

# Cómo perciben la evaluación por una prueba objetiva de la asignatura de LIC II, los estudiantes de segundo semestre de Biología de la FES Iztacala

Chirino Galindo Gladys<sup>1\*</sup>, Urbietta Ubilla Beatriz Rosalía<sup>1</sup>, Palomar Morales Martin<sup>1</sup>, Villanueva Santiago María Guadalupe<sup>1</sup> López Urrutia Roberto Eduardo<sup>2</sup>, Martínez García Martha<sup>2</sup>

<sup>1</sup>Universidad Nacional Autónoma de México, Carrera de Biología, Facultad de Estudios Superiores Iztacala, Avenida de los Barrios Número 1, Los Reyes Iztacala, Tlalnepantla, Estado de México, C. P. 54090, México.

<sup>2</sup>Universidad Nacional Autónoma de México, Laboratorio de Bioquímica Molecular, UBIPRO, División de Investigación y Posgrado, Facultad de Estudios Superiores Iztacala, Avenida de los Barrios Número 1, Los Reyes Iztacala, Tlalnepantla, Estado de México, C. P. 54090, México.

\*Autor para correspondencia: gchirinog@hotmail.com

## Recibido:

18/junio/2017

## Aceptado:

28/agosto/2017

## Palabras clave

Evaluación, Validez, Estrés

## Keywords

Evaluation, Validity, Stress

## RESUMEN

En éste trabajo se valoró la percepción que tienen los estudiantes de segundo semestre de la carrera de Biología de la FESI con respecto a la evaluación mediante una prueba objetiva en la asignatura de Laboratorio de Investigación Científica II, ya que se sabe que la evaluación debe ser objetiva y evidenciar el manejo de habilidades, destrezas y competencias. Se construyó un instrumento de 24 ítems, agrupados en 2 dimensiones: Percepción cognitiva del examen y reacciones del stress ante el examen. El instrumento fue validado por contenido, aplicación y constructo en una muestra de 68 alumnos voluntarios. El alfa de Cronbach dio un valor de 0.85, un índice KMO de 0.849, la esfericidad de Bartlett de 0.003. Con respecto a los estudiantes, éstos contestaron que los ítems muestran concordancia con la asignatura, el nivel de aprendizaje y los objetivos de su aplicación; y el nivel de estrés afecta el aprendizaje.

## ABSTRACT

In this work, the perception of students of the second semester of the career of Biology from the FESI with regard the evaluation across an objective test in the subject of Laboratorio de Investigación Científica II was valorated, since it is known the evaluation must be objective and evidence the management of skills, abilities and competencies. An instrument of 24 items, grouped in two dimensions was constructed: Cognitive perception of the test and stress reactions to the test. The instrument was validated by content, application and construct, in a sample of 68 voluntary students. The value of Cronbach's Alpha was 0.85, the KMO index was 0.849 and the Bartlett's sphericity was 0.003. With regard to the students, they answered that the items show concordance with the subject, the learning level and the objectives of its application, and the level stress affects learning.

## **Introducción**

Hasta hace pocos años, la investigación cualitativa, así como sus metodologías y técnicas eran inadvertidas, descalificadas o minimizadas por parte de la comunidad científica; sin embargo, la percepción de los científicos ahora ha cambiado, y la investigación cualitativa ha ido ganando prestigio en el mundo académico (Salgado, 2007).

Las técnicas cualitativas se utilizan para incrementar la fiabilidad o validez de un diseño cuantitativo, ya que aumentan la replicabilidad de las variables estudiadas, proporcionan un marco contextual y procesual para la manipulación experiencial, y refuerzan la validez de los resultados confirmando la relevancia de los constructos para situaciones reales; a su vez las técnicas cuantitativas, en estudios cualitativos, favorecen la posibilidad de realizar generalizaciones cuando la investigación se lleva a cabo en distintos contextos y contribuyen a la fiabilidad de los resultados cuando se emplean medidas estandarizadas para describir las variables de un contexto natural (Quecedo y Castaño, 2002).

El estudio a través de la percepción ha sido objeto de interés creciente dentro del campo de la Educación, porque hay juicios y opiniones en constante interacción y donde el individuo y la sociedad tienen un papel activo en la conformación de percepciones sobre el aprendizaje (Vargas, 1994). Hay estudios de percepción de los estudiantes en relación al desempeño de los docentes (Vásquez y Gabalán, 2006), sobre su actitud y su influencia en el aprendizaje (Tabera et al., 2015), y sobre el contexto escolar (Mateos, 2009); pero con respecto a un proceso evaluativo hay pocos estudios, aunque este aspecto comienza a tener importancia, dada la competencia en el desempeño laboral (Ricoy y Fernández-Rodríguez 2013).

Trillo y Porto (1999), reportan la percepción de los estudiantes en cuanto a la calidad de la evaluación, y expresan graves críticas sobre este proceso a nivel universitario. Ricoy y Fernández-Rodríguez (2013) mencionan que los estudiantes experimentan sensaciones contrapuestas, pues prevalecen las negativas como el nerviosismo, agobio, angustia, ansiedad, miedo, incertidumbre; sobre las positivas como son alivio y placer. Además, la percepción de los estudiantes, varía dependiendo del semestre en que se encuentren, debido a que los de primer semestre cuentan con menos elementos de juicio, y todavía no se encuentran familiarizados con el proceso de evaluación (Vázquez et al., 2012). De acuerdo a Tabera et al. (2015) los estudiantes señalaron que las actitudes y comportamientos docentes tenían un claro efecto sobre el clima de aprendizaje, e indicaron

la motivación y desmotivación como repercusiones principales en su evaluación.

Por ello, es necesario proponer estrategias para llevar a cabo la evaluación, así como contar con mecanismos que le permitan al docente promover en sus estudiantes la autoconfianza, para ir adquiriendo el dominio conceptual, de forma que les sea cotidiano, para que no se vea como una mera verificación del saber (Pereira-Chávez, 2015). Por estas razones, se planteó como objetivo del presente estudio construir y validar un cuestionario con el cual se valore la percepción que tuvieron los estudiantes del Laboratorio de Investigación Científica II de la carrera de Biología de la FES Iztacala, sobre la evaluación con una prueba objetiva aplicada en el semestre y sus reacciones (stress) ante ella.

