or chemical formulas are NOT considered numerical.

This criterion would be Not Applicable.

**If the answer choices are verbal:**

**Are the answer choices all approximately the same length?**

- An answer choice should **not** be substantially longer or shorter than any of the other choices. This may cue students to an answer without consideration of the item content.

Good Example

_____ 19. What is the purpose of standardizing a solution?

a. To determine its purity.
b. To determine its concentration.
c. To measure its volume.
d. To determine its molecular formula.
e. To determine the endpoint.

This item keeps all answer choices approximately the same length.

Poor Example

_____ 19. What is the purpose of standardizing a solution?

a. To determine its purity.
b. The purpose is to determine the concentration of the solution
c. To measure its volume.
d. To determine its molecular formula.
e. To determine the endpoint.

One answer choice is significantly longer than the others.

(This item includes phrasing cues as well (see page 8))

This is *Not Applicable* if an item is K-type

Symbolic answer choices, such as electron configurations or chemical formulas are NOT considered verbal.

This criterion would be *Not Applicable* if answer choices are symbolic.

**Does the item avoid "all of the above" as a possible answer choice?**

- Using "all of the above" as an answer choice can cue students to eliminate distractors.

Good Example

_____ 1. Which of the following contains a triple bond?

a. ethylene
b. ethane
c. propene
d. benzene
e. propyne

This item doesn't use all of the above or none of the above as answer choices

Poor Example

_____ 1. Which of the following contains a triple bond?

a. ethylene
b. ethane
c. propene
d. propyne
e. all of the above

The use of 'all of the above' is quickly eliminated when a student recognizes any molecule that doesn't contain a triple bond.

For K-type items, if an answer choice includes all of the possibilities, then it violates this guideline.

**Does the item avoid grammatical and phrasing cues?**

- A cue leads a student to the right answer or to eliminating a distractor.
- A grammatical cue is a difference in grammar between the stem and the answer choices or between answer choices.
- A phrasing cue is where a phrase from the stem is used in one distractor or in the correct answer.

Good Example (Grammatical cuing)

__ 19. Carbon has ____ proton(s).

(a) One
(b) Three
(c) Six
(d) Twelve

This item keeps the grammar of the stem consistent with the answer choices.

Poor Example (Grammatical cuing)

__ 19. Carbon has ____ protons?

(a) One
(b) Three
(c) Six
(d) Twelve

Answer choice A does not fit the grammatical structure of the stem. This may cue students to it being the incorrect answer.

Good Example (Phrasing cues)

__ 19. How many proton(s) does Carbon have?

(a) One
(b) Three
(c) Six
(d) Twelve

This item gives no phrasing cues to the correct answer.

Poor Example (Phrasing cues)

__ 19. How many proton(s) does Carbon have?

(a) One
(b) Three
(c) Six protons
(d) Twelve

This item using a phrase (protons) from the stem in the correct answer choice. This may cue students to choose this answer.

**Could the question be answered without looking at the answer choices?**

- It is important to write the stem of an item in a way that it could be answered without looking at the answer choices. This ensures that the central idea is included in the stem.

Good Example

__ 19. Hydrogen can have how many protons?

(a) 0 or 1
(b) 1
(c) 1 or 2
(d) 2

This item contains the central idea in the stem and can be answered without the answer choices.

Poor Example

__ 19. Hydrogen:

(a) can have 0 or 1 protons
(b) can only have 1 proton
(c) can have 1 or 2 protons
(d) can only have 2 protons

This item cannot be answered without looking at the answer choices. The stem doesn't contain the central idea.

**Does the item avoid complex K-type item format?**

- K-Type items have answer choices that contain combinations of other answer choices.
- K-Type Items have been shown to cue students to the correct answer
- Ordering items, such as the ordering of ion-size, are not considered to be K-type.

Good Example

_____ 12. What is the molality of a solution prepared by mixing 12.0 g benzene ($C_6H_6$) with 38.0 g $CCl_4$?

a. 4.05 *m*
b. 0.240 *m*
c. 0.622 *m*
d. 0.316 *m*
e. 0.508 *m*

This item avoids K-type format

Poor Example

_____ 5. Which of the following properties influence the frequency of a molecular vibration, seen in infrared absorption spectra?

i. Size (radius) of the atoms on each side of the bond
ii. Strength of the bond between atoms
iii. Mass of the atoms on each side of the bond

a. i only
b. ii only
c. iii only
d. i and ii
e. ii and iii

This is an example of a K-type question

Good Example

_____ 17. A double bond is composed of _____ bond(s) and _____ rotate.

a. Two sigma; cannot
b. Two pi; cannot
c. One sigma and one pi; cannot
d. One sigma and one pi; can
e. Two sigma; can

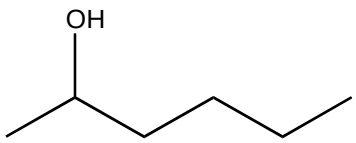
This is a fill-in-the-blank item. This is NOT in k-type format.

