



**Universidad Nacional Autónoma de México**

**Facultad de Estudios Superiores Iztacala**

**“Estimación de la validez de criterio del Cuestionario de Opinión de Estudiantes  
sobre la Práctica Docente (COPDE) en la carrera de Psicología”**

**ACTIVIDAD DE INVESTIGACIÓN-REPORTE**

**QUE PARA OBTENER EL TÍTULO DE  
LICENCIADO EN PSICOLOGÍA**

**P R E S E N T A**

**Oscar Sosa Pérez**

**Director: Dr. Sergio Manuel Méndez Lozano**

**Dictaminadores: Dra. Hortensia Hickman Rodríguez  
Mtra. Margarita Chávez Becerra**



**Los Reyes Iztacala, Edo de México, 2019**



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

## **Agradecimientos**

Quiero externar que este trabajo no hubiera sido posible sin el apoyo de los profesores Sergio, Hortensia y Margarita, que me han orientado, guiado y he aprendido mucho durante todo el proceso que me llevo llegar a la culminación del reporte que con mucho empeño he realizado y los conocimientos nuevos que me ha dejado.

También agradezco a mi familia por comprenderme, apoyarme y animarme durante el tiempo que dedique a la realización de este proyecto, que a lo largo de mi formación académica han estado presentes y me han aconsejado para tomar las decisiones adecuadas cuando ha sido necesario.

No me queda más que agradecer al equipo de investigación al que pertencí durante la realización del presente trabajo que también ha sido un apoyo para mí y que espero que en el camino que siga llegue a encontrarme de nuevo con ellos y ellas.

# ÍNDICE

INTRODUCCIÓN.....	1
MARCO TEÓRICO.....	6
1. Práctica docente.....	6
1.2 Evaluación de la práctica docente.....	7
1.3 Formas de evaluación.....	9
1.4 Relevancia del uso de cuestionarios para evaluar la práctica docente a nivel superior.....	10
2. Antecedentes de la evaluación de la práctica docente en la FES-Iztacala.....	12
2.1 Cuestionario de Opinión de Estudiantes sobre la Práctica Docente (COPDE).....	14
2.1.1 Pasos a seguir en la construcción de un instrumento.....	16
2.1.2 Propiedades psicométricas - confiabilidad y validez.....	18
3. Regresión lineal simple y múltiple.....	20
3.1 Métodos de selección de variables ( <i>stepwise, backward y forward</i> ).....	23
3.2 Corroboración de hipótesis para poder llevar a cabo la regresión lineal múltiple....	24
Justificación.....	27
Objetivo.....	27
MÉTODO.....	27
Participantes.....	27
Instrumento.....	29
Procedimiento.....	29
RESULTADOS.....	30
CONCLUSIONES.....	39
Referencias.....	43
ANEXOS.....	48

## INTRODUCCIÓN

El mundo está en constantes cambios, tanto económicos, como sociales culturales, entre otros. Para una adaptación a todos estos cambios son necesarias nuevas formas de enseñanza, por lo tanto, se presentan grandes desafíos para la educación, ya que se vuelve cada vez más determinante como inversión y valor importante tanto para los individuos como para las sociedades. Por tal razón, los sistemas educativos se esfuerzan en buscar la calidad mediante la adquisición y organización de conocimientos. Debido a esto, en los últimos años se han impulsado reformas educativas para desarrollar competencias y habilidades (Luengo, Luzón y Torres, 2008).

Aunque como tal, no existe una sola definición de competencia, se puede considerar como el conjunto de todos los saberes que van encaminados hacia una educación completa del ser, a partir de un aprendizaje significativo que le sirva para resolver los distintos problemas que se le presenten en su día a día (Gonczi y Athanasou, 1996; citado en Obaya y Ponce, 2010). Su naturaleza como concepto permite integrar las cuatro bases de la educación del presente siglo: aprender a conocer, aprender a hacer, aprender a vivir juntos y aprender a ser (Delors, 1999; citado en Obaya y Ponce, 2010).

Para el año 2002, se crea el Instituto Nacional para la Evaluación de la Educación (INEE), cuyo objetivo es contribuir a mejorar la educación por medio de una evaluación de la calidad del sistema educativo y los factores que la determinan. Asimismo, en el 2006, la Secretaría de Educación Pública (SEP) realiza la Evaluación Nacional del Logro Académico en Centros Escolares (ENLACE) con el fin de dar cuenta de los aprendizajes de los alumnos. Tanto las evaluaciones realizadas por el INEE como las de ENLACE coinciden en que los niveles de logro de los alumnos están por debajo de lo esperado y existen brechas entre las distintas modalidades educativas, que de acuerdo con los métodos de

enseñanza se pueden dividir en: escolarizada, no escolarizada y mixta (Ruiz, 2012).

Estos resultados son una base para la realización de una nueva reforma curricular para una mejora en el sistema educativo mexicano. Dicha reforma tiene como objetivo: “elevar la calidad de la educación para que los estudiantes mejoren su nivel de logro educativo, cuenten con medios para tener acceso a un mayor bienestar y contribuyan al desarrollo nacional” (p.53) (SEP, 2007 citado en Ruiz, 2012).

De acuerdo con la Secretaría de Educación Pública (2015), las reformas educativas pretenden dar al Sistema Educativo Nacional (SEN) elementos que mejoren la calidad de la educación pública obligatoria y fortalezcan la equidad al crear, por ejemplo, un Servicio Profesional Docente (SPD) y el establecimiento del INEE, visto como la principal autoridad en lo que a evaluación se refiere, con miras a realizar evaluaciones continuas de los profesores y verificar que estén en constante actualización.

La evaluación hay que diferenciarla de la medición. Este último término consiste en una comparación para obtener el grado o la amplitud de una característica relacionada, ya sea de un objeto o de una persona, a diferencia de la evaluación donde se emite un juicio a partir de dicha comparación. Además, la evaluación es considerada como un proceso continuo donde se reúne e interpreta la información recabada, que puede servir para emitir una valoración de las decisiones llevadas a cabo en el diseño de un sistema de aprendizaje que tiene una meta en específico. Dicha valoración sirve para encontrar una respuesta sobre la forma de mejorar, requiriendo el uso de instrumentos de medición exactos y adecuados, con el fin de obtener información sobre el progreso que se vaya obteniendo (Nader, 1995).

Cronbach (citado en Obaya y Ponce, 2010) entiende a la evaluación como la “recogida y uso de la información para tomar decisiones sobre un programa

educativo” (p. 32), y que permita retroalimentar al evaluado y no sólo como un elemento para realizar un juicio. Por otro lado, Stuffebcam y Shinkfield (citados en Obaya y Ponce, 2010) refieren que el propósito más importante de la evaluación es perfeccionar lo que se está evaluando.

Para clasificar a la evaluación, esta se puede dividir en diagnóstica cuando se quiere saber sobre las condiciones y posibilidades de aprendizajes obtenidos o de la capacidad para realizar una o varias tareas; formativa cuando se quiere saber si los objetivos de la enseñanza se están llevando a cabo o no y lo que se necesita para optimizar el desempeño; y sumativa cuando hay que aclarar la forma en que se medirá y juzgará el aprendizaje para dar constancia de éste, dar una calificación o, decidir la promoción o no del sujeto, entre otras actividades (Nader, 1995; Méndez y Tirado, 2016).

Las reformas en el plano curricular, tratan de impulsar prácticas de evaluación formativa que permitan brindarle al docente evidencias sobre el aprendizaje, y que se posibilite la identificación de logros o, dificultades y, ofrecer así, diferentes alternativas para mejorar su desempeño (Ruiz, 2012).

El profesor es una pieza clave en la implementación efectiva de las reformas educativas, pues tiene que traducir la reforma curricular con los objetivos y valores establecidos en ella; además debe ser un guía y acompañar en los procesos de aprendizaje a sus alumnos. Para el docente esto implica algunos desafíos, debido a que tiene que reorganizar su labor profesional, modificar sus formas de pensar, así como sus labores en función de los nuevos planteamientos que las reformas le demanden. Para poner en contexto al docente se crean una serie de programas nacionales y estatales de actualización del magisterio en servicio. Estos programas han recibido muchas críticas debido a que se dice que los educadores encargados de los cursos no están preparados, los programas no atienden a las verdaderas necesidades de los docentes o que los cursos que se otorgan están orientados a la teoría y no abordan la parte práctica (Ruiz, 2012).

Para poder verificar si hay un aprendizaje en el alumno, se requiere recurrir a las evaluaciones. Al respecto, Nader (1995) menciona que: “la evaluación es un conjunto de actividades programadas para recoger información sobre la que profesores y alumnos reflexionan y toman decisiones para mejorar sus estrategias de enseñanza y aprendizaje, e introducir en el proceso en curso las correcciones necesarias” (p.2). Además, las características de la evaluación en la educación deben incluir que esta sea: integral, orientadora, continua, reguladora del proceso educativo, y compartida democráticamente.

El panorama de la educación en México a nivel superior presenta distintos retos, que de acuerdo a Fernández (2017) se pueden resumir en los siguientes temas: cobertura y equidad, pertinencia, evaluación de la calidad, carrera académica, financiamiento y gobernanza. Sobre la temática relacionada con la evaluación de la calidad educativa, las iniciativas que se han propuesto son destinar recursos financieros para este fin por medio de programas, como el Programa Integral de Fortalecimiento Institucional (PIFI) integrado al programa de Fortalecimiento de la Calidad Educativa o el Programa Nacional de Posgrado de Calidad. Los indicadores usados para evaluar la educación en torno a su calidad que se ocupan pueden ser: el incremento de nivel de escolaridad del personal académico, la mejora del tiempo de dedicación y la consolidación de la infraestructura escolar. Sin embargo, estos elementos no abarcan completamente la ya mencionada calidad de la educación.

En la evaluación de la calidad educativa y la labor docente, se debe permitir tomar decisiones válidas y justificadas (Tejedor, 1997, citado en Calatayud, 2014). Al respecto de la evaluación, en la Facultad de Estudios Superiores Iztacala, la finalidad que se les ha dado a las evaluaciones, ha sido generar incentivos en los profesores con mejores puntuaciones (Tirado, Miranda y Sánchez, 2007). De acuerdo con Covarrubias (2003, citado en Covarrubias, 2017) el riesgo de otorgar una remuneración a partir de una evaluación favorable de los docentes, podría generar en estos actores y en los investigadores una respuesta en la que su trabajo vaya dirigido a realizar actividades para mejorar

sus ingresos y por ende se lleguen a desatender los objetivos del proyecto educativo institucional.

Para poder evaluar la labor o desempeño docente, es necesario contar con instrumentos confiables y válidos. Con base en esta premisa, del proyecto PAPIME PE302814 se desprende un trabajo donde se obtuvo la confiabilidad del Cuestionario de Opinión de los Estudiantes sobre la Práctica Docente (COPDE) que tiene como finalidad evaluar al constructo “práctica docente” (Aguirre, 2018).

Por tal motivo, si se requiere que los profesores atiendan los objetivos del proyecto educativo institucional en el que están inmersos, es necesario tener claro cuál es el objetivo de la evaluación, si ésta es sumativa o formativa. Esto con el fin de establecer programas o talleres donde se atiendan las áreas donde se obtuvieron resultados bajos y se corrijan las deficiencias que pueda tener el docente durante su desempeño al frente de su grupo en formación y que la evaluación no solo quede en incentivos para los que obtuvieron resultados favorables.

El instrumento del que se habló anteriormente cuenta con validez de constructo y validez de contenido, sin embargo, puede realizarse la estimación de la validez de criterio, que es precisamente el objetivo del presente reporte de investigación.

Para poder obtener la validez de criterio se siguió la metodología propuesta por Mazón, Martínez y Martínez (2009) donde se tomaron como variables criterio tres reactivos que no forman parte del COPDE. Dichos reactivos fueron: “Me interesa”, “Se me dificulta”, estos dos haciendo referencia a la materia impartida por el docente, ambos con tres opciones de respuesta (poco, regular, mucho) y “Asisto a clase” con cuatro opciones para poder responder (rara vez, algunas veces, casi siempre y siempre). Por tal motivo, se abordará el tema referido a la “práctica docente”, así como su evaluación través del uso del COPDE, refiriendo la importancia del uso de cuestionarios, de sus propiedades psicométricas y el uso del método por pasos sucesivos para la realización de una regresión lineal múltiple para la obtención de una validez de criterio.