## **Metodología**

El estudio fue cuantitativo descriptivo y de correlación porque proporcionó información sobre percepción, su finalidad fue describir la situación prevaleciente del fenómeno en el momento de realizarse el estudio y medir el grado de relación que existe entre dos o más conceptos o variables, la temporalidad fue transversal (Hernández-Sampieri et al., 2014).

### **Construcción del instrumento**

Los enunciados de los ítems, se elaboraron tomando en cuenta los siguientes parámetros: Que fueran claros, sencillos, fáciles de comprender, cortos, con relación con el tema de interés, redactados de forma personal y directa, sin palabras confusas ni valorativas (Murillo, 2006) y se elaboró el instrumento. Se concretaron 2 dimensiones, las cuales fueron: Percepción cognitiva de la prueba objetiva y las reacciones del stress ante la prueba objetiva de la asignatura experimental. Los enunciados fueron valorados de acuerdo a una escala Likert de 5 grados: 1=Totalmente en desacuerdo; 2=En desacuerdo; 3=Ni de acuerdo ni en desacuerdo; 4=De acuerdo y 5=Totalmente de acuerdo, para los enunciados del apartado de percepción cognitiva, y se asignaron puntuaciones de uno a cinco; mientras que para el de reacciones del stress, los grados de la escala fueron: 1=Nunca; 2=Algunas veces; 3=Frecuentemente; 4=Siempre; y se asignaron puntuaciones de uno a cuatro, quedando conformado un instrumento temporal de 30 ítems.

### **Validez de contenido**

La validez de contenido es una condición necesaria para ejecutar interpretaciones de los puntajes en los tests

(Kane, 2009). Asimismo, ésta no se refiere solamente a los ítems del instrumento de medida, también incluye las instrucciones para su gestión y los criterios para su corrección y puntuación (Abad et al., 2011). Por lo cual se realizó con el método, conocido como Panel de Evaluación del Contenido, basado en la evaluación individual de los ítems de un test por parte de expertos en la materia. Posteriormente, mediante la Razón de Validez de Contenido (RVC, Coefficient Validity Ratio), se comprueba qué ítems del instrumento son adecuados y deben mantenerse en la versión final del mismo. En este punto, se debe asignar a cada ítem una puntuación en base a tres posibilidades: 1). Que el ítem sea esencial para evaluar el constructo, 2). Que el ítem resulte útil pero prescindible o 3). Que el ítem se considere innecesario. Sobre esta valoración se aplica la siguiente expresión:

$$RVC = \frac{n - \left(\frac{N}{2}\right)}{\frac{N}{2}}$$

donde n es el número de expertos que otorgan la calificación de esencial al ítem y N, el número total de expertos que evalúan el contenido (Pedroza et al., 2014). Se utilizaron el criterio de 5 expertos, doctores pertenecientes al área de Investigación Educativa de la FES Iztacala, con el objetivo de discriminar si existía diferenciación clara de las dimensiones (percepción cognitiva de la prueba objetiva y reacciones de estrés ante la prueba objetiva) a través de los ítems, y si había correspondencia entre éstos, también si los enunciados en su conjunto medían la percepción hacia la prueba objetiva de la asignatura y si fueron elaborados con un lenguaje claro, preciso y comprensible para el nivel de los estudiantes, un ejemplo de los ítems se muestran en la tabla 1. Posteriormente, se analizaron los datos aportados por los expertos, lo que permitió seleccionar los mejores ítems. De aquí se conformó el instrumento definitivo y se procedió a la aplicación.

**Tabla 1.** Ejemplos de los ítems de las dimensiones de percepción cognitiva de la prueba objetiva (1 y 2) y reacciones de estrés ante la prueba objetiva (3 y 4).

1. "Las preguntas de la prueba objetiva son adecuadas para mi nivel de conocimiento".
2. "Los objetivos de la prueba objetiva son claros".
3. "La preocupación por la evaluación de la prueba objetiva me hace perder o aumentar el apetito".
4. "No puedo dormir, por pensar que puedo fracasar en la evaluación de la asignatura".

### Validación por aplicación

La aplicación del instrumento se realizó en la FES Iztacala en el mes de abril de 2017. Se tomó una muestra simple aleatoria de 63 alumnos voluntarios, con un rango de edad de 18 a 22 años, de segundo semestre de la Carrera de Biología. El cuestionario se aplicó a los estudiantes en una plataforma de uso libre *Online form Builder Jot-Form*, con un tiempo para su resolución de 20 minutos; se les dio un plazo de 4 días para que los alumnos entraran a la plataforma y resolvieran el cuestionario. La recopilación y ordenamiento de las respuestas se realizó con un Software especializado *Remark Office OMR Data Center Versión 2014*.

### Confiabilidad

Para calcular la confiabilidad de la escala, se empleó el Alfa de Cronbach de consistencia interna. El alfa de Cronbach permite estimar la fiabilidad de un instrumento de medida, asumiendo que los ítems (medidos en escala tipo Likert) miden un mismo constructo y que están altamente correlacionados (Welch y Comer, 1988; Bojórquez et al., 2013). Cuanto más cerca se encuentre el valor del alfa a 1 mayor es la consistencia interna de los ítems analizados. La fiabilidad de la escala debe obtenerse siempre con los datos de cada muestra para garantizar la medida fiable del constructo en la muestra concreta de investigación.

### Validez de constructo

Un constructo se define como un concepto formulado en forma deliberada con objetivos científicos, que tiene dos características: a) se vincula con otros constructos a través de hipótesis, entonces puede utilizarse en la investigación científica (aspecto relacional), y b) es sujeto de observación y medición (aspecto reductivo), La inteligencia, rendimiento, agresión, emoción, percepción, juicio, tendencia, son ejemplos de esto (Gras, 1880; Abreu, 2012). Un constructo tiene propiedades que son subyacentes, a las cuales no se les puede medir en forma directa, solo se miden a través de manifestaciones externas de su existencia, o sea, usando indicadores (Briones, 1996).

