



VICERRECTORADO DE CALIDAD E
INNOVACIÓN EDUCATIVA



Título del Informe:	Análisis de validez y fiabilidad del cuestionario de encuesta a los estudiantes para la evaluación de la calidad de la docencia
Fecha:	Diciembre 2011
Unidad Responsable:	ÁREA DE CALIDAD
Persona Responsable:	Pedro Gil Sopeña, TOC. Beatriz Porras, directora del Área de Calidad
Destinatario:	Comisión de Calidad e Innovación Educativa de la UC
Documentos que componen el informe:	
	<ul style="list-style-type: none">– <i>Análisis de validez y fiabilidad del cuestionario de encuesta a los estudiantes para la evaluación de la calidad de la docencia</i>– <i>Modelo de encuesta cursos 2007-2008 a 2010-2011</i>– <i>Modelo de encuesta curso 2011-2012</i>
Resumen:	
<p>La Comisión de Calidad e Innovación Educativa aprobó en septiembre de 2011 un nuevo modelo de encuesta dirigida a los estudiantes para recoger su opinión sobre la calidad de la docencia que reciben. El modelo aprobado parte de la propuesta realizada por el Área de Calidad, basado en un estudio sobre la fiabilidad del modelo anterior, y la incorporación de las observaciones y sugerencias recogidas a lo largo de los cuatro años en que estuvo vigente.</p> <p>La propuesta presentada proponía además la realización de un conjunto de pruebas de fiabilidad del nuevo modelo, a realizar en el primer trimestre del curso 2011-12.</p> <p>En este informe se presentan los resultados de estas pruebas de fiabilidad, y se recoge también la labor realizada en el seno del Área de Calidad y de la Comisión para aceptación de la validez del cuestionario de acuerdo con sus objetivos.</p> <p>Estos resultados respaldan la validez y fiabilidad del modelo de encuesta aprobado para la recogida de información sobre la opinión de los estudiantes acerca de la calidad de la docencia que reciben, y su utilización en la evaluación de la calidad de la enseñanza en el marco del SGIC de la UC.</p>	

Análisis de validez y fiabilidad del cuestionario de encuesta a los estudiantes para la evaluación de la calidad de la docencia

Resumen

La Comisión de Calidad e Innovación Educativa aprobó en septiembre de 2011 un nuevo modelo de encuesta dirigida a los estudiantes para recoger su opinión sobre la calidad de la docencia que reciben. El modelo aprobado parte de la propuesta realizada por el Área de Calidad, basado en un estudio sobre la fiabilidad del modelo anterior, y la incorporación de las observaciones y sugerencias recogidas a lo largo de los cuatro años en que estuvo vigente.

La propuesta presentada proponía además la realización de un conjunto de pruebas de fiabilidad del nuevo modelo, a realizar en el primer trimestre del curso 2011-12.

En este informe se presentan los resultados de estas pruebas de fiabilidad, y se recoge también la labor realizada en el seno del Área de Calidad y de la Comisión para la aceptación de la validez del cuestionario de acuerdo con sus objetivos.

Estos resultados respaldan la validez y fiabilidad del modelo de encuesta aprobado para la recogida de información sobre la opinión de los estudiantes acerca de la calidad de la docencia que reciben, y su utilización en la evaluación de la calidad de la enseñanza en el marco del SGIC de la UC.

ANTECEDENTES

La encuesta de opinión a los estudiantes sobre la actividad docente del profesorado utilizada por la Universidad de Cantabria llevaba cuatro años en uso, por lo que se vio necesario hacer una revisión del modelo en la que participasen los distintos grupos de interés (PDI, Estudiantes, Área de Calidad) con el fin de mejorarla aplicando la experiencia adquirida durante estos cuatro años de uso.

La nueva encuesta será uno de los instrumentos de información del Programa de Evaluación de la Actividad Docente (PEAD) que se pondrá en marcha en el curso académico 2011 – 12.

