ALGUNAS OBSERVACIONES SOBRE LOS RESULTADOS DE LOS EXÁMENES DELE (INICIAL E INTERMEDIO) A PARTIR DE PRUEBAS SIMULADAS

José Eugenio Borao Mateo

Universidad Nacional de Taiwán, Taiwán, R.O.C.

Resumen

La creciente importancia de los exámenes DELE ha hecho que también se expanda por los países orientales, como en Taiwan, en donde en la actualidad existen varios centros examinadores. Esta situación nos ha hecho preguntarnos en primer lugar por el significado de sus resultados, y por la equivalencia que éstos puedan tener con otros exámenes semejantes (como es el caso del FLPT, administrado en Taiwan por el LTTC). Para resolver esta pregunta hemos pasado varios de estos exámenes a diversos grupos homogéneos de estudiantes en un tiempo muy cercano para hacer algunas comparaciones, y tratar de extraer algunas conclusiones. Una primera observación de los resultados indica que el ránking de alumnos resultante tras un examen DELE Inicial y DELE Intermedio no posee una correlación suficientemente significativa. Por el contrario, la significancia sí que es relevante comparando los resultados del DELE Inicial con el examen FLPT. Pero ¿qué valor pueden tener éstas y otras observaciones? Para poder establecerlo habremos de diferenciar en primer lugar cuáles son los grupos de personas que se posicionan frente a estos exámenes y cuál es su interés por los mismos.

1. Los diversos sujetos de interés frente a los exámenes de E/LE

Pensamos que hay cuatro principales sujetos de interés frente a los exámenes de lengua: alumnos y centros examinadores, centros de enseñanza, administraciones y profesores, teniendo cada uno su propio interés (aunque, a veces parcialmente compartido) frente a dichos exámenes.

Como decimos, en primer lugar está el interés del "estudiante" que desea conocer su nivel desde un punto de vista objetivo, pues ello puede abrirle oportunidades educacionales o laborales. Su interés coincide en cierta medida

ACTAS XLIII (AEPE). José Eugenio BORAU MATEO. Evaluación de los exámenes DELE (i...
Centro Virtual Cervantes

con el de los centros de emisores de exámenes que se esmeran por ofrecer a los aprendientes dicho servicio. Posiblemente estos dos grupos, estudiantes y centros evaluadores, son los que más dinamizan la actividad evaluadora, pues aquéllos se necesitan mutuamente, y ésta funciona bajo las características del mercado, es decir, con oferta y demanda de un servicio, algo particularmente proverbial en los países orientales de tradición confuciana, en donde el ascenso social pasa por la criba de los exámenes. Como estos exámenes están más orientados al mercado laboral o educacional son más sumativos que formativos, y el nivel que se evalúa es independiente de la historia personal del aprendiente. El método de trabajo para ambos grupos (estudiantes y centros) es tradicional: preparación, ejecución de un examen y obtención/entrega de un resultado más o menos especificativo. Es conocido por todos, que actualmente, los exámenes DELE administrados por el Instituto Cervantes son los que mejor cumplen esta función, o al menos, los que dominan el mercado internacional de exámenes de castellano.

En segundo interés a considerar es el de los "centros de enseñanza", pues dichos exámenes pueden servirles como autoevaluación de su oferta didáctica. El método que utilizarán para ello es sencillo: observar los resultados de sus alumnos, tanto los generales (número real de aprobados), como los parciales, es decir los de las notas obtenidas en cada una de las destrezas evaluadas, conociendo de este modo cuáles son sus puntos fuertes y débiles. Naturalmente el método de observación debe tener algo de disciplina, y los resultados obtenidos deberían observarse bien dentro de un grupo de alumnos o comparativamente entre grupos[1]. Un ejemplo en Taiwan de una investigación en este sentido fue llevada a cabo por las universidades Wenzao y Tamkang (Sun, 2004).

Un tercer punto de vista a la hora de acercarse a los exámenes de lenguas es el interés de las "administraciones" cuando están interesadas en conocer el rendimiento de las escuelas. Pero no sólo la administración de un ministerio o consejería municipal, sino también la propia de las universidades, pues ésta puede mostrar interés por dichos niveles a la hora de firmar convenios de cooperación, de traslado de expedientes o de establecimiento de niveles que permita el acceso directo a niveles de enseñanza. En este caso el método de trabajo para establecer equivalencias pasa inevitablemente por la comparación de resultados en departamentos afines de español, es decir, observando el resultado de un mismo examen DELE en grupos de alumnos homogéneos de centros diferentes. El problema de este análisis es que puede generarse una situación comparativa incómoda, y que los centros muestren una cierta reserva de compartir o hacer públicos sus datos. Un ejemplo del interés por este punto de vista es el estudio que realizamos en nuestra universidad para identificar la

[1] Hay un claro problema metodológico y es que las notas que entrega el Instituto Cervantes no son inmediatas, y se entregan sólo a los examinandos. Conseguir, pues, estos resultados por parte de los centros no es una tarea fácil, y en algún caso imposible. Resultaría más fácil si el centro de enseñanza colaborara con el Instituto Cervantes en la administración de los exámenes, o si los examinandos fueran a continuar al menos un semestre más afiliados al centro en el que estudiaban cuando se examinaron. Esto es lo que, ocurre por ejemplo en la Escuela Universitaria Wenzao (Taiwan), y así lo presentaremos después al comentar un estudio que hicieron con sus propios alumnos en 2004.

