

## CAPÍTULO II

### SEGUNDA PONENCIA

#### ENCUESTAS EVALUADORAS DEL PROFESORADO

#### PRUEBA DE CARGO. PRUEBA LÍCITA. PRUEBA TASADA

JOAQUÍN MUÑOZ GARCÍA

MARÍA DOLORES JIMÉNEZ GAMERO

Departamento de Estadística e Investigación Operativa.

Universidad de Sevilla.

MARÍA MUÑOZ CONDE

Instituto de Estadística de Andalucía. Junta de Andalucía

### I. INTRODUCCIÓN

Desde hace ya años, los servicios que prestan las administraciones públicas se han incorporado a un mundo, donde prácticamente todo lo que se hace y se gestiona ha de ser medido o valorado. Ésta es una exigencia que desde la sociedad se hace a las administraciones públicas. A groso modo, el objetivo de tales valoraciones es doble: (1) nos permite abundar en la transparencia que se demanda a la gestión pública, y (2) está referido al hecho comparativo o/y competitivo que supone estar en un mundo globalizado, que en el caso particular de España se tiene una referencia global más próxima, la Unión Europea. Las valoraciones que se hacen de los servicios públicos han de servir de referencia a los propios servicios y a sus trabajadores, a sus usuarios y a los ciudadanos en general, ya que ello permite a la Administración corregir defectos y mejorar su gestión, facilitando la toma de decisiones en los correspondientes servicios públicos.

Por consiguiente, la Universidad española, que tiene entre sus funciones realizar el “servicio público de la educación superior mediante la investigación, la docencia y el estudio”<sup>1</sup>, tiene la obligación de estar sometida a evaluaciones, tanto en el ámbito de la investigación como de la docencia y el estudio. Es más el Título V de la citada ley orgánica 6/2001, es el “De la evaluación y acreditación”.

En este trabajo se abordan aspectos relacionados con el servicio de la docencia en la Universidad. Ello supone inicialmente tener que evaluar diferentes componentes, como son: la relativa al profesorado, como transmisor del conocimiento, la del Departamento, como organizador de la docencia entre el profesorado del

---

1. Ley Orgánica 6/2001, de 26 de Diciembre, de Universidades

mismo, la de la Facultad, como ámbito que proporciona al profesorado medios para realizar su labor docente, y la relativa a la Universidad, identificando a ésta como órgano donde se toman decisiones sobre cuestiones fundamentales para la docencia, que pueden ir desde el número de alumnos a admitir por grupo, por profesor o por asignatura, a la propia normativa docente.

En la valoración de los servicios públicos, es necesario comenzar citando el hecho de que los usuarios tienden a ser considerados clientes, previamente eran denominados administrados. En el caso de la Universidad, este concepto se complica ya que sus clientes en el ámbito de la docencia, van desde la Sociedad en general, pasando por las familias, instituciones u organizaciones que contribuyen a la matriculación de algunos de sus miembros en la Universidad, para que reciban la correspondiente formación (la concepción de clientes está íntimamente relacionada con aquellos que sufragan económicamente el servicio que recibe), aunque en lo que sigue y con el fin de precisar más, únicamente se hará referencia en el ámbito de la docencia, a los estudiantes o alumnos como los clientes de la Universidad, al ser ellos los receptores de la actividad docente.

En la actualidad un elemento fundamental, y quizás único, en la evaluación del servicio docente que se proporciona desde la Universidad, son las encuestas que responden los alumnos (“como clientes”) sobre el profesorado que les imparten las distintas asignaturas, y que se denominan aquí “Encuestas Evaluadoras del Profesorado”.

## II. ENCUESTAS EVALUADORAS DEL PROFESORADO COMO PRUEBA DE CARGO

Las Universidades españolas utilizan como medio básico para evaluar su servicio desde el punto de vista de la docencia<sup>2</sup>, y quizás el único y fundamental, la recogida de opiniones de los estudiantes a través de cuestionarios normalizados.

Los fines que pueden tener la recogida de la información de los alumnos sobre la efectividad de la enseñanza, pueden ser varios, como recoge Marsh (1987)<sup>3</sup>, y que son:

- Diagnosticar la efectividad de la enseñanza del profesorado, lo que resultará útil para la mejora de la enseñanza.
- Obtener una medida de la efectividad de la enseñanza, para ser utilizada en la toma de decisiones administrativas.