La Comisión de Calidad e Innovación Educativa de la UC encargó a la comisión permanente el análisis y revisión de la encuesta de calidad docente utilizada en el Sistema de Garantía Interno de Calidad. Para ello, el Área de Calidad preparó una propuesta de formulario acompañado de un dossier documental que fue debatido por la comisión permanente. Posteriormente, esta propuesta se elevó a la Comisión de Calidad e Innovación Educativa para su aprobación y finalmente fue ratificada por el Consejo de Gobierno de la Universidad de Cantabria.

Los aspectos de mejora que presenta la nueva encuesta a los estudiantes alcanzan los siguientes aspectos:

- Menor número de ítems.
- Redacción más escueta y clara de los enunciados.
- Adaptación a los criterios y directrices establecidos en el marco del Espacio Europeo de Educación Superior.

METODOLOGÍA DE ANÁLISIS

Para analizar la bondad del cuestionario debemos estudiar por un lado su **validez**, entendida como el hecho de que el instrumento mida aquello que realmente queremos medir, y por otro lado su **fiabilidad**, entendida como la ausencia de errores aleatorios e independencia de desviaciones producidas por errores causales.

La comprobación de la validez sigue métodos distintos que no se concretan en el cálculo de un coeficiente determinado. La fiabilidad por el contrario, es una característica de unos resultados, de unas puntuaciones obtenidas en una determinada muestra. En última instancia lo que nos va a decir un coeficiente de fiabilidad es, si el instrumento diferencia adecuadamente a los sujetos en aquello que pretendemos medir en función de las características particulares de dicho sujeto.

De lo dicho anteriormente se desprende que en primer lugar, nos debe preocupar la validez más que la fiabilidad: podemos estar midiendo muy bien aquello que no queríamos medir.

ANÁLISIS DE VALIDEZ

Un instrumento es válido si mide lo que decimos que mide y es realmente aquello que queremos medir. No hay una prueba estadística o matemática que pueda dar respuesta a esta cuestión, por lo que la comprobación de la validez seguirá otros métodos, como el análisis del contenido de los ítems, la verificación de hipótesis sobre el significado pretendido, etc.

Para nuestro cuestionario, la comprobación de la validez se realizó por medio de una valoración de expertos que atendieron a las siguientes consideraciones:

1. Se considera que un ítem es válido si estimula información exacta y relevante, por ese motivo se presta especial atención en la selección y redacción de los mismos.
2. Cuanto menos tenga que reflexionar el sujeto sobre el significado de un ítem, más válida será la respuesta puesto que no dará lugar a interpretaciones.
3. Algunas preguntas que son válidas para evaluar a un grupo de personas, pueden no serlo para otro grupo, por eso las preguntas deben ser válidas para el conjunto del profesorado de la Universidad y que la experiencia del estudiante durante el periodo de docencia sea capaz de percibir.

La comprobación de validez del cuestionario siguió tres etapas:

En primer lugar, el Área de Calidad realizó un estudio a partir de las evidencias recogidas durante los cuatro años de utilización del modelo de encuesta anterior, las sugerencias aportadas durante este tiempo por los distintos grupos de interés y los modelos utilizados por otras universidades nacionales e internacionales, presentando una propuesta para modificar el formulario de encuesta.

Esta propuesta fue presentada a los miembros de la Comisión de Calidad e Innovación Educativa para su estudio y aportación de nuevas sugerencias. El modelo propuesto, junto con las nuevas aportaciones, fueron analizadas y debatidas por la comisión permanente, que planteó algunas modificaciones.

El modelo acordado en la comisión permanente se elevó a la Comisión de Calidad e Innovación Educativa para su estudio y aprobación. Esta Comisión también propuso modificaciones que se tuvieron en cuenta.

De esta forma nos aseguramos que el instrumento tenga validez y que efectivamente la medición represente el concepto deseado.

ANÁLISIS DE FIABILIDAD

En el cálculo de la fiabilidad vamos a emplear dos métodos y tres pruebas diferentes. Los métodos son, el análisis del coeficiente de consistencia interna y el test – retest. Las tres pruebas realizadas han sido, en primer lugar y previo al lanzamiento del nuevo formulario, el análisis del coeficiente de consistencia interna de una selección de trece ítems del modelo de formulario precedente que mantienen en el nuevo modelo una redacción equivalente. En segundo lugar, el análisis del coeficiente de consistencia interna del nuevo formulario de encuesta. Y en tercer y último lugar, la repetición del test a un mismo grupo de sujetos bajo las mismas condiciones.