Para validar el constructo mediante el análisis factorial, se usó el programa SPSS (Pérez-Gil, 2000; abad et al., 2011). Para poder aplicar ésta técnica es necesario que exista una correlación entre las variables del instrumento, se utilizan dos medidas que permiten determinar la viabilidad de ésta: la prueba de adecuación muestral de Kaiser-Meyer-Olkin (KMO) y la prueba de esfericidad de Bartlett. Según Carvajal et al. (2011) y Abad et al., (2011), el KMO indica el grado de intercorrelación de las

variables, si es superior a 0.7 se considera viable, y entre más cerca de 1, excelente será la adecuación muestral. Por otro lado, la prueba de esfericidad de Bartlett, indica que existe correlación entre todas las variables, por lo que el análisis factorial puede llevarse a cabo. Un análisis factorial se considera bueno si se obtiene un nivel de significancia menor de 0.05. Con la finalidad de perfeccionar los resultados, se rotaron cada uno de ellos bajo el método de componentes principales con rotación Varimax (Lozares y López, 1991), excluyendo los reactivos que tuvieran un peso factorial por debajo de 0.30 (Ferrando y Anguiano-Carrasco, 2010).

### **Percepción**

Primeramente, se contabilizaron las respuestas y se graficaron todos los porcentajes por cada pregunta, para ver resultados preliminares. Posteriormente, una vez validado el instrumento, se ordenaron y sumaron los puntajes con las respuestas de los estudiantes por cada reactivo, y para realizar un análisis de discriminación entre las respuestas para medir la percepción con respecto a la evaluación, se tomó en cuenta la prueba de la mediana (Ruiz, 2005), la cual diferencia entre los puntajes altos y bajos de cada uno de los reactivos, y el puntaje total del instrumento, por lo cual se estableció un punto de corte, considerando la mediana de la sumatoria total de las respuestas, y los cuestionarios resueltos se agruparon en “arriba de la mediana” y “abajo de la mediana”. Consecutivamente, para cada respuesta se agruparon también los datos arriba y abajo de la mediana. Al final, se efectuó un análisis de chi-cuadrado por cada respuesta con respecto a las respuestas totales del área, con el fin de determinar, para cada respuesta, que tan alejado se encontraba el grupo con respecto a la calificación global del área.

### **Análisis estadístico**

El análisis de los datos se realizó con el paquete estadístico SPSS versión 23.0, con un índice  $\alpha$  de 0.05 de significancia. La significancia estadística es un parámetro que permite asegurar que, los resultados obtenidos no se deben al azar o a procedimientos de muestreo o de tratamiento inadecuados o sesgados (Daniel, 2002).

## **Resultados y discusión**

### **Construcción del instrumento y Validez de contenido**

Las estimaciones efectuadas por los expertos fueron objetivas, el 85% de las opiniones fueron de común acuerdo, concertando en la disminución del número de ítems y en la compilación de estos según las categorías.

De acuerdo a la Razón de Validez de Contenido (RVC), el 43.3% de los ítems, lograron los puntajes más altos, es decir que obtuvieron un 100% de respuestas en: “Que el ítem sea esencial para evaluar el constructo”; el 36.7% de los ítems, se encontraron en “Que resulte útil pero prescindible”, y las puntuaciones más bajas representan un 20% de los ítems en “Que el ítem se considere innecesario”; seguidamente de la validación de especialistas y siguiendo sus observaciones, se procedió a eliminar y/o cambiar otros aspectos de los ítems, pero sin que la estructura general se alterada. Al final resulto un instrumento de 24 ítems, en forma de enunciados, clasificados en 2 categorías con 12 proposiciones sobre percepción cognitiva de la prueba objetiva y 12 de reacciones del stress ante la prueba objetiva; se descartaron los 6 que obtuvieron puntuaciones bajas y se volvieron a numerar. Según Escofet et al., (2016) la validez de contenido de los ítems consistió en traducir las dimensiones en elementos medibles; es decir, pasar de las dimensiones a los indicadores y de los indicadores a las preguntas. La validación por los expertos permitió pasar de lo teórico a lo empírico para poder medir lo que se quiere. En el diseño de una escala, según Martín (2004), se considera que el número mínimo de ítems para evaluarla es de 6, pero el número total puede ir de 10 a 90, de manera que puedan abarcar de forma proporcional cada una de las dimensiones definidas a priori en el constructo.

### **Validación por aplicación**

Del total de la población de estudiantes de segundo semestre de la Carrera de Biología, la población encuestada corresponde al 22.7%; de los cuales, el 54% fueron mujeres y el 46% fueron hombres; el rango de edad fue de 18 a 22 años, el promedio de 20 años. Según Martín (2004), para llevar a cabo una validación por estudio piloto, se aconseja que lo mínimo sea de 30-50 personas, y que los individuos de la muestra, se parezcan a la población dirigida; en este trabajo, se cumple con esa condición, y nos permite identificar, si existen resistencias psicológicas o rechazo hacia algunas preguntas, si el ordenamiento interno es lógico; si la duración está dentro de lo aceptable por los encuestados.

### **Confiabilidad**

La primera vez que se realizó la prueba estadística Alfa de Cronbach, arrojó un valor de correlación de 0.69 con los 24 ítems, posteriormente, cuando se hicieron las correlaciones, se eliminaron 4 ítems, por su baja correlación, se realizó de nuevo el análisis de consistencia interna y dio un valor de 0.85, que representa una adecuada confiabilidad, e indica como buena la

consistencia interna en la percepción de los estudiantes con respecto a su evaluación. Según Carvajal et al. (2011), La consistencia interna mide la homogeneidad de los enunciados de un instrumento ya que indica la relación entre ellos, y esta consistencia interna en el intervalo 0,8-0,9 se califica como de un nivel bueno.