---

2. La valoración de la actividad docente, no debe considerarse como algo novedoso, algunos autores sitúan sus inicios hacia el año 340 en Antioquía.

3. Marsh, H.W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research*. Vol. 11, pp. 253 – 388.

- Obtener una información que pueda ser útil a los estudiantes para la selección de profesorado y cursos.
- Obtener una medida de la calidad de la asignatura, de interés para la mejora de la asignatura y del plan de estudios.

A estos cuatro fines, le añade además uno de carácter general sobre la utilidad que tiene la recopilación de este tipo de información para la investigación general en el campo de la enseñanza. En el trabajo citado se discuten los fines mencionados, haciendo notar que el primero es casi universal cuando se pretende recabar información de los estudiantes sobre la enseñanza recibida. Además, resalta el hecho de que la información que se obtenga de los estudiantes no suele tener validez para cualquiera de los fines descritos previamente.

En el caso de las Universidades españolas, uno de los fines de la recogida de información y quizás el principal, es para la toma de decisiones administrativas. En particular, para la evaluación del profesorado, lo que en algunas universidades<sup>4</sup>, supone que en el caso de que un profesor tenga la evaluación desfavorable, puede suponerle acciones punitivas tanto académicas como económicas, y que pueden ir desde posibles medidas correctoras en sus métodos de trabajo, limitaciones académicas en la posible docencia a impartir, o no ser propuesto para la percepción de complementos económicos ligados a los méritos docentes.

Se ha adoptado la terminología de encuestas evaluadoras del profesorado. Estas encuestas deben ser consideradas como “prueba” de lo que se desea constatar, de modo que, para el profesor que obtiene un informe favorable, podrían ser consideradas como “prueba de descargo”, y para el profesor que obtiene un informe desfavorable, se calificarían como “prueba de cargo”, ya que podría tener determinadas repercusiones en su actividad académica presente y futura, como se ha descrito previamente.

Y si se adapta la frase tan utilizada en el mundo de la judicatura, diciendo: “es preferible que un mal profesor obtenga un informe favorable, a que un buen profesor obtenga un informe desfavorable”, nos conduce a que la “prueba de cargo” que se debe utilizar para evaluar el profesorado, no puede ser cualquiera, sino que ha de tener la validez y la eficacia probatoria necesaria para que se sustente los fines y objetivos que fundamenten las encuestas evaluadoras del profesorado. En resumen, la encuesta evaluadora del profesorado, debería tener la concepción de “prueba lícita”. Por ello, a partir de ahora nos centraremos en los fundamentos que han de regir las encuestas evaluadoras del profesorado de modo que, cuando sean utilizadas como “prueba de cargo”, puedan ser consideradas como “prueba lícita”. Además, esto está previsto en las normativas universitarias, donde se expresa que las encuestas evaluadoras del profesorado han de realizarse con las debidas garan-

---

4. Ver Reglamento General de Actividades Docentes de la Universidad de Sevilla

tías técnicas, de forma que se asegure su validez y su viabilidad. Esto último puede verse en el Reglamento ya citado de la Universidad de Sevilla.

### III. APROXIMACIÓN AL DISEÑO DE LAS ENCUESTAS EVALUADORAS DEL PROFESORADO PARA SER PRUEBAS LÍCITAS

Al diseñar una encuesta para evaluar al profesorado, y que ésta se ajuste al fundamento y a la idea que hay tras el concepto de prueba lícita, es imprescindible seguir el esquema previsto para el diseño de cualquier encuesta por muestreo probabilístico o censo, pues es la única forma posible de asegurar que las conclusiones que se obtengan tengan la validez estadística adecuada. Nótese que esta validez resulta fundamental si se desea consolidar la información obtenida de diferentes elementos.

A continuación se desarrollará una aproximación metodológica al diseño de un muestreo probabilístico o un censo. Como en ambos casos hay etapas más o menos comunes, se utilizará el término común de “prueba lícita para la evaluación”, y en aquellos puntos en los que haya que establecer distinciones se hará notar.

