

INVESTIGACIONES Y EXPERIENCIAS

EXAMEN DE SELECTIVIDAD: EL ESTUDIO DEL TRIBUNAL PARALELO

TOMÁS ESCUDERO ESCORZA (*)
CONCEPCIÓN BUENO GARCÍA (*)

Los cambios en el panorama del acceso a la Universidad española, que van teniendo lugar en los últimos tiempos, han ido modificando el énfasis entre las funciones atribuidas a las pruebas de entrada. Tal como indica Muñoz (1992, 6), «... desde la implantación de la limitación de acceso a los centros universitarios, conocida habitualmente como *numerus clausus*, la prueba de acceso a la Universidad pasa a desempeñar, prioritariamente, una función distribuidora de alumnos...». Lo mismo puede decirse del expediente secundario, que junto con la calificación del examen de «selectividad», y con la misma ponderación, conforman el criterio de selección y distribución de los alumnos admitidos entre las distintas carreras y estudios.

En este contexto de competencia, donde a veces se accede a una plaza de estudio deseada por tener unas centésimas más que otro aspirante en el promedio aludido, aumenta, lógicamente, la presión por equiparar el significado de los expedientes secundarios, contruidos a partir de múltiples escalas diferentes, y, sobre todo, por aumentar la objetividad y fiabilidad de los exámenes de «selectividad».

En el presente trabajo describimos con cierta extensión los resultados de una experiencia llevada a cabo en la convocatoria de junio de 1992 de las pruebas de acceso, en la Universidad de Zaragoza, y más en concreto en un tribunal ubicado en la ciudad de Teruel.

La experiencia fue promovida y financiada por el propio Ministerio a través del CIDE, siendo coordinada y realizada por un equipo de miembros del ICE y del Servicio de Alumnos de la Universidad de Zaragoza.

(*) ICE. Universidad de Zaragoza.

El objetivo era el estudio de la fiabilidad entre juicios de dos tribunales paralelos en la prueba de acceso. El juicio de uno de los dos tribunales tuvo, lógicamente, valor legal a todos los efectos y el del otro solamente fue utilizado para la investigación que presentamos.

Los resultados de la experiencia, al menos en parte sustancial, han sido utilizados de manera interna por el Ministerio (CIDE, 1992; Consejo de Universidades, 1992), difundiendo en la prensa algunos resultados concretos (*El País Educación*, 1993). Nosotros, ahora, superado el plazo razonable de discreción alrededor del estudio y toda vez que ya ha sido tenida en cuenta para las recientes reformas de las pruebas de acceso (Orden Ministerial de 9 de junio de 1993, Real Decreto 807/1993), ofrecemos una descripción más completa de la experiencia y de los resultados, con una orientación más académica, especialmente para su conocimiento por parte de estudiosos e investigadores.

1. ALGUNOS ANTECEDENTES

En los últimos tres lustros, aproximadamente, los estudios sobre la prueba de selectividad y el proceso de acceso a nuestra Universidad han sido abundantes, pero no han sido los problemas relacionados con la fiabilidad y la objetividad de la prueba los que han recibido más atención por los investigadores; la validez predictiva de la misma, por ejemplo, ha sido un centro de interés mucho más habitual.

A pesar de lo dicho, también existen ejemplos varios de trabajos en el terreno de la fiabilidad, incidiendo sobre todo en el problema de las diferencias de criterio en los tribunales, los sorteos, las tandas, etc. (Escudero y Bernad, 1978; Escudero, 1984, Sans, 1991, CIDE, 1991*b*). En todos estos trabajos y en otros similares, siempre se suele percibir el factor tribunal como el elemento potencialmente más perturbador de la fiabilidad de la prueba, indicando que el procedimiento debe mejorarse en ese sentido. En esto mismo inciden Muñoz-Repiso y otros (1991), pero indican a la vez, que la prueba en su conjunto es bastante mejor que su imagen externa y que, ni mucho menos, se trata de una lotería.

Los problemas de corrección de las pruebas motivaron que el Ministerio promoviera recientemente un ensayo de utilización de pruebas objetivas en el examen de selectividad, a la búsqueda de una mayor fiabilidad y capacidad de discriminación (CIDE, 1991*a* y 1991*b*). De este trabajo se desprendieron algunos resultados interesantes para el problema que estamos tratando. Así, por ejemplo, se vio que con determinado tipo de preguntas abiertas se pueden obtener altos índices de fiabilidad si se precisan bien los criterios de corrección. También se observó que las pruebas objetivas mejoraban la validez de contenido y la consistencia de las calificaciones, pero se planteaban dificultades para evaluar determinados objetivos y para mejorar sustancialmente la discriminación.

