

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO.
FACULTAD DE PSICOLOGIA.

SEIS ESCALAS DE VALIDEZ DE DOS INVENTARIOS DE PERSONALIDAD
UTILIZADOS EN SELECCION DE PERSONAL.

TESIS QUE, PARA OBTENER EL TITULO DE PSICOLOGO,
PRESENTA

JULIETA DEL ARCO FLORES.

MEXICO, D. F.

1973.



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

I N D I C E

	Pág.
INTRODUCCION.	
I LAS PRUEBAS PSICOLOGICAS EN LA SELECCION DE PERSONAL.	7
1) Las pruebas psicológicas como instrumentos predictores de la conducta en la Industria.	9
2) Los inventarios de personalidad en Selección de Personal.	15
II MMPI e IPP : SUS ESCALAS DE VALIDEZ.	28
1) MMPI, un bosquejo general de la prueba.	29
2) Escalas de validez del MMPI. Las escalas L, F, - y K.	36
3) IPP, un bosquejo general de la prueba.	61
4) Escalas de validez del IPP. Las escalas Sb, Bi y Cm.	69
III ANALISIS DEL FUNCIONAMIENTO DE LAS ESCALAS L, F, K, Sb, Bi y Cm EN UNA SITUACION DE SELECCION DE PERSONAL.	87
1) La muestra.	88
2) Distribución de frecuencias, medias aritméticas, - desviaciones estándar, curtosis y asimetría.	89
3) Diferencia entre medias (clasificación múltiple de - Duncan).	97
4) Correlaciones entre las escalas.	100
5) Análisis de varianza: varianzas y fuentes de varia-- ción.	118

	Pág.
6) Análisis factorial (Método Centroides de - Thurstone): el método centroides, procedimiento de cálculo, rotación de los ejes de referencia y - comentarios.	125
7) Composición factorial de la varianza: confiabili- dad, comunalidad, varianza única, varianza es- pecífica, varianza de error y totalización de factorialización.	154
CONCLUSIONES.	157
BIBLIOGRAFIA.	161

I N T R O D U C C I O N .

La idea de realizar este estudio surgió después de haber observado que el funcionamiento de las escalas de validez del MMPI en la situación de selección de personal, operan con una configuración típica; esto es, se aprecia que cuando sujetos aparentemente normales, con estudios superiores contestan inventarios de personalidad del tipo del MMPI, como requisito obligatorio para tener la posibilidad de ser considerados como aspirantes a un empleo determinado, las escalas L, F y K se configuran, para la mayoría de los casos, en forma de "V", en donde la escala L alcanza generalmente 7 puntos (T60), la escala F cuatro puntos (T53) y la escala K 18 puntos (T61). Esto aparentemente indica que bajo situaciones de presión los sujetos tienden a adoptar una postura defensiva.

Si se analizan, por separado las elevaciones de estas escalas, se encuentra que el promedio alcanzado por la escala L queda clasificado como "elevación moderada" (Dahlstrom y Welsh, 1972). Pudiera pensarse, por ejemplo, que puntajes crudos dentro del rango 7-8 podrían ser el resultado de una forma indiscriminada de responder a los ítems, sin ninguna relación con el contenido de los reactivos, puesto que si un sujeto contestara la prueba esencialmente al azar, cerca de la mitad de los ítems de L aparecerían contestados como "Falso" y el puntaje T debería caer más o menos en este rango. Pero en vista de la configuración de los perfiles totales, que en promedio se mantienen dentro de los límites T52 a T60, resultaría inapropiada una interpretación en este sentido, pudiéndose pensar que en general, el nivel de elevación en esta escala refleja una postura defensiva causada en gran parte por la presión exterior al tomar la prueba; en cierta forma, como señalan Good y Bratner (1961), la situación en la que la prueba es aplicada, puede producir ideas de referencia en los sujetos, los que en cierta manera son inducidos a responder los ítems de manera defensiva con el objeto de mostrarse como sujetos adecuados, lo cual puede explicar el perfeccionismo que se atribuyen a sí mismos, más que una limitación sería de su capacidad de "insight". (11)

Por lo que toca a la escala F, puede decirse que el promedio de elevación de esta escala corresponde a lo que Marks y Seeman (1963) han denominado "rango medio" y corresponde al nivel de puntaje obtenido generalmente por sujetos normales. Esta escala no se ve, de hecho, afectada por la situación de prueba y aparentemente mantiene un nivel independiente de los esfuerzos de los sujetos por dar la mejor imagen posible de sí mismos. (25)

La escala K se mantiene, en promedio, dentro de lo que Dahlstrom y Welsh (1972) denominan "promedio alto". En este caso, para interpretar la escala K, deberá considerarse primero que el nivel alcanzado en ella por un sujeto determinado, es fundamentalmente un reflejo de la forma en que el individuo se aproxima a la prueba, más que de la clase de persona que el sujeto es realmente, es decir, la situación a la que tiene que enfrentarse, su "status" actual y el propósito o meta que piensa alcanzar realizando los exámenes. (11)

Calificaciones en este "promedio alto" son, según Welsh y Dahlstrom (1972) obtenidos generalmente por sujetos de la clase media. Los sujetos con escolaridad mínima de bachillerato califican típicamente dentro de este rango (puntaje crudo de 16 a 20). En la mayor parte de los casos, las calificaciones dentro de este intervalo corresponden a sujetos bien ajustados,

con confianza en sí mismos y con capacidad para enfrentarse a sus problemas cotidianos. Estos sujetos tienen generalmente un punto de vista favorable hacia sí mismos, lo que se refleja en respuestas apropiadas; aparentemente los sujetos no se están describiendo en esta forma meramente para lograr aceptación social; su vida está bajo control y sus asuntos son bien manejados, parecen tener pocas perturbaciones emocionales, o amenazas a la autoestima o al autocontrol y muy pocos conflictos interpersonales. Estos sujetos se muestran retraídos, prudentes y circunspectos en su conducta normal; son descritos por Block, Bailey y Gough (1955) como razonables, lúcidos, con iniciativa, emprendedores, hábiles, versátiles, con altas habilidades intelectuales, ingeniosos, entusiastas, sociables, que gustan de tener ascendencia sobre los demás y con amplitud de intereses. En resumen, estos sujetos son competentes, eficientes y con buen equilibrio interior y en general se muestran satisfechos de sí mismos. Si en estos sujetos existen conflictos interiores, procurarán mantener una imagen adecuada ante los demás, en especial con aquéllos con los que sólo mantienen relaciones superficiales y negándose a sí mismos la existencia de tales conflictos. Block y Bailey señalan que estos sujetos se muestran a menudo reservados y taciturnos en la entrevista y con lenguaje que se percibe como muy elaborado. (11)

Por cuanto se refiere a la configuración en "V" de las escalas L, F y K, puede decirse que la elevación de las escalas L y K corresponde a un proceso de los sujetos, por medio del cual éstos buscan mostrarse a los demás desde el mejor ángulo posible. Varios estudios al respecto del proceso de distorsión del inventario muestran con bastante claridad que se trata de un proceso complejo y multifacético que sólo ha sido comprendido de una manera parcial. (10)

Los sujetos ingenuos y poco complicados psicológicamente, cuando tratan de presentarse a sí mismos, de una manera favorable ante figuras que de alguna manera perciben con autoridad, generalmente evitan los ítems de contenido más "estruendoso" de la escala L, pero de alguna manera fallan al tratar de manejar los aspectos autodevaluativos más sutiles involucrados en los ítems de la escala K. Los sujetos más mundanos y con mayor conocimiento de sí mismos y de las debilidades humanas, eluden las increíbles y simplistas virtudes de la escala L y rápidamente reconocen los lugares comunes de los defectos caracterológicos; pero ante las oportunidades más sutiles y tal vez más convincentes de autoengrandecerse que proporcionan los ítems de la escala K, son capaces de racionalizar y equivocarse de tal manera que terminan concediéndose a sí mismos el beneficio de la duda, obteniendo elevadas puntuaciones en K, (Dahlstrom y Welsh, 1972). Sin embargo, y a partir de los puntajes obtenidos en la escala K, no resulta del todo claro el si un sujeto es totalmente consciente de esta clase de distorsión y está eludiendo un conflicto consciente por medio de sus respuestas a la prueba, o si solamente está dando respuestas de la forma en que se percibe a sí mismo, de una manera automática y desde el punto de vista más cómodo para él, independientemente de su capacidad de tener, o no, un "insight" adecuado sobre sus problemas emocionales. (11)

Hay, en las mujeres, una clara tendencia a describirse más severamente que los hombres, por lo que las elevaciones de la escala K, para el sexo femenino probablemente reflejarán más la eficiencia psicológica que una postura autodefensiva, en tanto que para sujetos masculinos tales elevaciones parecen corresponder fundamentalmente a intentos de distorsión. (33)

En base a estos antecedentes, los perfiles que muestran elevaciones simultáneas en L y K dan lugar a algunos problemas de interpretación. Una --

alternativa que debería ser considerada, aún cuando existe muy poca evidencia a partir de investigaciones en este sentido, es que ambos indicadores - de validez reflejan realmente aspectos psicológicos de los sujetos. Las elevaciones simultáneas de L y K se encuentran también en los casos de distorsión deliberada del MMPI, así, esta configuración aparece en los estudios, en los cuales se pide a sujetos, normalmente considerados brillantes, que respondan al inventario de manera de mostrar la imagen más favorable posible de sí mismos (Hunt, 1948; Cofer, Chance y Judson, 1949; Gough, 1950; Wiggins, 1959). Posteriormente esta configuración se ha encontrado también entre pacientes psiquiátricos sinceros que han sido instruidos para contestar la prueba en la forma en que ellos consideraran que lo haría una persona bien ajustada. (Rapaport, 1958; Boe y Kogan, 1964, y Lanyon, 1967) o en la forma en que ellos responderían al inventario una vez fueran dados de alta (Marks y Seeman, 1963). Debe considerarse, también, que algunos de los perfiles que se encuentran en la práctica, sean producto del mismo tipo de distorsión deliberada. (6, 10, 11).

En la situación de selección de personal que atañe a este trabajo, el MMPI se utiliza para obtener una apreciación general de las condiciones en que se encuentra estructurada la personalidad; y no es utilizado para reportar rasgos de personalidad como predictores particulares del desempeño en el trabajo. Sin embargo, a raíz del análisis del perfil pueden hacerse algunos pronósticos sobre la actuación futura del sujeto; perfiles limítrofes o con elevaciones significativas afectan el corte final de aceptabilidad para un aspirante a empleo, ya que cuando se presentan indicios de perturbaciones de la personalidad, más o menos severos, éstos tenderían a afectar el proceso adaptativo al trabajo.

* * *

El IPP es un inventario de personalidad diseñado para evaluar rasgos referidos fundamentalmente al patrón de ajuste social, en aspectos que generalmente operan en un sentido positivo, ("deseables") tal como más explícitamente se aclara en el inciso 1 del segundo capítulo de esta tesis.

El IPP es una prueba que ha sido recientemente traducida al español (en forma no comercial) y cuyo manejo se encuentra en proceso experimental. Esta versión ha sido aplicada a más de 3000 profesionales mexicanos, encontrándose algunas configuraciones típicas de las tres escalas de validez que tiene la prueba. Este inventario es una versión modificada del CPI (California Psychological Inventory) de Gough, y la denominación que se le da (IPP) en este trabajo obedece a los cambios y alteraciones que se han hecho del protocolo original, aunque las escalas de validez empleadas en la muestra de sujetos que se reporta en esta tesis, corresponde a una traducción directa de los ítems que aparecen en las mismas escalas del CPI. Los sujetos que contestan el inventario en calidad de aspirantes a empleo, responden, en promedio, 38 reactivos de la escala Sb (T51), 26 de la escala Bi (T60) y 22 para la escala Cm (T35). En general puede apreciarse que a elevaciones de las escalas L y K se producen elevaciones de Bi. Se observan también disminuciones en Cm aunque aparentemente sin mayor concordancia con las elevaciones en L y K; es decir, puede observarse que cuando, por ejemplo, Bi se acerca a un puntaje de T50, la escala L se mantiene en un nivel similar, pero Cm se mantiene casi siempre en un nivel inferior a T50, independientemente del nivel alcanzado por las escalas de - - - - -

distorsión en ambos inventarios y dado que, en la situación actual la mayor parte de los sujetos mantienen elevaciones en L, K y Bi, que rebasan la línea media del perfil, aparentemente estas elevaciones mantienen alguna relación con los bajos puntajes en Cm.

Por otra parte, la escala Sb mantiene su promedio al nivel de T50, aunque parece tener ligeros incrementos con las elevaciones de L, K y Bi y viceversa.

Con el fin de poder precisar la relación existente entre las escalas de validez de ambos inventarios y al mismo tiempo tratar de aislar los factores que pudieran intervenir en el proceso de distorsión como producto de la situación de presión en la que los sujetos se ven requeridos a responder a las pruebas, se planteó la utilidad de realizar este estudio; en él se hace, inicialmente, una delimitación de la estructura y utilización de las escalas de validez en cuestión, y en segundo término se efectúa un análisis para aclarar la forma en que tales escalas interactúan, con la presunción teórica de poder establecer criterios que permitan comprender la intensidad y dirección en que estas escalas operan en el contexto de la selección de personal, así como fundamentar la interpretación de las mismas en función de los rasgos -- subyacentes (factores) encontrados, de manera tal que si un número menor de factores explican la varianza de las seis escalas estudiadas, resulta erróneo considerar que cada una de las escalas evalúa una postura defensiva diferente en los sujetos, por lo que viene a ser necesario delimitar la explicación cuantitativa y cualitativa de las mismas, basándose en los resultados obtenidos.

C A P I T U L O I

I.- LAS PRUEBAS PSICOLOGICAS EN SELECCION DE PERSONAL.

La medición de las diferencias individuales tiene una historia corta - pero muy activa. Ninguna otra corriente de investigación dentro de la --- psicología puede igualar la contribución hecha por los tests psicológicos - para el entendimiento de la conducta humana. En un lapso de apenas un siglo los psicólogos han delineado las principales dimensiones de las capacidades y características mentales humanas. Estos logros tuvieron su inicio con las investigaciones de Galton en Inglaterra, Ebbinghaus en Alemania, Binet en Francia y Terman y Thurstone en Estados Unidos.

Galton en su libro "Hereditary Genius", publicado en 1869, presentó un esquema de clasificación de los diferentes niveles de la capacidad humana.- Se ocupó de estudiar las diferencias interpersonales que se manifestaban en diferentes contribuciones a la sociedad, en cuestiones tales como las artes y las ciencias. Para poder expresar la posición relativa de cualquier sujeto dentro de una curva normal de distribución, estableció el concepto de calificación estándar del tipo de los puntajes "z" comúnmente empleados en estadística. Al desarrollar Galton esta herramienta, permitió que la medición de las capacidades humanas pudiera ser expresada en términos de un solo índice basado en las puntuaciones de todos los demás sujetos del grupo - para el cual se establecía la comparación. Hasta antes de Galton se había concedido muy poca atención a la posibilidad de medir la individualidad. - De esta manera, Galton fundamentó el estudio y medición de las diferencias individuales, haciendo hincapié en su significado y proporcionando la primera herramienta estadística apropiada para su medición. (16)

Hacia 1885, Ebbinghaus observó el fracaso de muchos de sus contemporáneos al tratar de estudiar actividades mentales complejas, tales como el pensamiento y la memoria. En una época en que la mayor parte de la experimentación psicológica estaba dedicada al estudio de la agudeza sensorial y del tiempo de reacción, Ebbinghaus realizó una serie de estudios exhaustivos sobre la memoria humana; no hizo hincapié sobre las diferencias individuales, ya que la mayor parte de sus experimentos fueron hechos consigo mismo; sin embargo, influyó sobre sus contemporáneos al abrir la posibilidad de estudiar funciones mentales complejas en lugar de respuestas simples; se interesó por el estudio de los procesos mentales humanos tal como ocurren en la vida real, en vez de tratar de dividirlos en componentes elementales sensoriales y motores, como era característico en la mayoría de las investigaciones de la época. (16)

En 1895, Binet fundamentó los principios de su "Psicología Individual", en donde criticó acrememente las prácticas comunes de pruebas sensoriales, argumentando que debería estudiarse al individuo integralmente, e insistiendo en la necesidad de medir facultades superiores como la memoria, la formación de imágenes, la imaginación, la atención y la comprensión. Durante la década siguiente, Binet experimentó con un gran número de tareas diseñadas para medir dichas facultades; analizó cuidadosamente el proceso mental necesario para el aprendizaje escolar, por medio de patrones diferenciales entre alumnos brillantes y torpes, lo que le permitió contrastar sus característi

cas particulares. Con este método empírico Binet logró desarrollar pruebas - que diferenciaban a los niños brillantes de los niños torpes, señalando que éstas eran las medidas de las facultades superiores que había estado buscando. La primera prueba de Binet, publicada en 1905, consistió en 30 tareas que oscilaban desde muy fáciles hasta difíciles. (1, 16)

El trabajo de Binet encontró una amplia respuesta en el resto del mundo, - muchos investigadores pusieron a prueba sus tests y se dieron cuenta de que - proporcionaban estimaciones exactas de la posición mental de los niños y permitían efectuar predicciones de su capacidad de aprendizaje. (16)

En Estados Unidos, Terman, emprendió una revisión y extensión de las ---- pruebas de Binet, agregando nuevos tests, y los estandarizó cuidadosamente en- 2300 niños escolares. Su revisión, publicada en 1916, contenía 90 tareas --- arregladas en orden creciente de dificultad. La puntuación de esta prueba de- Stanford-Binet se expresó como cociente de inteligencia, que venía a ser el cociente de la "edad mental" del individuo y la edad cronológica del sujeto. Con esta prueba la medición de las diferencias individuales cobró un vigoroso- impulso después del examen y selección de casi 2 millones de reclutas del ejército durante la primera guerra mundial. (16)

Con la amplia difusión de las pruebas de inteligencia surgió el interro-- gante de la naturaleza de la misma; se sabía, desde luego, que las personas - que calificaban alto en tales pruebas aprendían por lo general más rápido y - eran más capaces de confrontar ocupaciones que involucraran problemas altamen- te complejos. Se despertó sin embargo la controversia de si la inteligencia- era un rasgo monodimensional, o bien estaba compuesta de muchos factores de aptitudes básicas. Debe tenerse en cuenta que Binet pensó siempre en medir las diferentes facultades que componen la inteligencia, y puede asegurarse que - Terman jamás pensó en un concepto monolítico de inteligencia, aunque el uso - del C.I. pareció implicarlo. Sin embargo, los psicólogos británicos cerraron los ojos a este hecho y encabezados por Charles Spearman, defendieron airada-- mente que la inteligencia debía considerarse como una extensa capacidad gene-- ral. (16)

Fue el trabajo monumental de Thurstone el que resolvió la cuestión; desa- rolló y usó el análisis factorial múltiple para estudiar los patrones de rela- ción entre puntuaciones de las capacidades humanas. La investigación de --- Thurstone lo llevó a la conclusión de que las capacidades humanas consistían - de siete factores básicos, prácticamente independientes. (16)

La investigación de Thurstone sirvió de base para los estudios de Guilford sobre la naturaleza del intelecto humano, mismos que han permitido estructurar un esquema fundamental para el estudio y comprensión de la inteligencia.

Sin embargo, uno de los aspectos que ha evolucionado menos dentro de la - métrica humana, es la evaluación de la personalidad; comenzando por la caren-- cia de acuerdo sobre el término "personalidad", aunque para los fines de este- trabajo, la definición dada por Dunnette y Kirchner (1972), resulta adecuada:

Personalidad es la conducta que el sujeto presenta para ajustarse ".....a

las demandas interpersonales y situacionales de su medio". (12)

El empleo industrial de la métrica de la personalidad se dirige, no al -- diagnóstico psiquiátrico o al tratamiento, sino a la evaluación de los modos - conductuales que, siendo relativamente estables, se reflejan en patrones norma les de ajuste.

1.- LAS PRUEBAS PSICOLOGICAS COMO INSTRUMENTOS PREDICTORES DE CONDUCTA EN LA INDUSTRIA.

Las pruebas psicológicas, tal como se utilizan en la industria, tienen - distintas finalidades, entre las que pueden señalarse las siguientes: (31,32)

- 1.- Selección de nuevos empleados.
- 2.- Ubicación de empleados en distintos puestos.
- 3.- Ascensos.
- 4.- Evaluación de la eficacia de los programas de capacitación y adiestramiento.
- 5.- Asesoría y consulta.

