Effect of Multiple Attempts Assumption On Quiz Scores

Team 13

Meet the Team

Amber Wu

Antonio Moral

(Sylar)Jiajian Guo

Yixuan Wang

Bosoo Kim

Manushi Patel

Prior Work

Do Multiple Homework Attempts Increase Student Learning? A Quantitative Study **Kathy K. Archer**

First Published May 18, 2018

Summary

Groups	Count	Sum	Average	Variance
Group A Quiz Grade	406	244.9902	0.603424	0.048917
Group B Quiz Grade	511	361.6115	0.707655	0.04344

ANOVA

Source of variation	*Sum of Squares*	*df*	*Mean Square*	*F*	*p* value	*F* crit
Between groups	2.4579	1	2.4579	53.5910	5.4E-13	3.85164
Within groups	41.9657	915	0.04586			
Total	44.4236	916				

Note. Group A: Microsoft Excel templates with a single graded homework submission. Group B: Multiple homework attempts administered through a web-web based homework management system.

ANOVA Single Factor-Exam Scores Group A Versus Group B.

Archer, Kathy. (2018). Do Multiple Homework Attempts Increase Student Learning? A Quantitative Study. The American Economist.

63. 056943451877479. 10.1177/0569434518774790.

Our Research Question and Hypothesis

Do people get lower scores in their initial attempt when they are given multiple attempts?

Alternative Hypothesis:

People tend to get lower scores in their initial attempt if they are given multiple attempts.

Treatment and Randomization

TREATMENT:

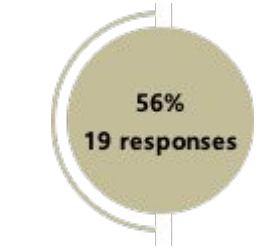
- Target people: Students of the MSBA
- No. of people surveyed: 68 students, 34 people in the treatment and the control.
- For our treatment group we decided to add one line before the survey page: ‘You will have multiple attempts to complete the questions.’ (but only consider their first attempt of the questions)

RANDOMIZATION:

- 50% of people into treatment and 50% into the control.
- Problem with this approach:
  - Did not control for the gender
  - Solution: Block randomization to control for the proportion of the gender

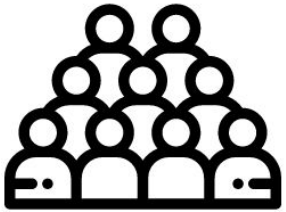
Data Collections and Outcomes

MSBA student list	Total: 68 (Excluding Team-members)
Randomization (50:50)	50% Treatment 50% Control
Distribute Survey	Through email: Opened for 7 days
Reminder email	Reminder: Opened for 2 days
Analyze Data	Closed survey



66% responded

45 results



Treatment

Control

Our Analysis

Outcomes:

- Scores
- Completion time (duration)

Covariates:

- Age
- Gender
- GPA
- Work experience

RANDOMIZATION CHECK

Age, Gender, GPA, Work experience

GRAPHS & CHARTS

Score: treatment vs. control
Duration: treatment vs. control

REGRESSION
DURATION~ TREATMENT

Controlling for age, gender, GPA and work experience

REGRESSION
SCORE ~ TREATMENT

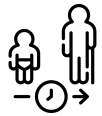
Controlling for age, gender, GPA and work experience

Our Results

Randomization check on pre-treatment Variables

Statistically Similar

Age



	Treatment	Control	P-value
Under 20	0.0384615	0.1052632	0.515617
Between 20-30	0.9230769	0.8947368	1.000000
Over 30	0.0384615	0.0000000	0.000000

Experience

	Treatment	Control	P-value
No Experience	0.3461538	0.3684211	1
Under 2 years	0.4230769	0.4210526	1
Between 2-5	0.1923077	0.2105263	1
Between 6-9+	0.0384615	0.0000000	0

Statistically Different

GPA

	Treatment	Control	P-value
Under 3.0	0.0384615	0.0000000	0.0000000
Between 3.0-3.5	0.1923077	0.5263158	0.0006283
Over 3.5	0.7692308	0.4736842	0.0027657

Gender

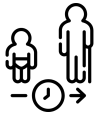
	Treatment	Control	P-value
Male	0.6153846	0.3157895	0.0022497
Female	0.3461538	0.6315789	0.0037487
NonBinary	0.0384615	0.0000000	0.0000000

Our Results

Treatment effects and regression findings

01 **Treatment Effects**

- Treatment group had a lower average score and shorter completion time
- Consistent even when adding variables
- **Not statistically significant**



02 **Age**

- Older demographic tended to score higher and take longer
- **Statistically significant**

03 **GPA**

- Reported GPA seemed to not have an effect on time taken
- Higher GPA tended to lead to higher scores
- **Statistically significant**

04 **Experience**

- Work experience had no significant effect on scores or timing
- **Not statistically significant**

05 **Gender**

- Non-binary gender tended to score higher
- Difficult to assess given the small sample.
- **Statistically significant**

No statistically significant difference between Treatment and Control

Limitations

Potential selection bias of some brain teaser haters aborted halfway.

Non-representative sample of US college students with only MSBA testers.

Excludability violation when some test takers directly google for answers.

Small number list of questions with few than ten leads to a lot noise.

THANK YOU

Q & A