

Vol. 18, Núm. 2, 2016

Instrumento para determinar los tipos de evaluación utilizados por los profesores universitarios¹

An Instrument to Determine the Types of Assessment Used by members of Faculty

Arturo De la Orden Hoz (1) afordenh@gmail.com
Julio Herminio Pimienta Prieto (2) julio.pimienta@anahuac.mx

(1) Universidad Complutense de Madrid

(2) Universidad Anáhuac

(Recibido: 11 de febrero de 2015; Aceptado para su publicación: 20 de agosto de 2015)

Cómo citar: De la Orden, A. y Pimienta, J. H. (2016). Instrumento para determinar los tipos de evaluación utilizados por los profesores universitarios. *Revista Electrónica de Investigación Educativa*, 18(2), 40-52. Recuperado de <http://redie.uabc.mx/redie/article/view/1088>

Resumen

En este artículo presentamos el proceso de diseño y validación de un cuestionario para determinar los tipos de evaluación que emplean los profesores universitarios. La metodología seguida incluye desde la definición conceptual y operativa del dominio, hasta el análisis de los resultados del piloteo para determinar, por un lado, la confiabilidad del cuestionario y de cada pregunta y, por otro, su validez de constructo. Los resultados de este estudio aportan evidencia suficiente acerca de: 1) la validez de contenido del cuestionario, con un índice aceptable según los expertos (coeficiente alfa de Krippendorff mayor a 0.7); 2) La confiabilidad global del cuestionario equivalente a $\alpha=0.87$ y la confiabilidad de las preguntas entre $\alpha 0.81$ y $\alpha 0.87$; 3) la validez de constructo del cuestionario; 4) la estructura factorial no se aparta significativamente de la teoría que ha servido de base para la especificación de las preguntas. Los datos sugieren que el cuestionario es una herramienta válida y psicométricamente adecuada para su propósito.

Palabras clave: Educación Superior, Confiabilidad del Cuestionario, Validez de Contenido, Validez de Constructo.

Abstract

This paper presents the design and validation process of a questionnaire to determine the types of assessment used by members of faculty. The methodology used begins with the conceptual and operational definition of the domain and is followed through to the analysis of the pilot results to determine, on the one hand, the reliability of the questionnaire and each question and, on the other hand, the construct validity. The results of this study provide sufficient evidence to

¹ La investigación para la realización de este artículo ha sido posible gracias al patrocinio de Pearson Educación, S.A. de C.V.

show 1) the validity of the questionnaire content, which shows an acceptable index of agreement according to experts (with Krippendorff's alpha coefficient greater than 0.7); 2) the global reliability of the questionnaire, equivalent to $\alpha = 0.87$ and the reliability of the questions, between $\alpha 0.81$ and $\alpha 0.87$; 3) the construct validity of the questionnaire; 4) that the factorial structure is not significantly different from the theory used as a basis to establish the questions. The data suggests that the questionnaire is a valid tool and psychometrically sound for its purpose.

Keywords: Higher Education, Reliability of questionnaire, content validity, construct validity.

I. Introducción

La idea de investigar diferentes facetas de la evaluación no es nueva. Desde el mismo comienzo de la estructuración de la evaluación como una disciplina y como profesión, a comienzos de la segunda mitad del siglo XX, se percibió la necesidad de contestar, mediante estudios sistemáticos y preferentemente empíricos, diferentes preguntas sobre evaluación que permanecían sin respuesta. Muchos investigadores están de acuerdo en la necesidad de la investigación sobre las relaciones entre la teoría y la práctica evaluativa (Worthen, 1990; Shadish, Cook y Leviton, 1991; Smith, 1993; Pophan, 2014; Stiggings, 2014). Los beneficios del estudio empírico de la práctica evaluativa diaria son muchos y, entre ellos, cabe destacar la generación de la información necesaria para refinar la práctica evaluativa y desarrollar enfoques alternativos de evaluación (Smith, 1993), así como el avance de las concepciones sobre la conexión entre la teoría y la práctica evaluativa y el incremento de la comprensión de la influencia del contexto en la práctica de la evaluación.

Existen estudios centrados en la relación entre ciertos aspectos teóricos específicos y la práctica evaluativa. Se han considerado facetas tales como el examen y la valoración por los evaluadores de su propia práctica (Willians, 1988), las prácticas evaluativas del gobierno federal de los Estados Unidos (Wholey et al., 1970) y el uso de la evaluación (Alkin, Daillak y White, 1979), siendo esta última faceta la estudiada con más frecuencia (Shadish, Cook y Leviton, 1991; Smith, 1993).

Tras la llamada de Shadish, Cook y Leviton (1991), Smith et al. (1993) y Alkin (2004), en pro del conocimiento empírico sobre la práctica evaluativa para su perfeccionamiento y como factor esencial para el desarrollo de teorías relevantes y útiles, sigue hoy siendo importante objeto de atención para muchos investigadores.