MÉTODO 1. COEFICIENTE DE CONSISTENCIA INTERNA.

Lo que expresa directamente este coeficiente es, hasta que punto las respuestas están lo suficientemente relacionadas entre si como para poder concluir que todos los ítems miden lo mismo, y por tanto son sumables en una puntuación final que mida un rasgo. Estos coeficientes se esgrimen como garantía de unidimensionalidad, es decir, de que un único rasgo subyace a todos los ítems.

Los coeficientes de consistencia interna más utilizados son el Kuder – Richardson 20 y el coeficiente α de Cronbach. En realidad se trata de la misma fórmula, una (Kuder – Richardson) empleada para ítems dicotómicos (división en dos partes: unos y ceros, Si o No, etc.) y la otra (Cronbach) empleada para ítems continuos. En nuestro caso hemos utilizado el coeficiente α de Cronbach.

A partir de las varianzas, el α de Cronbach se calcula así:

$$\alpha = \left[\frac{k}{k-1} \right] \left[1 - \frac{\sum_{i=1}^k S_i^2}{S_t^2} \right],$$

donde:

S_i^2 es la varianza del ítem i ,

S_t^2 es la varianza de los valores totales observados y

k es el número de preguntas o ítems.

Los valores que puede tomar α varían entre 0 y 1, aunque hay que advertir que puede llegar a alcanzar valores negativos de existir parejas de ítems negativamente correlacionadas.

El α de Cronbach no es un estadístico al uso, por lo que no viene acompañado de ningún p -valor que permita rechazar la hipótesis de fiabilidad de la escala. No obstante, cuanto más se aproxime a su valor máximo, 1, mayor es la fiabilidad de la escala. Además, en determinados contextos y por tácito convenio, se considera que valores de α superiores a 0,7 son suficientes para garantizar la fiabilidad de la escala.

Prueba 1. Cuando se planteó el nuevo modelo de encuesta carecíamos de series de datos de respuestas del mismo, pero sí realizamos una primera prueba de fiabilidad que nos permitiese seguir avanzando en el diseño e implantación de la nueva encuesta. Uno de los objetivos que se propusieron cuando nos planteamos la revisión del anterior formulario fue que existiese una línea de comparación entre el nuevo y el antiguo modelo de encuesta que nos permitiese contrastar los datos y ampliar, en lugar de perder, la serie de resultados de los últimos cuatro años.

Seleccionamos un total de trece ítems del anterior modelo de encuesta, que mantienen en el modelo actual un significado equivalente, en concreto fueron los ítems 1, 3, 4, 5, 7, 9, 10, 11, 13, 16, 18, 19 y 21. A continuación se escogieron nueve asignaturas entre los planes de estudio de grado del curso académico 2009 – 10. Para la selección de esta muestra únicamente se tuvo en cuenta la participación

en la encuesta, de modo que estuvieran representadas asignaturas con una amplia variedad de participación.

Los resultados obtenidos fueron los siguientes:

Tabla 1. Primer análisis del coeficiente de consistencia interna.

CÓDIGO IDENTIFICADOR DE ASIGNATURA	Nº DE ENCUESTAS REALIZADAS	COEFICIENTE α DE CRONBACH
G745	11	0.7798
G760	7	0.7971
G1075	99	0.9946
G1089	25	0.8642
G1095	48	0.9486
G1096	48	0.9751
G1097	49	0.9502
G1102	17	0.9902
G1104	19	0.9519

Si bien los valores de α más bajos se corresponden con las asignaturas que menor participación han tenido, los valores obtenidos se encuentran en todo momento por encima del umbral de validez. Estos resultados nos permitieron seguir adelante con la propuesta, a la espera de completar el estudio de validez y fiabilidad con pruebas realizadas, esta vez, a las series de datos obtenidas con el nuevo formulario.

