

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO
FACULTAD DE PSICOLOGIA



**Estandarización y Análisis Psicométrico de la Prueba de
Dominós en una Escuela de Capacitación Industrial**

TESIS

QUE PARA OBTENER EL TITULO DE

LIC. EN PSICOLOGÍA

PRESENTAN:

ESPERANZA NÚÑEZ LÓPEZ

MA. DE LOS ANGELES NÚÑEZ LÓPEZ

MEXICO, D. F.

1981



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



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.

Z5053.08
UNSM.148
1981

M-20339
tps. 833

2655

A nuestros Padres:

*Que con su sabiduría han sabido guiarnos
por el camino del estudio y la superación.*

A nuestros hermanos:

*Que con su ejemplo sembraron en nosotros
la semilla de la tenacidad y la constancia.*

A nuestras tías:

Por el entusiasmo que supieron infundirnos

Al maestro Javier Aguilar:

*Que gracias a su colaboración fue posible
la elaboración de esta tesis.*

INTRODUCCION

Es sabido que en nuestro país son muy escasas las pruebas psicológicas válidas y confiables.

Como las Instituciones al seleccionar personal emplean pruebas que han sido construidas en otros Países y sobre una población determinada (lo cual como es de esperar trae consecuencias negativas en la ubicación de personal, ya que se valoran y califican los sujetos basándose en normas obtenidas en la población estudiada) no podemos valorar en forma adecuada algún atributo de conducta de un individuo, sobre normas hechas para una población cuyo patrón sociocultural es distinto u opuesto al nuestro.

Esto conduce en la mayoría de los casos a errores en la toma de decisiones, lo cual es desventajoso para las propias Instituciones y para el campo de la Selección de Personal puesto que cada día se vuelve más necesaria la colocación de personal de acuerdo a sus aptitudes, intereses, personalidad y capacidad intelectual, debido en gran parte a que en los últimos años la Tecnología ha evolucionado de una manera sorprendente y la especialización del trabajo se hace cada día más indispensable, como consecuencia del gran incremento industrial.

Al analizar este problema decidimos estandarizar una de las pruebas de inteligencia, la prueba de "DOMINOS" en un Centro de Estudios Científicos y Tecnológicos (CECyTs).

La prueba de "Dominós" es sencilla en su aplicación puesto que está construida con material no verbal y también es una de las pruebas más utilizadas en nuestro país.

Esta prueba junto con otros tests como el Raven, Wechsler, Barsit, Otis, D. A. T., MMPI y el Kuder, se utilizan para ubicar al personal según su capacidad intelectual, aptitudes, intereses y personalidad.

Los Centros de Estudios Científicos y Tecnológicos (CECyTs) son escuelas de capacitación industrial dependientes de la Secretaría de Educación Pública cuya finalidad es la de capacitar Técnicos a nivel medio en diferentes ramas.

Dichos centros no cuentan con instrumentos estandarizados para seleccionar a los alumnos y evaluar su aprovechamiento.

Creemos que este esfuerzo puede motivar a otras personas interesadas

en ello para que en un futuro no muy lejano podamos contar con suficientes tests estandarizados para nuestra población.

Es, pues, una tarea ardua pero que reditúa sus beneficios ya que así se podrán evaluar en forma confiable las características de la población estudiada.

Esta aportación, aunque mínima, pensamos ayudará a tomar decisiones más adecuadas y confiables con respecto a la ubicación de los alumnos en las diferentes especialidades del CECyT.

CAPITULO I

CONFIABILIDAD Y VALIDEZ DE LOS TESTS PSICOLOGICOS

Introducción a los tests psicológicos

La psicometría tiene como finalidad llevar a cabo la medición de la conducta, esta es uno de los pilares de la ciencia psicológica ya que el hombre quiere comprender su propia naturaleza y la de los demás.

La medición psicológica, como en cualquier otra ciencia, se indica con la identificación de elementos del mundo real con elementos de un sistema lógico abstracto, al cual se le denomina modelo.¹

Específicamente, Medición: “es el procedimiento mediante el cual se asignan números a las propiedades, atributos o características de los objetos, estableciendo las reglas específicas sobre las que se fundamentan tales asignaciones”.²

Los instrumentos de medición psicológica son todos aquellos procedimientos u operaciones que nos permiten llegar a obtener, objetivamente y con mayor certeza posible, la información acerca de la expresión de la conducta humana.

Así la medición objetiva se refiere a que la puntuación y la interpretación de dichas puntuaciones son independientes del juicio subjetivo del examinador ya que en este tipo de tests el sujeto debe seleccionar una respuesta de entre varias (verdadero-falso o de elección múltiple, de completamiento), además para su valoración se aplican plantillas que como es obvio, dará siempre los mismos resultados, sea cual fuere el examinador.

Durante la primera mitad de este siglo surgió la necesidad de medir ciertos atributos de la conducta, los primeros fueron Esquirol, Pineal, Ribot, Itard, Charcot, Seguin, Weber, Fechner y Wundt.³

Los tests se hacen porque se admite que los individuos difieren entre sí y porque el papel de la educación debe ajustarse a esas diferencias, en forma tal que cada persona desarrolla su potencial idiosincrático.

Según Webster, el uso del tests en educación significa “cualquier serie de preguntas o ejercicios o cualquiera medios que sirven para medir

¹ MORALES M. L.; *Curso de Psicometría Aplicada*; Ed. Trillas, México 1974 p. 6.

² MORALES M. L.; *op. cit.*, p. 7.

³ MORALES M. L.; *op. cit.*, p. 9.

la habilidad, conocimiento, inteligencia, capacidades o aptitudes de un individuo o grupo".⁴

Por lo general; la palabra test se refiere a evaluar algo o a alguien con arreglo a criterios dados, para obtener datos que revelen la relación entre un *x* sujeto y un marco de referencia.

Tests estandarizados

En el campo de la medición educacional y psicológica, estandarización significa, procedimiento fijo o uniforme que se sigue al administrar y calificar tests.

Un test estandarizado puede administrarse, pues, a un grupo o a un individuo, pero siempre bajo las mismas condiciones. El tiempo regularmente es el mismo y las respuestas a las preguntas se califican de la misma manera. Un test debe aplicarse a una gran muestra que represente el tipo de sujetos a los que se destina. Este grupo, conocido "muestra de tipificación", sirve para establecer las normas, que no solo indican el promedio de actuación sino también la frecuencia relativa de los diversos grados de desviación por encima y por debajo del promedio.⁵

De esta manera, es posible valorar los diferentes grados de superioridad e inferioridad.

Las condiciones de examen, las instrucciones y las calificaciones son siempre iguales. La calificación del test será la misma independientemente del lugar en que se realice.

Las respuestas correctas son únicas, ni la cultura, ni el criterio de quien la hace, justifican su alteración o cambio.

El principal valor de un test estandarizado estriba en ser un instrumento objetivo para medir habilidades o logros sin error subjetivo.

Los tests psicológicos y su clasificación

E. L. Thordike expresó "si alguna cosa existe ella existe en alguna

⁴ KARMEL L. J.; *Medición y Evaluación Escolar*; Ed. Trillas México 1978, p. 20.

⁵ KARMEL L. J.; *op. cit.*, pp. 23-24.

cantidad” y “si existe en alguna cantidad es susceptible de medición”.⁶

Así pues los atributos de conducta también pueden ser medidos, y para ello se construyen instrumentos adecuados denominados tests.

“Un test psicológico constituye esencialmente una medición objetiva de una muestra de conducta”.⁸

“Una prueba psicológica o test es una técnica sistemática que compara la conducta de 2 o más personas”.⁸

Es conveniente aclarar que los instrumentos de medición psicológica no se restringen a los denominados “pruebas mentales”.

También es importante señalar los factores que intervienen en la constancia o estabilidad de las puntuaciones de las pruebas de inteligencia.

Los valores predictivos de las puntuaciones de los tests de inteligencia depende del intervalo entre el test y el retest así como de la edad en la que se aplica el test inicial, ya que las puntuaciones ganan estabilidad de la infancia a la madurez. Un factor importante en el aumento de estabilidad en las puntuaciones de los tests con la edad, es la naturaleza acumulativa del desarrollo de la conducta y la consiguiente superposición de las puntuaciones.

Clasificación de los tests

Las pruebas psicológicas según lo refiere Ma. Luisa Morales⁹ se clasifican de acuerdo a ciertos aspectos como son:

Medición, Tipo de respuestas, Formas de administración, Manera de dar las instrucciones, Tipo de material usado, etcétera.

Por su medición se divide en pruebas de potencia o ejecución máxima y pruebas de ejecución típica.

⁶ ANASTASI Anne; *Psychological Testing*; Ed. The McMillan Company. New York, 1966, p. 31.

⁷ ANASTASI Anne; *Fundamentos de los Tests Psicológicos*; Ed. Aguilar; Madrid, 1973, p. 21.

⁸ ANASTASI Anne; *Fundamentos de Exploración Psicológica*; Ed. Bibl. Madrid, 1963, p. 38.

⁹ MORALES M. L.; *op. cit.*, pp. 26-31.

Pruebas de potencia o de ejecución máxima.

Estas pruebas requieren la habilidad máxima del sujeto.

Se clasifican dentro de este tipo las pruebas de inteligencia: las de aptitudes o habilidades y las de rendimiento (perceptual, escolar, etcétera).

Pruebas de ejecución típica

Son pruebas más coplejas y menos obvias que las pruebas de potencia, puesto que mide conductas menos estables. Intenta precisar lo que un sujeto acostumbra hacer en determinada situación y en una gama amplia de circunstancias. Entre ellas se encuentran las de personalidad, intereses, valores, etcétera.

Se suelen estudiar también por medio de la observación y la autoobservación.

De acuerdo al *tipo de respuestas requerido* se dividen en:

a) pruebas objetivas, b) subjetivas y c) pruebas proyectivas.

Como ejemplo de las pruebas objetivas están las de inteligencia; de las subjetivas, las de personalidad, y básicamente proyectivas el T.A.T.; Rorschach.

Los test proyectivos presentan el material de tal manera que el individuo va proyectando sus contenidos psicológicos sin darse cuenta.

Por la *forma de administración* de la prueba se clasifican en: a) individuales, b) colectivas, y c) autoadministradas.

Los tests colectivos se pueden aplicar a grupos numerosos (de 30 a 40 personas) siempre que permita la realización individual de la prueba.

Por la forma de dar las *instrucciones* y de *esperar las respuestas* pueden ser escritas u orales.

Por el *tipo de material* que se utiliza, pueden ser de lápiz y papel, de ejecución por medio de aparatos verbales y no verbales.

En los de lápiz y papel el sujeto escribe sus respuestas en el mismo test o en una hoja especial para respuestas.

En los de ejecución el sujeto maneja diferentes elementos como dibu-

jos, cubos, aparatos, etcétera. Su aplicación se ha restringido a la forma individual.

Los test verbales plantean los reactivos de tal forma que las respuestas sean de manera oral o escrita.

Los tests no verbales surgieron de la necesidad de evaluar a individuos con deficiencia sensorial. Carecen por completo de material verbal; las instrucciones de acuerdo al caso pueden ser escritas, orales o mímicas.

Por la forma de *ejecutar la calificación*, pueden ser pruebas de calificación manual y de calificación de computadora.

En base al *fin que persiguen* son:

- a) pruebas para bebés
- b) " " pre-escolares
- c) " " escolares
- d) " " adolescentes
- e) " " adultos

Unidades de medición y normas de calificación

Las calificaciones crudas, o sea, el número real de unidades o puntos obtenidos por un individuo, no tienen en sí mismas significado.¹⁰

En muchas situaciones donde se aplican medidas psicológicas, los resultados se deben comparar por métodos diferentes.

Un puntaje individual en un test se obtiene sumando el número de items resueltos correctamente, sin embargo, el puntaje original individual obtenido por un sujeto mediante tal procedimiento depende no solamente de su propia ejecución sino también de las propiedades del test, como por ejemplo el número de items del test, la diferencia de éstos, etc.

Estas cantidades tienen significado en función de un estándar o parámetro de comparación.

A fin de facilitar la interpretación en pruebas psicológicas se elaboran tablas de normas.

Norma: es el promedio o calificación sobre una prueba dada elabora-

¹⁰ MORALES M. L.; *op. cit.*, p. 30.

da para una población determinada. Esto permite ordenar la ejecución del test de la muestra de tipificación.

De este modo, las normas se establecen empíricamente determinado lo que un grupo representativo de personas hace en el test. La puntuación directa de cualquier individuo se refiere a la distribución de las puntuaciones obtenidas por la muestra de tipificación para averiguar en qué lugar de esta distribución figura.

Existen varias formas de convertir las puntuaciones directas, éstas son: de edad o cronológicas, percentiles o típicas.

Puntuaciones de edad

En las escalas de edad los elementos individuales están agrupados en niveles de edad. Por ejemplo, aquellos elementos resueltos por la mayoría de los sujetos de 10 años de la muestra de tipificación se colocan en el nivel de los de 10 años y así sucesivamente. La puntuación de un niño en un test corresponderá al nivel más alto de edad que pueda completar satisfactoriamente.

La utilidad de las normas de edad se limita en gran parte a la infancia.

Puntuaciones percentilares

Las puntuaciones percentilares son las que se expresan en función del porcentaje de personas, en el grupo normativo, que queda por debajo de una puntuación directa determinada.

Un percentil indica la posición relativa del individuo en el grupo normativo. Así por ejemplo, el 28 por 100 de los sujetos resuelven correctamente menos de 15 problemas en un test de razonamiento aritmético, lo cual corresponde a una puntuación directa de 15 al percentil 28.

Percentil 50: Los percentiles por encima de 50 representan ejecuciones superiores al promedio y los que están por debajo de 50, ejecuciones inferiores. Estas puntuaciones son fáciles de calcular y bastante comprensibles, son universales y se puede emplear lo mismo para niños que para adultos y para cualquier tipo de test.

Puntuaciones típicas

Las puntuaciones típicas expresan la distancia del individuo a la media en función de la desviación típica de la distribución.¹¹

Cualquier norma, sea cual fuere el modo en que se exprese, se halla limitada a la población normativa determinada de la cual se ha derivado. Las normas de los tests psicológicos de ninguna manera son absolutas, universales o permanentes.

En el desarrollo y aplicación de las normas del test debe prestarse una considerable atención a la muestra de tipificación. Como es natural, la muestra sobre la que han de basarse las normas debe ser lo suficientemente grande para que pueda proporcionar valores estables.

Cualquiera otra muestra que se escoja de manera análoga en la misma población no habrá de dar normas que difieran apreciablemente de las obtenidas.

Requisito igualmente importante es el de que la muestra sea representativa de la población que se considera. Deben investigarse cuidadosamente los factores selectivos más sutiles que pudieran hacer que la muestra no fuera representativa.

Cómo valorar un test

Existen consideraciones generales en la evaluación de los tests, la principal, es la de si el test mide realmente lo que se afirma y pretende, la siguiente en importancia es la de si el test mide en forma consistente y exacta. Estos factores citados se refieren a la validez y confiabilidad respectivamente.

Validez

La variable más importante para juzgar la adecuación de un instrumento de medida reside en su validez, es decir, en el grado en que el test mide realmente lo que pretende medir.

El instrumento de medida que se usa en un caso y con propósitos dados

¹¹ ANASTASI Anne; *op. cit.*, p. 55.

debe realmente medir el rasgo que se intenta medir, a este rasgo se le llama variable de criterio.

No hay problemas cuando se miden propiedades físicas de los objetos, pero cuando se miden características de tipo psicológico es necesario probar empíricamente que el instrumento mida la variable que se intenta medir en cada caso específico.

Por ejemplo, si se quieren medir los grados que los puntajes de éste realmente distinguen grados de ansiedad y no otras diferencias.

La validez proporciona un control directo de la forma en que cumple su función. La validez es cuestión de grados.

Un test rara vez es completamente válido y casi nunca es completamente inválido.

Los tests se usan para diferentes propósitos evaluativos, y para cada propósito se requieren distintos métodos de investigación para establecer la validez.

Tipos básicos de validez

No se puede hablar de una validez general, sino de una validez relativa al tipo de decisión que se toma de acuerdo al test aplicado.

Validez predictiva (o de criterio)

Muy frecuentemente, sobre todo en clasificación o en selección, las decisiones se basan en la ejecución futura de una persona, de acuerdo a los datos que proporciona el test.

Si se confirman los datos obtenidos en el test significa que éste posee un gran valor informativo, en caso contrario no tiene ningún valor.

Para conocer la validez del test, se requiere continuar observando al individuo.

En la selección o clasificación se toman en cuenta ciertos datos futuros de la conducta del sujeto, se aplica un test, se hacen predicciones y se espera el comportamiento del sujeto. Entonces es cuando se obtiene un registro de la conducta futura del sujeto, a este registro se le llama criterio y se le compara con la predicción hecha.

Validez concurrente

Se compara el test con una fuente actual informativa, es decir, consiste en cotejar la conducta presente del sujeto con los resultados de los tests. Por ejemplo, se puede comparar el diagnóstico a través del test psicológico clínico, con el juicio del psiquiatra que explora al enfermo en cuestión.

Este tipo empírico de constatación recibe el nombre de validación concurrente porque las 2 fuentes de información datan aproximadamente de la misma fecha. Este tipo de validez se usa para tests en situaciones de diagnóstico, o sea que se dispone de la medida de la variabilidad de criterio en el momento en que se obtienen los resultados del test.

Validez de contenido

En este tipo de validación cabe hacerse la siguiente pregunta. ¿Este test representa el contenido o las actividades que yo intento medir?

En vez de comparar las puntuaciones en el test con alguna otra medida, se examinan los items en sí mismos y comparándolos con el contenido que se desea incluir. Así, por ejemplo, la validez del contenido del test de español tendría que haber sido estudiada comparando los items con el programa de estudios que han seguido los estudiantes de español.

Las aptitudes, destrezas y conocimientos requeridos del estudiante para considerar que ha tenido éxito en la ejecución de un test debe ser, precisamente, de los tipos de aptitudes destrezas y conocimientos que la escuela desea fomentar en los estudiantes y evaluar en términos de puntuaciones de un test.

El autor del test, al elaborar sus reactivos, deberá consultar todas las fuentes de información asequibles y confiables relacionadas con la materia de un test. Por lo tanto, se debe evaluar la validez de contenido de un test en términos del grado en que los reactivos del test se equiparan con sus objetivos particulares, y por el grado en que la muestra de items del test es representativa de la población total.

Validez de construcción o de hipótesis de trabajo

Siempre que un examinador se pregunta qué significa psicológicamente una puntuación, se pone en cuestión cuáles son los conceptos que pueden ser utilizados de manera adecuada para interpretar la ejecución en el test. A este tipo de concepción teórica se le llama hipótesis de trabajo y al proceso de validar tal interpretación se le denomina validación de la hipótesis de trabajo.

El concepto de validez de construcción es especialmente útil en relación a los tests que miden rasgos para los cuales no hay un criterio externo.

Crombach y Meehl (1966) en sus estudios sobre la validez de constructo, dicen: Un constructo es un atributo postulado que supuestamente se refleja en el desempeño de los individuos en un test.¹²

En la validación de un test, el atributo del que hacemos afirmaciones al interpretar esta última es un constructo.

Esperamos que una persona en algún momento posea o no cierto atributo cualitativo (amnesia) o una estructura, o posea en algún grado cierto atributo cuantitativo (alegría).

Un constructo tiene ciertos significados, expresados en afirmaciones de este carácter general: las personas que poseen este atributo en la situación *x*, actuarán en la forma *y*. Por ejemplo, la teoría de la ansiedad, aceptada por algún examinador, podría incluir las siguientes expectativas: si se expone a las personas a la amenaza de un choque eléctrico, su ansiedad aumentará.

Los neuróticos alcanzan más ansiedad que los no neuróticos, la administración de ciertas drogas disminuye la ansiedad, las personas ansiosas poseen un alto nivel de aspiración; cada una de estas expectativas puede ser probada mediante un experimento o un estudio estadístico de diferencias de grupo.

La validez se determina mostrando que las consecuencias que pueden predecirse sobre la base de la teoría con respecto a los datos del test pueden, en lo fundamental, confirmarse por una serie de pruebas.

¹² KARMEL L. J.; *op. cit.*, p. 125.

La mayor dificultad de los estudios empíricos consiste en obtener un índice adecuado para medir el criterio.

Para obtener una estimación sencilla acerca del valor de un test se obtiene una correlación. Está en un índice estadístico de la relación que existe entre 2 variables. Es el método más corriente para responder a ciertas preguntas.

¿Predice el test el rendimiento en este oficio? ¿Estos tests miden la misma cosa?