Pero el estudio, dentro de los realizados hasta ahora, más orientado al análisis de la fiabilidad entre juicios de los tribunales de la selectividad, es el que des-

cribimos en este trabajo. En un principio se planteó utilizando dos tribunales paralelos en vez de uno y varias Universidades, pero determinadas circunstancias como la premura de tiempo y las grandes dificultades organizativas condujeron a una experiencia realizada en los términos que presentamos.

Como consecuencia de investigaciones como las señaladas y de la propia presión social alrededor de las pruebas de selectividad, la Administración educativa y las propias Universidades han ido introduciendo durante los últimos años modificaciones con la intención de mejorar la fiabilidad y objetividad de las mismas. Las reformas más sustanciales tuvieron lugar en 1987, a raíz de las importantes protestas estudiantiles contra la selectividad (Orden Ministerial de 3 de septiembre de 1987 y Real Decreto 406/1988).

En estas reformas se eliminaron los sorteos de asignaturas, se potenció la eliminación de las tandas de exámenes, generalizando el llamamiento único y se redujo el número de tribunales, ampliando el número de especialistas. También se suprimió el ejercicio de «la conferencia», por ser demasiado abierto a sesgos de interpretación y de procedimiento.

En distintos momentos se ha estudiado la implantación de tribunal único, dentro de cada Universidad, pero no se ha visto clara su viabilidad, con carácter general, al menos entendida como una sustancial mejora a la situación actual. Esta modalidad organizativa sí que fue adoptada en algún caso, como en la Universidad de Santiago de Compostela (Rodicio, 1993), pero estableciendo dentro del tribunal único «Comisiones de evaluación» por disciplinas y «Comisiones delegadas» por grupos de centros. En la práctica, entendemos, que esta nueva organización no reduce de manera significativa el problema de la multiplicidad de escalas de calificación.

Las recientes reformas (Real Decreto 807/1993 y Orden Ministerial de 9 de junio de 1993) inciden sobre todo en la homogeneización y mejora de los criterios de corrección, el tipo de preguntas, la participación de especialistas en los tribunales y el número de ejercicios a corregir por cada miembro. Se sigue incidiendo, por tanto, en aspectos citados por los investigadores como fuentes de problemas para la objetividad y fiabilidad.

Naturalmente, y tal como señala Muñoz (1992), estas mejoras en los aspectos métricos aludidos, conllevan algunos riesgos o problemas en otros. Así, por ejemplo, con el tiempo se ha limitado bastante la posibilidad de evaluar la capacidad de expresión y de interpretación libre de los alumnos a lo largo de la prueba de selectividad en conjunto.

2. EL TRIBUNAL PARALELO (*)

Unas semanas antes de la realización de las pruebas de acceso en junio de 1992, el CIDE solicitó a la Universidad de Zaragoza el establecimiento de dos tribunales paralelos a uno de los tribunales de selectividad, con vistas a estudiar las diferencias de corrección entre ellos.

Según las previsiones ministeriales, la experiencia debía llevarse a cabo en cuatro o cinco Universidades de diversos tamaños y contextos, pero la premura de tiempo y algunas otras dificultades hicieron que el estudio se realizara sólo en la Universidad de Zaragoza y con un único tribunal paralelo y no con dos como era el propósito inicial.

Tras las adecuadas negociaciones entre el CIDE y el Vicerrectorado de Ordenación Académica de la Universidad de Zaragoza, el Servicio de Alumnos y el ICE de esta Universidad diseñaron, organizaron y llevaron a cabo la experiencia que describimos en este trabajo.

En la citada convocatoria, la Universidad de Zaragoza tenía doce tribunales de selectividad, por tanto, el tribunal paralelo sería el decimotercero. Para realizar la experiencia se eligió, por razones meramente técnicas y de control científico-metodológico, el décimo tribunal que actuaba en la ciudad de Teruel. Esta denominación inicial la mantenemos en nuestro trabajo y, a lo largo del mismo, seguimos hablando de los tribunales 10 (el oficial de selectividad) y 13 (el paralelo).

Los vocales (correctores) del tribunal 13 fueron elegidos siguiendo los mismos criterios que los restantes tribunales y fueron desconocedores de su característica especial hasta donde lo permitía la lógica del proceso. Lo mismo puede decirse de los miembros del tribunal 10, que fueron desconocedores de la experiencia prácticamente hasta el final de la misma; solamente el presidente fue partícipe y colaboró en la organización y realización del estudio.

Con relación a lo anterior, hay que decir que el sigilo y el control experimental con el que se llevó a cabo este proceso fue de tal naturaleza que, a pesar de la presión social y periodística existente en todo el país sobre ella, nunca se tuvo conocimiento por los medios de que la experiencia se realizaba en Teruel. El control fue total y puede decirse que el proceso realmente no trascendió más allá de los propios organizadores.