Resulta más frecuente en la industria, encontrar que las pruebas psicológicas se emplean para seleccionar y colocar a la gente, que por otros concep-- tos.

En virtud de que la experiencia ha mostrado con bastante claridad la im-- portancia de las diferencias individuales en el desempeño del trabajo, resulta ría ocioso encarecer la importancia de la colocación adecuada que debe dársele a cada empleado, por lo que se pasará a referirse a las pruebas psicológicas - como sistema suplementario de la selección de personal.

La investigación actual en la industria, ha demostrado que el empleo de - las pruebas psicológicas permite un importante avance sobre otros sistemas de selección de personal que prescinden de las mismas, aunque no por ello dejan - de ser un sistema suplementario que incluye ciertos márgenes de error, como - cualquier otra herramienta apoyada en procedimientos estadísticos. (21)

Las pruebas de aptitud y de rendimiento han sido consideradas como la par te fundamental de la selección de personal dado que, cuando son utilizadas sis temáticamente, permiten a la empresa optar por los aspirantes mejor califica-- dos para los puestos considerados. Las pruebas que se utilizan para la selec ción de futuros empleados pueden dividirse, de una manera muy gruesa, como - pruebas de grupo o individuales. Las primeras han encontrado en la industria mayores posibilidades de utilización, fundamentalmente por razones de índole - económica. (17)

Los principales tipos de pruebas que son utilizadas en la industria son -

las siguientes: (1, 8, 12, 31, 32)

1.- Pruebas de Actuación.- El tipo más sencillo y quizá el más utilizado es la muestra de trabajo, o prueba de actuación, en la cual el solicitante tiene que demostrar que puede realizar un trabajo determinado; por ejemplo, cuando una mecanógrafa solicita trabajo, se le pide que transcriba algunas páginas y luego se valúan su habilidad y su rapidez. En este caso, las características y aptitudes determinantes del éxito no son manifiestas, y las capacidades necesarias para desarrollar la tarea no son explícitamente medidas.

2.- Pruebas de "Inteligencia".- Dando por supuesto que las personas brillantes y rápidas en aprender pueden asimilar casi cualquier tarea más rápidamente y mejor que quienes están menos dotados, muchas empresas emplean pruebas de "inteligencia general"; aunque a la fecha no existe un consenso unánime al respecto del concepto de inteligencia, por lo que la métrica de esta característica o conjunto de características humanas resulta aún hoy confusa y ambigua.

3.- Pruebas de Aptitud.- Estas pruebas están generalmente enfocadas a medir la capacidad potencial de los aspirantes a empleo para aprender una nueva labor o comprender algún tipo específico de problemas; se han creado un gran número de pruebas de aptitud especializadas, cuyo intento es predecir las probabilidades que determinado solicitante puede tener para aprender cierto tipo de trabajo de una manera efectiva. Así pudiera presentarse el caso de que a una persona le resultara difícil realizar trabajos que comporten cálculos, pero que tuviera una buena capacidad para visualizar cuerpos geométricos en el espacio. Existen pruebas para valorar el nivel en que un empleado que carece de aptitudes para realizar una labor especializada, pueda aprovechar sus habilidades en otros renglones y resultar útil a la empresa que lo emplea.

4.- Pruebas de Personalidad.- Un gran número de empresas tienen gran interés en los ajustes personales de sus empleados, tanto sociales como interpersonales, ya que se piensa que, independientemente del bagaje intelectual del sujeto, sus patrones de ajuste al medio, determinan en gran medida el que el empleado resulte productivo. Cuando este punto de vista se enfoca al personal que debe ocupar puestos directivos o de supervisión de otros empleados, cobra fuerza la idea de que este tipo de aproximaciones métricas son necesarias, por cuanto estos empleados demostrarán su eficiencia para la industria, enfrentando un cúmulo de tareas que involucran relaciones interpersonales. Resultaría, sin embargo, absurdo pensar que tales pruebas pueden abarcar la "personalidad total" de un sujeto, pero no obstante las serias limitaciones de estos instrumentos, permiten efectuar aproximaciones válidas de algunos aspectos de los aspirantes, que pueden estar estrechamente relacionados con su desempeño laboral.

Existen diversos factores, tanto de índole práctica como teórica, que determinan la utilidad funcional de las pruebas psicológicas en las organizaciones industriales, entre las que se encuentran la validez, la confiabilidad y la proporción de selección.

La Validez.- En un sentido lato, la validez se refiere al grado en que -

una prueba es capaz de conseguir los fines o propósitos que con ella piensan alcanzarse; el proceso de comprobación del valor de una prueba determina la validez de la misma a efectos de la selección de personal. El nivel de validez de una prueba, afecta el valor u objetividad que se espera del test en cuanto a su finalidad de seleccionar y colocar personal. (32)

Los tipos de validez de una prueba, son fundamentalmente cuatro, y dependen primordialmente del método empleado para probarla. (24)

a).- Validez de contenido.- La validez de contenido indica hasta qué punto el contenido de la prueba abarca el objetivo y las situaciones que se intentan medir con el instrumento. Para poder estimar la validez de contenido, es necesario definir exactamente tales objetivos y situaciones, por lo que esta determinación debe hacerse en base a las opiniones de "expertos" sobre la exactitud de las cuestiones o problemas (el contenido de la prueba), y su capacidad para medir, por esta razón la validez de contenido no puede expresarse como un coeficiente de validez.

b).- Validez predictiva.- Este criterio proporciona un índice del grado en que las predicciones hechas gracias a la prueba, quedan confirmadas posteriormente; por lo general este tipo de validez se determina comprobando la relación existente entre los resultados de la prueba y un criterio de ejecución de la tarea. La validez predictiva computada en forma de coeficiente, resulta indispensable para propósitos de selección y clasificación de personal.

c).- Validez concurrente.- La validez concurrente indica hasta qué grado los resultados de una prueba se relacionan con algún criterio actual (o concurrente) de precisión y exactitud. La razón fundamental para construir una prueba que mida una variable de la cual ya se tienen datos, es que en la mayoría de los casos, el test ahorra tiempo y esfuerzo y proporciona el mismo resultado que la medida de criterio, y, como la validez predictiva, se expresa por un coeficiente de validez.

Muy frecuentemente, la validez concurrente está dada por el coeficiente de correlación entre la prueba y el criterio. En la validez concurrente, la cuestión es si el test mide lo que se pretende medir, y el coeficiente de validez indica qué tan adecuados son los datos de la prueba como base para el diagnóstico, en el sentido más amplio de la palabra, versus la validez predictiva, donde el problema es la exactitud con que el test predice lo que intenta predecir, aspecto especialmente importante en la selección y clasificación de personal. Es importante señalar aquí que una buena capacidad predictiva presupone una buena capacidad diagnóstica, si algo no puede medirse con exactitud, difícilmente podrá usarse como base para predicciones válidas.

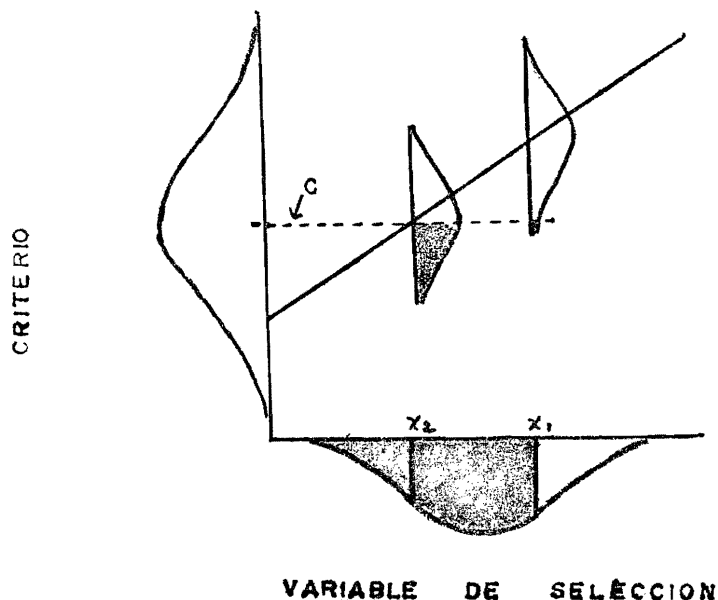
d).- Validez de construcción.- Este criterio establece cuáles son los aspectos psicológicos que trata de medir una prueba cuando no existe un criterio externo para medir tales rasgos. Esta validez suele determinarse comparando una prueba con otra que ha demostrado su utilidad para evaluar la cualidad de que se trata, o bien utilizando el análisis factorial que puede detectar el -

grado en que distintos tests pueden medir, en común, tal característica.

Otro aspecto que debe ser satisfecho por las pruebas psicológicas es el de la confiabilidad, esto es, que el instrumento mida de manera constante el aspecto que pretende valorar. Los datos deben ser confiables desde dos puntos de vista: deben ser significativos y reproducibles.

El primer requisito, se refiere a que el instrumento de medida evalúe -- realmente el rasgo que intenta medir, en tanto que el segundo aspecto indica que los resultados obtenidos con la prueba en determinado momento y bajo ciertas condiciones, deben ser los mismos si se vuelve a medir el mismo rasgo en condiciones idénticas. De hecho, la confiabilidad es la exactitud de la medición, independientemente de que realmente se esté midiendo lo que se pretende medir, por lo que puede ocurrir que una prueba con un alto coeficiente de confiabilidad, sea totalmente inútil para ser empleada en la selección de personal por existir muy poca o ninguna relación entre los resultados obtenidos y el éxito que el individuo en cuestión tenga posteriormente en el desempeño en el trabajo. En otras palabras, la prueba puede tener un alto coeficiente de confiabilidad, aunque su índice de validez (predictiva o concurrente), sea muy bajo aunque en la situación inversa (una prueba con baja confiabilidad) resultaría improbable encontrar una alta validez. (24)

Por otra parte, el empleo de las pruebas psicológicas en selección de personal tiene generalmente por objetivo elegir de un grupo más o menos extenso de aspirantes, una porción más pequeña que, en promedio, supere al grupo mayor en los aspectos sujetos a valuación, sean cuales fueren éstos. En la práctica, el sistema de selección de personal permite, a efecto de las pruebas, el ingreso de un cierto número de individuos que de hecho fracasarán en el empleo, y asimismo rechaza a una proporción similar de personas que hubieran dado resultados adecuados. Ahora bien, siempre que el porcentaje de aspirantes que ingresen a la empresa y que resulten buenos empleados, sea significativamente mayor mediante el empleo de pruebas psicológicas que sin las mismas, el sistema resultará benéfico y la discriminación de los sujetos vendrá a ser más eficiente y tendrá mayores probabilidades de reeditar ganancias a la compañía.



De acuerdo con la figura que aparece en la página anterior, para la cual se han dado dos límites de corte X_1 y X_2 ; con corte en X_1 aceptamos solamente un pequeño número de solicitantes; con el límite de corte en X_2 aceptamos un número de solicitantes considerablemente mayor. Se han dibujado las distribuciones de error alrededor de la línea de regresión para los límites de corte X_1 y X_2 ; si el puntaje C sobre el eje Y representa el nivel en la variable de criterio, para la cual el resultado puede considerarse como exitoso, puede concluirse que el riesgo de fracaso para los aceptados con el puntaje más bajo, en el límite de corte X_2 es aproximadamente el 50%. Por otra parte, el riesgo de fracaso resulta muy pequeño para los aceptados con límite de corte en X_2 . (9)

Por consiguiente, cuanto menor sea la proporción de selección, tanto mayor será la efectividad de la selección expresada como proporción de éxito. También, como puede verse en la figura anterior, cuanto más alto sea el coeficiente de validez, tanto mayor será la efectividad expresada como proporción de éxito; a mayor validez menor desviación de error de los puntajes de criterio alrededor de la línea de regresión, por lo tanto, para un nivel crítico dado la distribución de criterio, cuando la validez se incrementa, presenta menor riesgo de fracasar en los aceptados de cada nivel. (9)

En la mayoría de las situaciones de selección, la decisión final de aceptar o rechazar sujetos, se basa en los resultados de varias pruebas. Los distintos subpuntajes se combinan para tomar decisiones; aunque esto puede lograrse de diferentes maneras, consideraremos solamente las dos más comunes. (24)

El nivel crítico de aceptación puede obtenerse de varias maneras, el mismo puntaje total puede obtenerse como la suma de un puntaje alto en algunas pruebas y un puntaje bajo en otras, pero las pruebas que aparecen con calificaciones altas y calificaciones bajas, pueden diferir grandemente de un individuo a otro; en este caso, la suma de los puntajes de los subtests es lo que decide si un individuo es aceptado o no, sin importar en cuál o cuáles de los subtests ha obtenido el mejor resultado. Por este método, todos los sujetos para los cuales la suma de los puntajes estándar obtenidos en los subtests, es menor que el puntaje crítico señalado, serán rechazados; en tanto que todos aquellos que caigan por encima del mismo, serán aceptados; este procedimiento se denomina "modelo compensatorio", ya que es posible que quien obtiene bajos puntajes en un área determinada, pueda compensarlo siendo fuerte en otra.

En algunas situaciones, el modelo compensatorio de los puntajes no es el mejor método; este es el caso cuando un nivel crítico debe excederse en varias distribuciones de subtests, donde no es posible compensar la debilidad en ciertos aspectos mediante fuerza en otros; es decir, para aceptar a un sujeto, éste debe rebasar todos los límites de corte fijados, este procedimiento es denominado método de corte múltiple o "modelo conjuntivo".

No hay prueba que por sí sola sea capaz de medir todas las capacidades y características que se requieren para desempeñar un oficio o empleo; aún las actividades laborales más sencillas resultan complejas si se considera el conjunto de capacidades que se requieren para desempeñar adecuadamente un empleo.

Esto hace, por lo general, necesario el empleo de baterías de pruebas, - las que, gracias a los métodos estadísticos, pueden combinarse para dar un resultado compuesto, de tal manera que cada una de ellas se ajuste para dar las mayores correlaciones posibles entre el resultado de las pruebas y el criterio.

En muchos casos es necesario, además, utilizar una batería de pruebas más bien que una sola, puesto que lo más que puede cubrir una sola prueba por sí misma, es uno de los múltiples aspectos o cualidades que son necesarias para el desempeño del trabajo en cuestión. Se requiere una batería completa para obtener una visión amplia del conjunto de características que deben conjugarse para predecir un desempeño exitoso en el trabajo.

2.- LOS INVENTARIOS DE PERSONALIDAD EN LA SELECCION DE PERSONAL.

Cuando se habla de inventarios de personalidad, del tipo de lo que Cattell (1957) ha llamado "tests objetivos" (los que proporcionan medidas directas que no son modificadas por la interpretación), se pretende, con su aplicación, obtener una muestra de conducta típica de los sujetos, lo que, en otras palabras, refleja la orientación conductista de manejar respuestas observables. Esta -- consideración no deja de tener un alcance limitado, ya que la mensuración de la personalidad humana no puede eludirse de considerar aspectos psíquicos internos; ciertas expresiones, tales como "agresividad latente", "actitudes hacia los demás" o "tolerancia a la frustración", hacen referencia a estos procesos. Muchos fenómenos psicológicos importantes, tales como las emociones o los sueños, ocurren sólo dentro de la mente del sujeto, incluso gran parte de los procesos de ajuste de un individuo, están determinados más por la percepción que el sujeto tiene de los hechos, que por los hechos mismos. En consecuencia, muchos psicólogos prefieren estudiar las reacciones subjetivas, antes que las respuestas externas. (7)

Los primeros cuestionarios de personalidad fueron creados con la idea de estudiar el mundo interno de la percepción y de los sentimientos. A finales del siglo pasado, Galton ideó esta técnica ante la necesidad de utilizar un procedimiento estandarizado que pudiera ser aplicado a un gran número de sujetos en sus estudios sobre las imágenes mentales. Stanley Hall, en sus extensos estudios sobre el desarrollo de los adolescentes utilizó también cuestionarios, de donde consideró las respuestas conjuntas de un gran número de adultos, sin tomar en consideración a los individuos aislados. (8)

La técnica de cuestionario cumplía funciones diferentes para ambos científicos; para Galton, los inventarios tenían por objeto el de obtener información sobre hechos internos que sólo así podían encontrar una manifestación externa; en tanto que para Hall no eran más que un medio para simplificar el trabajo que hubiera implicado la observación directa de la conducta. (8)

El primer inventario para evaluar los procesos personales de ajuste al medio fue el Woodworth Personal Data Sheet, empleado por el ejército de los Estados Unidos durante la primera guerra mundial, con el fin de detectar a los soldados que presentaban posibilidades de desmoralización en el combate, cuando las entrevistas psiquiátricas eran impracticables en un proceso de reclutamiento que incluía miles de sujetos. Este cuestionario incluía una lista de síntomas parecida a la que los psiquiatras hubieran empleado, explícita o implícitamente, en una entrevista y las presentó en forma de cuestionario; difería de la entrevista fundamentalmente en que cierta exactitud del interrogatorio personal, era sacrificada en aras de la velocidad. Aquellos sujetos que respondían al cuestionario en forma tal que mostraban desajustes severos, eran remitidos posteriormente al servicio psiquiátrico. El cuestionario tenía la ventaja de detectar a los soldados desajustados en una situación en la que hubiese sido imposible la exploración individual. (8)

La escala de Woodworth fue precursora de un gran número de "inventarios de

ajuste" que consistieron fundamentalmente en listados de problemas, síntomas o dolencias, sin intentar una descripción detallada de la personalidad y proporcionando, en muchos casos, sólo una puntuación única, que representaba el nivel de ajuste. Algunas veces no se investigaba más que un tipo de síntomas, -cual es el caso del Cornell Medical Index, que estudiaba síntomas somáticos. En otras ocasiones los ítems se hallaban agrupados en categorías lógicas, como el inventario de ajuste de Bell, que presenta puntuaciones del ajuste familiar, físico, social y emocional. Este tipo de inventarios presume que los ítems -incluidos son capaces de diferenciar entre sujetos desajustados y aquéllos que no lo son y tienen por objeto identificar a los individuos que necesitan orientación o terapia. (1, 8)

Los inventarios de ajuste deben considerarse, en el mejor de los casos, - como instrumentos de clasificación que permiten seleccionar personas que reconocen sus propios síntomas; no son útiles para medir de manera definitiva ningún rasgo y su información es superficial.

Durante el período comprendido entre 1920 y 1945, los inventarios fueron considerados principalmente como un sustituto para la observación de la conducta, y las cuestiones se centraban más en lo que el individuo hacía que en lo que sentía o pensaba. Los cuestionarios fueron ampliados para describir tantos aspectos de la conducta como fuera posible y las respuestas fueron agrupadas de manera homogénea para proporcionar puntuaciones de rasgos o perfiles de respuestas. Durante este período se concibió la personalidad como un conjunto de hábitos; se consideraba la fuerza de tales hábitos o rasgos en función - de la frecuencia de respuestas que se presentaban.

En los primeros inventarios la lista de categorías de conducta era seleccionada de una manera arbitraria; algunos rasgos como el de "confianza en sí mismo", provenían de la experiencia y del sentido común y otros, en cambio, como la "introversión", estaban fundamentados en ciertas teorías de la personalidad. Se elaboraron docenas de estos instrumentos con sólo tomar los ítems de los predecesores, añadir unos cuantos nuevos y calificarlos según otras combinaciones. El instrumento más conocido de este tipo, fue el inventario de personalidad de Bernreuter, semejante al inventario de Bell, pero con reactivos - más variados, que calificaban tendencias neuróticas (patrón de ajuste), autosuficiencia, introversión y dominancia.

