

LA UTILIZACIÓN DE BASES DE DATOS COMO HERRAMIENTAS DE EVALUACIÓN

José Torreblanca López, Matilde Sancho Martínez y Antonio Ariza García

Universidad de Sevilla

En este artículo se comprueba la validez de utilización de las bases de datos para elaborar pruebas objetivas. Al comparar distintos grupos no se han encontrado diferencias estadísticamente significativas entre los alumnos que realizaban las pruebas elaboradas manualmente por el profesor y los que realizaban las pruebas preparadas por el ordenador. Sin embargo, la flexibilidad que permite este sistema produjo un descenso drástico de los alumnos no presentados al final de curso.

This article examines the validity of the data bases from which objective tests are elaborated. No statistically significant differences were found between the teacher created tests and those generated by computer. Nevertheless, the flexibility allowed by this system resulted in a drastic reduction in absenteeism for final examinations.

DESCRIPTORES: Evaluación asistida por ordenador, bases de datos, pruebas objetivas, evaluación.

1. Introducción.

En la actividad docente encontramos tres momentos bien definidos, la planificación o programación de la asignatura, la realización en sí (clases, seminarios, practicas, etc.), y por último el proceso de evaluación o control de los resultados obtenidos. Por tanto la evaluación es una actividad sistemática integrada en el proceso educativo, cuya finalidad es servir de diagnóstico, como señala De Ketele (1984), no solo del rendimiento del alumno, sino también del enseñante, los objetivos, la técnica, el método. En definitiva, de todo el proceso de enseñanza-aprendizaje. Metodológicamente, la evaluación resulta de relacionar dos realidades que son: lo observado y lo esperado.

La evaluación es la fase final de un proceso en donde se comprueba si el diseño de enseñanza que se ha llevado a cabo es correcto o no, es decir, si los objetivos que se marcaron al comienzo del proyecto se han logrado o no. Si el alumno ha conseguido los objetivos marcados al comienzo del proceso diremos que éste ha sido planteado de forma correcta; si no es así tendremos que replantear y corregir lo que sea necesario.

Son muchos los procedimientos de evaluación que podemos utilizar, siendo válidos todos aquellos que sean capaces de poner de manifiesto si las actividades del profesor y del alumno llevan al logro de los objetivos propuestos.

Una de las técnicas de evaluación utilizadas con más frecuencia en el campo de la enseñanza son las pruebas objetivas. Las pruebas objetivas se pueden considerar como un examen escrito con características muy concretas entre las que cabe destacar como más interesantes:

- Las preguntas han de ser muy breves.
- Han de estar enunciadas muy claramente.
- Deben permitir una calificación cuantitativa rápida.

La mayor dificultad de esta técnica de evaluación es que dada la exigencia de objetividad que tiene esta prueba, la redacción de sus ítems no es algo simple y elemental. Una redacción adecuada exige del profesor preparación, imaginación, esfuerzo y práctica.

Estas pruebas de tipo "test" permiten valorar fundamentalmente la información que posee el alumno sobre el contenido de la asignatura, pero no valoran otros aspectos como la creatividad, la capacidad de elaboración y expresión personal de los contenidos, la originalidad, etc. Las principales ventajas de este sistema radican en que puede abarcar gran parte de los temas tratados en clase, la evaluación es totalmente objetiva y es de gran facilidad de corrección.

Modelos de pruebas objetivas hay varios que, en esencia coinciden, aunque varían en el enunciado de los apartados. Los más conocidos son los modelos propuestos por Norman Gronlund (1971, 1974), García Hoz (1946), Fernández Huertas (1950) y Marín Ibáñez (1971) en sus respectivas obras.

Norman Gronlund propone los siguientes tipos de preguntas:

1.- Respuesta íntegra por el estudiante.

- La respuesta es una frase íntegra o texto.
- La respuesta es interpolar palabras en una frase o texto.

2.- Selección de respuesta entre varias.

- Contestar Verdadero o Falso.
- Elección de la respuesta entre varias dadas (Elección múltiple).