Índice de validez

La magnitud del índice de validez de un ítem depende de la magnitud de la correlación entre los puntajes del ítem y los del criterio, y de la magnitud de la desviación estándar del ítem.

Una expresión numérica directa del coeficiente de validez es la proporción entre la suma de los índices de validez de los ítems y la suma de sus índices de confiabilidad.

Confiabilidad

El conocimiento de la confiabilidad es necesaria para que los datos de los instrumentos de medición psicológica puedan usarse correctamente. ¿Cuál es la estabilidad de las puntuaciones en el test? Cuando se midan nuevamente a las mismas personas, ¿qué tan consistentes son las puntuaciones del test?

La confiabilidad de un test se refiere a la consistencia de las puntuaciones obtenidas por los mismos individuos cuando son examinados con el mismo test en diferentes ocasiones, con conjuntos distintos de elementos equivalentes o bajo otras condiciones de examen.

En un sentido más amplio la confiabilidad del test indica hasta qué punto atribuirse a errores de medida las diferencias individuales en las puntuaciones del test y hasta qué punto cabe atribuir las diferencias verdaderas de las características sometidas a consideración.¹³

¹³ ANASTASI Anne; *op. cit.*, p. 74.

Para establecer la confiabilidad cuando se miden las características de un objeto con la ayuda de instrumentos físicos sólo se requiere emplear los mismos instrumentos. Pero cuando se refiere a variables de tipo psicológico, a menudo es difícil estimar la confiabilidad igual que en el caso anterior ya que los individuos que fueron sometidos a la prueba fueron afectados por la primera aplicación del instrumento; este efecto es diferente en cada individuo. Por lo cual los resultados obtenidos en la primera medición, serán un tanto diferentes a los de la segunda aplicación, simplemente porque el material ya ha sido presentado una vez.

Si hay un espacio de tiempo entre las 2 aplicaciones, los individuos pueden haberse desarrollado de diferentes maneras.

En consecuencia, cualquier condición que no sea afin al propósito del test representa una varianza de error. Así cuando un examinador trata de mantener uniformes las condiciones de aplicación del test, controlando ambiente, instrucciones, límites de tiempo, rapport y otros factores similares, está reduciendo la varianza de error y haciendo más confiables las puntuaciones. Sin embargo, aún dándose las condiciones óptimas, ningún test es un instrumento perfectamente confiable.

Tipos de confiabilidad

Hay diferentes métodos para estimar la confiabilidad, entre ellos están el de Test-Retest, Tests Paralelos, División por mitades y el método de Kuder-Richardson.

Método Test-Retest

El método más obvio para hallar la confiabilidad de un test consiste en aplicarlo en una segunda ocasión, (es decir, que requiere de la aplicación del mismo test dos veces).

La confiabilidad del retest indica hasta qué punto las puntuaciones en un test pueden generalizarse en distintas ocasiones; cuanto más alta sea la confiabilidad, menos susceptibles son las puntuaciones a los cambios diarios casuales que se producen en la condición del sujeto o en el ambiente en que se aplica el test.

Esta técnica, aunque aparentemente sencilla y directa, presenta dificultades al ser aplicada a la mayoría de los tests psicológicos. La práctica originará probablemente valores diferentes de mejora en las puntuaciones del retest de varios sujetos. Además, si el intervalo entre los tests es bastante corto, los sujetos recordarán muchas de sus primeras respuestas. En otras palabras, pueden darse el mismo patrón de respuestas correctas y equivocadas al intervenir la memoria.

De esta manera, las puntuaciones de las dos aplicaciones del test no se obtienen independientemente, y la correlación entre ellas sería falsamente alta.¹⁴

Según Lindeman (1967), aunque la técnica del reexamen no sea apropiada a la mayoría de los tests psicológicos, hay algunos tests motores y de personalidad que no resultan muy afectados por la repetición y pueden tratarse por el método de reexamen.¹⁵

Método de tests paralelos

Una manera de evitar las dificultades con que se tropieza en la confiabilidad del retest se consigue mediante el uso de formas equivalentes del test. Puede darse a los sujetos una forma de test en la primera ocasión y otra forma paralela en la segunda.

En el desarrollo de formas equivalentes debe asegurarse que verdaderamente sean paralelas. Las formas paralelas de un test deberán ser tests contruidos independientemente y destinados a satisfacer las mismas especificaciones.

Los tests han de contener el mismo número de elementos y éstos tienen que expresarse de la misma forma y abarcar el mismo tipo de contenido. Asimismo la amplitud total y el nivel de dificultad de los elementos deben ser iguales.

Hay que controlar para su comparabilidad las instrucciones, los límites de tiempo, los ejemplos ilustrativos, el formato y todos los demás aspectos del test.

¹⁴ ANASTASI Anne; *op. cit.*, pp. 81-85.

¹⁵ KARMEL L. J.; *op. cit.*, p. 130.

El concepto de muestreo de los elementos, o muestreo de contenido, no sólo constituye la base de la confiabilidad de la forma equivalente, sino también de otros tipos de confiabilidad.

La confiabilidad de la forma equivalente deberá ir siempre acompañada de una indicación sobre el lapso transcurrido entre las diversas aplicaciones del test, así como una descripción de las experiencias significativas sucedidas.

Si se aplican las dos formas en sucesión inmediata, la correlación resultante sólo indica la confiabilidad existente entre las formas, no entre las diversas circunstancias. Aunque la confiabilidad de las formas equivalentes es mucho más ampliamente aplicable que la del retest, también tiene sus limitaciones. En primer lugar, si las funciones de la conducta que se está considerando están sujetas al efecto de una larga práctica, el uso de formas paralelas reducirá este efecto pero no lo eliminará; otra limitación está en la diferencia real entre las formas, además se debe tener en cuenta que administrar otra forma de un test es costoso y toma mucho tiempo.¹⁶

Método de división por mitades

Partiendo de una sola aplicación de una forma de test es posible llegar a la medida de su confiabilidad siguiendo varios procedimientos de división por mitades. Se obtienen dos puntuaciones para cada individuo el test en medidas comparables.

La confiabilidad de la subdivisión proporciona una medida de equivalencia o adecuación de las muestras de elementos.

En esta medida no entra la estabilidad temporal de las puntuaciones, puesto que sólo requiere una sola sesión.

Cualquier test puede dividirse en muchas formas: en la mayor de ellos, la primera y la segunda mitad no serán comparables, así como a causa de los efectos acumulativos de la agitación, la práctica, la fatiga, el aburrimiento o cualquier otro factor que varíe progresivamente desde al principio al fin del test.

¹⁶ ANASTASI Anne; *op. cit.*, p. 84.

Después de calificar el test, se colocan los items en la matriz de puntajes en orden de frecuencia de solución correcta luego se forma un test paralelo con los items numerados pares y otro con los items impares, para que los tests sean igualmente difíciles. Si los elementos estaban originariamente dispuestos en un orden de dificultad aproximadamente equivalentes de las mitades.

Método de Kuder-Richardson

Un cuarto método para hallar la confiabilidad de un test, también con una sola aplicación de una sola forma, se basa en la consistencia de las respuestas de los sujetos a todos los elementos del test. Así que este método supone que todos los items tienen la misma medida y la misma varianza.

Coefficiente de correlación

Puesto que todos los tipos de confiabilidad se refieren al grado de consistencia o concordancia entre 2 conjuntos de puntuaciones derivadas independientemente, pueden expresarse todos en función de un coeficiente de correlación.

Conociendo la confiabilidad de los instrumentos de medición se pueden interpretar los datos con un grado conocido de confianza.

Esencialmente, un coeficiente de correlación (r) expresa el grado de correspondencia o relación entre 2 conjuntos de puntuaciones. Así el individuo que puntúa más alto en la variable 1 obtiene también la puntuación más elevada en la variable 2, el que logra el segundo puesto en la variable 1 es también el segundo en la variable 2, así sucesivamente hasta el individuo de puntuación más baja del grupo existirá una correlación perfecta entre las variables 1 y 2 cuyo valor será más de 1.00.

Los coeficientes de correlación pueden calcularse de varios modos, según la naturaleza de los datos. El más común es el coeficiente de correlación producto momento de Pearson.

Este coeficiente no sólo toma en cuenta la posición del individuo en

el grupo, sino también la magnitud de su desviación por encima o por debajo de la media del grupo.

El coeficiente de confiabilidad del Test-Retest es simplemente la correlación entre las puntuaciones obtenidas por los mismos sujetos en las dos aplicaciones del test. La varianza de error corresponde a las fluctuaciones en la actuación debidas al azar desde una sesión de prueba a otra. Estas variaciones pueden ser debidas, en parte, a condiciones de examen no controladas.

Para el método de Tests Paralelos la correlación entre las puntuaciones obtenidas en las dos formas representan el coeficiente de confiabilidad del test. Este coeficiente combina, por tanto, dos formas de confiabilidad. Sin embargo, como ambos aspectos son importantes para la mayoría de los fines de la aplicación de tests, la confiabilidad de la forma paralela proporciona un índice útil para la valoración de muchos de aquéllos.

En el método de División por Mitades se computa el coeficiente de correlación entre los puntajes de los tests mitades. Este coeficiente de correlación puede considerarse como el coeficiente de confiabilidad de uno de los tests mitades, si son paralelas se estima el coeficiente de confiabilidad para el test entero por medio de la fórmula de Spearman-Brown.

El procedimiento más común para hallar la consistencia entre los elementos es el de Kuder y Richardson en el cual se hace un examen de la ejecución de cada elemento. La fórmula 20 de Kuder-Richardson es la más conocida y de más amplia aplicación, su representación es la siguiente:

$$r = \left(\frac{n}{n-1} \right) \frac{\sigma^2 - \sum pq}{\sigma^2}$$
 donde r es el coeficiente de confiabilidad de la totalidad del test; n , el número de elementos y σ la desviación típica de las puntuaciones totales.

$\sum pq$ se halla tabulando la proporción de personas que resuelven (p) cada elemento y de las que no lo resuelven (q). Se calcula el producto de p y q para cada elemento y se suman estos productos para todos los elementos, lo que da $\sum pq$.

Lo verdaderamente importante en la correlación consiste en saber si

¹⁷ ANASTASI Anne; *op. cit.*, p. 88.

es significativamente mayor que cero. Cuando se dice que una correlación es significativa al nivel de 1×100 (0.01) quiere decir que no hay más de una probabilidad de cada 100 de que la correlación en la población sea cero.

Los niveles de significación se refieren a los riesgos que se está dispuesto a afrontar al extraer conclusiones de los datos. Si se dice que una correlación es significativa al nivel del 0.05, la probabilidad de error es de 5 veces por cada 100. La investigación psicológica suele emplear los niveles del 0.01 ó del 0.05. Estos niveles se encuentran en las tablas de consulta de significación de correlaciones.

Análisis de items

El puntaje obtenido por un individuo en un test es por regla general el número de items que resolvió correctamente.

La seguridad de que este puntaje obtenido sea la estimación de su puntaje verdadero depende respectivamente de la confiabilidad y validez de los datos y por ende ellos determinan el valor del test.

La confiabilidad y la validez de los datos depende de las propiedades de los items individuales que forman el test.

El test total no tiene propiedades que no puedan derivarse de las que poseen los items individuales o de las relaciones entre ellos.

La elección de items debe basarse en un análisis de probabilidad de que cada item incrementa la confiabilidad y la validez. Este análisis de items puede hacerse por varios métodos. Un item contribuye a la confiabilidad del test cuando mide la misma clase de puntaje verdadero que los otros items del test, es decir, contribuye a la confiabilidad si el componente verdadero del item está determinado por el mismo factor que determina la magnitud de los componentes verdaderos, medidos por los otros items. Igualmente un item contribuye a la validez del test si mide la misma clase de puntaje verdadero que la medida de criterio. Si el item contribuye a la confiabilidad del test, tendrá una correlación positiva con los demás items si contribuye a la validez del test tendrá una correlación positiva con la medida de criterio.

Los métodos usados para el análisis de items se divide en dos grupos

principales: a) métodos de atajo que investigan las diferencias entre los grupos extremos en la distribución del test y del criterio con relación a su capacidad para resolver cierto ítem, y b) métodos que determinan el grado de la relación por un coeficiente de correlación.¹⁸

Métodos de correlación

Correlación continuo biserial. Esta correlación se utiliza con una variable de distribución de puntajes continua y otra con dos categorías de puntajes.

Esta correlación se obtiene con la siguiente fórmula:

$$r_{pbis} = \frac{M_p - M_y}{S_y} \sqrt{\frac{p}{q}}$$

El máximo valor del coeficiente depende de la relación entre las proporciones p y q , en la distribución. Aún para proporciones iguales de p y q el coeficiente máximo es menor que 1 y se reduce más cuando la diferencia entre estas proporciones aumenta.

Correlación biserial

Una suposición fundamental, a menudo usada en la teoría de los tests, es que los puntajes verdaderos de los individuos en variables psicológicas tienen una distribución normal. Sin embargo, se obtendrá una distribución dicótoma si el test consiste de un solo ítem y se califica con 1 la respuesta correcta y con 0 la incorrecta. Esta distribución con dos categorías será el resultado de dividir una distribución normal en 2 partes. Los individuos que se hallan arriba del nivel de pase en el continuo de ejecución se ponen en una categoría y los que están abajo se ponen en la otra.

Tanto el método continuo biserial como el biserial, se emplean para

¹⁸ MAGNUSON David; *Teoría de los Tests*; Ed. Trillas, México 1969, p. 238.

estimar el grado en que el ítem mide la misma variable que los demás ítems del test.

Análisis de ítems: diferencias entre grupos extremos

Si al hacer un análisis de ítems se desea probar la contribución de un ítem a la confiabilidad del test, cuanto mayor es la correlación entre las medidas del test y las medidas hechas con el ítem mejores esta distribución.

Se puede obtener una expresión para la correlación del ítem con el test comparando las proporciones de los individuos con puntajes en las mitades superior e inferior de la distribución de puntajes del test, quienes han sido capaces de resolver el ítem.

La diferencia $p_u - p_l$ (donde p_u es la proporción de la mitad superior de la distribución de puntajes de quienes resolvieron el ítem del test, y p_l es la proporción de la mitad inferior de la distribución de quienes resolvieron dicho ítem) es una medida del grado de relación entre el ítem y el test.

Correlación entre el ítem y el test

La relación entre el ítem y el test puede expresarse como un coeficiente de correlación.

Cuando se examina la contribución de un ítem a la confiabilidad, se busca conocer si ésta es mayor con el ítem que sin él, para esto se examina si el ítem mide el mismo factor que los otros, una correlación positiva entre el ítem y los demás indicará que esto ocurre.

Los coeficientes pueden usarse como base para decidir cuáles ítems de una preliminar del test deben conservarse y cuáles rechazarse.

Frecuencia de respuesta correcta y correlación entre los ítems individuales

La frecuencia de respuestas correctas es un factor importante en el análisis de ítems ya que a menudo se intenta obtener un tipo específico de distribución de puntajes.

La frecuencia de respuestas correctas para un ítem se expresa comúnmente como un puntaje p que da la proporción del número total de individuos examinados que resolvieron el ítem. Un alto puntaje p significa que el ítem fue fácil, por lo cual fue resuelto por la mayoría de los individuos; este alto puntaje p corresponderá a un bajo puntaje estándar. Un ítem con baja frecuencia de respuestas correctas tiene un alto puntaje estándar.

Una intercorrelación de 0 entre los ítems significa que cada uno mide algo diferente de los demás, y la variable definida por ítems tan heterogéneos es difícil de interpretar.

Además la confiabilidad del test sería forzosamente baja ya que si se supone que la confiabilidad de cada uno de los ítems es la misma, la confiabilidad del test no sería mayor que la de un solo ítem.¹⁹

¹⁹ MAGNUSON David; *op. cit.*, p. 262.

CAPITULO II

LA INTELIGENCIA Y SU MEDICION

Inteligencia

A. Antecedentes históricos y teóricos de su conceptualización

A pesar del uso tan extendido del término Inteligencia no ha sido posible unificar criterios al respecto, lo cual sin duda, se debe a la complejidad de las conductas asociadas con la función inteligente.

Haciendo un breve resumen de lo que algunos personajes han entendido por Inteligencia tratamos de dar un panorama evolutivo general del concepto de Inteligencia.

En primer lugar, atendiendo a sus raíces etimológicas, la palabra inteligencia proviene de las voces latinas Intus-Legere, que significa "leer dentro" o de Inter-Legere, que significa escoger entre varios aspectos lo principal dejando a un lado lo innecesario. Se considera que también procede del término latino Intelligentia, que significa facultad de conocer, comprender, entender, etcétera.

Precisamente a partir de su significado etimológico han surgido una serie de conceptos en relación a la Inteligencia, lo cual en la mayoría de los casos dista mucho de su definición real.

Los escolásticos consideran la Inteligencia como la facultad de conocer instintivamente y la disposición de obrar inmediatamente para poseer la idea (formar juicios inmediatos) y por razón, la facultad discursiva por excelencia.

Protágoras, discípulo de Demócrito, considera que "El hombre es la medida de todas las cosas, de las que son, para que sean; de las que no son, para que no sean". Es decir, que sólo la naturaleza humana por sí misma nos puede ofrecer la base de un conocimiento estable.¹

Para Kant, inteligencia es una función pasivo-activa (receptiva-abstractiva), que requiere de la función mediadora de la imaginación, o sea, es la facultad de representarse lo que por su naturaleza no puede ser dado por la experiencia sensible pasiva).²

¹ THONNARD F. J.; *Compendio de Historia de la Filosofía*, Ed. Pontificios, París 1949, p. 32.

² THONNARD F. J.; *op. cit.*, p. 655.

Por otro lado Enrique Bergson considera a la Inteligencia como la facultad de conocer analítica o abstractamente los objetos exteriores, es decir, la potencialidad de descomponer y volver a componer de acuerdo a un sistema determinado. El mismo filósofo establece la diferencia entre la inteligencia e instinto al afirmar que el instinto es la facultad de utilizar y aun de construir instrumentos organizados, por lo cual logra la perfección desde un principio (es esencialmente especializado y no puede salirse de un grupo cerrado de actividades).

La inteligencia es la facultad de fabricar instrumentos inorgánicos, distintos de nuestros miembros; queda por tanto abierta a un progreso indefinido.

Entre ambos, dice Bergson, media una diferencia, no solo de grado sino de naturaleza, como entre lo cerrado y la abierto. No obstante son complementarios, es decir, que al lado del instinto hay huellas de inteligencia y viceversa.³

Primeros estudios psicológicos relativos a la inteligencia

En el libro "El tratado del alma y de la vida" escrito en 1530 por Luis Vives se propone la investigación del estudio de la conducta y las causas que la motivan.

Ya en forma más concreta Juan de Dios Hurarte de San Juan, plantea la necesidad de estudiar a los educandos y de describir sus capacidades en su obra "El examen de los ingenios".

En 1662 De la Chambre publica en París un libro titulado "El arte de conocer a los hombres"; es con él como empiezan a ser más profundos los conocimientos relacionados con la unidad biopsicosocial del hombre.

En 1810 las diferentes investigaciones permitieron grandes adelantos. Sin embargo, no se puede decir que se había llegado a un conocimiento del hombre, a causa de la falta de métodos científicos adecuados.

Los estudios de Charles Darwin, acerca de la vida de los organismos, fue de gran influencia para la Psicología ya que se iniciaba un período

³ THONNARD F. J.; *op. cit.*, p. 965.

de gran actividad en la investigación científica.

Si Darwin planteaba "la supervivencia del más apto entre las plantas y los animales inferiores", debería existir también "la supervivencia del más apto entre los seres humanos".⁴

Así Francis Galton en Inglaterra, uno de los pioneros de la investigación psicológica, consideró necesario establecer los principios mediante los cuales los rasgos son transmisibles de padres a hijos, para lo cual se requería llevar a cabo su medición. Para referirse a la medición de algunos de los atributos humanos consideró la necesidad de estandarizar los métodos y los instrumentos de medición, así realizó una serie de estudios sobre las diferencias individuales en la discriminación sensorial. A fin de determinar las medidas o promedios de ejecución y dispersión de tales conductas; empleo métodos estadísticos.

Fundó en 1884 el primer laboratorio antropométrico en el Sur de Kensington Museum.

En 1885 Rieger amplió las investigaciones de Galton aplicando pruebas a deficientes mentales y utilizó por primera vez el cronómetro con lo cual proporcionó mayor exactitud y precisión a sus estudios.

James Mc Keen Cattel (1860-1944) asistente de Wundt en 1885, se interesó por las diferencias individuales, en 1888 entró en contacto con Galton en la Universidad Inglesa de Cambridge.

El mérito de Cattel consiste en que aplicó prácticamente en una escala importante todo lo que había aprendido junto a Wundt y Galton. En 1890 fue el primero en utilizar la palabra *test*, y realizó numerosos estudios, especialmente en el terreno de las diferencias individuales.