Aunque el diseño de la experiencia era metodológicamente muy simple, su realización con las debidas garantías de control de posibles variables perturbado-

(*) El tribunal paralelo fue dirigido por los miembros del ICE Agustín Ubieto (presidente) y Emilio Palacián (secretario). Ellos lógicamente, también participaron en la organización de la experiencia y nos han proporcionado la información del funcionamiento interno del tribunal. El resto de los miembros del tribunal fueron elegidos por los mismos procedimientos que todos los vocales de las pruebas de selectividad.

ras era muy complicada y costosa. Además de formar el tribunal paralelo, con iguales costes que uno convencional, hubo que diseñar un formato de «hojas de ejercicios» que permitiera a la vez todas las garantías de anonimato que prescriben las pruebas de acceso y su fotocopia rápida y completa, para permitir la doble corrección sin interferencias de un tribunal en otro y, además, sin demoras sobre los tiempos normales previstos por la organización general de las pruebas en la Universidad de Zaragoza.

Esta operación fue posible escribiendo sólo por una cara, en un cuadernillo de hojas en el que la primera era de datos de identificación y en cada una de las siguientes páginas estaba impreso el número de página, el nombre de la asignatura y un número de identificación del examen.

Se alquiló en exclusiva una empresa de fotocopias durante tres días y todo el proceso lo realizaron personalmente, incluso trabajando por las noches, el presidente y el secretario del tribunal paralelo, en estrecha colaboración con el presidente del tribunal oficial. La operación fue tremendamente laboriosa, pero no surgieron complicaciones graves, ni extravió ni confusión entre exámenes. Las mayores dificultades se presentaron con las fotocopias de las láminas de dibujo, pero al ser una asignatura optativa con un número de alumnos no muy alto, se pudieron solventar sin retrasar el plan previsto.

Los dos tribunales, el 10 y el 13, corrigieron los ejercicios en el mismo plazo de tiempo, pero no tuvieron ningún contacto entre ellos a lo largo de todo el proceso. Solamente estuvieron en Teruel los citados presidente y secretario del tribunal 13. El tribunal 10 actuó en todo momento como si no existiera el tribunal paralelo y sin ningún tipo de contacto con sus miembros.

Las correcciones del tribunal 10 tuvieron el tratamiento habitual, mientras que las del tribunal 13 se pasaron directamente al equipo coordinador del estudio que, tras recibir los correspondientes al tribunal 10 y organizar debidamente la información, remitió el bloque de datos al CIDE, para la prevista utilización ministerial, tal como hemos señalado con anterioridad.

Cumplida la función prevista y tras el debido período de confidencialidad, ahora podemos analizar y presentar en mayor detalle los resultados de este estudio.

3. MUESTRA, VARIABLES Y TRATAMIENTO DE DATOS

Los alumnos que participaron en nuestro tribunal, en tanda única, fueron 348, de los que 118 eran varones y 230 mujeres.

La distribución por opciones de COU, A (Científico-tecnológica), B (Biomédica), C (Ciencias sociales) y D (Humanística-lingüística) fue 130, 83, 102 y 33 alumnos respectivamente.

Las variables analizadas por sexo, opción de COU y, naturalmente, tribunal (oficial y paralelo) han sido los ocho ejercicios individuales de la prueba de selectividad (texto, lenguaje, filosofía, idioma, primera y segunda obligatoria de la opción y las dos optativas del alumno), la media del expediente académico secundario, la media del primer ejercicio (común \pm las cuatro primeras variables), la media de segundo ejercicio (opcional – las variables 5-8), la media de la prueba en conjunto y la media de las PAU (prueba y expediente).

Además de las trece variables básicas anteriores, también hemos analizado las distribuciones de dos variables diferencia, entre las puntuaciones de cada alumno, según los tribunales 10 y 13, en la prueba de selectividad (global) y en la prueba de acceso (selectividad y expediente).

El tratamiento de estas variables lo hemos llevado a cabo con los programas STATVIEW (Feldman y Gagnon, 1985) y SYSTAC (Wilkinson, 1986). Básicamente nos hemos centrado en el análisis de las distribuciones, frecuencias, rangos, medias, desviaciones típicas, Anova, hemos contrastado las diferencias en las características de las distribuciones de las variables entre tribunales y grupos con el test de Bartlett, X^2 , t y F y hemos calculado las intercorrelaciones entre las variables dentro de cada tribunal y entre tribunales.

En conjunto, las variables analizadas y su consiguiente tratamiento nos permiten estudiar en profundidad las diferencias que han podido producirse entre las calificaciones por efecto de la influencia de los tribunales participantes en la experiencia.

4. RESULTADOS

Los análisis realizados nos han proporcionado una gran cantidad de resultados que, lógicamente, no podemos reproducir en su totalidad. Debemos limitarnos, por tanto, a los resultados más importantes y relevantes para los propósitos básicos del estudio. En primer lugar presentamos los relativos a la comparación directa entre tribunales y, posteriormente, otros análisis diferenciados para cada tribunal, que nos pueden mostrar una panorámica diferente en cada caso.