ACTAS XLIII (AEPE). José Eugenio BORAU MATEO. Evaluación de los exámenes DELE (i...
Centro Virtual Cervantes

nota del DELE Inicial que permitiera el acceso directo a las clases de Español Avanzado (Borao, 2008).

En cuarto lugar tenemos el interés o punto de vista de los "profesores", o el de cualquier persona a la que le afecte el examen, singularmente las personas que confeccionan los exámenes, que sin duda buscan ser lo más objetivas posible. Tal vez la objetividad sea la parte más difícil de evaluar pues el método de trabajo presupone que un mismo alumno (o grupos de alumnos) haya de tomar dos exámenes diferentes, utilizando uno de ellos como paradigma. Aunque no conocemos estudios específicos al respecto sin duda debe haberlos, y de hecho algunos investigadores alemanes ya han señalado las dificultades de objetivación de este tipo de exámenes[2]. El presente trabajo se orienta a este tipo de análisis de la objetividad.

2. El origen de nuestra investigación

El presente trabajo tuvo su origen a principios del curso 2007-2008 en el que de alguna manera nos planteamos en la Sección de Español del Departamento de Lenguas Extanjeras de la Universidad Nacional de Taiwan resolver dos cuestiones.

La primera de ellas era la de conocer mejor el nivel general de nuestros alumnos de una manera homologable. En particular nos proponíamos ver si los de Español Avanzado (que se incorporan a esta asignatura con 240 horas de clase como mínimo) estaban en condiciones de tomar oficialmente el examen DELE Intermedio con garantías de éxito, o si, por el contrario, debían ser aconsejados para tomar el DELE Inicial (nivel del que se suponía sí podría aprobar la mayoría). El método a seguir para saberlo era en sí muy sencillo, pues había que familiarizar un poco a los casi 25 alumnos de este nivel (Español Avanzado) con dichos exámenes y pasárselos después de modo simulado (a partir de los modelos existentes en la página web del Instituto Cervantes) y analizar los resultados. Aunque –tras preparar cuatro modelos de DELE Inicial y uno de DELE Intermedio— el número de aprobados en el examen simulado de DELE Intermedio fue del 23%, en realidad nadie se atrevió a tomar oficialmente este examen, y sólo media docena de ellos tomaron el examen DELE Inicial en abril de 2008. Esta prueba simulada nos permitió conocer que al menos una cuarta parte de los alumnos (en realidad más[3]) podría ser aconsejada positivamente a realizar dicho examen. Naturalmente, una vez que los alumnos examinados en el DELE Inicial nos faciliten sus resultados podremos conocer mejor nuestra capacidad predictiva en las simulaciones.[4]

La segunda de nuestras cuestiones al inicio de nuestra investigación era la de encontrar una equivalencia del nivel de nuestros alumnos cuando acaban sus

[2] Angelika Loo, *Teaching and Learning Modern Languages in Large Classes*, Shaker Verlag, 2007.
[3] El 23% podría fácilmente haberse ampliado al 41% ya que hubo algunos alumnos que suspendieron con nota alta, es decir, superior a la de otros aprobados.
[4] En el momento de revisar el presente texto (principios de septiembre de 2008) sólo he podido conocer el resultado de tres alumnos. Dos de ellos mejoraron sensiblemente los resultados, mientras que otro los mantuvo exactamente idénticos. Es decir, en base a estos mínimos datos la simulación realizada parece algo fiable.

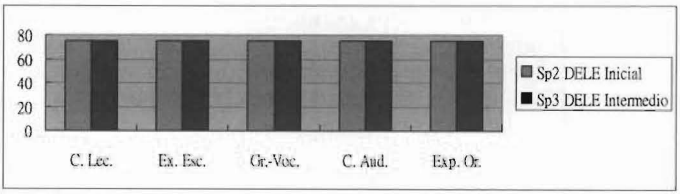
ACTAS XLIII (AEPE). José Eugenio BORAU MATEO. Evaluación de los exámenes DELE (i...
Centro Virtual Cervantes

cursos de Español I y Español II con los niveles del DELE Inicial, asi como del FLPT, un examen de lenguas realizado en Taiwan, que tiene una cierta oficialidad. La principal motivación de ello era la de poder situar en su nivel correspondiente a aquellos posibles alumnos que vinieran con un certificado de alguno de estos exámenes. El método para conseguir esta información también era sencillo. Bastaba con pasar a finales de curso a los grupos de Español I y de Español II el examen DELE Inicial de modo simulado, o bien de realizar de modo oficial el examen FLPT[5]. Así se podría obtener una nota referencial, que nos diera un valor representativo de los alumnos aprobados, para lo cual seleccionamos arbitrariamente la nota media. En realidad, al final se excluyó a los alumnos de Español I del experimento, por considerar que para muchos de ellos podría ser difícil y desanimante. Así pues, llegamos a la conclusión de que, tratándose de un alumno que acaba su curso de Español II (240 horas), la nota media de DELE Inicial correspondía a 67 puntos, y en el caso del FLPT a 45 puntos.