Construir una “prueba lícita para la evaluación”, supone desarrollar un trabajo previo que fundamente las siguientes cuestiones: la aplicación de la citada prueba, el análisis de los resultados que se obtengan de la misma y la conclusión a hacer sobre el profesorado. Este trabajo previo consiste en: estudiar, analizar y fijar los contenidos de las etapas que se recogen en el diagrama del Anexo I<sup>5</sup>, y que se describe seguidamente.

1. *Fines y objetivos para la “prueba lícita para la evaluación”*. En el diseño de la prueba, se debe comenzar identificando los fines, algunos ya citados previamente, y objetivos de la misma. Éstos no pueden ser definidos de una forma ambigua, sino que hay que hacerlo de la forma más precisa posible, es decir, si el objetivo es evaluar la calidad de la docencia, es necesario definir qué significa “calidad en la docencia”. La calidad puede ir referida a: (1) la demanda que del alumnado hacen las empresas, (2) el trato, confianza y dedicación del profesorado al alumnado, (3) la posible eficiencia del sistema docente, la cual habrá de ser especificada. Por ejemplo, podría fijarse como criterio que la máxima eficiencia se tiene, cuando el número de alumnos que entran en el sistema es idéntico al número de los que salen.

La precisión al decidir los fines y fijar los objetivos de la prueba lícita de evaluación, inciden de forma determinante en otras etapas del diseño de la prueba, como pueden ser: los elementos a entrevistar, los contenidos del cuestionario,

---

5. La ordenación de las etapas, no debe ser considerada como una ordenación estricta, ya que pueden abordarse etapas de una forma conjunta, e incluso existir cierta retroalimentación con las discusiones que se realicen en las distintas etapas.

la referencia temporal, etc. Por ejemplo, no coinciden los elementos a entrevistar, ni la temporalidad, ni los contenidos del posible cuestionario, si se deseara evaluar el efecto de la enseñanza recibida por el alumnado en una Universidad y la obtención de un trabajo; que si se deseara evaluar cada uno de los cursos de los grados, o evaluar los métodos de calificación del profesorado, o la confianza que en su formación le dan al alumnado los profesores de las asignaturas.

Es conveniente indicar que el número de objetivos que se fijen debe ser pequeño, si se quiere que la prueba sea efectiva. Estos objetivos deben ser acordados por expertos, y además, necesariamente han de ser susceptibles de la observación y el juicio de los alumnos.

2. *Población en la “prueba lícita para la evaluación”*. La población de la prueba es el conjunto de personas que deben proporcionar la información adecuada a los fines y objetivos propuestos en la etapa previa. Los elementos de la población deben ser identificables y accesibles, para lo cual se ha de disponer de la información necesaria (marco). En caso contrario, deberá rectificarse la población, con lo que ello pueda significar sobre los fines y objetivos considerados en la etapa anterior. Por ejemplo, no puede ser la misma población si desea obtener información por profesor y asignatura, que si se deseara obtener información únicamente por profesor y grado, o únicamente por profesor. En este último caso, se tendría una población por profesor, y en los ejemplos previos se tendrían varias poblaciones, ya que habría una población por grado en que impartiera el profesor, o una por cada asignatura<sup>6</sup>.

La población ha de ser fijada con antelación a la realización de la prueba. De no ser así, los resultados que se obtengan no tendrán una referencia estadística adecuada, pues se estaría alterando la población de referencia en el momento de realizar la prueba.

En esta etapa también es importante fijar la temporalidad de la prueba, considerando los dos tiempos siguientes: (1) el de referencia, que se identifica con el esfuerzo de memoria que ha de hacer el entrevistado, y (2) el de ejecución, que se refiere al momento de realización de la prueba. Estos tiempos deben ser tratados en esta etapa o en la anterior, ya que ello puede condicionar la población que se defina.

3. *Fijar las variables y escalas de la “prueba lícita para la evaluación”*. En estas pruebas es necesario expresar como variables la información a recoger, para conseguir los objetivos fijados en la primera etapa. Esta labor no es sencilla, ya que pueden haberse fijado términos o conceptos que no son expresables

---

6. En el caso de disponer de poblaciones, para cada población hay que seguir las mismas etapas, por tanto en lo que sigue, se hará referencia de forma genérica a una población.

mediante una variable, y en este caso se recurre generalmente a la construcción de escalas. Para la definición de escala puede tomarse la existente en pág. 210 de Corbetta (2007)<sup>7</sup>: “es un conjunto coherente de elementos que se consideran indicadores de un concepto más general”. La construcción de una escala no es tarea fácil, siendo necesario seguir una metodología adecuada, como se describe de forma resumida en el mismo Corbetta (2007).