Desde 1890 hasta los primeros años de nuestro siglo, los tests aplicados por Cattel y sus discípulos eran del mismo tipo que los utilizados por Galton, y derivaban, lo mismo que ellos, del material de laboratorio de Wundt: los procesos mentales (sensación, percepción, tiempo de reacción).⁵

En la última década del siglo pasado Munstenberg, Balton, Bour-

⁴ MORALES M. L.; *op. cit.*, p. 72.

⁵ REUHLIN Murice; *Historia de la Psicología*; Ed. Paidós, Buenos Aires 1973, p. 55.

don y otros, se ocuparon de estudios semejantes en relación a la inteligencia.

Jastrow en 1892 y *Wissler* en 1901, comprobaron que la correspondencia entre los resultados de los tests (como los aplicados por *Galton* y *Cattell*) es muy pequeña o nula.⁶

Pero no fue sino hasta que *Binet* y *Henry* publicaron en 1896 un artículo donde proponen la exploración de funciones complejas tales como: memoria, atención, comprensión, etc., como conductas implicadas en la Inteligencia.

Binet sostenía que la Inteligencia es una operación mental combinada que opera como un total unificado. La definió como "la tendencia a tomar y mantener una dirección; la capacidad de adaptación destinada a obtener la o las metas deseadas y el poder de autocrítica".⁷

Tomando en consideración lo anterior *Binet* y *Simón* elaboraron un instrumento que pudiera medir la inteligencia general. La finalidad de la construcción de esta prueba se basa en el principio de que se pueden identificar diferencias de mentalidad en grados de brillantez o subnormalidad con diferencias en niveles de desarrollo, representados por la capacidad promedio de niños de varias edades.

Posteriormente *Binet* hizo una revisión de su primera escala y reconoció ciertas fallas, por lo cual la escala fue modificada en cuanto a la estandarización por edades, la amplitud de la muestra, e incluyó el concepto de edad mental.

La aparición de esta prueba provocó interés en al ámbito internacional, lo cual dio como resultado un gran número de aplicaciones y evaluaciones acompañadas de sugerencias. Esto indujo a *Binet* a una nueva revisión ampliando las escalas por edades, formulando un nivel para adultos y omitiendo reactivos considerados dependientes del aprendizaje escolar.

Como era de esperarse nuevamente la prueba de *Binet* provocó un sinúmero de revisiones, traducciones, adaptaciones y estandarizaciones por diferentes investigadores, así tenemos por ejemplo a *Horning* y

⁶ REUHLIN Maurice; *op. cit.*, p. 56.

⁷ MORALES M. L.; *op. cit.*, p. 76.

Kuhlmann en E.U.A., la escala fue traducida por Goddard y Healy (1910-1911).

R. M. Yerkes. (Notable norteamericano conocido por sus obras *Introduction to Psychology* (1911) y *Mental Life of Monkeys and Apes* (1916).

Midió la inteligencia de los anormales por medio de tests; también colaboró en la selección psicológica en el ejército norteamericano durante la Primera Guerra Mundial.

El Army Test, en sus formas alfa y beta, surgió durante la Primera Guerra Mundial debido a la necesidad de seleccionar a los mejores elementos para enviarlos a la guerra, y así fueron establecidas las aplicaciones colectivas de dichas pruebas.

Terman en 1916 realizó estudios sobre la prueba de Binet a fin de adaptarla y estandarizarla a los E.U.A. Añadió una serie de pruebas para nivel adulto medio y otra para adultos superiores.

Introdujo el uso de lo que se conoce como coeficiente intelectual.

Spearman (1904) ha desarrollado una teoría acerca de la inteligencia denominada bifactorial que sostiene la presencia del factor "g" y de los factores específicos "e".

Su análisis factorial se basa en la suposición de que si dos habilidades no están relacionadas una con otra en cierta medida, se plantea la necesidad de la independencia de dichas habilidades de un factor común en esa misma medida, lo contrario también se considera posible, de aquí se deduce "g" y "e".

El factor "g" es un factor cuantitativamente único de la Inteligencia común y fundamental para todas las funciones cognoscitivas. Su magnitud es intraindividualmente constante e interindividualmente variable.

Una persona con un factor "g" elevado será suficiente en todas las actividades mentales. Incluye lo que se conoce como inteligencia general del ser humano.

El autor de esta teoría considera a "g" como la medida de algo "análogo" a la energía, esto es, a cierta fuerza capaz de ser transferida de una operación a otra diferente, y afirma que en el terreno fisiológico

existen algunas razones para esperar que se descubra una energía mental en el sistema nervioso, especialmente en la corteza cerebral.

Los factores específicos son aquellos que funcionan con cada actividad intelectual y cada una de éstas tiene su propio factor "e" que puede ser distinto de aquellos que funcionan en otra situación.

El factor "e" es cuantitativo e intraindividualmente variable, o sea que varía de persona a persona y en la misma persona, de acuerdo al tipo de actividad presente. Por lo tanto toda tarea intelectual incluye los dos factores: específico y general.

Factor de grupo "G": Es un factor común que relaciona unitariamente gran parte de un conjunto afin de habilidades. Se consideran factores de grupo: Verbal (V), Mecánico (M), Numérico (N), etc.

Así pues, a pesar de considerarse bifactorial la teoría de Spearman, se distinguen 3 factores: "e", "g" y "G".

Wible introdujo un factor considerado de grupo y denominado factor de voluntad. Garnett incluyó también un factor (H) habilidad para cambiar con rapidez de una a otra tarea.

Fueron aceptados por Spearman como factores de presencia el primero, y el segundo como de rapidez. (Spearman, 1957).

Charles Spearman define la inteligencia diciendo "que esta consiste en la educación de relaciones y la educación de correlatos".

La educación de relaciones se refiere a que antes un item o más las personas tienden a establecer relaciones entre ellos.

La educación de correlatos establece que ante un item y una relación toda persona tiende a concebir el item correlativo.⁸

Más concretamente, es la capacidad de cada individuo de establecer relaciones desde las más simples hasta las más complejas.

Esta concepción de la Inteligencia está basada en 3 leyes no genéticas que son: a) educación de relaciones b) educación de correlatos y c) el de autoconciencia o introspección.⁹

Piaget: Considera la inteligencia distinguiendo 3 aspectos de ella:

⁸ MORALES M. L.; *op. cit.*, p. 96.

⁹ RAVEN J. C.; *Tests de Matrices Progresivas*; Ed. Paidós, Buenos Aires 1957, p. 15.

contenido, función y estructura.

Contenido: manifestación observable de la conducta.

Función: actividad intelectual: conceptos y leyes (varían con la edad y el desarrollo del individuo).

Estructura: es el conocimiento que se desarrolla a través de la actividad (partiendo de la herencia y de la experiencia).

Por lo anteriormente expuesto, la inteligencia tiende como todas las demás funciones del orden sensoriomotriz y cognoscitivas al equilibrio entre el organismo (estructura biológica heredada) y el medio externo.

Para Stern la función de la inteligencia es una adaptación mental a las situaciones nuevas. Es así como contrapone inteligencia a instinto ya que éste es una forma de conducta heredada y siempre igual.

En 1949 Sir Cyril Burt identificó varios niveles de conductas intelectuales las cuales jerarquizó de la siguiente manera: En el nivel más alto las relaciones, en el siguiente (segundo nivel) las asociaciones, en el tercero percepciones y en el cuarto sensaciones.

Posteriormente subdividió las categorías que contenían más de dos factores.

Es importante señalar que antes de Burt ya Vernon había establecido un modelo jerárquico en el cual se fundamentan dos factores a partir del factor "g" 1) educación verbal y 2) K-M que implica aspectos pragmáticos. Este modelo se subdivide en 3 factores: habilidad especial, habilidad manual e información mecánica.

En los Estados Unidos de América entre los años de 1920 a 1930 se extendió el interés por estudiar el análisis factorial de Spearman. Destacando entre otros Kelly, Holzinger y especialmente Thurstone.

L.L. Thurstone (1897-1955) "Padre de la escuela factorialista americana" publicó en 1931 un artículo y en 1935 una obra acerca de su propio método (The Vectors of the Mind).

Rechazó la afirmación de Spearman y se dedicó a investigar los tipos de factores que intervienen en la Inteligencia. Con este fin empleó más de 60 pruebas psicológicas.

Al analizar dichas pruebas obtuvo los siguientes factores:

Factor V: Comprensión verbal

Factor W: Fluidez Verbal

Factor N: Numérico

Factor E: Espacial

Factor M: Memoria de asociación

Factor R: Inducción o razonamiento.

Guildford (1961) utilizando el análisis factorial, desarrolló una teoría a la que denominó "estructura del intelecto". Su clasificación se basa en el tipo de proceso o de operación ejecutada.

Así, obtuvo 5 grupos de habilidades intelectuales:

1. Factores de Cognición: Conocimiento, descubrimiento o redescubrimiento de lo conocido.
2. Memoria: Retención de lo conocido.
3. Pensamiento convergente: Conocimiento de información
4. Pensamiento divergente: Recuerdo de información.
5. Evaluación: Se clasifica mediante decisiones como "correcto", "bueno", etcétera.

De acuerdo a la clase de material o contenido que se utiliza los clasifica en:

Figuras: Material concreto, se percibe por los sentidos en forma directa.

Simbólico: Signos convencionales que representan objetos o situaciones.

Semántico: Se refiere al significado.

Conductual: Inteligencia social.

Por las unidades de los productos de información se dividen en:

Unidades: Cosas a las que se les aplican nombres.

Clases: Conjunto de objetos que participan de una o más características comunes.

Relación: Conexión entre dos cosas.

Sistemas: Patrones de interdependencia o interacción de partes.

Transformaciones: Cambios o modificaciones de información.

Implicación: Es algo esperado o predicho a partir de una información.

El número de factores del intelecto ha seguido un proceso de investigación, actualmente se considera que existen entre 50 y 100 factores.

David Wechsler psicólogo que se ha destacado en la exploración de la inteligencia, la define como "La capacidad global del individuo para actuar propositivamente, para pensar racionalmente, y para conducirse adecuadamente y eficazmente en su ambiente".¹⁰

Explica que la Inteligencia está compuesta de habilidades que no están totalmente independientes unas de otras pero que son diferenciable cualitativamente, por lo tanto al medir esas habilidades se evalúa la Inteligencia.

Fundamenta sus afirmaciones argumentando que el producto de la inteligencia no sólo está en función del número de habilidades sino también de sus combinaciones, existen otros factores independientes de las habilidades intelectuales que influyen en la conducta intelectual (motivaciones, intenciones etcétera), además la eficiencia de la conducta intelectual no depende de la cantidad de habilidades intelectuales.

Thorndike, define la Inteligencia como la capacidad de dar respuestas correctas desde el punto de vista de la verdad y del hecho.

¹⁰ MORALES M. L.: *op. cit.*, p. 116.

Fue el primero en considerar que al evaluar la inteligencia debía tomar en cuenta el número de los productos mentales, la eficiencia y la velocidad de realización.

Agrupó a la Inteligencia en 3 categorías puesto que considera que las habilidades son producciones mentales ordenadas en diferentes clases.

Inteligencia abstracta o verbal: Manejo fundamental de símbolos.

Inteligencia práctica: Manipulación de objetos.

Inteligencia social: Facilitadora de las relaciones humanas.

Donaiewsky. Considera la inteligencia biológicamente en 3 manifestaciones: comprobación de lo que falta, la búsqueda y la claridad. En esencia es la formación de hipótesis y finalmente la descarga o solución.

Stoddard. Conceptualiza la inteligencia como una habilidad en actividades: 1) difíciles; 2) complejas; 3) abstractas; 4) económicamente realizadas; 5) adaptadas a un fin; 6) de valor social y 7) que permiten originalidad y exigen cierta concentración de energía.

Holzinger. Se dedicó a determinar al mismo tiempo los factores de grupo, coincidiendo con el factor "g", llegando a aislar 9 factores comunes suplementarios, mediante el método de Spearman y su técnica bifactorial.

Bonnardel (1945), empleando el método de Thurstone y el de Spearman y con el procedimiento de Holzinger, analizó las intercorrelaciones de 26 pruebas (algunas de igual tipo y otras diferentes) pero no alcanza a obtener en 127 escolares de 14 años factor general, únicamente obtiene factores de grupo en las pruebas de igual tipo.

B. Desarrollo mental

El desarrollo mental de las capacidades mentales, según la mayoría de los autores tiene gran incremento durante los primeros años de la vida para llegar a su climax entre los 14 y los 16 años.

Thorndike considera que es hasta los 22 años.

Algunos psicólogos piensan que el desarrollo mental final y el decre-

cimiento de este varía según las diferentes capacidades y funciones.¹¹

Terman afirma que en el desarrollo mental la herencia juega un papel preponderante, aunque sin dejar a un lado la influencia del medio ambiente.

Binet ha observado que al colocar a los niños en un ambiente socioeconómico mejor, estos han aumentado considerablemente su nivel intelectual, por lo tanto defiende la influencia primordial del ambiente sobre el desarrollo mental.

Repaport expresa su teoría acerca del desarrollo mental diciendo que al nacer todo individuo trae consigo una capacidad latente "dotación natural" que es susceptible de desarrollarse si encuentra un medio adecuado.

Sin embargo, si la "dotación natural" es limitada también el desarrollo será limitado aunque el ambiente sea adecuado.¹²

Kühler encontró que los chimpancés (que manifiestan conductas semejantes en ciertos aspectos a la de los niños pequeños) tienen manifestaciones de inteligencia aunque carecen de razonamiento lógico.

Podríamos concluir que la herencia como el ambiente son factores esenciales en la capacidad mental.

Así, el régimen alimenticio, las oportunidades educativas, la salud, el nivel escolar de los padres, las enfermedades, los defectos físicos, los desórdenes emocionales, etcétera, influyen notablemente en el desarrollo de la inteligencia.

En el desarrollo de la Inteligencia intervienen en forma considerable la percepción y el lenguaje.

Percepción:

El recién nacido entra en un ambiente de estímulos mucho más complejos y muestra sensibilidad a muchos tipos de estimulación.

Como dijera William James "El niño es asaltado por sus ojos, oídos,

¹¹ SÁNCHEZ HIDALGO Efraín; *Psicología Educativa*; Ed. Universitaria, Río Piedras 1963, pp. 35-39.

¹² SZÉKELY Bela; *Los Tests*; Ed. Kapelusz, Buenos Aires 1960, pp. 356-358.

nariz, piel y entrañas al mismo tiempo, siente como si fuera un extraordinario torbellino de confusión".¹³

El desarrollo perceptual tiene 3 fases: 1) totalidad sin diferenciación, 2) sensaciones separadas de las partes, y 3) patrón integrado.¹⁴

Primeramente el niño capta los objetos sin que ellos tengan algún significado para él, después, poco a poco llega a comprender las partes como un todo significativo. De lo anterior se desprende que toda persona pasa por estas etapas para comprender el mundo que le rodea. En el adulto la comprensión es rápida gracias a la experiencia acumulada. El niño llega a interpretar en forma significativa las sensaciones cuando ya ha acumulado alguna experiencia.

Lenguaje

Schachtel (1959) afirma que no podemos recordar nada antes de adquirir el lenguaje, porque la memoria necesita de la codificación verbal.

El dominio del lenguaje en el niño se produce con lentitud.

Kaplan y Kaplan dividen la producción del lenguaje en 4 etapas que se traslapan pero que pueden diferenciarse por las vocalizaciones.

- 1) El desarrollo del lenguaje se inicia con el llanto que emite el niño y por la risa.
- 2) De 3 semanas a 5 meses el niño emite vocalización a manera de gritos variados.
- 3) Aproximadamente a los 7 meses se articulan sonidos vocales y consonantes, e imita la entonación de los adultos.
- 4) Hacia el final del primer año se inicia el habla estructurada.¹⁵

Posteriormente el niño es capaz de comprender el significado de muchas palabras que no emplea, que oye a otras personas o que encuentra en su lectura.

Mediante el uso "adecuado" del lenguaje el niño puede satisfacer

¹³ RUCH F. L.; *Psicología y Vida*; Ed. Trillas, México 1975, p. 95.

¹⁴ SÁNCHEZ HIDALGO Efraín; *op. cit.*, p. 358.

¹⁵ RUCH F. L.; *op. cit.*, p. 109.

mejor sus necesidades biológicas y sociales.

C. Investigaciones sobre la inteligencia

Diversas y variadas han sido las investigaciones que en este campo se han desarrollado, una vez que se estableció la cuantificación de esta forma de conducta al construir pruebas de esta índole.

Además se consideró la importancia de establecer relaciones entre la inteligencia y otras conductas, en América se nombraron comisiones investigadoras acerca de los métodos de aplicación de los tests.

Kelly en 1903 y Norsworthy en 1906 realizaron estudios con niños normales y otros con debilidad mental utilizando tests sensomotores y tests mentales sencillos.

También se ha explorado el efecto de la escolaridad sobre el desarrollo intelectual.

Recientemente en Suecia, Husen comparó la ejecución del test de 722 jóvenes examinados a su entrada en el servicio militar y con la obtenida dos años antes. Se subdividió la muestra en 5 grupos de acuerdo al nivel de educación, se observó que los grupos de educación inferior descendieron en su promedio y los otros grupos mostraron ganancias medias.

Owens en 1950 aplicó el test alfa del ejército a 127 hombres a los cuales ya se les había aplicado dicha prueba 30 años antes al ingresar al colegio oficial de Jowa.

Las puntuaciones del retest mostraron ganancias medias significativas en relación a la formación escolar adquirida después de la aplicación del test inicial, es decir, que su capacidad intelectual se vio favorecida (aunque no en una gran proporción) lo cual indica que la formación escolar contribuyó a su desarrollo.

Los estudios sobre el funcionamiento intelectual en la madurez y en la vejez han tratado de establecer el límite del desarrollo intelectual así como el comienzo y el índice de descenso.

Se asegura que la ejecución intelectual continúa mejorando por lo menos hasta los 30 años. Para personas que se dedican a actividades relativamente intelectuales, la mejora puede continuar toda la vida.

En base a estudios realizados se dice que el descenso en la puntuación se inicia cerca de los 50 años o poco después.

Sin embargo, existe la necesidad de una revisión sistemática de las investigaciones llevadas a cabo en los tests más utilizados y prometedores.

Dominós, prueba de inteligencia

El test de Dominós fue elaborado en 1944 por el psicólogo inglés Edgar Anstey para ser aplicado a los miembros de la armada británica como test paralelo al de las Matrices Progresivas de Raven y con la intención de superar algunos de los inconvenientes de dicha prueba.

En su versión original la prueba de Anstey está integrada por 44 items, y quedó restringida para uso oficial exclusivamente.

Para 1955 preparó una nueva versión del Dominós e introdujo ligeras diferencias, constando ahora de 48 items presentados en orden de dificultad creciente.

El Dr. Risso, Director del servicio médico pedagógico de la Universidad del Trabajo de Uruguay, estandarizó la segunda versión de la prueba de Anstey con una muestra de 1736 sujetos de 12 a 15 años, así obtuvo un baremo nacional para Uruguay, calculó los valores estadísticos de su confiabilidad ($r = 0.85$) y la validez fue contrastada con el Raven ($r = 0.55$), además elaboró el manual de la prueba.

Por el año de 1957 el Dr. Risso y J. Bernstein realizaron la estandarización regional de la prueba de Anstey.

La versión original del test de Dominós 1944 ha tenido dos adaptaciones: Una de ellas se realizó en el Instituto de Psicología Industrial de Londres, la otra fue realizada en Francia por Pierre Pichot (1950) y es conocida como Test D-48, se modificó en cuanto a la forma del item y en la distribución de las fichas.

El test D-48 ha sido estandarizado para sujetos de 12 años y medio a 39 años, proporciona normas para 6 niveles culturales: desde el nivel inferior (analfabetas) hasta el más alto (licenciatura, doctorado).

Se verificó su validez por factorización, correlación intertest y con jueces externos.

En América Latina F. del Olmo, Director de la sección de Psicología Industrial de la Universidad de Caracas, preparó una segunda forma paralela (B) en 1956 y hacia 1957 efectuó modificaciones en su presentación, en las normas según los niveles de instrucción de una muestra de 8,695 adultos de diferentes niveles escolares.

En Argentina, el Profesor Santiago J. Germano adaptó y aplicó el test D-48 al castellano según niveles de escolaridad.