Una vez acabados estos experimentos nos dimos cuenta de que disponíamos de una base de datos que podrían servir para llegar a otras conclusiones. En realidad se trataba de una base de datos imperfecta, ya que no había sido diseñada para análisis más complejos; aún así, poseía una originalidad, y era el hecho de que unos mismos alumnos habían tomado exámenes diferentes y eso podría permitir obtener conclusiones –necesariamente provisionales— acerca de los exámenes propiamente dichos.

Punto de partida

Antes de hacer comparaciones valga la pena señalar dos posibles modelos teóricos e ideales de comportamiento de los alumnos en los exámenes. El modelo I establecería que dos grupos progresivos de alumnos, por ejemplo, los que estudian Espanol II (Sp2) y Espanol Avanzado (Sp3), en condiciones ideales de estudio, y puestos enfrente de exámenes igualmente progresivos, deberían dar unos resultados globales más o menos semejantes entre sí y entre cada una de las destrezas. La situación paradigmática podría ejemplificarse idealmente del siguiente modo:

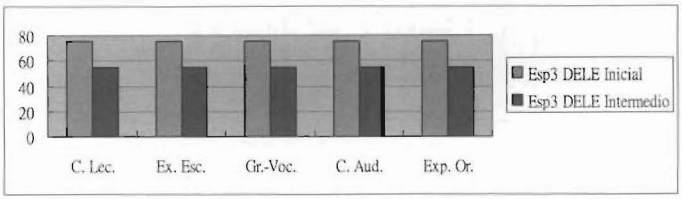


A su vez tendríamos el modelo 2, en el cual el mismo grupo de alumnos (por ejemplo Esp3) examinado en dos exámenes progresivos diferentes (por ejemplo, DELE Inicial y DELE Intermedio) en dos periodos de tiempo más o menos cercano, y con un diseño perfecto de los exámenes, deberían dar unos

[5] Para ello llevamos a cabo un convenio de investigación con el LTTC, la entidad administradora del mismo.

ACTAS XLIII (AEPE). José Eugenio BORAU MATEO. Evaluación de los exámenes DELE (i...
Centro Virtual Cervantes

resultados igualmente escalonados, no sólo en general sino en cada una de las destrezas en particular. La siguiente gráfica ejemplificaría la situación paradigmática ideal:



Consideramos pues estos dos paradigmas como punto de referencia en nuestro posterior análisis de nuestra universidad. Pero antes de ello haremos algunas observaciones al estudio comparativo mencionado entre Wenzao y Tamkang, orientado a informar al segundo tipo de interés, arriba citado.

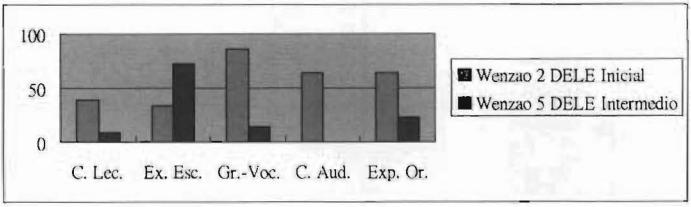
3. Observación de los resultados de los exámenes DELE en los Departamentos de Español de Wenzao y Tamkang desde el punto de vista del interés (II) del centro de lenguas

En el curso 2003-2004 el Colegio Universitario Wenzao (Taiwan), entidad que además colabora con el Instituto Cervantes en la administración del examen DELE en Taiwan, llevó a cabo un estudio de resultados de DELE en colaboración con el Departamento de Español de la Universidad Tamkang (Sun, 2004), con la finalidad de llegar a conclusiones que pudieran servir para mejorar el diseño curricular del centro. Es decir, estábamos ante un estudio de evaluación del segundo tipo de interés que mencionábamos arriba, el de introducir reformas para mejorar la calidad de enseñanza.