Flanagan realizó una investigación sobre el inventario de Bernreuter, introduciendo las puntuaciones de rasgos definidos de acuerdo a ciertas normas estadísticas; adoptó el principio de que, para merecer nombres distintos, los rasgos deben poseer entre sí una correlación muy baja. Correlacionó las puntuaciones de este inventario para 305 adolescentes y encontró que los rasgos - de introversión y de tendencia neurótica, tal como los mide el cuestionario, no eran independientes, ya que varios reactivos se hallaban saturados en ambas escalas y la correlación entre éstas era de .87. Al aplicar el análisis factorial, Flanagan observó que la "confianza en sí mismo" y la "sociabilidad" - bastaban para explicar las cuatro escalas; las puntuaciones entre estos dos rasgos apenas si correlacionaban entre sí, lo que condicionaba una independencia recíproca. (8)

Después del término de la segunda guerra mundial, siguió un período en el que se intentó estudiar estadísticamente las dimensiones que pudieran explicar la personalidad. Las intercorrelaciones de los ítems condujeron a Guilford a la idea de que la introversión podía ser dividida en introversión social, introversión mental, depresión, ciclotimia y rathimia (despreocupación). De acuerdo con sus hallazgos, Guilford desarrolló un inventario que incluía estos cinco factores, al que más adelante añadió otros ocho aspectos de la personalidad, aunque las escalas nunca llegaron a presentar correlaciones nulas. Otros investigadores utilizaron los datos obtenidos por Guilford dentro de configuraciones que consideraron más efectivas; así Thurstone pensó que los trece rasgos de Guilford podían reducirse a siete factores, a los que denominó reflexivo, sociabilidad, estabilidad emocional, vigor, ascendencia, actividad e impulsividad. (16)

Las posibilidades de combinación de ítems, conjuntos de respuestas, escalas, etc., hacen que este juego no tenga fin, máxime que hasta la fecha no -- existe ninguna teoría definitiva que permita un acuerdo acerca del número de factores de la personalidad, ni sobre su organización o sobre la nomenclatura que les debe ser aplicada.

La construcción de cuestionarios de personalidad de acuerdo con el "principio del signo" *, tal como se utiliza en el Strong, ha sido empleada muy rara vez en los inventarios de personalidad, debido a que los criterios en esta área, se encuentran sujetos a polémica. Un punto de partida con mayores posibilidades de utilización práctica, está dado por la clasificación psiquiátrica; sin embargo, hay algunos casos de utilización de claves empíricas para la construcción de inventarios de personalidad; cual es el caso de la escala Humm----Wadsworth de temperamento, que se utiliza para la distinción entre maniacos y paranoicos.

En esencia, este tipo de técnica se utilizó en la construcción del MMPI; - inventario que fue publicado en 1942 y cobró rápida aceptación siendo actualmente uno de los más utilizados e investigados.

Mientras que las corrientes factoriales y empiristas han creado pruebas - que intentan encontrar aspectos que correlacionen entre sí, otra tendencia con siste en estructurar un esquema teórico y preparar ítems que proporcionen información en base a tal esquema. Los primeros trabajos en este sentido fueron estimulados por la teoría de Jung sobre la personalidad; aunque esta teoría posee ya muy escasa influencia, encontramos sin embargo inventarios como el de Myers-Briggs, cuyos ítems y claves de puntuación se encuentran enfocados

* Principio teórico que señala que cierto tipo de respuestas son características de un grupo determinado de personas y permiten establecer la clasificación diferencial de los mismos.

explícitamente hacia la teoría jungiana. Otros instrumentos creados a la luz de esta orientación, son la escala de Edwards, la de Taylor de Ansiedad Manifiesta, la escala F de California para identificar personalidades "autoritarias". La validación de este tipo de pruebas exige por lo general una investigación intensa y concienzuda y presupone que el armazón teórico está debidamente fundamentada y es lo suficientemente sólida para permitir el desarrollo de un instrumento de medida.

Entre algunos de los inventarios de personalidad más importantes, citaremos: (5, 8)

- 1).- Billet-Starr Youth Problems Inventory. Es un inventario de problemas que cubre doce áreas a saber: salud y seguridad físicas, relaciones con los demás, relaciones heterosexuales, vida familiar, organización financiera personal, intereses y actividades, vida escolar, potencialidades personales, proyectos para el futuro, salud mental, moralidad y religión. Fue elaborado con el fin de obtener una clasificación general de sujetos que debían pasar a exploración individual, así como para identificar algunos aspectos que deben ser tomados en cuenta en la orientación profesional.
- 2).- California Personality Inventory. Es un amplio inventario que incluye la valoración de 18 rasgos, mismos que se mencionan en el siguiente capítulo. Los baremos de calificación fueron elaborados empíricamente, y aunque la interpretación se basa fundamentalmente en una integración psicológica del perfil completo, cubre más aspectos de la personalidad que la mayor parte de los inventarios que aquí se mencionan. Algunas de las calificaciones correlacionan bastante alto entre sí, lo que ocasiona algunos problemas en la interpretación de las escalas aisladas.
- 3).- California Test of Personality. Este cuestionario proporciona datos sobre el ajuste personal y social e incluye 16 escalas: confianza en sí mismo, sentido de autoestima, sentido de independencia, sentimientos de pertenencia, tendencias evasivas, síntomas nerviosos, valor total del ajuste personal, ajuste social, habilidades sociales, tendencias antisociales, relaciones familiares, relaciones escolares, relaciones ocupacionales, relaciones con la comunidad, ajuste social total, ajuste total.

Algunas de las escalas como el sentido de autoestima, síntomas nerviosos y relaciones familiares, arrojan distribuciones atípicas que muy difícilmente proporcionan información significativa sobre los patrones de ajuste; asimismo el intento del manual por simplificar la teoría y la utilización del inventario, propicia solamente dificultades en la aplicación e interpretación del mismo.

- 4).- Edwards Personal Preference Schedule. Este inventario proporciona puntajes de quince necesidades en base a 225 comparaciones apareadas: logro, aceptación, ascendencia, exhibicionismo, autonomía, afiliación, intracepción, ayuda, dominio, degradación, alimentación, cambio, persistencia, heterosexualidad y agresión. La prueba está diseñada de manera que minimiza los efectos de "fachada", aunque las respuestas son susceptibles de ser falseadas. Los datos proporcionados son útiles en problemas de orientación.

- 5).- Gordon Personal Profile y Gordon Personal Inventory. Cada una de las formas utiliza 20 ítems de elección forzada; el perfil mide dominancia, responsabilidad, estabilidad emocional y sociabilidad y el inventario, cautela, creatividad, relaciones personales y energía. La evidencia, en lo que se refiere al nivel de significancia de las puntuaciones, es extremadamente limitada, y los coeficientes de confiabilidad reportados por el autor (por encima de .80), parecen ser excesivamente altos.

- 6).- Guilford-Zimmerman Temperament Survey. Evalúa 10 rasgos relativamente independientes obtenidos mediante el análisis factorial: actividad general, inhibición, ascendencia, sociabilidad, estabilidad emocional, objetividad, relaciones amistosas, condescendencia, relaciones personales, masculinidad y 3 escalas de falsificación: sutil, ingenua y actitud negligente. Es una versión más eficaz de las primeras escalas de Guilford (Inventory of Factors STDCR, Guilford-Martin Inventory of Factors GAMIN y Guilford-Martin Personnel Inventory) y constituye un valioso instrumento descriptivo, aunque los datos sobre la validez del instrumento son restringidos.

- 7).- Kuder Preference Record (Forma Personal). Es un complemento del Inventario de Intereses Profesionales y está integrado por 168 ítems de elección forzosa que miden la preferencia entre la actividad en grupo, situaciones estables, trabajo teórico, evasión de conflictos, trabajo directivo y validez. Se han agrupado ciertas estructuras ocupacionales que mejoran la utilidad de esta escala en orientación profesional, aunque se conocen pocos datos acerca de su empleo como instrumento descriptivo o de diagnóstico. Esta prueba se halla prácticamente libre de los efectos de la deseabilidad social, aunque es susceptible de ser distorsionada; dado que los reactivos no poseen un sentido evidente de "bueno" o "malo", el inventario cumple adecuadamente con sus fines.

- 8).- Minnesota Counseling Inventory. Este inventario consta de 355 ítems que deben ser contestados como cierto o falso, y se enuncian como frases autodescriptivas que no son más que proposiciones del MMPI. Consta de 9 escalas: relaciones familiares, relaciones sociales, estabilidad emocional, conformidad, ajuste a la realidad, estado de ánimo, liderazgo, validez y no contestadas. Las escalas-

poseen validez, aunque no muy altas, para separar niños y preadolescentes que padecen un ajuste defectuoso en relación con aquéllos que lo poseen adecuadamente; los retests efectuados después de tres meses, muestran confiabilidades que oscilan entre .70 y .80.

- 9).- Minnesota Multiphasic Personality Inventory. Este inventario fue elaborado por un psicólogo, Starke Hathaway, y un psiquiatra, J. C. Mckinley, con el objeto de servir de complemento en el diagnóstico-psiquiátrico. Consta de 566 reactivos que permiten enunciar los datos diagnósticos en términos psiquiátricos; los reactivos deben ser contestados como cierto o falso; el contenido de éstos es muy diverso, algunos se refieren a aspectos conductuales, otros a sentimientos observables exteriormente y algunos a actitudes generales de tipo social. Una visión más amplia de este inventario se proporciona en el siguiente capítulo de este trabajo.

- 10).- Money Problem Check Lists. Proporciona evaluaciones para nueve áreas diferentes: salud, seguridad económica, mejoramiento personal, personalidad, hogar y familia, noviazgo, sexualidad, religión y ocupación. Las puntuaciones altas señalan a los sujetos que deben recibir orientación y los ítems proporcionan la base sobre la cual se establece la discusión de la prueba.

- 11).- SRA Youth Inventory. Es un inventario eficaz, que cubre los problemas típicos de los adolescentes en el terreno pedagógico y de orientación, así como el ajuste social y emotivo. Incluye nueve escalas: escuela, futuro, "yo mismo", gente, hogar, sexo, salud, aspectos generales y dificultades básicas. Es generalmente utilizado como punto de partida para la orientación individual o de grupo y puede emplearse también como un inventario de selección que permite detectar aquellos individuos que requieren un estudio intenso.

- 12).- Sixteen Personality Factor Questionnaire. Incluye 16 puntuaciones que miden 16 pares de características factoriales de la personalidad encontrados por Catell: Sizotimia vs. afectotimia, propensión a la neuroticidad vs. fuerza del yo, sumisión vs. dominancia, menos inteligente vs. más inteligente, prudente vs. impulsivo, oportunista vs. concienzudo, threctia vs. parmia, harria vs. premsia, alaxia vs. protensión, praxermia vs. autia, naturalidad vs. astucia, confianza en sí mismo vs. propensión a la culpa, conservatismo vs. radicalismo, dependiente del grupo vs. autosuficiente, autoconflictivo vs. autodisciplinado y baja tensión érgica vs. alta tensión érgica. Estas dimensiones son relativamente independientes y presentan ciertas ventajas por lo que se refiere a la investigación. Las escalas cortas poseen una confiabilidad muy baja (.45 a .55) y la información sobre las normas no es del todo satisfactoria.

- 13).- Study of Values. Es un inventario que se utiliza comunmente con jóvenes y alumnos universitarios y está compuesto por ítems de elección forzosa, para discriminar entre las actividades y las creencias preferidas; se califica de acuerdo a la doctrina de Spranger, que clasifica los valores en teóricos, económicos, políticos, estéticos, sociales y religiosos. Generalmente es utilizado como un suplemento de los inventarios de intereses en orientación profesional y ha sido empleado frecuentemente en la psicología social.

* * *

(En la selección de personal, los psicólogos suelen participar en la toma de decisiones vitales acerca de la aceptación, clasificación y pronóstico de los aspirantes. Los intentos de incrementar el número de clasificaciones y predicciones correctas, ha propiciado el desarrollo de un gran número de instrumentos psicométricos, así como límites de corte para las escalas o índices de selección. Dado que los sistemas de selección que no utilizan instrumentos psicométricos permiten establecer algunos requisitos de selección con alguna exactitud, un instrumento psicométrico, para ser eficiente, debe posibilitar un número mayor de decisiones correctas con respecto a la actuación futura de los aspirantes que el que pudiera lograrse en función de criterios basados solamente en la experiencia. Ahora bien, la eficiencia de un gran número de inventarios de personalidad utilizados en selección de personal, resulta difícil y a veces imposible de evaluar debido a las siguientes razones: (26)

a).- Los datos experienciales al respecto de la eficiencia de un empleado son difícilmente sistematizados por las empresas, y la información al respecto es generalmente muy restringida; resulta, por consiguiente, sumamente difícil poder determinar si un instrumento dado produce un número significativamente mayor de decisiones correctas del que sería posible obtener, basándose únicamente en las tasas de éxito de la experiencia previa; sin embargo, no resulta sorprendente, dado el estado actual de la métrica de la personalidad, encontrar que la eficiencia de los llamados inventarios "objetivos" de la personalidad no es significativamente mejor que la producida por criterios experienciales cuando estos datos pueden estimarse consistentemente.

b).- La mayoría de los datos acerca de la distribución de los rasgos de personalidad proporcionan distribuciones que son generalmente insuficientes para evaluar la eficiencia del instrumento en circunstancias donde éstos difieren significativamente de las tasas de éxito, basadas en la experiencia; además de que en muchos casos las muestras utilizadas para establecer la comparación, son demasiado pequeñas para establecer los límites de corte que permiten realizar decisiones adecuadas.

c).- La mayor parte de estos inventarios son publicados sin datos de validaciones transversales; si un instrumento de este tipo se aplica únicamente a los grupos criterio, a partir de los cuales fue desarrollado, su validez y confiabilidad suelen ser espuriamente altos.

d).- Hay una gran carencia de información al respecto del tipo de población en la que un inventario de personalidad puede ser aplicado con eficacia.

e).- La mayor parte de los resultados son publicados solamente en términos de pruebas de significancia para diferencias entre grupos, en vez de señalar el número de decisiones correctas para los individuos dentro de los grupos.

Pudiéramos decir que en la interpretación de un inventario de personalidad y en la evaluación de su efectividad, como instrumento de selección de personal, el psicólogo debiera conocer bastante bien a la población de donde provienen los aspirantes, sus estratos y las tasas de ocurrencia (en la población) de la conducta o rasgo que está tratando de medir. A pesar de las deficiencias en los instrumentos de métrica de la personalidad, y tal vez debido a ello, es afortunadamente difícil encontrar psicólogos que se fíen de una sola escala, por lo general la decisión de aceptar o rechazar un candidato por sus características de personalidad se encuentra basada en la configuración de varias escalas y datos adicionales, aunque resulta evidente que cualquier procedimiento complejo que conlleve un número cualquiera de variables, concluirá en una decisión, misma que tendrá una tasa de éxito determinada por la identificación de los casos seleccionados con el criterio de adecuación al trabajo.

Para los propósitos de esta tesis nos limitaremos a hacer algunos comentarios al respecto de la decisión de aceptar o rechazar un candidato en función de los resultados obtenidos por instrumentos psicométricos, independientemente de los componentes y de la complejidad de tal decisión. (26)

Si existen k condiciones antecedentes bajo las cuales puede ocurrir un evento determinado, y tales condiciones tienen las probabilidades antecedentes de P_1, P_2, \dots, P_k de suceder y la proporción del evento, o de aparición de un determinado rasgo o característica, es p_1, p_2, \dots, p_k dado que el evento ha ocurrido, la probabilidad de que haya surgido sobre una base específica, digamos j , de las condiciones antecedentes estaría dada por:

$$P_j = \frac{P_j p_j}{\sum_{i=1}^k P_i p_i}$$

Supongamos que tenemos 2 categorías diagnósticas: "normales" y "desajustados", y el procedimiento de selección es tal que la probabilidad antecedente de que elijamos un sujeto "normal" es de .90 y la de encontrar un sujeto "desajustado" es de .10. Supongamos también que nuestro instrumento de medición diagnosticara que el 5% del grupo normal son "desajustados" y el 85% del grupo de "desajustados" también lo son. Si eligiéramos al azar un protocolo de la prueba y éste resultara diagnosticado como "desajustado", la probabilidad de que realmente lo fuera sería entonces de:

$$\begin{array}{ll} P_1 = .10 & P_2 = .90 \\ p_1 = .85 & p_2 = .05 \\ P_j = \frac{.10 \quad (.85)}{.10 \quad (.85) + .90 \quad (.05)} = .65 \end{array}$$

Si se hace una práctica de inferencia en la circunstancia de que el protocolo salió del primer grupo, el juicio diagnóstico será correcto a la larga en un 65% de las veces; debe advertirse, sin embargo, que el término diagnóstico es "calificado" correctamente a favor del grupo número uno, ya que hay una diferencia de 80% en la tasa de "desajustados" entre ambos grupos, pero esta considerable disparidad en la tasa del diagnóstico es superada por una tasa base muy baja en este grupo (probabilidad antecedente de escoger un protocolo "desajustado"). De esta manera la inferencia hacia el segundo grupo como origen de protocolos "desajustados" será errónea 35 veces de cada cien. En el proceso de selección de personal la probabilidad de aparición de un rasgo o categoría diagnóstica equivale a la probabilidad antecedente de su tasa base y los protocolos "desajustados" en el primer grupo son la proporción de ciertos positivos, en tanto que esos mismos protocolos son la proporción de falsos positivos. La inspección y manipulación más adecuada para un caso como el anterior, quedaría dada por:

$$P_d = \frac{P p_1}{P p_1 + Q p_2}$$

P_d = Probabilidad de que un individuo esté "desajustado", dado que la puntuación observada en su prueba es positiva.

P = Tasa base de ciertos positivos en la población examinada.

$$Q = 1 - P$$

p_1 = Proporción de "desajustados" identificados por la prueba (tasa de ciertos positivos).

$$q_1 = 1 - p_1$$

p_2 = Proporción de "normales" erróneamente identificados por la prueba como "desajustados" (tasa de falsos positivos).

$$q_2 = 1 - p_2$$

De lo anterior podemos formular el siguiente cuadro:

DEFINICIONES *

DIAGNOSTICO DE LA PRUEBA	DIAGNOSTICO	VERDADERO
	POSITIVO	NEGATIVO
POSITIVOS	P_1 Tasa de ciertos positivos - (proporción de "desajustados" designados como "desajustados").	P_2 Tasa de falsos positivos (proporción de "normales" designados como "desajustados").
NEGATIVOS	q_1 Tasa de falsos negativos -- (proporción de "desajustados" designados como "normales").	q_2 Tasa de ciertos negativos (proporción de "normales" designados como "normales").

1.- Para que un término diagnóstico positivo sea probablemente más verdadero que falso, la proporción de las tasas base de positivos y negativos en la población debe ser mayor que la proporción de falsos positivos y ciertos positivos, esto es:

$$\frac{P}{Q} > \frac{P_2}{P_1}$$

Si esta condición no es satisfecha, la atribución de "desajuste" basada en el inventario, es probablemente más errónea que correcta, aunque el signo usado sea válido (es decir $P_1 \neq P_2$).

En el ejemplo citado anteriormente, el límite de corte identifica el 85% de los sujetos "desajustados" pero también diagnostica erróneamente al 5% de los "normales" evaluados, para que la decisión psicométrica "desajustado", sea más frecuentemente verdadera que falsa, la proporción de los realmente "desajustados" debe ser al menos de 1 a 17 (.059).

En vista de que las tasas base para cualquier conducta dada pueden diferir de una situación a otra, cual sería v.gr.: la postura defensiva en situaciones de exámenes a presión (caso de la selección de personal) vs. evaluaciones solicitadas directamente por el sujeto (caso de la orientación vocacional), un límite de corte inflexible no debe ser defendido para ningún inventario de personalidad. Esta proposición se aplica en lo general a aspectos tales como cla

* Meehl y Rosen: "La probabilidad antecedente y la eficiencia de los signos, los patrones o los puntos de corte psicométrico". Psychological Bulletin, 1955, -52 Págs. 194-216.

sificación, evaluación de rasgos, disimulación, etc. Cuando es pequeña, resulta aconsejable analizar la posibilidad de tratar con una población restringida, dentro de la cual la tasa base del atributo es examinado sea alto.

2.- Si las tasas de base son iguales, la probabilidad de que un diagnóstico positivo sea correcto, es la proporción de la tasa de ciertos positivos entre la suma de las tasas de positivos ciertos y falsos; es decir:

$$P_d = \frac{P_1}{P_1 + P_2} \quad \text{si} \quad P = Q = .5$$

Si la población de nuestro ejemplo estuviera dividida en proporciones --- iguales de "normales" y "desajustados", la decisión de diagnosticar a un sujeto como "desajustado" tendría una probabilidad de .94 de resultar acertada. Este caso especial es a menudo erróneamente generalizado para justificar el uso de pruebas en aquellos casos donde la simetría de la tasa base no existe ($P \neq Q$) lo que contrarresta la diferencia ($P_1 - P_2$) trayendo como consecuencia la paradoja de que decidir sobre la base de más información, puede deteriorar las oportunidades de alcanzar una decisión correcta. El evidente absurdo de tal idea provoca que muchos psicólogos tomen decisiones equívocas como si la posibilidad de discriminación (es decir, que $P_1 \neq P_2$) señalara el procedimiento adecuado para decidir.