En la Tabla 2, se observan los resultados de la estadística descriptiva del instrumento, las correlaciones de los ítems, son significativas (>0.3), entre la mayoría de ellos y el puntaje total, menos en 4 de ellos (9, 13, 16 y 19) que obtuvieron correlaciones por debajo de la significancia, (Arechabala y Miranda, 2002); Se observa que la escala se conserva estable entre 0.78 y 0.84 sin los ítems eliminados.

**Tabla 2.** Estadística descriptiva de los 24 ítems del instrumento de evaluación. I (Ítems), M (Media), DE (Desviación estándar), AC (Alpha de Cronbach con el ítem), CIT (Correlación ítem test). \*Los ítems muestran baja correlación y fueron eliminados.

I	M	DE	AC	CIT
1	3.80	0.77	0.80	0.568
2	4.12	0.65	0.83	0.526
3	3.92	0.74	0.80	0.589
4	3.87	0.92	0.82	0.578
5	4.22	0.72	0.83	0.573
6	3.80	0.82	0.81	0.461
7	3.66	0.84	0.82	0.436
8	2.92	0.90	0.83	0.591
9*	3.66	0.82	0.68	0.236
10	2.84	1.03	0.8	0.425
11	3.30	0.97	0.84	0.539
12	4.01	0.88	0.83	0.599
13*	2.65	0.84	0.63	0.254
14	2.60	0.73	0.82	0.547
15	2.06	0.98	0.81	0.485
16*	2.57	0.99	0.65	0.164
17	2.60	1.18	0.82	0.456
18	1.92	0.93	0.8	0.545
19*	2.93	0.91	0.6	0.214
20	2.73	0.91	0.82	0.605
21	2.42	0.96	0.83	0.561
22	2.39	1.08	0.82	0.587
23	3.01	0.97	0.79	0.481
24	2.42	0.68	0.78	0.467

**Validez de constructo**

En este estudio se obtuvo un KMO de 0.849 y una prueba de esfericidad de Bartlett de 0.003, esto indica la posibilidad de realizar de manera factible el análisis factorial. Se hizo el análisis con 20 reactivos (ya que se eliminaron 4), que obtuvieron un nivel adecuado de confiabilidad, con

el método de los componentes principales con rotación Varimax, la cual es una técnica estadística de análisis multivariado que permite seccionar la información contenida en un conjunto de p variables de interés en m nuevas variables independientes, cada una explica una parte específica de la información y mediante combinación lineal de las variables originales otorgan la posibilidad de resumir la información, total en pocas componentes que reducen la dimensión del problema (León et al. 2008, abad et al., 2011). Después de varias interacciones, se formaron 4 factores y una varianza explicada total de 54.68%; como se observa en la Tabla 3.

**Tabla 3.** Análisis multivariado de los ítems y los cuatro componentes principales a partir del método Varimax. Varianza explicada (VE), Varianza total (VT).

Componentes	1	2	3	4
Ítem 1	0.657			
Ítem 2	0.598			
Ítem 3		0.687		
Ítem 4		0.625		
Ítem 5		0.631		
Ítem 6	0.582			
Ítem 7		0.594		
Ítem 8	0.678			
Ítem 10		0.653		
Ítem 11	0.546			
Ítem 12	0.567			
Ítem 14			0.705	
Ítem 15				0.712
Ítem 17				0.651
Ítem 18				0.587
Ítem 20			0.654	
Ítem 21			0.583	
Ítem 22				0.614
Ítem 23				0.575
Ítem 24			0.726	
Varianza explicada	20.61%	15.24%	10.89%	7.94%
Varianza total 54.68%				
Alfa de Cronbach	0.89	0.86	0.84	0.81
Alfa de Cronbach total 0.85				

El primer factor se diferencia por contener información sobre la comprensión y habilidades en su evaluación, formado por 6 ítems (1, 2, 6, 8, 11 y 12), que contribuyen al 20.61% de la varianza total explicada y una confiabilidad de alfa de Cronbach de 0.890. El segundo factor, integra ítems relativos a las instrucciones del examen para evaluación, el cual quedó formado por 5 ítems (3, 4, 5, 7

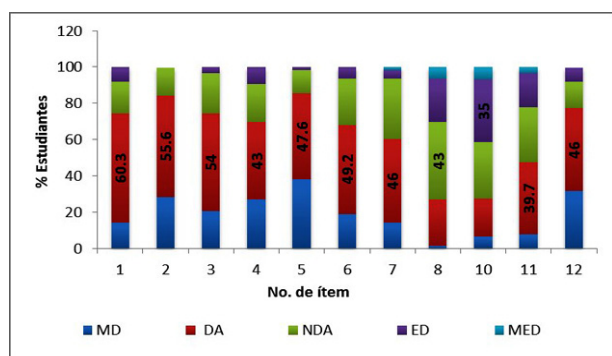


y 10), con una aportación a la varianza total explicada de 15.24% y un nivel de confiabilidad de alfa de Cronbach de 0.860. El tercer factor aborda la ansiedad que les produce el estrés por la evaluación; aquí se agruparon 4 ítems (14, 20, 21 y 24), que presentan una varianza de 10.89% y un nivel de confiabilidad de alfa de Cronbach de 0.840. El cuarto factor, referido a las reacciones físicas que produce el estrés ante la evaluación, se formó con 5 reactivos que tienen una varianza de 7.94% y un alfa de Cronbach de 0.810; como se observa en la tabla 3.

Según Martín (2004), la validez de constructo, es la más importante ya que evalúa el grado en que el instrumento refleja la teoría del fenómeno o del concepto que mide. Esta validez garantiza que las medidas que resultan de las respuestas del cuestionario pueden ser consideradas y utilizadas como medición del fenómeno que se quiere medir.