4.1. *Diferencias entre tribunales*

En la tabla 1 recogemos para cada una de las trece variables básicas de las pruebas de acceso y para cada uno de los tribunales, el rango de puntuaciones y los valores de las medias y las desviaciones típicas. Además, señalamos en otra columna el nivel de significación de la diferencia entre ambos tribunales, caso de que exista, y la correlación entre ambos.

El conjunto de los resultados de la tabla 1 nos destacan el alto poder compensatorio que existe entre las distintas partes de la prueba. A pesar de detectar

TABLA 1

Resultados diferencias y correlaciones entre tribunales

Variable	Tribunal 10			Tribunal 13			Signif. Dif.	π_{10-13}
	\bar{X}	S	Rango	\bar{X}	S	Rango		
Texto (X ₁)	5,307	0,937	1-8,5	4,624	1,249	1-8	**	0,406
Lenguaje (X ₂)	3,434	1,556	0,5-9,5	4,489	1,412	1-9,2	**	0,592
Filosofía (X ₃)	5,364	1,318	1-8	5,324	1,723	0,7-9	-	0,527
Idioma (X ₄)	5,488	1,953	0,5-9,3	5,189	1,371	0,5-9	*	0,587
Obligat. 1 (X ₅)	4,553	2,206	0,2-9,9	4,871	2,256	0,1-9,9	-	0,842
Obligat. 2 (X ₆)	4,952	1,901	0,3-9,5	4,770	1,993	0-9,9	-	0,662
Optativa 1 (X ₇)	4,337	2,096	0-9,9	4,671	1,941	0,3-9,9	*	0,787
Optativa 2 (X ₈)	4,550	1,736	0-8,8	4,776	1,858	0-9,5	-	0,760
Exp. Sec. (X ₉)	6,954	0,763	5,758-8,915	-	-	-	-	-

(Continúa)

algunas diferencias significativas entre los tribunales en algunos ejercicios parciales, en los promedios globales de la prueba y sobre todo en el promedio final con el expediente, los resultados en uno y otro tribunal son muy parecidos en todo y, además, la correlación entre ellos es altísima.

En las pruebas del primer ejercicio, el ejercicio común, salvo en la de Filosofía, encontramos diferencias significativas en los promedios y en las dispersiones de calificaciones y, además, correlaciones relativamente bajas entre los tribunales, incluido el ejercicio de Filosofía. Sin embargo, en el promedio de este ejercicio común [ejercicio 1 (X_{10})], las distribuciones de calificaciones son casi idénticas y la correlación entre ellas supera el valor 0,75. El efecto de compensación observado en nuestro estudio, entre las cuatro pruebas del primer ejercicio de la selectividad, es muy acusado.

En las pruebas del ejercicio 2, el diferenciado según la opción de COU, el panorama es más parecido entre tribunales y las correlaciones son sensiblemente más altas que en el ejercicio 1. Sólo se encuentran diferencias significativas en la Optativa 1, pero con una correlación alta (0,787). A pesar de lo anterior, el promedio en el ejercicio 2 en conjunto es algo más bajo en el tribunal 10 que en el 13, aunque esta diferencia no es significativa y la correlación es muy alta. En este segundo ejercicio, el tribunal paralelo fue, en promedio, algo más benévolo que el tribunal oficial.

En este ejercicio 2, aunque nosotros trabajamos con cuatro calificaciones (cuatro variables) debemos tener presente que corresponden a bastantes asignaturas más, dos obligatorias y cuatro optativas para cada una de las cuatro opciones de COU. Podemos pensar que, probablemente, un análisis diferenciado asignatura por asignatura, nos daría un paralelismo mayor entre tribunales, al menos en algunas asignaturas.

Con relación a los rangos de variación de las calificaciones, hemos de decir que se observa un gran paralelismo entre ambos tribunales.

Si analizamos la frecuencia acumulada de las variables Selectividad (X_{12}) y Acceso (X_{13}) podemos ver si con el tribunal paralelo hubieran aprobado más o menos alumnos que con el tribunal oficial. En concreto, los porcentajes de puntuaciones por debajo de 4 en la selectividad, en los tribunales 10 y 13 fueron, respectivamente, 19,539 y 19,251. Los porcentajes con puntuaciones en el acceso (selectividad y expediente) por debajo de 5 en los tribunales 10 y 13 fueron, respectivamente, 12,931 y 13,793.

Vemos que las diferencias entre los tribunales en cuanto a porcentajes de aptos son escasísimas. Curiosamente, el tribunal oficial, el 10, dejó promediar la prueba con el expediente a un alumno más que lo que hubiera dejado el paralelo, pero con este último tribunal hubieran sido tres alumnos más los que hubieran superado el 5 de promedio entre ambas puntuaciones.