Generalidades sobre el dominós

Analíticamente y por los problemas que comprende la prueba de Anstey, corresponde a los siguientes grupos:

Test de series y de educación de relaciones

Test de completamiento de series numéricas

Los principios que regulan las secuencias en las series son:

<i>Principios</i>	<i>Reactivos</i>
a) Simetría	1 - 6
b) Alternancia y Progresión simple	7 - 18
c) Asimetría	19 - 24
d) Progresión circular	25 - 30
e) Progresión compleja	31 - 36
f) Combinación de principios previos	37 - 42
g) Adición y sustracción	43 - 48

Es un test homogéneo perceptual de figuras abstractas aunque las fichas ofrecen un material concreto.

Puede ser autoadministrable, individual y colectivo.

Se fundamenta en la teoría bifactorial y Spearman.

Se clasifica dentro de los tests de potencia porque da el tiempo necesario para realizar el máximo de tareas de que sea capaz un sujeto.

Descripción del test de Anstey

Propósito

Valorar la capacidad de una persona para conceptualizar y aplicar el razonamiento sistemático a nuevos problemas.

Baines afirma que el test de Dominós es uno de los mejores tests del factor "g".

Son puntos básicos que valora el test de Dominós:

- a) Percibir exactamente el número de puntos de cada grupo de fichas.
- b) Descubrir el principio de organización del conjunto.
- c) Decidir la cantidad de puntos que debe colocar en cada una de las mitades en blanco para completar el diseño.

Utilización

Es aplicable a sujetos de 12 años en adelante, aunque puede administrarse en forma individual a niños de 10 años.

Es bastante sencillo de calificar, lo cual hace útil para la evaluación intelectual de grupos numerosos.

Administración

No es necesario preparar especialmente a los sujetos para la tarea del test, bastará con una buena motivación.

Tiempo de ejecución

Es un test de ejecución máxima, dice Anstey que los sujetos menos inteligentes nunca resolverán los problemas más difíciles por más tiempo que se les permita.

Agregó, sin embargo, que la determinación de un límite razonable

facilita la administración del test.

La experiencia indica que es conveniente no ser excesivamente riguroso ya que cuando se trabaja con adultos y con adolescentes (con sentido de autocrítica) el permitirles unos minutos más del tiempo establecido hace que la mayoría del grupo examinado termine el test.

Es aconsejable tomar un tiempo base de 30 minutos con una tolerancia de 12 ó 15 minutos.

Con respecto a los tests de poder, Guilford señala que una de las características de estos tests es la de que todos los sujetos tienen la oportunidad de resolver todos los problemas.

Clasificación de los resultados

Se clasifican con 1 cada respuesta correcta y con 0 las respuestas incorrectas. Entendiéndose por correctas aquellas respuestas que representan las cifras exactas para cada mitad de las fichas en blanco; y como erróneas, las respuestas que omiten el 0 o dejan en blanco una mitad de la ficha.

Anstey señala que hay una probabilidad de 1/49 de que la respuesta correcta se dé por adivinación.

Para calcular la eficiencia de cada sujeto en el test se divide el número de respuestas correctas entre el total de items que el sujeto resolvió, es decir, que únicamente se computan las correctas considerando el número total de las contestadas sin tomar en cuenta las que no contestó. Así tenemos:

$$\% \text{ Efic.} = \frac{C}{T}$$

Confiabilidad

Para el cálculo del coeficiente de confiabilidad en el estudio del Dr. Risso se emplearon las fórmulas propuestas por Kuder y Richardson, que son especialmente útiles cuando se mide un solo factor de la inteligencia como en el caso del Dominós.

El coeficiente de confiabilidad del test de Dominós tiende a crecer con la edad, yendo de 0.85 a 0.91.

Validez

Dentro de los criterios empleados con mayor frecuencia en el proceso de validación de nuevos tests, se emplean tests ya establecidos como válidos.

En el caso del Dominós se empleó como criterio el test de Matrices Progresivas de Raven.

Teniendo el Raven y el Dominós coeficientes de confiabilidad próximos a la unidad, el coeficiente de validez resulta aproximadamente igual al coeficiente de correlación entre ambos.

Tanto el Raven como el Dominós son tests que se formaron con el mismo fundamento, quedando demostrado que ambos miden factor "g" y un mínimo de factores específicos, aunque se considera que el Dominós mide menos factores específicos que el Raven.

Vernon demuestra lo anterior mediante un estudio realizado en 1947 y apoyado por Baines.¹⁶

Estudios realizados en México acerca del test de dominós

En el año de 1966 Ma. de la Cruz Samaniego Araujo realizó un estudio psicológico de la inteligencia juvenil con el propósito de estandarizar la prueba de Dominós en la Escuela Nacional Preparatoria para substituir al Test de Matrices Progresivas de Raven, ya que este había perdido su validez.

Para dicho estudio se tomó una muestra de 500 alumnos del 4o. año de la Escuela Nacional Preparatoria, 250 del turno matutino y 250 del vespertino.

¹⁶ BERNSTEIN Jaime; *Tests de Dominós* (manual); Ed. Paidós, Buenos Aires 1972, p. 18.

De acuerdo al análisis de items realizados en este estudio se propone un nuevo orden de los reactivos de acuerdo a su grado de dificultad: ¹⁷

1	17	35
2	18	36
3	19	34
4	21	33
5	20	39
6	22	40
7	24	41
8	23	42
9	26	37
10	25	38
11	27	45
12	28	43
13	29	44
16	30	46
17	32	48

El Dr. Luis Lara Tapia efectuó un estudio en 1967 en la Universidad del Edo. de México cuyo objetivo principal fue el de contribuir al desarrollo de la Orientación Vocacional en México mediante el estudio de varios tests, entre ellos el de Dominós.

En este estudio se utilizaron 3 muestras, cada una de ellas de aproximadamente 1 200 sujetos.

Hecho el análisis del grado de dificultad por cada reactivo se propone un nuevo orden, diferente al propuesto por Anstey.

¹⁷ SAMANIEGO ARAUJO Ma. de la Cruz; *Estudio Psicológico de la Inteligencia Juvenil*, UNAM, México 1966.

Orden Propuesto:

1	12	24	41
2	19	11	40
3	20	29	43
5	16	37	46
7	15	38	33
9	27	47	35
6	4	28	42
14	10	21	30
18	17	34	36
13	22	39	45
8	26	32	44
25	23	31	48

Carmen Moncada Ivar también llevó a cabo en 1969 un estudio titulado "El test de Dominós, un Estudio Psicométrico" con la finalidad de realizar un análisis estadístico acerca de la Confiabilidad, Validez y Grado de dificultad de los reactivos de la prueba; así como la obtención de Baremos.

Se tomó una muestra de 750 sujetos de la Comisión Federal de Electricidad, el 85% de ellos Ingenieros y el resto alumnos del 5º año de la Facultad de Ingeniería de la UNAM.

En este estudio también se propuso un nuevo orden de los reactivos de la prueba:¹⁸

¹⁸ MONCADA IVAR Carmen; *El Tests de Dominós un Estudio Psicométrico*; UNAM, México 1969.

1	10	45
2	15	42
3	20	28
4	22	35
7	5	33
19	23	34
11	25	38
8	12	37
21	32	43
9	17	29
26	31	47
6	44	48
14	18	46
13	41	40
16	36	39
24	27	30

En el estudio "Estandarización de la Prueba de Inteligencia de Dominós, en Alumnos de Secundarias Oficiales Diurnas del D. F.", efectuado en 1974 por Ma. Melania Guillén Rodríguez cuyo propósito era obtener el baremo general de la prueba para las secundarias oficiales diurnas del D. F., en cuanto a Edad, Sexo y Grado Escolar; con una muestra de 600 alumnos de 30 secundarias de un total de 102, 300 de ellos hombres y 300 mujeres, 200 por cada grado escolar: 1o., 2o. y 3o. de secundaria, se obtuvo entre otros aspectos el siguiente orden de acuerdo al grado de dificultad de los reactivos: ¹⁹

¹⁹ GUILLÉN RODRÍGUEZ Ma. Melania; *Estandarización de la Prueba de Inteligencia en Dominós en alumnos de Secundarias Oficiales Diurnas del D. F.*; UNAM, México 1974.

1	20	36
3	24	42
2	7	30
4	13	28
5	15	45
9	21	40
8	23	43
10	22	41
17	27	29
19	8	37
6	33	48
14	39	23
12	35	34
11	46	47
18	31	38
16	26	44

Test de matrices progresivas de Raven

J. C. Raven elaboró en colaboración con Penrose un instrumento para medir la capacidad intelectual en el año de 1936 asignándole el nombre de "Matrices Progresivas", posteriormente el test sufrió modificaciones y fue hasta el año de 1938 cuando se estructuró en forma definitiva.

Después elaboró otras formas para edades y capacidades específicas. Existe una escala para niños de 3 a 10 años y para adultos deficientes (Progressive Matrices 1947; Sets, A, AB, B). En 1956 apareció la escala Progressive Matrices; Sets, A, B, C, D y E, debido a la revisión de la escala de 1938, en esta revisión se reemplazaron algunos items y se modificó la ubicación de otros atendiendo a su complejidad.

Es un test que se caracteriza por el ahorro de tiempo, de material y de personal, debido a su estructuración. Esto ha permitido en cierta medida (y por otras razones que veremos más adelante) su adopción en diferentes países como Argentina, Chile, Brasil, Uruguay y Centroamérica.

Descripción del test

Propósito:

Es un instrumento destinado a “medir la capacidad intelectual para comparar formas y razonar por analogía, con independencia de los conocimientos adquiridos”.

Tareas que debe desarrollar el sujeto examinado:

- a) Descubrir la matriz de pensamiento implicada en cada figura mediante la observación, la comparación y el razonamiento analógico (educación de relaciones y educación de correlatos).
- b) y dar una respuesta a los problemas planteados completando figuras con el trozo que corresponda.

Utilización

Es aplicable para las edades de 12 a 65 años en su escala general y en la escala especial, para niños menores de 12 años o sea de 4 a 11 años y adultos dañados intelectualmente.

La escala general está constituida por 5 series A, B, C, D y E. Puede aplicarse en forma individual o colectiva.

Se presentan 60 láminas de figuras geométricas abstractas, cada serie consta de 12 items en blanco y negro. Su complejidad es creciente.

La escala especial está formada por 3 series: A, Ab y B, siendo A y B las mismas de la escala general.

Las 3 series se presentan en 36 láminas ya sea en forma de cuadernillo o de tablero. Se emplean 7 colores: negro, rojo, azul, verde, amarillo, celeste y rosa, mostrando una variedad de 58 combinaciones cromáticas, sirviendo de fondo el color más pálido en cada caso.

Material

Consta de los siguientes elementos para su aplicación individual:

- a) Colección encuadernada de láminas en blanco y negro (escala general) y de colores (escala especial).

Cada lámina tiene en la mitad superior un conjunto geométrico incompleto y en la parte inferior las opciones en figuras más pequeñas.

- b) Protocolo. Está diseñado en 3 partes: en la primera se registran los datos de identificación del sujeto. En la segunda se anota y se califican las respuestas y en la tercera se registran los datos del diagnóstico y la conducta observada por el sujeto durante el examen.

- c) Clave Matriz (plantilla de calificación).

Es una rejilla de cartulina que registra por cada serie los números de las soluciones acertadas.

Permite seguridad y velocidad al calificar.

- d) Tabla de valoración: contiene los percentiles y la tabla de diagnóstico.

- En su aplicación colectiva se forma de: a) 30 cuadernos de matrices, b) Aplicaciones de la matriz A1 y el protocolo.

Administración

Se presentan 60 problemas ordenados de acuerdo a su dificultad creciente y una serie de opciones mezclando entre ellas la respuesta correcta, en cada problema se presentan 6 u 8 opciones, 6 opciones para las series A y B, y 7 para las series C, D, y E.

Las respuestas se anotan en una hoja especial denominada protocolo.

Tiempo de ejecución

Se establece 60 minutos como tiempo límite para resolver la prueba, y 30 como mínimo.

Forma de evaluar la prueba

Se califica el acierto o el error por medio de la clave matriz de corrección donde se encuentran los números de las respuestas acertadas; se marcan en el protocolo en las casillas correspondientes a los signos más o menos, según sea la solución positiva o negativa. A cada respuesta positiva se le da un punto, los cuales al sumarse dan el puntaje total del sujeto (puntaje bruto).

La escala de puntajes es de 0 a 60. El puntaje bruto se transforma a percentilar consultando las tablas de normas percentilares para las diversas edades.

Los percentiles a su vez establecen un rango intelectual.

Validez

Se ha verificado su validez contrastado los resultados con la opinión de los Maestros, con el éxito o el fracaso en el aprendizaje y con el Terman-Merrill.

Las correlaciones obtenidas han sido satisfactorias. La validez con el Terman-Merrill es de .86.

También se han encontrado índices de saturación del factor "g" en relación con el Raven con las siguientes pruebas: Baines .75; Vernon .79; Anstey .82 y Binet .82.

Confiabilidad

Los resultados obtenidos en diversas aplicaciones del Raven en retest permiten establecer un alto índice de confiabilidad de la prueba. Sin embargo es variable con la edad, fluctúa de .85 a .95.

Características generales del Raven

Sencillo en su aplicación y evaluación.

Aplicable a niños y adultos.

Se puede aplicar a cualquier sujeto independientemente de su nivel escolar.

Interesante y agradable.

Por su material es de figuras geométricas abstractas.

Es un test perceptual y de opción múltiple.

Puede ser autoadministrado, individual y colectivo.

Es un test factorial.

Fundamentos teóricos del test

Se funda en la "teoría eclética de los 2 factores" del psicólogo inglés Charles L. Spearman. (analizados en hojas anteriores).

El Raven busca medir factor "g", es importante señalar lo que refirió Raven al respecto: "En sí mismo no constituye un test de "inteligencia general" y siempre será un error considerarlo como tal".

El test de Raven se construyó sobre la hipótesis de base "de que si los principios no genéticos de Spearman eran correctos, podrán disponerse de un test útil para comparar la gente con respecto a sus inmediatas capacidades de observación y pensamiento claro".

Raven consideró que los problemas del test implican educación de relaciones y de correlaciones.

Las dos primeras series (24 matrices) corresponden por su contenido (percepción de semejanzas, diferencias, simetría y continuidad) a educación de relaciones.

Las series C, D, y E (36 matrices) plantean problemas de razonamiento por lo cual requiere de educación de correlatos.

Relación del Raven con otros tests

Por la tarea que demandan se pueden comparar el Raven, con la prueba de 5 del Army Test Beta, con el test "Bellevue", prueba 1 de Wechsler, la prueba de Healey, la de Seguin-Goddard, Terman-Merrill, test de Richard Meili, etcetera.

Aplicaciones del Raven

Sus aplicaciones son amplias y variadas; se emplea para estudios diferenciales y sociales en edad, sexo, nivel socioeconómico, etc., con respecto a la capacidad intelectual.

A nivel educacional se utiliza para agrupar alumnos por grupo de acuerdo a su capacidad intelectual.

También se usa para la orientación vocacional y selección profesional y en las clínicas psicológicas para conocer el deterioro mental de los pacientes.

Estudios acerca del Raven

En Argentina en 1949 se evaluó el rendimiento en el test de Raven efectuando una evaluación por errores y otras por la determinación del modo de operar del sujeto.

Esta técnica fue utilizada en España (1955) por el Dr. Gonzalo Moya, en Francia por J. L. Loreche (1956) y D. Erpicum (1959) y R. Maistriaux (1959).

El Dr. Washington L. Risso (1958) construyó el primer baremo Uruguayo sobre una muestra de 2165 sujetos de 12 a 14 años.

J. Rueda (1958) también obtuvo un baremo al aplicar el Raven a 1532 estudiantes de secundaria de 15 a 21 años.

Madurga (1958) realizó un estudio experimental sobre el Raven.²⁰

Test de Wechsler

David Wechsler elaboró (1939) un instrumento de medición de la inteligencia organizado en once grupos homogéneos de pruebas denominadas subtests.

Wechsler dividió los subtests de su escala en cinco subtests verbales, cinco subtests de ejecución y un subtest de vocabulario. La suma de los puntajes pesados correspondientes a los subtests verbales, y la suma de los puntajes pesados correspondientes a los subtests de ejecución son transformados con ayuda de la tabla de C. I. de Wechsler, en el C. I. verbales y de ejecución, y la suma total de los 10 subtests, en un C. I. total.

A diferencia de la mayoría de los tests de inteligencia, se compone de varios grupos de pruebas homogéneas. En consecuencia cualquier subtest pone en juego una "función" específica o conjunto de "funciones".

²⁰ RAVEN J. C.; *op. cit.*, pp. 31-41.

Escala verbal

- a) Información
- b) Comprensión general
- c) Aritmética
- d) Retención de dígitos
- e) Semejanzas
- f) Vocabulario

Escala de ejecución

- a) Completamiento de Figuras
- b) Diseño de cubos
- c) Ordenamiento de dibujos
- d) Ensamble de objetos
- e) Símbolos de dígitos

En 1940 Wechsler inició la redacción de la forma II para comparar los resultados de la primera, sin embargo fue suspendida por la necesidad de diseñar (en ese momento histórico, Segunda Guerra Mundial) una batería denominada "Wechsler mental ability scale form B" que debía utilizarse en el ejército.

Reestructura su escala en 1955 y presenta el WAIS (Wechsler Adult Intelligence Scale) que contiene cambios en algunos items y la omisión de otros. Además se ampliaron las edades de aplicación en ambos extremos.

Antes en 1949 planteó la necesidad de evaluar la inteligencia en los niños de 5 a 15 años de edad, la cual construyó bajo las mismas bases de las anteriores, así surgió el WISC y posteriormente el WPPSI para preescolares que cubre las edades de 4 a 6 y medio años.

Es importante mencionar que el Dr. Enrique Cerda realizó la adaptación de la escala de Wechsler (W-B,II, Wisc, Wais) para su aplicación en países hispanoamericanos.

En México se han realizado estudios de investigación con la prueba de Wechsler con el fin de conocer el desarrollo de la personalidad del escolar mexicano bajo la dirección de Díaz Gurrero y Lara Tapia.

Los reactivos de la prueba se clasifican en 3 grupos:

- a) Aquellos que cualquier persona puede contestar independientemente del nivel escolar.
- b) Reactivos que se espera sean contestados por personas de nivel escolar medio.
- c) Son items que sólo pueden ser contestados por sujetos con cierta especialización.

A su vez los subtest se clasifican en varios grupos dependiendo de lo que mida cada subtest.

- | | |
|--|---|
| 1) Subtests Verbales: | Información, comprensión, semejanzas y vocabulario. |
| 2) Subtests de Atención y Concentración: | Atención y retención de dígitos. |

En la escala de ejecución hay 2 grupos de subtests:

- | | |
|------------------|--|
| 1) Visuales: | Completamiento de figuras y ordenamiento de imágenes. |
| 2) Visomotrices: | Ensamble de Objetos, diseño con cubos, símbolos de dígitos y, en la escala de niños, laberintos. |

Descripción de los subtests (escala verbal)

Subtest de Información

Ese subtest examina, en cierto sentido, la memoria. El caudal de in-

formación adquirido depende de la capacidad de dotación y de la riqueza del medio original, pero puede enriquecerse mediante la instrucción y la experiencia.

Consta de 29 ítems que permite obtener la cantidad de conocimientos que posea una persona. Se presentan en el mismo orden en que se realizó la estandarización.

Este subtest declina con la edad.

Subtest de comprensión

Examina la capacidad del sujeto para resolver y solucionar problemas de la vida cotidiana.

Consta de 14 situaciones problema en las que el sujeto debe comprender cuál es el problema y dar la o las respuestas adecuadas.

Se ha observado que el factor verbal contribuye al buen desempeño de estas tareas.

Subtest de aritmética

En el subtest de Aritmética el sujeto se concentra en el continuo de números y en el patrón de los cuatro cálculos aritméticos básicos.

La concentración se define como un esfuerzo voluntario para recibir el estímulo en forma discriminada y significativa, o bien para proyectarlo hacia afuera desde patrones existentes en el interior.

Subtest de semejanzas

En este subtest se pone a prueba la formación de conceptos verbales.

Consta de 13 pares de palabras y al sujeto se le pide que mencione la semejanza entre dos cosas lo cual demanda la búsqueda de conceptualizaciones generalizadas. Existen a grandes rasgos, tres planos diferentes en los cuales los sujetos buscan sus respuestas, son: el concreto, el funcional y el conceptual abstracto.

En el plano concreto se toma un rasgo específico común de las cosas en cuestión y se le convierte en contenido relacionante.

En el plano funcional es la función cumplida por ambas cosas. En el plano conceptual abstracto el sujeto escoge características esenciales que las cosas tienen en común y los convierte en el vínculo entre ambas. el funcional y el conceptual abstracto.

En el plano concreto se toma un rasgo específico común de las cosas en cuestión y se le convierte en contenido relacionante.

