

EVALUACION EXTERNA DE LA REFORMA DE LAS ENSEÑANZAS

Mariano ALVARO PAGE
Centro Nacional de Investigación
y Documentación Educativa (CIDE)
Ciudad Universitaria s/n
28040 MADRID
tn.: 449 77 00 Ext. 262

Resumen:

Versa sobre la Evaluación Externa del Primer Ciclo de la Reforma de las Enseñanzas Medias. El esquema seguido es el siguiente:

- Objetivo de la investigación: Comparar resultados de alumnos de la Reforma con los de la enseñanza vigente.
- Diseño: Cuasi experimental y longitudinal.
- Pruebas utilizadas, criterios de selección de las mismas y variables medidas: de rendimiento, aptitudes, actitudes, aspectos didácticos y expectativas académico profesionales.
- Aplicaciones o trabajos de campo: Octubre-noviembre de 1984, octubre-noviembre de 1985, mayo-junio de 1986 y abril-mayo de 1987.
- Resultados: Los alumnos de la Reforma consiguen resultados muy parecidos a sus compañeros de BUP con siete horas menos de clase a la semana y netamente superiores a los de FP con cinco horas semanales más de clase.

Palabras clave:

Evaluación. Reforma. Generación. Grupo experimental y grupos de control. Pretest y postest. Estudio longitudinal.

OBJETIVOS

El objetivo primordial, que comprende otros parciales que se relatarán más tarde, consiste en *ver si después de un año o dos de experiencia los alumnos experimentales obtienen mejores, peores o iguales resultados en actitudes cívico-sociales, sentimientos hacia el estudio, aptitudes y rendimiento que los alumnos de control.*

Este objetivo general puede desglosarse, por que los incluye, en estos otros tres parciales:

a) Comparación de los resultados obtenidos por el grupo experimental con los de los grupos de control en la primera aplicación de ambas generaciones (pretest), lo que nos indicará cual era el punto de partida, la línea base de todos los alumnos al inicio del experimento; es decir, cuando se comenzaba a someter a prueba un nuevo tipo de enseñanza.

b) Comparación de los resultados obtenidos por el grupo experimental con los obtenidos por los grupos de control en el primer y segundo postest lo que nos indicará cual es la situación, después de un año y/o dos de experiencia, de los tres grupos comprendidos en el diseño: experimental, control interno y control externo.

c) Estudio de la evolución de los tres grupos, lo que podremos hacer comparando los resultados del pretest con los del postest.

Todavía podría hacerse otra subdivisión de objetivos según las áreas de estudio: actitudes cívico-sociales, sentimientos hacia el estudio, aptitudes y rendimiento.

DISEÑO

El diseño de este estudio podría catalogarse de *cuasi-experimental*, siguiendo la terminología de Campbell y Stanley (1970) con medidas sucesivas, una pretest y varias postest. En esta comunicación sólo se adelantan datos del pretest y primer postest.

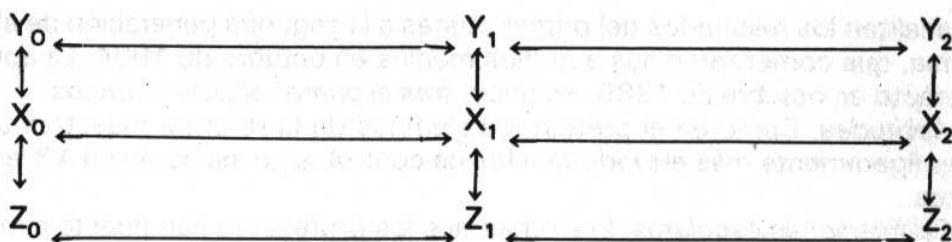
En el diseño existen tres grupos: uno *experimental* y dos de control. De estos dos últimos uno pertenece a los mismos centros que el grupo experimental y otro está construido a partir de una muestra representativa a nivel estatal. Al grupo de los centros experimentales hemos dado en llamarle *control interno* y al de los centros de la muestra, *control externo*.

El diseño no puede catalogarse de puramente experimental porque, por una parte, la elección de los centros experimentales y, por consiguiente de los alumnos no se hizo de manera aleatoria y, por otra, la asignación de alumnos a grupo control y a grupo experimental, dentro de los centros experimentales no estuvo basada en el mero azar.

Existen razones obvias y comprensibles desde el punto de vista social para haber alterado las leyes probabilísticas. Pues para la elección de los centros experimentales hubo que contar con la elección expresa de la Dirección de los centros y para pertenecer al grupo experimental dentro de estos centros era necesario el expreso consentimiento de los padres.

La asignación no aleatoria en muchos casos a un grupo o a otro dentro de los centros experimentales, la no existencia en alguno de estos centros de grupo de control y la vivencia, de alguna manera, de una situación o clima de experimentación en los centros experimentales por todos los alumnos, incluso los de control, llevó al equipo del CIDE a tomar la decisión de elegir otro nuevo grupo que sirviera de contraste.