## MARCO TEÓRICO

### 1. Práctica docente

De acuerdo con Tejedor (2012) es complejo definir la práctica docente, porque como tal no hay un consenso para delimitar el concepto de “buen profesor” o del constructo “calidad docente”, así como de “prácticas docentes”, lo que dificulta la medición del constructo.

Para Martínez (2012), la práctica docente consiste en toda aquella actividad que realizan los docentes en el aula con la finalidad de que el estudiante obtenga los aprendizajes establecidos en los programas y en los planes de estudio a lo largo del curso. A su vez, De Lella (1999, citado en García, Loredó, y Carranza, 2008) define a la práctica docente como aquellas acciones que realiza el profesor en el aula relacionadas con el propósito de enseñar.

Para García-Cabrero, Loredó, Carranza, Figueroa, Arbesu, Monroy y Reyes (2008, citados en García, et al., 2008), la práctica llevada a cabo por los profesores en un contexto institucional, además de la llevada a cabo dentro del aula, se le puede definir como práctica educativa, es decir, todas aquellas variables que se encuentran fuera del espacio áulico se encuentran incluidas en dicha práctica.

Sin embargo, Schoenfeld (1998, citado en García, et al., 2008) menciona que la distinción entre práctica docente y práctica educativa solo se puede hacer de manera conceptual, porque ambas prácticas se relacionan demasiado, una con la otra. Antes de realizar una actividad en el aula, se llevan a cabo planeaciones por parte del profesor que se ajustan de acuerdo a la interacción con el alumno durante la revisión de los temas, los debates o las discusiones que se generen durante la clase.

Por lo tanto, se podría decir que la práctica docente es toda aquella actividad llevada a cabo por el profesor en el aula que tenga que ver con el proceso enseñanza-aprendizaje con los alumnos, sin dejar de lado que, para poder realizar una actividad durante la clase, se tuvo que planear con anticipación.

Una vez que se tiene mayor claridad de en qué consiste la práctica docente es que se puede proceder a la evaluación de ésta. A continuación, se profundiza en esa temática.

## **1.2 Evaluación de la práctica docente**

Uno de los agentes principales en el proceso de enseñanza-aprendizaje es el docente, que debe estar constantemente actualizándose. Una forma de conocer su desempeño es por medio de la evaluación para poder detectar sus fortalezas y debilidades (Arce, 2010).

La evaluación del desempeño docente, de acuerdo con Tejedor (2012), se entiende como el conjunto de juicios emitidos sobre la labor del docente en aspectos como la enseñanza, el aprendizaje y el desarrollo académico que van dirigidos hacia los estudiantes, y que deben ser recolectados de manera válida, confiable y objetiva. Asimismo, el autor antes mencionado, señala que la evaluación docente consiste en una serie de pasos dirigidos a la obtención de la calidad de la enseñanza con el fin de contribuir a su mejora.

En el sentido de mejorar la práctica docente, la evaluación del profesorado puede ser de utilidad para obtener información que permita enriquecer la manera en que el docente realiza su tarea al tomar decisiones, tanto a nivel académico como administrativo. Por lo que, con estas mejoras, las estrategias que se puedan establecer van encaminadas a mejorar la calidad de la educación impartida, y a desarrollar la formación de los estudiantes (Márquez y Madueño, 2016).

Principalmente, la evaluación de la práctica docente puede ser realizada por medio de los estudiantes, por pares académicos, jefe inmediato, directivos, así como por medio de realizar autoevaluaciones. Estas formas sirven para integrar las distintas fuentes de información que se pueden obtener a partir de la práctica docente (Arce, 2010). Cabe resaltar que la evaluación debe ser un proceso sistemático y continuo con el fin de poder establecer programas de apoyo a los profesores para fomentar su desarrollo profesional como docente (Arce, 2010; Reyes, Luna y Caso, 2014).

Existen diversos países como Estados Unidos (EU), Australia, Colombia y Chile que se han enfocado a establecer diferentes estrategias para evaluar la práctica docente. Por ejemplo, en EU se han diseñado dos instrumentos y un portafolio para dicha evaluación. El TVASS (Tennessee Value Added Assessment System) que evalúa las aportaciones por parte del profesorado hacia los alumnos y que reporta la relación entre el desempeño del personal docente y los avances de los alumnos con respecto a su formación académica. Otro instrumento es el TECAT (Texas Examination of Current Administrator and Teachers), que debido a los costos que generaba, los pocos datos que se obtenían y el estrés que provocaba en los profesores, se optó por dejar de utilizarlo y se cambió por el EXCET, el cual sirve para determinar los conocimientos y competencias que posee el docente en las distintas etapas de su desempeño profesional. En el caso del portafolio NBPTS (The National Board for Professional Teaching Standards), lo que se evalúa es el desempeño del profesor en su zona de trabajo. Aunque estos instrumentos se han utilizado para evaluar la práctica docente, algunos presentan una deficiente realización en los reactivos de las pruebas; y suelen carecer de confiabilidad y validez (Vaillant, 2008), por lo que es menester trabajar las propiedades psicométricas de los instrumentos.

### 1.3 Formas de evaluación

Arce (2010) refiere que la evaluación por medio de los estudiantes es de las más empleadas. Usualmente se utiliza un cuestionario para evaluar al docente, donde se debe tomar en cuenta el contexto en el que se encuentra inmerso, para que de una manera sistemática se obtenga la opinión de los alumnos respecto al desempeño docente. Cabe aclarar que este no debe ser el único instrumento para evaluar la docencia.

Asimismo, Arce destaca que, en la evaluación por pares académicos, son los expertos de una materia específica quienes dan un juicio u opinión del o de los temas que son de su dominio. Por medio de esta evaluación la información obtenida va orientada a las siguientes dimensiones: conocimiento de la asignatura, planeación didáctica, evaluación de los aprendizajes y desarrollo de material didáctico.

Asimismo, dicho autor establece que la evaluación por medio del jefe inmediato, se enfoca en que éste emite una valoración cuyo fin es mejorar la eficacia del docente, así como incrementar su calidad educativa por medio de su desarrollo como profesional. En este sentido, con la evaluación del jefe inmediato se puede dar cuenta de los siguientes aspectos: capacidades pedagógicas; productividad y calidad de la enseñanza; relaciones interpersonales; y desempeño de las tareas asignadas.

Otro tipo de evaluación que destaca el autor nombrado con anterioridad, es la evaluación por medio de los directivos, que va encaminada a conocer las fortalezas y debilidades del docente, así como aquellos factores que impiden el desarrollo de su labor como docente. Por lo que durante la evaluación por parte de los directivos se deben tomar en cuenta los siguientes indicadores: planeación didáctica; evaluación de los aprendizajes; actualización profesional y docente; además de proyectos de investigación.

Aparte de los tipos de evaluación anteriormente mencionados, Tejedor (2012) resalta a la autoevaluación que básicamente consiste en hacer una recopilación de la opinión del propio docente sobre su desempeño al frente de un grupo. Por medio de esta recopilación se puede saber sobre distintos aspectos en torno a la planificación, el desarrollo y los resultados de la práctica que lleva a cabo el profesor. Las ventajas de este tipo de evaluación son la facilidad de obtención de los datos y la rapidez con la que se obtienen e interpretan los resultados obtenidos.

Como se ha visto, existen formas de evaluación con distintas finalidades y que facilitan el conocimiento de diversos aspectos relacionados con el docente. Entre ellas la más utilizada es la que se realiza por medio de los estudiantes, pero lo recomendable tal como lo menciona Arce (2010) no es utilizar un solo tipo de evaluación, sino que se lleven a cabo distintas formas de evaluación que sirvan para comprender mejor y tener mayor información sobre el constructo “práctica docente”.

#### **1.4 Relevancia del uso de cuestionarios para evaluar la práctica docente a nivel superior**

Como se había dicho, una de las estrategias de evaluación docente, es la que se basa en la opinión de los estudiantes y es uno de los mejores procedimientos, ya que el alumno es el permanente observador de la práctica docente (Tejedor, 2003; citado en Tejedor, 2012). La evaluación hacia tales agentes educativos puede realizarse a través de cuestionarios que tengan que ver con el contexto; dicha evaluación debe basarse en el desempeño percibido hacia sus profesores, y se puede optar por utilizar preguntas cerradas, abiertas o de ambos tipos (Arce, 2010).

Es precisamente en la evaluación de la práctica docente por medio de la aplicación de cuestionarios a los alumnos, donde ellos reportan su opinión sobre el desempeño de sus profesores a lo largo del curso. Al respecto Marsh (2007, como se citó en Reyes, et al., 2014) menciona que con el uso de estos

cuestionarios se cuenta con mayor confiabilidad y validez que al utilizar otra fuente de información. Además, se ha optado por el uso de este tipo de evaluaciones ya que su uso masivo implica poco tiempo durante la aplicación, así como un bajo costo (Reyes et al., 2014).

Siguiendo con el tema de los cuestionarios, Martínez (2012) menciona que dichos instrumentos buscan información a través de las respuestas que se emiten a las preguntas que los integran. En la práctica docente, los interrogados de acuerdo al tipo de evaluación que se realice, pueden ser los maestros, los directores, los alumnos o los padres de familia. Al aplicar los cuestionarios, se espera que los interrogados conozcan la información, estén dispuestos a otorgarla, y que sean capaces de entender completamente las preguntas que se les realizan.

Un cuestionario para alumnos deberá cuidar especialmente el contenido de las preguntas y la extensión de las mismas, ya que instrumentos largos o complejos producirán información deficiente e imprecisa. Cuando son preguntas cerradas, se debe considerar la calidad y exhaustividad de las opciones de respuesta. El tipo de cuestionario conocido como escala, busca explorar posturas subjetivas como opiniones o actitudes, y no el conocimiento de hechos objetivos. Para poder conocer dichas opiniones o actitudes entendidas como constructos latentes, se deben hacer inferencias a partir de aspectos que se puedan observar, a través de la expresión verbal de éstas, o por medio de conductas que las reflejen. Cuando todas las preguntas se refieren al mismo aspecto de la realidad, la escala es unidimensional. Si todos los ítems de una escala pertenecen a la misma dimensión, o se refieren al mismo constructo latente, será más probable que el conjunto de las respuestas manifieste lo que pretende. En suma, el optar por este tipo de instrumentos para realizar una evaluación del constructo se vuelve una de las mejores opciones al momento de querer realizar una evaluación de este tipo.

## **2. Antecedentes de la evaluación de la práctica docente en la FES-Iztacala**

En 1990, se tiene el antecedente de evaluar al posgrado por parte de los alumnos en la Facultad de Estudios Superiores Iztacala (FES-I) (Tirado, Nieto, Reynoso y Carrascoza, 1990; citado en Tirado et al., 2007). A continuación, se hará un recuento más detallado de los antecedentes de la evaluación por medio de la opinión de los alumnos descrita por los autores antes mencionados.

En el periodo de 1995 a 1999, se propuso en el Programa Académico de Desarrollo Institucional (PADI) de la FES Iztacala evaluar el desempeño docente por medio de la opinión de los alumnos con la finalidad de mejorar la calidad en la educación.

Conjuntamente, en 1996 el Consejo Técnico conformado tanto por profesores como por alumnos dio su consentimiento para el programa: Cuestionario de Opinión de los Alumnos sobre el Desempeño Docente (OASDE), donde las opiniones favorables motivarían al docente evaluado, mientras que las desfavorables servirían de reflexión al profesor y corregir aquellas limitaciones o carencias detectadas. Para elaborar el primer cuestionario de opinión de los alumnos el Consejo Técnico trabajó en conjunto con la Unidad de Programación y Evaluación (UPE). Se hizo la revisión de un conjunto de instrumentos previos al elaborado por las instancias antes mencionadas y finalmente el producto obtenido constó de un formato impreso con una escala de opción múltiple que pudiera revisarse con lector óptico, para fines prácticos, económicos, de mayor facilidad y de mayor rapidez en cuanto a su lectura.

