



Nivel de exigencia y poder de discriminación de pruebas pedagógicas de ejecución máxima

Edgar Aníbal Cárdenas Ayala*

Resumen

En el presente trabajo de investigación de tipo descriptivo correlacional se estudiaron las variables nivel de exigencia y poder de discriminación de los ítems de las pruebas pedagógicas de ejecución máxima que se aplican a los estudiantes de las instituciones educativas de la Región Junín. Se planteó como objetivo general, analizar estadísticamente e interpretar la relación existente entre las citadas variables. Y, como hipótesis: existe una relación de causalidad entre los estadígrafos psicométricos de nivel de exigencia y de poder de discriminación, de los ítems de las pruebas pedagógicas de ejecución máxima que se aplican a los estudiantes de las instituciones educativas de la Región Junín. Arribando a la conclusión, entre otras, que existe una correlación $r = 0,957$, alta y directa con significatividad promedio de 0,01 así también, el coeficiente de determinación pone de relieve que hasta el 92 % del valor del nivel de exigencia está determinado por el valor del poder de discriminación, y el 8 % se debe a otros factores; y, que no existe diferencia entre el grado de dependencia —del nivel de exigencia con respecto al poder de discriminación—; y el grado de influencia —del nivel de exigencia en el poder de discriminación— de las pruebas pedagógicas de ejecución máxima que se aplican a los estudiantes de las instituciones educativas de la Región Junín.

Palabras clave:

Edumetría, Estadística e investigación educativa, Psicometría

Demanding's level and power of discrimination of pedagogical testing of execution maximum

Abstrac

In the present research work of a descriptive correlational type, the variables demanding's level and power of discrimination of the items of the pedagogical tests of execution maximum that are applied to the students of the educational institutions of the Junín Region were studied. The general objective was to statistically analyze and interpret the relationship between the aforementioned variables. And, as a hypothesis: there is a causal relationship between the psychometric statistics of demanding's level and the power of discrimination, of the items of the pedagogical tests of execution maximum that are applied to the students of the educational institutions of the Junín Region. Arriving at the conclusion, among others, that there is a correlation $r = 0,957$, high, direct with average significance of 0,01 as well, the determination coefficient highlights that up to 92% of the value of demanding's level is determined by the value of the power of discrimination, and 8% is due to other factors; and, that there is no difference between the degree of dependence —of demanding's level with respect to the power of discrimination—; and the degree of influence —of demanding's level in the power of discrimination— of the pedagogical tests of execution maximum that are applied to the students of the educational institutions of the Junín Region.

Keywords:

Edumetric, Statistics and educational investigation, Psychometric

Recibido: 30-11-2023

Aceptado: 20-01-2024

*Filiación: Universidad Nacional del Centro del Perú

Datos del autor:

Edgar Aníbal Cárdenas Ayala. Investigador y docente de Ciencias Naturales, Métodos Cuantitativos para la Investigación Científica e Idiomas (inglés y portugués). Doctor en Ciencias de la Educación y Magíster en Didáctica Universitaria; ambos por la Universidad Nacional del Centro de Perú. Correos: ecardenas@uncp.edu.pe y anibalito.c@gmail.com. ORCID: <https://orcid.org/0000-0002-8888-2401>

Introducción

Desde hace más de una década y de manera casi ininterrumpida se vienen realizando de manera sistemática investigaciones centradas en averiguar las relaciones que existen entre los diferentes estadígrafos psicométricos de los ítems de las pruebas pedagógicas de ejecución máxima que se aplican a los estudiantes de las instituciones educativas de la Región Junín; pero aún no se ha agotado el estudio de todas las relaciones existentes, es por eso que en la presente investigación se realiza el estudio puntual de la relación entre los estadígrafos psicométricos de nivel de exigencia y de poder de discriminación de las mencionadas pruebas.

Por otro lado, el análisis cuantitativo y cualitativo de datos provenientes de la aplicación de las pruebas pedagógicas, implica, entre otros temas, el cálculo de los estadígrafos psicométricos nivel de exigencia y de poder de discriminación; quedando por determinar la relación que existe entre estos dos; es decir, existe primacía de uno de ellos sobre el otro, uno es causa o efecto del otro, por qué hay que calcular siempre los dos o es necesario solo uno de ellos.

Entonces, en el presente trabajo se calculó la correlación r de Pearson entre el nivel de exigencia y el poder de discriminación, la que tuvo un valor $r = 0,957$ alto, directo y con significatividad promedio de 0,01; así también, el coeficiente de determinación pone de manifiesto que hasta el 92,1% del valor del nivel de exigencia está determinado por el valor del poder de discriminación, y el 7,90% se debe a otros factores; y, que no existe diferencia entre, el grado de dependencia del nivel de exigencia con respecto al poder de discriminación; y el grado de influencia del nivel de exigencia en el poder de discriminación de las pruebas pedagógicas de ejecución máxima que se aplican a los estudiantes de las instituciones educativas de la Región Junín.