### Percepción

Con respecto a la percepción cognitiva del examen (factor 1 y 2), los estudiantes estuvieron de acuerdo con que los ítems reflejan la comprensión de la asignatura, su nivel de aprendizaje, y los objetivos de su aplicación, de manera que representan un porcentaje mayor al 50%. Los alumnos también están de acuerdo con que existan preguntas específicas por unidad, el grado de dificultad de los ítems y que sean de aplicación y análisis, además del motivo, validez y reconocimiento por su colaboración (entre el 40 y 50%). El 43% de ellos percibe que interno al cuestionario puede o no haber preguntas redundantes (Fig. 1).



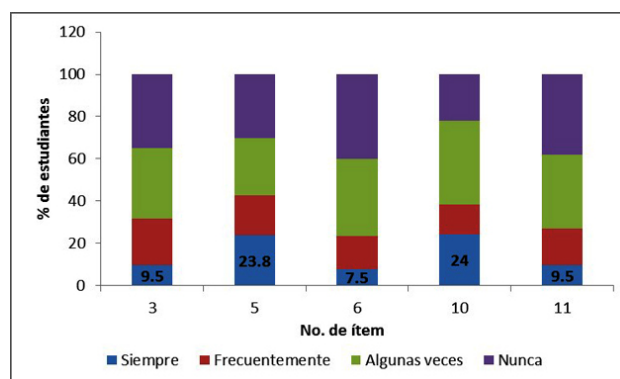
**Figura 1.** Percepción de los estudiantes de Biología con respecto a los exámenes (factor 1 y 2). Muy de acuerdo (MD), De acuerdo (DA), Ni en desacuerdo ni de acuerdo (NDA), en desacuerdo (ED) y muy en desacuerdo (MED).

Pérez (2005), nos dice que la evaluación debe presentar un vínculo entre el fin y los objetivos perseguidos desde sus variadas dimensiones, sin olvidar que es un

elemento clave en todos los procesos vinculados con la calidad educativa. Por ello, es necesario trabajar desde su concepción, su diseño, su ejercicio y el uso que se hace de la información recogida.

Por otra parte, Brown (2003) menciona que los procesos evaluativos son los medios para que el profesor aprenda a partir de la evaluación de sus estudiantes; con lo cual se espera que desarrolle habilidades personales, criticidad, capacidad de escuchar y adaptarse a los cambios cotidianos. La evaluación permite reflexionar acerca de la importancia conceptual, procedimental y actitudinal de la enseñanza-aprendizaje y resalta que los docentes utilizan variadas maneras de examinar y reflexionar sobre el trabajo de sus estudiantes dentro y fuera de sus aulas. Pero también los docentes han tenido la iniciativa de ir más allá de calificar, para poder reflexionar, por lo que es importante romper el aislamiento de la práctica docente en el aula y crear una comunidad de reflexión cooperativa (Allen, 2006).

Las reacciones de los estudiantes con respecto al estrés que les produce el examen como herramienta de evaluación, fueron diversas, entre el 7.5 y el 24% siempre pierden el apetito, la cercanía de la evaluación les produce ansiedad, insomnio, pierden la motivación y tienen sensación de falta de capacidad para terminar la carrera (Fig. 2).



**Figura 2.** Percepción de los estudiantes de Biología con respecto al estrés (factor 4) producto de la evaluación.

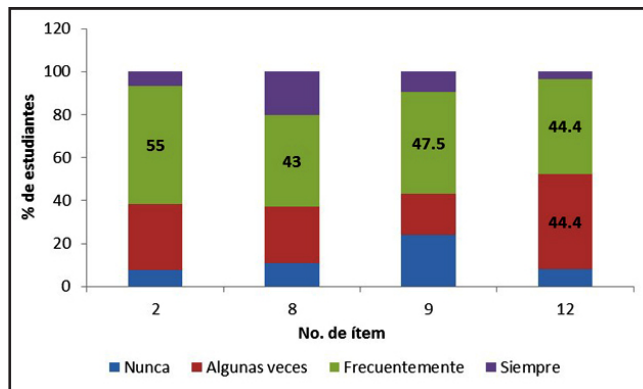
Según Andrade (2011), la ansiedad ante los exámenes, que es la forma más común de evaluación, afecta el rendimiento académico, debido a que éste estado de ánimo interfiere en la recuperación de la información que el alumno debe realizar cuando está contestando un examen. La característica de esta reacción es principalmente la preocupación por el posible fracaso,

o desempeño inadecuado en la tarea que se evalúa, que traería como consecuencia una calificación baja afectando la autoestima del alumno y seguramente una devaluación social que impacta primeramente a la familia y al grupo social al que pertenece el alumno.

Phinder-Puente (2014) indica que los niveles de estrés y ansiedad de estudiantes en medicina, son más elevados en época de exámenes, si bien también se incrementan en el tiempo fuera de exámenes.

En un estudio realizado en Escocia en estudiantes universitarios se observó un incremento significativo en niveles de depresión, síntomas obsesivos y pérdida de concentración en todos ellos tras seis semanas de permanencia en la Universidad, tiempo en el cual aún no han tenido oportunidad de realizar exámenes (Fisher y Hood, 1987), indicando que el fenómeno del estrés académico puede no ser debido exclusivamente a la realización de exámenes, aunque éste sigue siendo un elemento sumamente importante.