Si con los datos del ejemplo supusiéramos que las tasas base son de 50% - para cada grupo ("normales" y "desajustados"), pudiera pensarse que la selección correcta de términos diagnósticos a través del inventario, puede lograrse en un 94% de los casos, lo cual resulta exageradamente alto, ya que utilizando las - tasas base reales (.10 y .90), se mostró que su poder real de discriminación - lograba apenas el 65% de decisiones correctas.

3.- La tasa base de la clase más numerosa, a la que aquí llamaremos positiva, debe ser menor que el cociente de la proporción de ciertos negativos entre la suma de los ciertos negativos más los falsos negativos. Es decir, a menos que $P < q_2/q_1 + q_2$ las decisiones basadas en la prueba tendrían un efecto adverso. Una expresión alternativa estaría dada por $(P/Q) < (q_2/q_1)$ cuando $P > Q$, es decir, - la proporción de la clase más grande a la más pequeña debe ser menor que la proporción de la tasa de ciertos negativos entre la tasa de falsos negativos. Cuando $P < Q$ las condiciones para que la prueba resulte más efectiva, en relación a las tasas base son:

$$Q < \frac{P_1}{P_1 + P_2} \quad \text{y} \quad (Q/P) < (P_1/P_2)$$

4.- El mejoramiento de una escala y la posibilidad de trasladar el límite de corte de manera de incrementar el índice de ciertos positivos en función de un aumento de la tasa base de ciertos positivos, es una función proporcional al incremento de la tasa base de positivos y, análogamente, el incremento de ciertos negativos por el incremento de la tasa de ciertos negativos, es proporcional al aumento de la tasa base de negativos. Es decir, si se mejora el-

poder discriminativo de la prueba, el mejoramiento neto en la tasa de decisiones correctas resulta:

$$H'_T - H_T = \Delta p_1 P + \Delta q_2 Q,$$

donde H_T es la proporción original de aciertos totales, y H'_T es la nueva proporción total de aciertos.

5.- Un corolario de lo anterior es que al trasladar un límite de corte, la decisión mejorará solo si la proporción de mejoría Δp_1 de la tasa de cieros positivos y el empeoramiento Δp_2 de la tasa de falsos negativos exceden la proporción de negativos y positivos reales en la población:

$$\frac{\Delta p_1}{\Delta p_2} > \frac{Q}{P}$$

Así, supongamos que a partir de los datos de nuestro ejemplo, mejoramos la validez del instrumento de manera que ahora detectamos 12% más de "desajustados" que anteriormente, a expensas de solo un incremento del 4% de la tasa de falsos positivos. Esto parecería bastante alentador, sin embargo, el número de "desajustados en la población es solamente el 10%. Entonces puesto que:

$$\begin{array}{ll} \Delta p_1 = .12 & P = .90 \\ \Delta p_2 = .04 & Q = .10 \end{array}$$

Aplicando la fórmula vemos que

$$\frac{.12}{.04} < \frac{.90}{.10}$$

es decir, que la desigualdad requerida no se mantiene y el uso rutinario de este aparente índice de mejoría resultara en un incremento de la proporción de decisiones erróneas.

Como una conclusión derivada de los aspectos señalados, puede mencionarse que la toma de decisiones fundamentada en el uso de instrumentos psicométricos, solo puede ser adoptada sobre la base de elevar al máximo la frecuencia de decisiones correctas. Obviamente llevar a la práctica un procedimiento de selección de personal que considere las variables fundamentales para proporcionar un servicio eficiente, no resulta del todo sencillo en la práctica, dado que independientemente de las limitaciones actuales de los instrumentos psicométricos, existen en toda industria o empresa una serie de restricciones y limitaciones de índole política, administrativa y social, que los psicólogos industriales no pueden eludir o pasar por alto. (9, 26)

5

CAPITULO II

II.- MMPI E IPP: SUS ESCALAS DE VALIDEZ.

La validez es un concepto complejo. En psicometría se ofrecen una multitud de términos que son sustitutos o refinamientos de algún aspecto de la validez, esto incluye algunas alternativas que se consideran de utilidad para algunas aplicaciones particulares, como base confiable para alguna inferencia específica, tal como se presenta en el caso de las saturaciones factoriales o al analizar la varianza de diferentes medidas. En los últimos años han surgido poderosos argumentos para suprimir de la práctica científica todos aquellos términos cuyo significado o connotación tengan un sentido ambiguo y en consecuencia pierdan su precisión o denotación fundamental. Se podría arguir a este respecto que términos tales como "validez" o "confiabilidad" tienen connotaciones equívocas cuando se refieren a pruebas psicológicas. En la práctica debe, sin embargo, considerarse la conveniencia de que las cuestiones a las que se refiera un conjunto de datos, tengan relación con aspectos concretos y bien definidos del problema que se estudia; en el caso de las pruebas psicológicas la cuestión de la validez es básica y omnipresente. A pesar de que después de la segunda guerra mundial la Psicología ha sufrido drásticas transformaciones en su metodología y marcos teóricos de referencia, existen poderosas razones para mantener vigentes conceptos como el de validez.

Al dominar a algunos de los indicadores del MMPI e IPP bajo el rubro de "escalas de validez", se corre el riesgo de ser ambiguo y de caer en equívocos potenciales. En la teoría psicométrica, el término validez ha sido reservado, casi por completo, para designar la propiedad de una escala o instrumento que permita legitimar alguna inferencia psicológica particular que puede ser extraída de los valores generados por la apropiada administración del instrumento. Así, una escala de aptitudes musicales, tendrá una validez alta cuando pronostique con exactitud la habilidad musical de los sujetos, aún cuando éstos no tengan ninguna instrucción musical formal.

La validez, en el sentido especial a que se refieren las escalas objeto de este trabajo, atañe a la aceptabilidad de los perfiles obtenidos de la administración de los inventarios. Cada administración de un inventario suscita la cuestión de la validez del mismo y con frecuencia resulta difícil definirla con exactitud o responder de manera totalmente cierta a este complejo problema. A pesar de que tradicionalmente se ha denominado a estos índices con el nombre de "escalas de validez", parecería más adecuado hablar de la determinación de la utilidad de la prueba en general a partir de un resultado específico (el de tales escalas) así como de la determinación de la aceptabilidad de un grupo particular de datos para un sujeto dado por medio del inventario.

Este capítulo pretende ser, sólo, un antecedente que permita encuadrar esta tesis, por lo que casi resulta innecesario decir que la revisión que aquí se hace del MMPI y del IPP, así como de sus escalas de validez no es, ni pretende ser, ni definitiva, ni exhaustiva, aunque sí necesaria para la comprensión del problema particular de que se trata.

1).- MMPI. Un bosquejo general.

Desde hace veinticinco años comenzaron a publicarse estudios sobre el "Inventario Multifásico de la Personalidad de Minnesota", comunmente llamado-MMPI (Minnesota Multiphasic Personality Inventory). En 1943 la Editorial de la Universidad de Minnesota publicó el material del Inventario y su primer manual. (27)

Existen, desde 1946, un manual adicional y un Inventario en forma de folleto, mismo que en la adaptación para la lengua española realizada por el Dr. Rafael Núñez es utilizado en este estudio.

Aún no existe suficiente información sobre la validez o confiabilidad de las escalas originales. "Espero que ningún psicólogo sobrevalore la validez del MMPI, ya que ésta no provee una predicción precisa de la conducta y, además existen todavía muchas discrepancias en las calificaciones que no han podido ser estimadas convenientemente y, como he dicho, la prueba será también menos satisfactoria en sus resultados si es aplicada en otros países sin que antes haya sido investigada en la forma debida". (Hathaway, 1968). (27)

Hasta el momento, la prueba no reúne los requisitos de validez que serían deseables, pero su empleo se encuentra justificado por las limitaciones metodológicas para la construcción de inventarios de personalidad, que a la fecha siguen restringiendo la evaluación de estos aspectos de la conducta humana.

El criterio de mayor importancia para cualquier test es su validez, pero este no es un concepto tan simple como parece. La validez se refiere siempre a algún fin, por lo que es necesario indagar qué es lo que el test mide antes de preocuparse por conocer qué tan bien lo mide. En el área de la personalidad el problema de las dimensiones involucradas es especialmente importante. Sin embargo, todavía no hay un esquema que permita poner en concordancia las dimensiones fundamentales de la personalidad, por lo que para un test dado lo más importante es considerar los aspectos que valúa. (19)

En relación a lo que los tests tratan de medir, los de personalidad caen dentro de dos grandes categorías: aquéllos que hacen el intento de abarcar la personalidad completa en una dirección sistemática y los que se ocupan por algunos objetivos "ad hoc". (1)

El primer tipo está basado generalmente en estudios factoriales y no puede validarse por procedimientos simples de correlación. El segundo tipo puede ser restringido a una medida única, simplificando el problema de la validez, pero el aspecto principal, es la disponibilidad de un criterio de control para elegir los ítems de la prueba y para evaluar la validez del test.

El MMPI cae dentro del segundo grupo. No pretende proporcionar las dimensiones básicas de la personalidad, pero predice las categorías psiquiátricas comúnmente aceptadas. El MMPI tiene una validación empírica que es la base del tratamiento de la prueba. En la fase actual de la psiquiatría, las categorías del MMPI son las más significativas para un punto de vista diagnóstico, y el problema de la medición se simplifica por la posibilidad de definir el criterio, aunque éste no sea perfectamente confiable, el problema de qué es medido se ve, por lo tanto, simplificado para el MMPI.

Las implicaciones de esta situación no son, sin embargo, siempre bien comprendidas, aunque el MMPI es uno de entre varios tests multidimensionales; algunas personas tienen la idea de que la prueba es útil como un examen general de la personalidad, cosa para la que no fue diseñado. El inventario pone de relieve los posibles estados de incapacidad del sujeto e indica el tipo de desorden mental, pero si la configuración del perfil puede ser significativa cuando ninguno de los puntajes cae fuera del rango normal, éste es un problema al que no nos avocaremos en este trabajo; no obstante la extensa literatura desarrollada alrededor de la prueba, hay poca información sobre estos aspectos.

Debido probablemente a la vaguedad con que son definidos un gran número de desórdenes mentales, el inventario ha visto proliferar una enorme variedad de nuevas escalas; de entre todas ellas, sólo la escala Si ha sido incluida entre las escalas básicas del MMPI.

Esta escala no es, en rigor, una escala clínica pero tiene probablemente mayor valor diagnóstico que la mayoría de las nuevas escalas. El gran número de reactivos de la prueba pudiera hacer pensar que el inventario proporciona un amplio rango de medidas generales pero, en consecuencia, la aplicación de los procedimientos apropiados para derivar tales medidas se haría más difícil, para lo cual se requeriría un análisis factorial de los reactivos, sin embargo, aunque se han hecho numerosos análisis, éstos han sido de escalas individuales o de los puntajes de las escalas.

El estudio de las escalas indica que usualmente involucran factores de segundo orden encontrados en el estudio de la personalidad. Kassebaum, Couch y Slater analizaron 32 escalas derivadas del MMPI y encontraron tres factores de los que los dos primeros parecen ser Fuerza del Yo e Introversión-Extroversión que corresponden con los principales factores encontrados por Eysenck y a los factores de segundo orden de Catell. Hay un acuerdo general respecto a la existencia de estos dos factores pero no acerca de la naturaleza del primero de ellos. Eysenck lo llama Emocionalidad, Catell habla de Ansiedad y Kassebaum, Couch y Slater hablan de Fuerza del Yo. En el tratamiento de pacientes una elección correcta de tales interpretaciones podría ser relevante. Puede ser que cada interpretación sea completamente correcta en base al material disponible, pero surgen serias dudas de que la dificultad real radique de hecho en un error de apreciación de las complejas interrelaciones de las respuestas emocionales. Cuando una batería de pruebas proporcione

una apropiada diferenciación entre Fuerza del Yo, Emocionalidad y Ansiedad, - se tendrá una fundamentación más sólida para establecer una discriminación de tales conceptos. En vista de que no existe acuerdo al respecto de la naturaleza de las dimensiones básicas de la personalidad, la postura que se apoya - estrictamente en aproximaciones empíricas a la prueba, parecería la más razonable. Pero esto puede no ser una justificación para el uso de las nuevas - escalas como sustitutos de las escalas básicas de personalidad y más aún se - pueden suscitar algunas cuestiones inadecuadas, tales como la pertinencia de las mismas escalas originales. La Psicastenia y la Esquizofrenia son apreciadas como entidades distintas aunque ambas aparecen altamente correlacionadas. El manual no proporciona datos al respecto pero Kassebaum y otros encuentran una correlación de .83. La revisión de los datos sobre la confiabilidad sugiere que la varianza específica de estas dos escalas puede ser insignificante. La alta correlación de las escalas F (Validez) y K (Corrección) - con Pt y Es suscita algunas dudas y ha ocasionado que durante mucho tiempo, - no haya sido posible aclarar los conceptos psiquiátricos implicados. (5)

Un ejemplo significativo de cómo pueden presentarse problemas en este renglón, lo dá el estudio de Rosen quien diseñó 5 nuevas escalas psiquiátricas a partir del MMPI. La organización de los reactivos se basó en la capacidad - de éstos para diferenciar sujetos "normales" de aquéllos que presentaban de - una manera u otra, conductas "patológicas", lo cual tiene una gran importancia para propósitos de clasificación diagnóstica.

La escala de reacción conversiva Cr reemplaza a la anterior escala Hi, - pero tiene con ella una correlación de -.24. Rosen argumenta que ésto se explica por un factor común de "normalidad", pero debe señalarse que esta investigación no incluye ningún análisis factorial que permita precisar la relación con tal factor.

Hasta este momento solamente es posible afirmar que los reactivos que diferencian a los histéricos de otros cuadros "anormales", podrían ser interpretados en forma inversa si en la evaluación se incluyen las escalas Hi y Cr. Este es solamente un ejemplo de situaciones que pueden presentarse cuando se utiliza la prueba para cualquier otro propósito que no fuera el original.

La escala de somatización de Rosen (Sm) correlaciona significativamente con la escala Cr (ligeramente por encima del coeficiente de confiabilidad) y muestra también una correlación negativa, aunque insignificante, con Hi. La esquizofrenia paranoide (Pz) correlaciona .81 con Es, lo que parece mostrar - el mejor funcionamiento de Pz con respecto a la escala original; pero la escala de reacción depresiva (Dr) no muestra una correlación significativa con la escala D. Para la escala de reacción de ansiedad (Ar) no hay equivalente entre las escalas básicas.

Desafortunadamente Rosen eligió categorías psiquiátricas nuevas, utilizando un grupo de sujetos normales, para el establecimiento de los criterios diferenciales; resultaría interesante saber qué resultados hubiera obtenido de haber conservado el grupo psiquiátrico original para la validación de las escalas.

El ejemplo anterior nos lleva directamente a la cuestión de la validez para pruebas de este tipo misma que pudiera ser expresada simplemente como un conjunto de coeficientes de correlación. Curiosamente, los autores no optaron por el uso de este método, sino que se apoyaron en aspectos tales como la capacidad de las escalas para predecir una nueva categoría diagnóstica que presuntamente tiene éxito en el 60% de los casos, aunque dos grandes aspectos derivados de esta posición no resultan del todo claros. En primer lugar, resulta incierto hasta qué punto un diagnóstico psiquiátrico se ve influido por el resultado de la prueba. Si dos categorías psiquiátricas presentan una alta correlación, como es el caso de las escalas Pt y Es ($r=.84$), el sistema de clasificación de la prueba podría jugar un importante papel, dando la apariencia de un resultado adecuado, incluso más allá de lo que podría ocurrir si las escalas no estuvieran significativamente relacionadas.

En segundo lugar, debe apreciarse que la población de donde fue derivada la validación del inventario, está ya seleccionada por su necesidad de tratamiento psiquiátrico y la prueba se ocupa simplemente de clasificarla. Parece justificado el empleo de la prueba para estos fines, pero desafortunadamente resulta también muy fácil para el usuario del inventario centrar su atención en los aspectos calificados por la prueba, perdiendo de vista los datos de validez de los mismos. Si bien los autores advierten que, en una población promedio, la mayoría de las desviaciones de los perfiles pueden referirse a características de personas normales, más que a problemas que requieran tratamiento psiquiátrico, no por ello el manejo rutinario de la prueba toma siempre en cuenta tal advertencia, máxime cuando las calificaciones son obtenidas mecánicamente y el examinador confía en el acierto de su diagnóstico. Consideramos que debería existir mayor interés en el estudio de la validez de la prueba, antes de pretender establecer diagnósticos acertados entre sujetos de la población normal. Esta información es absolutamente necesaria cuando la prueba se utiliza con propósitos generales de clasificación, como de hecho sucede frecuentemente, y el manual bien podría proporcionar esta información. El trabajo de Rosen mencionado antes, independientemente de los defectos que pueda tener, enfatiza la necesidad de estudiar estas áreas. (5)

Otro aspecto planteado por los autores que justifica un comentario adicional, es el hecho de que las personas clasificadas dentro de una cierta categoría pueden no tener necesariamente su más alto puntaje en esa categoría; por ejemplo, un esquizofrénico puede obtener un puntaje tan alto para depresión como para esquizofrenia. Esto es bastante comprensible, pero constituye otra trampa para el examinador no entrenado, además de suscitar algunas otras dudas acerca de los principios de clasificación psiquiátrica y específicamente surge

la pregunta de si el MMPI no comete el error de agrupar un número de términos psiquiátricos sin una consideración adecuada de su "status" lógico. Es de suponerse que masculinidad - feminidad o depresión tienen un "status" diferente al de esquizofrenia. Este es un ejemplo de los aspectos de validez que deben ser considerados en relación con la prueba. (5)

Un estudio reciente de estos aspectos realizado por Exner y colaboradores, muestra que un grupo de sujetos a los que se les pidió que deliberadamente se mostraran como anormales, pero no tanto como para requerir hospitalización, reportaron elevaciones significativas en todas las escalas, excepto en L. Los intentos por dar buena impresión fueron poco exitosos; las diferencias más importantes (significativas estadísticamente) se obtuvieron en las escalas L, F, K y Dp, en ese orden. El rango de puntajes de esas categorías, no obstante, mostró demasiada superposición entre los esfuerzos por mentir y por contestar honestamente como para poder establecer una discriminación individual útil. Surgen algunas dudas adicionales acerca del uso de las escalas de validez al considerar algunas intercorrelaciones del estudio de Rosen.

En ese estudio F tiene una correlación de .82 con Pz, y K una correlación de .72 con Cr. En vista de las probables confiabilidades de estas medidas, surgen serias dudas acerca de la cantidad de varianza referida a la función de corrección. Hay una afluencia constante de estudios referidos a esta prueba, pero la mayoría de ellos supone, más que demuestra, la validez del test. Resultan interesantes algunas comparaciones con otras pruebas. En particular es importante notar que las medidas de masculinidad-feminidad del MMPI, el Guilford-Zimmerman y el Strong no correlacionan entre sí, y probablemente representan diferentes conceptos. (5)

Taylor y Mc Lachlan encontraron que la escala Mf del MMPI discriminaba satisfactoriamente un grupo de transvestistas (homosexuales activos) de los que no lo son, cuestión que puede, además, ser psiquiátricamente significativa, pero los autores no señalan más datos de la aplicabilidad de otras medidas de masculinidad-feminidad en este grupo.