Una primera clasificación o taxonomía de las facetas evaluativas objeto de las preguntas del cuestionario propuesto, y que ayudaría a clarificarlas, deriva de la consideración de los componentes lógicos de los modelos sistémicos de evaluación, como el Modelo de Calidad Educativa (De la Orden, 1997; De la Orden et al., 2002, y De la Orden, 2009) o el Modelo Contexto, Insumo, Proceso y Producto (CIPP, por sus siglas en inglés) (Stufflebean, 1982), esto es, las cuatro amplias categorías de *contexto* (circunstancias en que la evaluación tiene lugar), *entradas* (recursos humanos, económicos y materiales asignados a la evaluación), *procesos* (procedimientos empleados en la planificación, realización, disseminación y uso de los resultados de la evaluación) y *productos* (resultados de la evaluación en términos de información, juicios de valor y consecuencias de la propia evaluación; es decir, las decisiones sobre el objeto evaluado), dentro de las cuales caben todas las posibles facetas evaluativas susceptibles de ser objeto de investigación.

En este estudio, sin embargo, estamos interesados en lo que denominamos *tipos de evaluación*, o formas diferentes de evaluar asociadas a las grandes sub-funciones en que se manifiesta la función general de la evaluación, es decir, *la optimización del objeto evaluado*; en nuestro caso, el rendimiento académico de los estudiantes universitarios, valorado por los docentes en términos del aprendizaje logrado (conocimientos, habilidades, actitudes, o integradamente en las competencias) como efecto de la intervención educativa. De hecho, la evaluación cumple, como

manifestación diferencial de su función básica, diversas funciones específicas (sub-funciones en el contexto educativo). Así podemos hablar en educación de *evaluación diagnóstica*, *evaluación pronóstica*, *evaluación de acreditación*, *de certificación de calidad de un producto*, *de rendición de cuentas*, etc.

En general, se distinguen –según la función específica dominante– dos grandes categorías de uso que se han utilizando tradicionalmente, o tipos de evaluación que se conocen como *evaluación sumativa* y *evaluación formativa* (Scriven, 1967; Bloom 1971; Stufflebeam, 1971); sin embargo, recientemente se ha acuñado el término *assessment for learning* o *evaluación para el aprendizaje* (Stiggings, 2014) para hablar de la intención de la evaluación de involucrar a los estudiantes en la misma y como mecanismo de autorregulación.

Los factores determinantes de la diferenciación son, por una parte, la modalidad de las decisiones a que dan lugar, es decir, el uso que se pretende hacer de la evaluación en cada caso y, por otra, el *componente sistémico* en que se centra la evaluación (contexto, entrada, proceso o producto) que, a su vez, determina el *momento de su aplicación* en el continuo de la intervención educativa.

Las decisiones que definen la evaluación sumativa apuntan a la selección, clasificación, promoción de los alumnos y a la certificación de sus logros de aprendizaje. Este tipo de evaluación, que tiene generalmente carácter final, se centra básicamente en los resultados inmediatos, productos mediatos e impacto de la acción educativa.

Según los evaluadores de tendencia formativa, las decisiones específicas propias de este tipo de evaluación educativa apuntan directamente al incremento de la aproximación de la acción del profesor (enseñanza) de los logros de los alumnos (aprendizaje), lo que podría traducirse en un mayor logro académico, más altos estándares y un conocimiento más profundo. Este tipo de evaluación se centra más en el contexto, las entradas y en especial en el proceso de la enseñanza y el aprendizaje. En contraste con la evaluación sumativa, de carácter generalmente final, la evaluación formativa se aplica en los distintos estadios del proceso de enseñanza-aprendizaje y adopta formas muy diversas. El uso adecuado de sus resultados, a través de un *feedback* inmediato, directo y permanente, se convertiría en el mejor instrumento para asegurar la coherencia entre objetivos, proceso y producto de la enseñanza y el aprendizaje y, en consecuencia, para su optimización y calidad.

Por su parte, la evaluación para el aprendizaje, además de considerar los aspectos de una evaluación formativa, pone de relieve el papel activo del alumno durante todo el continuo para el logro de sus procesos autorregulativos tanto individualmente como con sus pares (Pophan, 2014; Stiggings, 2014)

Además de las decisiones, usos, componentes evaluativos en que se centran y el momento de su aplicación en el continuo educativo, citados como los principales factores que determinan la diferenciación de los tipos de evaluación, existen situaciones, procedimientos y otros elementos de la conducta evaluativa de profesores y alumnos asociados también a la diferenciación de los modos de evaluar, que resultó necesario conocer para elaborar las preguntas del cuestionario y analizar las respuestas de los profesores de modo que permitieran contestar satisfactoriamente a las cuestiones que plantea la investigación e interpretarlas correctamente como indicadores fundamentales de los distintos tipos de evaluación.

