

Engaging Teachers with Technology Increased Achievement, Bypassing Teachers Did Not

El objetivo de esta tarea es evaluar los efectos de la incorporación de videos producidos por expertos basados en el plan de estudios en:

1. Las puntuaciones de los alumnos en los 1 exámenes de matemáticas y ciencias
2. La probabilidad rendir el examen de continuidad del gobierno

1. Sobre el siguiente texto, comente utilizando los conceptos de validez externa e interna

“Our study took place within the Lahore, Multan, and Rawalpindi districts of Punjab province, Pakistan, the most populous province in Pakistan, home to over half of Pakistan’s 208 million residents. To be eligible for our study, schools had to appear in the Punjab School Census, include grades 1 through 10, and have a boundary wall, electricity, and physical classrooms—basic amenities in the Punjabi context. (. . .) From eligible schools, we selected schools separately for eLearn Classrooms and eLearn Tablets. For eLearn Classrooms, we selected 60 schools, an equal number of boys’ and girls’ schools, for the sample.20 Randomization at the school level was stratified by district and gender.21 One control school dropped out by the endline stage, leaving us with 29 control schools and 30 treatment schools.”

El objetivo es conocer el impacto de la inclusión de vídeos en el plan de estudios de los alumnos, y para ello los investigadores realizaron una evaluación experimental, mediante una selección aleatoria. Para ello, de toda la población de escuelas, seleccionaron sólo la población elegible (según los grados y la infraestructura, etc.). Esto significa que no será posible extrapolar los resultados del estudio a toda la población de escuelas, ya que serán fundamentalmente diferentes de las escuelas elegibles.

De la población elegible se tomó una muestra aleatoria de 60 escuelas, estratificada por género. Al tomar la muestra aleatoriamente, según la ley de los grandes números, para una muestra suficientemente grande no habría diferencias estadísticamente significativas en las medias de las variables observables y no observables entre la muestra el resto de la población, lo que garantiza la validez externa, es decir, es posible extrapolar los resultados de la evaluación a toda la población elegible.

Por último, se realizó el mismo proceso para asignar las escuelas a los grupos de control y de tratamiento, estratificando las escuelas por género y distrito. Esto garantiza la validez interna, es decir, que no haya diferencias estadísticamente significativas entre los grupos de control y de tratamiento, en todas las características, excepto la asignación al grupo de tratamiento. Esto es muy importante para asegurarse de que la diferencia en los resultados se debe únicamente al tratamiento y nada más.

Sin embargo, en ambos niveles la selección no fue puramente aleatoria, sino que se estratificó: por género, 50/50, para la muestra, y tanto por género como por distrito para los

grupos de tratamiento y control. Esto podría conllevar un sesgo de selección, si la población no sigue la misma proporción de distrito y género que la seleccionada para la estratificación. Por ejemplo, si el total de la población elegible tiene un 80% de niños en edad escolar, habría una diferencia estadísticamente significativa entre la muestra y la población, lo que significaría que la evaluación es externamente inválida.

2a. Indique si existen diferencias en cada una de las variables para los grupos de tratamiento y control. Comente si son comparables en estas variables. Sea explícita/o respecto al nivel de significancia que está utilizando en su análisis.

Para avaliar se existem diferenças estatisticamente significativas entre as mesmas variáveis de diferentes grupos, fazemos um teste t para os meios de distribuição dessas variáveis para os diferentes grupos, con las siguientes hipótesis:

- H0: No hay diferencias entre las medias
- H1: Hay diferencias entre las medias

En el caso de las variables que rechazan la hipótesis nula, i.e. valor $p < \alpha$ (0.05), hay pruebas estadísticamente significativas de que la media de estas variables es diferente entre los grupos, por lo que no se pueden comparar los grupos de tratamiento y control. Por lo tanto, hay que controlar estas variables: **zirttotalbl**, **attendancebl**, **classprobsummi**

Los resultados del test t se pueden ver en la siguiente tabela (para un nivel de significancia de 5%):

Variable	P-Value ($\Pr(T > t)$)	Interpretación
mothereducationbl	0.1294	No rechaza H0
homeworkbl	0.1497	No rechaza H0
classsize	0.4827	No rechaza H0
schoolfemale	0.1374	No rechaza H0
zirttotalbl	0.0007	Rechaza H0
attendancebl	0.0011	Rechaza H0
classprobsummi	0.0000	Rechaza H0

2b. ¿Qué implica la presencia de diferencias significativas?

Significa que los grupos no son comparables, es decir, que no son estadísticamente iguales. En términos prácticos, significa que el impacto en los resultados puede no deberse únicamente al tratamiento, y tal vez a las diferencias inherentes entre los grupos, lo que hace que la evaluación no sea válida internamente. Para superar este problema, podemos controlar las distintas variables observadas incluyéndolas en el modelo de regresión, de modo que sus efectos no se confundan con el efecto del tratamiento, lo que nos permite confiar más en los resultados.

Sin embargo, para aumentar la solidez de la evaluación, habría que hacer una nueva selección aleatoria, hasta que no se produzcan diferencias.

3. Determine el efecto de la participación del programa. Sea cuidadosa/o en la interpretación de sus resultados.

Para un nivel de significancia de 5%, como $p\text{-value} = 0.001$ es menor que 0.05, el resultado es estadísticamente significativo. Así, la participación en el programa Elearn Classrooms aumenta las puntuaciones de los alumnos en los exámenes de matemáticas y ciencias en 15.31%

Controlando por las variables *zirttotalbl*, *attendancebl* y *classprobsummi*, se concluye que el resultado sigue siendo significativo ($p\text{-value} = 0.000$) y que la participación en el programa Elearn Classrooms aumenta las puntuaciones de los alumnos en los exámenes de matemáticas y ciencias en 27.3%

Para un nivel de significancia de 5%, como $p\text{-value} = 0.000$ es menor que 0.05, el resultado es estadísticamente significativo. Así, la participación en el programa Elearn Classrooms aumenta la probabilidad rendir el examen de continuidad del gobierno en 4,68%

Controlando por las variables *zirttotalbl*, *attendancebl* y *classprobsummi*, se concluye que el resultado sigue siendo significativo ($p\text{-value} = 0.000$) y que la participación en el programa Elearn Classrooms aumenta la probabilidad rendir el examen de continuidad del gobierno en 4.8%. Sin embargo, las variables *zirttotalbl* y *classprobsummi* presentan un resultado no significativo ($p\text{-value} > 0.05$), por lo que se considera que no tienen impacto en el resultado obtenido.