Prueba 2. En esta segunda prueba se va a analizar, igual que en la anterior, el coeficiente de consistencia interna, pero esta vez sobre las series de datos obtenidas con el nuevo formulario de encuesta.

Se han seleccionado al azar un conjunto de diecisiete unidades de evaluación (par asignatura – profesor) de los nuevos planes de estudio de grado, cuya docencia se ha impartido durante el primer cuatrimestre del curso académico 2011 – 12.

Tras analizar los datos los resultados han sido los siguientes:

Tabla 2. Segundo análisis del coeficiente de consistencia interna.

CÓDIGO IDENTIFICADOR DE ASIGNATURA	PROFESOR	Nº DE ENCUESTAS REALIZADAS	COEFICIENTE α DE CRONBACH
G1211	P1	30	0,9700
G21	P1	70	0,9321
G21	P2	68	0,9771
G26	P1	77	0,9410
G329	P1	80	0,9658

G329	P2	20	0,9576
G329	P3	17	0,9298
G339	P1	32	0,9377
G339	P2	26	0,9186
G339	P3	48	0,9567
G339	P4	40	0,9549
G339	P5	97	0,9611
G357	P1	78	0,9562
G375	P1	47	0,8818
G378	P1	45	0,9140
G947	P1	21	0,9339
G947	P2	16	0,9273

Los resultados obtenidos con el nuevo formulario de encuesta son aún mejores que los primeros. El nuevo modelo arroja valores de α muy por encima del umbral de validez (0,7) y en todos los casos salvo uno, por encima de 0,91.

Podemos concluir que la nueva encuesta tiene una gran consistencia interna y es unidimensional y por tanto indicativa de un único rasgo.

MÉTODO 2. TEST – RETEST.

Los sujetos responden dos veces al mismo test, dejando entre las dos veces un intervalo de tiempo. El coeficiente de correlación entre las dos ocasiones es lo que se denomina coeficiente de fiabilidad test – retest. El intervalo de tiempo puede ser de días, semanas o meses, pero no tan grande como para que los sujetos puedan alterar la percepción que tienen de aquello que queremos medir.

Este método corresponde a un concepto más intuitivo de fiabilidad: un instrumento es fiable si en veces sucesivas aporta los mismos resultados.

Prueba 3. Seleccionamos a un grupo de 22 sujetos de segundo curso del Grado en Economía que estuvieron matriculados de la asignatura G355 el curso anterior, a los que, a primeros de noviembre y en un intervalo de una semana, realizamos dos encuestas de opinión sobre la actividad docente del profesorado. La asignatura se impartió durante el segundo cuatrimestre del curso 2010 – 11. De este modo, los sujetos no alterarían la percepción que ya se habían formado del profesor que les dio docencia en su momento, al no mediar más aprendizaje ni contacto con dicho profesor entre una encuesta y otra.

A los sujetos se les advirtió de que estaban realizando una prueba estadística, pero no en que consistía dicha prueba, ni que una semana después iban a tener que repetirla.

Por último, la encuesta se realizó en las mismas condiciones en las que se va a formalizar para el resto de estudiantes y docentes: se efectúa al inicio o final de la clase, durante un tiempo de 10 o 15 minutos, sin la presencia en el aula del profesor al que se está evaluando y en el más completo anonimato.

Los resultados obtenidos fueron estos:

Encuesta realizada el 11/11/11

Tabla 3. Resultados de la primera encuesta.

