



Presented at the DLSU Research Congress 2019  
De La Salle University, Manila, Philippines  
June 19 to 21, 2019

## Does Flexible Assessment Lead to Greater Student Engagement? Evidence from a Randomised Trial

Gerardo L. Largoza\*, Jashia Caila S. Chua<sup>1</sup>, Marinelle R. Macalindong<sup>1</sup>, Nicole B. Tajanlangit<sup>1</sup>

<sup>1</sup> School of Economics, De La Salle University

\*Corresponding Author: gerardo.largoza@dlsu.edu.ph

**Abstract:** Flexible assessment is the practice of allowing students some autonomy over how their grades are calculated, say by allowing them to assign the highest weights to the type of assessment they favour. Self-determination theory predicts that when students participate more actively in their own evaluation, they end up more engaged within their classes and improve their learning outcomes. In this paper, we report the results of a small-scale experiment we conducted, in which similar classes under the same professor were randomly assigned either to a flexible assessment treatment or traditionally assessed (with the professor exclusively deciding the weights). We found that flexible assessment – in this case allowing students to transfer up to five percentage points to their favoured output – raised their grades and academic performance...but had no significant impact on their engagement, as measured by Canvas' online metrics.

**Key Words:** flexible assessment; pedagogy; classroom experiments

### 1. INTRODUCTION

Within higher education, there has been increasing interest in “flexible assessment”, the practice of allowing students to choose when, where, and how they accomplish their course requirements (Rideout, 2017).

Part of it derives from self-determination theory, which argues that having autonomy, competence, and relatedness in an activity is crucial to strengthening one’s motivation, engagement, and performance. On the other hand, imposing deadlines,

fixed evaluations, or any restrictions upon class output, diminishes autonomy and possibly performance and engagement.

In this paper, we investigate claims about the link between flexible assessment, performance and engagement using a small-scale randomised controlled trial. The mode of flexible assessment is straightforward: students randomly assigned to the treatment group were permitted to add or deduct five percentage points from the preset or default assessment weights of their professor. Doing so allowed students to determine for themselves which assessments to focus on, and consequently how



Presented at the DLSU Research Congress 2019  
De La Salle University, Manila, Philippines  
June 19 to 21, 2019

engaged they would be throughout the course. If, as suspected, the type of assessment dominates how students allocate study time and effort, flexible assessment may create a virtuous cycle by which particular student strengths are magnified (compared to if weights were standardised), which leads to greater engagement, and which may in turn improve outcomes across the board.

A common choice among students is to decrease the weights of examinations (the most traditional of assessments) and increase the weights of performance-based tasks. The latter are often termed more “authentic”: i.e., evidence of learning based on the student’s ability to apply class concepts to real-world situations.

While some educators favour flexible assessment for its potential positive effects on learning outcomes — and studies such as those by Guest (2005) report students being highly satisfied with flexible assessment practices — doubts remain. Critics of flexible assessment argue that autonomy given to students, if not carefully restricted, could be abused. Instead of increasing student engagement, excessive autonomy may unduly shift attention toward obtaining higher marks at the expense of learning. Critics also argue that flexible assessment may contribute to grade inflation.

Our paper seeks clarity between these views, not from observational studies or correlational data, but from a randomised controlled trial which, although small in scale, allows us to calculate the causal effect of flexible assessment on both performance and student engagement.

## 2. METHODOLOGY

We first address a few key obstacles to running a randomised trial in a natural academic setting.

First, we must ensure that variation in the teacher input — experience, teaching style, course

content, etc. — is minimised. So we selected one professor teaching two sections of the same course. We also considered the nature of the course; we did not think it ideal to select a course already regarded as “high-effort” (e.g., a major course, say in Accountancy), as it would likely understate the impact of flexible assessment on already high levels of engagement. The course also had to be available to students of varying degree programmes, genders, abilities, and year levels to increase robustness and external validity — but we were constrained by reasons of fairness to randomise at the level of the class rather than the individual student.

To measure engagement, we used metrics available from the University’s Learning Management System, powered by Canvas: page views, number of contents viewed, time spent online. To proxy for student performance, we used final grades in the course.

We then coordinated with Professor XY and secured permission to allow a randomly-selected class the opportunity to add or subtract five points to the preset grade weights.

While analysing data, we also considered “Highest Allocated Assessment Grade” or the grade of the student on the assessment where he or she allocated most of the allowed weights. This was to measure the effect of flexible assessment when engagement is interacted with performance. Afterwards, we verified whether the two classes are statistically similar by running a difference-between-means test (t-test) on their demographics. Once we established their statistical similarity, we ran Ordinary Least Square (OLS) to obtain the effect of the flexible assessment on average, Average Treatment Effect (ATE) to show the exact treatment effect, Average Treatment Effect on Treated (ATET) to see the effect of the intervention on the treatment group, and 0.5 Quantile Regression to obtain the median effect of the flexible assessment, removing the effects of any outliers.