Para caracterizar más ajustadamente los tipos de evaluación, podríamos transformar los dos tipos generales *evaluación sumativa* y *evaluación formativa*, en que se manifiesta la función evaluativa general, en los siguientes cuatro tipos:

1. *Evaluación sumativa total*, caracterizada por:

- Certificación de los aprendizajes logrados por los estudiantes y procedencia de su

- promoción de nivel.
- Centrada en el producto final.
- Evaluación terminal.
- Uso de instrumentos de máxima objetividad.

2. *Evaluación sumativa parcial*, caracterizada por:

- La evaluación del logro educativo final como integración de logros parciales y procedencia de su promoción de nivel.
- Evaluación en períodos establecidos y terminal.
- Uso de instrumentos objetivos, exámenes escritos y orales, para la prueba final y las parciales, y análisis de productos.

3. *Evaluación Formativa*, caracterizada por:

- Certificación de aprendizajes y perfeccionamiento de procesos.
- Evaluación continua.
- Uso de instrumentos diversos: observación directa, análisis de productos, rúbricas, escalas, listas de cotejo, pautas evaluativas y preguntas orales, portafolios.

4. *Evaluación para el aprendizaje*, caracterizada por:

- Integración de enseñanza, aprendizaje y evaluación.
- Implicación de los alumnos en su propio progreso.
- *Feedback* inmediato y permanente.
- Apoyándose permanentemente en las respuestas de los alumnos, se va ajustando continuamente la instrucción.
- Implica regularmente a los alumnos en autoevaluación y coevaluación, manteniendo estándares constantes, de modo que puedan seguir su propio crecimiento a lo largo del tiempo y sentirse responsables de su propio éxito.

La tabla de especificaciones sobre modos y tipos de evaluación (tabla I) sintetiza y representa el dominio del cuestionario, es decir, los contenidos de referencia utilizados para elaborar las preguntas y recolectar la información prevista en el plan de investigación y como criterio que utilizaron los expertos para valorar las preguntas en la validación de contenido del instrumento. La tabla también sirvió como base del modelo teórico con el cual se contrasta el modelo de la estructura interna del cuestionario resultante del análisis factorial confirmatorio de los resultados de la aplicación piloto.

Tabla I. Especificaciones para la elaboración de las preguntas del cuestionario

	Componentes asociados tipos de evaluación	Sumativa	Formativa	Para el aprendizaje
1	Diagnóstico		X	X
2	Pronóstico		X	X
3	Rendición de cuentas	X		
4	Acreditación	X		
5	Evaluación de entrada		X	X
6	Evaluación de procesos		X	X
7	Evaluación de producto	X		
8	Evaluación terminal	X		
9	Evaluación continua		X	X
10	Referencia normativa	X		
11	Referencia criterial	X	X	X
12	Evaluación individual	X	X	X
13	Evaluación grupal	X	X	X
14	Implicación de los alumnos			X
15	Autoevaluación y Coevaluación			X
16	Heteroevaluación	X	X	
17	Integración evaluación-enseñanza		X	X
18	Evaluación en períodos establecidos	X	X	
19	Uso de un solo instrumento	X		
20	Uso de distintos tipos de instrumentos	X	X	X
21	Pruebas de ensayo	X	X	X
22	Pruebas de opción múltiple	X	X	X
23	Pruebas de realización	X	X	X
24	Información individual	X	X	X
25	Información colectiva	X	X	X
26	Retroalimentación individual	X	X	X
27	Retroalimentación colectiva	X	X	X
28	Criterios implícitos	X		
29	Criterios explícitos	X	X	X
30	Planificación	X	X	X

II. Metodología

Se utilizó la encuesta como método general, mediante la aplicación de un cuestionario rigurosamente elaborado (del que daremos cuenta posteriormente), a una muestra de docentes de las disciplinas y universidades previamente seleccionadas. En principio, se invitó a responder al cuestionario vía internet a todos los profesores universitarios de México y Centroamérica cuyo correo electrónico conocíamos por haber estado involucrados en procesos comunes de capacitación. La respuesta fue masiva, 900 profesores remitieron el cuestionario cumplimentado; sin embargo, las singulares características del muestreo nos impiden utilizar los datos para responder a nuestras preguntas con el rigor y objetividad necesarios. Es evidente, también, que estos datos son suficientemente válidos como excepcional ensayo piloto y, por consiguiente, utilizables como base para la validación del instrumento mediante el cálculo de algunos índices importantes de confiabilidad y validez.

