A JMCQ Design Example

1. The prompt will be the desired effect of a correctly applied concept. One can also provide a template or structure as a given (which shortens the code for the multiple-choice options).

   Example: For our example, our prompt can be:

   ```
   Given the following two lines of code:
   ```

   ```
   1    str1 = "Hello"
   2    name = "Rose"
   ```

   ```
   Which choice causes the following printout:

       Hello, my name is Rose.
   ```

2. The three MCQ options are a) the correct code, b) a plausible program containing one of the noted common errors, and c) a plausible program containing another one of the noted common errors. Note that this uses only two of the most common errors. The others can be used for justification distractors. Note that in the actual assessment, the options appear randomized.

   Example: the following are the "cause" (choices) for MCQ.

   a)

   ```
   print(str1 + ", my name is " + name + ".")
   ```

   b)

   ```
   print("{str1}, my name is {name}.")
   ```

   c)

   ```
   print(str1,", my name is", name, ".")
   ```

3. To construct the 8 justification options, we include the following:

   (a) "this is correct" (to be used as the justification for the correct choice);

   (b) "I don't know";

   (c) "other" (which required a comment);

   (d) the (sound) justification for why the first distractor is incorrect;

   (e) the (sound) justification for why the second distractor is incorrect;

(f-h) three plausible but unsound justifications. These three are crafted from the remaining common mistakes noted in step 1.

It is important to note the following caveats:

- The justifications should be written in plain language (as used in class).
- The plausible but unsound justifications are invalid reasons of why an incorrect choice is wrong. As such, the important detail is that the reason of why the choice is wrong must not be expressed in the unsound justification. To craft these, you can express justifications for why other common mistakes are made (and that are not expressed in the MCQ options).
- Only one of the justifications should match each distractor.

For our example, The justifications could be the following (randomized in the actual assessment):

i. "this is correct" (applies to choice a);
ii. "I don't know";
iii. "other";
iv. an `f` is missing before the print string to make it an f-string (applies to choice b);
v. this statement adds unnecessary spaces (applies to choice c);
vi. this statement is missing necessary spaces (does not apply);
vii. this statement adds unnecessary commas (does not apply);
viii. this statement is missing necessary commas (does not apply).

Effort of crafting questions: In practice, each question was created from scratch in 10 to 15 minutes, or by starting from an existing MCQ question, in 5 to 10 minutes.

Grading insights: When grading, a mismatch might hint at a misconception. For example, distractor (v.) may be erroneously picked for option (a) if the student thinks all spaces in the statement affect spaces in the printout.

B Stats per section

course	sem	prof-sec	quiz	treat	G	M	J	p
prog	s24	A1	1	jmcq	76.67	73.33	63.33	0.0030
prog	s24	A1	2	jmcq	70.00	73.33	66.67	0.0370
prog	s24	A1	3	jmcq	46.67	46.67	40.00	0.0110
prog	s24	A1	4	jmcq	63.33	73.33	53.33	0.0000
prog	s24	A1	5	jmcq	76.67	73.33	73.33	0.1900
prog	s24	A1	6	jmcq	56.67	73.33	46.67	0.0100
prog	f24	A1	1	jmcq	90.00	100.00	83.33	0.0560
prog	f24	A1	2	jmcq	90.00	100.00	83.33	0.0250
prog	f24	A1	3	jmcq	56.67	73.33	53.33	0.0230
prog	f24	A1	4	jmcq	86.67	100.00	80.00	0.1810
prog	f24	A1	5	jmcq	100.00	100.00	100.00	0.0570
prog	f24	A1	6	jmcq	83.33	100.00	66.67	0.0020
prog	f24	A2	1	mcq	100.00	100.00	–	–
prog	f24	A2	2	mcq	100.00	100.00	–	–
prog	f24	A2	3	mcq	66.67	66.67	–	–
prog	f24	A2	4	mcq	66.67	66.67	–	–
prog	f24	A2	5	mcq	100.00	100.00	–	–
prog	f24	A2	6	mcq	100.00	100.00	–	–
prog	f24	B3	1	jmcq	93.33	100.00	86.67	0.2350
prog	f24	B3	2	jmcq	80.00	100.00	73.33	0.0590
prog	f24	B3	3	jmcq	63.33	73.33	53.33	0.0510
prog	f24	B3	4	jmcq	76.67	73.33	80.00	0.7970
prog	f24	B3	5	jmcq	93.33	100.00	86.67	0.0900
prog	f24	B3	6	jmcq	63.33	73.33	60.00	0.3970
prog	s25	A1	1	jmcq	80.00	100.00	80.00	0.2580
prog	s25	A1	2	jmcq	93.33	100.00	86.67	0.0020
prog	s25	A1	3	jmcq	76.67	73.33	70.00	0.2060
prog	s25	A1	4	jmcq	93.33	100.00	86.67	0.0210
prog	s25	B2	1	jmcq	76.67	73.33	80.00	0.5280
prog	s25	B2	2	jmcq	76.67	73.33	73.33	0.0560
prog	s25	B2	3	jmcq	76.67	73.33	66.67	0.2720
prog	s25	B2	4	jmcq	70.00	73.33	66.67	0.0460