Aproximaciones teóricas y conceptuales

Pruebas de ejecución máxima

Son aquellas que miden las características de la persona de acuerdo a una calificación de correcto o errado y pueden ser de rendimiento, aptitud y habilidad.

a.- De rendimiento

Miden el aprendizaje que se ha producido como resultado de las experiencias en un programa de estudios o entrenamiento. No tienen un tiempo fijo establecido, puede referirse a lo aprendido en una semana, un mes, un año, etc. Se utilizan para tomar decisiones en relación al avance de una estudiante.

b.- De aptitud

Señalan capacidades para adquirir ciertas conductas o habilidades, dadas las oportunidades adecuadas. Miden los resultados de experiencias de aprendizaje generales e incidentales y su marco de referencia se enfoca hacia el futuro. Tienen por finalidad predecir lo que puede aprender en el futuro.

c.- De habilidad

Miden las experiencias de aprendizaje más amplias y tienen su marco de referencia en el presente e indican el poder para realizar una tarea en el aquí y el ahora.

Estadígrafos psicométricos de las pruebas de ejecución máxima

Todo instrumento de recolección de datos debe reunir los siguientes requisitos esenciales: confiabilidad, validez, poder de discriminación, nivel de exigencia y sensibilidad; el presente trabajo de investigación halló la relación que existe entre los estadígrafos psicométricos de nivel de exigencia y poder de discriminación de las pruebas de ejecución máxima de rendimiento, por lo tanto, las aproximaciones teóricas están orientadas a estos dos estadígrafos psicométricos.

Nivel de exigencia, facilidad-dificultad, inteligibilidad, estadígrafos de deformación (asimetría)

El nivel de exigencia no es otra cosa que el índice de facilidad-dificultad, cuando se trata de pruebas pedagógicas, pero, cuando se trata de otro tipo de instrumento de recolección de datos es mejor llamarle grado de inteligibilidad-ininteligibilidad. Todos estos, pueden ser calculados tal como se calculan los estadígrafos de deformación.

La deformación consiste en analizar la simetría o asimetría (ausencia de simetría) de las distribuciones (Nunnally y Berstein, 1995, p. 128). La simetría puede ser positiva o negativa. La deformación estadística se deduce comparando la distribución con la forma de la curva normal que corresponde a una distribución simétrica. Si el valor calculado de la asimetría es +/- 0,37 o menor, la curva es --para efectos prácticos-- simétrica. Entre +/- 0,37 y +/-1,00 la curva tiene una asimetría aceptable. Cuando la asimetría es mayor de +/-1,00 la curva es claramente asimétrica.

Para efectos de análisis estadístico de los instrumentos de recolección de datos --pruebas pedagógicas-- a través de la deformación es pertinente introducir en concepto de facilidad o dificultad (Shiefelbein, 1999, p. 286), es decir, un ítem componente de un instrumento de recolección de datos o todo el instrumento será fácil si presenta una asimetría negativa; en cambio será todo lo contrario si

presenta una asimetría positiva.

Por ejemplo, en una prueba pedagógica, la asimetría positiva indica que el nivel de conocimiento o comprensión que está midiendo el instrumento es demasiado alto, en cambio si la asimetría es negativa indica que el nivel de conocimiento o comprensión que está midiendo el instrumento es demasiado bajo, para los niveles reales del grupo. También estará midiendo otros factores que no son parte del objetivo de la prueba pedagógica pero que influyen en los resultados.

Tabla 1.

Clasificación de la asimetría y nivel de exigencia

VALORES DE LA ASIMETRÍA	NIVEL DE EXIGENCIA
De: -2,001 a más	Muy fácil
De: -1,001 a -2,000	Fácil
De: -0,201 a -1,000	Medianamente fácil
De: -0,001 a -0,200	Tendencia a la facilidad
0,000	Dificultad intermedia
De: 0,001 a 0,200	Tendencia a la dificultad
De: 0,201 a 1,00	Medianamente difícil
De: 1,001 a 2,000	Difícil
De: 2,001 a más	Muy difícil

La dificultad también puede ser calculada por la proporción de los sujetos que respondieron correctamente el ítem, siempre que el instrumento de recolección de datos (prueba pedagógica) sea dicotómico:

$$D = \frac{\text{Nº de participantes que respondieron bien el ítem}}{\text{Nº total de participantes evaluados.}}$$

La dificultad también puede ser calculada por la proporción de los sujetos que respondieron correctamente el ítem, siempre que el instrumento de recolección de datos (prueba pedagógica) sea dicotómico:

$$P = \frac{R - \frac{W}{(O - 1)}}{N}$$

Donde:

P: Dificultad corregida

R : N° de participantes que marcaron correctamente el ítem

W: N° de participantes que marcaron incorrectamente el ítem

O: Número de alternativas de respuesta que tiene el ítem

N: N° total de participantes evaluados

Tabla 2.

