Optimizing Learning Strategies: Evaluating the Impact of Active vs. Passive Learning on Student Engagement and Comprehension

Achinthya Sreedhar, Chaitali Deshmukh, Kendall Sims, Neha Shastri

Introduction

This study examines the effectiveness of active learning versus passive learning in student engagement, perceived preparedness, and material helpfulness. Specifically, it evaluates whether interactive learning materials lead to higher engagement and better perceived learning outcomes compared to traditional passive methods like pre-written notes. The motivation behind this research is to optimize educational practices to enhance student learning experiences. By comparing active and passive learning strategies, this experiment aims to provide evidence-based insights for educators on the most effective teaching approaches. Beyond the classroom, the findings could also inform workplace training programs and marketing strategies that leverage interactive content to boost engagement and retention.

Research Question

Does active learning lead to higher student engagement and perceived preparedness as compared to passive learning?

Null Hypothesis

There is no difference in the effect between active and passive learning: Students who participate in active learning report similar levels of perceived preparedness and classroom engagement as those who participate in passive learning.

Experimental Design

The experiment was randomized at the student level, with surveys distributed one day before the case discussion. To ensure balance across key demographics, we used stratified random assignment with gender-based blocking. Within each gender group, students were randomly assigned to the control or treatment condition, ensuring an even distribution. This approach minimized confounding effects and improved the comparability of outcomes.

Survey Design

The experiment was conducted using Qualtrics, providing a controlled and structured survey process. To ensure participant eligibility, access was restricted to students in our MSBA class and our professor, requiring authentication through Boston University email IDs. This login mechanism prevented external responses and that only eligible participants could access the survey. Randomization was implemented at the student level before survey distribution, with students pre-assigned to either the control group (passive learning) or the treatment group (active learning). Both groups accessed the survey through the same login page, ensuring a consistent initial experience.

I. *Pre-Experiment Survey*
Participants received the StubHub case survey a day before the discussion. Upon logging in, all participants completed the same pre-experiment survey, which collected key covariates, including gender, prior exposure to the StubHub case, self-assessed knowledge in causality, and likelihood of participating in class. After completing these questions, they were directed to their assigned learning experience.

II. *Learning Experience*

- Control Group (Passive Learning): Received pre-written notes summarizing the key points of the case discussion.
- Treatment Group (Active Learning): Engaged with the same content in a quiz format. Students answered multiple-choice questions with immediate feedback:
    - Correct answers were highlighted in green.
    - Incorrect answers were shown in red with a red "X" mark.
    - Questions were scored, adding an incentive for engagement.

III. *Post-Experiment Survey*
After completing their assigned learning tasks, both groups were directed to the same post-experiment survey, gathering self-reported data on understanding, engagement, preparedness, and learning preferences.

IV. *Post-Class Survey*
After the case discussion, students completed a post-class survey assessing their recall of the material, participation in the discussion, and whether the information provided increased their willingness to participate.

Outcome Variables

Our experiment focused on three key outcome variables:

1. *Post_Class_Participation_Val*: A self-reported measure of students' participation in the case discussion after class, rated on a scale from 0 to 100.
2. *Post_Class_Recall_Val*: A self-reported measure of students' ability to recall case study material during the discussion, rated on a scale from 0 to 100.
3. *RQ1_Understand_Material_Val*: A self-reported measure of students' comprehension of the study material before attending class, rated on a scale from 0 to 100.

Power Analysis

The target effect size for our three outcome variables was 10% (10 points on a 100-point scale).
After collecting data, we calculated the standard deviations for each metric and used them to determine Cohen’s d. A subsequent power analysis revealed that our experiment was underpowered, indicating an insufficient sample size to reliably detect the desired effect.

Outcome Variable	Sample Size Used	Sample Size Needed
Post Class Recall	35	128
Post Class Participation	35	183
Pre Class Understanding of Material	55	344

Covariates Considered

1. *Gender (Male, Female, Other):* Used to assess differences in engagement and preparedness across demographic groups.
2. *Prior Exposure to the Case Study (Yes/No):* Students were asked if they had read the StubHub case. Since prior familiarity could impact comprehension and engagement regardless of the learning method, this variable helps control for baseline knowledge differences.
3. *Self-Assessed Knowledge in Causality (Scale: 0-100):* Since the study involves educational material, prior knowledge of causality could influence how much students benefit from the materials. Participants rated their understanding on a 0 to 100 scale, allowing us to account for baseline differences in subject matter expertise.
4. *Self-Assessed Likelihood to Participate in Class (Scale: 0-100):* Since classroom participation varies by individual tendencies, students rated their likelihood of participating on a 0 to 100 scale. This allows us to control for pre-existing engagement levels and isolate the impact of the learning method.