Subtest de dígitos

Algunos consideran que este subtest es de memoria inmediata; pero otros, que es un test de atención.

La habilidad requerida es la repetición de números.

Consta de una serie de dígitos que varían de 3 a 9 en orden directo y hasta de 8 dígitos en orden inverso.

Este subtest es de los más sensibles de la escala.

Subtest de vocabulario

Este subtest examina el caudal "adquirido" automáticamente en el curso de una maduración sin trabas de la "dotación natural". La riqueza e integración del vocabulario depende, en parte, de la capacidad de dotación natural y de la riqueza del medio educacional de los primeros años.

Está formado por 40 palabras tomadas al azar de un diccionario y en orden de acuerdo al grado de dificultad.

Es de las escalas menos sencibles.

Escala de ejecución

Subtest de completamiento de figuras

Es un test de concentración visual.

Está formado por 21 reactivos que consiste en tarjetas que muestran figuras a las cuales les falta una parte importante.

El examinador exhibe los dibujos y el sujeto debe descubrir qué es lo que falta en ellos.

Diseño con cubos

Es un test de coordinación visomotriz. Aquí el papel visual consiste en una diferenciación progresiva del patrón visual el objeto de la acción motriz es yuxtaponer las caras de las piezas, bajo la guía de la diferenciación visual ya obtenida; y esta parte construida inicia luego, a su vez, la diferenciación del modelo original.

Consta de 10 diseños diferentes y son progresivos en grado de dificultad.

Cada diseño está constituido por 9 cubos de colores, sólo que 4 de ellos se usan en la primera parte de la prueba.

Ordenamiento de figuras

Es un test de la aptitud de anticipación y planeamiento. En él hay que ordenar en una sucesión significativa varias series de dibujos.

Está constituido por 8 series de figuras y cada una se presenta en forma desordenada para que el sujeto las ordene y forme una historia.

Ensamble de objetos

Se pide el ordenamiento de las distintas piezas de un muñeco, un perfil y una mano. Estas figuras se presentan en forma de rompecabezas para ser acomodadas por el examinado de tal manera que representen un objeto familiar.

Mide la percepción visual, la integración visomotora y la anticipación.

Símbolos de Dígitos

Esta prueba consta de un rectángulo dividido en 9 cuadros: en los de la parte superior se colocan los números del 1 al 9 y en la parte inferior se halla un símbolo.

En seguida se encuentran 4 rectángulos divididos en 100 celdillas dobles.

El sujeto tiene que aparear los números con los símbolos que aparecen en la parte superior.

Mide factores tales como: memoria inmediata, integración visomotora, capacidad de reproducción e imitación.

Validez de la prueba

Se realizaron interrelaciones de los subtests, correlaciones con la escolaridad, correlaciones con el juicio de maestros respecto a los alumnos incremento y decremento de las calificaciones en función de la edad y correlaciones con la escala de Stanford-Binet.

Confiabilidad

Se obtuvieron coeficientes de estabilidad temporal tanto en subtests como en la escala completa, logrando un coeficiente de confiabilidad comparable a la de Stanford-Binet aproximadamente de .90.

En la ciudad de México el Instituto Nacional de Ciencia del Comportamiento y la Actitud Pública, Asociación Civil (INCCAPAC), bajo la dirección del Dr. Rogelio Díaz Guerrero, inicia en 1964 la estandarización de la prueba de Wechsler en sus formas WAIS y WISC.²¹

²¹ MORALES M. L.; *op. cit.*, p. 148.

CAPITULO III

APTITUDES Y RASGOS DE PERSONALIDAD

Aptitudes

A. Antecedentes históricos

a) Enfoque filosófico

Ya desde la antigüedad (Platón 429 - 348 a.c.) se trató de establecer una actividad de acuerdo a la capacidad del individuo.

Así, por ejemplo, Platón en su obra "La República" señala las condiciones más adecuadas para asentar en la tierra el dominio perfecto del Bien y de la Justicia; mediante la definición de un sistema denominado colectivismo racional, porque se funda no en el número y la igualdad de los ciudadanos sino en las distintas capacidades de cada uno, los divide en 3 clases correspondientes a las 3 necesidades de la sociedad: artesanos, guerreros y jefes.¹

Aristóteles (384 - 322 a.c.) dice que la inteligencia es "participación de la divinidad" y que para poseer la virtud se requiere una actuación perfecta de todas las facultades restantes, cada una en su rango y en armoniosa dependencia jerárquica.

Las condiciones humanas responden a diferencias sociales, de raza y de sexo.²

También Sto. Tomás de Aquino y San Agustín se interesaron en el terreno psicológico de las "facultades".

b) Enfoques psicológicos

En el siglo XVI diversos investigadores se interesaron por las diferencias de capacidad en el hombre, así Huarte de San Juan realizó estudios sobre la relación del trabajo con la "naturaleza" del hombre, se pueden mencionar también a Camus, Lavoisier, etcetera.

¹ THONNARD F. J.; *op. cit.*, p. 83.

² THONNARD F. J.; *op. cit.*, p. 125.

Binet, siguiendo en cierta medida a Galton, trató de medir las habilidades mediante el estudio de la quiromancia y la discriminación sensorial.

Posteriormente, considerando que era imposible medir todas las habilidades simples, trató de medir en conjunto las habilidades humanas elaborando una prueba de inteligencia.

Los trabajos de este investigador se encuentran resumidos en su obra "Las habilidades del hombre" (1927).

Spearman, interesado también en las habilidades humanas, realizó un estudio de éstas, mediante un modelo matemático.

Concluyó sus investigaciones señalando que existen varias actividades o habilidades mentales o factores independientes uno de otro, así fue como propuso su teoría factorial.

Thurstone se concretó en el desarrollo matemático del análisis factorial a fin de explotar las habilidades humanas, consideraba la existencia de habilidades primarias y la influencia constante y recíproca de diferentes factores. Así encontró 6 factores predominantes: numérico, verbal, espacial, fluidez verbal, memoria y razonamiento.

Thorndike se interesó por los productos mentales o habilidades, las cuales no podían ser explicadas por un solo factor.

Se puede concluir en base a las aportaciones de estos investigadores que la inteligencia está compuesta de muchas habilidades y que difieren de un individuo a otro, es decir, que hay diferencias individuales notables en el campo de las aptitudes.

Frederick W. Taylor despertó el interés por estudiar habilidades en el campo industrial.

Observando a las personas en el trabajo se dio cuenta de que sólo algunos sujetos eran capaces de realizarlo, debiéndose esto probablemente a una propiedad psicofísica determinada.³

³ BAUNGARTEN F.; *Exámenes de Aptitud Profesional*; Ed. Labor, Barcelona 1957, pp. 7-9.

Mediante un método psicofisiológico Camus y Nepper evaluaron en Francia las aptitudes para aviadores y choferes.

Con los tests Army alfa y el test Army de clasificación se estudiaron las disposiciones de 1.750.000 soldados para el manejo de diferentes armas y para los diversos empleos en el ejército.

En Alemania, Inglaterra, Francia y E.U. se vieron en la necesidad de conocer más ampliamente y de evaluar más eficazmente las aptitudes, ya que la situación económica y social de estos países estaba deteriorada a consecuencia de la segunda guerra mundial.

Por consiguiente se crearon los talleres-escuela (en Alemania), se elaboraron pruebas para medir habilidades y se ubicó a las personas en los trabajos de acuerdo a sus aptitudes.

Toda esta serie de pruebas llevaron al reconocimiento de la medición de aptitudes específicas mediante la creación de tests contruidos ex-profesamente y no en la forma usual de apreciar las habilidades administrando tests de puntaje único.

El uso muy desarrollado del agrupamiento de pruebas separadas se ha venido disminuyendo actualmente gracias a la elaboración de pruebas integradas, lo cual incluye ciertas ventajas en cuanto a una integración satisfactoria y adecuada para medir lo que se intenta medir, además se elabora con un propósito determinado y se cuida en forma especial su estandarización.

B. Definición de las Aptitudes

Al igual que al definir la inteligencia se presentan diferentes enfoques, así también al hablar de aptitudes se plantean una serie de desacuerdos en cuanto al propio término y a la definición misma.

¿Las aptitudes son capacidades, habilidades, disposiciones destrezas o dones adquiridos o heredados?

Es importante señalar que el hombre es un ser biopsicosocial y que en esta medida las aptitudes son el resultado de una interacción de la herencia y del ambiente. Por ende, un individuo al nacer trae consigo ciertas potencialidades que le permiten aprender al entrar en contacto con el medio ambiente.

Debemos considerar que todas las personas cuentan con aptitudes para resolver problemas, pero sólo se desarrollarán aquéllas que se ejerciten y para ello se requiere de las oportunidades que proporcione el medio.

Las aptitudes de acuerdo a algunas definiciones como la de Bingham incluyen cualquier característica que predisponga al aprendizaje, como son los intereses, la personalidad, la inteligencia y destrezas especiales.

Las aptitudes deben estar de acuerdo con el gusto del sujeto por realizar una actividad y a su vez con la capacidad general y específica de las personas y de acuerdo a su individualidad.

Es necesario aclarar algunos términos con los que se suele designar a las aptitudes.

“Don”: aptitud de dar sin esfuerzo un rendimiento productivo.⁴

“Disposición” es emplear aptitudes de menor predominancia.⁵

“Habilidades” son las que conforman a las aptitudes.

“Destrezas” son el desarrollo refinado de las capacidades.

“Capacidades” son el ejercicio de las aptitudes, se refiere a lo que una persona hace.

Cualquier aptitud requiere de numerosos sentidos y sobre esta base se pueden desarrollar distintas capacidades.

Los tests de aptitud se destinan a medir las capacidades, es decir, a predecir lo que puede adquirirse con el adiestramiento. Los tests que dicen lo que una persona puede hacer “ahora” son pruebas de realización.

Los tests de aptitud, por definición, predicen sobre realizaciones todavía no logradas. Pero sus elementos pueden consistir en muestras de realización, es decir, de lo que puede lograrse ahora.⁶

⁴ BAUNGARTEN F.; *op. cit.*, p. 52.

⁵ BAUNGARTEN F.; *op. cit.*, p. 52.

⁶ HILGARD E. R.; *Introducción a la Psicología*; Ed. Morata, Madrid 1975, p. 56.

La distinción entre un test de aptitud y uno de realización no está basada en el contenido de las preguntas o elementos sino en los propósitos de las dos clases de pruebas, es decir, en aquello para lo que están destinados. Los tests de aptitud predicen la realización futura; los de realización miden el logro presente.

Los tests de rendimiento se utilizan principalmente en las escuelas. Cualquier examen sobre lo que se ha estudiado, puede considerarse como prueba de rendimiento.

Una vez aclarados estos puntos, presentamos algunas definiciones que sobre el concepto de aptitudes se han elaborado.

Warren define las aptitudes como un conjunto de características que expresan una habilidad para aprender.

Piéron. La aptitud es la condición congénita de una determinada modalidad de eficiencia del sustrato congénito de una capacidad.

Bingham: Es una condición o serie de características vistas como sintomáticas de la habilidad de un individuo para adquirir entrenamiento (usualmente especificado), algún conocimiento, destreza, o una serie de respuestas, tales como la habilidad para hablar un lenguaje, producir música, etcétera.

Christiaens: La aptitud es una disposición hereditaria que facilita el aprendizaje o el trabajo.

C. Estudios sobre las aptitudes

Munsterberg desarrolló métodos para seleccionar a los más aptos conductores de tranvías, puesto que debido al gran desarrollo de los transportes en E.U. el número de accidentes fue mayor y se planteó la necesidad de evitarlos.

En Europa se emplearon varios instrumentos para seleccionar los diferentes aspectos de las aptitudes de los candidatos para conductores de tranvías, como son: comprensión, memoria, atención, etc. Entre otros,

Stern y Lahy utilizaron dichos instrumentos. Spierlein y Skorodenskij en Moscú en el año 1922 realizaron un método para los conductores de tranvías.

Posteriormente se realizaron estudios en otras ocupaciones como la de telefonista, impresores y oficinistas.

La gran demanda de mano de obra especializada de los países industrializados y la revolución industrial inglesa dieron un gran auge al estudio de las aptitudes para diferentes oficios y profesiones así como técnicas de medición de tiempos de reacción y los tiempos de movimientos para cada miembro del cuerpo.

D. Origen de las aptitudes en el hombre

Una respuesta a la pregunta acerca del origen de las aptitudes la encontramos en los estudios de genética de G. Mendel (1865) sobre las leyes de la hibridación.

En esencia la transmisión genética se realiza a través de los genes del padre y de la madre en el momento de la concepción. Esto nos indica que cada sujeto nace con sus propias características a excepción de los gemelos idénticos (producto de un mismo huevo) aunque se consideran algunas diferencias en cuanto a la nutrición intrauterina.

La aparición de aptitudes musicales en edades tempranas y la proliferación de este talento en varios miembros de una familia, hacen pensar que se deben a la herencia.⁷

Esto pareciera ser cierto en el caso de Juan Sebastián Bach, no así en Shuman único músico de su familia constituida por 156 miembros.

El medio en que se desarrolla una familia, la interacción familiar y los factores sociales hacen que sus miembros sean semejantes.

No se han podido determinar en qué proporción obra cada uno de

⁷ PIÉRON H.; *Tratado de Psicología Aplicada; La Psicología diferencial* (libro 1); Ed. Kapelusz; Buenos Aires 1960, pp. 25-30.

ellos. Lo cierto es que la base de las diferencias individuales han de encontrarse en la herencia de cada individuo y en las condiciones ambientales en las que se ha desarrollado. Más precisamente se puede decir que las aptitudes son el resultado de una multiplicidad de influencias estructurales y funcionales, hereditarias y ambientales.⁸

E. Desarrollo de las Aptitudes.

Puede ser considerado en 3 etapas:

- a) En la niñez la aptitud aparece en forma espontánea, desarrollándose con lentitud hasta la madurez (dependiendo del medio ambiente).
- b) Ya en la madurez la aptitud desarrollada sufre un período de estancamiento.
- c) En esta última etapa se observa un retroceso de la aptitud desarrollada como consecuencia del envejecimiento.⁹

F. Pruebas de Aptitudes.

Las pruebas que miden aptitudes se suelen clasificar en tests de aptitudes diferenciales y en tests de aptitudes específicas.

El propósito de medir habilidades particulares o específicas, radica en descubrir el "nivel" en el cual puede realizarse un individuo, por lo general no requiere de grandes procesos intelectuales superiores, más bien requiere de rapidez perceptual y destreza motora.

En términos generales en los tests de aptitud diferencial se requiere que se opere a niveles intelectuales altos.

Se hizo necesario decidir qué aptitudes debían medir estos instrumentos ya que son tantas las aptitudes que no sería posible (en función de tiempo y economía) incluirlas todas.

⁸ ANASTASI Anne; *op. cit.*, p. 54.

⁹ BAUNGARTEN F.; *op. cit.*, p. 56.

Es conveniente indicar que estos instrumentos tendrán que ser válidos y confiables, además deben ser tests independientes para poder ser utilizados en la orientación y en la selección de personal, contar con normas adecuadas, con material práctico y fácil de administrar e interpretar sus resultados.

Dentro de este tipo de pruebas podemos mencionar el D. A. T. (test de aptitud diferencial) y el G. A. T. B. (batería de tests de aptitud general).

Esta batería se utiliza para la orientación vocacional y para la selección de personal que ya ha terminado el ciclo medio básico (secundaria). Su construcción se funda en el análisis factorial de Thurstone.

Los tests de aptitudes especiales fueron construidos en un principio para completar los datos de los tests de inteligencia general. Estos tests en cierta medida abarcan algunas funciones de las baterías de aptitud múltiple.

Entre los tests de aptitudes específicas podemos citar:

El test de aptitudes Burocráticas de Minnesota.

Test de habilidades Mecánicas de Minnesota (Paterson).

Test de aptitudes para el dibujo (Salvador Navarro).

Test de Juicio Artístico de Meier.

Test Digital de O'Connor.

SRA; Aptitudes Mentales Primarias.

FACT: Test de Clasificación de las aptitudes de Flanagan.

Aptitudes Mecánicas

Lo que comúnmente se llama aptitud mecánica es en realidad una combinación de inteligencia general y ciertos aspectos de coordinación

muscular. Entre los test de habilidad mecánica se hallan algunos que miden principalmente el aspecto mental de las relaciones mecánicas, mientras que otros miden principalmente ciertas habilidades musculares o motoras, es decir, que requieren de la manipulación real de objetos mecánicos.

Se han desarrollado muchos tipos de tests de aptitud mecánica. La mayoría pertenece a una de estas dos clases. Es necesario, por consiguiente, distinguir entre dos tipos de tests de aptitud mecánica.

Como ejemplo de los primeros tenemos el test de Comprensión Mecánica de Bennett. Este test se compone de reactivos que tiene dibujos que miden la aplicación de principios físicos. Esta prueba es especialmente útil para predecir éxito en tareas que requieren la capacidad para entender el funcionamiento de maquinaria.

El test de Mc. Quiarrie de capacidad mecánica es una batería formada por 8 subtests, los primeros tres miden destreza manual: trazar, extraer y puntear; los siguientes tres: copiado, colocación y bloques son de percepción espacial, el último, búsqueda, mide rapidez y exactitud perceptual.

De los tests de aptitud mecánica del segundo tipo podemos mencionar el test de Rapidez de Manipulación de Minnesota. Consiste en un tablero con 4 hileras de 15 agujeros cada uno y 60 discos idénticos que encajan en los agujeros.

El examinador coloca los discos en sus posiciones correctas los quita y le pide al sujeto que los coloque.

Se dan 4 oportunidades. El tiempo total de examen es de 6 a 8 minutos. Es aplicable a sujetos de 13 a 50 años de edad.

Otro ejemplo de este tipo de tests es el de destreza de dedos y pinzas de O'Connor.

El test de destreza digital consiste en una vasija casi plana y en una placa metálica con 100 agujeros dispuestos en 10 hileras de 10 agujeros cada una. El sujeto debe colocar 3 alfileres en un agujero. El de destreza

con pinzas usa el lado opuesto más largo con lo que permite al sujeto colocar uno en cada agujero.

Tablero de Prude: Consiste en una tabla rectangular con 4 tazas que contienen los materiales del test en un extremo y 2 hileras de agujeros en medio. Se le pide al sujeto que coloque los alfileres en los agujeros con la mano derecha, después con la mano izquierda y posteriormente con ambas manos. La tarea final es armar el alfiler, la arandela y el collar con la mano derecha, con la izquierda y simultáneamente con ambas manos. Se le califica en base al número de alfileres colocados en 30 segundos y el número de montajes hechos en 60 segundos.¹⁰

Aptitudes Motoras

Las funciones motoras como la destreza, la actividad manipulativa y el control muscular en general no se relaciona con las funciones mentales. No es sorprendente, por consiguiente, hallar que una persona muy inteligente pueda enredarse o confundirse al intentar ensamblar una pieza sencilla de maquinaria o manejar un torno.

Aunque no se conoce comúnmente, es un hecho que los diversos tests motores muestran escasa o ninguna relación entre ellos. En un estudio se halló que la correlación media entre en gran número de tests motores era sólo de + 0.5.

Evidentemente, estas funciones son muy específicas y cada persona puede estar dotada de una combinación de aptitudes específicas completamente distintas. Por consiguiente no se puede hablar de una aptitud motora en general. Un alto grado de coordinación muscular puede limitarse a ciertas partes del cuerpo: La coordinación puede así ser superior en lo que concierne a las manos, e inferior en lo que concierne al uso de los pies. La coordinación visomanual puede desarrollarse en un grado distinto que la coordinación entre la utilización de las manos y el oído. La eficacia del control muscular puede variar de un trabajo pesado a

¹⁰ KARMEL Louis J.; *op. cit.*, pp. 230-245.

uno ligero, de un trabajo fino a uno basto y de un trabajo rápido a uno lento.

Por estas razones, los tests de funciones motoras deben seleccionarse con más cuidado que los mentales. Puesto que las aptitudes motoras son muy específicas, los tests individuales son relativamente sencillos y su administración no requiere una gran cantidad de tiempo .

Cuando lo que se quiere es probar la aptitud, la tarea exigida no duplica normalmente ninguna forma de trabajo. Cuando más se parece el test a un tipo de trabajo, más se acerca a un test de realización.