Para su análisis, el centro eligió a dos grupos de alumnos, de segundo año y de quinto año a los que se les pasó los exámenes DELE Inicial e Intermedio respectivamente. Cada grupo estaba formado por 36 alumnos seleccionados de un total correspondiente a 101 (segundo año) y a 95 (quinto año). El estudio no se fijó tanto en los resultados finales de apto o no apto, sino en los resultados parciales que miden las destrezas básicas. El examen fue realizado simultáneamente con la Universidad Tamkang, pues de este modo contrastivo es más fácil identificar líneas de mejora. No es del caso mencionar ahora las conclusiones a las que llegaron, sino que –para el caso que nos ocupa—vimos que otro tipo de comparaciones eran posibles. Por ejemplo, la de la respuesta de los alumnos de segundo año evaluados por el DELE Inicial, comparada con la respuesta de los alumnos de quinto año del mismo centro examinados por el DELE Intermedio. Aunque se trataba de dos grupos de alumnos diferentes y de dos exámenes diferentes, ambos pertenecían al mismo centro docente y al mismo centro emisor de exámenes, y se situaban en niveles predeciblemente homogéneos, pues nivel de enseñanza y nivel de aprendizaje habrían crecido paralelamente. Es decir, cabría esperar una homogeneidad en las proprociones de los cinco parámetos. Sin embargo, la comparación ofrecía los siguientes resultados:

ACTAS XLIII (AEPE). José Eugenio BORAU MATEO. Evaluación de los exámenes DELE (i...
Centro Virtual Cervantes

Gráfica 1: Comparación de los exámenes DELE Inicial e Intermedio en Wenzao

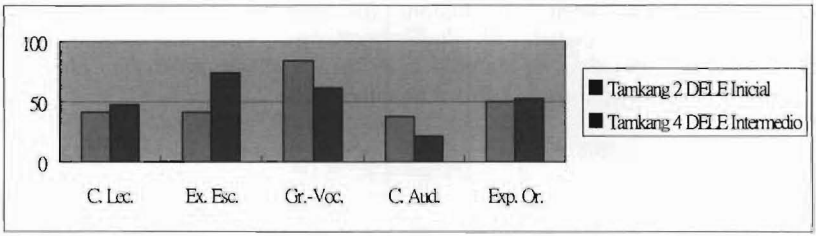


Fuente: Elaboración propia a partir de los datos de Sun (2003)

Ciertamente una cierta homogeneidad existe en cuatro de los parámeros, aunque en unos casos (Comprensión Lectora y Expresión Oral) se muestra de modo más radical que en otros (Gramática-Vocabulario Comprensión Auditiva). Sin embargo, esa regularidad inversamente proporcional en el quinto de los parámetros, la Expresión Escrita. Por consiguiente, tal vez éste pudiera ser un factor de distorsión en los exámenes.

Lo mismo podía hacerse con los dos grupos de estudiantes de la Universidad Tamkang, de segundo año (DELE Inicial) y cuarto año (DELE Intermedio), que habían tomado el examen a la vez que los de Wenzao. Estos eran los resultados comparativos:

Gráfica 2: Comparación de los exámenes DELE Inicial e Intermedio en Tamkang



Aquí los resultados de los parámetros 2, 3 y 4 mantienen la homogeneidad del caso anterior, aunque de forma no tan dispar. Por el contrario, los parametros 1 y 5 son diferentes al caso de Wenzao, aunque muestran lo que debería ser una situación ideal de compenetración entre escuela y centro externo evaluador: valores iguales próximos a la nota media.

¿Cómo interpretar ambas gráficas? Extraer conclusiones es la parte más difícil, pero si consideramos que las irregularidades predominan sobre las regularidades; y si se adopta la presuposición de que los exámenes del DELE tienen una progresión bien diseñada, entonces la explicación de la disparidad habría que buscarla dentro de la institución de enseñanza, y suponer que los alumnos del centro que llevó a cabo la evaluación no tienen un progreso homogéneo en las diversas destrezas de la lengua. Por el contrario, si se presupone que el centro educativo sí posee una progresión regular, entonces el examen del DELE no habría sido suficientemente representativo, al menos para este grupo de aprendientes.

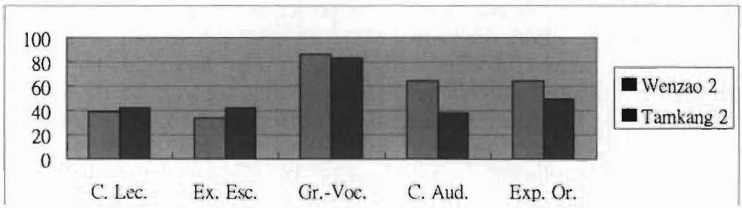
ACTAS XLIII (AEPE). José Eugenio BORAU MATEO. Evaluación de los exámenes DELE (i...
Centro Virtual Cervantes

Tal vez por el hecho de la importancia y prestigio que van adquiriendo los exámenes DELE y por el hecho de que el Colegio Universitario Wenzao sea administrador de los mismos, una de las conclusiones del artículo de Sun era que los profesores de niveles inferiores deberían prestar más atención a la Expresión Escrita y a la Comprensión Lectora (Sun, 2003: 23). Pero ¿hay que "autoinculparse" cuando unos resultados parciales aparecen supuestamente deficientes? ¿Es mejor la evaluación sumativa que la formativa para un centro universitario? ¿Cuál es la validez de los exámenes DELE?