A partir de 1997, luego de una prueba piloto, el Consejo Técnico ayudado por la UPE, hizo las valoraciones y ajustes necesarios de acuerdo a las observaciones de profesores y alumnos con respecto a la aplicación del instrumento. El cuestionario de papel se pasó a una versión computarizada en línea cuidando el

anonimato de los alumnos que lo contestaban (Miranda y Sánchez, 1999; Miranda y Bustos, 2001; citados en Tirado et al., 2007).

Fue en el 2001 que el Consejo Técnico tomó la decisión de otorgar un reconocimiento a aquellos profesores que se encontraran en el 10% superior de opiniones favorables, dicha posición sería tomada en cuenta para cualquier promoción o programa del que pudiera ser partícipe. Mientras que a los profesores ubicados en el 10% inferior y que no cumplieran con las responsabilidades que se les demandan, se les incitaba a que no siguieran con estas prácticas.

Tirado et al. (2007) llevaron a cabo un trabajo en el que se aplicaron cuestionarios a 26010 alumnos de la FES-Iztacala donde se evaluaron a 1011 docentes. Dicho cuestionario constó de cuatro preguntas de tipificación y 27 preguntas de opción múltiple en una escala de tipo Likert, repartidas en cuatro componentes. En el primero se obtuvo el grado de aprecio del alumno por la asignatura, en el segundo se hizo referencia a las características propias de la docencia y en los componentes tercero y cuarto se hicieron preguntas tanto para asignaturas teóricas como aplicadas por medio de preguntas equivalentes, es decir, se preguntaba lo mismo, pero se hacía el ajuste a la pregunta para que concordara con los dos tipos de asignaturas. Este estudio permitió delimitar aquellas características propias de lo que se puede considerar como un buen docente. Se concluye que los alumnos le dan un peso importante a la motivación recibida por parte del profesor al mostrar interés por lo que está enseñando, si es claro en lo que explica y si resuelve dudas en clase. Sin embargo, en este caso específico no se menciona la confiabilidad y validez que presenta el instrumento ocupado.

Por otro lado, en el 2016 se llevó a cabo el piloteo de un instrumento que sirviera para evaluar la opinión de los alumnos sobre la práctica docente de los profesores de la carrera de psicología de la FES-I, el cual sirvió para la construcción de un instrumento que evaluara dicho constructo y estimar su

validez. Al tener construido el instrumento, se sometió a una primera evaluación por un grupo de investigadoras que llegaron a un acuerdo sobre la pertinencia de todos los ítems, para construir una versión preliminar del instrumento que fue evaluado por un grupo de expertos sobre la docencia, para darle validez de contenido al instrumento (Olvera, 2016).

## **2.1 Cuestionario de Opinión de Estudiantes sobre la Práctica Docente (COPDE)**

El Cuestionario de Opinión de los Estudiantes sobre la Práctica Docente (COPDE) cuenta con 52 reactivos y está dividido en cuatro factores: el primero denominado “Metodología y Satisfacción General” cuenta con 28 reactivos, el segundo titulado “Cumplimiento del programa y evaluación” con 13, el tercero llamado “Actitud hacia los estudiantes” con 5 y, el último, con el nombre de “Ética profesional” con 6 (Aguirre, 2018). Dicho instrumento posee una alta confiabilidad con un coeficiente de  $\alpha = 0.975$  y con una validez de constructo donde la varianza total explicada de los factores es de 59.894%, acercándose al 60% lo cual es aceptable para las ciencias sociales, tal y como lo señalan Hair, Anderson, Tatham y Black (1999).

Al respecto, estos autores mencionan que la varianza total está compuesta por tres tipos de varianza: la común, la específica o única y del error. La primera se define como la varianza que se comparte en una variable con otras dentro del análisis. La segunda es la varianza que se asocia solamente con una variable en específico; mientras que la tercera se debe a poca confiabilidad en el proceso de recolección de datos, a un error de medición o a un componente aleatorio del fenómeno en cuestión.

Molero y Ruiz (2005) realizaron una investigación similar sobre las dimensiones y variables más relevantes en la evaluación de la docencia universitaria. En dicha investigación, las principales dimensiones que encontraron fueron la interacción con el alumnado, metodología, obligaciones docentes-evaluación y, medios y recursos, que explican el 64.77% de la varianza total. Al respecto, Hair et al.

(1999) señalan que, en ciencias naturales el porcentaje aceptado para la varianza total explicada es de un 95%, mientras que, en las ciencias sociales, donde la información que se obtiene es menos precisa, el porcentaje aceptable a obtener es del 60%. En el caso del COPDE se cumple con este criterio y las dimensiones de este instrumento son muy parecidas a las que se tomaron en cuenta en la investigación antes mencionada.

Al respecto, Mazón et al. (2009) mencionan que, para la elaboración de un instrumento que evalúe la práctica docente por medio del alumnado, se deben tomar en cuenta aspectos como el aprendizaje centrado en el estudiante, el empleo de las nuevas tecnologías de la información y la comunicación, entre otros; y no sólo considerar los modelos tradicionales relacionados con la educación. Todo esto con miras a optimizar el desempeño de los profesores en las aulas y a incrementar su calidad de práctica a través de la retroalimentación puntual y específica en cuanto a algunas deficiencias detectadas en su desempeño profesional.

Parte de los objetivos que se persiguen con el COPDE es que permita identificar el puntaje de la práctica docente total, aunado con cada uno de los cuatro factores que lo componen (Metodología y Satisfacción General; Cumplimiento del Programa y Evaluación; Actitud hacia el estudiante; y Ética profesional) (Méndez et al., 2017; Hickman, Alarcón, Méndez, Cepeda y Aguirre, 2019).

El COPDE puede dar cuenta de las dimensiones de la práctica docente en relación al desempeño adecuado o inadecuado del profesor. Para la realización del instrumento en cuestión se tomaron en cuenta cada uno de los pasos mencionados por Muñiz y Fonseca (2008) que se abordarán en el siguiente apartado. Asimismo, se ha visto que el cuestionario cuenta con validez de constructo y presenta confiabilidad alta (Aguirre, 2018). Sin embargo, se puede obtener la validez de criterio aunado a la validez que actualmente presenta el instrumento.

### 2.1.1 Pasos a seguir en la construcción de un instrumento

Muñiz y Fonseca (2008) señalan que el proceso de evaluación universitaria es complejo por la cantidad de aspectos diferentes que pueden ser evaluados. Estos mismos autores mencionan que, para poder desarrollar un instrumento de medida objetivo y riguroso se deben seguir una serie de pasos:

- 1) Marco general del instrumento de medida, en este paso se incluye la justificación de la construcción del instrumento, se delimita la variable a medir, así como el contexto y las circunstancias en las que se va a aplicar el instrumento, sin dejar de lado la interpretación y las decisiones que se van a tomar a partir de las puntuaciones o resultados. Dicho paso se tomó en cuenta para aclarar que el COPDE va enfocado a alumnos que se encuentran en el nivel de educación superior y que están inmersos en la carrera de psicología.
- 2) Definición operativa de la variable en cuestión. Este paso sirve para facilitar la identificación de las conductas implicadas en la medición, además se debe especificar la definición semántica (identificación de las facetas o dominios de la variable), y la definición sintáctica (la relación entre facetas o dominios con otras variables que sean de interés). La tesis realizada por Murillo (2016), puede dar cuenta de este paso donde por medio de la técnica de redes semánticas se definió el concepto de un buen profesor.
- 3) Especificaciones del instrumento. En este tercer paso se realiza la descripción minuciosa de aquellos aspectos relacionados con la administración del instrumento, número, longitud, tipo, distribución de los ítems, instrucciones y aspectos que tengan que ver con la seguridad del mismo. Las características de este paso se pueden encontrar en la tesis realizada por Aguirre (2018) y en el método del presente reporte.
- 4) Construcción de los ítems. Los principios básicos para construir un ítem son: claridad, sencillez, representatividad, relevancia y diversidad. Este paso se llevó a cabo por Hickman et al. (2016) donde en un primer

momento se construyeron 100 ítems cuidando cada uno de los principios básicos antes mencionados.

- 5) Producción, base de datos, normas de puntuación y corrección. En este paso se obtiene una versión preliminar del instrumento, se construye la base de datos para el vaciado de las puntuaciones, se especifican las normas en caso de que haya uno o algunos ítems que deban recodificarse y en caso de que se requiera, especificar la forma en que se van a distribuir las puntuaciones.
- 6) Estudio piloto cualitativo y cuantitativo, en este paso el estudio piloto es una representación de lo que será posteriormente el estudio de campo, pero con una cantidad menor de participantes, el estudio cualitativo sirve para detectar aquellos errores en la construcción de los ítems, es decir, aquellos que son confusos, que no se entienden o que sean incongruentes, mientras que con el estudio piloto cuantitativo se procede a la obtención de las propiedades métricas del instrumento, específicamente de cada ítem, dejando aquellos que sirvan para maximizar las propiedades estadísticas del instrumento en construcción. Este paso se llevó a cabo mediante un jueceo primero con estudiantes que cumplieran con las características del estudio y con profesionales que tuvieran conocimientos sobre la práctica docente y la selección de ítems (Aguirre, 2018).
- 7) Selección de otros instrumentos de medida convergentes, con la finalidad de obtener mayores evidencias de validez y que permitan una mejor interpretación de las puntuaciones obtenidas con los participantes. Para este paso, se está trabajando con el banco de datos de la Institución, ya que en cada semestre se lleva a cabo una evaluación por medio del cuestionario institucional con el que cuenta la FES Iztacala, para posteriormente realizar su comparación.
- 8) Estudio de campo. En este paso se procede a la aplicación de la versión corregida del instrumento a la muestra que se estableció previamente. El tamaño de la muestra y la aplicación del instrumento se menciona en el

método, de igual manera en el trabajo realizado por Aguirre (2018), se detalla el estudio de campo para la aplicación del instrumento COPDE.

- 9) Estimación de las propiedades métricas. En este paso se obtienen las propiedades métricas del instrumento para constatar estadísticamente que cuenta con confiabilidad y validez. La obtención de la confiabilidad, así como la validez de contenido y de constructo ya han sido estimadas (Aguirre, 2018), quedando pendiente la obtención de la validez de criterio que es el objetivo del presente reporte de investigación.
- 10) Finalmente, la construcción de la versión definitiva, informe final o manual del instrumento de medida. En este último paso, se les envían los resultados a las partes implicadas en la investigación y se realiza una manual con las instrucciones necesarias para la aplicación del instrumento. Posterior a la de la versión preliminar, obtenida por medio de una primera depuración de los ítems, se continúan realizando ajustes en los ítems para tener una versión resumida del instrumento y llegar a la versión definitiva del mismo.

### **2.1.2 Propiedades psicométricas - confiabilidad y validez.**

La confiabilidad consiste en la exactitud y la precisión con la que un instrumento mide un objeto o la ausencia de errores de medición. También se refiere a la capacidad que tiene una medida obtenida de poder generalizarse a otras situaciones. Algunos sinónimos de este término son: estabilidad de la medida y consistencia interna. La primera se refiere a que, al medir un atributo psicológico con un determinado instrumento, éste será confiable si al evaluar a los mismos sujetos con el mismo instrumento o con uno equivalente, las medidas obtenidas en la segunda aplicación son muy similares a las obtenidas en la primera. La segunda se refiere a que los reactivos de un instrumento son consistentes entre sí en la forma en que evalúan el atributo psicológico propuesto (Aragón, 2015).

Por su parte, la validez se refiere a lo que una prueba pretende medir, su eficacia y lo que se puede inferir de los puntajes obtenidos. En función del tipo de prueba

y el uso al que esté destinada se pueden obtener tres tipos de validez: de constructo, de contenido y de criterio.

La primera consiste en que a medida que una variable es abstracta se puede hablar de un constructo; lo que se busca es validar la teoría subyacente al sistema de evaluación y a la medida misma (Aragón, 2015). Respecto a este tipo de validez, es sabido que muchas de las variables psicológicas son difíciles de observar directamente, pero pueden formar parte de un constructo hipotético que trata de explicar la conducta humana, entonces lo que se busca con este tipo de validez es constatar que el registro de las conductas para la variable en cuestión, sean indicadores válidos del constructo al que hacen referencia (Argibay, 2006).