SUJETO	ITEM1	ITEM2	ITEM3	ITEM4	ITEM5	ITEM6	ITEM7	ITEM8	ITEM9	ITEM10	ITEM11	ITEM12	ITEM13	ITEM14	ITEM15	ITEM16	ITEM17
1	5	4	4	4	5	4	6	4	5	4	2	5	4	4	5	5	4
2	5	5	5	4	4	4	5	5	5	4	5	5	5	4	4	4	5
3	4	3	4	4	4	5	5	4	4	5	4	5	5	5	4	5	5
4	5	4	5	5	4	5	4	4	6	4	6	6	6	6	6	5	6
5	3	5	4	3	4	5	5	4	3	3	6	5	5	4	5	4	5
6	6	2	5	5	6	5	6	6	6	6	6	6	6	5	4	6	6
7	5	5	5	4	5	5	4	6	5	4	4	5	6	6	4	4	5
8	5	2	4	5	4	3	5	6	5	5	5	4	5	5	5	4	5
9	4	5	4	3	4	4	5	5	5	5	4	5	4	3	4	4	4
10	4	5	3	3	5	3	5	4	5	5	5	5	5	5	5	5	5
11	5	4	5	6	5	6	6	6	6	4	6	5	5	6	6	6	6
12	4	4	3	4	3	3	5	3	4	3	2	4	4	4	4	4	4
13	5	4	5	5	4	4	5	4	5	5	4	4	4	5	5	4	5
14	4	4	4	4	4	4	4	4	5	5	5	5	5	5	5	5	5
15	3	4	4	4	4	3	4	3	3	3	4	4	3	4	4	4	3
16	4	5	5	3	4	3	4	4	4	4	5	4	4	4	3	4	4
17	3	4	3	4	4	3	4	3	4	3	4	4	3	4	4	3	3
18	4	2	3	3	3	4	5	4	4	3	4	4	4	3	4	6	4
19	5	4	3	4	4	3	5	5	5	3	3	4	3	3	5	3	5
20	4	4	4	4	4	5	5	4	5	4	4	4	4	4	5	3	4
21	5	5	4	5	5	5	6	5	4	4		6	4	4	5	6	5
22	4	4	4	3	4	3	4	3	5	3	4	5	3	4	4	5	4

Encuesta realizada el 18/11/11

Tabla 4. Resultados de la segunda encuesta.

SUJETO	ITEM1	ITEM2	ITEM3	ITEM4	ITEM5	ITEM6	ITEM7	ITEM8	ITEM9	ITEM10	ITEM11	ITEM12	ITEM13	ITEM14	ITEM15	ITEM16	ITEM17
1	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
2	3	4	3	4	3	4	4	3	4	3	4	4	3	4	4	3	3
3	4	4	4	4	4	4	4	4	3	4	4	4	4	4	4	4	4
4	6	2	6	5	6	5	6	6	6	6	6	6	6	5	5	6	6
5	4	4	4	3	4	3	4	3	4	4	4		3	3	4	4	4
6	3	4	3	2	4	4	5	3	3	4	4	5	5	3	4	5	4
7	4	4	4	3	4	3	4	3	4	4	4	4	3	3	3	3	4
8	5	4	5	5	5	4	5	5	5	4	5	5	5	6	5	5	6
9	5	5	5	5	5	6	5	6	5	5	4	5	5	6	5	5	5
10	5	4	5	5	5	5	5	6	5	5	6	5	5	5	5	5	6
11	4	3	4	5	5	4	5	4	4	4	4	5	6	6	6	4	5
12	5	4	5	5	5	6	5	5	5	4		5		5	5	5	4
13	4	3	3	5	5	4	4	3	5	5	4	4	3	3	4	5	4
14	5	4	4	4	5	6	5	4	4	5	5	4	5	4	4	4	5
15	5	5	5	4	5	5	6	5	6	5	5	6	5	5	5	5	6
16	5	4	4	4	4	3	4	4	5	4	2	4	3	4	4	4	5
17	4	3	3	4	3	3	4	2	3	3	2	5	4	4	5	4	4
18	5	4	4	3	4	3	4	4	5	4	5	5	4	4	4	5	5
19	4	4	4	4	4	5	4	4	4	5	4	3	4	4	4	3	4
20	4	4	3	4	4	3	4	4	3	3	4	4	4	4	5	5	5
21	5	4	5	5	5	5	5	6	5	5	4	4	4	5	5	5	6
22	3	3	5	4	4	3	4	6	4	1	4	3	3	4	3	3	4