Table 5: Stats summary (medians) for all programming sections.

course	sem	prof-sec	quiz	treat	G	M	J	p
theory	s24	A1	1	jmcq	86.67	100.00	83.33	0.1130
theory	s24	A1	2	jmcq	90.00	100.00	83.33	0.8520
theory	s24	A1	3	jmcq	100.00	100.00	100.00	0.0990
theory	s24	A1	4	jmcq	76.67	73.33	80.00	0.5520
theory	s24	A1	5	jmcq	93.33	100.00	86.67	0.0070
theory	s24	A1	6	jmcq	65.00	73.33	60.00	0.3180
theory	s24	A1	7	jmcq	76.67	73.33	80.00	0.4020
theory	s24	B2	1	jmcq	70.00	80.00	60.00	0.0620
theory	s24	B2	2	jmcq	70.00	73.33	60.00	0.1190
theory	s24	B2	3	jmcq	86.67	100.00	73.33	0.0820
theory	s24	B2	4	jmcq	50.00	46.67	40.00	0.0250
theory	s24	B2	5	jmcq	75.00	73.33	66.67	0.1040
theory	s24	B2	6	jmcq	40.00	46.67	33.33	0.0250
theory	s24	B2	7	jmcq	46.67	66.67	40.00	0.1330
theory	f24	A1	1	jmcq	86.67	100.00	86.67	0.2760
theory	f24	A1	2	jmcq	90.00	93.33	86.67	0.8290
theory	f24	A1	3	jmcq	100.00	100.00	100.00	0.2510
theory	f24	A1	4	jmcq	90.00	100.00	83.33	0.6930
theory	f24	A1	5	jmcq	78.33	86.67	80.00	0.5220
theory	f24	A1	6	jmcq	58.33	63.33	56.67	0.2740
theory	f24	A1	7	jmcq	76.67	73.33	80.00	0.9050
theory	f24	A1	8	jmcq	85.00	100.00	76.67	0.0670
theory	s25	A1	1	jmcq	93.33	100.00	86.67	0.0660
theory	s25	A1	2	jmcq	93.33	100.00	86.67	0.0160
theory	s25	A1	3	jmcq	100.00	100.00	100.00	0.5730
theory	s25	A1	4	jmcq	86.67	100.00	80.00	0.2700
theory	s25	A1	5	jmcq	80.00	80.00	80.00	0.5780
theory	s25	A1	6	jmcq	53.33	56.67	43.33	0.1220

Table 6: Stats summary (medians) for all theory sections.