La segunda busca demostrar que los reactivos de la prueba en cuestión, son una muestra representativa de un universo y asegurar un muestreo cuidadoso de un dominio de contenido relevante. Argibay (2006) menciona que en la realización de un instrumento, la validez de contenido consiste en evaluar que los ítems usados para construir el test, sean relevantes para los objetivos con los que se construyó el test, dicho de otra forma, si estos ítems se encuentran dentro del dominio de interés. Sin dejar de lado que los ítems formen parte de una muestra representativa del universo de conductas que se eligieron como indicadores de la variable a medir.

Por último, el tercer tipo de validez es aquél que se utiliza para pronosticar una conducta a futuro, a la que se le llama criterio; una vez que se determinó éste, se elaboran los reactivos que estarán relacionados con él. Para establecer dicha validez se correlacionarán los puntajes de la prueba con los puntajes de la variable denominada como criterio (Aragón, 2015). Argibay (2006) indica que lo que se busca con la validez de criterio es relacionar las puntuaciones del instrumento con otra variable que lleva el nombre de variable criterio. Este tipo de validez sirve cuando lo que se quiere es inferir, a partir de los puntajes obtenidos en el instrumento, la repercusión que tienen éstos en otra variable que sea de interés para el investigador. Cabe destacar que de acuerdo a la temporalidad del criterio se pueden distinguir dos tipos de validez de criterio.

Cuando las puntuaciones del instrumento son utilizadas para predecir una medida del criterio que se va a llevar a cabo en un futuro, se conoce como validez predictiva. En cambio, si se relacionan las puntuaciones del instrumento con una medida del criterio de manera simultánea, se le llama validez concurrente.

Mazón et al. (2009) proponen una forma de obtener la validez de criterio, donde se ve reflejada la satisfacción por parte del alumnado hacia el desempeño docente de sus profesores y que éste se deba a los factores implicados en el instrumento, y no a otros factores que no forman parte del mismo. Cabe mencionar que la metodología que siguieron estos autores será retomada para la realización de este trabajo, donde se pretende relacionar y explicar el interés del alumno hacia la materia, su asistencia y la dificultad que pueda percibir ante ésta, tomando en consideración los dominios inmiscuidos en el constructo de práctica docente.

Para poder relacionar las dimensiones del instrumento COPDE con las variables seleccionadas como criterio, se optó por utilizar el método *stepwise* o por pasos sucesivos, pero antes de describir este método, es necesario precisar en qué consiste la regresión simple y la regresión múltiple.

### **3. Regresión lineal simple y múltiple**

Para poder realizar un análisis de regresión lineal múltiple, se deben observar correlaciones entre la o las variables dependientes o criterio con las variables independientes o predictoras. Al análisis bivariado, es decir, solo entre dos variables se le llama correlación de Pearson que sirve para encontrar la multicolinealidad entre dos variables por eso se le denomina bivariada. Mientras que a la correlación entre muchas variables se le llama correlación multivariante (De la Garza, Morales y González, 2013).

El coeficiente de correlación de Pearson ( $r$ ) es una medida que sirve para establecer la relación lineal entre dos variables cuyos datos se encuentran en una escala de intervalo o de razón. Este coeficiente va de -1 a +1, cuando se encuentran valores cercanos a -1 o +1 se dice que hay una relación lineal fuerte

y cuando se obtienen valores cercanos a cero se habla de una relación lineal débil. Cuando se cuentan con datos ordinales se obtiene el coeficiente de correlación por rangos de Spearman (Anderson, Sweeney y Williams, 2008).

Siguiendo con los mismos autores, cuando se habla de regresión, a la variable que se va a predecir se le conoce como variable dependiente, y a la variable o las variables que se utilizan para obtener el valor de la variable dependiente se les conoce como variables independientes. A continuación, se hace una descripción de la regresión lineal simple y múltiple de acuerdo con los autores antes mencionados.

Al análisis que se realiza donde una variable dependiente “y” se asocia con una variable independiente “x” mediante una aproximación por medio de una línea recta, se le conoce como análisis de regresión lineal simple. Cuando en el análisis de regresión participan más de una variable independiente se le conoce como análisis de regresión múltiple.

La regresión lineal simple se puede representar con el modelo  $y = \beta_0 + \beta_1x + \epsilon$ . De este modelo, se puede determinar la ecuación de regresión lineal simple que es  $E(y) = \beta_0 + \beta_1x$ ; cuando es representada en una gráfica,  $\beta_0$  es la intersección entre el eje “y” y la recta de regresión,  $\beta_1$  se puede considerar como la pendiente y  $E(y)$  sirve para establecer la relación entre la media o valor esperado de “y” cuando se tiene un valor específico de “x”.

La ecuación de regresión estimada que se obtiene por el método de mínimos cuadrados es  $\hat{y} = b_0 + b_1x$ , donde se emplean datos muestrales, es decir,  $b_0$  y  $b_1$  se utilizan para obtener los parámetros desconocidos del modelo  $\beta_0$  y  $\beta_1$ . Con el método de mínimos cuadrados se puede establecer la ecuación que representa de la mejor manera, la relación lineal entre la variable dependiente e independiente. El coeficiente de determinación ( $R^2$ ) es una medida de la bondad de ajuste de la ecuación de regresión estimada, que se entiende como la proporción de la variación en la variable dependiente que es explicada por dicha ecuación. En cuanto al índice de correlación se entiende como la medida de la intensidad lineal que existe entre dos variables.

Para poder determinar si la relación entre dos variables es significativa, se utilizan las pruebas t y F, en la primera se espera que, si “x” y “y” se relacionan linealmente,  $\beta_1 \neq 0$ . En el caso de la segunda, cuando se tiene una sola variable independiente se llega al mismo resultado que con la prueba t, pero esta prueba se utiliza cuando se tiene más de una variable independiente para determinar que hay una significancia general en la relación entre la variable dependiente con las variables independientes.

De la Garza, Morales y González (2013) añaden que la prueba F se utiliza como prueba global del modelo de regresión, cuando se rechaza la  $H_0$ , es decir que  $\beta_1 = 0$ , el valor de la variable independiente no se anula, en el caso del pronóstico de la variable dependiente aumentará o decrementará de acuerdo al valor que tenga  $\beta_1$ , es decir, si es positivo o negativo, lo mismo sucederá con el valor de X, y por lo tanto X deberá formar parte de la ecuación. Si se acepta la  $H_0$ , esto quiere decir que cualquier valor que se le dé a X se convertirá en cero, por lo que la variable tiene un efecto sobre Y, entonces X no forma parte de la ecuación, se da por terminado el análisis y esto quiere decir que el promedio sería el pronóstico adecuado para la variable dependiente. Mientras que la prueba t se utiliza como prueba individual para la pendiente donde igual hay dos hipótesis, la  $H_0$  donde se espera que  $\beta_1 = 0$  y la  $H_a$  donde  $\beta_1 \neq 0$  y se sigue la misma lógica que con la prueba F para rechazar o aceptar una de las hipótesis así como su interpretación.

Un coeficiente Beta (coeficiente de regresión parcial estandarizado) sirve para obtener la ecuación de regresión luego de una estandarización de las variables originales, donde se convierten las puntuaciones directas en típicas. El hecho de que estos coeficientes se encuentran en puntuaciones típicas, permite que se puedan comparar entre sí. Por lo que se puede establecer la cantidad de cambio que hay en la variable dependiente por cada unidad en la variable dependiente, mientras que las demás variables independientes permanecen constantes (Pardo y Ruiz, 2005).

### 3.1 Métodos de selección de variables (*stepwise*, *backward* y *forward*)

De la Garza et al. (2013) mencionan que, se pueden llevar a cabo ciertos métodos para seleccionar variables que realmente sean significativas y desechar aquellas que no lo son, dichos métodos o procedimientos se conocen como *stepwise*, *backward* y *forward* o bien, traducidos al español, serían por pasos sucesivos, hacia adelante y hacia atrás, respectivamente.

De acuerdo con Anderson, Sweeney y Williams (2008), el procedimiento de eliminación hacia atrás o *backward*, se empieza con un modelo que contiene cada una de las variables independientes. Posteriormente se van eliminando de una en una las variables independientes con el mismo procedimiento que se utiliza en la regresión por pasos. La diferencia con el procedimiento antes mencionado, es que, con la eliminación hacia atrás, aquella variable que ya fue eliminada no puede ingresar de nuevo al modelo. Este procedimiento concluye cuando ya no se encuentra alguna variable independiente en el modelo que presente un valor-p mayor que el nivel de significancia  $\alpha=0.05$ . Este método sirve para dejar solamente aquellas variables que son realmente significativas y que sirven para poder explicar cómo se comporta la variable dependiente (De la Garza et al., 2013).

El procedimiento de selección hacia adelante o *forward* inicia sin incluir ninguna variable independiente en el modelo establecido, posteriormente se van integrando variables con el mismo procedimiento utilizado en la regresión por pasos, con el fin de poder establecer cuál variable independiente debe agregarse al modelo. Con este procedimiento no se puede eliminar del modelo aquella variable que ya ha sido agregada. Este procedimiento concluye cuando las variables que no se agregaron al modelo presentan un valor-p mayor al nivel de significancia  $\alpha=0.05$  (Anderson et al., 2008). Para De la Garza et al. (2013) este método tiene la finalidad de que el modelo obtenido sea significativo por eso es que los pasos antes explicados se repiten hasta que ya no haya variables que puedan entrar y que sean significativas.

El procedimiento *stepwise* o por pasos sucesivos, consiste en repetir una misma ejecución que consiste en eliminar alguna de las variables que ya desde un inicio forman parte de un modelo establecido. Se debe calcular el estadístico F y el valor-p para cada una de las variables independientes contempladas para conformar dicho modelo. Si el valor-p de una de las variables independientes es mayor que al nivel de significancia  $\alpha$  se elimina del modelo y el proceso de regresión comienza de nuevo (Anderson et al., 2008). De la Garza et al. (2013) añaden que este método consiste en comparar, eliminar e introducir variables y que el proceso continúa hasta que ya no haya variables que comparar o bien que ya no puedan entrar o salir de los modelos construidos.

### **3.2 Corroboración de hipótesis para poder llevar a cabo la regresión lineal múltiple**

De acuerdo con Vilá, Torrado y Reguant (2019) cuando se requiere crear un modelo de regresión lineal múltiple se deben cumplir ciertos supuestos o hipótesis. Dichos supuestos son: linealidad, es decir, se espera que al graficar la relación entre las variables independientes y la variable dependiente sea lineal; independencia, con este supuesto se pretende que los errores obtenidos de la medición de las variables que explican la variable dependiente sean independientes entre cada una de ellas; homocedasticidad, se refiere a que cuando se obtengan las varianzas, estas sean constantes; normalidad, que como su nombre lo indica, se espera que las variables cumplan con la ley normal; y no colinealidad, con esta hipótesis se esperaría que las variables independientes no se encuentren correlacionadas entre sí de manera muy alta.

Según la hipótesis que se requiera corroborar, se pueden llevar a cabo pruebas, estadísticas, coeficientes, índices o bien se pueden obtener gráficas o tablas según sea el caso (De la Garza et al., 2013).

Según Lizasoain y Joaristi (2003) refieren que para la hipótesis de linealidad se pueden obtener diagramas de dispersión con la variable dependiente y cada una de las variables independientes, o bien, se puede obtener la gráfica con los valores predichos tipificados ZPRED y residuos tipificados ZRESID. Pardo y Ruiz (2005) mencionan que un residuo no tipificado es la diferencia entre la frecuencia observada y la esperada, un residuo tipificado es un residuo no tipificado dividido por la raíz cuadrada de su correspondiente frecuencia esperada; su valor esperado vale 0, pero su desviación típica es menor que 1. En suma, para poder decir que se cumple esta hipótesis, en la gráfica no se debe observar ningún tipo de tendencia (Lizasoain y Joaristi, 2003).