These covariates were selected to ensure that variations in student engagement and preparedness stem from the learning method rather than individual characteristics. By accounting for these factors, we aim to isolate the causal effect of active learning on student outcomes.

Randomization Checks

The Z-test yielded a p-value of 0.1154, which exceeds the 0.05 threshold, confirming that randomization was properly conducted.

	Gender | Read Stubhub Case? | Knowledge of Causality | Likelihood to Participate			
	(1)	(2)	(3)	(4)
Treatment	-0.052	0.150	0.748	2.507
	(0.179)	(0.119)	(5.078)	(7.772)
Observations	35	35	35	35
R^2	0.003	0.037	0.001	0.003
Adjusted R^2	-0.028	0.008	-0.030	-0.027
Residual Std. Error	0.509	0.381	15.506	23.291
F Statistic	0.086	1.585	0.022	0.104
Note:				*p<0.1; **p<0.05; ***p<0.01

The balance check confirms that randomization was properly conducted, as no significant differences exist between the control and treatment groups. The treatment variable does not account for any meaningful variance in the dependent variables, indicating that the groups are well-balanced. This suggests that the randomization process effectively created comparable groups for analysis.

Interpretation of Results

I. *Cohen's d & power of our experiment*

Outcome Variable	Effect Size (Cohen's d)	Power
Post Class Recall	-0.11	0.07
Post Class Participation	0.13	0.08
Understanding of Material	0.24	0.17

The Cohen's d values indicate small effect sizes for all outcome variables, suggesting that the treatment had minimal impact. The largest effect was on Understanding of Material (d = 0.24), though still small, while Post Class Recall (d = -0.11) suggests a negligible or slightly negative effect, and Post-Class Participation (d = 0.13) shows a very small positive impact.

II. *ATE*

Outcome Variable	ATE	P-Value
Post Class Recall	-2.65	0.75
Post Class Participation	4.34	0.71
Understanding of Material	5.03	0.20

All p-values exceed 0.05, indicating no statistically significant ATE.

III. *CATE: Gender*

Outcome Variable	ATE (Male)	ATE (Female)
Post Class Recall	-4.28	-1.49
Post Class Participation	10.78	1.59
Understanding of Material	-9.56	16.51

The Conditional Average Treatment Effect (CATE) shows that the treatment increased participation for male students but did not improve their understanding or recall. In contrast, female students benefited more in understanding, though their participation and recall remained largely unaffected.

IV. *CATE: Whether Student Read the Case*

Outcome Variable	ATE (Read the Case)	ATE (Did Not Read the Case)
Post Class Recall	4.53**	-35
Post Class Participation	-5.05	53.2
Understanding of Material	4.13	-5

*(** - statistically significant p<0.05)*

The Conditional Average Treatment Effect (CATE) by case reading suggests that students who read the case experienced slight improvements in recall and understanding, with minimal impact on participation. However, those who did not read the case showed a sharp decline in recall and understanding but participated more actively in the discussion, possibly to compensate for their lack of preparation.

Regression Analysis

I. *Post Class Recall ~ Treatment * Read StubHub Case*

	Dependent variable: Post Class Material Recall
	(1)
CVQ2_read_stubhub_case_bool	-27.529**
	(12.430)
Intercept	78.000***
	(10.898)
Treatment	-35.000***
	(10.898)
Treatment:CVQ2_read_stubhub_case_bool	39.529***
	(13.957)
Observations	35
R^2	0.156
Adjusted R^2	0.074
Residual Std. Error	23.359 (df=31)
F Statistic	5.150*** (df=3; 31)
Note:	*p<0.1; **p<0.05; ***p<0.01

The regression analysis indicates that the treatment had a significant ($p<0.01$) negative effect on recall performance overall. However, students who read the case beforehand performed better in the treatment group, suggesting prior engagement enhanced the quiz-based intervention's effectiveness. The model explains limited variance ($R^2 = 0.113$), implying other factors influence recall. The treatment's impact was not uniform, potentially disadvantaging unprepared students due to cognitive overload or the quiz structure. These findings emphasize the role of prior preparation in active learning and suggest instructional designs should better support students with varying knowledge levels.