4. Análisis de los resultados de los exámenes DELE en Wenzao College y Tamkang University desde el punto de vista de interés (IV) de los profesores

La interpretación de resultados siempre será un fenómeno complejo por la serie de variables que intervienen en cada centro, pero es muy útil para el día a día de la mejora de la institución. Más me ha interesado en este trabajo prestar atención a las posibles disparidades internas que pueda haber dentro de los exámenes DELE, y que ya han quedado apuntadas en el apartado anterior. Veamos ahora los resultados de DELE Inicial en grupos afines de estas dos escuelas:

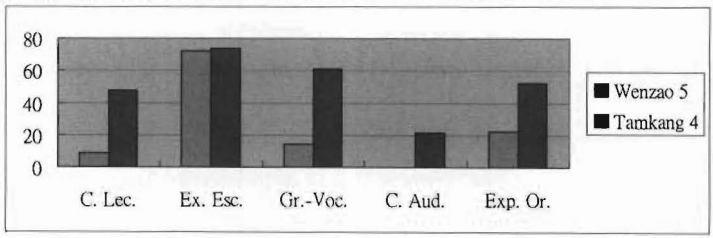
Gráfica 3: Comparación de los exámenes DELE Inicial en Wenzao y en Tamkang



Aquí podemos ver cómo los resultados en ambas escuelas son próximos, reflejando aspectos menores, como que en una están mejor preparados para la Expresión Escrita, mientras que en otra lo están para la Comprensión Auditiva. Pero, se observa además con cierta sorpresa que una de las pruebas, Gramática y Vocabulario, destaca por sus resultados positivos. ¿Se debe esto a un excesivo énfasis en la enseñanza2 de la lengua en dichas instituciones, o a una falta de rigor en el diseño de este particular examen?

Sin embargo, al mirar ahora cuáles son los resultados para el DELE Intermedio se observa una gran disparidad de valores entre cuatro de los parámetros y el segundo de ellos:

Gráfica 4: Comparación de los exámenes DELE Intermedio en Wenzao y en Tamkang



Fuente: Elaboración propia a partir de los datos de Sun (2003)

ACTAS XLIII (AEPE). José Eugenio BORAU MATEO. Evaluación de los exámenes DELE (i...
Centro Virtual Cervantes

Estos datos nos llevan a dos consideraciones. Por un lado, en cuatro de los valores (el primero y los tres últimos) la proporción de la disparidad se mantiene de modo indicando que el nivel de los alumnos de una escuela está en apariencia homogéneamente por encima del de la otra[6]. Esto parece que nos llevaría a hablar en favor del diseño del examen DELE Intermedio, pero nuevamente vemos que no. Por un lado las diferencias no son coherentes, pues van desde los 25 puntos en Expresión Oral a los 45 de Gramática y Vocabulario. Y lo más importante es que el quinto valor, el de Expresión Escrita es idéntico en ambos casos.

Por otro lado, el hecho de que la alta concordancia que existía en el DELE Inicial en el parámetro de Gramática y Vocabulario se traslade a una alta disonancia en el DELE Intermedio (con notas bajas en Wenzao) nos lleva a la sospecha de que los exámenes de Gramática y Vocabulario resultan proporcionalmente, y en relación a los otros parámetros, más fáciles en el DELE Inicial que en el DELE Intermedio. Otra suposición podría ser que el DELE Inicial evalúa más objetivamente a los alumnos que el DELE Intermedio.

6. Comparación de resultados DELE y FLPT en NTU y en NCCU

Intentemos ahora una nueva comprobación de los exámenes DELE, en la que incluiremos un examen realizado específicamente en Taiwan, el FLPT. Empezaremos haciendo obervaciones semejantes a las realizadas en en/entre Wenzao y Tamkang, ahora en/entre NTU (National Taiwan University) y NCCU (National Cheng Chi University). Estas dos universidades no tienen Departamento de Español, pero la primera tiene una sección de español dentro del Departamento de Lenguas Extranjeras que ofrece un programa compacto de esta lengua. La segunda universidad inició hace un par de años un sólido programa de estudios europeos, con sección de español, de manera que la enseñanza de este idioma se aproxima bastante al de un departamento. Consideramos como punto de partida (que resultó ser cierto) que los alumnos de Español Avanzado de NTU eran homologables a los de segundo año de NCCU, lo cual vimos que nos permitiría llevar a cabo comparaciones. Los alumnos fueron examinados de acuerdo a las siguientes pruebas:

DELE Inicial (simulado):
- 18 alumnos de Español Avanzado (NTU): 17/03/2008
- 20 alumnos de Español II (NTU): 24/05/2008

DELE Intermedio (simulado): 21/04/2008
- 17 alumnos de Español Avanzado (NTU) (los mismos que en DELE Inicial): 21/04/2008
- 19 de Segundo año del Programa de Estudios Europeos (NCCU): 21/04/2008

FLPT (real):
- 23 alumnos de Español Avanzado (NTU) (los mismos que en los dos casos anteriores): 12/04/2008

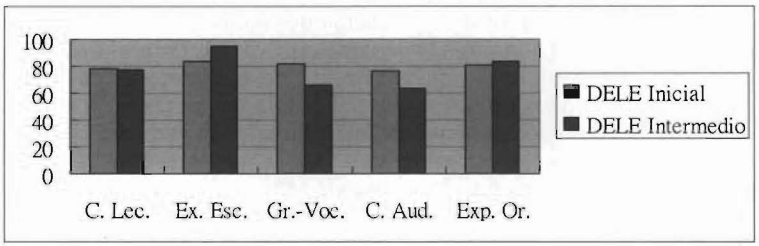
[6] En realidad habría que tener en cuenta el hecho de que un buen número de estudiantes de Tamkang han pasado un año en España; y algunos de los mejores alumnos de Wenzao continúan luego en la Universidad de Tamkang.

