Can multiple-choice questions replace constructed response test as an exam form in business courses? Evidence from a business school

Professor Leiv Opstad,

Norwegian University of Science and Technology

Business schools Norway

- A traditional 4-hours written school exam decides the letter grade in most business courses

- There is a pressure to substitute the traditional constructed- response (CR) with multiple-choice (MC) questions.

The purpose of this article

- Can Multiple-choice questions replace traditional essay questions?

- Method:

- To compare the students' performance in a business course (Macroeconomics) where one has used both kind of tests.

- This paper does not explore if the students will change learning style by introducing MC-based exam

The reasons for MC format exam are :

- Low administrative costs
  - The score can be determined by a computer instead of two examiners
  - No students will complain or ask for validation of the given grade
- Students are equally treated
- Easy to check the answer and find the score
- Special advantages with large scale courses (for instance with more than 500 students with identical exam, but belonging to gegrahically separate campuses.

But there are some disadvantages:

- It can be easier to cheat (in an exam hall one can notice the anwers of fellow students)
- The students' learning style can change
- The female students underperform compared to males
- Difficult to construct a good test (design of question etc.)

What do we want to measure with the final exam?

- Students' academic skill
- Qualification for further studies or for working marked
- According to Bloom's taxanomy there are 6 educational objectives:
1. Knowledge
2. Comprehension
3. Application
4. Analysis
5. Synthesis
6. Evaluation

According to the litterature:

- It is easy to create MC-based questions from the two first level of Bloom's taxonomy , but it can difficult for the 4 others (Hickson & Reed, 2011).

- Using CR-format exam makes it easier to include higher levels of Bloom's taxonomy

- If one wants to measure sophisticated dimensions of knowledge, there is a substantial difference between MC and CR

An important article

- Walstad and Becker (1994) reported a minor difference between a composite score of the two test methods and a pure MC test. Consequently, due to the high testing cost, one recommended MC-based exam.

The sample

- From a Norwegian business school (period of 5 years)
- The paper exam consists of:
- 75 percent CR-questions
- 25 percent MC-questions

The data

	Min	Max	Mean	St. Dev.	N
MC-score (scale 0 to 100)	13.3	100	60.5	15.4	1190
CR-score (scale 0 to 100)	4.3	100	49.9	19.2	1190
Macroeconomics (grade)	0	5	3.10	1.19	1190
Gender (0:F, 1:M)	0	1	0.41	0.49	932
GPA	44.0	65.4	51.9	3.39	613
Microeconomics (grade)	0	5	3.24	1.20	887
Business mathematics (grade)	0	5	2.85	1.62	887
Management (grade)	0	5	2.88	1.26	927
Compulsory midterm MC-based test (macroe-conomics), 32 questions	8	31	19.6	4.3	911

Note 0:F, 1:E, 2:D, 3:C, 4:B, 5:A

Correlation coefficients for all 5 years

	MC	CR	MC-midterm
MC		0.62	0.43
CR	0.62		0.36
MC-midterm	0.43	0.36	

For each year

	2012	2013	2014	2015	2016
Correlation coefficient	0.50	0.54	0.71	0.66	0.67

The correlation coefficient

- The value depends heavily on the design of question. That might explain the variation between 0.5 and 0.7

- Even with an identical test method, there will be a considerable variation in the students' achievement (Becker 1994)

- Notice the link (correlation coefficient) between two MC-tests (taken at different times) is lower than the connection between MC- and CR- test (taken at the same time) for the same course and group of students.

The gender difference

	Females	Males	Diff.	St. Dev.	T-value	Sign. level
MC	52.7	64.7	-7.4	0.98	-7.5	0.000
CR	48.3	55.6	-7.3	1.22	-6.0	0.000

Linear regression model comparing MC and CR

- Independed variables:

- Gender (0:F, 1:M)

- : GPA score

- : Performance in Business mathematics (0:F, 1:E, 2:D, 3:C, 4:B, 5:A).

: Performance Microeconomics (0:F, 1:E, 2:D, 3:C, 4:B, 5:A)

: Performance introduction course in management (0:F, 1:E, 2:D, 3:C, 4:B, 5:A)

	MC model			CR model		
	Coeffi-cient (B)	T-value	Sig. level	Coeffi-Cient (B)	T-value	Sig. level
Constant	10.9 (8.52)			6.6 (9.73)		
GPA	0.48 (0,11)	1.27	0.20	0.17 (0.19)	0.90	0.366
Gender	5.01 1.12)	2.89	0.004 (***)	2.49 (1.28)	194	0.053 (*)
Business Math	1.45 (0.42)	4.46	0.000 (***)	1.24 (0.48)	2.58	0.010 (**)
Microeconomics	5.66 (0.61)	3.44	0.000 (***)	8.01 (0.70)	11.40	0.000 (***)
Management	0.13 (0.46)	0.27	0.787	2.22 (0.53)	4.20	0.000 (***)
	Adj. R^2=0.376, N=473			Adj. R^2=0.442, N=473		

Similarities

- The influence of Business maths is almost the same
- GPA has no statistical significant effect
- The R –square has value around 0,4 , (but slightly higher for CR)

Differences

- Gender (stronger impact for MC in favour males)
- The impact of microeconomics is relative stronger for the CR-model
- Performance in management is significant positive linked to CR-model, but not to MC-model.

Explanation

- *Performance in microeconomics* is a good indicator for success in macroeconomics, therefore CR capture more the academic skills than MC. On the other hand the reason might be the exam in microeconomics is CR. If the exam in microeconomics was MC-questions, the result might be different.

- Performance in management captures better the writing and presentation ability.

There is a gender gap, why?

- Different attitudes towards maths
- Women in general are more risk-averse than men are
- Sociocultural differences
- Student-specific characteristics
- Others?

Conclusion

- This paper shows that there are some differences between CR- and MC-based tests
- One needs more investigation to determine if MC- can substitute CR
- There are many arguments against excluding essay questions
- May be combination of CR- and MC- is a good solution

Reference

- Hickson, S., Reed, W. R., & Sander, N. 2012. Estimating the effect on grades of using multiple-choice versus constructive-response questions: Data from the classroom. *Educational Assessment, 17*(4), 200–213. https://doi.org/10.1080/10627197.2012.735915

- Walstad, W. B., & Becker, W. E. 1994. Achievement differences on multiple-choice and essay tests in economics. *The American Economic Review, 84*(2), 193–196. https://doi.org/10.2307/1182910