Para determinar la validez de contenido del cuestionario se recurrió al juicio de expertos. En este caso, fueron seleccionados 10 profesores universitarios con formación teórica y experiencia en evaluación del rendimiento académico de los estudiantes en instituciones de educación superior, a quienes se les solicitó valorar las 50 preguntas del instrumento (tabla II). Para facilitar su tarea y estandarizar sus respuestas, se les remitió la matriz siguiente y las instrucciones.

Tabla II. Matriz con las preguntas del cuestionario a valorar por los expertos

CUESTIONARIO		1	2	3
1	La evaluación que utilizo está centrada en los resultados de aprendizaje.			
2	La evaluación que utilizo está centrada en los procesos.			
3	Utilizo tanto evaluación centrada en resultados como en procesos.			
4	Si al final del semestre se realiza un examen, lo elaboro sin la participación de otras personas.			
5	Si al final del semestre se realiza un examen, este es determinado por la institución.			
6	Si al final del semestre se realiza un examen, este es elaborado con mi participación.			
7	Comparo el resultado del examen, prueba u observación del desempeño, con criterios claros y objetivos, previamente establecidos y conocidos por los estudiantes.			
8	Comparo el resultado del examen, prueba u observación de su desempeño, con criterios claros y objetivos previamente establecidos; pero no conocidos por los estudiantes con anterioridad.			
9	Comparo los resultados de los alumnos con alguno de estos promedios: del grupo, de la generación, del equipo en el que trabajaron, nacional, de la universidad, de la facultad, de la carrera, del año escolar.			
10	Los porcentajes (la ponderación) para integrar la calificación final, son determinados por mí.			
11	Los porcentajes (la ponderación) para integrar la calificación final, no son determinados por mí.			
12	Los porcentajes (la ponderación) para integrar la calificación final, son determinados por la institución y por mí.			
13	Los porcentajes (la ponderación) para integrar la calificación final, son determinados conjuntamente por los alumnos y por mí, siguiendo las políticas institucionales.			
14	Tomo en cuenta, además de los exámenes que voy realizando, todos los trabajos desarrollados durante un período.			
15	Divido el período en cortes y voy asignando calificaciones parciales, mismas que se promedian para obtener la calificación final.			
16	Tomo en cuenta tanto los exámenes y trabajos, como las calificaciones parciales obtenidas en cada corte para integrar la calificación final.			
17	Utilizo con mis estudiantes la evaluación continua integrada en la propia enseñanza, con retroalimentación directa a los alumnos durante la realización de las tareas de aprendizaje.			
18	Utilizo con mis estudiantes la evaluación al terminar períodos específicos de aprendizaje, seguida de retroalimentación sobre el grado de logro de los objetivos.			
19	Utilizo con mis estudiantes tanto la evaluación continua integrada en la enseñanza, como la evaluación al terminar períodos específicos.			
20	Con la evaluación, promuevo la reflexión de los propios estudiantes sobre sus realizaciones.			
21	Con la evaluación continua contribuyo, al mismo tiempo, al aprendizaje de los estudiantes.			
22	Primero enseño y luego evalúo.			
23	Planifico al mismo tiempo las actividades de enseñanza y de evaluación, considerando los aprendizajes esperados y las evidencias de su logro.			
24	Primero planifico la enseñanza para un período y posteriormente las de evaluación, considerando los aprendizajes esperados y las evidencias.			
25	Cuando han trabajado en equipo, asigno una calificación al equipo en su conjunto.			
26	Cuando han trabajado en equipo, asigno una calificación a cada uno de los estudiantes por separado.			
27	Cuando han trabajado en equipo, en algunas ocasiones asigno una calificación al equipo en su conjunto y en otras, asigno una calificación a cada uno de los estudiantes por separado.			
28	Al evaluar, tomo en consideración las opiniones de los estudiantes sobre sus propios trabajos.			
29	Tomo en consideración las opiniones de los estudiantes sobre los trabajos de sus compañeros.			
30	Al evaluar, asigno calificación sin tomar en consideración las opiniones de los estudiantes.			
31	Para evaluar, utilizo exámenes tipo ensayo (de desarrollo) y preguntas abiertas: orales y escritas.			
32	Para evaluar, utilizo pruebas objetivas de elección de respuesta.			
33	Para evaluar, utilizo pruebas de realización (ejercicios prácticos).			
34	En mis actividades evaluativas, utilizo un solo instrumento (generalmente examen escrito).			
35	En mis actividades evaluativas, utilizo varios tipos de instrumentos.			
36	Considero que la evaluación debe ser objetiva al máximo.			
37	Acepto cierto grado de subjetividad en la evaluación.			
38	Considero que la evaluación debe ser objetiva; pero acepto cierto grado de subjetividad.			
39	Comunico a cada uno de mis estudiantes, individualmente, la información y el juicio valorativo resultantes de la evaluación.			
40	Comunico a los estudiantes, la información y el juicio valorativo grupalmente, pero mencionando los resultados de cada uno públicamente.			
41	Comunico los resultados, indiferentemente, en grupo o individualmente.			
42	Cuando trabajan en equipo, retroalimento individualmente a cada estudiante de forma oral o escrita.			
43	Cuando trabajan en equipo, retroalimento a los estudiantes mediante comentarios dirigidos a todo el equipo, de forma oral o escrita.			
44	Cuando trabajan en equipo, retroalimento tanto individualmente como al equipo en general.			
45	Devuelvo los trabajos evaluados con comentarios y sin calificación.			
46	Devuelvo los trabajos evaluados sin comentarios, pero con calificación.			
47	Devuelvo los trabajos evaluados con comentarios y calificación.			
48	Involucro a los alumnos para que hagan comentarios que contribuyan a la mejora de sus trabajos.			
49	No involucro a los alumnos en la realización de comentarios para la mejora de sus trabajos.			
50	Dejo en libertad el que hagan o no, comentarios a sus trabajos.			