Clasificación de índice de dificultad por proporción y el nivel de exigencia

VALOR DE LA PROPORCIÓN	NIVEL DE EXIGENCIA
de 0,75 a 0,99 o (de 75 a 99)	Muy fácil
de 0,55 a 0,74 o (de 55 a 74)	Fácil
de 0,45 a 0,54 o (de 45 a 54)	Intermedio
de 0,25 a 0,44 o (de 25 a 44)	Difícil
de 0,05 a 0,24 o (de 5 a 24)	Muy difícil

Poder de discriminación

Mide el grado con el cual el ítem es capaz de establecer diferencias entre las personas con niveles

altos y bajos de una habilidad, aptitud o conocimiento que está siendo evaluado. Para calcular, se separan los resultados de las pruebas considerando el grupo superior (27%) y el grupo inferior (27%), luego se obtienen separadamente para cada ítem el porcentaje de participantes que responden correctamente, ambos datos se restan y el resultado final es la discriminación que tiene el ítem. La fórmula es la siguiente:

$$\text{Dis} = \text{GS} - \text{GI}$$

Donde:

GS: % del grupo superior que contestó correctamente el ítem.
 GI: % del grupo inferior que contestó correctamente el ítem.

El valor obtenido debe ser positivo y para aceptar el ítem debe ser igual o mayor a 0,30.

El poder discriminativo de las pruebas pedagógicas representa el grado de diferenciación, separación, distanciamiento o dispersión de los resultados observados mediante las respuestas que el estudiante manifieste ante cada pregunta planteada. Desde la perspectiva psicométrica, se espera que, mayoritariamente, las respuestas acertadas, correctas o buenas provengan de aquellos estudiantes que también muestran un rendimiento o desempeño general adecuado en el total de la prueba, y que las respuestas erradas u omitidas provengan del grupo de estudiantes que presente un rendimiento general también deficiente en la prueba. Es decir, se espera que cada ítem sea capaz de permitir una diferenciación lo más nítida posible entre los postulantes de alto y bajo rendimiento, lo que, cuantitativa e idealmente, equivaldría a 100%, o en proporciones a 1,00. La magnitud de esta capacidad diferenciadora de los ítems se expresa por índices de discriminación, cuyo cálculo se realiza mediante el método correlacional, coeficiente de correlación punto biserial ($r_{p.bis}$).

$$P. D. = r_{p.bis} = (p_i - X_t) / s_t \cdot (p/q)^{1/2}$$

Donde:

P.D.: Poder de discriminación del ítem.
 P_i: Promedio de las puntuaciones totales de los estudiantes que respondieron bien el ítem analizado.
 X_t: Promedio de las puntuaciones totales de todos los estudiantes considerados en el análisis.
 S_t: Desvío de las puntuaciones totales de todos los estudiantes considerados en el análisis.
 P y Q: proporciones de los estudiantes que, respondieron bien, mal u omitieron el ítem analizado.

Tabla 3.

Clasificador del poder de discriminación del ítem

TIPO DE DISCRIMINACIÓN	VALOR
Discriminación muy buena	de 0,40 a 0,99
Discriminación aceptable	de 0,30 a 0,39
Discriminación intermedia	de 0,20 a 0,29
Discriminación inaceptable	de 0,05 a 0,19

Fuente: (Santibáñez, 2001, p. 187)

Metodología

Población

Se tomó como población de la investigación a todas las pruebas pedagógicas objetivas de ejecución máxima que aplicaron a sus alumnos, en las diferentes instituciones educativas del nivel primario, secundario y tecnológico de la Región Junín, los estudiantes de la Maestría y Doctorado en Educación, en los años académicos 2020 al 2021, en la asignatura de Análisis de datos cuantitativos y cualitativos; las mismas que se consignan en los anexos del presente trabajo de investigación.

Muestra

Para la elección de la muestra se utilizó un diseño muestral no aleatorio dirigido, de tal manera que, la muestra estuvo conformada por diez pruebas pedagógicas de ejecución máxima (todas dicotómicas),

que aplicaron --los estudiantes de la Maestría en Educación-- a sus alumnos en las instituciones educativas de la Región Junín en los dos últimos Años Académicos Escolares: 2020 al 2021.

Tipo de investigación

Dentro del campo de la investigación educativa y de acuerdo a las características de la hipótesis y los objetivos se enmarcó dentro del tipo de investigación descriptivo correlacional.