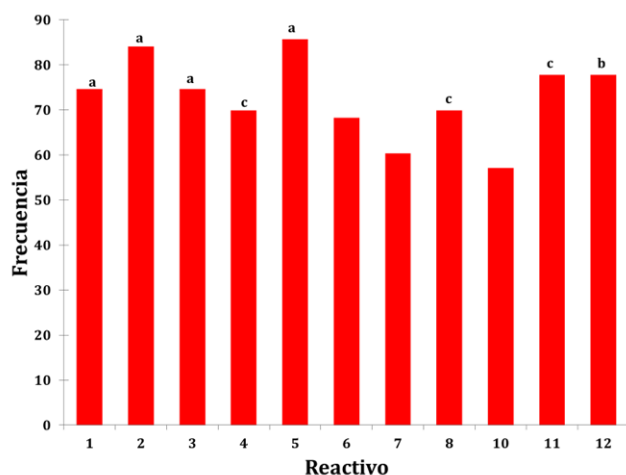
Los estudiantes valoraron el examen y mencionan que frecuentemente (40 al 55%) sus calificaciones son bajas por no tener adecuados método de estudio, falta de tiempo para estudiar, concentración para realizar las tareas y sobrecarga de trabajo académico (Fig. 3). Es de importancia que el estudiante identifique sus errores, las dificultades de aprendizaje y pueden desarrollar y aplicar estrategias posteriormente para superarlos; facilitando el proceso de enseñanza-aprendizaje como lo menciona Porto (2009) en su estudio sobre la percepción de los universitarios de Santiago, encuentra que los encuestados afirman que es necesario evaluar el conocimiento adquirido, pero de ese conocimiento lo que se evalúa más frecuentemente es la solución de los ejercicios o la reproducción de la información, con independencia de que se haga memorísticamente, ya que hay división de opinión sobre la frecuencia con la que se evalúan la calidad de los procesos de razonamiento, creación, síntesis, etc. Prácticamente nunca o en pocas ocasiones se considera necesario evaluar las condiciones de estudio de los alumnos (espacio, tiempo, recursos, etc.); sus conocimientos previos; esfuerzo; la actitud e interés del alumnado; ni las capacidades de cada persona (madurez, formación, etc.). De esta forma se comprueba como la evaluación, además de considerarse únicamente como la realización de un examen y su calificación, se sigue reduciendo únicamente a los resultados finales (resultados buscados y reducidos a contenidos conceptuales), y se presta más atención a los errores que a los aciertos.



**Figura 3.** Percepción de los estudiantes de Biología con respecto al estrés (factor 3) producto de la evaluación.

En relación a la prueba de la mediana, para el primer grupo de preguntas, 52.4% de los estudiantes contestaron por arriba de la mediana, y 47.6% por abajo. Las preguntas con diferencia significativa por la prueba de la mediana, fueron la pregunta 1 “Que tanto comprendo la asignatura”, la 2 “mi nivel de conocimiento es adecuado a la evaluación”, la 3 “los objetivos de la evaluación son claros”, la 4 “se explican los motivos del examen”, la 5 “se indica si el examen tiene o no validez”, la 8 “si hay o no preguntas redundantes en el examen”, la 11 “el examen es de aplicación y análisis” y la 12 “el grado de dificultad de las preguntas son adecuadas a lo que se revisó en clase”; en las cuales contestaron por encima de la media el 74.6, 84.1, 74.6, 69.8, 85.7, 69.8, 77.8 y 77.8 de los estudiantes, y como se observa en la figura 4, son las respuestas que contrastan con la mediana.

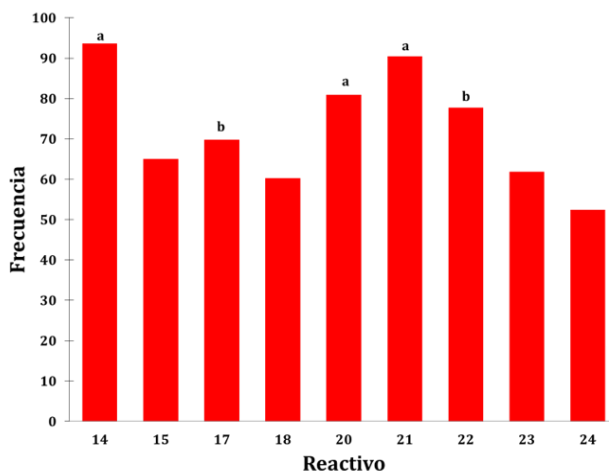
Foster (1991) asevera que el planeamiento y la evaluación de la enseñanza permiten al docente observar si el estudiante comprende la materia de clase y la relaciona con destrezas que le permiten ascender a niveles más elevados: ésta es la esencia de la docencia y el aprendizaje. El nivel de habilidad, conocimiento y de destrezas, ha de ser evaluado a intervalos regulares para determinar el progreso de los estudiantes y también, para estudiar otras acciones que le permitan al estudiante superar las áreas o los capítulos más problemáticos de la clase. Los docentes se deben presentar imparciales en sus ponderaciones sobre el rendimiento de los estudiantes. Además, las pruebas han de mantenerse relacionadas con el contenido estudiado en clase y con el tiempo de estudio de los estudiantes (Acevedo, 2004).



**Figura 4.** Frecuencia de respuestas del cuestionario sobre la percepción cognitiva de los estudiantes (primeras 12 preguntas). Por claridad, sólo se representa la frecuencia de alumnos que contestaron por arriba de la mediana. <sup>a</sup>Diferencia ( $p < 0.001$ , evaluada por chi-cuadrada) entre la respuesta a la pregunta individual, y la respuesta global a las 12 preguntas (33 por arriba de la mediana). <sup>b</sup>Diferencia ( $p < 0.01$ ) evaluada por chi-cuadrada) entre la respuesta a la pregunta individual y la respuesta global a las 12 preguntas. <sup>c</sup>Diferencia ( $p < 0.05$ ) evaluada por chi-cuadrada) entre la respuesta a la pregunta individual y la respuesta global a las 12 preguntas. Se excluyeron las preguntas que no cumplieron los criterios de validez.

De igual manera, para el segundo grupo de preguntas, 52.4% de los estudiantes contestaron por arriba de la mediana, y 47.6% por abajo. En cuanto a las preguntas 14 “mi evaluación es baja, porque no tengo un método de estudio”, la 17 “me provoca ansiedad no obtener una buena evaluación”, la 20 “no tengo tiempo para estudiar”, la 21 “no me concentro en actividades de estudio”, la 22 “me falta motivación por parte de los profesores”, los estudiantes respondieron 93.6, 69.8, 81.0, 90.5, y 77.8% por arriba de la mediana, respectivamente, como se muestra en la figura 5.

