

# Estandarizar la adquisición del conocimiento, ¿beneficia o no?

Figuroa Rodríguez, Jorge Luis

2017-03

---

<http://hdl.handle.net/20.500.11777/2586>

<http://repositorio.iberopuebla.mx/licencia.pdf>



# Estandarizar la adquisición del conocimiento, ¿beneficia o no?

J.L. Figueroa-Rodríguez<sup>(1)</sup>

<sup>(1)</sup> Prepa Ibero Puebla

Academia de Matemáticas e Informática

[jorgeluis.figueroa@iberopuebla.mx](mailto:jorgeluis.figueroa@iberopuebla.mx)

*En el siguiente trabajo se expondrán las primeras conclusiones que se obtuvieron al introducir evaluaciones parciales y finales “estandarizadas” en cuatro grupos de la materia de Probabilidad y Estadística, materia en la que se observaba cierta dispersión en las calificaciones finales. Se encontró que al implementar dichas evaluaciones de control en los grupos, los promedios finales tendieron a igualarse en los cuatro grupos.*

**Palabras clave** — pruebas estandarizadas, seguimiento, evaluación continua.

## I. INTRODUCCIÓN

La evaluación de los conocimientos adquiridos a lo largo de un curso ha sido, es y seguirá siendo uno de los puntos centrales en la impartición de un curso. Este tema se vuelve aún más complejo si hablamos que de un mismo curso existen muchos grupos con diferentes profesores. ¿Cómo evaluamos si los alumnos están adquiriendo los mismos conocimientos a un nivel suficiente?

En cada grupo la dinámica es única. El aprendizaje puede ser más lento, más rápido, más profundo, con mayor o menor debate, sólo por citar algunos puntos de la transmisión del conocimiento. Además tenemos que añadir el factor de qué temas nos apasionan a nosotros en cada curso. En resumidas palabras, la cantidad de variables que están involucradas en la impartición de un curso es inmensa. Sin embargo, al finalizar un curso debe haber un número de temas cubiertos a un cierto nivel. Es decir, la aplicación de pruebas estandarizadas muchas veces se vuelve en la única herramienta que se tiene.

La aplicación de pruebas estandarizadas ha sido controversial, teniendo sus defensores y opositores. Popham (1999) sostiene que las pruebas estandarizadas no permiten evaluar la calidad de la educación por tres razones: desajustes entre lo enseñado y lo medido, la tendencia psicométrica a eliminar ítems importantes de la prueba y causalidades confundidas. Martínez-Rizo (2016) señala que una consecuencia positiva de la

aplicación de pruebas a gran escala es la atención de la sociedad sobre la relevancia de hacer esfuerzos para mejorar los niveles de aprendizaje.

La aplicación de algún tipo de prueba estandarizada se previó en los grupos de Probabilidad y Estadística de nivel universitario (cuarto semestre). Para esta materia existían seis grupos con cuatro profesores diferentes. Inicialmente los promedios de los grupos rondaban entre 7.3 y 8.9 con un nivel de aprobación de cerca del 85%.

## II. DESARROLLO DEL ARTÍCULO

El proyecto se dividió en tres etapas: la primera etapa fue reunir a los profesores involucrados en la impartición de la materia para discutir que estrategias se implementarían de acuerdo a las herramientas disponibles para este propósito; la segunda etapa fue preparar tareas y exámenes en línea de cada uno de los temas que cubría el temario (esta decisión fue tomada ya que se detectó que no todas las tareas y evaluaciones del curso tenían el mismo nivel de complejidad y los temas preguntados no eran homogéneos); la tercera etapa consistía en aplicar exámenes departamentales (no idénticos en contenido pero si idénticos en temas y complejidad).

Para medir el desempeño de los alumnos se tomaron los promedios finales del curso como variable de medición, ya que el promedio engloba las evaluaciones parciales, tareas y exámenes en línea.

## III. RESULTADOS Y CONCLUSIONES

En la primera etapa del proyecto se discutió que herramientas había disponibles y que estrategias se iban a implementar en estos grupos. Se decidió generar una batería de tareas y exámenes en línea, utilizando la plataforma institucional para dicha aplicación.

Para la segunda etapa del proyecto se prepararon 12 tareas que cubrían los temas del curso en un nivel intermedio. Además se generaron pruebas en línea que seguían la misma temática y un nivel

ligeramente más alto de las tareas. Las tareas se dejaban cuando se habían dado las bases de cada tema y se entregaban unos días después de haber terminado el tema. En este momento se aplicaba una prueba en línea, la cual podía realizarse 3 veces. Cabe mencionar que cada intento de las pruebas en línea se generaba de forma aleatoria por lo que los conceptos tenían que estar totalmente claros para poder contestar cada intento y obtener una mejora. Para la generación aleatoria de las pruebas en línea se cargaron bancos de ejercicios de más de 50 preguntas por prueba de las cuales se tomaban 10 de forma aleatoria.

En la figura 1 se muestran los resultados antes de la aplicación de los exámenes en línea. Se muestran los promedios de 4 grupos en una escala de 0 a 100.

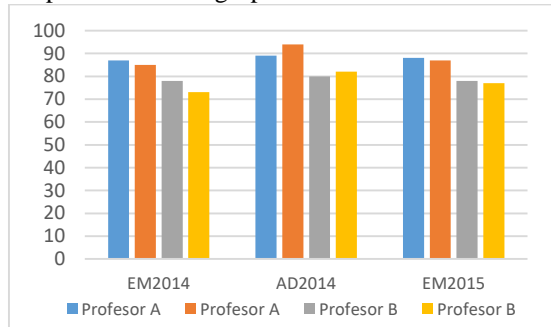


Figura 1. Calificaciones de 4 grupos con 2 profesores antes de la aplicación de evaluaciones en línea.