Con el tribunal 10, el oficial, el número de aprobados fue 278, el 79,89 por 100 de los 348 presentados. Con el tribunal 13, el paralelo, esta cifra hubiera sido 279, el 80,17 por 100 de los 348 presentados.

Pero la diferencia no es lógicamente un aprobado más en un tribunal que en otro, puesto que existe alguna discrepancia entre los tribunales en cuanto a quiénes son los aprobados.

En la siguiente tabla aparece el cruce de los aprobados y no en ambos tribunales en números brutos (entre paréntesis los porcentajes correspondientes).

TABLA 2

Número y porcentajes (entre paréntesis) de alumnos aptos y no aptos por uno o por los dos tribunales simultáneamente

		Tribunal 13		
		No apto	Apto	Total
Trib. 10	No apto	53 (15,23)	17 (4,89)	70 (20,11)
	Apto	16 (4,60)	262 (75,29)	278 (79,89)
	Total	69 (19,83)	279 (80,17)	348 (100)

Los datos anteriores nos dicen que 262 alumnos, el 75,29 por 100 de los presentados, habrían aprobado con cualquiera de los dos tribunales y que 53 alumnos, el 15,23 por 100, habrían suspendido con ambos. Pero quedan algunos alumnos que aprobaron con un tribunal y con otro no. En concreto, 16 alumnos, el 4,60 por 100, aprobaron con el tribunal 10, pero no lo hubieran hecho con el tribunal paralelo. Alternativamente, 17 alumnos, el 4,89 por 100, sí que hubieran aprobado con este último tribunal, pero no lo hicieron con el oficial.

Observando estos datos desde otra perspectiva y limitándolos solamente a los suspensos en la prueba de selectividad (alrededor de 20 por 100 del total de presentados), diremos que las caras de éstos habrían variado entre un tribunal y otro en un 23 por 100 aproximadamente.

Pero, según el contraste χ^2 simétrico de McNemar (Sachs, 1978), esta diferencia entre los aprobados y los suspensos en cada tribunal, no llega a ser estadísticamente significativa a los niveles habituales.

Si analizamos estas discrepancias por centro, sexo y opción de COU, nos encontramos lo siguiente:

- a) Las discrepancias entre tribunales afectan a todos los centros. Consecuentemente, no existe diferencia en razón del centro secundario de origen.
- b) Entre los no aptos con el tribunal oficial que hubieran aprobado con el paralelo se acentúan algo más los varones. Lo contrario habría sucedido en el caso alternativo; con el tribunal paralelo habrían suspendido algunas mujeres más. Vemos que se produce una cierta interacción entre los tribunales en cuanto al sexo de los aprobados por uno y suspensos por el otro, pero que puede estar relacionada con las diferencias que apuntamos a continuación según opciones de COU.
- c) Los no aptos del tribunal oficial, aptos en el paralelo, son todos de las opciones de ciencias (A y B), casi en la misma proporción. Los no aptos del paralelo, aptos en el oficial, son la mitad de la opción A y el resto de las opciones C y D.

En conjunto, este análisis de interacciones de las discrepancias entre tribunales según centro de procedencia, sexo y opción de COU, no ofrece ninguna tendencia especialmente relevante, al menos de manera clara. Las diferencias parecen responder a circunstancias no sistemáticas, relacionadas con las características particulares de los vocales de los tribunales.

Como ya hemos señalado en el apartado anterior, también hemos calculado dos nuevas variables para cada uno de los 348 alumnos del estudio. Se trata de la diferencia entre sus notas en la selectividad (Dif-S) en los tribunales 10 y 13 y lo mismo, pero con la nota final del acceso (Dif-A). Las características básicas de estas dos variables aparecen en la tabla 3.

TABLA 3

Características de las variables-diferencia entre los tribunales 10 y 13

<i>Variable (Trib. 10-Trib. 13)</i>	\bar{x}	<i>S</i>	<i>Rango</i>
Diferencia-Selectividad	-0,091	0,472	(-1,562)-1,462
Diferencia-Acceso	-0,046	0,236	(-0,781)-0,731

Como era de esperar de los resultados analizados anteriormente, la media de ambas diferencias es cercana a cero, pero levemente negativa, porque ya hemos dicho que el tribunal 13 fue ligeramente más benevolente que el tribunal 10. Como es lógico también, los parámetros de la Diferencia-Acceso son la mitad que los de la Diferencia-Selectividad, por el modo en que se construye la varia-

ble acceso, a partir de la variable selectividad, promediándola con el expediente secundario del alumno.

Probablemente, lo más significativo de estos datos es la existencia de alguna discrepancia de un punto y medio entre un tribunal y otro en el global de la prueba de selectividad en algunos alumnos. Sin embargo, estos casos extremos son muy escasos. En concreto, los casos cuya discrepancia en la puntuación en la selectividad entre ambos tribunales es mayor que 1,1 son 6, esto es, el 1,7 por 100. La gran mayoría de estas diferencias se concentran en valores entre -0,5 y 0,5. En este intervalo se ubican el 69 por 100 de los casos y la gran mayoría de los restantes en valores cercanos a los extremos de este rango.