Todo esto parece indicar que, aunque el MMPI es un excelente instrumento para clínicos entrenados que han dominado muchos de los intrincados aspectos de su manejo, puede, sin embargo, ser altamente peligroso en manos del usuario que carece de la fundamentación teórica adecuada y puede verlo como "la varita mágica" para el establecimiento de diagnósticos, y considerarlo libre de problemas de calificación subjetiva. (5)

* * *

Los datos del Manual y del Atlas se apoyan en más de 1000 referencias, una formidable cantidad de material que cubre casi todos los aspectos concebibles de la construcción, confiabilidad, validez y uso de la prueba, se encuentran a disposición del usuario. La proliferación de escalas (más de 200 en-

las 2 últimas décadas), presuntamente evalúan diferencialmente muchas clases importantes de conducta, y va más allá de los más optimistas sueños que se hayan formulado los autores cuando eligieron el calificativo de "multifásico" para su inventario. Los niveles de validez de muchas de estas escalas permanecen en la oscuridad, pero aquéllas que han resistido las validaciones cruzadas (por ejemplo, la escala de Fuerza del Yo de Barron, la de Prognosis para Tratamiento por Electrochoque de Feldman y las escalas de los factores-A y R de Welsh han aumentado considerablemente el poder de esta prueba y si no han rebasado completamente los objetivos de los autores en el terreno de medir todas las fases importantes de la personalidad, al menos han proporcionado un avance en este sentido. (5)

Si tuvieran que resumirse todas las investigaciones acerca de la permanente cuestión de la validez diagnóstica de la prueba, tal sumario podría mostrar algo como lo siguiente: (5)

El MMPI puede diferenciar bastante bien entre aquellos sujetos que sufren o que no sufren problemas emocionales o de ajuste en una amplia variedad de categorías y puede, asimismo, servir como un excelente instrumento de clasificación; sin embargo el uso de este instrumento para tales propósitos requiere de decisiones en relación a aspectos tales como costos, estrategias, presupuesto, utilidad, y, como consecuencia, el usuario potencial deberá familiarizarse con tales aspectos antes de adoptar el uso del instrumento.

Bien puede ser que para problemas muy específicos un instrumento simple sea suficiente; aunque no puede refutarse el valor del MMPI para diferenciar entre sujetos que provienen de poblaciones normales o anormales, existen serias dudas al respecto de si la prueba es lo suficientemente sensible para establecer una adecuada clasificación diagnóstica de los sujetos anormales.

Problemas neuróticos o psicóticos, desórdenes de carácter y posiblemente algunos trastornos psicómaticos, como grandes grupos diagnósticos, pueden ser separados de manera confiable, sino en forma fina, si por lo menos en grandes categorías. Las distinciones precisas dentro de cada uno de estos grupos nosológicos han sido, sin embargo, improductivas. Uno de los estudios más importantes en el área de los diagnósticos diferenciales, fue realizado por Dahlstrom y Meehl, quienes utilizaron las reglas de diferenciación del perfil por medio de pasos sucesivos para clasificar perfiles de neuróticos, psicóticos o indeterminados. En un estudio de validación cruzada de 988 casos seleccionados de 8 grupos diagnósticos, se logró un 76% de aciertos de entre un grupo cuyo diagnóstico ya había sido determinado; sin embargo, no existe otro instrumento de esta clase que haya sido totalmente investigado; como una prueba general de personalidad, el MMPI tiene cierto número de puntos débiles, de los cuales no es el menos importante el que se refiere a la exclusión o atonía de algunas variables consideradas importantes para las actuales teorías de la personalidad.

Desafortunadamente otras pruebas que pueden considerarse competidoras del MMPI tienen iguales o mayores limitaciones; como instrumento clínico usado con juntamente con otras pruebas y, como balanza para las diferencias, el MMPI aporta una contribución definida que hasta el momento no ha sido igualada. Sin embargo, el MMPI es aún deficiente en evaluaciones de la personalidad dentro de rangos normales de ajuste.

La siguiente sección de este trabajo contiene una revisión de las escalas de validez del MMPI, de las cuales se excluye el primer indicador de validez, - la escala de frases omitidas, que no ha sido considerada debido, en primer término, a su naturaleza y, en segundo lugar, a que en la situación específica - que nos ocupa, los sujetos regularmente proporcionan un número poco significativo de frases sin contestar (2.35 en promedio, N=1950).

E S C A L A L

El segundo indicador de validez del perfil regular del MMPI es la escala de quince ítems, diseñada para identificar esfuerzos intencionales o deliberados para evadir responder el inventario de manera franca y honesta. Esta escala es denominada escala de mentira, o simplemente escala L. Hathaway y McKinley confeccionaron esta escala después de que Hartshorne y May (1928) habían desarrollado una escala similar en sus estudios de mentira entre niños de edad escolar. En su trabajo, Hartshorne y May encontraron que hay cierto número de debilidades comunes o defectos personales, los que sus sujetos indicaban que son generalmente malos pero que reconocían también como verdaderos acerca de sí mismos. Estos autores encontraron, sin embargo, que algunos sujetos intentan, sistemáticamente, aparecer ante los demás de la mejor manera posible, negando tales atributos negativos aún cuando sea totalmente verosímil y probable que tales características sean tan aplicables a ellos como a cualquier otro. Utilizando esta aproximación general, Hathaway y McKinley reestructuraron algunos de los mismos contenidos y escribieron nuevos reactivos en el mismo sentido, hasta producir quince ítems que fueran consistentes con respecto al conjunto de las restantes declaraciones del MMPI, tanto en formato como por las palabras empleadas. Estos ítems fueron introducidos como un grupo dentro del conjunto y calificado arbitrariamente como un índice de la tendencia a falsear al responder la prueba. Los ítems de la escala L se presentan en la tabla siguiente, junto con la indicación de los ítems que se superponen con otras escalas básicas, así como la dirección crítica de los mismos: (11)

FALSA: L

- 45.- No siempre digo la verdad.
- 60.- No leo todos los editoriales del periódico diariamente.
- 90.- De vez en cuando dejo para mañana lo que debiera hacer hoy.
- 135.- Si pudiera entrar a un cine sin pagar y que estuviera seguro de no ser visto, probablemente lo haría.
- 150.- Prefiero ganar a perder en un juego.
- 165.- Me gusta conocer gente de importancia porque eso me hace sentir importante.
- 195.- No me agradan todas las personas que conozco.
- 225.- A veces murmuro o chismeo un poco de la gente.
- 255.- En las elecciones algunas veces voto por candidatos acerca de quienes conozco muy poco.

FALSA: LF

75.- Algunas veces me enojo.

FALSA: L2

285.- A veces me río de chistes sucios.

FALSA: LK23

30.- A veces siento deseos de maldecir.

FALSA: L59

120.- Mis modales en la mesa no son tan correctos en casa como cuando salgo a comer fuera en compañía de otros.

FALSA: L

VERDADERA: (678)

15(314).- De vez en cuando pienso en cosas demasiado malas para hablar de ellas.

FALSA: L9

105.- Algunas veces, cuando no me siento bien, estoy malhumorado.

NOTA:- En esta y en las subsecuentes tablas que muestran los ítems que integran las escalas, un asterisco (*) antes de un ítem señala que éste es un reactivo de corrección en la escala K.

La inspección de los ítems listados en la tabla anterior mostrará el contenido referente a la negación de la agresión, de malos pensamientos, de la debilidad de carácter o resolución, de la pobreza de auto-control, de los prejuicios, así como deshonestidad. Los datos de la tabla revelan también que todos los ítems son calificados en este sentido si son contestados como falsos. Esta última característica de la escala ha sido criticada, aduciéndose que es excesivamente susceptible a la aquiescencia de quienes contestan el test.

Muchos de los ítems corresponden exclusivamente a la escala L, pero algunos otros aparecen también en una o más de las otras escalas del perfil, gene

ralmente en la misma dirección crítica. Sin embargo, la mayor superposición es solamente de dos ítems para las escalas 2 y 9, los cuales no limitan en grado importante el rango de calificación o de covarianza.

Los estudios e investigaciones sobre la escala L han indicado generalmente que los supuestos hechos en la integración de los reactivos que conforman la escala fueron esencialmente correctos. Los datos de los grupos normativos de Minnesota y de estudios tales como los de Gravitz (1970) han mostrado que la mayoría de los ítems de la escala L son contestados en dirección no defensiva por un mínimo de 50% de los sujetos normales.

Los defectos personales señalados en los ítems 15, 135, 165 y 255, no son, sin embargo, reconocidos por la mayoría de los estudiantes de preparatoria (college) de Minnesota o por los adultos de la muestra de aspirantes a empleo utilizada por Gravitz. Algunos de estos ítems están enfocados a atributos conductuales que pueden emerger precisamente en adultos jóvenes y en consecuencia ser poco apropiados para sus auto-descripciones comunes y corrientes, v.gr.: el escuchar una conversación de manera automática y emitir opiniones acerca de los tópicos tratados.

Estos estudios señalan también que los ítems 60, 75, 90 y 285 (para sujetos masculinos) raramente son contestados en dirección crítica. Estos reactivos no contribuyen de manera apreciable a la utilidad de la varianza de la escala L y técnicamente reúnen los requisitos para ser incluidos en la escala F (actualmente sólo el ítem 75 aparece en tal escala), como se aclara en el capítulo referente a la escala F. Otro aspecto digno de mención de los estudios de las características de los ítems que componen la escala L: muchos adultos omiten contestar estos ítems. Esto no es cierto, sin embargo, para sujetos masculinos del bachillerato; aunque tal vez las mismas razones señaladas arriba, referentes a las omisiones al responder defensivamente algunos de estos ítems, sean aplicables a algunos sujetos preparatorianos que encuentran difíciles o imposibles de contestar varios de los ítems señalados, tanto en dirección de falso como de verdadero. Es altamente probable que las pautas de responsabilidad y compromisos sean suficientemente diferentes para estudiantes que aún son solteros, que no tienen capacidad legal para participar totalmente en muchos de los aspectos de la ciudadanía, no trabajan todavía en sus carreras profesionales, las que generalmente son para hombres y mujeres casados, en un país (Estados Unidos) que engendra importantes diferencias en el marco de referencia con el que los sujetos se aproximan a una prueba como el MMPI. Por medio de los ítems de la escala L, pueden apreciarse muchos aspectos de la conducta producida por la experiencia o patrones de vida de los post-adolescentes en nuestra cultura, considerablemente diferentes de la dirección en que la conducta se ve influenciada por los compromisos, preocupaciones y problemas encarados por los adultos maduros. (10, 11, 20)

E S C A L A F

El tercer indicador de validez rutinariamente registrado en el perfil estándar del MMPI, es la escala F. Esta escala ha sido designada indirectamente como escala de frecuencia (o infrecuencia), escala de confusión y algunas veces simplemente como escala de validez.

Ha sido diseñada para detectar respuestas desusuales o formas atípicas de responder a los enunciados de la prueba. Investigaciones recientes provenientes de la hipótesis de la desviación de Berg emparentada directamente con la investigación de las desviaciones de conducta por la prueba, han confirmado la presencia de muchos factores psicológicos diferentes que llevan a modos desusados de aprobar los ítems. Las primeras aproximaciones de Hathaway y McKinley al respecto de las diversas fuentes de invalidez de los puntajes del inventario han sido ampliamente confirmadas. Estos estudios han ayudado a aclarar muchas implicaciones de la personalidad por altos puntajes en la escala F.

Hathaway y McKinley han escogido los sesenta y cuatro ítems que componen la escala F simplemente en base a la distribución asimétrica de las respuestas de los hombres y mujeres que integraron los grupos normativos del Minnesota; aquellos ítems en los que no más del 10% de tales sujetos adultos (y con frecuencia menos del 5%) respondieron en una dirección particular. Una respuesta significativa a estos ítems es aquella que está dada en la dirección infrecuente. Con la excepción de las respuestas al ítem 112, hay una gran consistencia entre los adultos que normaron el Minnesota, por cuanto toca a las convicciones, actitudes, sentimientos y experiencia expresadas en estos enunciados. Generalmente, el contenido es inequívoco en sus alusiones. (Harris y Baxter, 1965). Como puede verse en la lista de ítems, el contenido es también de gran diversidad, fluctuando de sensaciones grotescas, pensamientos extraños y experiencias peculiares a sentimientos de enajenación y aislamiento de los miembros de su familia, de otros o de instituciones sociales, con actitudes atípicas con respecto a las leyes, la religión, o la autoridad y a ciertas convicciones, expectativas y auto descripciones inverosímiles o contradictorias. Bajos puntajes en este grupo de ítems, son, en consecuencia, indicadores de conformidad general con los patrones de respuestas de la población estándar, sobre un amplio rango de contenido experiencial. Por el contrario, respuestas a un gran número de estos ítems en dirección crítica, reflejan una desviación del grupo normativo en cierto número de áreas con diferente contenido. (Comrey, 1958).

ESCALA F: COMPOSICION, DIRECCION CRITICA Y SUPERPOSICION CON
OTRAS ESCALAS BASICAS.

C I E R T O: F.

- 14.- Tengo diarrea una vez al mes o más frecuentemente.
- 31.- Me dan pesadillas con mucha frecuencia.
- 34.- Tengo tos la mayor parte del tiempo.
- 48.- Cuando estoy con gente me molesta el oír cosas muy extrañas.
- 49.- Sería mejor si casi todas las leyes fueran descartadas.
- 50.- Mi alma a veces abandona mi cuerpo.
- 53.- Un sacerdote puede curar enfermedades rezando y poniendo las manos -
sobre la cabeza de usted.
- 56.- Cuando muchacho (a) me suspendieron de la escuela una o más veces -
por hacer travesuras.
- 66.- Veo cosas, animales o gente a mi alrededor que otros no ven.
- 85.- Algunas veces me siento tan atraído por artículos personales de otros,
como calzado, guantes, etc., que quiero tocarlos o robarlos aunque -
no haga uso de ellos.
- 139.- Algunas veces siento el impulso de herirme o herir a otros.
- 146.- Me siento impulsado hacia la vida errante y nunca me siento feliz a-
menos que esté viajando de un lado a otro.
- 184.- Frecuentemente oigo voces sin saber de donde vienen.
- 197.- Alguien ha estado tratando de robarme.
- 200.- Hay personas que quieren apoderarse de mis pensamientos o ideas.
- 205.- Algunas veces me ha sido imposible evitar el robar o llevarme algo -
de una tienda.
- 206.- Soy muy religioso (más que la mayoría de la gente).
- 209.- Creo que mis pecados son imperdonables.
- 211.- Puedo dormir de día pero no de noche.
- 218.- No me molesta mucho el ver sufrir a los animales.
- 227.- Me han dicho que camino dormido.

- 246.- Con frecuencia me salen manchas rojas en el cuello.
- 247.- Tengo motivos para sentirme celoso de uno o más miembros de mi familia.
- 252.- A nadie le importa mucho lo que le sucede a usted.
- 256.- La única parte interesante del periódico es la página cómica.
- 269.- Con facilidad puedo infundirle miedo a otros y a veces lo hago por diversión.
- 286.- Nunca estoy tan contento como cuando estoy solo.

F A L S A: F

- 54.- Le agrado a la mayor parte de la gente que me conoce.
- 83.- Cualquier persona capacitada y dispuesta a trabajar fuerte tiene buenas posibilidades de obtener éxito.
- 113.- Creo que la ley debe hacerse cumplir.
- 169.- No tengo miedo de manejar dinero.
- 185.- Aparentemente oigo tan bien como la mayoría de las personas.
- 199.- Se debe enseñar a los niños la información básica sobre la vida sexual.
- 257.- Por lo general espero tener éxito en las cosas que hago.
- 258.- Creo que hay un Dios.

F A L S A: LF

- 75.- Algunas veces me enojo.

F A L S A: KF2

- 272.- A veces estoy lleno de energía.

VERDADERA: F 123

- 23 (288).- Sufro de ataques de náusea y de vómito.

VERDADERA: F 4

42.- A mi familia no le gusta el trabajo que he escogido (o el trabajo - que pienso escoger para el resto de mi vida).

215.- He bebido alcohol en exceso.

245.- Mis padres y familiares me encuentran más defectos de los que debieran.

VERDADERA: F 468

35 (331).- Si la gente no la hubiera cogido conmigo yo hubiera tenido mucho más éxito.

FALSA: F 48

20 (310).- Mi vida sexual es satisfactoria.

FALSA: F 5

112.- Con frecuencia siento la necesidad de luchar por lo que creo que es justo.

115.- Creo en otra vida después de ésta.

VERDADERA: F 6

27.- Los espíritus malos se posesionan de mí a veces.

123.- Creo que me están siguiendo.

151.- Alguien ha estado tratando de envenenarme.

275.- Alguien controla mi mente.

293.- Alguien ha tratado de influir en mi mente.

FALSA: F 7

164.- Me gusta leer y estudiar acerca de las cosas en que estoy trabajando.

VERDADERA: F 8

40.- La mayor parte del tiempo preferiría soñar despierto antes que hacer cualquier otra cosa.

168.- Mi mente no está muy bien.

210.- Todo me sabe igual.

FALSA: F 8

- 17.- Mi padre fue un buen hombre.
- 65.- Yo quise a mi padre.
- 177.- Mi madre fue una buena mujer.
- 196.- Me gusta visitar lugares donde nunca he estado.
- 220.- Yo quise a mi madre.
- 276.- Me gustan los niños.

VERDADERA: F 89

- 156.- He tenido épocas durante las cuales he hecho cosas que luego no he recordado haber hecho.

La respuesta de algún ítem de la escala F en la dirección crítica puede, - por supuesto, ser altamente probable y comprensible si el reactivo se refiere - a algún aspecto especial de la vida personal o problema particular con que el - sujeto haya tropezado. Veintiún ítems de la escala F aparecen en la lista de - Items Críticos de Grayson, un conjunto de treinta y ocho enumerados que pueden - tener algún significado patológico especial en las circunstancias vitales de un - sujeto. Gravitz (1968) ha mostrado evidencia adicional de que estos ítems son - raramente contestados en la dirección crítica por hombres o mujeres en una mues - tra de sujetos normales. No obstante, respuestas a uno o varios de estos ítems - pueden reflejar válidamente problemas importantes de la situación personal que - requiera de atención especial (ataques, alcoholismo, evasiones, tumores del ló - bulo temporal, etc.) (11)

En el empleo de los ítems de la escala F como un conjunto de respuestas - raras, el clínico no debe perder de vista la posibilidad de que una o todas res - puestas a estos enunciados de algunos de los sujetos que son sometidos a la - prueba sean totalmente ciertos, clínicamente relevantes y dignos de investiga - ción especial.

La utilidad de la escala F como un indicador válido de distorsión de la - prueba con respecto al conjunto de factores que pueden manejarse en una prueba - determinada radica en que las escalas clínicas y los altos puntajes de F aparez - can, simultáneamente, en configuraciones desusadas. Como consecuencia, la in - terpretación completa del MMPI, requiere que las elevaciones de la escala F sean - previamente examinadas, dadas las implicaciones al respecto de la aceptabilidad - de cada protocolo, antes de examinar cualquier implicación adicional que pueda - atribuirse al estado clínico o a las características de personalidad del sujeto. - Esta admonición es particularmente importante en encuadres clínicos en los que - un extenso número de clientes reciben el servicio y cuando los procedimientos - normales requieren que el MMPI se utilice para proporcionar resúmenes prelimina

res de sujetos antes de que se tenga un amplio conocimiento acerca de ellos. Cuando más se conozca al respecto del estado actual de un sujeto, tanto más -- las interpretaciones de la elevación de la escala F pueden afinar la validez-- del protocolo.

Un gran número de factores pueden interferir con la atención y cuidado -- con que un sujeto toma la prueba o responde los enunciados individuales de un cuestionario de personalidad. El sujeto puede ser incapaz de comprender los ítems con claridad suficiente si no los lee correctamente, ya sea por tener limitaciones más o menos serias en su agudeza visual, o porque las condiciones físicas en el área de trabajo de que se dispone, sean inadecuadas para sustentar la lectura y respuesta del inventario (iluminación, formato o impresión del cuadernillo o de la hoja de respuestas, reflejos, ruidos o distracciones.

Algunos de estos problemas pueden afectar seriamente la ejecución de la - prueba. En particular, estos factores pueden conducir a una elevación del - puntaje en F como consecuencia de que el sujeto no perciba con claridad los - ítems del test, o bien que los interprete en sentido opuesto al que los res-- ponde.

Si la forma en que el sujeto percibe los enunciados de la prueba es clara e inequívoca, puede ser, sin embargo, posible que pueda no ser capaz de leer o comprender el contenido de los ítems en un grado suficiente como para respon-- derlos significativamente. Sujetos con deficiente capacidad de lectura, personas clasificadas como fronterizas o deficientes mentales, o quienes tienen - un limitado nivel de escolaridad (sexto grado o menos), pueden no comprender - algunos de los ítems del MMPI. En consecuencia, estos sujetos pueden anotar-- sus respuestas sin comprender del todo lo que dicen acerca de sí mismos. Pues to que algunos de los ítems más difíciles del MMPI aparecen en la escala F, - una deficiente capacidad de lectura del sujeto resulta terreno fértil para co-- sechar elevados puntajes en esta escala.