En algunos casos se diseñan a propósito tests motores para duplicar algunas de las actividades requeridas por una tarea determinada. Estos son tests de realización y se pueden utilizar para seleccionar hombres experimentados. Puede utilizarse, sin embargo, para probar la aptitud, si se administran a hombres sin experiencia o con la misma cantidad de experiencia. Cuando tales tests se usan para distinguir entre buenos y malos operarios con el mismo grado de experiencia, se convierten en tests de aptitudes muy satisfactorios.¹¹

Tests de Aptitudes

D. A. T. Test de Aptitud Diferencial

A. Antecedentes

En 1947 se desarrolló en los E. U. una prueba para medir habilidades de una manera integrada, científica y estandarizada. La construcción de esta prueba fue hecha por Bennett con 46 000 estudiantes (hombres y mujeres) del octavo al décimo segundo años con el propósito de dar orientación educativa y vocacional. Ya que una serie de gentes conectadas con estos campos habían expresado la necesidad de contar con una batería

¹¹ MAIER Norman R. F.; *Psicología Industrial*; Ed. Rialp; Madrid 1973, p. 268-474.

práctica y confiable; y que sirviera también para el campo industrial. (Bennett, 1968).

Como es de suponerse un estudio de tal envergadura no puede ser estructurado en una primera instancia como realmente integrado y completo, por lo cual fueron necesarias las revisiones hechas en 1952 y en 1959.

Durante una nueva revisión (1962) surgieron las formas "L" "M" debido a la necesidad de disminuir el costo y de simplificar su administración y calificación.

Debido a las nuevas demandas en el campo de la orientación se hizo necesario revisar nuevamente la prueba D.A.T. en el año de 1963, habiendo hecho las siguientes modificaciones:

Re-estandarización sobre más de 50 000 casos.

Cambios en el formato para facilitar aún más su administración y calificación.

Inclusión de nuevos puntajes para Razonamiento verbal y Habilidad Numérica.

B. Descripción del D.A.T.

Los tests de aptitud diferencial han sido elaborados para apreciar las aptitudes intelectuales básicas; sus contenidos (en gran parte) no dependen de los conocimientos escolares.

Está constituido por 7 folletos distintos:

Razonamiento Verbal

Habilidad Numérica

Razonamiento Abstracto

Relaciones Espaciales

Razonamiento Mecánico

Velocidad y Exactitud

Uso del Lenguaje (ortografía y oraciones)

Test de razonamiento verbal

Mide la habilidad para razonar en forma verbal; es decir, mide la capacidad para comprender conceptos, para abstraer y generalizar así como para medir el reconocimiento de vocabulario.

Predice el éxito en ocupaciones tales como: supervisores, jefes de quejas, empleos administrativos; en carreras como: Lic. en Contaduría, en Relaciones Internacionales, en Derecho, en Periodismo, etcétera, ya que todas las ocupaciones y carreras que mencionamos requieren de relaciones basadas en el lenguaje y en el uso de conceptos.

Esta prueba está formada por 50 oraciones a las que les falta la primera y la última palabra. Para completar la oración tiene que buscarse el par de palabras que aparecen abajo de cada reactivo.

Habilidad numérica

Está formada por 40 problemas numéricos, encontrando para cada problema 5 posibles respuestas.

Mide la habilidad para razonar con números, o sea, para operar inteligentemente con materiales cuantitativos.

El lenguaje puede jugar un papel importante en este test. Predice el éxito en ocupaciones y profesiones relacionados con las ciencias físicas, como por ejemplo: Contaduría, Química, Ingeniería, Estadigrafía, etcétera.

Se pueden combinar los tests de razonamiento verbal y de habilidad numérica para obtener una medida de la inteligencia general.

Se han obtenido coeficientes de correlación con el Wechsler y estos dos tests (en individuos de 16 a 17 años) de .79, para los de 18 a 19 de .75 y de .74 con los de 20 a 24 años, con la batería de contenido de High School se obtuvo una correlación de .80 para muchachos y .70 para

muchachas.

Razonamiento abstracto

Contiene 50 series de diagramas grandes y claros, dispuestos de tal manera que sean obvias las diferencias entre los diagramas sucesivos.

Mide la capacidad para establecer relaciones entre patrones y figuras abstractas, habilidad para manipular y deducir principios en forma mental, más concretamente mide la habilidad para resolver problemas de tamaño, forma, posición o de cantidad sin emplear un lenguaje numérico o verbal.

Este tipo de habilidad es requerida por el Químico, Físico, Biólogo, Ingeniero, etcétera, y por todas aquellas personas que requieren comprender procesos no visibles.

Se puede agrupar con los tests espacial y mecánico, verbal y numérico.

Relaciones espaciales

Esta formado de 60 modelos que representan cuerpos geométricos los cuales se pueden unir y formar un cuerpo. En las opciones de respuesta se representan 4 cuerpos armados, debiendo el sujeto elegir la figura correcta.

Mide la habilidad para formar un plano tridimensional figuras de objetos sólidos que se representan en un plano bidimensional (planos rasos). Además mide la habilidad de visualidad y manipulación mental de objetos.

Es una habilidad necesaria para áreas como: diseño, arquitectura, ingeniería civil, artes plásticas, decoración.

También en cierta medida el dentista, el maquinista, el carpintero, el cirujano, necesitan el sentido de las formas y posiciones de los objetos en el espacio.

Razonamiento Mecánico

68 dibujos que representan situaciones mecánicas son los elementos formativos del test de Razonamiento Mecánico

Mide la habilidad para entender principios de mecánica y física así como las leyes que rigen los movimientos de herramientas o instrumentos simples.

Esta habilidad es necesaria en ingenieros, mecánicos, técnicos electromecánicos, técnicos en maquinas herramientas, encargados de mantenimiento, ensambladores y otros más que requieren del manejo de los principios comunes de física y mecánica.

Una persona con nivel intelectual alto comprende con facilidad los principios de operación de recursos complejos.

Velocidad y exactitud

El test está formado por 2 partes que contienen series de combinaciones de números y letras. La primera parte tiene como finalidad entrenar al individuo para que pueda ejecutar la segunda parte con mayor facilidad.

Mide la capacidad de percepción y retención momentánea así como la rapidez en la respuesta.

Es necesaria esta habilidad en los trabajos de oficina, almacén, tiendas y en otros empleos similares.

Uso del Lenguaje (ortografía y gramática)

Es una habilidad necesaria para distinguir el uso correcto del lenguaje en actividades como Secretaria, Periodista, Escritor y en cualquier trabajo donde se requiera del lenguaje escrito.

Su inclusión en la batería se funda en la necesidad de representar habi-

lidades básicas para los campos académicos y vocacionales.

Los puntajes de ortografía y gramática se obtienen por separado (aunque es difícil requerir cualquiera de las dos en forma separada) ya que no correlacionan altamente, es decir, un sujeto puede obtener un puntaje alto en ortografía y no necesariamente lo obtendrá también en gramática.

La prueba de ortografía consta de 100 palabras correctas e incorrectas mezcladas entre sí.

La prueba de gramática mide el uso adecuado de la puntuación y de las palabras.

Nota: Esta prueba no existe en castellano.

Material de Test:

7 folletos para la forma A y 7 para la B.

Hojas de respuesta para cada prueba. En la prueba de velocidad y exactitud se requiere de hojas de respuesta diferentes para las formas A y B, en las demás pruebas se emplean las mismas hojas de respuesta.

Claves de calificación (a mano o en máquina).

Se califica a base de plantillas perforadas que se corresponden con las hojas de respuesta. Estas son las que se emplean para la corrección a mano, y para calificar en máquina IBM son necesarias otras plantillas.

Tablas de normas separadas por sexo ya que existen variaciones en las habilidades de hombres y mujeres.

Las normas fueron obtenidas en percentiles y por grado. Estas normas se obtuvieron con 47 000 estudiantes.

Hojas de perfil: sirven para representar el puntaje individual obtenido en cada prueba y para comparar los resultados entre las pruebas y dar una interpretación.

Administración

Se puede aplicar en forma individual o colectiva, es conveniente que en su aplicación se tomen en cuenta las condiciones físicas del local y también el estado físico del examinado. Además se requiere la aplicación de las pruebas en cierto orden dependiendo del número de sesiones para mantener el interés y evitar la monotonía.

Por lo general se recomienda su aplicación en sesiones matutinas o en forma combinada mañana y tarde.¹²

Personalidad

Al parecer todo mundo sabe qué es la personalidad, en nuestras conversaciones diarias solemos hablar de este término, pero en realidad son muy pocas aquellas personas que pueden dar una explicación amplia y exacta de su significado.

Al término personalidad se le suele relacionar con el de "persona" del Latín clásico cuyo significado primitivo era "máscara", también significó el ocultamiento del conjunto de características o cualidades internas verdaderas y personales al relacionarlo con las máscaras que usaban los actores en el teatro.

Si bien es cierto que la personalidad tiene significados diferentes para distintos individuos, también es evidente que es un término amplio y general lo cual dificulta el definirlo con precisión.

Así pues, podemos señalar algunas de las definiciones que al respecto se han formulado.

"La personalidad es una unidad dinámica multiforme" (Stern, 1923), aunque refiere que siempre se tiene por objetivo alcanzar esa unidad y

¹² BENNET K. G., SEASHORE G. H., WESMAN G. A.; *Test de Aptitud Diferencial* (manual); Ed. The Psychological Corporation, New York 1968, pp. 5-9.

que no siempre se logra de una manera perfecta.¹³

“Personalidad es el conjunto de todas las disposiciones, impulsos, tendencias, apetencias e instintos biológicos innatos del individuo, unido a las disposiciones y tendencias adquiridas por experiencias.” (Prince, 1924)¹⁴

“Un esquema unificado de experiencias, una organización de valores consistentes entre sí”. (Lecky, 1954).¹⁵

“Es la configuración de las características y formas de conducta del individuo que determina sus adaptaciones únicas al ambiente propio.” (Hilgard, 1972.)¹⁶

“Personalidad es la organización dinámica en el interior del individuo de los sistemas psicofísicos que determinan su conducta y pensamiento característicos.” (Allport, 1977.)

Allport analiza su definición en sus puntos esenciales y expresa que: “Dentro de la definición de personalidad debe considerarse la organización ya que de esta manera se puede comprender su integración y estructura”.

Dicha organización requiere de la “mente” y del “cuerpo” (de lo interno y de lo externo) en unidad, bajo un sistema que interaccione diversos elementos tales como: costumbres, sentimientos, etcétera, que mueven o dirigen las actividades y pensamientos específicos dándole un matiz personal, que permita la adaptación al medio y en algunos casos su modificación.¹⁷

La personalidad incluye características tales como la honestidad, la jovialidad, la perseverancia, el dominio, la emocionalidad, la adaptación

¹³ ALLPORT Gordon W.: *La Personalidad su configuración y desarrollo*: Ed. Herder; España 1977, p. 41.

¹⁴ ALLPORT Gordon W.: *op. cit.*, p. 45.

¹⁵ ALLPORT Gordon W.: *op. cit.*, 46.

¹⁶ ALLPORT Gordon W.: *op. cit.*, p. 47.

¹⁷ ALLPORT Gordon W.: *op. cit.*, pp. 41-49.

a la vida, la sociabilidad, la relación entre la emoción y la razón y la cooperación.

Es indudable que los rasgos mentales y físicos influyen sobre muchos rasgos de la personalidad, pero las modificaciones de la personalidad surgen debido a que este grupo de rasgos cambia mucho con la experiencia, además los factores mentales y físicos influyen sobre esta experiencia.

Los rasgos de personalidad dependen también de la herencia y de la condición funcional de nuestros cuerpos. La herencia determina las posibilidades de desarrollo de la personalidad, pero la experiencia puede influir sobre su curso dentro de ciertos límites. No se ha logrado determinar, sin embargo, una estimación precisa de la influencia limitadora de la herencia. Se está de acuerdo, generalmente, en que algunas llegan a convertirse en neuróticas o mal adaptadas, a través de la herencia, con más facilidad que otras.

Las funciones de las glándulas endocrinas desempeñan un papel importante en el modo de ser general del hombre y determinan su desarrollo físico y mental. Un funcionamiento inadecuado de la glándula tiroides, por ejemplo, hace que el individuo se cansa fácilmente, le convierte en perezoso e incapaz de concentrarse, mientras que un funcionamiento excesivo de ésta glándula produce inquietud, irritabilidad y angustia.

El funcionamiento extremadamente deficiente de esta glándula produce marcados síntomas físicos e impide tanto el desarrollo mental como el físico. Las glándulas paratiroides tienen una función calmante, de tal forma que su secreción excesiva produce lasitud, mientras que la deficiencia sobreexistencia.

En general las glándulas regulan la química del cuerpo y es preciso que tengan un equilibrio apropiado para que el hombre presente unos rasgos normales de personalidad.

Instrumentos de Medición de la Personalidad

De los instrumentos de medición de la personalidad contamos con los llamados inventarios que son auto-informes objetivos donde el examinado responde a aseveraciones o preguntas concretas generalmente. Estas aseveraciones se refieren a aspectos personales como emociones, necesidades personales, sentimientos, etcétera.

Las técnicas proyectivas son instrumentos que se emplean en la medición de la personalidad. Se presenta una serie de estímulos que el individuo debe interpretar, lo cual implica subjetividad.

Generalmente se aplican en forma individual.

A diferencia de los inventarios que presentan una serie de afirmaciones estructuradas, los instrumentos proyectivos emplean un material indefinido y vago en su contexto. Además, mediante una técnica proyectiva, el examinado no puede falsificar deliberadamente sus respuestas ya que no sabe cómo van a ser evaluadas éstas.

Sus fundamentos están basados en que las respuestas que da el sujeto dependen de sus antecedentes, lo que implica que varían en cada individuo y que cada uno se proyecta a sí mismo, es decir, que estructurarán los estímulos ambiguos según su propia forma de percibir una situación.

Sir. Francis Galton elaboró hace más de medio siglo el primer instrumento proyectivo consistente en una lista de palabras de alto grado emocional, mismas que se presentan al sujeto de una en una y ante las cuales debe responder lo que le venga a la mente, a esta prueba se le denomina de asociación de palabras.

Inventarios de Personalidad

Podemos considerar de entre los inventarios de personalidad los siguientes: TDOT (Dimensiones para el temperamento de Thorndike).

16 PF de Catell, Lista de preferencia personal de Edwards, MMPI (Inventario Multifásico de la Personalidad de Minnesota).

El instrumento de Thorndike es un inventario donde la persona debe describirse a sí misma con respecto a 10 dimensiones del temperamento.

Consta de una hoja de respuestas y un folleto que presenta una serie de aseveraciones divididas en 20 grupos con 10 afirmaciones cada una. Se le pide al sujeto que lea las aseveraciones en forma rápida y que seleccione las 3 más aplicables a su persona, después debe marcar el espacio correspondiente a las inaplicables.

Thorndike (1966) afirma que "la confiabilidad del TDOT es menor que la que se espera habitualmente en los tests de capacidad, pero resulta favorecida cuando se compara con muchos otros inventarios de personalidad".¹⁸

16 PF de Catell

Catell antes de estructurar su cuestionario comenzó a reunir categorías de características y conducta personal, realizó estudios factoriales y formuló las cuestiones del inventario en 16 medidas relativamente independientes, tales como el predominio, la estabilidad emocional, etc (Catell, 1957).

Lista de Preferencia Personal de Edwards PPS.

Este instrumento está formado por 16 motivos o necesidades básicos. Está estructurada de tal forma que se deben seleccionar entre pares de cuestiones consideradas igualmente deseables desde un punto de vista social.

Inventario Multifásico de la Personalidad de Minnesota MMPI.

"A fines de 1958, los Drs. Hathaway y Mckinley comenzaron a reu-

¹⁸ HILGARD E.; *op. cit.*, pp. 56-61.

nir frases para una prueba que crearon convencidos de la necesidad existente de una técnica para evaluar la personalidad desde diferentes ángulos, multifásicamente por medio de escalas o perfiles que fueran tan útiles tanto para la investigación como para la práctica clínica . . .”. “Las frases se basaron en las preguntas que se hacen para elaborar historias clínicas de medicina general, neurológicas y psiquiátricas, como también en escalas publicadas en esa época con fines de encontrar datos sobre la personalidad y escalas que ofrecían informes sobre orientación vocacional”. (Núñez, 1898).¹⁹

Costa de 550 enunciados breves que se refieren en su mayoría a problemas psiquiátricos, también abarca cuestiones religiosas, políticas, educativas, sociales, etcétera.

El sujeto tiene que contestar: “verdadero”, “falso” o “no puedo decir”.

Está estructurada por escalas, 9 de las cuales son consideradas clínicas, una que mide la tendencia a la introversión o a la extraversion social (Si) y 4 de validez.

Escalas Clínicas

1. Hipocondriasis (Hs)
2. Depresión (D)
3. Histeria (Hi)
4. Desviación psicopática (Dp)
5. Masculinidad - feminidad (Mf)
6. Paranoia (Pa)
7. Psicastenia (Pt)

¹⁹ NÚÑEZ R.; *Aplicación del Inventario Multifacético de la Personalidad a la Psicopatología*; Ed.; El Manual Moderno, México 1968, pp. 1-3, 27-56.

8. Esquizofrenia (Es)

9. Hipomanía (Ma)

Hipocondriasis

Busca medir estados de preocupación somática excesiva, es decir, la canalización somática a conflictos psicológicos.

Depresión

Se le suele llamar a esta escala del "estado de ánimo" ya que mide estados emocionales momentáneos y transitorios, que suelen expresarse con baja moral, tristeza y sentimientos de desesperanza.

Histeria

Se formuló para identificar a pacientes con síntomas de histeria de conversión, aunque en sujetos normales puede medir características de la personalidad como: idealista, sentimental, perseverante, afectuoso, etcétera.

Desviación Psicopática

Tiene como objetivo principal medir la aceptabilidad o conformidad con el medio ambiente social: normas éticas, autoridad, relaciones interpersonales, etcétera.

Masculinidad - feminidad

Busca medir intereses que puedan diferenciar a hombres y a mujeres.

es decir, características propias de cada sexo.

Paranoia

Fue diseñada con el fin de encontrar pensamientos o ideas paranoides como: ideas de referencia, ideas de persecución, ideas de grandeza, etcétera.

Psicastenia

Se caracteriza por medir reacciones fóbicas y obsesivo-compulsivas que se manifiestan con dudas excesivas, temores, actos ritualistas, etcétera.

Esquizofrenia

Esta escala es la de las más débiles de la prueba debido a la complejidad que reviste esta enfermedad. Algunos rasgos que se miden son: incongruencia de afecto, fraccionamiento del pensamiento, apatía, indiferencia.

Hipomanía

Fue estructurada para diagnosticar los estados de ánimo de los maníaco-depresivos que se caracterizan por una actividad intensa, insomnio, distracción, actitud de sospecha, etcétera.

Escalas de validez

Las escalas de validez son 4, su objetivo es conocer el grado de confianza de la prueba ya que busca descubrir el descuido, la tendencia a dar respuestas socialmente aceptables o a fingir enfermedad.

Dichas escalas son:

Frasas omitidas

Representa el número total de enunciados que el sujeto no contestó con "Verdadero" o "Falso".

Escala L

Mediante esta escala se puede conocer el grado de veracidad con que el sujeto contestó la prueba.

Escala F y K

Son escalas de validez que miden las actitudes del sujeto al contestar el test, puede ser que asuma actitudes especiales para aparecer como persona inadecuada, incompetente, finge enfermedad, busca evadir responsabilidades (escala F), o trata de dar una buena imagen social, es decir, que se asume una actitud defensiva, (escala K).

Interpretación de la prueba

Se debe anotar primero el número de respuestas omitidas, después hay que obtener las puntuaciones en bruto de cada una de las escalas para convertirlas en puntuaciones T (puntuaciones estándar) con las que se traza el perfil. Una vez elaborado éste debe ser interpretado en su configuración total, es decir, que aunque la escala de mayor elevación sirve de referencia debe buscarse su relación con las demás escalas. También es importante considerar la edad y el sexo de la persona estudiada.

Tests proyectivos

Existen varios tests proyectivos, entre los más utilizados se encuentran el TAT (Test de Apercepción Temática), el Rorschach, el Machover y el Bender.

A continuación mencionaremos los aspectos más importantes de cada uno de ellos.

TAT (Test de apercepción temática)

Fue construida esta prueba por H.A. Murray y C.D. Morgan en 1935, consta de 19 cuadros para una edad y un sexo particulares y una tarjeta en blanco. Los cuadros presentan figuras humanas en diferentes posturas y actividades. El sujeto debe hacer un relato acerca de cada cuadro.

Al interpretar se trata de obtener un cuadro completo con todas las tarjetas administradas, para así obtener las áreas de conflicto: madre-hijo, identidad sexual, hostilidad, etcétera.

El supuesto básico de esta prueba radica en que el examinado expresará sus necesidades y problemas de personalidad proyectándolos en el héroe o heroína del relato. (Mc. Connell,) 1978).