ACTAS XLIII (AEPE). José Eugenio BORAU MATEO. Evaluación de los exámenes DELE (i...
Centro Virtual Cervantes

- 20 alumnos de Español II (NTU) (los mismos que en DELE Inicial): 7/06/2008
- 46 alumnos de Segundo Año del Programa de Estudios Europeos (NCCU): 12/04/2008

Una primera comparación que se estableció fue entre los resultados de los alumnos de Español Avanzado que habían obtenido en sus exámenes DELE Inicial (modelo del 24 de noviembre de 2006) y DELE Avanzado (modelo del 10 de noviembre de 2007), tomándose el segundo sólamente seis semanas después del primero. El resultado gráfico fue el siguiente:

Gráfica 5: Comparación de los exámenes DELE Inicial e Intermedio en Español Avanzado (NTU)



Fuente: Resultados de exámenes simulados (NTU)

La gran diferencia con respecto a los exámenes de Tamkang-Wenzao es que aquí los estudiantes son los mismos, por ello se parte de la idea de que el Examen DELE Intermedio ha de ser sustancialmente más difícil que el DELE Inicial [7]. En este caso sólo parecen normales los resultados de "Gramática-Vocabulario" y "Comprensión Auditiva" que descienden en una proporción semejante. Quizás ello sea debido a que esta parte del exámen sea más controlable por el examinador, dado su caracter objetivo. Sin embargo, no es lógico que la "Comprensión Lectora" y la "Expresión Escrita" (juzgadas también de modo objetivo) se mantengan al mismo nivel la primera, o a un nivel superior la segunda. Igual podría decirse de la "Exposición Oral", pero a ello no le damos tanta importancia porque aquí es más fácil que hayan podido influir elementos subjetivos[8].

7. El examen FLPT como paradigna

Otra observación que podemos hacer es la de comparar los dos resultados anteriores (NTU DELE Inicial e Intermedio) con el examen *Foreign Languages Proficiency Test* (FLPT). Este examen es ofrecido desde hace bastantes años por una entidad de larga tradición en Taiwan, el *Language Training and Testing Institute* (LTTC). Se trata de exámenes con un alto grado de retroalimentación

[7] Más aún si consideramos que el alumno no ha podido tener un progreso alto en las seis semanas que separa un examen de otro, y su única preparación había sido un examen de DELE Intermedio y cuatro de DELE Inicial.

[8] La sorpresa podría ser mayor si comparamos los valores similares para los casos de Wenzao y Tamkang. Los parámetros a comparar no son idénticos, pero las dos escuelas muestran una mayor capacidad de respuesta en la Expresión Escrita en los alumnos de cursos superiores, en comparación con los de cursos inferiores. Tal vez por ello, los exámenes de DELE Intermedio deberían ser más exigentes en esta prueba.

ACTAS XLIII (AEPE). José Eugenio BORAU MATEO. Evaluación de los exámenes DELE (i...
Centro Virtual Cervantes

en el sentido de que los resultados se analizan para ver qué preguntas contienen elementos viciados que no sean suficientemente discriminantes[9]. Al no ser públicos, se pueden analizar y mejorar con regularidad. Además están planteadas para un público local, por profesores locales. Por tanto, es de esperar que el resultado clasifique bien a los estudiantes. Comparar aquí los exámenes DELE y FLPT es más complejo ya que las notas que se ofrecen no se corresponden de modo idéntico, pero sí podemos observar el ranking en el que se sitúa cada alumno en ambos tipos de examen, y ver si hay una correspondencia. Además, al tratarse de muestras de pocos alumnos se puede ver el comportamiento de cada caso particular, ofreciendo una radiografía de la situación más clara que unos datos estadísticos.

La tabla que presentamos a continuación nos indica claramente el modo en que cada alumno se posicionó tras los exámenes. Nos muestra que al comparar los exámenes FLPT con DELE Inicial (flechas de la izquierda) el ránking de estudiantes es similar, pues las líneas que indican variación pequeña son las que predominan. A su vez si comparamos el DELE Intermedio con el FLPT (flechas de la derecha), vemos mayor irregularidad, por la mayor presencia de fuertes subidas y bajadas de ránking. Comparando ahora DELE Inicial y DELE Intermedio (flechas del centro) las irregularidades son nuevamente patentes, habiendo pocos alumnos que conserven más o menos su ranking inicial.