Las investigaciones han puesto, también, de manifiesto que un adecuado ni vel de comprensión en la lectura, si bien necesario, puede no ser suficiente - para una interpretación adecuada de los ítems de la prueba. Un sujeto con - buena capacidad para leer y escribir, con antecedentes bilingües, y en quien - el castellano haya sido aprendido como segunda lengua, que puede ser capaz de leer los ítems sin tener necesariamente el requisito de familiaridad con el - idioma, no comprenderá claramente las connotaciones emocionales de las frases de las pruebas, o bien será incapaz de modificar o interpretar los enunciados en el mismo sentido en que lo hizo el grupo normativo. Una parte importante de la interpretación de los ítems de la prueba, parece ser una cierta habilidad pa ra participar en los aspectos comunes de la cultura de los grupos originales - estudiados por los autores de la prueba. Muchos sujetos que carecen del meollo esencial de estos patrones culturales muestran atributos atípicos y auto-des-- cripciones desusadas al responder los ítems de la escala F; sin embargo, algu-- nas personas con tales características pueden no ser incluidas en esta situa-- ción. Esto es, - - - - -

aunque sus patrones de respuesta verbal aparecen lo bastante anómalos como para afectar los puntajes de las otras escalas de la prueba, no alteran la escala a un grado tal que descubra este aspecto particular del MMPI. Para este propósito son necesarias las comprobaciones especiales de validez. (Ver la discusión sobre variables moderadoras Manual: Págs. 148 a 154). (11)

Otra fuente de dificultad al contestar la prueba es una extensa desorganización de la personalidad proveniente de reacciones psicóticas agudas. Aunque pacientes con problemática de esta naturaleza puedan, generalmente, ser capaces de leer o interpretar los reactivos de la prueba antes de su crisis conductal y pueden recuperar tal aptitud después de algún grado de remisión, durante la fase aguda se encuentran tan fuera del contacto de los estímulos circundantes que las instrucciones son incomprensibles, las proposiciones carecen de sentido o las hojas de respuesta resultan tan confusas que son incapaces de organizar sus respuestas en un protocolo utilizable. Como resultado, estos sujetos pueden confundir Ciertos y Falsos, poner dos marcas en un mismo ítem, perder la numeración y señalar ítems equivocados, dando como resultado un protocolo incalificable. Como se señala en la primera parte del Manual, los pacientes en estado de confusión pueden ser capaces de arreglárselas con la forma de tarjetas del MMPI o bien puede encontrarse que una administración con cinta grabada se encuentra dentro de sus posibilidades. Las personas que padecen de reacciones tóxicas, aquéllos que se encuentran bajo el efecto de las drogas o pasan por una fase aguda de reacción esquizofrénica indiferenciada, hipermanía o reacciones de delirio maniaco, o bien una acentuada reacción depresiva; asimismo aquéllos que son sometidos a la prueba poco después de un tratamiento por choques eléctricos, pueden ser incapaces de ofrecer un perfil interpretable. Puntajes elevados en la escala F pueden indicar la inhabilidad para discernir el contenido o significado de los ítems de la prueba y el hecho de no poder responder diferencialmente a las frases del test. Es digno de mención que la habilidad para responder a los ítems del MMPI puede ser recuperada una vez que los sujetos son capaces de mostrar habilidades interpersonales adecuadas o suficiente autocontrol para mostrar el impulso a abandonar la remisión clínica. Muchos de estos pacientes pueden ser capaces de dar protocolos válidos del MMPI mientras son declarados socialmente irresponsables y reclusos o bien mientras el habla espontánea es incoherente o inarticulada.

Con una prueba como el MMPI, estos indicadores conductuales gruesos pueden no ser guías confiables de la capacidad de un sujeto para resolver la prueba.

Otro aspecto del estado clínico del sujeto que puede originar dificultades en la situación de prueba se presenta por la introducción de orientaciones especiales al respecto de la forma de tomar la prueba o componendas que sean una viación a las instrucciones generales de la prueba. Marks y Seeman (1963), por ejemplo, han reportado como una paciente puede imaginar significados especiales de la sesión de prueba y del empleo del MMPI, no como una oportunidad para una cuidadosa autodescripción de su situación normal y manera de ver las cosas, sino por el contrario como un clamor de ayuda o una súplica para recibir atención especial. Esto se hace particularmente probable en posturas exhibicionistas en las que el paciente pueda sentir que, a menos que dramatice su condición, no recibirá atención adecuada. También cuando el sujeto está descargan-

do deliberadamente, por no tener oportunidad de expresar de alguna otra manera la falta de preparación que siente para tomar decisiones con respecto a la prueba o expresar su opinión respecto de los ítems. El perfil resultante puede - mostrar puntajes muy altos y casi uniformemente elevados, en donde también la escala F aparece notoriamente alta. En este caso puede ser que otros grupos de respuestas funcionen en forma similar durante la prueba a los que aparecen cuando el sujeto responde a toda la prueba como cierto o como falso o a algunos ciclos alterados de contestaciones de este tipo, dando respuestas socialmente inadecuadas o atípicas a cada ítem, admitiendo cada síntoma o dificultad sugerido o pintando un cuadro muy estereotipado de competencia y eficacia social. Muchas de estas formas de aproximación a la prueba generan patrones reconocibles de configuración de las escalas que frecuentemente involucran los puntajes de la escala F.

Otra dificultad general para la obtención de un perfil válido y utilizable de un sujeto, puede asimismo aparecer como una elevación de la escala F. Personas que se aproximan a la prueba con hostilidad, resentimiento o accediendo sólo superficialmente a la tarea de resolver la prueba, pueden ser tan negligentes que incurran en frecuentes errores mecánicos al anotar sus respuestas en la hoja respectiva. Puesto que la escala F es larga y los ítems se encuentran dispuestos a través de toda la prueba, una pobre cooperación y falta de atención puede detectarse al responder a los ítems de la escala en la dirección crítica. En la forma de tarjetas debe advertirse también que algunos de los errores mecánicos del protocolo corren al cargo del administrador o calificador de la prueba. Los errores mecánicos de parte de quien califica la prueba pueden en esta forma distorsionar el puntaje de la prueba, y las elevaciones de la escala F pueden muy bien revelar esta clase de errores (por ejemplo, cuando las tarjetas de Cierto y de Falso están puestas, inadvertidamente, al revés). No todos estos aspectos pueden, sin embargo, ser plenamente identificables por una dramática elevación de la escala F. Algunas veces el sujeto comienza la prueba con razonable atención y cuidado, para al cabo de un rato comenzar a fastidiarse y distraerse. Si se usa la forma de cuadernillos, una forma de comprobar esta posibilidad es examinando la proporción de ítems de la escala marcados con al dirección crítica en los primeros cuarenta ítems (alrededor de los primeros doscientos ítems del cuadernillo) y los últimos veinticuatro (los que aparecen en el siguiente centenar de reactivos).

Hay también tres ítems de la escala F que se repiten (primero como 20, 23 y 35 y después como 310, 288 y 331); estos ítems pueden ser revisados para ver si el sujeto los contestó en la misma forma en cada ocasión. Adicionalmente, puede también servir de ayuda al examinar todos los ítems repetidos de la prueba como una forma de comprobar esta fuente de error, más que de restringir la atención a un subgrupo de reactivos aislados de la escala F. (11)

E S C A L A K

Mientras que los tres primeros índices de validez (? , L y F) fueron introducidos desde la publicación del MMPI, una cuarta medida, la escala K, fue -- agregada después que la experiencia con el complemento inicial de validez fue -- comprobado en cierto número de grupos clínicos diferentes. Tal experiencia -- demostró que la inclusión de los indicadores previos de validez funcionaban -- fundamentalmente para detectar aspectos gruesos de la invalidez del protocolo, -- permitiendo que muchas formas importantes de distorsionar la prueba fueran i-- -- dentificables. En adición a este desarrollo de la experiencia clínica con el MMPI, algunos estudios empíricos de la falsificación y asimetría de la prueba -- han sido efectuados en Minnesota y en otros lugares, los que han servido para -- poner de manifiesto tanto las ventajas como los puntos débiles de las escalas -- de validez de que se disponía. La investigación encaminada a desarrollar la -- escala K fue consagrada a incrementar la sensibilidad de los índices de vali-- -- dez de la prueba identificando el impacto de una instrumentación sutil que au-- -- mentara o disminuyera el valor de los factores evaluados, así como a proporció -- nar un elemento con significado estadístico para la corrección de los valores -- de las escalas clínicas y que compensara por sí mismo los efectos de estos fac -- tores en el perfil clínico.

En el esquema que se muestra adelante, aparecen las dos clases de errores de que se han ocupado los autores de la prueba, en relación con la aplicación -- clínica del MMPI en varios grupos diagnósticos.

Perfil del MMPI	ESTADO CLINICO	
	Normal	Anormal
Anormal	Falso Positivo	Cierto Positivo
Normal	Cierto Negativo	Falso Negativo.

Tomando el perfil como un todo, puede considerarse que tiene elevaciones -- suficientes (en una o más de las escalas componentes) como para poder determi -- nar que el sujeto muestra un nivel significativo de psicopatología, o bien que -- el perfil puede considerarse dentro de los límites normales. Si esta dicoto -- mía de los perfiles en configuraciones normales y anormales se evalúa también -- por medio de la evidencia del estado clínico actual de cada persona que ha rea -- lizado la prueba, esto es, por los métodos corrientes distintos del MMPI, de -- evidenciar psicopatología o bien, se demuestra el funcionamiento normal, se ge -- nera el cuadrante de clasificación que aparece arriba. La intención general -- en la derivación y aplicación de la escala K fue minimizar ambas clases de -- errores, falso-negativo y falso-positivo en las respuestas a la prueba, sin re -- ducir el número de cualquiera de las dos clases de respuestas en direcció -- n opuesta, los cierto-positivos y los cierto-negativos. La aproximación impli --

cada, define a los aciertos y los errores en términos del perfil total (más que en términos de un diagnóstico particular de subgrupos identificados por escalas separadas) buscando un factor o medida de corrección general. Esta clase de variable se ha denominado, desde entonces, factor supresor (Meehl y Hathaway,-- 1946).

Como el esquema sugiere, esta búsqueda avanza por diferentes caminos. Por ejemplo, grupos de perfiles significativamente elevados pueden ser obtenidos de pacientes y normales ya conocidos, y, por contraste, intentar aislar al algún indicador de cuándo tales elevaciones del perfil son sospechosas y cuándo son honestas. Esta aproximación fue usada por Meehl (1945) en la derivación de la escala N, que fue uno de los precursores de la escala K. El trabajo realizado con la escala N fue positivo, pero parece identificar solamente cierta clase de falso-positivos (por ejemplo, aquéllos con elevaciones engañosas en la triada neurótica, pero no a los que muestran elevaciones en la tétada psicótica). En contraste con la aproximación de la escala N, los trabajos con la escala K comenzaron con grupos de perfiles "normales" extraídos de anormales conocidos y presuntos normales (algunos de los casos normativos del Minnesota). Se usó el mismo número de casos para cada sexo y se hizo un esfuerzo por excluir las diferencias atribuibles al sexo en la escala resultante; aunque durante el desarrollo de la escala supresora K, se encontró que envuelve dos fases en su derivación. (11)

Veinticinco perfiles de sujetos masculinos y otros tantos de personas del sexo femenino en un grupo de internos psiquiátricos sirvieron de base, encontrándose tales perfiles dentro de los límites normales (ningún puntaje por encima de T70), y habiendo evidencia de que la prueba se respondió defensivamente (puntajes en L de cuando menos T60). La evaluación psiquiátrica de estos 50 sujetos indican que éstos presentaban una gran variedad de diferentes desórdenes, fundamentalmente problemas conductuales y caractereológicos (como alcoholismo) más que desórdenes de tipo neurótico o psicótico.

A juzgar por su estado clínico, estos sujetos debieran arrojar perfiles anormales al tomar la prueba, pero por la naturaleza de su problemática, podrían no ser sorprendidos, si realmente asumieran una postura altamente defensiva acerca de sí mismos, por la situación de prueba. La elevación de la escala L incrementa la credibilidad y la convicción de que estos pacientes han fracasado en conseguir perfiles del MMPI lo suficientemente elevados a causa de una postura defensiva general en la descripción de sí mismos.

No se hicieron esfuerzos adicionales por aparear cada patrón o elevación de los perfiles de estos casos psiquiátricos falso-negativos con perfiles de sujetos normales. Obviamente que en los perfiles que solo muestran variaciones dentro de los límites normales para cada uno de los supradichos grupos, la fluctuación y las variaciones configuracionales quedan considerablemente restringidas. Adicionalmente, como un comentario subsecuente que permita aclarar algunas configuraciones de las escalas clínicas, debe señalarse que las configuraciones del MMPI que ocurren con mayor frecuencia en muestras de sujetos normales, son aquéllas que involucran aspectos caractereológicos más que neuróticos -

o psicóticos. Por lo tanto, excepto por la estipulación hecha para la muestra falso-negativa en la que la escala L aparece apreciablemente elevada, las elevaciones y configuraciones de los casos verdadero-negativos y falsos-negativos utilizados para derivar la escala K son considerablemente semejantes.

El análisis sistemático de los ítems dentro de cada grupo sexual (hombr**es** y mujeres), condujo a la identificación de veintidos ítems comunes a ambos sexos los cuales dividen los perfiles negativos, ciertos y falsos, por su contestación a los ítems para cuando menos, el 30% de la población. Cualquiera ítem que ya aparezca en la escala L fue excluido de este grupo, tal conjunto de ítems fue designado como L6 y aparece (sin asteriscos) en la próxima tabla. (10, 11)

Los ítems fueron calificados de manera que puntajes elevados señalaran gran defensividad y alta probabilidad de que un puntaje normal en la prueba fuera producto del punto de vista hacia el test, más que de una auto-evaluación y auto-descripción escrupulosa o concienzuda.

ESCALA K: Composición, Dirección Crítica y Superposición con otras escalas básicas.

FALSA: (K)

374.- Durante ciertos períodos mi mente parece trabajar más despacio que de costumbre.

397.- Algunas veces he sentido que las dificultades se acumulaban de tal modo que no podía vencerlas.

FALSA: K

406.- Frecuentemente he conocido personas a quienes se suponía expertos y que no eran mejores que yo.

*461.- Me es difícil el dejar de lado la tarea que he emprendido aun cuando sea por poco tiempo.

*502.- Me gusta hacerle saber a la gente lo que pienso acerca de las cosas.

FALSA: LK23

* 30.- A veces siento deseos de maldecir.

FALSA: KF2

*272.- A veces estoy lleno de energía.

FALSA: (K2)

39.- A veces siento deseos de destruir cosas.

FALSA: (K23)

*160.- Nunca me he sentido mejor que ahora.

FALSA: (K235)

89.- Se necesita discutir mucho para convencer a la mayor parte de la gente de la verdad.

FALSA: K240

296.- Tengo épocas en las que me siento muy alegre sin que exista una razón especial.

CIERTO: 27

FALSO: (K)

142.- Decididamente a veces siento que no sirvo para nada.

CIERTO: 20

FALSO: (K)

138.- La crítica o el regaño me hiere profundamente.

FALSO: K3

* 71.- Creo que mucha gente exagera sus desdichas para que se conduelan de ellos y les ayuden.

FALSO: (K3)

129.- A menudo no puedo comprender porqué he estado tan irritable y malhumorado.

234.- Me molesto con facilidad, pero se me pasa pronto.

FALSO: K34

*170.- No me preocupa lo que otros piensen de mí.

FALSO: 4 CIERTO: (K)

96.- Tengo muy pocos disgustos con miembros de mi familia.

FALSO: (K4)

*183.- Estoy en contra de dar dinero a los mendigos.

FALSO: K4 CIERTO: (59)

134.- A veces los pensamientos pasan por mi mente con mayor rapidez que -
lo que puedo expresarlos en palabras.

CIERTO: 57 FALSO: (K)

217.- Frecuentemente me encuentro preocupado por algo.

FALSO: (K8)

322.- El dinero y los negocios me preocupan.

FALSO: (K9)

148.- Me impacienta que me pidan consejo o que me interrumpen cuando estoy
trabajando en algo importante.

CIERTO: 0 FALSA: (K)

383.- La gente me desilusiona con frecuencia.

398.- Frecuentemente pienso: "quisiera volver a ser niño".

CIERTO: 0 FALSO: (K349)

180.- Encuentro difícil entablar conversación con alguien que conozco por -
primera vez.

267.- Cuando estoy en un grupo de gente tengo dificultad pensando las cosas
apropiadas de qué hablar.

CIERTO: 0 FALSO: (K36)

124.- La mayor parte de la gente se vale de medios algo injustos para obte-
ner beneficios o ventajas antes que perderlos.

CIERTO: 0

FALSO: (K49)

171.- Me siento incómodo cuando tengo que hacer una payasada en una reunión, aun cuando otros estén haciendo lo mismo.

CIERTO: 0

FALSO: (K6)

316.- Creo que casi todo el mundo mentiría para evitarse problemas.

La inspección de los ítems que conforman la escala L6, muestran que el contenido es totalmente heterogéneo, comprendiendo descripciones de aspectos tales como la salud mental, estabilidad y control, sentimientos y expectativas acerca de los demás y varios aspectos de las relaciones familiares. Resulta interesante el hecho señalado por Comrey (1958) en el sentido de no haber encontrado ningún índice relevante de coherencia en las intercorrelaciones de los ítems de la escala K, en la misma forma, el subgrupo de ítems de corrección no emerge del análisis factorial como un factor único. Esta carencia de unidad condujo a Comrey a desacreditar el uso de esta escala como una medida de la forma de tomar la prueba. Una interpretación alterna es que sólo una postura de defensividad total trae como consecuencia que estos ítems, de tan diverso contenido, sean contestados consistentemente en dirección defensiva (o exhibicionista). Entre una muestra de sujetos que se comportan de manera abierta y franca ante la prueba, habrá pocas razones para esperar que estos reactivos sean contestados con alguna uniformidad. De ahí que las diversas cargas factoriales encontradas por Comrey resulten ciertas. (5)

Partiendo de la forma en que se derivó la escala L6, la expectativa habría sido que los sujetos que adoptan posturas defensivas, deberían dar perfiles normales en el MMPI y altos puntajes en la escala L6, mientras que los sujetos normales con una actitud apropiada para tomar la prueba obtendrían perfiles con bajos puntajes en esta escala especial de defensividad. En sus estudios preliminares de las características de la escala L6, Meehl y Hathaway (1946) encontraron que la mayor parte de los sujetos normales que se emplearon en la muestra para derivar el Minnesota no cayeron en la parte inferior de la distribución de esta escala sino más bien al centro de la distribución de frecuencia, ninguno muy alto o muy bajo. La exploración de las características psicológicas de los bajos puntajes en L6, condujo al descubrimiento de que la escala es bipolar en sus implicaciones psicológicas. Mientras los altos puntajes identifican algunas clases de posturas defensivas al responder el MMPI, los puntajes del extremo opuesto parecen revelar lo que Florencia Goodenough describiría más tarde como "una clase peculiar de exhibicionismo, el que toma la forma de un impulso que expone algunas perturbaciones y confiesa algunas debilidades" (1949). Los sujetos que respondieron la prueba con sinceridad y con un grado apropiado de apertura evitaron caer en cualquiera de los extremos de la escala y obtuvieron puntajes al centro de la distribución de L6. Estudios posteriores sobre la aquiescencia al responder, han ampliado nuestra comprensión de este proceso. (10)

Estas mismas exploraciones revelaron una característica adicional de la escala L6 que los autores percibieron como una seria deficiencia: los bajos -

puntajes no solamente se encuentran entre sujetos normales que parecen estar exagerando su psicopatología (falso-positivos), sino que también aparecen entre algunos pacientes psicóticos, fundamentalmente en aquéllos que muestran marcada depresión o reacciones esquizofrénicas. Puesto que una de las clases de puntajes bajos en la escala L6 parece ser una tendencia a falsear la anomalía en la prueba, los autores pensaron sobre la necesidad de llevar a efecto una adecuada corrección que controlara los aspectos psicóticos, por medio de ítems que no fueran afectados por los esfuerzos deliberados de ocultar la anomalía en el MMPI. Los trabajos previos sobre el falseamiento en los patrones anormales del MMPI por sujetos ostensiblemente normales realizados por investigadores como Gough (1947), Hunt (1948) y Bird (1948), proporcionaron datos acerca de qué ítems resultaban apropiados para ser contestados en sentido diferente cuando el sujeto estaba tratando de mostrar un cuadro de psicopatología y los cuales resultaron relativamente inafectados por esta forma peculiar de tomar la prueba, pero que fueron también contestados de la misma manera en las administraciones regulares del MMPI, por estos mismos sujetos cuando tomaban la prueba de manera normal. Por ejemplo, dentro del sub-grupo de los ítems que habían sido cambiados, Gough (1947) identificó aquéllos ítems los cuales en realidad no superponen el sentido real en que los pacientes se describen a sí mismos. A partir de estos datos, Gough derivó su escala de disimulación, Ds. Dentro de los ítems estables de estos estudios sobre el falseamiento, Meehl y Hathaway se dieron a la búsqueda de ítems que permitieran diferenciar entre pacientes normales y psicóticos. Ocho de tales ítems fueron identificados y agregados a la escala L6, el conjunto total de treinta ítems fue nominado como escala K (corrección). Estos ítems han sido señalados con un asterisco en la tabla precedente. El subgrupo de corrección para psicoticismo es calificado de manera que los sujetos psicóticos obtengan altos puntajes en esta escala, teniendo este puntaje complementario el efecto de mostrar una elevación de la calificación por encima del promedio en la escala K, la cual no resulta exagerada por sujetos normales y separa los puntajes de los sujetos falso-positivos. Esta última clase de sujetos presuntamente obtienen bajos puntajes en la escala K porque se encuentran "parados" sobre el extremo inferior de L6 y en realidad no expresan pensamientos o necesidades patológicos, como para obtener un puntaje elevado en el subconjunto de psicoticismo. La naturaleza resultante de la escala K es compleja (resulta obvio que algunas investigaciones posteriores con el MMPI, que han realizado un análisis de los ítems de la escala K han malentendido esta "corrección" especial de la escala de corrección como es el caso de Hanley, (1956)).