Presented at the DLSU Research Congress 2019  
De La Salle University, Manila, Philippines  
June 19 to 21, 2019

### 3. RESULTS AND DISCUSSION

In this section, we report results on the effectiveness of flexible assessment (student autonomy) on academic performance and engagement, having fitted OLS and median regression models on performance and engagement outcomes respectively.

On average, autonomy increases final grades by 0.1889 ( $t=1.87$ ,  $p=0.066$ ), and grades on highest allocated assessment by 0.3893 ( $t=3.19$ ,  $p=0.002$ ). The effect of the treatment is magnified by 0.2 on highest allocated assessment grade, where performance interacts with engagement. However, student autonomy does not affect engagement via any of the available measures.

Controlling for confounders, we used adjustment, weighting, and matching methods to determine the causal effect of autonomy on performance and engagement. Autonomy positively affects final grades and grades on highest allocated assessment by 0.2125 ( $z=2.09$ ,  $p=0.036$ ) and 0.4099 ( $z=3.39$ ,  $p=0.001$ ), respectively. The effect of autonomy on increasing performance is also magnified by 0.2 when performance is combined with engagement. However, autonomy does not affect student engagement.

Applying the same methods we used to infer causal effects of autonomy, students from the treatment group tallied increases of 0.1832 ( $z=1.85$ ,  $p=0.064$ ) in their final grades, and 0.3916 ( $z=3.15$ ,  $p=0.002$ ) on their highest allocated assessment grades. The effect of autonomy was lower for the treatment group than in general — though we could see that the treatment effect also magnified when performance is combined with engagement.

On the median, autonomy is effective in increasing student performance as well as engagement, but only in terms of contents viewed. Autonomy increases final grades by 0.23 ( $t=2.68$ ,  $p=0.010$ ), and grades on highest allocated assessment

by 0.57 ( $t=4.03$ ,  $p=0.000$ ). As it is also effective in increasing student engagement, autonomy increases the number of contents viewed in Canvas by 2 ( $t=1.77$ ,  $p=0.082$ ). Students also experience greater effect of autonomy, by 0.34, on highest allocated assessment grade, where performance is combined with engagement.

### 4. CONCLUSIONS

Our results indicate that allowing students to add or deduct weights of up to five percentage points from their professor's preset assessment weights increases their overall performance. However, it has no significant impact on their available measures of engagement.

Why would giving students the autonomy over their assessment weights increase their final grades? While it is possible that by modifying the assessment weights students were able to inflate their grades, we do not believe that this is the cause for our finding. In our data analysis, we adjusted the final grades to the preset weights of the professor, thus eliminating the grade differences caused by the treatment. Furthermore, when we computed the students' final grades using the weights they assigned to every assessment, half of the class even got a lower final grade.

Instead, we believe that allowing students to change their assessment weights increases their overall performance because it gives them control over how they will be evaluated. If a student reckons that he is not good at traditional exams, he could deduct percentage points from this assessment and transfer them to an assessment where he feels he would more likely excel. This creates an environment in which students more easily receive greater rewards for doing things that they are naturally good at, encouraging them to exhibit their abilities and be rewarded for what they can do, and not by what they cannot do.



Presented at the DLSU Research Congress 2019  
De La Salle University, Manila, Philippines  
June 19 to 21, 2019

However, giving them autonomy over grading weights does not necessarily increase their engagement because students may perceive higher cost for their efforts than the incentive they could get from adding or deducting up to five percentage points from the professor's preset assessment weights. It is also possible that the study efforts exerted by students to increase their grades were not captured by our data, since we were limited only to the engagement metrics that Canvas could provide. If students prepared for their assessments without using Canvas, those efforts would of course remain unrecorded.

All told, the results of this trial encourage further work, whether by scaling up the experiment, using sharper measures of performance and engagement, or perhaps even stronger versions of flexible assessment. Debates on teaching, learning, and assessment are ripe for evidence-based approaches that prioritise careful design and control.

## 5. REFERENCES

- Guest, R. (2005). Will flexible learning raise student achievement? *Education Economics*, 13(3), 287–297. <https://doi.org/10.1080/09645290500073761>
- Lang, J. M. (2018, March 13). 'How Much Do You Want Your Final to Count?' Retrieved December 1, 2018, from <https://www.chronicle.com/article/How-Much-Do-You-Want-Your/242802>
- Pacharn, P., Bay, D., & Felton, S. (2013). The Impact of a Flexible Assessment System on Students' Motivation, Performance and Attitude. *Accounting Education*, 22(2), 147–167. <https://doi.org/10.1080/09639284.2013.765292>
- Rideout, C. A. (2018). Students' choices and achievement in large undergraduate classes using a novel flexible assessment approach. *Assessment and Evaluation in Higher*

*Education*, 43(1), 68–78.

<https://doi.org/10.1080/02602938.2017.1294144>