Las instrucciones a los expertos fueron las siguientes:

- Por favor, para cada una de las 50 preguntas del cuestionario, que aparecen en la segunda columna de la matriz, señale Ud. el tipo de evaluación a que se refiere poniendo una X en la casilla correspondiente a la intersección de la fila de la pregunta con la columna que encabeza el tipo de evaluación que Ud. ha elegido como referencia de dicha pregunta.
 - Tipo de evaluación: 1= Evaluación sumativa
 - Tipo de evaluación: 2= Evaluación formativa
 - Tipo de evaluación: 3= Evaluación para el aprendizaje
- Solamente podrá elegir un tipo de evaluación como referencia de cada pregunta. Si Ud. opina que podría referirse a más de una, elija aquella en que le parezca más clara o frecuente dicha referencia; es decir, sólo debe clasificarse a cada pregunta en una categoría.

III. Resultados

El análisis de las valoraciones de los expertos se centró en la determinación del acuerdo entre jueces, para lo cual se aplicó el coeficiente alfa de Krippendorff tomando como criterio para mantener las preguntas en el cuestionario un nivel de acuerdo igual o mayor a 0.7. Con este criterio fueron eliminadas las preguntas correspondientes a los números 31, 32, 33, 49 y 50 del instrumento original.

Procedimos a calcular, por un lado, la confiabilidad del instrumento y de cada una de sus preguntas aplicando a los resultados un índice equivalente del alfa para escalas ordinales; y, por otro, acercarnos a la validez de constructo del cuestionario a través de un análisis factorial confirmatorio para contrastar la congruencia de la estructura factorial de los resultados con la explícita en la teoría que sirvió de base a la justificación del instrumento y, por tanto, a la tabla de especificaciones para la elaboración de las preguntas.

El alfa global del cuestionario es aceptable: 0.87 y, como muestra la tabla II, la confiabilidad de las preguntas varía entre 0.81 y 0.87. Por otra parte, la eliminación de alguna de las preguntas de más baja confiabilidad no se traduce en incremento significativo en la confiabilidad general del instrumento.

Tabla III. Confiabilidad de las preguntas del cuestionario

Ítem	Alfa	Ítem	Alfa	Ítem	Alfa
1	0.82	18	0.81	38	0.87
2	0.81	19	0.81	39	0.87
3	0.81	20	0.81	40	0.87
4	0.82	21	0.81	41	0.87
5	0.82	22	0.82	42	0.86
6	0.82	23	0.81	43	0.87
7	0.81	24	0.82	44	0.86
8	0.82	25	0.82	45	0.87
9	0.82	26	0.87	46	0.87
10	0.82	27	0.87	47	0.86
11	0.82	28	0.86	48	0.86
12	0.82	29	0.86		
13	0.82	30	0.87		
14	0.81	34	0.87		
15	0.82	35	0.86		
16	0.81	36	0.87		
17	0.81	37	0.87		

Para el análisis de componentes principales se ha utilizado el software “R” y, como los datos son producto de una medida ordinal, en lugar de aplicar el análisis basado en la correlación producto–momento de Pearson– se recurrió a un equivalente basado en la determinación de una correlación policórica, más coherente con el carácter ordinal de los datos de nuestro cuestionario. El valor de ajuste del modelo, 0.93, resulta aceptable para los propósitos del estudio.

El resultado del análisis tras rotaciones Varimax, Oblimin y Promax, identifican entre dos y tres factores que, en líneas generales, se acercan a la estructura señalada en la teoría en que se basa la construcción del cuestionario.