Una forma de comprobar que se cumple la hipótesis de independencia de los errores, es utilizando la prueba  $d$  de Durbin-Watson. Con este estadístico se esperaría obtener valores cercanos a 2 para poder decir que los residuos son independientes entre sí. Los valores pueden variar de 0 a 4, un valor cercano a 0 indica una autocorrelación positiva y un valor cercano a 4 indicar una autocorrelación negativa (Anderson et al., 2008).

Con la hipótesis de no colinealidad se puede ocupar el índice de tolerancia (T) que se define como el inverso del FIV (factor de inflación de la varianza). La tolerancia siempre tiene un valor entre 0 y 1. Un valor cercano a 1 indicaría que la variable es tolerada y un valor cercano a 0 significaría lo contrario; como límite se considera que sea mayor a  $1/3$  por lo que un valor menor a éste y cercano a 0 indicaría que no es bueno que dicha variable se encuentre en el modelo de regresión (De la Garza et al., 2013). Estos mismos autores, también mencionan que se puede incluir el factor de inflación de la varianza (FIV), el cual es un indicador de cuánto las varianzas de los  $\beta$  crecen y se vuelven imprecisas. El VIF oscila entre 1 e infinito. Un valor de 1 significa que la variable X es independiente, es decir, que no depende de ninguna otra que se encuentre en el análisis; un valor mayor a 1 significa que el coeficiente ( $\beta$ ) que acompaña a la VI se vuelve inestable, ya que su varianza ha sido inflada por estar relacionada con otra variable X. Como máximo se deberá tener un valor menor a 3. Un valor

del VIF mayor o igual a 10 implica problemas severos de multicolinealidad, por lo tanto, con los valores intermedios de entre 3 y 10 se tienen problemas de multicolinealidad no tan severos o débiles. Se puede decir que cuando se presenta una elevada proporción de la varianza de dos o más coeficientes se encuentra asociada a un mismo valor, esto es un indicador fuerte de que hay multicolinealidad entre las variables independientes (Lizasoain y Joaristi, 2003).

Para la hipótesis de normalidad se pueden obtener distintas pruebas, pero de todas las posibles, con verificar solo una es suficiente. La prueba de Kolmogorov-Smirnov-Lilliefors permite medir qué tanto se parecen o coincide una distribución teórica específica y una variable. Esta prueba se utiliza para muestras grandes, cuando el tamaño de la muestra es mayor o igual a 50. Cuando no se rechaza la hipótesis nula ( $H_0$ ), se concluye que los errores siguen una distribución normal. Para rechazar la  $H_0$  es necesario que p-valor sea menor a  $\alpha$ . Si no se rechaza la  $H_0$ ; se concluye que hay evidencia estadística que indica que los errores siguen una distribución normal. Las gráficas P-P plot son llamadas así por los percentiles que se necesitan encontrar. Estas gráficas sirven para verificar que los errores siguen o no una distribución normal. Si se grafican los percentiles versus los errores, a la gráfica se le denomina P-P plot; los errores siempre van en el eje Y (De la Garza et al., 2013).

Por último, la palabra homocedasticidad hace referencia a la homogeneidad de varianzas, lo que significa que existe constancia en ellas. Para poder verificar que las varianzas son constantes, se puede obtener: una gráfica de los residuales estandarizados, la prueba de Levene o la prueba M de Box transformada a una F. El procedimiento más común es el primero; en la gráfica se habla de heterocedasticidad cuando las varianzas no son constantes y se pueden observar algunos patrones en forma de abanicos o rombos; en cambio, cuando hay homocedasticidad se puede distinguir en la gráfica una distribución aleatoria (De la Garza et al., 2013).

## **Justificación**

El COPDE puede dar cuenta de las dimensiones de la práctica docente en relación al desempeño adecuado o inadecuado del profesor. Para la realización del instrumento en cuestión se tomaron en cuenta cada uno de los pasos anteriormente mencionados por Muñiz y Fonseca (2008). Asimismo, se ha visto que cuenta con validez de constructo y con alta confiabilidad en estudiantes de la carrera de Psicología (Aguirre, 2018). Aunado a esto, se puede obtener la validez de criterio a parte de la validez de constructo que actualmente presenta el instrumento.

## **Objetivo**

Este trabajo se desprende del proyecto PAPIME PE302814, que tiene como objetivo general conocer la opinión de los estudiantes de psicología de la Facultad de Estudios Superiores Iztacala sobre las prácticas docentes de los profesores.

El presente trabajo tiene como objetivo general obtener la validez de criterio del instrumento COPDE, y como objetivo específico establecer la relación entre los factores que conforman el instrumento antes descrito y el interés del alumno por la materia, el hecho de asistir a clase y el grado de dificultad hacia la materia que presentaron los participantes de este estudio.

# **MÉTODO**

## **Participantes**

La muestra que evaluó la práctica docente a través de la aplicación del instrumento COPDE perteneció al ciclo escolar 2016-1 de la carrera de Psicología de la Facultad de Estudios Superiores Iztacala. Esta generación estuvo constituida por 2795 alumnos inscritos en la carrera de Psicología (Secretaría de Planeación y Cuerpos Colegiados, 2017).

La muestra se seleccionó por medio de un procedimiento no probabilístico por conveniencia, esto es, el investigador seleccionó la muestra bajo su propio juicio (Hernández, Fernández y Baptista, 2010). Sin embargo, se tomaron en cuenta los siguientes criterios de inclusión para la selección de la muestra: 1) formar parte de la matrícula de la universidad pública seleccionada; 2) cursar los semestres 1°, 3°, 5° y 7°, siendo el número de cuestionarios contestados por semestre de 755, 816, 832 y 627 respectivamente. Cabe señalar que en algunos casos, un alumno pudo contestar hasta dos instrumentos por lo que en total se obtuvieron 3030; 3) estar inscritos en las materias correspondientes a los semestres antes mencionados (ver Tabla 1), extraída de la tesis de Aguirre (2018).

Una vez cumplidos los criterios antes descritos, es que se pudieron aplicar 3030 instrumentos, es decir se pudo evaluar a 222 docentes que en porcentaje equivale al 75.51% del total de 294 profesores que en ese ciclo escolar laboró.

**Tabla 1**  
*Asignaturas consideradas por cada semestre.*

Semestre	Asignatura
Primero	Psicología Experimental Teórica Psicología Experimental Laboratorio Métodos Cuantitativos Psicología Aplicada Laboratorio
Tercero	Psicología Experimental Teórica Psicología Experimental Laboratorio Métodos Cuantitativos Psicología Aplicada Laboratorio Teoría de las Ciencias Sociales Metodología de la Investigación Teórica y Aplicada
Quinto	Psicología Experimental Laboratorio Métodos Cuantitativos Psicología Aplicada Laboratorio Psicología Clínica Teórica Psicología Social Teórica Educación Especial y Rehabilitación Teórica Desarrollo y Educación Teórica

Séptimo

Psicología Experimental Laboratorio  
Psicología Aplicada Laboratorio  
Psicología Clínica Teórica  
Psicología Social Teórica  
Educación Especial y Rehabilitación Teórica  
Desarrollo y Educación Teórica

---

Nota: Recuperado de “Confiabilidad y validez del Cuestionario de Opinión de Estudiantes sobre la Práctica Docente (COPDE) en la carrera de Psicología”, Aguirre, L., 2018, p. 30.

### **Instrumento**

El COPDE evalúa la opinión de los estudiantes sobre la práctica docente; luego de la revisión de la literatura realizada previamente, el instrumento se construyó siguiendo lo expuesto por Muñiz y Fonseca (2008). Los pasos a seguir ya fueron descritos con anterioridad en el apartado 3.1 del presente reporte de investigación. Por lo que haciendo una recopilación de los distintos pasos que se han seguido, se puede decir, que luego de un estudio piloto, y la realización de un análisis factorial exploratorio (Aguirre, 2018); el instrumento quedó constituido por 52 reactivos distribuidos en un total de 4 dimensiones.

### **Procedimiento**

Se llevó a cabo una metodología con base en la propuesta por Mazón et al. (2009), donde se obtuvo la validez de criterio de tipo concurrente. Para ello, se eligieron tres reactivos externos que respondieron los estudiantes, ya que como menciona Aragón (2015) la consigna para realizar la validez de criterio es que se elija uno o varios fuera de la estructura del instrumento mismo.

Las preguntas que se seleccionaron fueron: 1) Me interesa, que consta de tres opciones de respuesta (poco, regular, mucho); 2) Se me dificulta, que también presenta tres opciones de respuesta (poco, regular, mucho); y 3) Asisto a clase, la cual refiere cuatro opciones (Rara vez, algunas veces, casi siempre, siempre). Cabe resaltar que estas preguntas habían sido creadas con la finalidad de ser utilizadas para la obtención de validez de criterio y que dieran indicios sobre la práctica educativa con respecto a la relación entre los reactivos seleccionados y los factores que constituyen el instrumento COPDE.

Por medio de un análisis de regresión lineal múltiple utilizando el método por pasos sucesivos se analizó la relación y la explicación que las variables predictoras o independientes realizan de manera significativa sobre estos reactivos.

Previo al análisis de regresión se obtuvieron las correlaciones entre las variables dependientes con las independientes, ya que este es un requisito para proceder con un análisis de este tipo. Se corroboraron las hipótesis de linealidad, normalidad, multicolinealidad, independencia y homocedasticidad necesarias para dar paso a la realización de una regresión lineal múltiple (Mazón et al., 2009).

## **RESULTADOS**

Se utilizó el programa SPSS v. 22, para elaborar la base de datos y, posteriormente, obtener la validez de criterio. Dicha base de datos corresponde al vaciado de los 3030 cuestionarios aplicados al ciclo escolar 2016-1 de los semestres 1°, 3°, 5° y 7° de ambos turnos. Mediante el método de regresión múltiple por pasos sucesivos, se analizó la contribución de cada uno de los factores del COPDE para analizar los tres reactivos que se les preguntaron a los estudiantes, y así estimar la validez de criterio del COPDE.

Al llevar a cabo el análisis de correlación de Pearson se muestra que en el caso de la variable criterio "Me interesa", ésta se correlaciona con los factores 1, 2 y 3 que se refieren a la Metodología y Satisfacción General; Cumplimiento del programa y evaluación; y Actitud hacia el estudiante, sin embargo, las correlaciones son bajas, en el caso de la correlación con el factor 4 es muy baja a casi nula. Con el reactivo "Se me dificulta" las correlaciones también son muy bajas, pero son inversas, es decir, a mayor dificultad de la materia percibida por el alumno, será menor la percepción que tiene sobre cada uno de los factores. Tampoco hay correlación entre el reactivo "Asisto a clase" y cada uno de los cuatro factores, porque las correlaciones son muy bajas a casi nulas (véase

Tabla 2). Cabe aclarar que todas las correlaciones son significativas y por eso se procedió a realizar el análisis de regresión lineal múltiple por pasos sucesivos para analizar que hubiera un grado de explicación significativo de las variables explicativas sobre cada una de las variables explicadas.

**Tabla 2.**

*Correlaciones de las variables dependientes con las independientes.*

VI/VD	Me interesa	Sig.	Se me dificulta	Sig.	Asisto a clase	Sig.
Factor 1	.413	.000	-.143	.000	.121	.000
Factor 2	.328	.000	-.113	.000	.147	.000
Factor 3	.221	.000	-.171	.000	.071	.000
Factor 4	.063	.000	-.084	.000	.073	.000

Por medio del método *stepwise* (por pasos sucesivos), se utilizó la regresión lineal múltiple para establecer la asociación entre cada una de las variables dependientes sobre los tres reactivos con los otros factores que constituyen al COPDE, es decir, *Metodología y Satisfacción General* (Factor 1), *Cumplimiento del programa y Evaluación* (Factor 2), *Actitud hacia el estudiante* (Factor 3) y *Ética profesional* (Factor 4).

A partir de la selección de cada una de las tres variables, se obtuvieron distintos modelos que explican dicha asociación entre la variable dependiente y las variables independientes. El método antes mencionado se utilizó con cada uno de los tres reactivos y a continuación se describen los modelos obtenidos para cada uno.