Método básico de investigación

El método utilizado en la investigación fue el descriptivo.

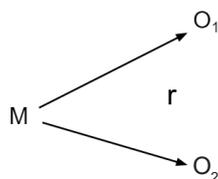
Diseño de la investigación

El diseño utilizado en esta investigación fue el descriptivo correlacional, porque resulta imposible el control experimental riguroso de las variables. Además, este diseño se adecua perfectamente a la presente investigación.

Dentro de los diseños descriptivos se encuentra el diseño descriptivo correlacional, y es el que se aplicó para determinar la relación entre los estadígrafos psicométricos, nivel de exigencia y poder de discriminación, de los ítems de las pruebas pedagógicas de ejecución máxima que se aplican a los estudiantes de las instituciones educativas de la Región Junín.

En el caso concreto de la presente investigación, se aplicaron las pruebas pedagógicas, luego con los datos recolectados, para cada ítem, se calcularon los estadígrafos psicométricos de, nivel de exigencia y poder de discriminación, los que fueron analizados. (Sánchez, 2006, p. 105).

El diseño puede diagramarse de la siguiente manera:



Donde:

M: Ítems de la prueba pedagógica que se aplican a los estudiantes de las instituciones educativas de la Región Junín.

O₁, O₂: Estadígrafos psicométricos de, nivel de exigencia y poder de discriminación, respectivamente.

R: Coeficiente de correlación

Variables de investigación

Variable 01: (V-01)

Ítems de las pruebas pedagógicas de ejecución máxima que se aplican a los estudiantes de las instituciones educativas de Región Junín.

Variable 02: (V-02)

Estadígrafos psicométricos, nivel de exigencia y poder de discriminación, de los ítems de las pruebas pedagógicas de ejecución máxima que se aplican a los estudiantes de las instituciones educativas de Región Junín.

Técnicas e instrumentos de recopilación de datos

Las diez pruebas pedagógicas que se seleccionaron como muestra fueron las que aplicaron a sus estudiantes los docentes.

Estadígrafos psicométricos, nivel de exigencia y poder de discriminación, de los ítems de las pruebas pedagógicas de ejecución máxima que se aplican a los estudiantes de las instituciones educativas de Región Junín.

Procedimientos de recopilación de datos

Los docentes --mencionados en el anterior acápite-- fueron capacitados (oportunamente) en el diseño, elaboración y aplicación de pruebas; así como, en el análisis estadístico y el cálculo de los estadígrafos psicométricos de nivel de exigencia y poder de discriminación. Los diez instrumentos de recolección de datos, fueron aplicados por los docentes a sus estudiantes en sus respectivas instituciones educativas.

Técnicas y procedimientos de procesamiento y análisis de datos

Siendo la investigación de carácter cuantitativo, para el cálculo se utilizó el paquete estadístico IBM SPSS Statistics 24, con los menús correspondientes: estadísticos descriptivos --media, suma y desvío--. El cálculo de los estadígrafos se realizó para cada uno de los ítems y para el puntaje total obtenido en cada una de las diez pruebas. Los datos, pueden observarse en los Anexos respectivos.

-El nivel exigencia --teniendo en cuenta que todas las pruebas fueron dicotómicas--; se determinó para cada ítem, mediante la división, del número de estudiantes que respondieron bien el ítem entre el número total de estudiantes que rindieron la respectiva prueba, para así hallar la proporción. Los resultados de éstas, para cada uno de los ítems de las diez pruebas, estuvieron comprendidos entre 0,05 y 0,99; dando lugar al comentario apreciativo singular en términos de facilidad y dificultad, (Shiefelbein, 1999, p. 286).

-El poder de discriminación, se hizo mediante el cálculo de medias, desvíos y la aplicación de la fórmula respectiva. Para la interpretación de los valores de los coeficientes de poder de discriminación de cada uno de los ítems de las pruebas, se tuvo en cuenta el rango de valores: superiores a 0,30 --de 0,30 a 0,39 cualitativamente indican poder de discriminación aceptable y de 0,40 a 0,99 cualitativamente indican poder de discriminación muy bueno--.

Resultados

Tabla 4.

Resumen de correlaciones de pearson entre los estadígrafos psicométricos de nivel de exigencia y poder de discriminación.