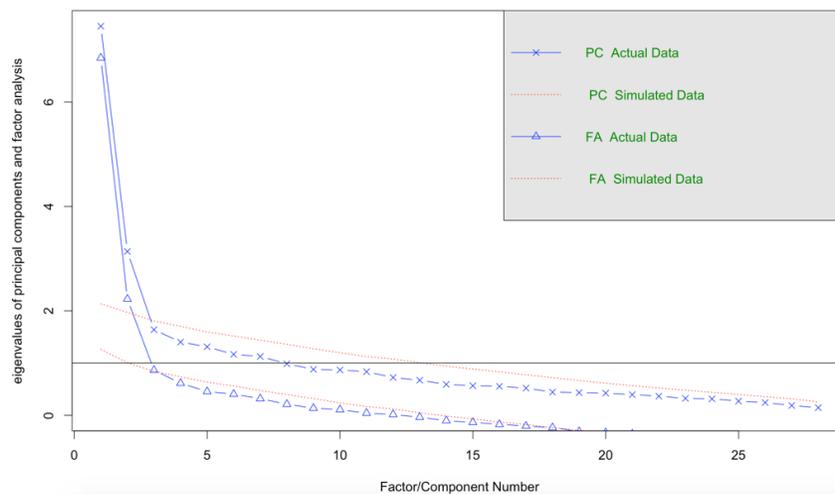


Figura 1. Número de factores/componentes

Tabla IV. Matriz factorial resultado de la rotación (Promax)

Ítem	Factor 1	Ítem	Factor 2	Ítem	Factor 3
21	0.7	45	0.5	34	0.6
17	0.7	8	0.5	24	0.5
19	0.7	40	0.5	30	0.4
20	0.7	46	0.4	1	0.4
18	0.6	26	0.4	22	0.3
7	0.5	29	0.4		
48	0.5	42	0.4		
3	0.5	41	0.4		
2	0.5	37	0.4		
14	0.4	27	0.4		
28	0.4	9	0.3		
23	0.4	38	0.3		
47	0.4				
16	0.4				
44	0.4				
39	0.4				

Tabla V. Proporción de varianza explicada por cada factor

	MR1	MR2	MR3
Cargas ss	4.47	2.22	1.42
Proporción de varianza	0.16	0.08	0.05
Varianza acumulada	0.16	0.24	0.29
Proporción explicada	0.55	0.27	0.18
Proporción acumulada	0.55	0.82	1.00

Tabla VI. Ajuste del modelo

La media de la complejidad (dificultad) del ítem= 1.2			
La prueba de la hipótesis que tres factores son suficientes			
Los (márgenes) grados de libertad para el modelo nulo son 378 y la función objetiva fue 7.38			
Los (márgenes) grados de libertad para el modelo son 297 y la función objetiva fue 1.91			
El valor cuadrático medio de los residuos (SMSR) es 0.05			
El grado de libertad del valor cuadrático medio corregido de los residuos es 0.6			
Ajuste basado fuera de los valores diagonales= 0.93			
Medidas de adecuación de puntuación del factor			
	MR1	MR2	MR3
Correlación de las puntuaciones con los factores	0.91	0.92	0.86
Raíz cuadrada múltiple de las puntuaciones con factores	0.88	0.83	0.75
Correlación mínima de posibles puntuaciones de los factores	0.75	0.65	0.49