Tabla 1: Ranking de estudiantes tras los exámenes FLPT y DELE (Inicial e Intermedio)

FLPT		DELE INICIAL		DELE INTERMEDIO			FLPT	
		Edu	92,61	1	Jua	90,63		
San	70,67	Cla	88,84	2	Sar	90,13	San	70,67
Cla	68,33	Mer	88,48				Cla	68,33
Edu	67,33	Jua	87,26	3	Gui	87,35	Edu	67,33
Sar	66,33	Sar	86,83	4	Cel	85,53	Sar	66,33
Mer	64,33	Gui	86,27	5	Ben	84,05	Mer	64,33
Gui	63	Nur	85,02	6	Nur	84,03	Gui	63
Ben	62,33	Ben	81,39				Ben	62,33
Cel	60,33	Glo	81,18				Cel	60,33
Car	56,33	Cel	80,88	7	Edu	80,43	Car	56,33
Elv	49,33	San	78,55	8	Sof	78,92	Elv	49,33
Glo	48	Ven	76,73	9	Ant	76,62	Glo	48
Son	46,67	Car	76,35	10	Cla	75,12	Son	46,67
Sof	46	Elv	72,38	11	Nie	73,88	Sof	46
Nie	44,67	Sof	71,2	12	Glo	73,52	Nie	44,67
Ali	42	Ant	69,2	13	Ven	71,38	Ali	42
Jua	38,67	Jua	67,36	14	Jua	69,5	Jua	38,67
Ven	36,67	Ali	66,65	15	Car	67,98	Ven	36,67
Ant	35	Son	63,67	16	Mer	62,33	Ant	35
				17	Ali	57,92		

Elaboración propia a partir de los resultados de los exámenes

[9] Por ejemplo, se estudia si en las preguntas hay correspondencia entre los aciertos o fallos y el nivel de los examinandos. Así una pregunta en la que los alumnos de nivel supuestamente alto obtuvieran resultados semejantes a los del nivel supuestamente bajo estaría mal planteada.

ACTAS XLIII (AEPE). José Eugenio BORAU MATEO. Evaluación de los exámenes DELE (i...
Centro Virtual Cervantes

En otras palabras, si el DELE Intermedio es el único que presenta irregularidades de ranking en ambas comparaciones, es de suponer que el diseño de su examen sea más problemático.

Esta divergencia nos lleva adicionalmente a ponderar nuevamente la relación entre los exámenes locales, ya existentes como el FLPT, y los internacionales como el DELE. En el congreso de la Lengua Española celebrado en 2004 en Rosario (Argentina), Rosa Filipchuk, responsable del centro de examen DELE de la Universidad Ricardo Palma (Perú), advertía de la multiplicación de certificaciones expedidas por las universidades e instituciones académicas que acreditan la competencia de la lengua española, amparadas en distintos dispositivos legales, para lo cual proponía que una solución a ello podría ser el examen DELE, ya que se trata –decía— de una certificación eminentemente técnica, fiable y de reconocimiento internacional, que, año tras año, extiende su área de acción a nivel mundial. Pero también añadía que era necesario un acuerdo entre instituciones académicas de América Latina y el Instituto Cervantes con el objetivo de unificar criterios del español de América Latina frente al modelo peninsular y señalar un "reconocimiento mutuo de las certificaciones de conocimientos de la lengua española"[10]. Esto, hoy por hoy, sólo nos suscita preguntas. ¿Es posible esa homologación sin que su significado se difumine? ¿Estaría el Instituto Cervantes dispuesto a compartir parte de su autoridad con otras instituciones? ¿Entraría el FLPT de Taiwán a formar parte de ese reconocimiento mutuo?

Conclusiones

Aunque lo más claro que podría concluirse es la idea general de que el proceso evaluador es una tarea ardua[11] vamos a intentar precisar algunas conclusiones más específicas.

Una primera conclusión es que la distancia que separa el examen DELE Inicial del DELE Intermedio, no es tan grande como pudiera pensarse inicialmente, e incluso una misma persona hasta podría obtener notas muy cercanas en ambos exámenes (e incluso superiores en el examen más difícil de los dos, al menos tal como quedó expuesto en el caso particular de la tabla 1).[12]

---

[10] Rosa Filipchuk de Romero, "Hacia una certificación concertada", *Actas del III Congreso Internacional de la Lengua Española*, Rosario, noviembre de 2004.

[11] Sobre esta afirmación considero de gran importancia las discusiones habidas en el citado congreso de Rosario (2004). Allí se dedicó una sesión a la certificación de la competencia lingüística en E/LE, reconociéndose las dificultades técnicas que entrañaba el llevar a cabo los exámenes del DELE, especialmente la falta de formación de los examinadores. Véase Neus Figueras Casanovas, "Estándares y calidad en la elaboración y administración de pruebas y exámenes. Criterios mínimos para el reconocimiento y la comparabilidad", *Actas del III Congreso Internacional de la Lengua Española*, Rosario, noviembre de 2004.