En esta etapa es conveniente resaltar, el hecho de que todos los autores dedicados al estudio de metodologías para la evaluación del profesorado coinciden en lo siguiente: cualquiera que sea el objetivo que se fije para evaluar la actividad docente, nunca es posible resumirlo en una respuesta a una pregunta simple, es decir, la valoración que los estudiantes hagan del profesor siempre será multidimensional. Por ejemplo, para un determinado objetivo sobre la docencia, puede ocurrir que, un profesor tenga su trabajo muy estructurado, ser claro en sus exposiciones, y sin embargo no tener la capacidad adecuada de entusiasmar al alumnado.

4. *Instrumento y método para la recogida de la información.* El instrumento tradicional en “la prueba lícita de evaluación” para recoger la información es un cuestionario estructurado. Para su construcción deben seguirse las recomendaciones existentes para la elaboración de cuestionarios. En este punto resulta imprescindible la participación de expertos.

También es necesario conocer el método que se utilizará para la recogida de información. Aunque existen diversos métodos, el más aconsejable en este tipo de prueba es el de la entrevista directa. Por la experiencia existente, otros métodos, que van desde la entrevista a través de internet, a la que puede realizarse por correo postal o por teléfono, suelen presentar, entre otros defectos, una alta tasa de no respuesta.

En la construcción del cuestionario para este tipo de prueba, debe plantearse o discutirse, desde el punto de vista de los objetivos que se planteen en la primera etapa, el hecho de la homogeneidad del cuestionario para todo el alumnado de la Universidad, dada la diferencia existente entre los grados que se impartan en la misma, debido a los diferentes entornos de docencia en que se desarrolla la formación del alumnado.

5. *Métodos para obtener los elementos de la Población en la “prueba lícita para la evaluación del profesorado”.* Como ya se ha comentado existen dos métodos para conseguir los miembros de la población que deben proporcionar la información que se desea: uno es el censo y el otro es la realización de encuestas por muestreo.

---

7. Corbetta, P. (2007). Metodología y Técnicas de Investigación Social. (Edición Revisada). McGraw-Hill.

En el caso del censo, se pretende que respondan al cuestionario elaborado todos los elementos de la población. En este caso existe la posibilidad de que se presente el error de cobertura, del que se hablará en la etapa siguiente.

Lo ideal en este tipo de prueba lícita para la evaluación del profesorado, es hacer un censo. Si resulta costoso, o el número de elementos de la población que se considera es elevado, o es imposible realizarlo por cualquier otra razón, se realizaría una encuesta por muestreo, es decir, se entrevistaría a parte de la población, donde los elementos a participar en la encuesta deberán ser elegidos probabilísticamente, ya que como indica Särndal, Swensson and Wretman (1992)<sup>8</sup> en pág. 7, “Únicamente en situaciones muy afortunadas, la selección no probabilística producirá estimaciones precisas”. Lo que se fije en etapas previas precisará de forma general el tipo de muestreo que debe aplicarse.

La decisión sobre censo o encuesta por muestreo probabilístico, delimitará la forma de realizar las pruebas pilotos correspondientes. Usualmente, lo observado en la prueba piloto es de gran utilidad en las etapas siguientes, o para la valoración del cuestionario.

6. *Errores en la “prueba lícita de evaluación del profesorado”*. Cuando se desea que la prueba lícita de evaluación del profesorado tenga rigor y significado estadístico, es necesario fijar los errores que se está dispuesto a admitir al aplicar dicha prueba. Esto permite conocer la precisión y fiabilidad de las conclusiones que se obtengan de la misma.

Para el análisis de los errores hay que distinguir los dos métodos citados: el censo y las encuestas por muestreo probabilístico.

En el caso de aplicar un censo hay que fijar con antelación a la obtención de la información, el llamado error de cobertura, que nos da la divergencia que se presenta entre la población fijada como objetivo y las personas que participan en la prueba. Este error se expresa mediante el cociente entre el número de elementos que participan en la prueba y el número de elementos de la población. Este error delimita la validez de las conclusiones que se obtengan para la población objetivo.