C Plots of Scores per section

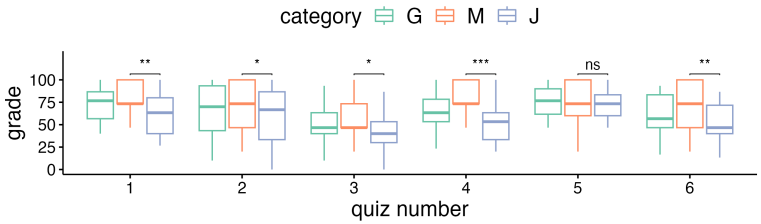

Figure 3: Prog S24 A1: Correctness vs Justification Grades per quiz

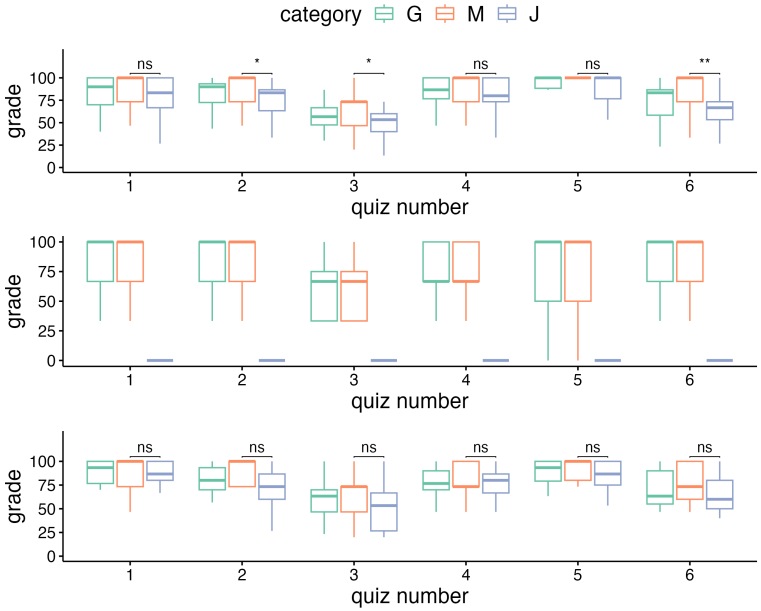

Figure 4: Prog F24 A1 (top), A2 (middle), and B3 (bottom): Correctness vs Justification Grades per quiz

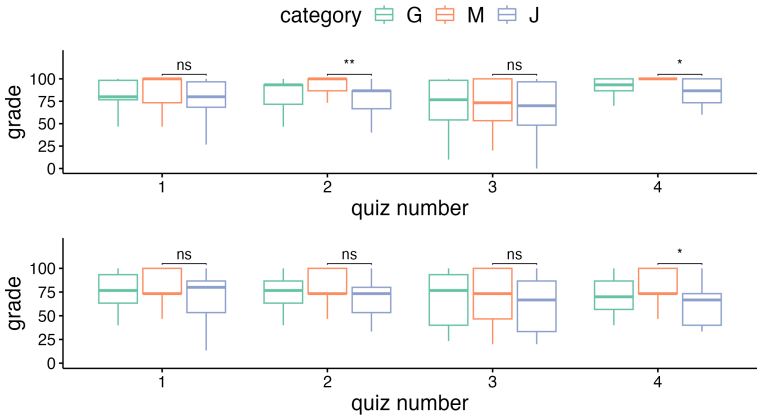

Figure 5: Prog S25 A1 (top) and B2 (bottom): Correctness vs Justification Grades per quiz

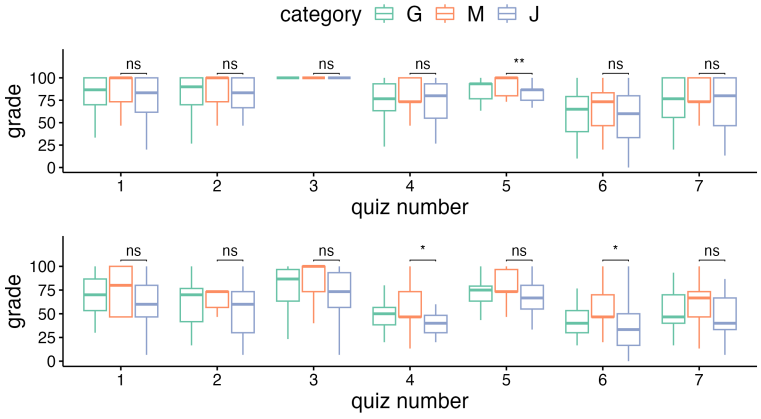

Figure 6: Theory S24 A1 (top) and C2 (bottom): Correctness vs Justification Grades per quiz