PRUEBA PEDAGÓGICA	Correlación de Pearson r	Coefficiente de determinación r ²	Variabilidad explicada r ² (%)
N°01	0,811	0,658	65,8
N°02	0,986	0,972	97,2
N°03	0,995	0,990	99,0
N°04	0,980	0,960	96,0
N°05	0,992	0,984	98,4
N°06	0,993	0,986	98,6
N°07	0,997	0,994	99,4
N°08	0,842	0,709	70,9
N°09	0,992	0,984	98,4
N°10	0,985	0,970	97,0
PROMEDIO	0,957	0,921	92,1

Para las diez pruebas pedagógicas (dicotómicas), se tiene una correlación de Pearson promedio $r = 0,957$; correlación alta y fuerte, con significatividad bilateral promedio de 0,01; por ser positiva, indica que existe una correlación directa entre los estadígrafos psicométricos de nivel de exigencia y poder de discriminación; es decir, a mayor valor del nivel de exigencia mayor valor del poder de discriminación de las pruebas. Por otro lado, el coeficiente de determinación promedio es $r^2 = 0,921$; lo que se interpreta --mediante el promedio de la variabilidad explicada-- que hasta el 92,1% del valor del nivel de exigencia de las pruebas está determinado por el valor de poder de discriminación, y el 7,90% se debe a otros factores.

Tabla 5.

Resumen de regresiones lineales entre los estadígrafos psicométricos de nivel de exigencia y poder de discriminación.

PRUEBA PEDAGÓGICA	Grado de dependencia del nivel de exigencia con respecto del poder de discriminación b	Grado de influencia del nivel de exigencia en el poder de discriminación k
N°01	0,853	0,771
N°02	0,925	1,051
N°03	1,023	0,969
N°04	0,938	1,023
N°05	0,967	1,017
N°06	0,988	0,998
N°07	0,941	1,056
N°08	1,036	0,685
N°09	0,939	1,048
N°10	0,967	1,002
PROMEDIO	0,958	0,962

Para las diez pruebas pedagógicas, en promedio, no existe diferencia entre el grado de dependencia, del nivel de exigencia con respecto al poder de discriminación; y el grado de influencia del nivel de exigencia en el poder de discriminación de las respectivas pruebas.

Docimasia de hipótesis de la correlación promedio

-**Hipótesis nula (Ho):** La correlación entre los estadígrafos psicométricos de nivel de exigencia y poder de discriminación de las pruebas pedagógicas de ejecución máxima no es significativa.

$$H_0: p = q$$

-**Hipótesis alterna (H1):** La correlación entre los estadígrafos psicométricos de nivel de exigencia y poder de discriminación de las pruebas pedagógicas de ejecución máxima es significativa.

$$H_1: p \neq q$$

-**Distribución de la muestra:** la distribución de la muestra es la distribución t de student con grados de libertad $gl = N - 2$; teniendo en cuenta que son diez pruebas pedagógicas de ejecución máxima consideradas en el espacio muestral, será:

$$Gl = 10 - 2 = 8$$

-**Nivel de confianza:** 95%; nivel de significación $\alpha = 0,05$; y un error de 5%.

-**Prueba estadística:** se eligió la prueba t de student bilateral, en función de la $r = 0,876$ de Pearson promedio, para las diez pruebas pedagógicas de ejecución máxima.

$$t = \frac{r}{\sqrt{1-r^2} \sqrt{N-2}}$$

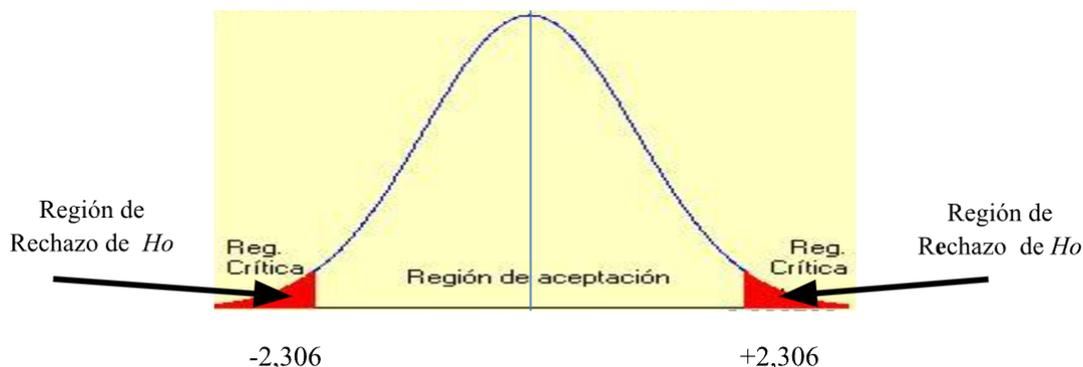
$$t = \frac{0,957}{\sqrt{1-0,916} \sqrt{8}} = \frac{0,957}{0,103} = 9,291$$

$$t_{\text{calculada}} = -6,796$$

-**Región crítica:** $t_{\text{tabulada}}: \pm 2,306$. Puesto que H1 es bilateral, la región crítica abarca todos los valores de $t \geq +2,306$ o $t \leq -2,306$

Figura 1:

Campana de Gauss



-**Decisión:** Puesto que la t calculada está ubicada en la región de rechazo del H_0 ; es decir, la $t_{calculada} > t_{tabulada}$, se acepta la H_1 .