Las tablas de resultados obtenidas están en bruto, es decir, tal cual salen de la lectora óptica, por ese motivo las puntuaciones varían entre 1 y 6 en lugar de entre 0 y 5. Tras un primer tratamiento de los datos, se transforman de la primera a la segunda escala, que será la utilizada para la evaluación del profesorado. La lectura de datos se hace de este modo para que la lectora óptica pueda discriminar aquellos ítems que no se contestan y queden en blanco de los que tienen una valoración. La lectora debe asignar un valor a cada registro, por ese motivo el 0 es asignado a “sin contestar”, el 1 equivale a “totalmente en desacuerdo”, el 2 a “en desacuerdo” y así sucesivamente hasta el 6 “totalmente de acuerdo”. Posteriormente al exportar los datos de la lectora, ésta devuelve una tabla con celdas en blanco donde había 0 y un valor entre 1 y 6 para cada uno de los demás registros. El Área de Calidad resta un punto a cada valor para transformar la escala 1 – 6 a la 0 – 5 y que de esta forma puedan equipararse y contrastarse con los resultados de los Autoinformes del Profesorado y los Informes del Responsable Académico con los que se completa la evaluación del profesorado.

Análisis:

El estudio de los resultados obtenidos con la prueba del Test – Retest se va a centrar en los análisis de los promedios, la desviación estándar, la frecuencia absoluta y por último la correlación entre ambas encuestas.

Análisis del promedio:

El promedio es la suma de todos los valores de una variable dividido por el número de valores o respuestas existentes.

Un análisis de los promedios obtenidos en ambas encuestas (Tabla 5) nos revela que son muy similares ítem a ítem, lo que es reflejo de que los sujetos de la muestra entendieron los enunciados de la misma forma en las veces sucesivas en las que respondieron a la encuesta. El promedio final, obtenido de los promedios de cada ítem, desvela una desviación entre una encuesta y otra de tan solo 7 centésimas.

Tabla 5. Promedios por ítem.

ITEM	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	Promedio
Encuesta 1	4,36	4	4,09	4,05	4,23	4,05	4,86	4,36	4,68	4,05	4,38	4,73	4,41	4,41	4,55	4,5	4,64	4,37
Encuesta 2	4,36	3,82	4,18	4,14	4,41	4,18	4,55	4,27	4,36	4,14	4,19	4,48	4,19	4,32	4,41	4,36	4,68	4,30

Tabla 6. Promedios por sujeto.

ENCUESTA 1	ENCUESTA 2
5,4706	5,5294
5,4118	5,1765
5,1176	5,1176
4,8750	5,1176
4,8235	4,9412
4,5882	4,8824
4,5294	4,8667
4,5294	4,5882
4,5294	4,5294
4,5294	4,2353
4,4118	4,0000
4,3529	4,0000
4,2941	4,0000
4,2353	3,9412
4,1765	3,9412
4,0000	3,9412
3,9412	3,8235
3,8824	3,6875
3,7647	3,5882
3,6471	3,5882
3,5882	3,5294
3,5294	3,5294

Un análisis de los promedios obtenidos, esta vez estudiando los valores resultantes por sujeto (Tabla 6), muestra también semejanzas muy altas entre ambas encuestas, indicativo de la claridad y estabilidad de los enunciados.

Análisis de la desviación estándar:

La desviación estándar es una medida de dispersión que nos dice cuánto tienden a alejarse los valores del promedio en una distribución. En concreto es “el promedio de la distancia de cada punto respecto del promedio”. La desviación muestra la variabilidad de una distribución, indicando por medio de un número, si las diferentes puntuaciones de una variable están muy alejadas del promedio. Cuanto mayor sea ese valor, mayor será la variabilidad, cuánto menor sea, más homogénea será con respecto al promedio. Valores altos en la desviación son indicativos de que ha habido disparidad en la forma de contestar de los sujetos, no han opinado lo mismo unos y otros.

Nosotros no vamos a centrar nuestro análisis en el valor de la desviación obtenido como tal, ya que no importa que los sujetos hayan opinado en distinto sentido sobre un determinado ítem, sino en la comparación entre los resultados de las dos encuestas para cada ítem, ya que eso significaría que las diferencias de opinión se han mantenido en las dos ocasiones.