**Is this item linked to one or more objectives of the course?**

- Test items should test one or more objectives of the course.

Good Example:

_____ 3. What is the correct formula for the organic molecule shown below?




a. $C_7H_{14}O$

b. $C_6H_{14}O$

c. $C_7H_{13}O$

d. $C_6H_{13}O$

e. $C_4H_{10}O$

Hypothetical Course Objectives

Students Should Be Able to:

1. Interconvert between skeleton structures and chemical formulas.
2. Determine the number of atoms in a molecule based on various representations.
3. Draw Lewis Dot Diagrams from molecular formulas.

This item directly assesses course objective 1 and indirectly assesses objective 2.

Poor Example:

_____ 2. Alkenes by definition contain a _____.

a. C=C bond
b. C≡C bond
c. C–C bond
d. C=H bond
e. C≡H bond

This item doesn't assess any of listed the course objectives.

**Are all answer choices plausible?**

- **All** distractors should be made by using common student errors or misconceptions. Even if only one distractor is not, then the item is in violation of this guideline.

- Each distractor should have been chosen by more than 5% of the students tested.

Good Example

_____ 6. What kind of electromagnetic radiation is able to break bonds?

**a. Ultraviolet**
b. Infrared
c. Visible
d. Microwave
e. Radiowaves

All the answer choices are likely to be chosen. They are all viable forms of electromagnetic radiation

Poor Example

_____ 6. What kind of electromagnetic radiation is able to break bonds?

**a. Ultraviolet**
b. Infrared
c. Visible
d. Microwave
e. The bonds of friendship are too strong to break.

Answer choice E is not plausible.

**Are there six or less thinking steps needed to solve this problem?**

- A thinking step is a small cognitive process that must be taken to solve a problem (Johnstone & El-Banna, 1986).
- The thinking steps should be based on the average student taking the exam.

The following is an example of the thinking steps that may exist in an item. Reproduced from (Johnstone & El-Banna, 1989) with permission from the Royal Society of Chemistry.

'What volume of molar hydrochloric acid would be exactly neutralized by ten grams of chalk?'

Thinking Steps:

1. chalk---calcium carbonate (recall)
2. calcium carbonate =-$CaCO_3$ (recall or deduce)
3. Formula weight of $CaCO_3$= 100 g (calculate)
4. When it reacts with hydrochloric acid, what are the products? (recall)
5. Write a balanced equation (transformation)
6. Recognize that 1 mole $CaCO_3$~ 2 moles HCl (deduce)
7. = 2 litres of molar HCl (recall)
8. 10 g $CaCO_3$ ~ 1/10 mole ~ 1/5 mole HCl (deduce)
9. ~ 1/5 litre molar HCl (recall) = 200 ml molar HCl

Because this item can be viewed as having nine thinking steps, it may be measuring working memory capacity along with the students understanding of chemistry. This negatively effects the validity of the item.

**Does the exam avoid placing three or more items that assess the same concept or skill next to each other?**

- Placing three or more similar questions next to each other may cue students to what the correct answer may be.
- A concept or skill is defined as the same learning objective.

**Does the exam avoid placing three or more difficult items next to each other?**

- A difficult item is defined as an item that you believe less than 50% of students will get correct.

**Is there an approximately even distribution of correct answer choices?**

- Correct answer choices should be approximately evenly distributed. No two distractors should have a difference of greater than two in frequency appearing in the key.

$$Even\ Distribution = \frac{i_t}{a} \pm 1\ (for\ each\ answer\ choice)$$

$i_t$ = Total number of items in the exam
a = Answer choices per item

Good Example

| Answer Key | 4 choices per item |
|---|---|
| 1) A | |
| 2) C | |
| 3) C | |
| 4) D | |

Poor Example

| Answer Key | 4 choices per item |
|---|---|
| 1) A | |
| 2) C | |
| 3) C | |
| 4) C | |

**Does the exam avoid linking performance on one item with performance on others?**

- Items should be independent of one another on an exam.

Good Example

1. What is the molar mass of $C_6H_{12}O_6$?

a) 168.2

b) 180.2

c) 200.2

2. How many moles of water are there in a 14.0 gram sample of $H_2O$?

a) 1.00

b) 0.780

c) 0.550

Students can do well on each item independent of each other.

Poor Example

1. What is the molar mass of $C_6H_{12}O_6$?

a) 168.2

b) 180.2

c) 200.2

2. Based your answer to question one, would a 0.5 mole sample of $C_6H_{12}O_6$ weigh more or less than 90.0 grams?

a) More

b) Less

c) Not enough information to tell

Students **cannot** succeed on item two if they don't succeed on item one.

References

Johnstone, A., & El-Banna, H. (1986). Capacities, demands and processes - a predictive model for science education. *Education in Chemistry*, *23*, 80–84.

Johnstone, A., & El-Banna, H. (1989). Understanding learning difficulties—A predictive research model. *Studies in Higher Education*, *14*(2), 159–168. http://doi.org/10.1080/03075078912331377486