II. *Understanding of Material ~ Treatment * Gender*

	Dependent variable: Understanding of Material
	(1)
Gender[T.Male]	-10.632
	(10.755)
Intercept	73.556***
	(8.970)
Treatment Group	-9.556
	(10.595)
Treatment:Gender[T.Male]	26.061**
	(13.116)
Observations	35
R^2	0.107
Adjusted R^2	0.021
Residual Std. Error	20.575 (df=31)
F Statistic	2.106 (df=3; 31)
Note:	*p<0.1; **p<0.05; ***p<0.01

The regression analysis indicates that while the treatment (quiz-based preparation) did not have a significant overall effect on students' understanding of the material, it had a heterogeneous impact based on gender. Male students in the treatment group reported significantly higher understanding compared to their female counterparts ($p < 0.05$), suggesting that the quiz format may have been more effective for them. The model explains only a small portion of the variance ($R^2 = 0.107$), implying that other factors influence student learning. These findings suggest a need to tailor learning interventions to different student groups, potentially modifying the quiz format to enhance its effectiveness for all students.

While only the above regressions showed a significant treatment effect, additional regression results and discussions are provided in the appendix.

Limitations

1. *Limited Participant Pool:* The study was restricted to MSBA students due to the specific case study and survey content, limiting generalizability to other programs or disciplines.
2. *Non-Compliance and Survey Completion:* Students in the treatment group were more likely to leave the pre-class survey incomplete (coefficient = 0.183, p = 0.10), suggesting the quiz format may have been overwhelming or time-consuming.
3. *Low Post-Class Survey Completion:* Only 35 out of 55 participants (63%) who completed the pre-class survey also completed the post-class survey, potentially affecting result reliability.
4. *Time Constraints and Survey Fatigue*: A busy academic schedule and limited data collection time may have impacted both response rates and data quality.
5. *Discrepancy Between Expected and Actual Recall:* Students' predicted recall did not align with their actual recall, highlighting the limitations of self-reported data due to response bias or projection bias. *EDA results are provided in the appendix.*
6. *Potential Spillover:* Students in the treatment group could have discouraged others from taking the survey or shared the correct answers, thus reducing the effectiveness.

Future Steps & Conclusion

This study examined the impact of active versus passive learning on student engagement, preparedness, and material comprehension. While the overall treatment effect was not statistically significant, heterogeneous effects emerged, particularly with gender and prior case reading. Male students in the treatment group reported significantly higher understanding, while students who did not read the case beforehand struggled with recall and comprehension despite increased participation. These findings suggest that active learning interventions, such as quizzes, may not be universally effective but can be optimized based on student characteristics and prior knowledge.

These findings have significant business implications, particularly in corporate training and employee development. Organizations can enhance learning outcomes by tailoring interactive training methods to different demographics and incorporating structured pre-learning to boost engagement and retention. Adaptive learning tools like quizzes and gamification can further improve knowledge retention, making them valuable for onboarding, leadership training, and professional development programs.

Future research should examine the long-term effects of active learning and how demographic factors like age and cultural background influence outcomes. Investigating learning styles could help tailor instructional methods, while controlling for external factors like workload may improve result accuracy. Incorporating objective performance measures alongside self-reports could enhance data reliability, and exploring technology-driven learning tools may further optimize educational strategies.

Appendix

Google Drive Link

BA830_A1_Group7

Literature Review

In their meta-analysis, Freeman et al. (2014) evaluated the effectiveness of active learning versus traditional lecturing in STEM undergraduate courses by analyzing 225 studies that reported data on examination scores or failure rates when comparing student performance. The results revealed that active learning led to significant improvements in student performance, with examination scores increasing by approximately 6%, and failure rates decreasing by 55% as compared to traditional lecturing. Active learning methods, which ranged from group problem-solving to interactive activities such as clickers and peer instruction, showed consistent benefits across various STEM disciplines and class sizes. The study emphasized that active learning not only enhances student engagement and performance but also reduces failure rates, suggesting that implementing these methods could help address the challenges of STEM retention and degree completion. These findings provide strong, clear support for integrating active learning strategies into STEM education to improve learning outcomes and student success.

EDA

I. *Comparison of Survey Completion Time: Treatment vs. Control*

Statistic	Treatment: Time Spent (in seconds)	Control: Time Spent (in seconds)
Mode	137	263
Min	8	10
Max	69706	52950

Some participants likely started the quiz and completed it the next day after the professor shared the QR code. Using the mode might provide a more accurate measure of time differences. The participants in the treatment group spent twice as much time on the task as those in the control group.