Roschach o test de manchas de tinta

Consta de 10 tarjetas, cada una con una mancha diferente, 5 en blanco y negro y 5 en color. El sujeto expresa lo que ve cuando le muestran una a una las tarjetas.

Esta prueba fue construida considerando la tendencia de "ver" rostros imaginarios, animales, escenas de guerra o figuras de hadas en las formaciones de nubes.²⁰

Existen varias formas de calificar la prueba, ya sea siguiendo un método cuantitativo o cualitativo, o ambos.

Su interpretación es un proceso difícil que requiere de un adiestra-

²⁰ HILGARD E.; *op. cit.*, pp. 137-149.

miento riguroso y de experiencia clínica.

Este complicado pero valioso instrumento para la medición de la personalidad fue creado por el psiquiatra suizo Hermann Rorschach en 1921.

Test de la figura humana de karen machover

Al igual que muchos otros, Karen Machover cree que el dibujo de la figura humana está vinculado con la personalidad del individuo que hace el dibujo. Afirma que: "el cuerpo o el yo es el punto de referencia más íntimo de cualquier actividad. En el curso del crecimiento llegamos a asociar diversas sensaciones, percepciones y emociones con ciertos órganos del cuerpo... Consecuentemente, el dibujo de una persona, al implicar la proyección de la imagen del cuerpo, proporciona un vehículo natural de la expresión de las necesidades y conflictos del propio cuerpo...²¹

Aplicación

Se emplea un papel blanco sin rayas de aproximadamente 8.5 por 11 cms. y un lápiz con una goma de borrar.

Se le dice: "Dibuje una persona", si dibujó un hombre se le pide que dibuje a una mujer, y viceversa.

Cuando los sujetos se resisten a hacer el dibujo se les explica que la belleza del dibujo no importa. También se les suele decir: haga una historia o un cuento acerca de esta persona, si el examinado se resiste se le incita mediante algunas preguntas como: ¿es fuerte?, ¿es bien parecido?, ¿es casado o soltero?

Para su interpretación se analizan diferentes rasgos del dibujo como

²¹ PORTUONDO Juan A.; *Test Proyectivo de Karen Machover*; Ed. Pax; Madrid 1979, pp. 11-12.

son: la cabeza, la cara, los labios, los ojos, brazos y manos, las piernas y los pies, ropas, etcétera.

Test gúestáltico visomotor de bender

Lauretta Bender construyó entre los años 1932 - 1938 su test gúestáltico visomotor.

El test se funda en la psicología de la "gestalt" que sostiene que "las unidades organizadas o las configuraciones estructuradas constituyen las formas primarias de las reacciones biológicas, al menos al nivel psicológico de la conducta animal, y que, en el campo sensorial estas unidades organizadas o "gestalten" corresponden a configuraciones del mundo exterior . . . Hay una tendencia a percibir las "gestalten" de acuerdo con principios biológicos determinados por el patrón sensomotor de acción. Cabe esperar que éste varíe en los diferentes niveles de maduración o crecimiento y en los estados patológicos orgánicos o funcionalmente determinados." ²²

En sentido estricto es un test visomotor ya que la tarea que le solicita al sujeto es gráfica con los modelos a la vista, pese a llamarse "gúestáltico" se le suele designar proyectivo debido a que se emplea la asociación libre sobre las figuras dibujadas.

El test está formado por 9 diseños originales de Wertheimer y se le pide al sujeto que copie las figuras, estas se analizan y se evalúan las reproducciones según como el sujeto haya estructurado esos estímulos perceptuales.

Este test tiene numerosas aplicaciones tanto psicológicas como psiquiátricas. Así pues, es un instrumento valioso en la exploración del desarrollo de la inteligencia infantil y en el diagnóstico de los diversos síndromes clínicos de eficiencia mental, desórdenes cerebrales orgánicos, etcétera.

²² BENDER Lauretta; ; *Test Gúestáltico Visomotor*; Ed. Paidós, Buenos Aires 1979, pp. 11-26.

CAPITULO IV
LA MEDICION DEL CRITERIO
Y LA UTILIDAD DEL PREDICTOR

La medición del criterio y la utilidad de predictor

El criterio es fundamental para todas las mediciones de la psicología del trabajo. La magnitud de la contribución de la psicología se ve determinada por lo adecuadas que sean las medidas de criterio obtenidas.

El criterio se define como "una medición del desempeño de un trabajador". En términos generales, el criterio es una norma de evaluación para medir el desempeño de una persona, sus actitudes, sus motivos, etcétera. En el campo de la psicología industrial pueden encontrarse otras definiciones del criterio, así por ejemplo, Guion (1965) la define como "lo que debe predecirse".¹

Tipos de criterios de acuerdo a la visión clásica

El criterio es una medida del grado de desempeño, de ejecución de una u otra forma. La ejecución no es independiente del tiempo.

Una persona que en la actualidad haya alcanzado cierto grado de desempeño de acuerdo con alguna norma, no quiere decir que esta misma persona alcanzará ese mismo grado de ejecución, de acuerdo con la misma norma, en un tiempo futuro.

Además, las normas pueden cambiar y con frecuencia cambian con el tiempo.

Los criterios que se obtienen acerca del momento en que se utilizan los instrumentos de predicción se denominan criterios próximos, los obtenidos en alguna otra fecha se conocen como criterios distantes. Los criterios que se obtienen al mismo tiempo que se administra el instrumento de predicción se llaman criterios inmediatos.

Los criterios obtenidos un año o más después de obtener las medidas de predicción puede considerarse como distante.

Criterios reales: cualquier medida del éxito que se utilice es el cri-

¹ BLUM M., NAYLOR J.; *Psicología Industrial*; Ed. Trillas; México 1976, p. 252.

terio real.

El criterio final es un criterio teórico o ideal que por lo general no se da en la realidad.

Los criterios finales tienen características distantes, sin embargo, deben considerarse inmediatos.

Conceptos principales del criterio

La deficiencia, la pertinencia y la contaminación son conceptos relativos al criterio.

Deficiencia del criterio es el punto hasta el cual nuestro criterio carece de cierta varianza necesaria para el criterio final.

Pertinencia del criterio es el punto hasta el cual el criterio real coincide con el final o corresponde a él. No podemos obtener una medida empírica de la pertinencia de un criterio puesto que es raro que podamos medir verdaderamente el criterio final.

Contaminación del criterio es la varianza en el criterio real que no tiene relación con el final. Puede considerarse que consta de dos partes distintas: primero la contaminación sistemática del criterio que puede considerarse como la distorsión de un criterio; y segundo, como la contaminación aleatoria que puede decirse que es el error del criterio.

Tanto la deficiencia como la contaminación son características indeseables de un criterio.

Extensión del criterio y de la población

La extensión del criterio se refiere a la validación de la prueba en función de un nuevo criterio en otra muestra de la misma población, así la eficiencia de una prueba "se extiende" a un nuevo criterio.

La extensión de la población se refiere a la situación en la cual el criterio sigue siendo el mismo, pero la muestra se toma de una población

diferente. Así la eficiencia de la prueba puede extenderse a un nuevo grupo de personas.

Equivalencia de criterios

Puede decirse que dos criterios son equivalentes cuando son completamente intercambiables sin que se produzca ninguna pérdida de la eficiencia de predicción del proceso de selección.

Distorsión de oportunidad

Este tipo de distorsión del criterio se refiere a aquellas situaciones en que hay factores que se encuentran fuera del control del trabajador e influyen considerablemente en la cantidad producida.

Otra fuente importante de distorsión del criterio se refiere a las características del grupo al que pertenece una persona.

Si una persona pertenece a un grupo que limita artificialmente la productividad de sus miembros, habrá una clara distorsión de los datos de criterio.

Así mismo, si una persona pertenece a un grupo que ha sido preseleccionado de acuerdo con alguna base a priori y si a continuación se compara ese grupo con algún otro no seleccionado en esa forma, cualquier variable a priori podrá tener una validez falsa.

El problema de la edad y la duración en el empleo es otra muestra de la distorsión de oportunidad.

Distorsiones en las calificaciones

Uno de los criterios empleados con mayor frecuencia en la industria, las calificaciones, están también sujetas a distorsiones.

El efecto de desviación que se cita con mayor frecuencia en las calificaciones es el fenómeno conocido como "Efecto de halo", esto se refiere

a aquella situación en que un supervisor considera que una persona es sobresaliente en todos los rasgos simplemente porque posee una característica relevante.

Conocimiento de las distorsiones del predictor

Si una persona conoce los resultados obtenidos por las personas en la variable del predictor, es muy probable que deje que sus juicios de criterio se vean afectados por ese conocimiento. Ese error puede causar un incremento de naturaleza totalmente impura en la validez obtenida.

Requisitos de los criterios

Weitz (1961) señala que hay 3 cuestiones principales relativas a los criterios: tiempo, tipo y nivel.

Tiempo ¿cuándo debemos obtener nuestros datos de criterio?

Tipo De entre todas las medidas posibles del desempeño, ¿cuál o cuáles deben seleccionarse?

Nivel ¿Qué nivel de desempeño se considerará aceptable?

¿Cuántas unidades debe producir una persona para que se le pueda considerar como un buen trabajador?

Bellows (1954) da una lista de características que son necesarias y/o convenientes en cualquier criterio: ²

² BLUM M., NAYLOR J.; *op. cit.*, p. 262.

1. Confiable
2. Realista
3. Representativo
4. Relacionado con otros criterios
5. Aceptable para el analista de puestos
6. Aceptable para la gerencia
7. Consistente de una situación a otra
8. Predecible

Se sugiere una lista de otros conceptos, como:

9. No costoso
10. Inteligible
11. Mensurable
12. Pertinente
13. No contaminado y libre de distorsiones
14. Diferenciador

Guion (1961) estableció las etapas apropiadas en el desarrollo de un criterio: ³

1. Analiza el puesto y/o las necesidades de la organización por medio de técnicas nuevas que todavía deben investigarse.
2. Desarrolla medidas de la conducta real en relación a la conducta esperada, tal como se identifican en el análisis del puesto y las necesidades. Esas medidas deben suplementarse a las medidas de

³BLUM M., NAYLOR J.; *op. cit.*, pp. 263-264.

las consecuencias del trabajo.

3. Identifica las dimensiones de criterio que se encuentran implícitas en esas medidas por medio del análisis factorial, del análisis de grupos o el de esquemas.
4. Desarrolla medidas confiables, cada una de ellas con una alta validez de construcción de los elementos identificados en esa forma.
5. Determina la validez predictiva de cada variable independiente (predictor) para cada una de las medidas anteriores de criterio tomándolas de una en una.

Criterios compuestos

A principios de la historia de la selección, el mejor criterio era una etiqueta colocada en la medida del éxito en el trabajo que el psicólogo consideraba más satisfactoria, de acuerdo con el conjunto de criterios que tenía a la mano.

Esta visión fue desapareciendo gradualmente al aparecer el concepto más complejo, el criterio compuesto, este punto de vista sostiene que el éxito real en el trabajo es de naturaleza multidimensional.

Al aceptarse algunas medidas compuestas del éxito como método correcto para desarrollar criterios es preciso establecer un índice compuesto de evaluación. Hay numerosos procedimientos alternativos.

Nagle (1953) subrayó esas diversas alternativas: ⁴

1. *Ponderación de sub-criterios, utilizando juicios de "expertos".*
Este procedimiento implica hacer que varios expertos examinen los criterios separadamente y asignen pesos relativos a cada uno de ellos de acuerdo con sus propias percepciones de la importancia

⁴ BLUM M., NAYLOR J.; *op. cit.*, pp. 264-266.

respectiva de las medidas individuales.

Ejemplo:

C 1 = número de unidades producidas por unidad de tiempo.

C 2 = calidad promedio de las unidades producidas.

C 3 = número de accidentes por unidad de tiempo.

C 4 = calificación dada por el supervisor inmediato, y

C 5 = número de retrasos a la hora de entrada al trabajo.

La ecuación básica para el criterio compuesto:

$$C = W_1C_1 + W_2C_2 + W_3C_3 + W_4C_4 + W_5C_5$$

La tarea de cada experto consiste en decidir los valores que deben asignar a W_1 , W_2 , W_3 , W_4 y W_5 .

La selección de "expertos" es un problema importante al utilizar este procedimiento, con frecuencia hay pocos individuos disponibles que pueden considerarse como suficientemente competentes y neutrales para confiarles la tarea de asignar pesos a los subcriterios.

2. *Ponderación de subcriterios proporcional a sus confiabilidades.*

Este procedimiento de más importancia a las medidas individuales del desempeño que tienen mayor confiabilidad.

3. *Ponderación de subcriterios proporcional a su correlación promedio con otras variables.*

Se basa en un proceso puramente estadístico y no es una técnica muy defendible.

4. *Ponderación de subcriterios, de acuerdo a su predicabilidad.*

Este método se ha denominado procedimiento "de criterio más predicable" (Hatelling, 1955). Se ponderan deliberadamente los subcriterios a fin de obtener la máxima correlación entre el predictor y el criterio compuesto.

5. *Selección y ponderación de subcriterios por medio del análisis factorial.*

El análisis factorial es un método que agrupa variables en conjunto homogéneos de acuerdo a sus intercorrelaciones.

Ha sido un método utilizado por muchos investigadores para determinar cuáles deben ser las dimensiones básicas del criterio compuesto.

6. *Ponderación de subcriterios para elevar al máximo la diferencia entre individuos.*

Este método trata de hacer aumentar la gama de los resultados de criterio tanto como sea posible, haciendo que en esa forma resulte más fácil de diferenciar de acuerdo con la dimensión del criterio.

Estadísticamente, el método implica que se dé mayor importancia a los subcriterios que tengan intercorrelaciones altas o que ponderen cualquier factor general básico en los subcriterios.

Probablemente sólo tendrá valor cuando el éxito en el trabajo se componga de un único factor básico.

7. *Ponderación igual a los subcriterios.*

Un modo de evitar el problema de decir qué subcriterios deben recibir el mayor peso en el criterio compuesto es darles la misma importancia a todos los subcriterios.

Aunque indudablemente esa suposición será siempre incorrecta, puede sostenerse que a la larga, sólo se puede tratar de adivinar

los pesos correctos, cometeremos menos errores si les atribuimos a todas las variables la misma importancia.

8. *Ponderación en base al costo de dinero.*

Brodgen y Taylor (1950) abogaron por el "criterio del dinero" para resolver el problema de la ponderación al combinar varios aspectos del éxito en el trabajo en una medida compuesta de criterio.

Ante todo, Brodgen y Taylor sugieren que el criterio final para cualquier organización lucrativa debe expresarse en dinero. Así todos los subcriterios deben transformarse en valores en dinero. Hay 2 grandes dificultades para el uso de este tipo de criterio: en primer lugar, la tremenda dificultad de convertir todas las medidas del éxito en el trabajo a una base monetaria. La segunda dificultad es que algunas medidas del éxito en el trabajo no pueden expresarse en términos monetarios.

La ventaja del método es para la obtención de una medida de criterio compuesto y se limita probablemente a aquellas situaciones en que el trabajo es bastante básico y los subcriterios pueden relacionarse fácilmente con los beneficios para la compañía en unidades monetarias.

Criterios múltiples

Ghiselli (1956) fue probablemente el primer investigador que se rebeló contra los criterios compuestos, Abogó firmemente por el tratamiento separado de cada una de las dimensiones del éxito en el trabajo.

Ghiselli señaló que, por lo común, el éxito no sólo es un concepto multidimensional, sino que hay por lo menos 5 tipos diferentes de multidimensionalidad de los criterios.⁵

⁵ BLUM M., NAYLOR J.; *op. cit.*, p. 271.

Los denominó “dimensionalidad estática”, “dimensionalidad dinámica” y “dimensionalidad individual”.

Dimensión Estática de Criterios

Si se considera el problema del desarrollo de criterios en cualquier punto del tiempo, casi siempre el éxito en el trabajo es multidimensional, o sea, que hay varias formas distintas de definir el éxito en ese trabajo y en ese momento.

Dimensiones Dinámicas de Criterios

Este criterio se ocupa adecuadamente del hecho de que el desempeño de los individuos en las diversas dimensiones cambiarán con el tiempo, a medida que aquellos vayan aprendiendo y desarrollando en el trabajo.

Si las correlaciones entre dimensiones cambian en función de la práctica o la experiencia, esto indicará que también cambian las dimensiones básicas.

Dimensiones Individuales de Criterios

Ghiselli sugiere que varios trabajadores en el mismo empleo pueden considerarse como igualmente buenos, no obstante, la naturaleza de sus contribuciones a su organización puede ser muy diferente.

Sostiene que quienes trabajan en el mismo puesto pueden tener que evaluarse de acuerdo con distintas dimensiones de criterios.

El problema de los criterios compuestos vs. los múltiples tiene una gran variedad de facetas diferentes y ambas posturas poseen cierto valor, aunque su rechazo parece derivarse de una de dos fuentes importantes o de ambas: ponderación y la necesidad de progresar.

La primera objeción al criterio compuesto es la ponderación de sub-criterios para constituir una medida general del éxito en el trabajo. Al-

gunos sostienen que sólo se pueden combinar legítimamente las mediciones cuando se refieren a una misma cosa.

De acuerdo al estudio de Naylor y Wherry sus resultados indican claramente que el problema de la ponderación no es quizá la principal objeción para el uso del criterio compuesto, dado que es posible combinar de manera confiable aspectos independientes del éxito en el trabajo en un criterio compuesto.

Una segunda objeción importante es que ha habido muy pocos progresos en la eficiencia de las técnicas de selección al correr de los años.

Wallace sugiere que sólo al examinar de manera individual todos los tipos de subcriterios para lograr una mayor comprensión, independientemente de la pertinencia de esos subcriterios, logramos ver progresos en el desarrollo de mejores fundamentos para la selección y colocación.

Utilidad de los Instrumentos de Medición

La utilidad de un dispositivo de predicción es el punto hasta el cual su empleo hace mejorar la calidad de las personas seleccionadas por encima, de lo que hubiera sido posible, en caso de que no se utilizara ese dispositivo.

La validez y la confiabilidad son características que desempeñan un papel primordial en la determinación de la utilidad de cualquier instrumento de selección. Además, existen otros factores que resultan igualmente críticos en cualquier situación que implique la selección de grupos. Estas variables son: a) confiabilidad de criterio, b) pertinencia de criterio, c) proporción de selección, y d) porcentaje de los empleados actuales a los que considera eficientes.

La validez de un instrumento de predicción es un índice estadístico primario que influye en su utilidad.

Como primer principio general en la utilidad de los tests instrumentos de predicción, tenemos que, dado cualquier punto de corte arbitraria-

mente definido en una prueba, cuanto mayor sea la validez tanto mayor será el incremento del puntaje promedio en el criterio para el grupo seleccionado, por encima del observado para el grupo total:

$$(X \text{ grupo seleccionado}) - (X \text{ grupo total})$$

Otras dos variables que desempeñan un papel primordial en la determinación de la utilidad de un predictor son la proporción de selección y el porcentaje de empleados actuales a los que se considera eficientes.

En párrafos anteriores se definió la utilidad de un predictor como el mejoramiento en la calidad de las contrataciones realizadas mediante el empleo de un dispositivo de predicción.

La calidad se define en términos de: a) resultado promedio del criterio del grupo o b) de acuerdo a la proporción de personas de ese grupo que tienen resultados en el criterio por encima de algún valor que considera como mínimo para que una persona pueda ser eficiente como empleado.

El manejo de la proporción de selección y/o un cambio en el porcentaje de empleados actuales a los que se considera eficientes dará cambios en la calidad de los empleados contratados.

La proporción o razón de selección se suele expresar de la siguiente manera: $n/N = RS$

donde:

n = número de vacantes de un puesto

N = número de solicitantes de empleo disponible para colocación

Cuando la RS es igual o mayor que 1.00 el uso de cualquier dispositivo de selección tendrá poco sentido.

Si la RS es menor que 1.00 habrá más solicitudes de empleo que vacantes, por lo cual podrá haber selección del personal que se quiera contratar.

Un segundo principio general se refiere a que una baja razón de selección dará siempre como resultado empleados de mejor calidad en tanto la relación entre el predictor y el criterio tenga algún valor superior a 0.

Porcentaje de empleados actuales que resultan eficientes

En el análisis de la validez y de la RS se ha dado por sentado que el criterio es continuo y que, por ende, mientras más alto sea el puntaje en el criterio más satisfactorio se considerará al empleado.

Para cualquier validez y cualquier punto de corte de criterio una reducción de la RS dará como consecuencia un aumento de la validez efectiva.

Para cualquier validez estadística y RS dada, mientras menor sea el porcentaje de empleados actuales a los que se considera como satisfactorios mayor será el aumento del porcentaje de los candidatos satisfactorios obtenidos mediante la utilización del predictor.

$$\text{Unidad} = \frac{A}{A + D} - \frac{A + B}{A + B + C + D} = \text{porcentaje de aumento de la efectividad.}$$

A = son aquellos sujetos que en la prueba indica deben tener éxito.