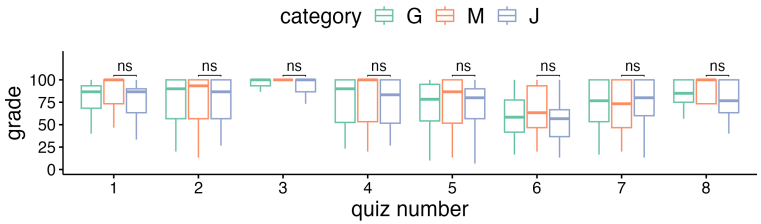

Figure 7: Theory F24 A1: Correctness vs Justification Grades per quiz

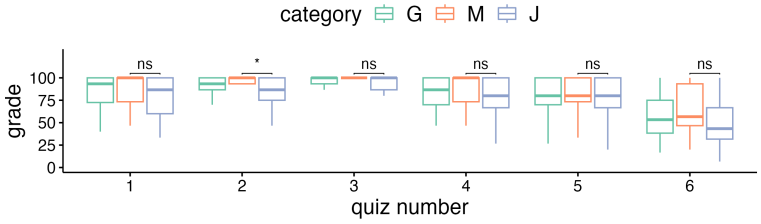

Figure 8: Theory S25 A1: Correctness vs Justification Grades per quiz

quiz	$A1_{M\%}$	$B3_{M\%}$	$A2_{M\%}$	pA1B3	pA1A2	pB3A2
1	83.3	88.3	63.1	1.000	0.013	0.004
2	82.1	78.3	57.1	1.000	0.002	0.041
3	54.4	60.3	29.8	1.000	0.005	0.001
4	82.7	71.4	42.0	0.472	0.000	0.002
5	92.2	90.7	52.2	1.000	0.000	0.000
6	78.2	71.4	57.9	1.000	0.057	1.000

Table 9: Comparison of full correctness per question in the A1, B3 and A2 sections of the programming course in Fall 2024 (First-quiz p-values)

quiz	$A1_{M\%}$	$B3_{M\%}$	$A2_{M\%}$	pA1B3	pA1A2	pB3A2
1	66.7	77.8	63.0	1.000	1.000	1.000
2	85.4	97.0	58.3	0.547	0.033	0.001
3	92.4	100.0	53.3	0.411	0.000	0.000
4	87.2	100.0	62.7	0.159	0.055	0.000
5	72.2	100.0	57.8	0.660	1.000	0.124
6	57.1	100.0	45.8	0.422	1.000	0.159

Table 10: Comparison of full correctness per question in the A1, B3 and A2 sections of the programming course in Fall 2024 (Recovery-quiz p-values)