Como se puede apreciar en la Tabla 3, todos los modelos obtenidos son significativos estadísticamente hablando ( $p=.000$ ), es decir, los efectos generados por las variables predictoras sobre la variable criterio son significativos. Para que un modelo sea considerado como significativo, el valor de la significancia no debe exceder a .05, lo que quiere decir que las variables que conforman a dicho modelo son relevantes para la explicación de la variable criterio (Brace, Kemp y Snelgar, 2009; Anderson, Sweeney y Williams, 2008).

**Tabla 3**  
 ANOVAS de los tres reactivos utilizados como variables predictoras.

Reactivo	Modelo	Suma de cuadrados	Grados de libertad	Media cuadrática	F	Sig.	
<b>Me interesa</b>	2	Regresión	161.521	2	80.760	325.533	.000
		Residuo	742.772	2994	.248		
		Total	904.293	2996			
<b>Se me dificulta</b>	1	Regresión	34.957	1	34.957	89.655	.000
		Residuo	1166.978	2993	.390		
		Total	1201.935	2994			
<b>Asisto a clase</b>	1	Regresión	16.832	1	16.832	65.783	.000
		Residuo	764.541	2988	.256		
		Total	781.373	2989			

Para el reactivo - “Me interesa”-, se obtuvieron dos modelos distintos (véase Tabla 5). En el primer modelo se incluyó el factor 1, mientras que en el segundo modelo se incluyeron el factor 1 y 3; cabe destacar que los factores 2 y 4 no se incluyeron en ninguno de los modelos obtenidos. De los dos modelos, el segundo cuenta con un valor  $R^2$  de .179, lo cual indica que casi el 18% de la varianza explicada para este reactivo se debe a los factores 1 y 3 solamente. Para el reactivo “Me interesa”, los factores 1 y 3 son los que explican significativamente parte de la varianza de la puntuación obtenida en el COPDE ( $R^2=0,179$ ;  $F_{(80,760)} = 325.533$ ,  $p=0,000$ ). La razón por la que los factores 2 y 4 no se incluyeran en ningún modelo obtenido es debido a que no cumplieron con el criterio de entrada (que la variable independiente tuviera un valor p menor o igual a .05) para poder ser considerados en alguna de las ecuaciones que explican la variable dependiente (De la Fuente, 2011).

Así, en relación con los coeficientes obtenidos para el reactivo “Me interesa”, se puede establecer la siguiente ecuación:

$$\hat{Y} = 1.830 + .011x_1 - .016 x_3$$

Dicha ecuación indica que por cada punto en la variable factor 1 ( $x_1$ ) se incrementa el interés de los alumnos .011 puntos en promedio y por cada punto en la variable factor 3 ( $x_3$ ) disminuye el interés .016 en promedio. Cabe aclarar que en el caso del factor 3 habría un decremento ya que se presenta un signo negativo (De la Fuente, 2011).

**Tabla 4**

*Modelos obtenidos por el método de pasos sucesivos de los reactivos criterio.*

Reactivo	Modelo	Predictores	R	R cuadrado	R cuadrado ajustado
Me interesa	2	C, F1, F3	.423	.179	.178
Se me dificulta	1	C, F3	.171	.029	.029
Asisto a clase	1	C, F2	.147	.022	.021

Nota: C=Constante, F1=Factor 1, F2=Factor 2, F3=Factor 3, F4=Factor 4.

En el reactivo “Se me dificulta”, solo el factor 3, explica significativamente parte de la varianza obtenida ( $R^2=0,029$ ;  $F_{(34.957)}=89.655$ ,  $p=0,000$ ). En este reactivo solamente se obtuvo un modelo donde se incluyó el factor 3, es decir, que los factores 1, 2 y 4 no fueron incluidos en el único modelo obtenido. En este modelo, el valor de  $R^2$  fue de .029, lo que indica que casi el 3% de la varianza explicada es por la asociación del factor 3 relacionado con la variable en cuestión, que se refiere al grado de dificultad de la materia percibida por el alumno (véase Tabla 4).

En este caso, tres de los cuatro factores no fueron incluidos en el único modelo obtenido, es decir, los factores 1, 2 y 4 fueron excluidos porque ninguno cumplió con el criterio de entrada por tener un valor p menor o igual a 0.05.

Con los coeficientes obtenidos se puede establecer la siguiente ecuación para el modelo que se obtuvo:

$$\hat{Y} = 2.155 - .025 x_3$$

Para este caso, se puede establecer que por cada punto en la variable factor 3 ( $x_3$ ) disminuye .025 en promedio, la dificultad que percibe el alumno sobre la materia (De la Fuente, 2011). Como se puede observar en la Tabla 3 de los análisis de varianza, el modelo fue significativo a nivel global ( $p = .000$ ).

Para el reactivo "Asisto a clase", es el factor 2 que explica significativamente parte de la varianza obtenida ( $R^2 = 0,022$ ;  $F_{(16,832)} = 67.783$ ,  $p = 0,000$ ). Al igual que con el reactivo anterior, solo se obtuvo un modelo, pero a diferencia de lo sucedido con el otro reactivo, en este caso el factor se incluyó en el modelo fue el factor 2 y no el factor 3 como había sucedido anteriormente, por ende, los factores que no forman parte del modelo son los factores 1, 3 y 4. El  $R^2$  en este caso, tuvo un valor de .022, lo que significa que el 2.2% de la varianza explicada se debe a la asociación del factor 2 con la variable criterio (véase Tabla 4).

La ecuación que se puede obtener a partir del modelo es la siguiente:

$$\hat{Y} = 3.330 + .007 x_2$$

Siguiendo la lógica de la ecuación, se puede decir que por cada punto en la variable factor 2 ( $x_2$ ) se incrementa .007 puntos en promedio en la asistencia del alumno a la materia impartida por el docente en cuestión (De la Fuente, 2011).

Con relación a los coeficientes Beta tipificados obtenidos, se puede decir que para el reactivo "Me interesa", el factor que tiene mayor peso es el factor 1 con un valor coeficiente Beta tipificado con .501, le sigue el factor 2 con un valor coeficiente Beta tipificado negativo de -.127, entonces la variable que más ayuda a explicar al reactivo antes mencionado y con mayor peso es la que tiene que ver con "Metodología y Satisfacción General" (véase Tabla 5). Con el reactivo "Se me dificulta" sólo un factor que se encuentra en relación directa con éste, el valor del coeficiente Beta tipificado es de -.171, el cual es bajo y hay una relación negativa entre el reactivo y el factor 3 sobre la actitud hacia el estudiante, esto

quiere decir que a una mayor dificultad percibida por el estudiante de la materia, es menor la percepción que tiene el estudiante sobre la actitud del docente (véase Tabla 5). Por último, con el reactivo “Asisto a clase”, el factor que ayuda a explicar a este reactivo es el factor 2 que tiene que ver con el “Cumplimiento del programa y evaluación” donde el valor del coeficiente Beta tipificado fue de .147, el cual también es un valor bajo, lo que quiere decir que al haber una mayor percepción del alumno sobre la evaluación y la planificación llevada a cabo por el docente, mayor será la probabilidad de asistencia a clase de dicho alumno (véase Tabla 5).

**Tabla 5.**

*Valores de los coeficientes B y Beta de cada uno de los reactivos.*

Reactivo	Modelo	Predictores	Valor de B	Valor de Beta	Sig.
Me interesa	2	Constante	1.830		.000
		Factor 1	.011	.501	.000
		Factor 2	-.016	-.127	.000
Se me dificulta	1	Constante	2.155		.000
		Factor 3	-.025	-.171	.000
Asisto a clase	1	Constante	3.330		.000
		Factor 2	.007	.147	.000

A pesar de que solamente en el caso del reactivo “Me interesa” hay un valor de .501 con el factor 1, mientras que en los demás reactivos los valores no son mayores a .2, todos los valores fueron significativos ( $p=.000$ ) por lo que aunque la mayoría de los factores explican poco del constructo “práctica docente”, solamente el factor 4 no se encuentra asociado con ninguno de los reactivos, lo cual indica que no está relacionado con ninguno de los reactivos y por ende no influye en los reactivos para explicar ninguna de las variables.

Siguiendo con la metodología propuesta por Mazón et al. (2009) se deben corroborar las hipótesis de linealidad, normalidad, independencia, multicolinealidad y homocedasticidad. En el caso de la hipótesis de linealidad se espera que exista una asociación de cada una de las variables independientes

con la variable dependiente de forma lineal. Esto no se observa claramente en la matriz de correlaciones obtenida para ninguno de los tres reactivos con cada uno de los factores como se muestra a continuación:

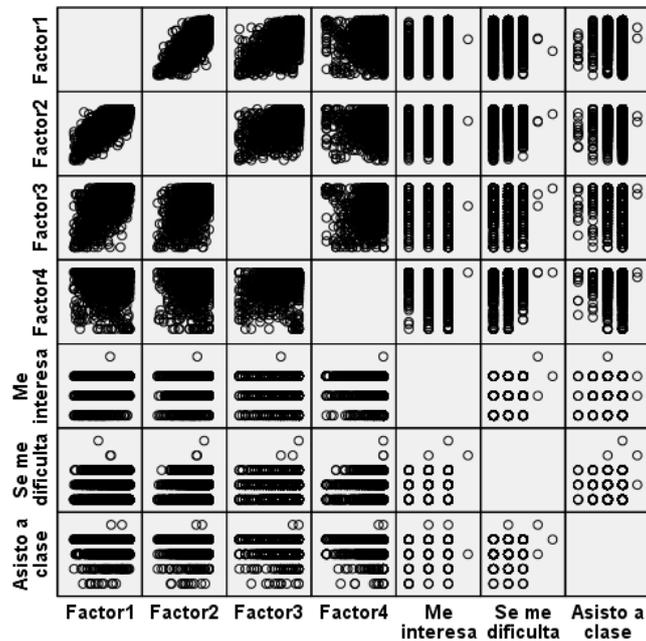


Figura 1. Matriz de correlaciones entre todas las variables tanto independientes como dependientes.

Para corroborar la hipótesis de normalidad se obtuvo la prueba Kolmogorov-Smirnov-Lilliefors para los tres reactivos, donde se espera que los valores de los residuales no sean significativos, es decir que sean mayores a 0.05, esta hipótesis no se cumple porque todos los valores fueron significativos (véase Tabla 6). También se obtuvieron los histogramas con los residuos estandarizados para los tres reactivos, donde se espera una distribución normal, pero en ninguno se pudo observar dicha distribución (véase Anexo 1). Además de la prueba antes mencionada y los histogramas, se obtuvieron las gráficas P-P para los tres reactivos, y en los tres se presentaron bolsas que se forman hacia abajo (sesgo a la derecha) y hacia arriba (sesgo hacia la izquierda) por lo que se puede concluir que los errores no siguen una distribución normal (véase Anexo 2).

**Tabla 6.**

*Prueba de normalidad Kolmorov-Smirnof de residuos estandarizados para cada una de las variables dependientes.*

Variable	Valor Estadístico	Significancia
Me interesa	.252	.000
Se me dificulta	.257	.000
Asisto a clases	.404	.000

Para corroborar la hipótesis de independencia se utilizó la prueba de Durbin-Watson, con la cual se espera obtener un valor cercano a 2, para poder decir que no hay autocorrelación en los errores de las variables independientes (De la Garza, et al. 2013). Esta hipótesis se cumple porque con cada uno de los reactivos se obtuvieron valores cercanos a 2 en el modelo número 2 en el reactivo de “Me interesa”, y en los modelos número 1 en los reactivos de “Se me dificulta” y “Asisto a clase” (véase Tabla 7).