En la derivación de L6, los autores deliberadamente excluyeron los ítems ya identificados en la escala L estándar, en atención a que uno de los criterios para la selección de casos para L6, fue un alto puntaje en L.

Uno de los ítems de L fue incluido en la escala K en la segunda etapa; sin embargo, sólo como parte de la corrección por psicoticismo. Alrededor de la tercera parte de los ítems de la escala K son calificados en dirección opuesta a la de su dirección clínica, o bien sin relación con la dirección crítica de cualquiera de las otras escalas que componen el perfil básico. La mayoría de los ítems de la escala K también aparecen, sin embargo, en algunas escalas clínicas calificados en el mismo sentido que la escala o escalas en cuestión.

Puede señalarse aquí que los esfuerzos hechos para corregir las escalas - individuales por medio de subconjuntos de ítems de corrección, generalmente - han tenido como consecuencia la identificación de algunas clases de ítems que - son muy similares a aquéllos que componen la escala K. (10, 11).

LA CORRECCION DE LAS ESCALAS CLINICAS POR MEDIO
DE LOS PESOS DE K.

Si bien los autores de la escala K sugieren que los puntajes de este re-- ciente indicador de validez sean utilizados como las preexistentes medidas de - validez que incluye la prueba; esto es, su aplicación juiciosa altera las in-- ferencias que pueden obtenerse del perfil básico; la inclusión de la escala K - en tales inferencias avanzó rápidamente hasta convertirse en un grupo de pro-- cedimientos psicométricos formales que son aplicados rutinariamente en todos - los perfiles del MMPI. Hathaway y Meehl (1948) dieron por sentado el proce-- dimiento por medio del cual los pesos de corrección por medio de puntajes K - son agregados a los puntajes crudos de determinadas escalas del perfil antes - de determinar sus puntajes T. En 1947, Hathaway y Meehl prepararon un suple-- miento para el Manual de la prueba, en el que incorporaban los nuevos procedi-- mientos. The Psychological Corporation modificó posteriormente su impresión de las formas de perfil, incluyendo la escala de corrección K dentro de la ru-- tina de calificación, e integrando el material de los procedimientos y pesos - de la escala en las revisiones posteriores del Manual. Como se indica en la primera parte de dicho Manual, para determinar el valor del puntaje T cuando - la escala K no se califica, por lo que se refiere a las escalas que se ven -- afectadas por esta corrección, deben consultarse tablas especiales. Debe sub-- rayarse que no es posible obtener puntajes T no corregidos a partir de los pun-- tajos crudos simples (sin su respectiva corrección K) en las columnas de cali-- ficación del perfil regular. (19)

Puesto que la escala K intenta mejorar la diferenciación entre perfiles - normales y anormales, la dispersión crucial de las variaciones de puntaje en - cualquiera de las escalas se considera sobre una sigma arriba y abajo del pun-- taje de T70 y, cuando mucho, dos sigmas alrededor de este valor (ver la discu-- sión sobre elevaciones fronterizas en la sección de Interpretación del MMPI, - en el Manual). (11)

En consecuencia, la región en la que los pesos específicos de la escala - K deben efectuar la discriminación de los perfiles fue definida por Hathaway, - McKinley y Meehl (1948) como la que se encuentra comprendida entre los punta-- jes T50 y T80. Los puntajes abajo de este rango, son agrupados arbitraria-- mente como "50", mientras que los puntajes que van más allá del supradicho ran-- go, se agrupan como "80".

Esta decisión es completamente consistente con el propósito inicial de -- la escala K, pero los autores impulsaron la depuración de rangos, métodos y -

pesos alternos de discriminación con sus propias áreas de aplicación.

Hasta la fecha, estos intentos no han recibido el apoyo teórico o de investigación empírica que merecen y necesitan. La razón de pensar que tales límites son arbitrarios es simplemente que; los puntajes inferiores a este rango son tan bajos que una corrección estadística simple puede no producir la elevación de la escala al rango de interpretación clínica, mientras que los puntajes superiores a este rango, son tan altos que una reducción no estadística puede originar que caigan dentro de los límites normales. Los pesos K son aplicados por lo general a todos los puntajes crudos posibles, no obstante, dado que los puntajes T son interpretados de manera muy diferente dentro de regiones diferentes y a lo largo de toda la escala, y no solamente dentro de esta arbitraria región "fronteriza", persiste la no resuelta cuestión de si los pesos sugeridos por los primeros estudios de Minnesota, son en la actualidad óptimos para todas las rutinas de aplicación. Un segundo e importante punto, que también permanece sin solución es que los cambios en la configuración del perfil introducidos por correcciones "escala por escala", en las cinco escalas afectadas, ¿aumentan o disminuyen la precisión de las inferencias y juicios clínicos que se apoyan en estas configuraciones?. En vista de esto, debería esperarse que los ajustes a las escalas que integran la prueba, además de incrementar el poder discriminativo de una escala particular, mejoraran también el ajuste de la configuración total. Por otra parte, la variación del puntaje está siendo agregada a varias escalas, lo que, tal vez, reduce la información independiente proporcionada por las escalas que integran el perfil y posiblemente vuelve confusas algunas discriminaciones decisivas. Comparaciones adicionales de las configuraciones de los perfiles del MMPI, con y sin la corrección K, junto con información directa de las limitaciones empíricas que cada grupo puede proporcionar sobre las características relevantes de los sujetos que se han sometido a la prueba, son absolutamente necesarios. (11)

La determinación de los valores óptimos de K para cada escala, con el fin de obtener la separación de los casos con reacciones neuróticas o psicóticas fue hecha por el simple método de ensayo y error.

No se establecieron diferencias sexuales en la aplicación de la corrección K; los pesos resultantes son los mismos para ambos sexos.

Los incrementos decimales de los valores crudos de K fueron agregados sistemáticamente a los puntajes crudos originales de cada escala obtenidos por los pacientes-criterio y a los de los adultos normales de Minnesota. Se derivaron nuevos puntajes T para cada escala y los puntajes resultantes de cada grupo fueron comparados en términos de un criterio diferencial que los autores llamaron la mitad de la suma de cuadrados del remanente que aparecería después de ajustar la distribución a puntajes T entre 50 y 80. El peso que generó el valor máximo de este criterio diferencial en varias muestras experimentales, fue elegido como peso estándar para una escala. En el caso de las escalas 7 y 8, el puntaje crudo total de K demostró ser el mejor en la discriminación entre sujetos normales y neuróticos obsesivo-compulsivos y casos esquizofrénicos respectivamente. Para las escalas 1, 4 y 9, se encontró, sin embargo, que una fracción

del puntaje crudo (0.5, 0.4 y 0.2 respectivamente) discriminaba mejor, mientras que fracciones mayores o menores del puntaje crudo de K sólo proporcionaban discriminaciones pobres entre los grupos de referencia y los sujetos--criterio en los casos de hipocondriasis, desviación psicopática y reacciones depresivas. McKinley, Hathaway y Meehl (1948) estudiaron los efectos de la corrección K sobre las escalas 2, 3, 5 y 6 encontrando que cualquier corrección de las mismas tiende a reducir la discriminación proporcionada por los valores originales de la escala. Estas escalas no son corregidas en la rutina de aplicación del MMPI, aunque Heilburn (1963) ha recomendado correcciones para algunas de ellas en el contexto de la consulta colegial.

McKinley, Hathaway y Meehl (1948) también proporcionan evidencia basada en validaciones cruzadas de casos cuando la corrección K se efectúa para las escalas 1, 4, 7, 8 y 9, las que fueron mejoradas estadística y pragmáticamente en su posibilidad de diferenciar entre pacientes y normales.

Se computaron nuevas medias aritméticas y desviaciones estándar a fin de obtener los puntajes T de las escalas en las que se efectúa la corrección K, y se encontró que los grupos clínicos y normales mostraban una superposición significativamente menor en su distribución de puntajes. Ya sea que tal superposición haya sido evaluada por medio de un puntaje límite sobre la escala T (por ejemplo, 70), o bien por medio de un porcentaje fijo de sujetos normales, (por ejemplo el percentil 95), de todas maneras resulta palpable que cada una de las escalas clínicas que incluye la corrección K, proporciona una mejor separación entre sujetos normales y un criterio específico de anormalidad, que la versión que no incluye la corrección.

Sus tabulaciones muestran que este decremento de la superposición de las distribuciones es obtenido algunas veces por reducción del número de puntajes bajos en las nuevas muestras de casos clínicos, algunas otras reduciendo en el grupo normal de referencia el número de sujetos con puntajes elevados en estas escalas y otras veces operando ambas situaciones de manera simultánea. Estos hallazgos y la experiencia subsecuente con las correcciones K en los perfiles de varios servicios clínicos de los Hospitales de la Universidad de Minnesota han llevado a los autores a recomendar que las escalas clínicas con corrección K sean consideradas como las versiones básicas de las escalas que componen el inventario. Al mismo tiempo, la fuerte presión de los usuarios del MMPI, han hecho posible realizar comparaciones de puntajes corregidos y no corregidos con K en la identificación de pacientes y clientes en sus propios grupos clínicos y también manejar otros pesos de los valores de los puntajes K. (11)

Los datos publicados al respecto de los esfuerzos de la investigación, por medio de validaciones cruzadas de K y variaciones en el pesaje de la misma, son limitados y no concluyentes. Obviamente, muchas de las situaciones en que se usa la prueba no pueden proporcionar un criterio ideal que sea completamente independiente de la información o decisiones apoyadas en la prueba. A pesar de haber transcurrido hasta hoy ya más de 20 años desde la inclusión de la escala K en el MMPI, las investigaciones realizadas con la prueba, no han dado mayor atención a este problema.

Los esfuerzos de investigaciones que se encuentran disponibles han sido dedicados a varias cuestiones afines concernientes al uso de la corrección K: ¿resultaría más apropiado sustraer fracciones de K más que adicionarlas a los puntajes crudos de las escalas clínicas?, ¿puede la corrección K compensar deliberadamente los puntajes clínicos falseados? y ¿puede usarse la corrección K para mejorar las discriminaciones de algunas de las otras escalas componentes del perfil que ahora no son objeto de tal corrección?. Mucho menos atención se ha puesto a la cuestión medular: ¿es posible que la introducción de la corrección K permita un incremento en la capacidad discriminativa de las escalas, de forma tal que redunde en una mejor aproximación a la interpretación del perfil total?. (11)

Hasta la fecha, parece ser que el efecto supresor de una escala de corrección como K, no se ve incrementado por el hecho de cambiar el signo algebraico de los pesos que son adicionados a ciertas escalas.

La consideración primordial es que la escala K refleja una gran proporción de su varianza sobre la escala clínica que recibe la corrección, y que tal incremento no se refiere al atributo clínico que la escala en cuestión pretende medir. Si esta relación se mantiene, entonces, la parte del puntaje que corresponde a la contribución de la variable supresora, puede ser pesado y combinarse con la escala clínica, de manera tal que la distribución de puntajes resultante reflejara más claramente la varianza del atributo que sirve como criterio para tal escala. Ni los esfuerzos de Fricke (1956) ni los de Heilbrun (1963) para mejorar la capacidad diagnóstica de la escala Hi, por ejemplo, por medio de la substracción de los pesos de K han servido para sensibilizar la escala de manera apreciable.

De manera similar, los pocos estudios que se han realizado para tratar de compensar los efectos de la distorsión deliberada, tanto en sentido positivo como negativo, no indican que los pesos de K empleados normalmente en la calificación del MMPI puedan servir para esta función. En un estudio anterior (1948) Schmidt, pidió a 11 sujetos, con severas perturbaciones psiconeuróticas, y de quienes conocía su historial clínico, que resolvieran nuevamente el MMPI, asumiendo una postura que ellos consideraran correspondería a los sujetos normales. Bajo estas condiciones, los perfiles de tales sujetos fueron bastante similares a los del patrón que corresponde a sujetos normales, especialmente en las escalas Hs, D, Hi, Dp, Pt y Es. Al introducir la corrección K en tales perfiles, se observó que el efecto de la postura compensatoria quedó eliminado para las escalas Dp y Es, en tanto que la escala Pt sólo se afectaba parcialmente. Para la escala Hi el efecto de la corrección K, fue mínima, en tanto que la escala Ma no fue afectada, ni por los intentos de distorsión ni por la corrección K.

Por cuanto toca al problema de separar los perfiles distorsionados de los no distorsionados, Schmidt notó que la escala K se vió afectada por las instrucciones de distorsionar la prueba, prácticamente, al mismo nivel de la escala L. La forma en que estos sujetos respondieron el MMPI bajo instrucciones-

de mostrarse psicológicamente, lo mejor posible, se pusieron de manifiesto por la elevación de las escalas L y K, aunque esta postura no contribuyó a decrementar sus puntajes en la escala F. (11)

Puesto que la derivación de L y la integración de la escala K estuvieron basadas en comparaciones generales de grupos con perfiles normales y perfiles anormales, resulta un tanto sorprendente encontrar que el efecto supresor de K pudiera ser demostrado pragmáticamente sólo en algunas de las escalas del perfil; Hathaway y Meehl (1946) atribuyeron esta configuración del efecto de K a las características particulares de las escalas, ya fuera por la existencia de componentes correctivos previos en la escala (p. ej. la escala D), o bien por la inclusión de ítems sutiles en las escalas, los que provienen directamente del criterio-base de la escala (p. ej.: las escalas Hi y Pa). El trabajo de Heilburn (1963), mencionado arriba, fue dedicado a la búsqueda, escala por escala, de nuevos pesos para la corrección K, en ambos sentidos: por adición o por sustracción, los que pudieran mejorar el poder discriminativo de la prueba, entre estudiantes universitarios normales y anormales conocidos. Heilburn esperaba, también, que algunas de las escalas, que normalmente no incluyen la corrección K, podrían recibirla, en tanto que algunos de los pesajes actuales podrían ser disminuídos o eliminados para mejorar tal capacidad diferencial. Los casos que fundamentaron su criterio, fueron definidos en términos globales de normalidad o anormalidad, más que en términos de subgrupos específicos por diagnóstico clínico. En este sentido, el criterio fue similar al contraste de grupos que utilizaron Meehl y Hathaway en 1946 en la derivación de las escalas originales, más que la derivación de pesos específicos para cada subgrupo diagnóstico. Heilburn recomienda el uso de diferentes pesos para hombres y para mujeres, sugiriendo, además, que tales pesos deben usarse en sentido negativo en la escala Hi (en ambos sexos), que los pesos para las escalas 1, 4 y 9 sean reducidos a cero, y que los pesos para las escalas 7 y 8 sufran una reducción diferente para cada sexo. Asimismo, elaboró nuevas tablas de puntajes T para ser usadas en la población de estudiantes universitarios. Hace también una advertencia al respecto del uso de estos datos, en tanto no se lleven a cabo estudios a fondo, por medio de validaciones cruzadas, de las modificaciones que propone; aunque hasta la fecha no ha publicado ningún trabajo al respecto, pese a que los estudios originales a que se ha referido este comentario fueron iniciados en 1963.

La investigación del impacto total de la corrección K sobre la configuración del MMPI, ha sido también desilusionante en cuanto a calidad y precisión. Los dos estudios principales que se han realizado al respecto se basaron en clasificaciones "a ciegas" realizadas por clínicos que estudiaron la configuración de cierto número de perfiles; en tanto que un tercer estudio (Wiener, 1948) utilizó un índice psicométrico crudo (cualquier escala del perfil superior a un puntaje T de 70) en muestras de veteranos de la Administración (en los Estados Unidos), pacientes hospitalizados que no tenían ningún empleo que correspondiera con los del grupo normativo del MMPI. Wiener encontró que había más perfiles elevados significativamente en sus dos grupos clínicos, cuando los perfiles fueron calificados con la corrección K que cuando se empleaban puntajes T sin la corrección K, reduciendo los perfiles falsos-negativos en un 80%. No se incluyó en este estudio la información referente a las bases empleadas para estimar cuáles perfiles se consideraron como falsos-negativos o como ciertos-negativos. (11)

Hunt y colaboradores (1948) utilizaron los perfiles de 114 Veteranos de la Administración, todos casos hospitalizados y con problemas neuro-psiquiátricos, los que fueron evaluados con y sin la corrección K, para las escalas que normalmente la incluyen, y estudiaron la precisión con la que cada perfil quedó clasificado dentro de una de tres categorías: psicótico-neurótico o desórdenes de la conducta. La clasificación fue obtenida por un grupo de especialistas en los tres grupos clínicos y todos los casos sobre los que el grupo no pudo ponerse de acuerdo fueron eliminados de la tabulación. Se excluyeron, así mismo, algunos casos de la categoría desórdenes de la conducta para los que la decisión final para clasificarlos se encontraba basada en la dicotomía psicótico versus no psicótico. Solamente 89 de los 114 casos iniciales, alcanzaron el requisito de juicio unánime por parte de los expertos y de éstos, sólo 53 (58%) estuvieron totalmente de acuerdo con el diagnóstico psiquiátrico independiente que se hizo de estos casos. Cuando los 89 casos fueron calificados con la corrección K, la precisión del diagnóstico se elevó de manera insignificante (al 61%), aunque no en un grado estadísticamente significativo. Estos investigadores analizaron también el efecto de la corrección K en muestras que incluían un menor número de casos de la categoría desórdenes de la conducta, y encontraron que la precisión diagnóstica en la dicotomía psicótico versus neurótico, era esencialmente la misma lograda con la muestra total. Se estudiaron también 40 casos de elevación fronteriza (cuando menos una de las escalas por encima de 65, aunque ninguna de ellas por encima de 80), perfiles sobre los que la corrección K podría resultar más adecuada, con resultados similares a los anteriores. Hunt y sus colaboradores, llegaron a la conclusión de que la corrección K no mejoraba la capacidad del MMPI para establecer un diagnóstico diferencial. Si existía más información útil que pudiera ser derivada de esta investigación, este grupo de especialistas, cuando menos, fue incapaz de capitalizarla adecuadamente.