-**Conclusión estadística:** La correlación entre los estadígrafos psicométricos de nivel de exigencia y poder de discriminación de las pruebas pedagógicas de ejecución máxima es significativa.

Discusión final de resultados

Para dar solución en parte al problema de la investigación: ¿Qué relación existe entre los estadígrafos psicométricos de nivel de exigencia y de poder de discriminación, de los ítems de las pruebas pedagógicas de ejecución máxima que se aplican a los estudiantes de las instituciones educativas de la Región Junín?

En primer lugar, cuando se trata de pruebas pedagógicas el nivel de exigencia es el índice de facilidad-dificultad (Shiefelbein, 1999, p. 286), es decir, en un ítem componente de un instrumento de recolección de datos, prueba pedagógica, la asimetría positiva indica que el nivel de conocimiento o comprensión que está midiendo el instrumento es demasiado alto, es difícil; en cambio, si la asimetría es negativa indica que el nivel de conocimiento o comprensión que está midiendo el instrumento es demasiado bajo, para los niveles reales, es fácil. Como estadígrafos de deformación consisten en analizar la simetría o asimetría (ausencia de simetría) de las distribuciones (Nunnally y Berstein, 1995, p. 128). que puede ser positiva o negativa y se deduce comparando la distribución con la forma de la curva normal que corresponde a una distribución simétrica.

En segundo lugar, hay que poner de manifiesto que el poder de discriminación está explicado mediante el estadígrafo de relación, correlación punto biserial; para cada uno de los ítems de las pruebas, en los que se tuvo en cuenta el rango de valores: superiores a 0,30 --de 0,30 a 0,39 cualitativamente indican poder de discriminación aceptable y de 0,40 a 0,99 cualitativamente indican poder de discriminación muy bueno--; tal como lo prescribe (Santibáñez, 2001, p. 187).

En la presente investigación, se han observado estas consideraciones teóricas y los valores calculados de nivel de exigencia y de poder de discriminación de los ítems de las pruebas pedagógicas de ejecución máxima conformantes del espacio muestral; están enmarcados entre los requeridos.

El valor del coeficiente de correlación de Pearson promedio $r = 0,957$; entre los estadígrafos psicométricos de nivel de exigencia y de poder de discriminación es alta y fuerte, con una significatividad bilateral de 0,01; que por ser positivo pone de manifiesto que existe una relación directa entre los estadígrafos; es decir, a mayor nivel de exigencia mayor poder de discriminación, de las pruebas y viceversa. Esta forma de interpretar, es estrictamente matemática, y no estaría revelando relaciones de causa-efecto, como demarca el espíritu de la investigación.

Que existe relación entre el nivel de exigencia y el poder de discriminación, es un hecho; por eso se hizo necesario calcular el coeficiente de determinación promedio, cuyo valor es $r^2 = 0,921$; lo que quiere decir es que, en promedio el 92,1 % del valor del nivel de exigencia de las pruebas está determinado por el valor del poder de discriminación, y el 7,9% se debe a otros factores. Por cierto, que, para emitir cualquier juicio valorativo sobre lo expuesto, habría que analizar no solo diez pruebas pedagógicas, como es el caso de la presente investigación, sino muchas (muchísimas más); no perder de vista que a la presente investigación es exploratoria y sería un punto de partida para otras investigaciones que se

podieran realizar.

Observando que, los resultados de los coeficientes de correlación de Pearson -- entre las pruebas de ejecución máxima utilizadas como espacio muestral del presente estudio-- son altos; entonces, se puede aseverar que la causalidad buscada entre el nivel de exigencia y el poder de discriminación; por lo tanto se podría generalizar la mencionada causalidad, para todas las pruebas pedagógicas de ejecución máxima que pudieran haber; es más, esta generalización no sería riesgosa, aun considerando que la muestra para la presente investigación no fue tan vasta; tampoco aleatoria.

Por otro lado, observando los resultados --para las diez pruebas pedagógicas de ejecución máxima-- de los coeficientes de regresión lineal, se tienen iguales grados de dependencia y de influencia; es decir no existe subordinación, del nivel de exigencia con respecto al poder de discriminación de las pruebas, ni viceversa.