En el caso de las encuestas por muestreo probabilístico, los errores que se presentan proceden del hecho que no todos los elementos de la población objetivo son observados, sólo se observa una parte de ella, a la que se le denomina muestra, y si es seleccionada por métodos probabilísticos, se le denomina probabilística.

Una forma de proceder en este caso es: fijar los errores y a partir de ellos delimitar el número de elementos a seleccionar probabilísticamente de la pobla-

---

8. Särndal, C.-E., Swensson, B. and Wretman, J. (1992). Model Assisted Survey Sampling. Springer-Verlag

ción. Mientras menores sean los errores, mayor será el número de elementos a tomar de la población, y más precisión y fiabilidad tendrán las conclusiones. En el muestreo probabilístico los errores se fijan de acuerdo al método de muestreo que se aplique, a las variables y estimadores que se consideren, y al análisis estadístico que se desee realizar.

7. *Aproximación al análisis de datos de “la prueba lícita de evaluación del profesorado”*. Esta etapa hay que abordarla necesariamente después de haber cerrado todas las etapas anteriores de la metodología. En ésta, hay que delimitar el tratamiento a dar a la información recogida en los cuestionarios que se le han pasado a los elementos de la población objetivo. Este tratamiento depende inicialmente de las variables y/o escalas definidas en etapas previas. Generalmente se trabaja con variables de naturaleza ordinal y con relación multivariante entre ellas. Inicialmente se debe hacer un análisis exploratorio de los datos, con el objetivo de depurar los mismos, tanto desde el punto de vista de las posibles observaciones extremas, como el hecho de evitar posibles incongruencias en las respuestas. A continuación, se realizan los correspondientes cálculos y análisis estadísticos. En todos ellos se ha de ser extremadamente riguroso para que los resultados a obtener tengan una interpretación adecuada a lo planteado<sup>9</sup>. En el cálculo estadístico para las pruebas lícitas de evaluación del profesorado, hay que considerar las ponderaciones a que deben ser sometidas las respuestas que se dan a las preguntas del cuestionario cumplimentado por cada elemento de la población.

#### IV. PONDERACIONES

El alumno que responde a las encuestas que se hacen sobre el profesorado, no debe ser interpretado como un cliente pasivo frente a un producto de consumo, ya que en el caso de la docencia se establecen unas relaciones profesor-alumno y alumno-profesor. A esto hay que unir las relaciones de los alumnos entre sí, todo ello hace que el alumno deba ser considerado un “cliente proactivo”. Por tanto, hay que incorporar en la encuesta que se haga, aspectos de la actividad del alumno que deben ser tenidos como referencia para delimitar las ponderaciones que puedan afectar a las respuestas que ellos den a las preguntas que les hagan.

---

9. No ser riguroso desde un punto de vista estadístico, nos puede llevar a conclusiones peregrinas, como ilustración de lo que puede ocurrir si no se respeta el formalismo estadístico, baste citar los libros:

Huff, D. (1991). *How to Lie with Statistics*. Penguin Books.

Zeisel, H. (1998). *Dígalo con Números*. Fondo de Cultura Económica.



En Marsh (1987), se recogen determinados factores que inciden en la opinión que tienen los alumnos sobre sus profesores, y que se han puesto de manifiesto en diversos estudios. Entre éstos pueden citarse los siguientes: la dificultad del curso, la popularidad del profesor, la benevolencia de este último en sus evaluaciones, el interés del estudiante por la materia, la carga de trabajo y el número de alumnos. Además se pueden considerar cuestiones como la asistencia a clase, su reincidencia en cursar la materia, su hábito de estudio, el efecto Facultad, el efecto universidad, etc.