Las desviaciones obtenidas en ambas encuestas muestran que los resultados están muy relacionados entre las dos y discurren de forma pareja si se comparan ambas (Tabla 7).

Los valores más próximos los encontramos entre los ítems 1, 4, 7, 13, 14, 15, cuya desviación en ambos casos no difiere más allá del 0,04.

La desviación más alta comparando ambas encuestas las encontramos en los ítems 2, 8, 11, con valores de desviación superiores a 0,18.

Tabla 7. Desviación estándar.

ITEM	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Encuesta 1	0,79	0,98	0,75	0,84	0,69	0,95	0,71	1,00	0,84	0,90	1,16	0,70	0,96	0,91	0,74	0,96	0,85
Encuesta 2	0,79	0,66	0,85	0,83	0,73	1,05	0,67	1,20	0,90	1,04	0,98	0,81	0,98	0,95	0,73	0,85	0,89

Análisis de la frecuencia absoluta:

La frecuencia absoluta es el número de veces que se repite un determinado valor de una variable en una muestra o población.

Tabla 8. Frecuencias absolutas.

	ITEM	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
ENCUESTA 1	Frec. 1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Frec. 2	0	3	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0
	Frec. 3	3	1	5	6	2	8	0	4	2	7	1	0	4	3	1	3	2
	Frec. 4	9	11	10	10	14	6	7	10	6	8	9	9	8	10	10	9	7
	Frec. 5	9	7	7	5	5	7	11	4	11	6	5	10	7	6	9	6	10
	Frec. 6	1	0	0	1	1	1	4	4	3	1	4	3	3	3	2	4	3
ENCUESTA 2	Frec. 1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
	Frec. 2	0	1	0	1	0	0	0	1	0	0	2	0	0	0	0	0	0
	Frec. 3	3	4	5	3	2	7	0	5	4	3	0	2	6	4	2	4	1
	Frec. 4	9	15	9	10	10	7	12	8	8	10	13	9	7	10	10	7	10
	Frec. 5	9	2	7	8	9	5	8	3	8	7	4	8	6	5	9	10	6
	Frec. 6	1	0	1	0	1	3	2	5	2	1	3	3	2	3	1	1	5

Análisis de la correlación:

Para nuestro análisis de correlación vamos a emplear el **coeficiente de correlación de Pearson**. Es un índice que mide la relación lineal entre dos variables aleatorias cuantitativas. Es decir, mide el grado de relación de dos variables siempre y cuando ambas sean cuantitativas.

En cada ítem se ha calculado el coeficiente de correlación entre los valores obtenidos en la primera (X) y la segunda (Y) encuesta.

La fórmula de cálculo de este coeficiente es

$$r_{xy} = \frac{\sum XY - \bar{X}\bar{Y}}{S_x S_y}$$

donde S_x y S_y son las desviaciones típicas de X e Y, y $\bar{X}\bar{Y}$ es el producto de las medias de ambas variables.

Tabla 6. Coeficiente de correlación de Pearson.