[12] No estamos en condiciones de analizar la razón de ello, pero tal vez podría ser que, por ejemplo, la parte de Comprensión Escrita del DELE Intermedio en que los alumnos tienden a obtener notas altas, en realidad corresponde a sólo 12 preguntas. Las tres primeras (texto 1) se contestan a través de la fórmula verdadero o falso; y en las ocho restantes (textos 2 a 4) hay que elegir una de las tres opciones que se presentan. Si consideramos que para esta parte hay un tiempo moderadamente suficiente, el resultado puede reflejar una situación más favorable de la capacidad real del alumno. Además el tiempo puede rentabilizarse mucho más, si los alumnos son conscientes de que por regla general cada una de las tres preguntas se refiere ordenadamente al contenido de cada uno de los tres párrafos en que suele estar dividido el texto.

ACTAS XLIII (AEPE). José Eugenio BORAU MATEO. Evaluación de los exámenes DELE (i...
Centro Virtual Cervantes

Esta conclusión nos lleva a la pregunta de si los resultados podrían o no estar viciados por los examinandos. Algunos de los exámenes –DELE Intermedio y FLPT (no pudimos incluir DELE Inicial)— fueron hechos por dos grupos diferentes pero considerados de nivel semejante, y el resultado fue muy parecido en ambos grupos, es decir, que una segunda conclusión sería que la falta de homogeneidad en los resultados podría depender de la confección de los exámenes, y no necesariamente ser achacada a los examinandos. Acerca de si se podría elegir o no un examen como variable fija sobre el que comparar los otros, creemos que el FLPT tenía más garantías de ser un examen que refleje mejor el nivel de los alumnos, ya que tiene más años de experiencia, no es tan diversificado culturalemente (es más homogéneo en cuanto que predomina más el español peninsular), y sobre todo, porque los exámenes no son públicos lo que permite modificar las preguntas que se consideran viciadas, es decir que no son suficientemente discriminantes. Estas ideas nos preparan el camino de la tercera conclusión y es que la comparación del FLPT con el DELE Inicial resultó ser mucho más significativa que con el DELE Intermedio.

Observando las gráficas anteriores, y viendo el caso de que hay alumnos que obtienen muy buena nota en el DELE Intermedio, pero que no aprueban por errores en la sección de Gramática, nos hace pensar como cuarta conclusión que la parte de Gramatica y Vocabulario es la menos equilibrada, de modo que en el DELE Inicial es la parte más fácil, mientras que en el DELE Intermedio es la más difícil. Lo citado nos llevaría a pensar en una posible quinta conclusión, y es que a medida que los estudiantes progresan en su conocimiento de la lengua es más difícil encontrar sistemas objetivos para evaluarles. Esto, sin mencionar otro factor que pudimos constatar en preguntas en particular a alumnos tras conocer el resultado de sus exámenes, y es que hay circunstancias personales que pueden hacer imprevisibles algunos resultados (Borau: 2008).

BIBLIOGRAFÍA

BORAU MATEO, José Eugenio (2008), "La medición del progreso de los estudiantes de lenguas extranjeras". En: *Actas del IX Simposio sobre la didáctica, la cultura y la traducción del español*, Universidad de Tamkang, pp. 1-14.

ALDERSON, Charles J.; CLAPHAM, Caroline; WALL, Dianne (1998), *Exámenes de idiomas. Elaboración y evaluación*, Cambridge University Press.

CÁRDENAS BERNAL, Francisca, et al. (2006), "Evaluación". En: *La enseñanza del español como lengua extranjera a la luz del Marco Común Europeo de Referencia. Diseño Curricular para Extranjeros de la Universidad de Málaga*, Universidad de Málaga, pp. 157-166.

CUSHING WEIGLE, Sara (2002), *Assessing Writing*, Cambridge University Press.

DEAN BROWN, James (1996). Testing in Language Programs : A Comprehensive Guide To English Language Assessment, London.

HERNÁNDEZ, M. Rosario (2000), "La evaluación como estrategia docente en el desarrollo de la competencia escritora en E/LE". En: Actas del XI Congreso Internacional de ASELE, Universidad de Zaragoza, pp. 431-437.

LADO, Robert (1961), *Language Testing. The Construction and Use of Foreign Language Tests*, Longman.

MCNAMARA, T. F. (1996), *Measuring Second Language Performance*, Longman.

REA-DICKINS, Pauline & GERMAINE, Kevin (1992). *Evaluation*, Oxford University Press.

SUN, Su-ching (2004), "Importancia de la evaluación en la enseñanza del español: El DELE y sus resultados". En: *Actas del VII Congreso de Didáctica del Español en la República de China*, Universidad de Tamkang, pp.13-36.

VALCÁRCEL PÉREZ, María Soledad & VERDÚ JORDÁ, Mercedes (1988), *Observación y evaluación de la enseñanza comunicativa de lenguas modernas*, Ministerio de Educación y Ciencia, Madrid.

ACTAS XLIII (AEPE). José Eugenio BORAU MATEO. Evaluación de los exámenes DELE (i...
Centro Virtual Cervantes