Lógicamente, las diferencias en el acceso son la mitad de las de la selectividad; dos casos alrededor de 0,75 de diferencia, pero solamente un 5 por 100 superan la diferencia de 0,5 puntos en un sentido o en otro. El 69 por 100 ofrece diferencias menores que 0,25 a favor de uno u otro tribunal.

Estos datos nos indican que los datos globales de los dos tribunales son prácticamente idénticos, pero también que se producen unas ciertas fluctuaciones entre los alumnos en ambos sentidos, aunque el resultado global sea compensado. Lógicamente, estas fluctuaciones deben ser más acusadas en aquellas materias de la selectividad donde hemos observado menores correlaciones entre las calificaciones de ambos tribunales, en concreto, en las pertenecientes al primer ejercicio, el ejercicio común.

4.2. *Otras relaciones e interacciones en cada tribunal*

Hasta aquí hemos comparado directamente ambos tribunales, ahora vamos a estudiar algunos aspectos importantes dentro de cada tribunal y, *a posteriori*, veremos si el panorama es diferente en ambos.

Dentro de cada tribunal hemos llevado a cabo un análisis completo de la distribución de resultados según sexo y según la opción de COU, contrastando la existencia o no de diferencias significativas entre varones y mujeres y entre opciones.

Los resultados de la comparación según el sexo son los que aparecen en la tabla 4.

En relación con las diferencias según sexo, ambos tribunales muestran una conducta muy similar. No existen diferencias significativas en ninguna de las variables en ambos tribunales salvo en Filosofía (X_3) a favor de las mujeres y en el Texto (X_1) a favor de los varones en el tribunal 13.

El análisis de varianza por opciones de COU dentro de cada tribunal nos ofrece el panorama de diferencias que aparece en la tabla 5.

TABLA 4

Diferencias por sexo en ambos tribunales

Variable	Tribunal 10		Tribunal 13	
	Signif. Difer.	A favor de:	Signif. Difer.	A favor de:
Texto (X ₁)	-	-	*	Varones
Lenguaje (X ₂)	-	-	-	-
Filosofía (X ₃)	**	Mujeres	°	Mujeres
Idioma (X ₄)	-	-	-	-
Obligator. 1 (X ₅)	-	-	-	-
Obligator. 2 (X ₆)	-	-	-	-
Optativa 1 (X ₇)	-	-	-	-
Optativa 2 (X ₈)	-	-	-	-
Exped. Sec. (X ₉)	-	-	-	-
Ejercicio 1 (X ₁₀)	-	-	-	-
Ejercicio 2 (X ₁₁)	-	-	-	-
Selectiv. (X ₁₂)	-	-	-	-
Acceso (X ₁₃)	-	-	-	-
* ($\alpha \leq 0,05$)				
** ($\alpha \leq 0,01$)				

Por opciones de COU se observan algunas diferencias entre los tribunales, pero en asignaturas concretas. En los promedios globales, la similitud entre tribunales es mucho más alta. De nuevo se observa el efecto compensatorio que se produce por la estructura de la prueba y la composición de los tribunales.

Entre todas las variables analizadas existe una que es común para ambos tribunales, el expediente secundario, y que por tanto es un elemento de referencia importante para analizar las distintas conductas de uno y otro tribunal. En la tabla 6 ofrecemos, para cada tribunal, las correlaciones de todas las variables con el expediente. Naturalmente, eliminamos de este análisis la variable Acceso (X₁₃)

TABLA 5

Diferencias por opciones de COU en cada tribunal

Variable	Tribunal 10			Tribunal 13		
	Significac.	Opción más alta	Opción más baja	Significac.	Opción más alta	Opción más baja
Texto (X ₁)	**	D	B	—	D	B
Lenguaje (X ₂)	*	D	B	*	D	B
Filosofía (X ₃)	—	A	B	—	D	B
Idioma (X ₄)	**	A	D	—	A	B
Oblig. 1 (X ₅)	**	C	B	**	C	B
Oblig. 2 (X ₆)	**	D	A	**	B	D
Optat. 1 (X ₇)	**	C	D	**	C	B
Optat. 2 (X ₈)	**	D	B	**	A	B
Ex. Sec. (X ₉)	**	A	D	—	—	—
Ejerc. 1 (X ₁₀)	**	A	B	**	D	B
Ejerc. 2 (X ₁₁)	**	C	B	**	C	B
Select. (X ₁₂)	**	C	B	**	C	B
Acces. (X ₁₃)	**	A	B	*	A	B
* ($\alpha \leq 0,05$)						
** ($\alpha \leq 0,01$)						

porque no es independiente del expediente secundario, ya que se construye promediando éste con la prueba de selectividad.