**Tabla 7.**

*Pruebas y estadísticos descriptivos para corroborar las hipótesis de independencia y multicolinealidad.*

Reactivo	Modelo	Predictores	Valor de Durbin-Watson	Índice de colinealidad	Tolerancia	Valor de VIF
Me interesa	2	Constante	1.763	1.000	.517	1.934
		Factor 1		10.656		
		Factor 3		15.408		
Se me dificulta	1	Constante	1.811	1.000	1.000	1.000
		Factor 3		10.285		
Asisto a clase	1	Constante	1.793	1.000	1.000	1.000
		Factor 2		10.530		

Siguiendo con la hipótesis de multicolinealidad, se espera que los índices de condición sean menores a 30 (Valderrey, 2010), lo que indica que los modelos presentan ausencia de multicolinealidad. Así, como se observa en la Tabla 7 se puede indicar que los modelos cumplen con esta hipótesis, misma que se

corroborar con los índices de factor de inflación de varianzas (VIF), el cual oscila entre 1 e infinito, y que debe ser menor a 3 para que apoye la hipótesis de ausencia de multicolinealidad, al igual que el valor de Tolerancia que debe ser superior a 1/3 (De la Garza, et al. 2013).

De igual manera se obtuvieron las correlaciones entre las variables independientes donde se espera que no haya correlaciones con valores mayores a .70 (Anderson et al., 2008), dicho criterio se cumple porque, aunque una colinealidad de .834 entre el factor 1 y 2, pero para que haya multicolinealidad, debe haber una alta colinealidad en dos o más variables independientes, como lo expresa la siguiente tabla:

**Tabla 8.**

*Correlaciones entre los factores del instrumento COPDE para corroborar la hipótesis de multicolinealidad en las variables independientes.*

	Factor 1	Factor 2	Factor 3	Factor 4
Factor 1	1.000			
Factor 2	.834	1.000		
Factor 3	.696	.640	1.000	
Factor 4	.243	.267	.432	1.000

Nota: Todas las correlaciones son significativas al nivel 0.01.

Por último, en el caso de la hipótesis de homocedasticidad se espera que en la gráfica de residuales de las variables se presente una distribución aleatoria de estos o bien que no se presente algún tipo de patrón. Dicha distribución no se presenta en las gráficas obtenidas, debido a que en los tres reactivos se forman patrones de líneas separadas una de la otra, por lo que no hay homocedasticidad en la distribución de los datos, es decir, se presenta heterocedasticidad en dicha distribución (véase Anexo 3).

## CONCLUSIONES

A partir de los resultados obtenidos, se puede concluir que, con los tres reactivos tomados como criterios, no se puede presentar la validez de criterio con todos los factores como predictores del instrumento completo para su explicación. De igual forma, ya que la varianza explicada es baja en estos tres reactivos, no sería pertinente señalar que se logró estimar la validez, aunque los modelos sean significativos, además de que no se cumplen todas las hipótesis que se requieren para sustentar el uso de una regresión múltiple.

Además, como se pudo observar en dos de los modelos (con las variables explicadas “Asisto a clase” y “Se me dificulta”), que fueron producto después de realizar el método de pasos sucesivos, finalizaron en regresiones lineales simples y no en múltiples, presentándose sólo esta última con el reactivo “Me interesa”, el cual incluyó dos de los cuatro factores que conforman el instrumento COPDE.

Así, se puede decir que los dos factores que se incluyeron en el segundo modelo están asociados con el reactivo “Me interesa”. Es decir, el hecho de que un alumno se encuentre interesado por una materia, correlaciona de manera positiva a la opinión que tiene el estudiante sobre el Cumplimiento del Programa y a la Evaluación del docente, así como a la actitud que éste emite hacia sus estudiantes. Pero debido a que la varianza es apenas del 18%, se puede señalar que influye de manera muy pequeña.

Con el reactivo “Se me dificulta”, solo se obtuvo un modelo donde se incluye uno de los factores, es decir, solo el factor 3 se incluye en el único modelo obtenido. Entonces, se puede decir que cuando a un alumno se le dificulta la materia, se debe a la asociación con la actitud que el docente presente ante sus estudiantes. Sin embargo, la varianza es casi del 3%, por lo que la influencia del factor 3 sobre este reactivo es apenas perceptible.

Para el reactivo “Asisto a clase”, de igual manera a lo sucedido con el reactivo anterior, solo se pudo obtener un modelo, pero la diferencia es que el factor que

se incluyó en dicho modelo fue el factor 2. Esto quiere decir que el hecho de que un alumno asista a clase, tiene que ver con la forma en que el docente cumpla con el programa establecido y la evaluación acordada. Cabe señalar que la varianza es de apenas el 2.2%, por lo que el factor antes mencionado influye muy poco en la explicación del reactivo.

El factor 4 no se incluyó en ninguno de los modelos obtenidos para los reactivos establecidos como variables criterio, esto permite suponer que no tiene asociación con ninguna variable dependiente y por lo tanto no tiene una función de variable predictora.

El hecho de que no se cumplieran todas las hipótesis requeridas para la utilización de una regresión lineal múltiple, da pie a pensar en la utilización de otro tipo de regresión como la logística, ordinal, etc. para la obtención de la validez de criterio y reflexionar sobre la escala de medición que presenta la variable dependiente. Esto es importante revisar posteriormente para conocer cuál sería el patrón que siguen los datos para su correcta utilización. Además, también sería pertinente preguntarse qué tan útil será realizar distintos métodos de regresión con las mismas variables dependientes, al considerar que las relaciones son demasiado bajas en la explicación de los factores hacia las variables utilizadas.

Se considera también necesario que las investigaciones que utilicen el procedimiento de la regresión lineal múltiple tomen en cuenta el cumplimiento de las hipótesis que en este trabajo se analizaron, ya que, al no ser contempladas, los datos pudieran ser mal interpretados. Además, no basta con decir si cumplen o no las hipótesis como lo reportaron Mazón et al. (2009), también se deberían reportar evidencias del cumplimiento de éstas. De igual forma, no basta con señalar después de un análisis de regresión que los modelos obtenidos son significativos, sino que debe mencionarse cuánto es la varianza que explican, y en cuanto contribuye cada variable independiente sobre la dependiente para analizar el efecto que cada una tiene sobre la variable explicada.

En el caso de la hipótesis de linealidad, se puede decir que el tipo de escala influyó para que la asociación de las variables no fuera lineal, lo que hace pensar en que el tipo de análisis para obtener la validez de criterio podría ser uno de tipo ordinal por la manera en que se distribuyen los datos, es decir, de forma vertical (véase Figura 1).

Una de las razones por las que no se cumple la hipótesis de normalidad, pudiera ser porque la mayoría de los docentes obtuvieron resultados favorables luego de la aplicación del instrumento, y dado que las opciones de respuesta son en una escala tipo Likert, al momento de obtener una distribución de los datos, estos tendrían una distribución normal ya que la curva se inclinaría a la derecha o bien estaría sesgada a la derecha. Por otra parte, si quisiera tener una distribución normal, la mayoría los profesores tendrían que ser evaluados regularmente, situación que al momento de realizar la evaluación no se presentó.

Para la hipótesis de homoscedasticidad, también se puede decir que el tipo de escala influyó en la distribución de los datos y que se forman patrones de distribución de líneas horizontales.

Una dificultad que se presentó desde un inicio al seguir la metodología de Mazón, et al. (2009) fue que el criterio utilizado en su investigación era la Satisfacción General, pero en el caso del presente estudio, no se podía utilizar éste porque es parte de uno de los factores que conforman el instrumento, por lo que fue necesario utilizar otros criterios que no estuvieran incluidos en ninguno de los factores, esto debido a que el criterio utilizado para obtener la validez debe ser externo a la prueba misma (Aragón, 2015), por lo que ninguno de los cuatro reactivos sobre Satisfacción General que presenta el COPDE pudieron ser utilizados como variables criterio.

Otro punto a destacar es que al realizar la evaluación de la “práctica docente”, no se pierda de vista el objetivo principal que es mejorar la educación, ya que, al perderse dicho propósito, el docente puede incurrir en perseguir una evaluación con puntajes altos con el fin de obtener incentivos económicos tal como lo ha mencionado Covarrubias (2003).

El hecho de contar con instrumentos que muestren datos confiables y válidos, asegura que los resultados atiendan al constructo evaluado, por tal motivo se debe cuidar que durante la construcción de un instrumento se consiga que cuente con las propiedades psicométricas antes mencionadas. Esta situación se destaca ya que durante el desarrollo del capitulado del presente reporte de investigación diversos autores han hecho mención de la falta de validez y confiabilidad en los instrumentos que se han utilizado para evaluar la “práctica docente” o bien reportan han reportado la creación de instrumentos sin mencionar si cuentan con las dos propiedades psicométricas mencionadas con anteriormente (Tirado, Miranda y Sánchez, 2007; Vaillant, 2008).

Sin embargo, tal como lo sugiere Arce (2010) no basta con una sola forma de evaluación, porque se debe realizar más de una evaluación por medio de pares académicos, del jefe inmediato, de directivos o la autoevaluación (Tejedor, 2012). El hecho de contar con distintas formas de evaluación enriquece el conocimiento del constructo evaluado.

Se sugiere que se puedan utilizar otros reactivos, basados en la literatura, que pudieran ser explicados por los factores del COPDE para estimar su validez de criterio. Un ejemplo de estos reactivos, pudiera ser el promedio de los alumnos. O bien, de manera concurrente poder utilizar otra prueba que evalúe el quehacer del docente para analizar las relaciones con los reactivos del COPDE.

## Referencias

- Aguirre, L. (2018). *Confiabilidad y validez del Cuestionario de Opinión de Estudiantes sobre la Práctica Docente (COPDE) en la carrera de Psicología*. Tesis de Licenciatura. México: FES-Iztacala, UNAM.
- Anderson, D., Sweeney, D. y Williams, T. (2008). *Estadística para administración y economía*. México: Cengage Learning, 10a edición.
- Aragón, L. (2004). Fundamentos psicométricos en la evaluación psicológica. *Revista electrónica de Psicología Iztacala*, 7 (4), 23-43.
- Aragón, L. (2015). *Evaluación psicológica: historia, fundamentos teórico-conceptuales y psicometría*. México: Manual Moderno, 2da edición.
- Arce, J. (2010). Fines y modalidades de la evaluación docente en el nivel superior en México. *Revista Digital Universitaria*, 11 (6), 1-11.
- Argibay, J. (2006). Técnicas psicométricas. Cuestiones de validez y confiabilidad. *Subjetividad y Procesos Cognitivos*, 8, 15-33.
- Brace, N., Kemp, R. y Snelgar, R. (2009). *SPPS for Psychologists*. New York: Palgrave Macmillan, Four Edition.
- Calatayud, M. A. (2014). Evaluación de la práctica docente y calidad educativa: una relación encadenada. *Revista Iberoamericana de Educación*, 65, 1-12.
- Covarrubias, P. (2017). Situación Actual de los Programas de Evaluación Académica de la Educación Superior Mexicana: Sus Efectos en el Trabajo Académico. *Revista Iberoamericana de Evaluación Educativa*, 10 (2), 187-209.

De la Fuente, S. (2011). *Regresión múltiple*. Documento inédito. Madrid: Universidad Autónoma de Madrid, 1-53. Recuperado de [http://www.fuenterrebollo.com/Economicas/ECONOMETRIA/MULTIVARIANTE/REGRE\\_MULTIPLE /regresion-multiple.pdf](http://www.fuenterrebollo.com/Economicas/ECONOMETRIA/MULTIVARIANTE/REGRE_MULTIPLE /regresion-multiple.pdf)

De la Garza, J., Morales, B. y González, B. (2013). Análisis estadístico multivariante. Un enfoque teórico y práctico. México: Mc Graw Hill.

Fernández, E. (2017). Una mirada a los desafíos de la educación superior en México. *Innovación educativa*, 17 (74), 183-207.

García-Cabrero, B., Loredo, J. y Carranza, G. (2008). Análisis de la práctica educativa de los docentes: pensamiento, interacción y reflexión. *Revista Electrónica de Investigación Educativa, Especial*.

Gómez, D. (2016). Validación de un instrumento de evaluación sobre prácticas docentes. Tesis de Licenciatura. Universidad Nacional Autónoma de México.

Gregory, R. (2001). Evaluación psicológica. Historia, principios y aplicación. México: Paidós.

Hair, J., Anderson, R., Tatham, R. y Black W. (1999). Análisis Multivariante. Madrid: Prentice Hall Iberia, 5ta Edición.

Hickman, H., Alarcón, M. y Cabrera, R. (2016). Piloteo de un instrumento de evaluación de prácticas docentes de profesores de la carrera de psicología. *Debates en Evaluación y Currículum/ Congreso internacional de Educación*, 2 (2), 2249-2256.