Así también, contrastando los resultados del presente trabajo, con los obtenidos por Cárdenas en trabajos de investigación realizados los años 2008, 2010, 2014, 2015 y 2017, se observa que hay que continuar realizando una mayor cantidad de observaciones respecto a las relaciones existentes entre los diferentes estadígrafos psicométricos, porque en esos trabajos demuestra que, existe una relación de causalidad entre la facilidad-dificultad y la sensibilidad; así como, entre la validez, el poder discriminativo y nivel de dificultad; igualmente existe una correlación alta, fuerte, significativa e inversa entre la confiabilidad y la validez; así como, una correlación moderada, significativa e inversa, entre la confiabilidad y el poder de discriminación. En cambio, en los resultados de Cárdenas (2016), éste obtiene una correlación baja, inversa y con poca significatividad entre el nivel de exigencia y la confiabilidad; así también, Cárdenas (2018) concluye que la correlación entre la validez y la sensibilidad es débil e inversa en las pruebas pedagógicas de ejecución máxima que se aplican a los estudiantes de las instituciones educativas de la Región Junín. Cárdenas (2019) arriba a la conclusión que la correlación entre el nivel de dificultad (inteligibilidad) y la validez es débil y directa. Finalmente Cárdenas (2020) concluye que existe una correlación alta y directa entre el poder de discriminación y la validez en las pruebas de ejecución máxima aplicadas a los estudiantes de las instituciones educativas de la Región Junín.

Todo lo expresado líneas arriba, se escapa de los alcances de una buena práctica pedagógica, porque ésta tiene otras características, y los estadígrafos psicométricos investigados no son los únicos, sino que existen otros que también influyen en la elaboración de una prueba pedagógica; tales como los estadígrafos edumétricos. Santibáñez (2001).

Conclusiones

1.-El valor cuantitativo del nivel de exigencia de los ítems de las pruebas pedagógicas de ejecución máxima que se aplican a los estudiantes de las instituciones educativas de la Región Junín, fluctúa entre 0,28 y 0,81; lo que en términos cualitativos estarían en el rango de difícil a muy fácil.

2.-El poder de discriminación de gran parte de los ítems de las pruebas pedagógicas de ejecución máxima que se aplican a los estudiantes de las instituciones educativas de la Región Junín, en términos cualitativos, está en el rango de aceptable y muy bueno.

3.- La correlación $r = 0,957$ es alta y directa, con una significatividad bilateral promedio de 0,01; el coeficiente de determinación pone de manifiesto que hasta el 92,1% del valor del nivel de exigencia, —de las pruebas pedagógicas de ejecución máxima que se aplican a los estudiantes de las instituciones educativas de la Región Junín— está determinado por el valor del poder de discriminación, y el 7,90% se debe a otros factores.

4.- El análisis de regresión lineal, de las pruebas pedagógicas de ejecución máxima que se aplican a los estudiantes de las instituciones educativas de la Región Junín, pone de manifiesto que, no existe diferencia entre el grado de dependencia, del nivel de exigencia con respecto al poder de discriminación; y el grado de influencia del nivel de exigencia en el poder de discriminación, de las respectivas pruebas.

Referencias

- Aiken, L. (1995). Test psicológicos y evaluación. Prentice Hall.Hispanoamericana.
- Ary; Jacobs y Razaviev. (1992). Introducción a la investigación pedagógica. McGraw-Hill.
- Calvo, E. (2015). Estadística: software estadísticos versus fórmulas estadísticas. [Trabajo de investigación].

Centro de Producción de Documentos-UPSMP.