El conjunto de correlaciones ofrece una situación muy similar para ambos tribunales en la mayoría de las variables. Además, las correlaciones son sensiblemente altas, especialmente con los promedios. Con la prueba global (X₁₂), la correlación es altamente significativa y prácticamente idéntica en ambos tribunales. Esto contrasta bastante con los hallazgos de nuestras primeras investigaciones (Escudero y Bernad, 1978) y demuestra que los cambios que se han ido introdu-

ciendo para mejorar la fiabilidad y objetividad de la prueba, a los que anteriormente nos hemos referido, parecen haber tenido el efecto deseado.

En ambos tribunales, las correlaciones más bajas con el expediente secundario se dan con el Texto (X_1), Obligatoria 1 (X_5) y Optativa 1 (X_7).

TABLA 6

Correlaciones con el expediente de secundaria para cada tribunal

<i>Variable</i>	<i>Tribunal 10</i>	<i>Tribunal 13</i>
Texto (X_1)	0,263	0,326
Lenguaje (X_2)	0,459	0,497
Filosofía (X_3)	0,453	0,438
Idioma (X_4)	0,502	0,398
Obligatoria 1 (X_5)	0,254	0,301
Obligatoria 2 (X_6)	0,398	0,452
Optativa 1 (X_7)	0,332	0,346
Optativa 2 (X_8)	0,457	0,478
Ejercicio 1 (X_{10})	0,635	0,597
Ejercicio 2 (X_{11})	0,509	0,571
Selectividad (X_{12})	0,651	0,675

En los dos casos el ejercicio 1 (X_{10}) ofrece una correlación más alta que el ejercicio 2 (X_{11}) sobre todo en el tribunal 10. De hecho, en dicho tribunal, el promedio de ambos ejercicios (X_{12}) correlaciona con el expediente secundario muy poco más que el ejercicio 1 (X_{10}) por sí solo. Esta situación no se produce en el tribunal 13, donde los dos ejercicios ofrecen una correlación más cercana y, además, el promedio mejora sensiblemente la correlación con cada ejercicio.

De confirmarse y acentuarse la tendencia correlacional observada en el tribunal 10, habría que decir que el expediente secundario y el primer ejercicio de la selectividad ofrecen una alta concomitancia y, en todo caso, más alta que la que

se produce con el segundo ejercicio. Además, la correlación entre los ejercicios 1 y 2 (X_{10} y X_{11}) es más baja. Curiosamente, en ambos tribunales tiene el valor de 0,475 y, además, este valor sería muy parecido (0,505 ó 0,464) si intercambiáramos los tribunales, corrigiendo uno la parte común y otro la específica y viceversa. En otras palabras, lo que mide el expediente secundario se parecería más a lo que mide el ejercicio común, que a lo que mide el específico, que tendría un carácter menos confirmatorio y más complementario que aquél, sobre todo si es cierto que la relación entre ambos ejercicios es menor que la de cada uno de ellos, sobre todo el primero, con el expediente secundario.

5. DISCUSIÓN FINAL

Los hallazgos de este estudio del doble tribunal en la selectividad son, sobre todo, una demostración de que nuestro sistema de selección de universitarios tiene algunos problemas de consistencia entre tribunales, pero ni mucho menos es una especie de lotería «según el tribunal que toque». El procedimiento es mucho más consistente de lo que se suele decir, o al menos puede serlo, si se utiliza razonablemente, siguiendo las previsiones e indicaciones legales. Éste, desde luego, ha sido el caso de la experiencia realizada en su contexto real, tal como se hace con todos los tribunales, en la Universidad de Zaragoza.

Los distintos contrastes y análisis comparativos llevados a cabo nos conducen a unas situaciones muy similares para ambos tribunales con las variables globales y, en concreto, con la prueba de selectividad (X_{12}) y con su promedio con el expediente secundario (X_{13}).

En cuanto a resultados estadísticos con estas variables globales, que son las que se utilizan como criterio de selección, puede decirse que hubiera sido prácticamente igual la utilización de uno u otro tribunal. La conducta estadística global de ambos tribunales, el oficial y el paralelo, fue sorprendentemente similar.

En conjunto, los distintos ejercicios de la prueba juegan un papel compensatorio entre sí, si los vocales de los tribunales se eligen aleatoriamente, dándole a la misma un alto nivel de estabilidad en cuanto a los resultados globales. Además, respecto a elementos de interacción externos, como el sexo, las opciones de COU o el expediente secundario, los tribunales ofrecieron conductas muy similares. En definitiva, una consistencia entre tribunales mejor de lo que podía esperarse.

Sin embargo, sí que se observan diferencias particulares, por debajo de este efecto compensatorio global antes aludido, que pueden afectar a los ejercicios aislados de la prueba de selectividad y a individuos concretos que la realizan.