García Hoz propone el siguiente:

1.- Evocación de un recuerdo.

- Respuesta simple a una pregunta.
- Texto mutilado.

2.- Reconocimiento de un recuerdo.

- Respuestas sugeridas.
- Elección entre ambas.
- Elección única.
- Elección múltiple.
- Apreciar la mejor respuesta.
- Verdadero o Falso.
- Asociar conocimientos.

Las normas para la elaboración de ítems con respuestas de Verdadero o Falso que propone Gronlund son las siguientes:

- Cada uno debe expresar solo una idea.
- Evitar información de tipo general.
- Si es una opinión, teoría o hipótesis de algún autor o escuela debe decirse.
- Deben evitarse las afirmaciones no significativas en el conocimiento de la materia.
- No deben hacerse formulaciones negativas ni excesivamente largas.
- Las formulaciones verdaderas o falsas deben ser de la misma longitud.
- Repartir el número de enunciados de forma homogénea por el texto.

Para la formulación de ítems de elección múltiple, que tiene la siguiente estructura:

Enunciado = Base

Posibles respuestas = Opciones

Opciones falsas = Distractores

Opción verdadera = Respuesta

propone para su elaboración:

- Formulación de un solo problema en la base, que debe ser claro y sencillo, incluyendo más palabras en la base, descargando a la respuesta de información.
- No incluir en la base información negativa.
- Formulación muy precisa de la base.
- Evitar formulación en donde se incluyan pistas que conduzcan a la respuesta sin saberla.
- Formular los distractores de forma que resulten verosímiles y homogéneos con la respuesta.

En el caso de grandes grupos la utilización de pruebas objetivas es muy adecuada, aunque su preparación es compleja, permite una fácil corrección.

Debido a la bondad de este tipo de pruebas se han desarrollado proyectos de investigación como the Cooperative project on Evaluation of Results of Training (CERT) (Alonso 1984). Como indican De Pablos y otros (1993) en el seno de este proyecto se desarrolló un programa informático que apoya y facilita el uso de pruebas objetivas en la evaluación educativa, posibilitando gestionar gran cantidad de información recogida a través de pruebas objetivas de opción múltiple y analizarla, agilizando el proceso de manipulación y cálculos manuales de los formadores para evaluar aprendizajes. Una de las características principales de este programa es la inclusión de los denominados niveles de confianza que son una estimación que el alumno realiza en base a una escala de 0-5 del grado de confianza o seguridad con que responde a cada una de las cuestiones de la prueba.

Las capacidades del ordenador pueden utilizarse en un amplio rango de actividades relacionadas con los test. En concreto el ordenador se puede utilizar para escribir ítems, construir bancos de ítems, construir y administrar test, así como puntuar, analizar e informar de los resultados obtenidos en los test. Para llevar a cabo estas actividades en la actualidad se dispone de un elevado número de recursos tecnológicos (Rodríguez, García y Gil, 1996).

Como destacan Arriaga y otros (1996) el desarrollo del software educativo está llegando a un punto en el que los nuevos programas no mejoran sustancialmente el interés de los profesores por incorporarlos a la docencia, normalmente por sus entornos cerrados en los que no es posible la modificación y se obliga a una aceptación tal cual.

Intentando salvar estos inconvenientes nuestro primer objetivo ha sido verificar la validez de la utilización de las bases de datos como un primer paso en la elaboración de un programa de evaluación asistido por ordenador.

2. Material y método.

Se ha utilizado un programa de base de datos (dbase IV) para la gestión de un conjunto de preguntas de respuesta múltiple que, con anterioridad había sido empleado en cursos previos. Estas bases de datos se pueden ir actualizando y aumentando continuamente. La objetividad del sistema será mejor cuanto mayor sea la base de datos.