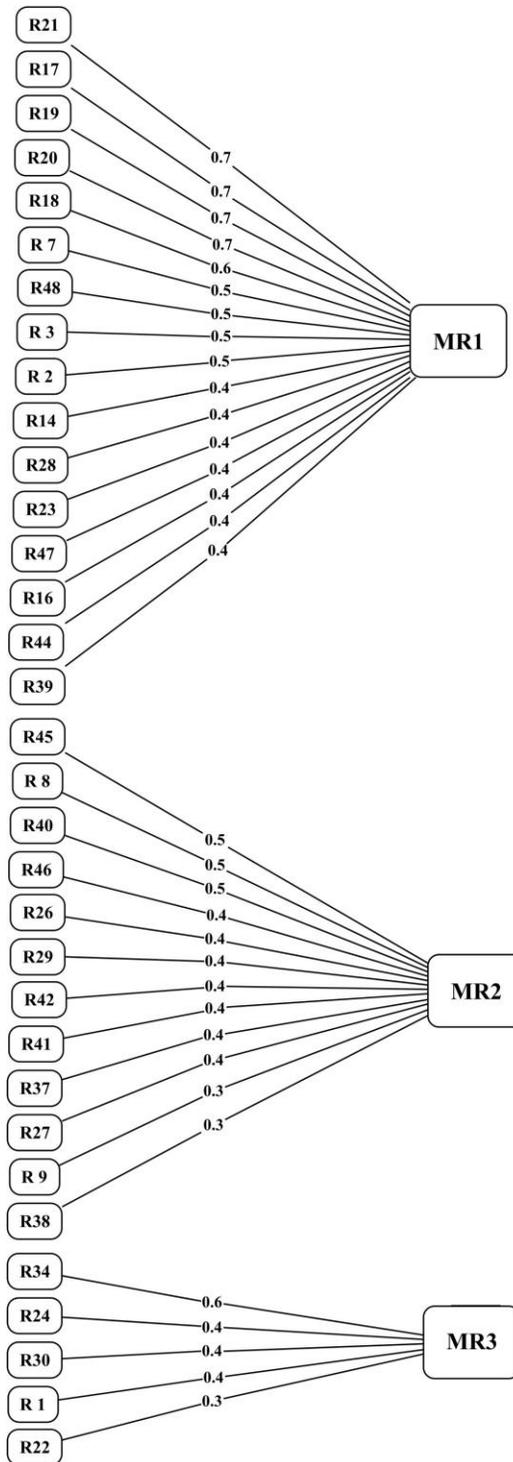


Figura 2. Representación del modelo con rotación Promax

Se han obtenido similares ajustes para las rotaciones Varimax, Oblimin y Promax, en la tabla VI se muestran los ajustes con rotación Promax. Del análisis confirmatorio se han obtenido con claridad tres factores (también es posible visualizarlos en la figura 1): el primero apunta una tendencia hacia una función formativa extrema, lo que se entiende como la aceptación del nuevo movimiento conocido como Evaluación para el Aprendizaje (*Assessment for learning*). Sin

embargo, las puntuaciones también confirman que los dos factores restantes apuntan respectivamente a las tradicionales funciones formativa y sumativa.

Tabla VII. Reactivos integrados a las dimensiones

Evaluación para los Aprendizajes	Evaluación Formativa	Evaluación Sumativa
<p>2. La evaluación que utilizo está centrada en los procesos.</p> <p>3. Utilizo tanto la evaluación centrada en resultados como en procesos.</p> <p>7. Comparo el resultado del examen, prueba u observación del desempeño, con criterios claros y objetivos, previamente establecidos; pero no conocidos por los estudiantes con anterioridad.</p> <p>14. Tomo en cuenta, además de los exámenes que voy realizando, todos los trabajos desarrollados durante un período.</p> <p>16. Tomo en cuenta tanto los exámenes y trabajos, como las calificaciones obtenidas en cada corte para integrar la calificación final.</p> <p>17. Utilizo con mis estudiantes la evaluación continua integrada en la propia enseñanza, con retroalimentación directa a los alumnos durante la realización de las tareas de aprendizaje.</p> <p>18. Utilizo con mis estudiantes la evaluación al terminar períodos específicos de aprendizaje, seguida de retroalimentación sobre el grado de logro de los objetivos.</p> <p>19. Utilizo con mis estudiantes tanto la evaluación continua integrada en la enseñanza, como la evaluación al terminar períodos específicos.</p> <p>20. Con la evaluación, promuevo la reflexión de los propios estudiantes sobre sus realizaciones.</p> <p>21. Con la evaluación continua contribuyo, al mismo tiempo, al aprendizaje de los estudiantes.</p> <p>23. Planifico al mismo tiempo las actividades de enseñanza y evaluación, considerando los aprendizajes esperados y las evidencias de su logro.</p> <p>28. Al evaluar, tomo en consideración las opiniones de los estudiantes sobre sus propios trabajos.</p> <p>39. Comunico a cada uno de mis estudiantes, individualmente, la información y el juicio valorativo resultantes de la evaluación.</p> <p>44. Cuando trabajan en equipo, retroalimento tanto individualmente como al equipo en general.</p> <p>47. Devuelvo los trabajos con calificación.</p> <p>48. Devuelvo los trabajos con comentarios.</p>	<p>8. Comparo el resultado del examen, prueba u observación de su desempeño, con criterios claros y objetivos previamente establecidos; pero no conocidos por los estudiantes con anterioridad.</p> <p>9. Comparo los resultados de los alumnos, con alguno de estos promedios: del grupo, de la generación, del equipo en el que trabajaron, nacional, de la universidad, de la facultad, de la carrera, del año escolar.</p> <p>26. Cuando han trabajado en equipo, asigno una calificación a cada uno de los estudiantes por separado.</p> <p>27. Cuando han trabajado en equipo, en algunas ocasiones asigno una calificación al equipo en su conjunto y en otras, asigno una calificación a cada uno de los estudiantes por separado.</p> <p>29. Tomo en consideración las opiniones de los estudiantes sobre los trabajos de sus compañeros.</p> <p>37. Acepto cierto grado de subjetividad en la evaluación.</p> <p>38. Considero que la evaluación debe ser objetiva, pero acepto cierto grado de subjetividad.</p> <p>40. Comunico a los estudiantes, la información y el juicio valorativo grupalmente, pero, mencionando los resultados de cada uno públicamente.</p> <p>41. Comunico los resultados, indiferentemente, en grupo o individualmente.</p> <p>42. Cuando trabajan en equipo, retroalimento individualmente a cada estudiante de forma oral o escrita.</p> <p>45. Devuelvo los trabajos evaluados con comentarios y sin calificación.</p> <p>46. Devuelvo los trabajos evaluados sin comentarios; pero, con calificación.</p>	<p>1. La evaluación que utilizo está centrada en los resultados de aprendizaje.</p> <p>22. Primero enseño y luego evalúo.</p> <p>24. Primero planifico la enseñanza para un período y posteriormente las de evaluación, considerando los aprendizajes esperados y las evidencias.</p> <p>30. Al evaluar, asigno calificación sin tomar en consideración las opiniones de los estudiantes.</p> <p>34. En mis actividades evaluativas, utilizo un solo instrumento (generalmente el examen escrito).</p>