Omar (2004) encuentra resultados parecidos al nuestro, señala que en su estudio los alumnos que se consideran de mal rendimiento y los “regulares” atribuyen su rendimiento empobrecido a sus actitudes personales negativas (no tengo métodos de estudio, no me concentro) y a sus actitudes sociales negativas (tengo mala conducta, soy charlatán, soy vago, pierdo tiempo ayudando o pidiendo ayuda, integro grupos de estudio y en lugar de estudiar espero que los demás hagan las cosas por mí).



**Figura 5.** Frecuencia de respuestas del cuestionario sobre la percepción de estrés de los estudiantes (últimas 12 preguntas). Por claridad, sólo se representa la frecuencia de alumnos que contestaron por arriba de la mediana. <sup>a</sup>Diferencia ( $p < 0.001$ , evaluada por chi-cuadrada) entre la respuesta a la pregunta individual, y la respuesta global a las 12 preguntas (33 por arriba de la mediana). <sup>b</sup>Diferencia ( $p < 0.01$ ) evaluada por chi-cuadrada) entre la respuesta a la pregunta individual y la respuesta global a las 12 preguntas. Se excluyeron las preguntas que no cumplieron los criterios de validez.

Para mejorar el aprendizaje y el rendimiento académico de los estudiantes, sobre todo en la educación superior, la motivación juega un papel esencial en el proceso de desarrollo de la sociedad. Por lo cual es crucial en ellos, por lo que hay que desarrollar un mejor entendimiento de los factores de la motivación académica (Hernández et al. 2012).

Coincidimos con Sánchez (2014), en que para algunos docentes, estudiar e impartir conocimientos, resulta algo reflexivo y constructivo, pero para la mayoría de estudiantes, es sinónimo de abandono y desinterés; esta percepción puede ser entendida como una falta de pasión hacia lo que hacemos, y si no sentimos entusiasmo al realizar nuestra docencia, nuestro trabajo no tendrá significado y esto no lo transmitiremos a nuestros estudiantes.

### Conclusiones

El instrumento de percepción de la evaluación con una prueba objetiva, es confiable y válido por contenido, constructo y aplicación; compuesto por cuatro factores, con una varianza explicada de 54.68%, esto significa que mide de forma adecuada la percepción de los estudiantes



y que puede aplicarse con otros alumnos de nuevas generaciones y servirá de retroalimentación en la triada enseñanza-aprendizaje-evaluación.

La percepción que tienen los estudiantes de la prueba objetiva fue que están de acuerdo en que el instrumento refleja los saberes que se deben adquirir en la asignatura; y que el estrés afecta de manera importante el aprovechamiento de los estudiantes.

### **Agradecimientos**

Al programa PAPIME clave PE-205215, por el financiamiento otorgado.

A los alumnos de segundo semestre de Biología de la FES Iztacala.

A los profesores Ricardo Mejía Zepeda, Ricardo Ortega Ortega, Norma Laura García Saldívar y Guillermo Elías Fernández, del Área de Laboratorio de Investigación Científica, por permitirnos aplicar la encuesta en sus grupos.

### **Referencias**

Abad F.J., Olea J., Ponsoda V., García C. (2011). Medición en ciencias sociales y de la salud. 1ª Ed. Síntesis. 560 pp.

Abreu J.L. (2012). Constructos, variables, dimensiones, indicadores y congruencia. *Daena: International Journal of Good Conscience*. 7: 123-130.

Acevedo Á.R. (2004). La percepción de los estudiantes universitarios en la medida de la competencia docente: validación de una escala. *Rev. Educ.*, 28: 145-166.

Allen J. (2006). Emotional intelligence in classrooms and in schools: What we see in the educational setting. En K. R. Murphy (Ed.), *A critique of emotional intelligence: what are the problems and how can they be fixed?* 1ª Ed. Lawrence Erlbaum Associates. p. 125-139.

Andrade C.A. (2011). Las emociones ante los exámenes y su relación con el aprovechamiento escolar en alumnos de primero y segundo grado de educación. Tesis de Maestría. Tecnológico de Monterrey.

Arechabala M.M.C, Miranda C.C. (2002). Validación de una escala de apoyo social percibido en un grupo de adultos mayores adscritos a un programa de hipertensión de la región metropolitana. *Cien. Enfer.*, 8: 49-55.

Bojórquez M.J.A., López A.L., Hernández F.M.E., Jiménez L.E. (2013). Utilización del alfa de Cronbach para validar la confiabilidad de un instrumento de medición de satisfacción del estudiante en el uso del software Minitab. 11th Latin American and Caribbean Conference for Engineering and Technology. 14-16 de agosto, Cancún, Mex.

Briones G. (1996). Metodología de la investigación cuantitativa en las ciencias sociales. 1ª Ed. ARFO Editores e Impresores Ltda. p. 29.

Brown S. (2003). Estrategias institucionales en evaluación. En S. Brown y Á. Glasner (Eds.), *Evaluar en la universidad. Problemas y nuevos enfoques*. 1ª Ed. Narcea. p. 23-33.

Carvajal A., Centeno C., Watson R., Martínez M., Sanz R. A. (2011). ¿Cómo validar un instrumento de medida de la salud? *An. Sist. Sanit. Navar.*, 34: 63-72.

Daniel W.W. (2002). Bioestadística. Base para el análisis de las ciencias de la salud. 4ª. Ed. Limusa. 923 pp.

Escofet A., Folgueiras P., Luna E., Palou B. (2016). Elaboración y validación de un cuestionario para la valoración de proyectos de aprendizaje-servicio. *Rev. Mex. Invest. Educ.*, 21:929-949.

Ferrando P.J., Anguiano-Carrasco C. (2010). El análisis factorial como técnica de investigación en psicología. *Pap. Psic.*, 31: 18-33.

Fisher S., Hood B. (1987). The stress of the transition to university: a longitudinal study of vulnerability to psychological disturbance and home-sickness. *British J Psychol.*, 79: 1-13.

Gras J.A. (1980). Psicología experimental. Un enfoque metodológico. 2ª Ed. Trillas. 460 pp.