- Hickman, H., Alarcón, M., Méndez, S., Cepeda M. y Aguirre, L. (2019). Propiedades psicométricas de un instrumento para medir prácticas docentes. *Enseñanza e Investigación en Psicología*, 1 (1), 74-86.
- Hernández, R., Fernández, C. y Baptista, P. (2010) Metodología de la investigación. México: Mc Graw Hill.
- Luengo, J., Luzón, A. y Torres, M. (2008). Las reformas educativas basadas en el enfoque por competencias: una visión comparada. *Revista de currículum y formación del profesor*. 8 (1), 27-46.
- Lizasoain, L. y Joaristi, L. (2003). Gestión y análisis de datos con SPSS. Versión 11. Madrid: Thomson.
- Luna, E. y Torquemada, A. (2008). Los cuestionarios de evaluación de la docencia para los alumnos: balance y perspectivas de su agenda. *Revista electrónica de investigación educativa*, 10, 1-15.
- Márquez, L. y Madueño, M. (2016). Propiedades psicométricas de un instrumento para apoyar el proceso de evaluación del docente universitario. *Revista Electrónica de Investigación Educativa*, 18 (2), 53-61.
- Martínez, F. (2012). Procedimientos para el estudio sobre las prácticas docentes. *Revista Electrónica de Investigación y Evaluación Educativa (RELIEVE)*, 18 (1), 1-22.
- Mazón, J., Martínez, J. y Martínez, A. (2009). La evaluación de la función docente mediante la opinión del estudiante. Un nuevo instrumento para nuevas dimensiones: COED. *Revista de la Educación Superior*, 38 (1), 113-140.
- Méndez, S., Hickman, H., Alarcón M., Cepeda M., Torres, X., Cabrera R. y Aguirre, L. (2017). ¿Qué opinan los estudiantes de psicología sobre la

actividad docente de sus profesores? *Memorias del Congreso Mexicano de Psicología*, FES-Iztacala, UNAM, 694-696.

Méndez Lozano, S. M., & Tirado Segura, F. (2016). Pensar históricamente: una estrategia de evaluación formativa. *Revista electrónica de investigación educativa*, 18(2), 62-78.

Molero, D. y Ruiz, J. (2005). La evaluación de la docencia universitaria. Dimensiones y variables más relevantes. *Revista de Investigación Educativa*, 23 (1), 57-84.

Muñiz, J. y Fonseca, E. (2008). Construcción de instrumentos de medida para la evaluación universitaria. *Revista de Investigación en Educación*, (5), 3-25.

Nader, J. (1995). Concepto y características de la evaluación. *Revista del Club Militar*, 320, 1-12.

Nunnally, J. (1987) Validez. En: *Teoría Psicométrica* (99-131). México: Trillas.

Obaya, A. y Ponce, R. (2010). Evaluación del aprendizaje basado en el desarrollo de competencias. *ContactoS*, 76, 31-37.

Olvera, T. (2016). *Juicio de expertos en un instrumento de evaluación docente*. Tesis de Licenciatura. México: FES-Iztacala, UNAM.

Pardo, A. y Ruiz, M. (2005). *Análisis de datos con SPSS 13 base*. España: Mc Graw Hill.

Reyes, E., Luna, E. y Caso, J. (2014). Propiedades psicométricas del Cuestionario de Evaluación de la Competencia Docente. *Revista Iberoamericana de Evaluación Educativa*, 7 (2), 117-130.

- Ruiz, G. (2012). La Reforma Integral de la Educación Básica en México (RIEB) en la educación primaria: desafíos para la formación docente. *Revista Electrónica Interuniversitaria de Formación del Profesorado*, 15 (1), 51-60.
- Secretaría de Planeación y Cuerpos Colegiales (2017). Estadística de inscripción por grupo. Recuperado de: <http://132.247.1.7:8180/sice/rptInscOrdinariosGrupos.jsp>
- Secretaría de Educación Pública (2015). Resumen Ejecutivo de la Reforma Educativa. *México Gobierno de la República*, 1-12.
- Tejedor, F. (2012). Evaluación del Desempeño Docente. *Revista Iberoamericana de Evaluación Educativa*, 5 (1), 318-327.
- Tirado, F., Miranda, A. y Sánchez, A. (2007). La evaluación como proceso de legitimidad: la opinión de los alumnos. Reporte de una experiencia. *Perfiles educativos*, 29 (118), 7-24.
- Vaillant, D. (2008). Algunos marcos referenciales para la evaluación del desempeño docente en América Latina. *Revista Iberoamericana de Evaluación Educativa*, 1 (2), 7-22.
- Valderrey, P. (2010). SPSS 17. Extracción del conocimiento a partir del análisis de datos. Madrid: Alfaomega Ra-Ma.
- Vilá, R., Torrado, M. y Reguant, M. (2019). Análisis de regresión lineal múltiple con SPSS: un ejemplo práctico. *REIRE, Revista d'Innovació i Recerca en Educació*, 12 (2), 1-10.

## **ANEXOS**

## Anexo 1

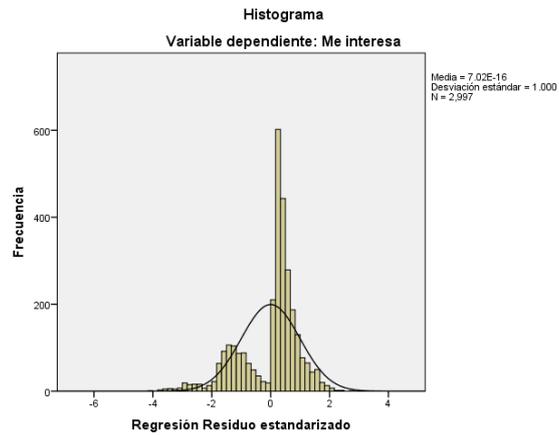


Figura 2: Histograma para corroborar la prueba de normalidad con los valores de los residuos estandarizados para el reactivo “Me interesa”.

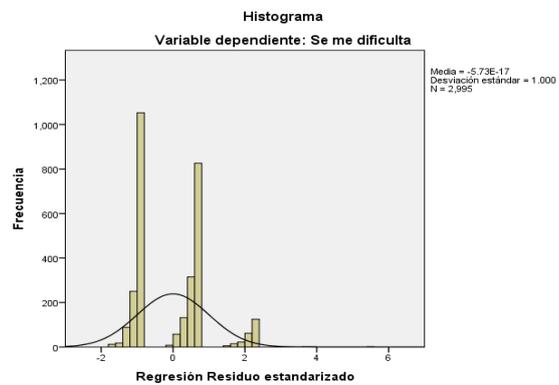


Figura 3: Histograma para corroborar la prueba de normalidad con los valores de los residuos estandarizados para el reactivo “Se me dificulta”.

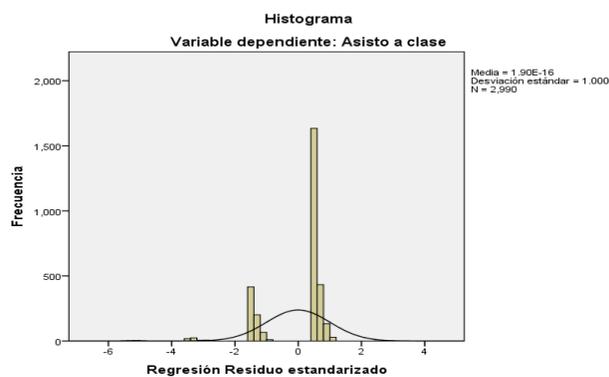
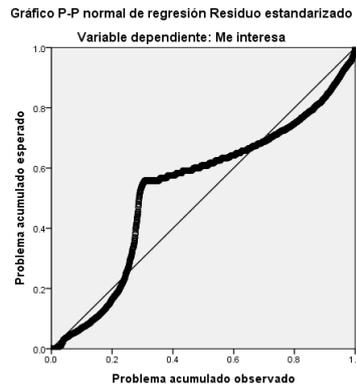
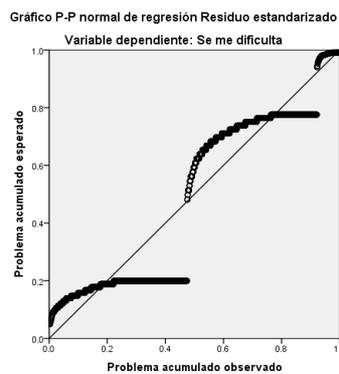


Figura 4: Histograma para corroborar la prueba de normalidad con los valores de los residuos estandarizados para el reactivo “Asisto a clase”.

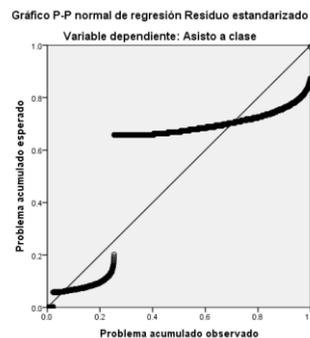
## Anexo 2



*Figura 5:* Gráfica donde se representan los percentiles de los residuos estandarizados para el reactivo “Me interesa” donde se pueden apreciar los sesgos o bolsas que se forman a la derecha y a la izquierda.



*Figura 6:* Gráfica donde se representan los percentiles de los residuos estandarizados para el reactivo “Asisto a clase” donde se aprecian los sesgos o bolsas que se forman hacia la izquierda.



*Figura 7:* Gráfica donde se representan los percentiles de los residuos estandarizados para el reactivo “Asisto a clase” donde se pueden apreciar los sesgos o bolsas que se forman hacia la derecha.

## Anexo 3

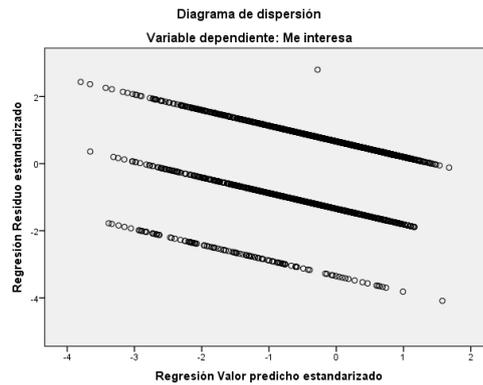


Figura 8. Gráfica utilizada para corroborar la hipótesis de homocedasticidad que representa la asociación entre los valores ZPRED Y ZRESID para el reactivo “Me interesa”.

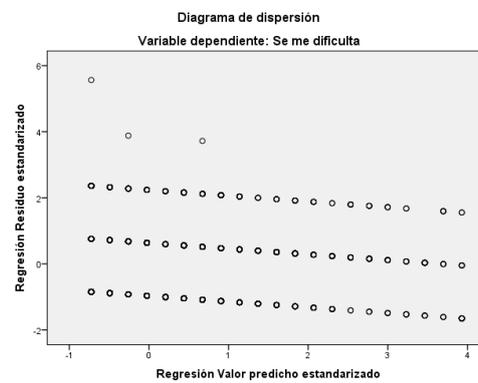


Figura 9. Gráfica utilizada para corroborar la hipótesis de homocedasticidad que representa la asociación entre los valores ZPRED Y ZRESID para el reactivo “Se me dificulta”.

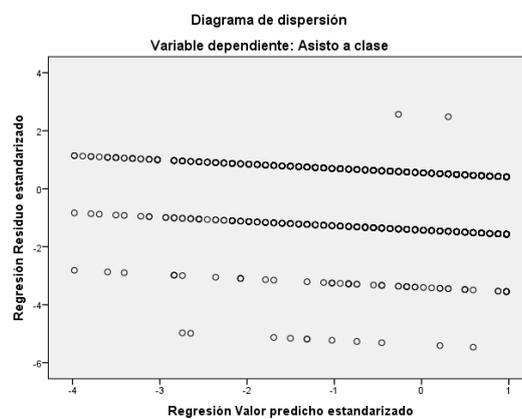


Figura 10. Gráfica utilizada para corroborar la hipótesis de homocedasticidad que representa la asociación entre los valores ZPRED Y ZRESID para el reactivo “Asisto a clase”.