El diseño completo podría esquematizarse de la siguiente forma:



- X_0, Y_0, Z_0 : Medidas pretest
- X_1, Y_1, Z_1 : Medidas de postest, después de 1 año escolar
- X_2, Y_2, Z_2 : medidas postest después de 2 años escolares.
- X = Grupo experimental
- Y = Grupo control, dentro de los centros experimentales.
- Z = Grupo control, de la muestra representativa.

←→ Comparación entre distintas medidas de un mismo grupo.



↑ ↓ Comparación entre grupos.

Por tanto, el diseño permite:

a) Establecer una línea base de cada uno de los tres grupos del estudio, lo que permitirá comparar el grupo experimental con los otros dos grupos antes de comenzar la experiencia.

b) Comparar las distintas medidas obtenidas de cada uno de los grupos.

De esta forma, podrá saberse no sólo si existen diferencias entre grupos, sino también cual ha sido la evolución de cada uno de ellos.

De lo anterior se deduce que el presente estudio también podría catalogarse de *longitudinal*. Pues se estudian dos cohortes de alumnos a lo largo de dos años.

RESULTADOS

Hasta la fecha sólo se han acabado los análisis referidos a la 1ª y 2ª aplicación de la II generación. En las fechas de la celebración de las Jornadas estará ultimado el análisis de la evaluación de los tres grupos después de un año de experiencia.

Resultados del pretest o línea base

En todas las variables dependientes estudiadas apenas se aprecian diferencias relevantes entre los tres grupos de alumnos comparados: experimental, control interno y control externo, aunque en muchos casos existan diferencias estadísticamente significativas.

En concreto, tan sólo en tres variables se observan diferencias significativas entre los tres grupos, siendo además el grado de asociación superior al 1%: Matemáticas-Cálculo (2,25%) en el caso de BUP; estudios querrías terminar y profesión futura con un porcentaje de varianza explicada de 3,24 y 2,89 respectivamente en el caso de FP.

En las restantes variables con diferencias significativas con respecto al grupo, el porcentaje de varianza explicado no llega al 1%, por lo que no se consideran relevantes. Este fenómeno puede explicarse por el hecho de que al contar con un N tan grande, cualquier pequeña diferencia entre los grupos puede resultar significativa.

A partir de estos resultados nos atrevemos a concluir que el punto de partida, es decir, la línea base es muy semejante en las variables estudiadas en los tres grupos, especialmente entre el experimental y el control interno, lo cual garantiza la validez de nuestro diseño.

Resultados del primer postest

Se analizan los resultados del primer postest a la segunda generación de alumnos de la Reforma, que comenzaron sus estudios medios en octubre de 1984. La aplicación se llevó a efecto en octubre de 1985, es decir, *tras el primer año de estudios*.

1. *Aptitudes*. Como en el pretest los alumnos de la reforma muestran un nivel de aptitudes ligeramente más elevado que los de control, algo así como un 4% en términos numéricos.

2. *Calificaciones escolares*. Las notas que los profesores han puesto a los alumnos

son prácticamente las mismas en todos los grupos, pero los criterios de calificación siguen siendo más severos en los centros de BUP.

3. Pruebas objetivas de rendimiento.

En los centros de BUP, los alumnos de la Reforma consiguen resultados más o menos iguales a sus compañeros con siete horas menos de clase a la semana; en los FP se consiguen resultados netamente superiores, pero con cinco horas semanales más de clase. En conjunto, la Reforma no produce ningún descenso en el nivel de conocimiento de sus alumnos en relación a los alumnos de los planes vigentes y muy probablemente es la causa de ligeras mejoras.

La Reforma no reduce ni aumenta las desigualdades de rendimiento de los alumnos.

4. Repeticiones y abandonos. Los alumnos de la reforma repiten y abandonan menos que los de los planes vigentes.

5. Actitudes cívico-sociales. No se han apreciado diferencias entre los alumnos experimentales y los de control.

6. Reacciones emocionales ante el estudio. La reforma parece causar menos ansiedad, motivar más y suponer un trabajo más satisfactorio para sus alumnos en los centros de BUP, pero no en los centros de FP.

7. Didáctica. Los profesores de la reforma practican una metodología mucho más activa que los de control, pero la satisfacción de los alumnos no es mayor.

8. Mantenimiento de las «subculturas» de BUP y FP. Pese a la Reforma, en los grupos experimentales de ambos tipos de centros siguen pautas de calificación distintas, y también los criterios de repetición de curso, las tasas de abandono y repetición y las actitudes cívico-sociales de los alumnos muestran pautas diferenciales en los centros de BUP y de FP.

BIBLIOGRAFIA

Existe en nuestro archivo bibliográfico un número ingente de referencias bibliográficas sobre rendimiento, pero estudios que pretendan evaluar un sistema educativo o un nivel educativo son muy pocos. Creemos que, como ejemplos, los más representativos son los siguientes:

CARABAÑA MORALES, J. *Educación, ocupación e ingresos en la España del siglo XX*, Madrid: Servicio de Publicaciones del MEC, 1983.

COLEMAN, J.S. et. al. *Report on equality of educational opportunity*. Washington: V.S. Government Printing Office for Department of Health Education and welfare, 1966.

JENKS, C. et. al. *Inequality, A Reassessment of the effects of family and Schooling in American*. Harper and Row, N. Y., 1972.