En un primer análisis, nos planteamos comprobar la validez del uso del ordenador para la realización de estas pruebas, para ello, estudiamos las posibles diferencias con los resultados que se habían obtenido previamente con exámenes elaborados manualmente por el mismo profesor. Así, el primer punto que nos fijamos para verificar la posibilidad de utilización de este sistema de exámenes fue comprobar la diferencia que podría existir con los cursos de años anteriores en los que la preparación del examen la hacía el profesor escogiendo un mismo número de preguntas de los diferentes temas en contraste con el ordenador que escoge las preguntas de manera aleatoria. Para esta comprobación se utilizaron listas de preguntas que habían servido para exámenes de cursos anteriores. Con el fin de que las condiciones experimentales fueran lo más homogéneas posibles, los cursos fueron seleccionados para que el temario de cada parcial y la fecha de realización de éste fueran similares. Los

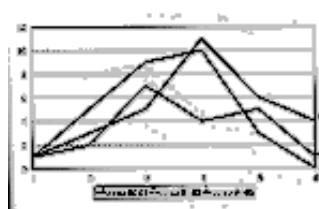
resultados se analizaron con las correspondientes pruebas estadísticas de análisis de diseño experimental en el que se comparan dos tratamientos midiendo una variable de respuesta cuantitativa.

3. Resultados.

Como se puede observar las notas (sobre un máximo de 6 puntos) de los diferentes cursos son algo distintas, pero si se analizan estadísticamente, utilizando el test T de Student-Fisher para comparación de medias, estas diferencias no son significativas como se puede observar en la siguiente tabla y su correspondiente gráfico.

| Primer Parcial | Media \pm E. Std. | N |
|----------------|---------------------|----|
| curso 90-91 | 3.49 \pm 0.22 | 30 |
| curso 91-92 | 3.12 \pm 0.26 | 20 |
| curso 94-95 | 2.95 \pm 0.18 | 28 |

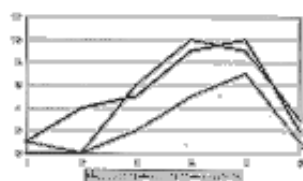
Primer Parcial (parte objetiva)



Al realizar el seguimiento de los datos que se obtuvieron en el segundo parcial se puede comprobar que el sistema es totalmente válido pues al aplicar las pruebas estadísticas, antes comentadas, no hay ninguna diferencia significativa entre los grupos, como se puede observar en la tabla:

| Segundo Parcial | Media \pm E. Std. | N |
|-----------------|---------------------|----|
| curso 90-91 | 3.36 \pm 0.23 | 31 |
| curso 91-92 | 3.83 \pm 0.27 | 16 |
| curso 94-95 | 3.87 \pm 0.17 | 28 |

Segundo Parcial (parte objetiva)



Esta igualdad entre los tres diferentes cursos se mantiene al analizar los resultados de la convocatoria de Junio (nota sobre 10):

| Convocatoria de Junio | Media \pm E. Std. | N |
|-----------------------|---------------------|----|
| curso 90-91 | 6.17 \pm 0.47 | 27 |
| curso 91-92 | 6.47 \pm 0.34 | 17 |
| curso 94-95 | 6.78 \pm 0.21 | 29 |

Convocatoria de Junio

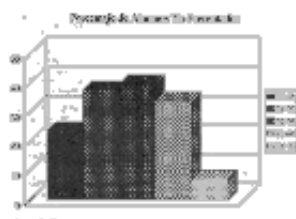


Una vez que parece establecida la igualdad entre los dos sistemas de generación de pruebas, debemos de resaltar que la utilización del ordenador permite una flexibilidad mucho mayor en la elaboración y realización de éstas. Actualmente la estamos utilizando, para ofrecer a los alumnos la posibilidad de presentarse al examen, previa cita, en día diferente al concertado con todo el grupo de clase. Esto no es una novedad, ya se podía hacer antes, pero no con la facilidad actual, que realmente pone a nuestro al-cance esta posibilidad y con el efecto motivador que tiene para el alumno.