La estructura obtenida no se aparta significativamente de la teoría que ha servido de base en la tabla de especificaciones y por consiguiente en la elaboración de las preguntas. Lo que en última instancia supone que no se corre mucho riesgo al afirmar que la estructura factorial derivada del análisis confirmatorio y la estructura factorial explícita en la teoría previa son similares o, al menos, compatibles.

Podríamos afirmar entonces, que el instrumento presenta un nivel aceptable de validez de constructo. Producto de la aplicación del análisis factorial, se eliminan las preguntas que realizan una escasa aportación a la diferenciación de los tipos de evaluación.

IV. Conclusión

En resumen, el estudio sugiere que el cuestionario es una herramienta válida y psicométricamente adecuada para coleccionar información pertinente de los profesores sobre los modos/tipos de evaluación del rendimiento académico de los estudiantes en las universidades. Se advierte que la evaluación para el aprendizaje y la formativa tienen cuestiones que las diferencian claramente, esencialmente la transmisión de la función de evaluación como una responsabilidad que compartiría el profesor con los estudiantes.

Referencias

- Alkin, M. C. (Ed.) (2004). *Evaluation roots. Theorists' views and influences*. Thousand Oaks, CA: Sage.
- Alkin, M. C., Daillak, R. y White, B. (1979). *Using evaluations: Does evaluation make a difference?* Thousand Oaks, CA: Sage.
- Bloom, B. S. (1969). Some theoretical issues relating to educational evaluation. En R. Tyler, *Educational evaluation: new roles, new Means. The sixty-eighth yearbook of the National Society for the Study of Education part II*. University of Chicago Press.
- Bloom, B. S., Hastings, J. T. y Madaus, G. F. (1971). *Handbook on formative and summative evaluation of student learning*. Nueva York: McGraw-Hill.
- De la Orden, A., Asensio, I., Carballo, R., Fernández Díaz, J., Fuentes, A., García Ramos, J. M. et al. (1997). Desarrollo y validación de un modelo de calidad universitaria como base para su evaluación. *RELIEVE*, 3(1). Recuperado de http://www.uv.es/RELIEVE/v3n1/RELIEVEv3n1_2.htm
- De la Orden, A., Garduño, L., Asensio, I. Mafokozi, J., González, C. y Muñoz, A. (2002). Funciones de la Universidad. *Revista de Educación*, 328, 267-281.
- De la Orden, A. (2009, junio). Evaluación y calidad: análisis de un modelo. *Estudios sobre Educación*, 16, 17-36.
- Pophan, W. J. (2014). *Classroom assessment. What teachers need to know*. Nueva Jersey: Pearson.
- Scriven, M. (1967). The methodology of evaluation. *AERA Monograph Series on Curriculum Evaluation*, 1.
- Scriven, M. (1991). *Evaluation Thesaurus* (4a. ed.). Thousand Oaks, CA: Sage.

Shadish, W. R. Jr., Cook, T. D. y Leviton, L. C. L. (1991). *Foundations of program evaluation*. Thousand Oaks, CA.: Sage.

Smith, N. L. (1993). Improving evaluation theory through the empirical study of evaluation practice. *Evaluation Practice*, 14(3), 237-242.

Stiggins, R. (2014). *Revolutionize assessment. Empower students, inspire learning*. Thousand Oaks, CA: Corwin.

Stufflebeam, D. L. y PDK National Study Committee on Evaluation. (1971). *Educational Evaluation and Decision Making*. Itasca, ILL: F. E. Peacock Publishers.

Stufflebeam, D. L. (1982). CIPP model for program evaluation. En G. F. Madaus, M. S. Scriven, y D. L. Stufflebeam (Eds.), *Evaluation models: viewpoints on educational and human services evaluation*, Boston: Kluwer-Nijhoff.

Wholey, J. S. et al. (1970). *Federal evaluation policy: analyzing the effects of public programs*. Washington, DC: Urban Institute.

Worthen, B. R. (1990). Program Evaluation. En H. J. Walberg y G. D. Haertel (Eds.), *International Encyclopedia of Educational Evaluation*. Nueva York: Pergamon.