## V. MEDIDA SINTÉTICA DE LA CALIDAD

Dentro de los objetivos que se plantean en las pruebas lícitas de evaluación del profesorado, suele contemplarse la posibilidad de simplificar lo observado en una única medida, de forma que los valores que se generen, produzca interpretaciones útiles a diferentes niveles. Estas medidas o valores no suelen ser medidas estadísticas descriptivas simples, sino medidas sintéticas, que no son más que funciones de ciertas variables latentes, que a su vez son funciones de variables observables en las entrevistas realizadas. En la construcción de medidas sintéticas es necesario considerar determinados efectos, como puede ser el jerarquizado de la propia información. Esta necesidad la explican claramente Rampichini, Grilli y Petrucci (2004)<sup>10</sup>, al afirmar que los alumnos presentan actitudes distintas según los cursos donde se encuentren, y concluyen en la pág. 364 que “cualquier comparación entre cursos que se base en indicadores descriptivos, por tanto, serían erróneos”.

Como indican Balzano y Trinchera (2011)<sup>11</sup>, no existe una solución metodológica simple para construir funciones de variables observadas en los cuestionarios cumplimentados por los elementos de la población. Estos autores reconocen que existen diversas aproximaciones para la construcción de medidas sintéticas. No obstante, ellos proponen utilizar modelos de ecuaciones estructurales, ya que estos modelos son utilizados para describir estructuras conceptuales, con presencia de variables latentes y variables observadas experimentalmente, y como ya se ha indicado, este es el caso de las pruebas lícitas de evaluación del profesorado. En el trabajo citado se aplican los modelos de ecuaciones estructurales, a encuestas pasadas a estudiantes de una Facultad de Humanidades del sur de Italia.

---

10. Rampichini, C.; Grilli, L. and Petrucci A. (2004). Analysis of university course evaluations: from descriptive measures to multilevel models. *Statistical Methods & Applications*.

11. Balzano, S. and Trinchera (2011). Structural equation models and student evaluation of teaching: A PLS path modeling study. En Attanasio M. and Capursi, V. (eds). *Statistical Methods for the Evaluation of University Systems. Contributions to Statistics*. Springer-Verlag.

## VI. REALIZACIÓN DE LA PRUEBA LÍCITA PARA LA EVALUACIÓN DEL PROFESORADO

Una vez se han cubierto las etapas anteriormente descritas se ha de pasar a la realización de la prueba lícita para la evaluación del profesorado. Esto conlleva contratar a los entrevistadores, proporcionarles la formación necesaria y pasar al trabajo de campo. Una vez recogida la información se introduciría en el sistema informático adecuado para ejecutar el análisis estadístico y obtener las medidas sintéticas que cubran los objetivos que inicialmente se propusieron.

La realización de la prueba puede suponer, según los objetivos que se marquen, una actividad planificadora y de gestión importante, para llevar a cabo el trabajo de campo deseado.

## VII. CONCLUSIONES

Las encuestas evaluadoras del profesorado han de ser consideradas, según se están planteando en la actualidad por muchas Universidades, como pruebas que permiten juzgar la actividad docente del profesorado, pudiéndose llegar en ciertos casos a poder imponer “penas” académicas y/o económicas. Por ello hay que considerar que, las encuestas evaluadoras del profesorado, deben tener consideración de prueba de cargo.

Como consecuencia, la encuesta evaluadora del profesorado que realizan las Universidades, ha de ser una prueba lícita, es decir, la prueba debe ajustarse, en este caso, a la exigencia científico – técnica de cualquier encuesta que pretenda obtener conclusiones con un determinado rigor y precisión estadística. Para ello, han de respetarse las etapas descritas, que han de ser siempre previas a la realización de las encuestas.

En las etapas descritas existen elementos que deben ser consensuados por las partes intervinientes en el proceso de evaluación: la precisión que hay que conseguir en la definición de los objetivos, los errores que han de fijarse en la prueba, la determinación de ponderaciones, la medida sintética, etc.

Por tanto, las encuestas evaluadoras del profesorado, deben ser pruebas lícitas, con sometimiento a la ciencia y a la técnica que fundamentan los procesos de evaluación mediante encuestas, y con ciertos elementos consensuados por las partes, todo ello con el fin de evitar que el proceso evaluador dependa de posibles criterios subjetivos. Así, la conclusión última es que las encuestas evaluadoras del profesorado realizadas en muchas Universidades deben tener la concepción de Prueba Tasada.

Y se deja la siguiente pregunta a responder ¿Basta la Prueba Tasada descrita para valorar la actividad docente de un Profesor?

## ANEXO I

Diagrama de las Etapas a Seguir en las Pruebas Lícitas de Evaluación del Profesorado