- Cárdenas Ayala, A. (2008). Análisis estadístico de los instrumentos de recolección de datos a través de los estadígrafos de deformación y apuntamiento. *Revista Horizonte de la ciencia*, 04, 79-88. <http://revistas.uncp.edu.pe/index.php/horizontedelaciencia>
- Cárdenas Ayala, A. (2009). Poder discriminativo, nivel de exigencia y capacidad predictiva de los exámenes de selección [Investigación no publicada] I.I.FPH-UNCP.
- Cárdenas Ayala, A. (2010). Relación entre la validez, el poder discriminativo y nivel de dificultad de las pruebas pedagógicas. *Revista Horizonte de la ciencia*, 01, 119-126. Huancayo, Perú. <http://revistas.uncp.edu.pe/index.php/horizontedelaciencia>
- Cárdenas Ayala, A. (2011). Índice de discriminación de los ítems politómicos en las pruebas pedagógicas. *Revista Prospectiva universitaria*, V- 8, 01, 119-126. UNCP. <http://www.uncp.edu.pe/universidad/investigación/index.php>
- Cárdenas Ayala, A. (2012). Análisis estadístico del índice de dificultad de los ítems politómicos en las pruebas pedagógicas. *Revista Horizonte de la ciencia*, 02, 74-81. <http://revistas.uncp.edu.pe/index.php/horizontedelaciencia>
- Cárdenas Ayala, A. (2013). Dificultad y sensibilidad de los instrumentos de recolección de datos. Editora Imprenta Ríos S.A.C
- Cárdenas Ayala, A. (2013). Índice de dificultad y asimetría de los ítems en las pruebas pedagógicas. [Investigación no publicada] I.I.FE-UNCP.
- Cárdenas Ayala, A. (2014). Confiabilidad y validez de las pruebas pedagógicas de ejecución máxima. [Investigación no publicada] I.I.FE-UNCP.
- Cárdenas Ayala, A. (2015). Confiabilidad y poder de discriminación de las pruebas pedagógicas de ejecución máxima. *Revista Prospectiva universitaria*, V-13, 01 y 02, 85-90. UNCP. <http://www.uncp.edu.pe/universidad/investigación/index.php>
- Cárdenas Ayala, A. (2016). Nivel de exigencia y confiabilidad de pruebas pedagógicas de ejecución máxima. [trabajo de investigación no publicado]. I.I, FE-UNCP.
- Cárdenas Ayala, A. (2017). Sensibilidad y confiabilidad de pruebas pedagógicas de ejecución máxima. *Revista Educación y ciencia*, V-3, 09, 71-78. UNCP.
- Cárdenas Ayala, A. (2018). Validez y sensibilidad de pruebas pedagógicas de ejecución máxima. [Investigación no publicada] I.E.I., FE-UNCP.
- Cárdenas Ayala, A. (2019). Nivel de dificultad (inteligibilidad) y validez de pruebas pedagógicas de ejecución máxima. *Revista Horizonte de la ciencia*, 19, 209-219. <http://revistas.uncp.edu.pe/index.php/horizontedelaciencia>
- Cárdenas Ayala, A. (2020). Poder de discriminación y validez de pruebas de ejecución máxima. I.E.I., FE-UNCP, *Revista Investigación y educación*, 1, 8-21. <https://revistas.uncp.edu.pe/index.php/invest>
- Cárdenas Ayala, A. (2021). Nivel de exigencia y confiabilidad de pruebas pedagógicas de ejecución máxima. *Revista educanatura*. <http://revistas.uncp.edu.pe/index.php/educanatura>
- Davis, F. (1989). *Analyze des items*. Nauwelaerts.
- Delgado; Ecurra y Torres. (2006). *La medición en psicología y educación*. Hozlo.
- Fermín, M. (1995). *La evaluación, los exámenes y las calificaciones*. Kapeluz.
- Gonzalvo, G. (1994). *Diccionario de metodología estadística*. Morata S.A.
- Hambleton y Novik. (1989). *Toward an integration of theory and method for criterion referenced test*. Measurmet.
- Hernández; Fernández y Baptista. (2003). *Metodología de la investigación*. McGraw-Hill.
- Lozada y Sánchez. (2010). *Estadística inductiva*. Universidad de Holguín "Oscar Lucero Moya".

- Nunnally y Bernstein. (1995). *Teoría psicométrica*. Mc Graw-Hill.
- Piris y Zacharzewski. (2017). *Sociedad estadística*. Universidad Nacional de Misiones.
- Quezada, N. (2010). *Metodología de la investigación-Estadística aplicada en la investigación*. MACRO.
- Rubio, A. (2018). *Cuidado al calcular las medidas de forma con el software Excel*. [Trabajo de investigación]. UNALM.
- Sánchez y Reyes. (2006). *Metodología y Diseños en la Investigación Científica*. Visión Universitaria.
- Santa, M. (2015). *Glosario Español-Inglés y English-Spanish de terminología estadística*. Universidad Nacional Abierta y a Distancia.
- Schiefelbein, E. (1999). *Teoría, técnicas, procesos y casos en el planeamiento de la educación*. Ateneo.
- Santibañez, J. (2001). *Manual para la evaluación del aprendizaje estudiantil (conceptos, procedimientos, análisis e interpretación para el proceso evaluativo)*. México: Trillas.
- Santisteban, C. (2009). *Principios de psicometría*. Síntesis.
- Tavella, N. (1991). *Análisis de los ítems en la construcción de instrumentos Psicométricos*. Trillas.
- Tembrick, T. (1981). *Evaluación*. Narcea.
- Thorndikee y Hagen. (1991). *Medición y evaluación en psicología y educación*. Trillas.
- Visauta y Martori i Cañas. (2003). *Análisis estadístico con SPSS para Windows, (Volumen II)*. McGraw-Hill.
- Schiefelbein, E. (1999). *Teoría, técnicas, procesos y casos en el planeamiento de la educación*. Ateneo.
- Santibañez, J. (2000). *Manual para la evaluación del aprendizaje estudiantil (conceptos, procedimientos, análisis e interpretación para el proceso evaluativo)*. Trillas.
- Santisteban, C. (2009). *Principios de psicometría*. Síntesis.
- Tavella, N. (1991). *Análisis de los ítems en la construcción de instrumentos Psicométricos*. Trillas.
- Tembrick, T. (1981). *Evaluación*. Editorial Narcea.