E Gains between First and Recovery quizzes

sem	sec_prof	quiz	n_students	First	Rec	Gain	sig
s24	A1	1	19	76.1	98.6	29.5	***
s24	A1	2	15	60.9	96.4	58.4	**
s24	A1	3	20	52.0	94.7	82.1	***
s24	A1	4	14	64.3	88.6	37.8	ns
s24	A1	5	9	58.5	91.1	55.7	ns
s24	A1	6	12	57.8	71.1	23.1	ns
f24	A1	1	13	66.9	90.0	34.5	ns
f24	A1	2	14	62.6	84.5	35.0	ns
f24	A1	3	21	52.1	86.0	65.2	***
f24	A1	4	9	67.0	85.6	27.6	ns
f24	A1	5	6	62.8	76.1	21.2	ns
f24	A1	6	7	61.4	64.3	4.7	ns
f24	A2	1	17	68.6	92.2	34.3	ns
f24	A2	2	11	54.5	90.9	66.7	ns
f24	A2	3	19	50.9	86.0	69.0	*
f24	A2	4	15	57.8	95.6	65.4	**
f24	A2	5	9	40.7	81.5	100.0	ns
f24	A2	6	8	58.3	62.5	7.1	ns
f24	B3	1	1	76.7	40.0	-47.8	ns
f24	B3	2	10	63.3	93.0	46.8	*
f24	B3	3	16	55.8	90.8	62.7	***
f24	B3	4	14	71.2	94.3	32.4	***
f24	B3	5	1	76.7	100.0	30.4	ns
f24	B3	6	2	81.7	100.0	22.4	ns
s25	A1	1	14	70.0	91.0	29.9	*
s25	A1	2	8	64.6	72.9	12.9	*
s25	A1	3	16	61.5	79.8	29.8	ns
s25	A1	4	3	63.3	65.6	3.5	*
s25	B2	1	5	54.7	98.7	80.5	*
s25	B2	2	8	65.8	79.2	20.3	ns
s25	B2	3	7	46.7	74.8	60.2	ns
s25	B2	4	2	63.3	90.0	42.1	ns

Table 11: Gains between the First round and recovery quizzes

F Student Responses

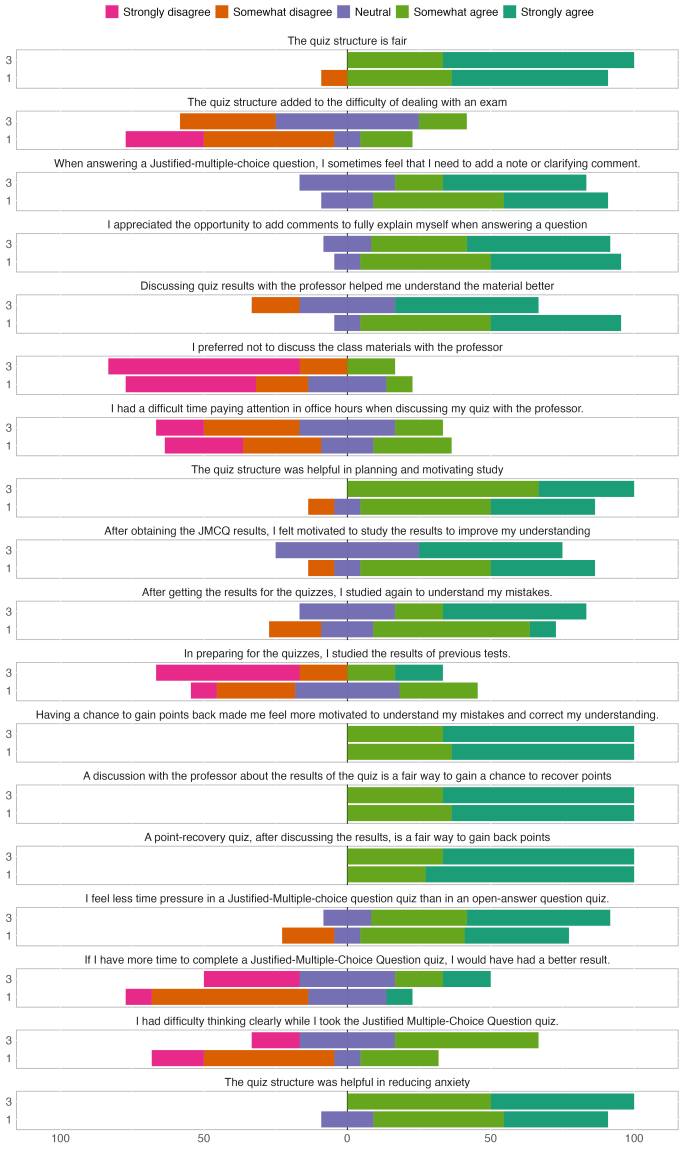

Figure 9: Survey responses for the Programming course in Fall 2024