Este simple cambio es el que se ha mostrado muy importante, y como consecuencia de su mas fácil adaptación al ritmo de aprendizaje del alumno, es el que provoca que el porcentaje de alumnos no presentados en los últimos cursos impartidos sea tan diferente con el actual.

| Curso | Alumnos no presentados (%) |
|-------|----------------------------|
| 90-91 | 22.85 |
| 91-92 | 37.03 |
| 92-93 | 39.13 |
| 93-94 | 33.33 |
| 94-95 | 6.45 |

Porcentaje de Alumnos No Presentados



Si comparamos los datos mediante el test T de datos apareados la diferencia es significativa para $p = 0.005$.

4. Discusión.

Tras el análisis de estos datos se puede concluir que el sistema de elaboración de exámenes por extracción aleatoria de preguntas de las bases de datos existentes, ofrece los mismos resultados que los exámenes realizados manualmente por el propio profesor que ha generado la base de datos. Con lo que se podría concluir que la ventaja del sistema recae sobre todo del lado del profesor al permitirle generar con gran facilidad exámenes diferentes. Sin embargo el alumno también se ve favorecido por esta herramienta, no sólo por el hecho de evitar cualquier temor de subjetividad por parte del profesor, sino que también al flexibilizar el periodo de presentación del alumno al examen, éste lo hace cuando realmente se siente preparado y no cuando le obligue la fecha establecida por todo el grupo de clase, intentando una fecha consensuada, que generalmente no es la mejor para cualquiera de los alumnos individualmente.

Aparte de lo descrito anteriormente este sistema favorece la responsabilidad personal de cada alumno sobre su proceso de enseñanza-aprendizaje al tener un control mucho mayor sobre los momentos que más tensión causan en el alumnado.

Se puede observar que existe una clara mejora de resultados en los alumnos al tener la posibilidad individualizada de examen. La mejora creemos que residiría no sólo en un mayor tiempo para poder afianzar los conocimientos, si no que además el alumno, al elegir el día en que va a examinarse escoge éste con todas las garantías para poder desarrollar su aprendizaje en las mejores condiciones. No estando supeditado al día que elige el curso que no es otro que el menos malo de todos los posibles. En esta simple diferencia creemos que reside la disparidad de resultados que se ha podido poner de manifiesto en este estudio. Los resultados obtenidos

son tan esperan-zadores que esperamos continuar con el desarrollo del sistema y con la posible elaboración de un programa de ordenador que permita no sólo elaborar sino corregir las diferentes pruebas.

Referencias bibliográficas.

ALONSO, C.y OTROS (1992). **Principios comunes para la evaluación de los resultados cognitivos de la formación.** Comisión de las Comunidades Europeas. Programa Eurotecnet.

ARRIAGA, J.y OTROS (1996). Sistemas de autor orientados a un fin educativo específico. **Pixel-Bit. Revista de Medios y Educación, 6.** 5-13.

DE KETELE, J.M.(1984). **Observar para Educar. Observación y evaluación en la práctica educativa.** Madrid, Visor Libros.

DE PABLOS PONS, J. y OTROS(1993). La evaluación del alumno en la Universidad: el proyecto CERT. **Revista de Enseñanza Universitaria, 6.** 49-71.

FERNÁNDEZ HUERTA, J.(1950). **Las pruebas objetivas en la escuela primaria,** Madrid, C.S.I.C.

GARCÍA HOZ, V.(1946). **Normas elementales de Pedagogía Empírica.** Madrid, Escuela Española.

GRONLUND, N.: (1971) **Medición y evaluación en la enseñanza.** México, Pax-México.

GRONLUND, N. (1974) **Elaboración de test de aprovechamiento.** México, Trillas.

MARÍN IBÁÑEZ, R.(1971). **Las pruebas objetivas,** Universidad de Valencia, Facultad de Filosofía y Letras, Valencia.

RODRÍGUEZ GÓMEZ, G., GARCÍA JIMÉNEZ, E. y GIL FLORES, J.: Recursos tecnológicos en la investigación educativa **Pixel-Bit. Revista de Medios y Educación, 6.** 55-75.