No se puede olvidar que se trata de pruebas abiertas y que su corrección está sujeta a diferencias de criterio. Esto se nota en el estudio, aunque luego unas asignaturas compensan con otras.

Estas diferencias de criterio particulares producen intercambios de individuos en los resultados, de forma que algunos alumnos aprobarían con un tribunal y no con otro y viceversa, aunque ambos terminen aprobando al mismo número de ellos. Pero tampoco estas discrepancias llegan a ser estadísticamente significativas.

Estos problemas aludidos no se pueden evitar con una prueba abierta como la nuestra y como la que tienen en la mayoría de los países de nuestro entorno (González y Valle, 1990; Escudero, 1991 y Moreno Olmedilla, 1992), pero se pueden minimizar, diseñando de manera precisa las cuestiones y los criterios de corrección y potenciando los sistemas de coordinación entre tribunales y áreas de corrección. Las Universidades deben seguir mejorando este proceso, con la confianza de que muchas de las críticas que recibe el sistema de acceso y su trabajo en este terreno, no siempre están suficientemente fundamentadas.

BIBLIOGRAFÍA

- CIDE (1991a): *Estudio experimental de pruebas objetivas para el acceso a la Universidad. Informe final*. Madrid, CIDE, policopiado.
- (1991b): *Más allá de los datos empíricos: consideraciones sobre la utilización de pruebas objetivas para el acceso a la Universidad*, manuscrito de difusión limitada.
 - (1992): «Estudio de doble corrección de pruebas de aptitud para el acceso a la Universidad». Madrid, informe inédito.
- Consejo de Universidades (1992): *Las pruebas de aptitud para el acceso a la Universidad: Problemática actual y propuesta de modificación*. Madrid, Secretaría General, informe policopiado.
- El País Educación* (1993): «Pruebas fiables», martes 22 de junio, p. 5.
- Escudero, T. (1984): «Características de la actual prueba de acceso», en I. Aguirre de Cárcer, (Ed.), *La selectividad a debate*. Madrid, Universidad Autónoma de Madrid, pp. 234-249.
- (1991): *Acceso a la Universidad: Modelos europeos, vías alternativas y reformas en España*, Informes 86. Zaragoza, ICE-Universidad de Zaragoza.
- Escudero, T. y Bernad, J. A. (1978): «Pruebas de acceso a la Universidad y su relación con los rasgos de madurez del alumno», *Revista de Educación*, 258/259, pp. 56-89.
- Feldman, D. y Gagnon, J. (1985): *Statview. The Graph Statistics Utility for the Macintosh*. Calabases, Ca., Brain Power Inc.
- González, B. y Valle, J. M. (1990): *El sistema de acceso a la educación superior en seis países de la CE*. Madrid, CIDE.
- Moreno Olmedilla, J. M. (1992): *Los exámenes: un estudio comparativo*. México, Fondo de Cultura Económica.

- Muñoz-Repiso, M. y otros (1991): *Las calificaciones en las pruebas de aptitud para el acceso a la Universidad*. Madrid, MEC/CIDE.
- Muñoz, F. (1992): «El sistema de acceso a la Universidad en España». Universidad Complutense de Madrid, tesis doctoral inédita.
- Orden Ministerial de 3 de septiembre de 1987 (*BOE* de 7 de septiembre), sobre pruebas de aptitud para el acceso a las Facultades, Escuelas Técnicas Superiores y Colegios Universitarios.
- Orden Ministerial de 9 de junio de 1993 (*BOE* de 10 de junio) sobre pruebas de aptitud para el acceso a las Facultades Técnicas Superiores y Colegios Universitarios.
- Real Decreto 406/1988 de 29 de abril (*BOE* de 30 de abril) sobre organización de las Pruebas de Acceso a las Facultades, Escuelas Técnicas Superiores y Colegios Universitarios y composición de los tribunales.
- Real Decreto 807/1993, de 28 de mayo (*BOE* de 8 de junio), por el que se modifica el Real Decreto 406/1988, de 29 de abril, sobre organización de las pruebas de aptitud para el acceso a las Facultades, Escuelas Técnicas Superiores y Colegios Universitarios.
- Rodicio, M. L. (1993): «El acceso a la Universidad en Galicia: Análisis de las disfuncionalidades de la prueba de selectividad y alternativas». Universidad de Santiago de Compostela, tesis doctoral inédita.
- Sachs, L. (1978): *Estadística aplicada*. Barcelona, Editorial Labor, S. A.
- Sans, A. (1991): «Fiabilidad y consistencia del proceso de selectividad. Un gigante con los dos pies de barro», en Latiesa, M. y otros (Rec. y Ed.), *La investigación educativa sobre la Universidad*. Madrid, CIDE, pp. 210-228.
- Wilkinson, L. (1986): *Systac: The System for statistics*. Evanston, Syxtac Inc.