	ITEM1		ITEM2		ITEM3		ITEM4		ITEM5		ITEM6		ITEM7		ITEM8		ITEM9		ITEM10		ITEM11		ITEM12		ITEM13		ITEM14		ITEM15		ITEM16		ITEM17		
Encuesta	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	
6	6	5	5	5	6	6	5	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	
5	5	5	5	5	5	5	5	5	5	5	5	6	6	6	6	6	6	6	5	5	6	6	6	6	6	6	6	5	6	5	6	6	6		
5	5	5	4	5	5	5	5	5	5	5	5	6	6	5	6	6	6	5	5	5	6	5	6	5	6	5	6	5	6	5	6	6	6		
5	5	5	4	5	5	5	5	5	5	5	5	5	6	5	6	6	5	5	5	5	6	5	5	5	5	5	5	5	5	6	5	5	6		
5	5	5	4	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5		
5	5	5	4	5	5	4	5	4	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	5	5	5	5	5	5	5	5	5		
5	5	4	4	4	5	4	5	4	5	5	5	5	5	5	5	5	5	4	5	5	4	5	5	5	5	5	5	5	5	5	5	5	5		
5	5	4	4	4	4	4	4	4	4	5	4	4	5	5	4	4	5	5	4	4	5	4	5	5	4	5	4	5	5	5	5	5	5		
5	5	4	4	4	4	4	4	4	4	4	4	4	5	4	4	4	5	4	4	4	4	4	5	4	4	4	4	4	4	5	4	4	5	5	
4	4	4	4	4	4	4	4	4	4	4	4	4	4	5	4	4	4	5	4	4	4	4	4	5	4	4	4	4	4	4	4	4	5	4	
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	5	4
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
4	4	4	3	3	3	3	3	4	4	4	3	3	4	4	4	3	4	4	3	4	4	4	4	4	4	4	3	4	4	4	4	4	4	4	4
4	4	4	3	3	3	3	3	4	4	4	3	3	4	4	4	3	4	4	3	4	4	4	4	4	4	4	3	4	4	4	4	4	4	4	4
4	4	3	3	3	3	3	3	3	4	4	3	3	4	4	3	3	4	3	3	3	3	3	3	3	3	3	3	3	4	4	3	3	3	4	4
3	3	2	3	3	3	3	3	3	4	4	3	3	4	4	3	3	4	3	3	3	2	2	4	3	3	3	3	3	3	3	3	3	3	4	4
3	3	2	3	3	3	3	3	3	3	3	3	3	4	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	4
3	3	2	2	3	3	3	2	3	3	3	3	3	4	4	3	2	3	3	3	1			4		3		3	3	3	3	3	3	3	3	3
Pearson	1,0000		0,8077		0,9406		0,8031		0,8475		0,9436		0,7631		0,9412		0,8525		0,8106		0,8992		0,8232		0,8838		0,9504		0,8862		0,8744		0,8457		

El valor del índice de correlación de Pearson (r) varía en el intervalo $[-1, 1]$:

Si $r = 1$, existe una correlación positiva perfecta. El índice indica una dependencia total entre las dos variables denominada *relación directa*.

Si $0 < r < 1$, existe una correlación positiva.

Si $r = 0$, no existe relación lineal aunque pueden existir todavía relaciones no lineales entre las dos variables.

Si $-1 < r < 0$, existe una correlación negativa.

Si $r = -1$, existe una correlación negativa perfecta.

De los valores obtenidos con este índice podemos concluir que existe una correlación positiva muy alta entre las dos encuestas y que hay por tanto un coeficiente de fiabilidad test – retest muy alto. De hecho todos los valores excepto el correspondiente al ítem 7 (0,7631), están por encima de 0,8 y en varios casos, incluso, por encima de 0,9. Cabe destacar el ítem 1 donde la correlación ha sido perfecta.

CONCLUSIÓN

La validez del cuestionario para los objetivos de evaluación se considera establecida por la revisión de expertos, al aportarse:

- Las sugerencias recibidas de profesores y estudiantes a lo largo de cuatro años de funcionamiento del cuestionario anterior.
- El análisis de otros modelos de cuestionarios utilizados en el mismo contexto de evaluación: encuestas de calidad docente de universidades nacionales e internacionales.
- El análisis de la comisión permanente de calidad e innovación educativa y del pleno de la comisión.

En cuanto a la fiabilidad, los resultados obtenidos con el coeficiente α de Cronbach son indicadores de la homogeneidad de los ítems, es decir, de que todos miden lo mismo y de que existe una gran consistencia interna, por ello podemos suponer que si las respuestas están relacionadas es porque los ítems expresan o son indicadores del mismo rasgo.

Los resultados obtenidos en la prueba del test – retest, los entendemos como un indicador de la estabilidad y no ambigüedad de la encuesta, en la medida en que en ambas ocasiones los resultados han sido muy parecidos. Los sujetos entendieron lo mismo, de la misma manera y respondieron de forma idéntica o casi idéntica.