Hernández-Sampieri R., Fernández-Collado R., Baptista-Lucio P. (2014). Metodología de la investigación. 6ª Ed. McGrawHill. 634pp.

Hernández H. C. A., Rodríguez P. N., Vargas G. Á. (2012). Los hábitos de estudio y motivación para el aprendizaje de los alumnos en tres carreras de ingeniería. *Rev. Educ. Sup.*, XLI: 67-87.

Kane, M.T. (2009). Validating the interpretations and uses of test scores. En R. W. Lissitz (Ed.), *The concept of validity. Revisions, new directions and applications*. 1ª ed. Information Age Publishing. p. 39-64.

- León G.Á., Llinás S.H., Tilano J. (2008). Análisis multivariado aplicando componentes principales al caso de los desplazados. *Rev. Cient. Ing. Des.*, Disponible en: <http://rcientificas.uninorte.edu.co/index.php/ingenieria/article/viewArticle/2098/4467>.
- Lozares C.C., López R.P. (1991). El análisis de componentes principales: aplicación al análisis de datos secundarios. *Papers*, 37: 31-63.
- Martín A.M.C. (2004). Diseño y validación de cuestionarios. *Matr. Prof.*, 5: 23-29.
- Murillo T.F.J. (2006). Cuestionario y escala de actitudes. Facultad de formación de profesorado y educación. Universidad Autónoma de Madrid. Disponible en: [https://www.uam.es/personal\\_pdi/stmaria/jmurillo/Met\\_Inves\\_Avan/Materiales/Apuntes%20Instrumentos.pdf](https://www.uam.es/personal_pdi/stmaria/jmurillo/Met_Inves_Avan/Materiales/Apuntes%20Instrumentos.pdf).
- Omar A. G. (2004). La evaluación del rendimiento académico según los criterios de los profesores y la autopercepción de los alumnos. *Revista Latinoamericana de Estudios Educativos*, XXXIV: 9-27.
- Pedrosa I., Suárez-Álvarez J., García-Cueto, E. (2014). Evidencias sobre la validez de contenido: Avances teóricos y métodos para su estimación. *Acción Psicológica*, 10: 3-20.
- Pereira-Chávez J. (2015). Evaluación, medición o verificación de los aprendizajes en el aula: Un estudio de caso en el Colegio Humanístico Costarricense de Heredia. *Rev. Electr. Educ.*, 19: 405-428.
- Pérez-Gil J.A., Chacón M.S., Moreno R.R. (2000). Validez de constructo: el uso de análisis factorial exploratorio-confirmatorio para obtener evidencias de validez. *Psicothema*. 12: 442-446.
- Pérez J. R. (2005). Calidad de la educación, calidad en la educación. Hacia su necesaria integración. *Educación XXI*, 8: 11-33.
- Phinder-Puente M.E., Sánchez-Cardel A., Romero-Castellanos F., Vizcarra-García J., Sánchez-Valdivieso E. A. (2014). Percepción sobre factores estresantes en estudiantes de Medicina de primer semestre, sus padres y sus maestros. *Inv. Ed. Med.*, 3: 139-146.
- Porto C. M. (2009). Evaluación de estudiantes en la universidad de Santiago de Compostela: percepciones de los propios alumnos. *Rev. Docen. Univ.*, 3: 1-22.
- Quecedo R., Castaño C. (2002). Introducción a la metodología de investigación cualitativa. *Rev. Psicodid.*, 14: 5-39.
- Ricoy M.C. y Fernández-Rodríguez, J. (2013). La percepción que tienen los estudiantes universitarios sobre la evaluación: Un estudio de caso. *Educación XX1*, 16 (2), 321-342. doi: 10.5944/educxx1.16.2.2645
- Ruiz B.R. (2005). Capítulo VIII: métodos no paramétricos para análisis químico, Disponible en: [http://www.ugr.es/~rruizb/cognosfera/sala\\_de\\_estudio/estadistica/tests%20nparametricos.PDF](http://www.ugr.es/~rruizb/cognosfera/sala_de_estudio/estadistica/tests%20nparametricos.PDF).
- Salgado L.A.C. (2007). Investigación cualitativa: diseños, evaluación del rigor metodológico y retos. *Liberabit. Rev. Psicol.*, 13: 71-78.
- Sánchez C. F. A. (2014). El docente frente al reto de motivar al alumno. *Revista Iberoamericana Producción Académica y Gestión Educativa*, Disponible en: <https://www.pag.org.mx/index.php/PAG/article/viewFile/134/182>.
- Tabera M.V., Álvarez M.J., Hernando A., Rubio M. (2015). Percepción de los estudiantes universitarios de ciencias de la salud sobre las actitudes de los docentes y su influencia en el clima de aprendizaje. *Rev. Compl. Educ.*, 26: 275-293.
- Trillo A.F., Porto C.M. (1999). La percepción de los estudiantes sobre su evaluación en la Universidad. Un estudio en la Facultad de Ciencias de la Educación. *Innov. Educ.*, 9: 55-75.
- Vargas M.L.M. (1994). Sobre el concepto de percepción. *Alteridades*, 4: 47-53.
- Vásquez R.F.E. y Gabalán C.J. (2006). Percepciones estudiantiles y su influencia en la evaluación del profesorado. Un caso en la Universidad Autónoma de Occidente, Cali-Colombia. *RELIEVE*, 12: 219-245.
- Vásquez C.M., Cavallo M.A., Aparicio S.N., Muñoz B.L., Robson C.M., Ruiz L.I., Secreto M.F., Sepiarsky P.A., Escobar M.E. (2012). Factores de impacto en el rendimiento académico universitario. Un estudio a partir de las percepciones de los estudiantes. *Decimoséptimas Jornadas "Investigaciones en la Facultad" de Ciencias Económicas y Estadística*. Universidad Nacional de Rosario.
- Welch S., Comer J. (1988). Quantitative methods for public administration: techniques and applications. 1ª Ed. Dorsey Press, 377pp.