B = son aquellos que la prueba indica que no tendrán éxito.

C = aquellos que no se contratarían si la selección se basara en la prueba.

D = representa a solicitantes de empleo a los que se contrataría pero posteriormente no resultarían satisfactorios en su trabajo.

Taylor y Russell (1959) prepararon una tabla detallada de la relación exacta entre el tamaño del coeficiente de validez, la razón de selección y el porcentaje de empleados actuales satisfactorios. En condicio-

nes dadas de validez, razón de selección y porcentaje de empleados satisfactorios sus tablas permiten determinar el porcentaje de contrataciones que serán satisfactorias, utilizando el predictor al mismo tiempo que los métodos comunes.⁶

La tabla de Naylor-Shine parece tener ciertas ventajas sobre las de Taylor/Russell: la primera utiliza diferencias en el porcentaje de éxito entre el grupo escogido el original en el puntaje promedio del criterio, mientras las segundas utilizan diferencias en el porcentaje de éxito entre el grupo escogido y el original, además requiere que los empleados se separen en dos grupos: "eficientes" e "ineficientes".

Sin embargo ambas tienen limitaciones muy importantes: se basan en la suposición de que: a) la relación entre el predictor y el criterio es lineal y b) el coeficiente de validez utilizado se obtiene por procedimientos de validez concurrente.

Confiabilidad del Predictor y del Criterio

La confiabilidad del predictor y la del criterio son también importantes ya que establecen límites sobre el tamaño del coeficiente de validez que puede obtenerse.

Hay una relación algebraica básica entre la validez y la confiabilidad del predictor y del criterio, ésta es:

$$rpc \text{ (obtenido)} = rpc \text{ (real)} rpp \times rcc$$

en donde: $rpc \text{ (obtenido)}$ = correlación observada (validez) entre el predictor y el criterio.

$rpc \text{ (real)}$ = correlación real (validez) entre el predictor y el criterio.

rpp = confiabilidad del predictor, y

rcc = confiabilidad del criterio.

⁶ BLUM M., NAYLOR J.; *op. cit.*, p. 81.

Teoría de las decisiones

El punto de partida es el reducir al mínimo el número de errores cometidos en el proceso de decidir si una persona debe ser contratada o rechazada, estableciendo un punto de corte adecuado.

Parece ser que el punto de corte que reduce al mínimo el error total es el punto en que se interceptan las distribuciones.

$$\text{Error total} = A + B + C$$

Rorer, Hoffman, La Forge y Hsieh (1966) señalan ciertos aspectos interesantes y muy importantes del problema general de predicción sobre la relación de la validez de la prueba con su utilidad.

1. Tanto las medidas como las varianzas de los dos grupos tienen diferencias entre sí.
2. Los grupos tienen iguales medidas pero distintas varianzas.
3. Las medidas de los grupos son considerablemente distintas pero los tamaños de los grupos son también muy diferentes.

CAPITULO V

DESCRIPCION DEL ESTUDIO REALIZADO

Planteamiento del problema

Es sabido que las pruebas o instrumentos de evaluación psicológica son en una gran proporción tests elaborados en otros países para medir atributos de conducta de la población en la cual se construyeron, lo cual trae limitaciones para su aplicación directa en otros países cuya población tiene diferencias marcadas en aspectos culturales, sociales, políticos, económicos, etcétera.

Por lo anteriormente expuesto y habiendo analizado la importancia de contar con instrumentos de evaluación válidos y confiables para nuestra población, consideramos que es importante abocarse a la tan importante tarea de elaborar instrumentos de medición propios o de adaptar aquellos que por sus características pueden resultar útiles para nosotros.

Por ello consideramos conveniente obtener las normas percentilares de la prueba de Dominós aplicada a alumnos del CECyT No. 29 de Jiquilpan, Mich. ya que las escalas utilizadas se han estandarizado en otros países por lo cual resultan inapropiadas para los estudiantes mexicanos.

Objetivo

Obtener las normas generales de la prueba de Dominós para la población del CECyT No. 29 ubicados en el estado de Michoacán. La obtención de las correlaciones entre los puntajes obtenidos en la prueba de inteligencia de Dominós y las calificaciones obtenidas en las materias siguientes: Inglés I y II, Filosofía I y II y Matemáticas I y II, ya que estas materias son comunes a todas las especialidades en los dos primeros semestres; así como establecer el coeficiente de validez y de confiabilidad de la prueba de Dominós en la población estudiada.

Para el caso contemplamos las siguientes tareas:

1. Análisis de Reactivos.

2. Determinación de la confiabilidad de la prueba.
3. Obtención de las normas percentilares para su calificación.
4. Determinación de la validez de la prueba.

Sujetos

Se aplicó la prueba de Anstey (dominós) a 500 aspirantes al CECyT No. 29 de la cual se tomó una muestra al azar de 200 alumnos en la cual están representadas las diferentes especialidades que los alumnos solicitaban en su ingreso (técnico laboratorista clínico, técnico en contabilidad, técnico electromecánico y técnico topógrafo).

En la muestra de alumnos se pueden observar diferentes niveles socio-económicos ya que a este centro concurren individuos de diferentes poblaciones que forman la zona de influencia del CECyT No. 29, aunque en todos los casos un mismo antecedente escolar, secundaria, que es requisito para ingresar a estos centros.

Sus edades fluctúan entre los 16 y los 27 años de edad.

El total de sujetos de sexo femenino es 52 y 148 de sexo masculino.

Escenario

El CECyT es una institución educativa a nivel medio superior (Bachillerato) que imparte capacitación técnica en 3 áreas diferentes: Área de Ciencias Químico Biológicas, Área de Ciencias Sociales Administrativas y Área de Ciencias Físico-Matemáticas.

Esta Institución cuenta dentro del Área de Ciencias Químico Biológicas con la especialidad de Técnico Laboratorista Clínico, en la de Ciencias Sociales Administrativas con la de Técnico en Contabilidad, y con las especialidades de Técnico Topógrafo y Técnico Electromecánico en el Área de Ciencias Físico-Matemáticas.

El plan de estudios de las diferentes especialidades se cubre general-

mente en 6 semestres, aunque puede cubrirse en 5 como mínimo y 8 como máximo.

Actualmente la Institución cuenta con 800 alumnos repartidos en especialidades y grados, en la de Técnico Laboratorista Clínico son 280 alumnos, en la de Técnico en Contabilidad, 240 alumnos, en la de Técnico Topógrafo, 135 alumnos, y en la de Técnico Electromecánico, 145 alumnos.

Se trabaja con el sistema de programación por objetivos.

Instrumento utilizado

Se empleó la prueba de Dominós que consta de 48 reactivos de orden de dificultad creciente en el que hay que completar series numéricas.

PROCEDIMIENTO

Aplicación y Calificación de la Prueba

Se aplicó el instrumento en forma colectiva a grupos de 25 alumnos en las aulas del propio CECyT, seleccionando aquellas que reunieran las condiciones ambientales necesarias para una situación de examen, como son iluminación adecuada, alejadas del ruido, etcétera.

La aplicación en todos los casos se realizó en el período que antecede a la iniciación de cursos (Agosto) y en sesiones matutinas de 8 y 10 de la mañana.

Se dieron las instrucciones marcadas por el manual de Dominós y se marcó con tiempo de ejecución 30 minutos.

Las calificaciones de la prueba se obtuvieron conforme a las instrucciones marcadas en el manual del test, así obtuvimos un puntaje crudo que corresponde al número de items correctos, su percentil y rango correspondientes.

Ordenación de puntajes

Una vez calificadas las pruebas se ordenaron de acuerdo al número de respuestas correctas de menor a mayor.

Se aplicó el método abreviado en el cual se emplea el 27% de sujetos de cada extremo de la distribución de los puntajes totales, en este caso 54 sujetos ya que la muestra tomada fue de 200.

El grupo superior se formó con los 54 puntajes más altos y el inferior con 54 puntajes más bajos.

Determinación de la frecuencia de estudiantes en cada grupo que contestaron correctamente cada ítem.

Se obtuvieron las frecuencias de aciertos para cada ítem tanto en el grupo superior como en el inferior, posteriormente se consultaron las tablas de Nuttal y Skuik (1971) para determinar el porcentaje de aciertos y conocer el grado de dificultad de la prueba.

También se obtuvo el coeficiente de correlación biserial puntual con la fórmula:

$$r_{pbis} = \frac{M_p - M_y P}{S_y y}$$

Después se obtuvo el producto de los porcentajes de aciertos y errores de cada uno de los reactivos, obteniéndose la tabla de valores respectiva.

Así mismo se determinó el coeficiente de consistencia interna de Kuder-Richardson (fórmula KR-20).

$$r = \left(\frac{n}{n-1} \right) \frac{\sigma^2 - \sum pq}{\sigma^2 T}$$

También se obtuvo el coeficiente de correlación de Spearman Brawn en las materias de Inglés, Filosofía y Matemáticas de primero y segundo semestres.

$$r = \frac{N \Sigma x y - (\Sigma x) (\Sigma y)}{\sqrt{[N \Sigma x^2 - (\Sigma x)^2] [N \Sigma y^2 - (\Sigma y)^2]}}$$

Determinación de los valores percentilares

Se ordenaron los percentiles de 5 en 5 hasta el percentil 95 de acuerdo con el número de sujetos que obtuvieron el mismo puntaje en la prueba de Anstey (Dominós).

RESULTADOS

Los resultados obtenidos los damos a continuación:

En la tabla No. 1 se presenta la distribución de frecuencias de puntajes.

En la tabla No. 2 se encuentran las normas percentilares para la población.

En la tabla No. 3 se presentan las frecuencias de aciertos para cada reactivo de los grupos superior e inferior y la proporción total de aciertos, así como el coeficiente de correlación biserial.

Valores obtenidos en la media, la desviación estándar, el coeficiente de confiabilidad y el error estándar.

Media = 19.96

Desviación Estándar = 8.14

DISTRIBUCION DE FRECUENCIAS DE PUNTAJES

TABLA NUMERO 1

<i>Puntaje</i>	<i>Frecuencia</i>
1	1
3	2
4	4
5	3
6	3
7	3
8	8
9	3
10	3
11	6
12	1
13	5
14	2
15	4
16	4
17	1
18	6
19	7
20	5

<i>Puntaje</i>	<i>Frecuencia</i>
21	5
22	13
23	9
24	15
25	15
26	12
27	9
28	12
29	7
30	16
31	5
32	5
33	4
34	1
36	2
37	1

TABLA DE VALORES PERCENTILARES
TABLA NUMERO 2

<i>Percentil</i>	<i>Puntaje</i>
5	5
10	8
15	11
20	15
25	16
30	18
35	20
40	21
45	22
50	23
55	24
60	25
65	26
70	27
75	28
80	29
85	30
90	32
95	36

**TABLA DE PROPORCIONES DE ACIERTOS Y
COEFICIENTES DE CORRELACIONES BISERIAL
PARA CADA REACTIVO**

TABLA NUMERO 5

<i>Item</i>	<i>Grupo Sup.</i>	<i>Grupo Inf.</i>	<i>F.</i>	<i>r</i>	<i>pq</i>
1	100	95	97	.27	.05
2	98	72	85	.53	.13
3	96	52	74	.59	.19
4	96	46	71	.63	.21
5	96	48	72	.62	.20
6	95	37	66	.66	.22
7	100	87	95	.45	.07
8	95	37	66	.66	.22
9	93	50	62	.66	.24
10	96	41	69	.66	.21
11	98	28	63	.78	.23
12	91	33	62	.61	.24
13	98	24	61	.79	.24
14	98	54	76	.65	.18
15	87	11	49	.73	.25
16	96	32	64	.71	.23
17	96	46	71	.63	.21

<i>Item</i>	<i>Grupo Sup.</i>	<i>Grupo Inf.</i>	<i>F.</i>	<i>r</i>	<i>pq</i>
18	78	17	48	.60	.25
19	96	52	74	.59	.19
20	85	33	59	.53	.24
21	98	26	62	.79	.24
22	87	11	49	.73	.25
23	89	6	48	.80	.25
24	89	17	53	.70	.25
25	85	11	48	.72	.25
26	96	28	62	.73	.24
27	72	4	38	.73	.24
28	46	2	24	.70	.18
29	55	6	21	.44	.17
30	0	0	01	.00	.01
31	52	2	17	.56	.14
32	80	7	44	.73	.25
33	61	0	31	.77	.21
34	28	0	15	.60	.13
35	41	0	21	.68	.17
36	52	2	27	.68	.20
37	24	2	15	.49	.11
38	11	0	06	.41	.06

<i>Item</i>	<i>Grupo Sup.</i>	<i>Grupo Inf.</i>	<i>F.</i>	<i>r</i>	<i>pq</i>
39	0	0	01	.00	.01
40	2	0	02	.11	.02
41	24	0	15	.57	.11
42	15	0	08	.47	.07
43	4	0	05	.25	.05
44	22	0	12	.55	.11
45	7	0	04	.33	.04
46	2	0	02	.11	.02
47	2	0	02	.11	.02
48	2	0	02	.11	.02
	29.94	9.99			7.78
	D	E			F

Coefficiente de confiabilidad de Kuder-Richardson (KR-20) $r = .90$.

Error estándar de medida = 2.52

Finalmente presentamos los valores obtenidos en el coeficiente de correlación de Pearson en 5 diferentes materias que se cursan en el primero en el segundo semestre con las puntuaciones en la prueba de Dominós. Estas correlaciones son significativamente mayores que cero ($p = .01$)

Inglés I	$r = 0.23$
Inglés II	$r = 0.52$
Filosofía I	$r = 0.52$
Filosofía II	$r = 0.37$
Matemáticas I	$r = 0.50$
Matemáticas II	$r = 0.23$

CONCLUSIONES

Esta investigación demostró que la prueba de Dominós es muy confiable en las condiciones empleadas.

El coeficiente de confiabilidad de Kuder-Richardson fue de .90.

Los coeficientes de validez para estimar las calificaciones escolares en Filosofía, Inglés y Matemáticas fueron las siguientes: Filosofía I $r = 0.52$, Filosofía II $r = 0.57$, Inglés I $r = 0.25$, Inglés II $r = 0.29$, Matemáticas I $r = 0.50$, Matemáticas II $r = 0.25$.

La tabla de valores percentilares obtenidas puede emplearse para determinar la posición que guarda cualquier estudiante del CECyT con respecto a la muestra estudiada.

La medida de las calificaciones en la prueba fue 19.96, lo cual indica que fue relativamente difícil para estos estudiantes. Los items fáciles, contestados correctamente por más del 80% de los alumnos, fueron: 1, 2, 7.

Los items difíciles contestados incorrectamente por menos del 20% de alumnos fueron: el 50, 51, 54, 57, 58, 59, 40, 41, 42, 43, 44, 45, 46, 47, y 48.

Los items cuyo poder discriminativo fue deficiente dado que su coeficiente de correlación biserial fue significativamente menor que 0, fueron los siguientes: 50, 59, 40, 46, 47, 48.

Tanto los items muy fáciles como los muy difíciles, aún los que tienen poco poder discriminativo, deberán reemplazarse por otros items especialmente elaborados para incrementar el coeficiente de confiabilidad de la prueba.

Por lo que toca a la validez será necesario, en estudios posteriores, correlacionar las puntuaciones en la prueba con las calificaciones en otras materias, así como correlacionarla con otras pruebas de aptitudes como por ejemplo el D.A.T., para determinar con mayor precisión su validez y emplearla junto con otras pruebas psicométricas y de conoci-

miento en el examen de admisión, para realizar predicciones más precisas del aprovechamiento escolar.

Deberán emplearse como medidas de criterio tanto el promedio de calificaciones en las diferentes materias cursadas en los semestres sucesivos de la carrera como la aprobación en determinadas materias y la deserción y permanencia en la escuela.

BIBLIOGRAFIA

- ALLPORT GORDON W. *La personalidad su configuración y desarrollo*, Ed. Herder España, 1977.
- ANASTASI ANNE; *Fundamentos de los Tests Psicológicos*; Ed. Aguilar Madrid, 1973.
- ANASTASI ANNE; *Psicología Diferencial*; Ed. Aguilar Madrid, 1974.
- ANASTASI ANNE; *Psychological Testing*; The Mc. Millan Company New York, 1966.
- BAUMGARTEN FRANZISKA; *Exámenes de Aptitud Profesional, Teoría y Práctica*; Ed. Labor Barcelona, 1957.
- BENDER LAURETTA; *Test Gestáltico Visomotor*; Ed. Paidós Buenos Aires, 1979.
- BENNETT, K. G., SEASHORE G H, WESMAN G.A. *Test de Aptitud Diferencia (manual)*; The Psychological Corporation New York, 1968.
- CRONBACH LEE JOSEPH; *Fundamentos de Exploración Psicológica*; Ed. Bibl. Madrid, 1963.
- HILGARD, ERNEST R. *Introducción a la Psicología*; Ed. Morata Madrid, 1975.
- KARMELO LOUIS J. *Medición y Evaluación Escolar*; Ed. Trillas México, 1978.
- MC. CONNELL JAMES V. *Psicología*; Ed. Interamericana México, 1978.
- MAGNUSON DAVID; *Teoría de los Tests*; Ed. Trillas México, 1969.
- MAIER NORMAN R.F. *Psicología Industrial*; Ed. Rialp Madrid, 1975.
- MORALES MA. LUISA; *Curso de Psicometría Aplicada*; Ed. Trillas México, 1974.
- NAYLOR J., BLUM M. *Psicología Industrial*; Ed. Trillas México, 1976.

- NÚÑEZ RAFAEL; *Aplicación del Inventario Multifásico de la Personalidad. (MMPI) a la Psicopatología*; Ed. El Manual Moderno México, 1968.
- PIERON H. *Tratado de Psicología Aplicada, La Psicología Diferencial, Libro 1*; Ed. Kapelusz, Buenos Aires, 1960.
- PORTUONDO JUAN A. *Test Proyectivo de Karen Machover*; Ed. Pax Madrid, 1979.
- RAPAPORT DAVID; *Test de Diagnóstico Psicológico*; Ed. Paidós Buenos Aires, 1974.
- RAVEN J.C. *Test de Matrices Progresivas (Manual)*; Ed. Paidós Buenos Aires, 1957.
- REUHLIN MAURICE; *Historia de la Psicología*; Ed. Paidós Buenos Aires, 1973.
- RUCH FLOY L. *Psicología y Vida*; Ed. Trillas México, 1975.
- SÁNCHEZ HIDALGO EFRAÍN; *Psicología Educativa*; Ed. Universitaria Río Piedras, 1963.
- SZÉKELY BELA; *Los Tests*; Ed. Kapelusz Buenos Aires, 1960.
- THONNARD F.J.; *Compendio de Historia de la Filosofía*; Ed. Pontificios París, 1949.

INDICE

INTRODUCCION	8
I. Confiabilidad y Validez de los Tests Psicológicos	11
<i>Introducción a los tests psicológicos</i>	13
<i>Clasificación de los tests</i>	15
<i>Normas de Calificación</i>	17
<i>Tipos de validez</i>	20
<i>Métodos para estimar la confiabilidad</i>	24
<i>Análisis de Items</i>	29
II. La Inteligencia y su Medición	35
<i>Antecedentes históricos y teóricos</i>	35
<i>Desarrollo mental</i>	44
<i>Investigaciones sobre la inteligencia</i>	47
<i>Pruebas de inteligencia</i>	48
a) <i>Test de inteligencia de Dominós</i>	48
b) <i>Test de Matrices Progresivas de Raven</i>	56
c) <i>Test de Wechsler</i>	62
III. Aptitudes y Rasgos de Personalidad	69
<i>Antecedentes históricos</i>	71
<i>Definición de las aptitudes</i>	73
<i>Estudios sobre las aptitudes</i>	75
<i>Pruebas de aptitudes</i>	77
<i>Aptitudes mecánicas y motoras</i>	78
<i>D. A. T. Test de aptitud diferencial</i>	81

<i>Personalidad</i>	87
<i>a) Inventarios de personalidad</i>	90
<i>b) Instrumentos Proyectivos</i>	95
IV. <i>La Medición del Criterio y la Utilidad del Predictor</i>	99
<i>Tipos de criterios</i>	101
<i>Requisitos de criterios</i>	104
<i>Utilidad de los instrumentos de medición</i>	111
<i>Confiabilidad del predictor y del criterio</i>	114
V. <i>Descripción del Estudio Realizado</i>	118
<i>Resultados</i>	125
<i>Conclusiones</i>	135
<i>Bibliografía</i>	136