II. *Comparison of Time Spent (Mode) by Students Who Read the Case Before the Quiz: Treatment vs. Control*

	Read the Case: Time Spent (in seconds)	Didn't Read the Case: Time Spent (in seconds)
Treatment	334	2367
Control	160	205

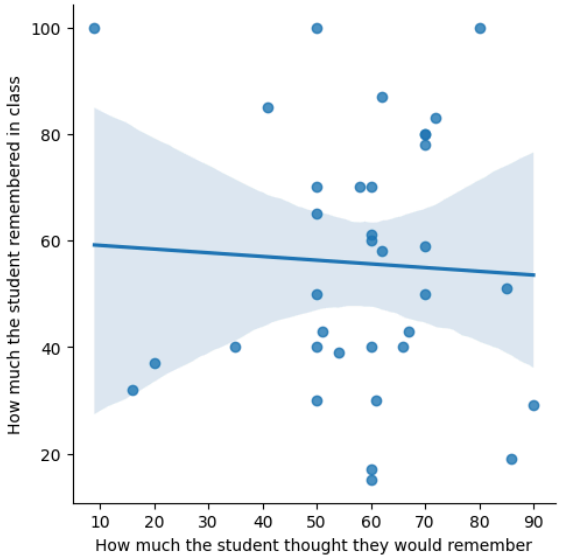
Participants who read the case beforehand generally spent less time on the survey compared to those who had not read it prior to taking the survey.

III. *Student Completion Rates for the Initial Survey*

	Percent of Complete Surveys
Read the Case	83%
Didn't Read the Case	65%

Participants who read the case beforehand were more likely to complete the survey.

IV. *How Much the Student Thought They Would Remember vs. How Much They Actually Remembered (Response Bias)*



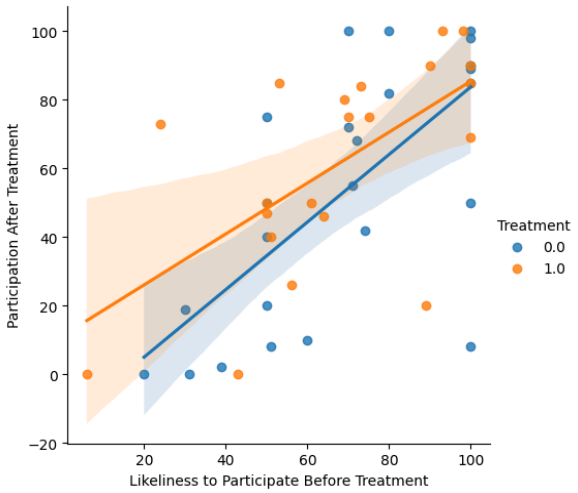
Regression Analysis (cont'd)

III. *Understanding of Material ~ Treatment + Read StubHub Case + Knowledge in Causality*

	Dependent variable: Understanding of Material
	(1)
CVQ2_read_stubhub_case_bool	-5.304
	(9.970)
CVQ3_knowledge_causality_val	0.872***
	(0.249)
Intercept	17.636
	(15.612)
Treatment	5.180
	(5.670)
Observations	35
R^2	0.372
Adjusted R^2	0.311
Residual Std. Error	17.255 (df=31)
F Statistic	4.601*** (df=3; 31)
Note:	*p<0.1; **p<0.05; ***p<0.01

The regression analysis reveals that prior knowledge of causality significantly enhances students' understanding, highlighting its importance in learning outcomes. However, reading the case before class showed no clear benefit, suggesting that pre-reading alone may not be sufficient for deeper comprehension. While the treatment group performed slightly better, the effect was not statistically significant, indicating that the intervention's impact remains uncertain. Factors such as cognitive fatigue or the complexity of the treatment may have influenced its effectiveness. These findings suggest that educational strategies should consider students' baseline knowledge and the potential for varying intervention effects based on individual learning needs.

IV. *Class Participation ~ Treatment + Likeliness to Participate*



	Dependent variable: Post Class Participation
	(1)
CVQ4_likeliness_participate_val	1.047***
	(0.148)
Intercept	-15.124
	(11.491)
Treatment	1.715
	(8.613)
Observations	35
R^2	0.522
Adjusted R^2	0.492
Residual Std. Error	23.781 (df=32)
F Statistic	24.997*** (df=2; 32)
Note:	*p<0.1; **p<0.05; ***p<0.01

The regression analysis highlights that students' pre-experiment self-reported likelihood to participate was the strongest indicator of actual class participation, suggesting that individuals who are already inclined to engage remained the most active. While the treatment group showed higher average participation, the effect was not statistically significant, indicating that the experiment may have had some influence but lacked strong evidence. These findings suggest that participation tendencies are largely pre-determined, and simply introducing a structured learning activity may not be enough to change participation behavior. To encourage broader engagement, educators may need to design interventions specifically targeting students who are less likely to participate.