En la figura 2 se muestran los resultados de la aplicación de la aplicación de exámenes en línea. Se empiezan a observar movimientos a la baja y al alta de los grupos.

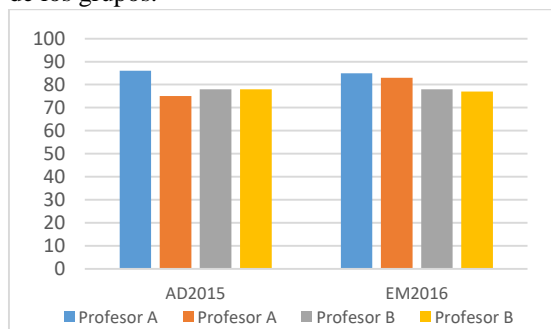


Figura 2. Calificaciones de los grupos después de la aplicación de los exámenes en línea

Para la tercera etapa se propuso aplicar exámenes departamentales. Estos exámenes se generaban aleatoriamente a partir de bancos de ejercicios contruidos para evaluar cada uno de los temas consensuados al iniciar el proyecto. Esto con la finalidad de obtener que temas son los que

representaban problemas para los alumnos y cuáles representaban fortalezas.

La medición de dichas variables requiere un gran esfuerzo por parte de los profesores. Sin embargo cada uno de estos esfuerzos va ligado a una mejora de los cursos. Uno de los primeros resultados obtenidos fue disminuir la dispersión en los promedios de los alumnos, obteniendo un valor de 8.2. A pesar de ver una “disminución” en el promedio de algunos grupos, la disminución de la variabilidad nos indica que los intentos de estandarizar el conocimiento han dado resultados positivos. En la figura 3 se tiene el único periodo registrado tanto con la aplicación de exámenes en línea como la aplicación de examen final departamental. Se observa como los promedios ya son muy similares.

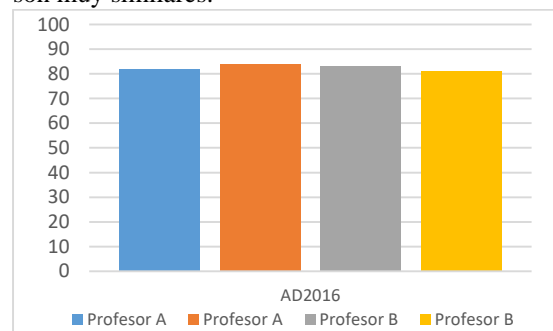
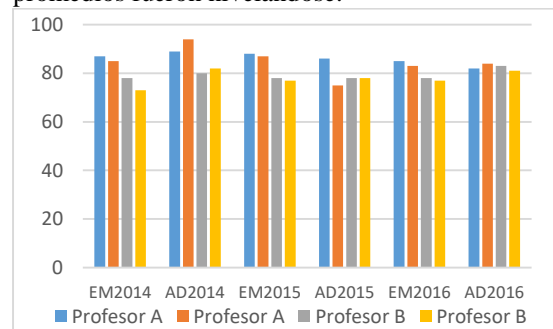


Figura 3. Calificaciones después de la aplicación de exámenes en línea y examen departamental.

En la figura 4 se observa un comparativo a lo largo del tiempo del desempeño de los grupos conforme se fue aplicando cada una de las evaluaciones. Observamos como en el último periodo los promedios fueron nivelándose.



Las razones posibles del porqué se observa una disminución en algunos grupos, puede deberse a que los alumnos se habían acostumbrado a un cierto tipo de ejercicios, es decir que la redacción de problemas por parte de diferentes personas pudo influir; otra posible causa es que la complejidad de



los problemas no era el adecuado. El observar la nivelación de las calificaciones tampoco puede ser concluyente ya que solo se tiene un periodo registrado. Las conclusiones todavía no son totalmente claras por lo que el proyecto debe continuar.

En lo particular la aplicación de exámenes “estandarizados” provee información valiosa, tanto a alumnos como a docentes, para la mejora continua de los cursos. Como se ha dicho una prueba estandarizada por sí misma no es la solución a la calidad en la enseñanza, sino una herramienta más para la mejora. La información obtenida es vital para detectar posibles fallos en la transmisión del conocimiento en áreas específicas o la detección de casos críticos.

#### IV. AGRADECIMIENTOS

Agradezco a la Mtra. Aída Ignacia Salázar Compañ por su colaboración en este proyecto, ya que sin su apoyo no hubiera podido realizar este estudio. La Mtra. Salázar ahora está a cargo del proyecto y le deseo mucho éxito para la continuación del mismo.

#### V. REFERENCIAS

- [1] Martínez-Rizo, Felipe (2016). Impacto de las pruebas en gran escala en contextos de débil tradición técnica: Experiencia de México y el Grupo Iberoamericano de PISA. RELIEVE, 22 (1), art. M0. DOI: <http://dx.doi.org/10.7203/relieve.22.1.8244>
- [2] Popham, W. James (1999) ¿Por qué las pruebas estandarizadas no miden la calidad educativa? PREAL-Grade. Recuperado de [http://www.oei.es/evaluacioneducativa/pruebas\\_estandarizadas\\_no\\_miden\\_calidad\\_educativa\\_popham.pdf](http://www.oei.es/evaluacioneducativa/pruebas_estandarizadas_no_miden_calidad_educativa_popham.pdf)