Silver y Sines (1962) realizaron un estudio similar con una muestra de casos psiquiátricos de un hospital estatal, aunque su clasificación se hizo solamente sobre dos aspectos, estos investigadores estudiaron un amplio rango de categorías diagnósticas (psicóticos, esquizofrénicos, neuróticos y desórdenes de la personalidad) con mayores variaciones que las evaluadas por las escalas clínicas del perfil. Mientras los especialistas que fungieron como jueces en el estudio de Hunt y colaboradores no informan acerca de la presencia o ausencia de las correcciones K, Silver y Sines reportan sistemáticamente sobre esta clase de información. Los perfiles del MMPI son presentados con y sin el puntaje T correspondiente a la escala K y con y sin la corrección K hecha sobre el perfil, señalándose además cuál perfil corresponde a cada caso. (11)

En consecuencia, si un juicio particular fue obtenido por una forma, diferente a la común de interpretar el MMPI, con o sin correcciones, pueden utilizarse tales "reglas" interpretativas para algún perfil en particular. Estos resultados indican, sin embargo, que existieron una serie de circunstancias ligeramente favorables para establecer la clasificación, a partir de la siguiente información: para los perfiles, con o sin corrección K, los juicios para establecer la clasificación señalada, funcionan igualmente bien. La probabilidad de acierto fue de aproximadamente 50% para clasificarlos adecuadamente en las cuatro categorías señaladas arriba, la cual se encuentra bastante por encima del nivel de azar para los 100 hombres y 110 mujeres de quienes se reportaron datos, aunque no deja de ser desilusionante el bajo nivel total de precisión que reportan estos autores. En consecuencia, los juicios emitidos al respecto de un aumento (o disminuación) en la utilidad del MMPI por la introducción-

de la corrección K en cinco de las escalas componentes, no puede ser demostrado por medio de este particular criterio externo.

Debe señalarse que la gran mayoría de correlaciones del perfil, así como de los datos de interpretación de que en la actualidad se dispone, se encuentran basados en puntajes que incluyen la corrección K. Con el objeto de poder capitalizar adecuadamente estos hallazgos y reglas de interpretación, los usuarios de la prueba debieran calificar los perfiles del MMPI incluyendo la corrección K, de acuerdo con el procedimiento estándar que especifica el Manual. En la medida de lo posible, cada usuario debiera, sin embargo, investigar al respecto de la utilidad de la corrección K en los perfiles obtenidos en la situación que se maneja, a fin de encontrar los pesos particulares de K que resulten más apropiados para corregir las escalas en situaciones especiales. (11)

IPP.- BOSQUEJO GENERAL DE LA PRUEBA. (5, 15)

I N T R O D U C C I O N.

El IPP fue creado con la esperanza de conseguir dos metas en la medida de la personalidad. La primera de ellas, de naturaleza más bien teórica, ha sido utilizada para desarrollar conceptos descriptivos que poseen gran importancia en aspectos sociales y de la personalidad. Muchos de los tests comunes de personalidad, así como otros instrumentos de medición disponibles anteriormente, habían sido diseñados para usarse en situaciones especiales, tales como la psiquiatría clínica, o contruídos para tratar de resolver algún problema particular, como es el caso de la orientación vocacional.

Este inventario ha sido contruído con el objeto de evaluar las características de la personalidad que tienen una amplia y directa aplicación en la conducta humana y las que adicionalmente están relacionadas con los aspectos favorables y positivos de la personalidad, más que a los aspectos mórbidos y patológicos.

La segunda de estas metas del IPP ha sido el desarrollar un instrumento breve, preciso y con subescalas confiables para la identificación y medida de las variables incluídas en el inventario. Un propósito ulterior ha sido que el instrumento sea de uso sencillo y apropiado para aplicaciones en gran escala.

La consecuencia de estos dos propósitos ha sido la creación del IPP en su forma actual. El cuadernillo de prueba contiene 480 ítems * (12 de los cuales están repetidos para facilitar la calificación a máquina), y producen 18 puntajes estándar. La prueba incluye una hoja de respuestas y otra especial para trazar el perfil.

El inventario fue creado, en principio, para ser usado con sujetos "normales" (sin disturbios psiquiátricos). Las escalas están dirigidas principalmente a las características de personalidad importantes para la interacción social y la vida. Así, aunque se ha encontrado que tiene especial utilidad en algunos grupos problema, como son delincuentes, individuos con tendencias antisociales, etc., su uso más general se encuentra en escuelas, negocios, industrias y en agencias clínicas y de consejo, sitios en donde la mayoría de los sujetos tienen un funcionamiento social adecuado.

* Aproximadamente 200 de estos ítems aparecen también en el MMPI. Estos ítems fueron incluídos en el IPP con autorización del Dr. Hathaway y bajo un convenio especial con University Of Minnesota Press y con The Psychological Corporation.

LAS ESCALAS DEL IPP. (15)

El IPP incluye las escalas estándar que aparecen adelante. Cada escala está destinada a cubrir una importante faceta de la psicología interpersonal y el conjunto total de las 18 escalas pretende proporcionar una apreciación de cada individuo, desde el punto de vista de su interacción social. Las escalas están agrupadas, por conveniencia, en cuatro amplias categorías, que tienen entre sí implicaciones similares. La lógica subyacente es aquí interpretativa, no estadística; no se pretende que estas cuatro categorías constituyan factores psicométricos.

CLASE I: Medidas de Equilibrio, Ascendencia, Seguridad de sí Mismo y Adecuación Interpersonal.

1.- Do Dominancia. Evalúa factores de habilidad para el liderazgo, dominancia, persistencia e iniciativa social.

2.- Cs Capacidad para adquirir posición ("status"). Índice de la capacidad individual para adquirir "status" (no su "status" actual o logrado). La escala intenta medir las cualidades y atributos personales subyacentes y conducentes a la adquisición de "status".

3.- Sd Sociabilidad. Identifica personas de temperamento abierto, sociable y participante.

4.- Ps Presencia Social. Evalúa factores tales como el equilibrio, espontaneidad y auto-confianza en las interacciones personales y sociales.

5.- Aa Auto-aceptación. Evalúa factores tales como el sentido de estima personal, auto-aceptación y la capacidad para el pensamiento y la acción independientes.

6.- Sb Sentido de bienestar. Identifica personas que minimizan sus preocupaciones y quejas, y que están relativamente libres de dudas y desilusiones.

CLASE II: Medidas de Socialización, Madurez, Responsabilidad y Estructuración Interna de Valores.

7.- Re Responsabilidad. Identifica personas de temperamento y disposición íntegra, responsable y confiable.

8.- So Socialización. Identifica el grado de madurez social, integridad y rectitud que el individuo ha alcanzado.

9.- Ac Auto-control. Evalúa el grado y adecuación de la auto-regulación y del auto-control, el estar libre de la impulsividad y el egocentrismo.

10.- To Tolerancia. Identifica personas con actitudes y convicciones - sociales permisivas, aceptantes y no críticas.

11.- Bi Buena impresión. Identifica personas capaces de crear una impresión favorable y que están interesadas acerca de cómo otros reaccionan hacia - ellos.

12.- Cm Sentido común. Identifica el grado en el que las reacciones y - respuestas de un individuo corresponden al patrón modal ("común") establecido - por el inventario.

CLASE III: Medidas de Potencial de Logro y Eficiencia Intelectual.

13.- La Logro vía adaptación. Identifica aquellos factores de interés - y de motivación que facilitan el logro en cualquier situación donde la adapta - ción es una conducta positiva.

14.- Li Logro vía independencia. Identifica aquellos factores de inte - rés y de motivación que facilitan el logro en cualquier situación donde la au - tonomía y la independencia son conductas positivas.

15.- Ei Eficiencia intelectual. Identifica el grado de eficiencia inte - lectual y personal que el individuo ha logrado.

CLASE IV: Medidas de las Formas Intelectuales y de los Intereses.

16.- Pi Disposición psicológica. Mide el grado en el que el individuo - está interesado en las necesidades internas, motivos y experiencias de los de - más, respondiendo a ellas.

17.- Fx Flexibilidad. Indica el grado de flexibilidad y de adaptabili - dad del pensamiento de una persona y de su conducta social.

18.- Fe Feminidad. Evalúa la masculinidad o feminidad de los intereses. (Altos puntajes indican intereses femeninos, bajos puntajes, intereses masculi - nos).

A D M I N I S T R A C I O N .

El IPP es auto-administrable. Las preguntas están impresas en un cuader - nillo que puede ser empleado varias veces. Las respuestas se marcan en una - hoja especialmente diseñada para las calificaciones que han de obtenerse manual - mente.

El inventario puede ser utilizado con personas de edades que van desde - los 12 ó 13 años, hasta los 65 ó 70 años.

Con sujetos de limitada escolaridad, algunos de los ítems aparecerán muy difíciles o irrelevantes. A pesar de esos problemas, los resultados del test en la mayoría de los casos son significativos y fácilmente interpretables. Con sujetos de educación media o superior este tipo de problemas disminuye considerablemente.

Hasta ahora, tal como ha podido determinarse por la precisión de los perfiles obtenidos y por los indicadores de confiabilidad y exactitud del test; los resultados son satisfactorios bajo cualquier condición de aplicación.

EL PERFIL Y LAS NORMAS.

Las normas para el IPP fueron desarrolladas a partir de una muestra en la que ambos sexos fueron tabulados por separado. Las estandarizaciones de puntajes para sujetos del sexo masculino están basadas en más de 6 000 casos, mientras las normas femeninas incluyen más de 7 000. Estos totales son suficientemente grandes, e incluyen un amplio rango de edades, grupos socioeconómicos y áreas geográficas, pero la muestra no fue obtenida como una verdadera muestra al azar de la población general.

El perfil, con normas para hombres y mujeres, se construyó en forma tal que proporcionara una conversión automática de los puntajes crudos a puntajes estandarizados, en donde la media aritmética es de 50, con una desviación estándar de 10. Con estas unidades de medidas, alrededor del 95% de la población puede ser localizada, al menos teóricamente, entre los puntajes estándar de 30 y 70 en cada escala.

La hoja del perfil incorpora los cuatro grupos de escalas previamente descritos. El propósito de estos grupos es facilitar la interpretación clínica del perfil y no definir categorías psicométricas o factoriales. Las escalas de la Clase I comparten un énfasis común sobre sentimientos de adecuación interpersonal e intrapersonal. Las de la Clase II, se refieren principalmente a las normas y valores sociales y a la disposición a observar o rechazar tales valores. Las escalas de la Clase III no constituyen una dimensión psicométrica, pero son, sin embargo, presentadas juntas por sus implicaciones sobre aspectos generales de manejo del medio ambiente. La Clase IV incluye escalas que funcionan más bien independientemente una con respecto a otra y a las otras 15 escalas, y las cuales parecen reflejar actitudes con respecto a la vida en general y a su significado a largo plazo.

Desde que el IPP fue usado por primera vez en una investigación de gran envergadura, en 1951, ha sido administrado a más de un millón de sujetos. No todos estos casos estuvieron disponibles para la estandarización de la prueba, pero los datos que han sido publicados, son suficientes para dar una idea de la tendencia y variaciones que deberán esperarse en diferentes grupos.

INTERPRETACION DEL IPP.

El monto de información proporcionado por la interpretación de una prueba psicológica es una función del alcance y adecuación del instrumento; de las habilidades y bagaje general del interpretador, y del conocimiento y experiencia del usuario con el test mismo. Algunos psicólogos están predispuestos a continuar usando instrumentos obsoletos simplemente por su experiencia anterior, cerrando la puerta al uso de instrumentos técnicamente más avanzados y fundamentados en hallazgos recientes de la psicología. Esta postura debiera ser reconsiderada, toda vez que en la actualidad, muchos de los inventarios de personalidad de reciente edición proporcionan evidencia y guía para su interpretación, de forma que el examinador puede empezar, de una vez, a desarrollar su conocimiento interpretativo y habilidades con un nuevo instrumento.

LAS ESCALAS INDIVIDUALES.

Los nombres de las escalas fueron cuidadosamente elegidos para describir tan aproximadamente como fuera posible, el tipo de conducta que están destinadas a reflejar. Por lo tanto el nombre proporciona la primera guía para el significado de una escala. Por ejemplo, de una persona que califique alto en Do (Dominancia), podría esperarse que impresione a otros como enérgica, persistente, segura de sí: una persona dominante; de un individuo que califique bajo, podría esperarse que fuera apartado, modesto, quizá inhibido y carente de autoconfianza.

El significado de una escala es ahondado mediante un conocimiento de la definición o propósito de la escala, el que, en la mayoría de los casos, sugiere en cierto grado el tipo de grupos de criterio usados para desarrollarla.

Según los autores de la prueba, las escalas no están destinadas para definir rasgos o especificar factores psicométricos, por lo que la validación deberá dirigirse al grado en el que el inventario puede predecir la conducta e identificar individuos que son percibidos en formas características. Según los autores, los análisis factoriales, aunque de interés para delinear la estructura interna del inventario, no son de relevancia muy particular para las demandas de validez requeridas por el test.

INTERACCIONES ENTRE LAS ESCALAS.

La mayoría de los sujetos, por supuesto, tendrán más de un solo puntaje alto y más de un puntaje bajo. La interpretación por lo tanto, debe tratar con patrones y combinaciones de puntajes altos y bajos. Esos patrones oscilan desde las combinaciones más simples hasta los perfiles complejos que involucran interacciones inusuales de altos y bajos puntajes; la efectividad de la prueba radica en gran parte en la habilidad del usuario, para comprender la interacción de las escalas individuales.

A la fecha, la discusión sobre tales patrones e interacciones en el IPP - debe ser tentativa. Reconociendo esta necesidad, he querido, empero, ofrecer algunos ejemplos del tipo de análisis de dos escalas que pueden y que deberían hacerse. La investigación adicional podría muy bien alterar o mejorar esos - diagramas, pero indudablemente son útiles en la consideración de este problema.

Se considera que las escalas Re (Responsabilidad) y Cs (Capacidad de status), generalmente intercorrelacionan casi $+0.35$. Este tipo de correlación - significa que muchas personas calificarán arriba del promedio en una y abajo - en la otra, aunque se piense que la tendencia predominante es calificar alto - o bajo en ambas. Desde nuestro punto de vista de "interacción de las escalas", cuatro posibilidades son de gran importancia: (a) calificación alta en ambas - escalas; (b) calificación alta en Re, pero baja en Cs; (c) calificación baja - en Re, pero alta en Cs; y (d) calificación baja en ambas escalas. ¿Qué carac - terísticas psicológicas debería uno esperar encontrar entre personas de cada - una de las cuatro categorías?. El listado de abajo proporciona la respuesta - que la teoría de las escalas y la evidencia de investigación pertinente a ellas, como dimensiones simples, podría sugerir:

CARACTERISTICAS.

Re alta			
	justo	eficiente	
	apacible	enérgico	
	paciente	perseverante	
	sincero	equilibrado	
	modesto	hábil	
Cs baja			Cs alta
	apático	agresivo	
	torpe	manipulador	
	desconfiado	oportunista	
	indiferente	egoísta	
	desconfiable	escéptico	
Re baja			

Desde el punto de vista de la técnica social y de la teoría de los procesos del pequeño grupo, la interacción de las escalas Do (Dominancia) y Sd (Sociabilidad) es de cierto interés; la correlación típica entre esas dos escalas es de $+0.65$. Las predicciones para sus interacciones son éstas:

CARACTERISTICAS.

Do alta	
analítico	consejero
crítico	coordinador

	reprobador	dirigente	
	juzgador	conductor	
	opositor	iniciador	
Sd baja			Sd alta
	evitador	consentidor	
	otorgador	conciliador	
	renunciador	participante	
	alejado	cooperador	
	evasivo	complaciente	
	Do baja		

Las escalas So (Socialización) y Ac (Auto-control), ofrecen otro interesante grupo de variables para análisis. Esas dos escalas correlacionan en $+0.50$ - una con otra.

CARACTERISTICAS.

		So alta	
	obstructivo	serio	
	crítico	sereno	
	dominante	auto-controlado	
	persistente	estable	
Ac baja			Ac alta
	agresivo	mentiroso	
	exigente	defensivo	
	excitable	terco	
	refractario	desconfiable	
	So baja		

Como un último ejemplo de las interacciones de las escalas, se consideran - los dos índices de logro, La (Logro vía adaptación) y Li (Logro vía independencia). Esas dos escalas dan una intercorrelación típica de $+0.40$.

CARACTERISTICAS.

		La alta	
	condescendiente	eficiente	
	diligente	maduro	
	moderado	organizado	
	tranquilo	estable	
Li baja			Li alta
	torpe	exigente	
	rudo	dominante	
	auto-defensivo	independiente	
	superficial	mordaz	
	La baja		

INTERPRETACION DE LOS PERFILES. (15)

En el uso operacional de un inventario como el IPP, el usuario está ordinariamente más interesado en la imagen total, global, de la persona que está siendo examinada que en el resultado de una escala simple. La mejor guía para una apreciación global organizada es proporcionada por el perfil completo del inventario.

La interpretación del perfil o interpretación configurativa, como frecuentemente es llamada, es en algo un arte que depende de la sutileza psicológica y de la experiencia práctica con el instrumento de examen que está siendo usado, así como del conocimiento técnico y de investigación del test. El dominio de los detalles específicos de información concerniente al test, es una precondición necesaria para la interpretación apropiada, pero más allá de esto el usuario debe depender de su propia experiencia y conocimientos psicológicos conforme vaya observando el funcionamiento de la prueba en su práctica cotidiana.

En beneficio de los usuarios potenciales del IPP que no hayan tenido una amplia experiencia con test multidimensionales de personalidad, los autores sugieren una aproximación a la interpretación del perfil que puede ser útil e ilustrar su uso. No se indica particularmente que sea superior a cualquier otra aproximación y se ofrece como un útil punto de partida desde el que se pueden desarrollar técnicas particulares para la interpretación del perfil.

PASO 1.- Se analiza la elevación global del perfil. Si casi todos los puntajes se encuentran por arriba de las líneas del puntaje estándar medio, las probabilidades son de que la persona está funcionando en forma efectiva tanto social como intelectualmente. Por el contrario, si la mayoría de los puntajes están por abajo de la media, hay buenas probabilidades de que el individuo esté experimentando dificultades significativas en un ajuste interpersonal.

Este juicio inicial, normalmente se hace en relación a la línea base general de 50. Los usuarios del test que trabajan con ciertos grupos también pueden hacer comparaciones del perfil de un individuo con normas específicas. Por ejemplo, un consejero de Preparatoria podrá basarse en los puntajes medios para estudiantes de Preparatoria, dado que sus puntajes promedios caen cinco o diez puntos por abajo de la línea base de 50. Los puntajes medios de diversos grupos educacionales y ocupacionales pueden ser vertidos a perfiles y referencias específicas.

PASO 2.- Se observa la elevación diferencial de los cuatro grupos de escalas. Habiendo hecho una rápida apreciación del nivel general de funcionamiento, se puede entonces buscar los grupos de escalas altas y bajas. Por ejemplo, si las seis escalas de la Clase I (Do a Sb) tienden a ser mayores que las escalas de la Clase III (La, Li y Ei), se puede esperar un individuo cuyas habilida

des sociales están altamente desarrolladas, pero cuyas fuerzas intelectuales y académicas son débiles.

PASO 3.- Se registran las escalas más altas y más bajas. Para lograr una interpretación más específica se examinan los puntos altos y bajos estudiando los adjetivos asociados y las descripciones conductuales de las páginas anexas. Mientras más extremos son esos puntajes, más probable es que un conjunto particular de adjetivos de los sumarios caracterice adecuadamente a una persona. Aquí deben considerarse los principios de interacción de las escalas. Cuando las conductas sugeridas por dos o más puntajes extremos parecen ser opuestas o contradictorias, es un dato que puede servir para contrarrestarlas o disminuirlas recíprocamente, es éste el aspecto en que la experiencia con el test capacita al usuario para desarrollar mejores hipótesis y para hacer predicciones más válidas.

PASO 4.- Se estudian las características únicas del perfil. Es axiomático que no hay dos seres humanos idénticos; similarmente, la mayoría de los perfiles del IPP revelarán ciertas características únicas o idiosincráticas; por ejemplo, una combinación rara de altos o bajos puntajes, alguna desviación poco común de la norma, algún "pico" excepcional de alguna escala, y así por el estilo. Algunas de esas variaciones se resistirán a la explicación o a la interpretación, pero aún cuando no puedan ser inmediatamente comprendidas, deberán ser observadas y consideradas.

PASO 5.- Se considera la variabilidad interna del perfil. Muchos perfiles tienen una línea de puntaje medio que difiere apreciablemente de la media del perfil o de la media de cualquier grupo particular. Por ejemplo, la mayoría de los puntajes puede oscilar entre T60 y T70, dando una media personal de T65. En tales casos, las desviaciones de las escalas individuales de esta "media personal" pueden tener una significación creciente para aclarar las características conductuales predominantes del individuo.

Es en este punto también, en los que cualesquiera propósitos especiales de la prueba deberían ser considerados. Si, por ejemplo, se están seleccionando candidatos para trabajar en Ingeniería de Ventas, un puntaje de T60 en Sd o en Ps puede no ser suficientemente alto, aunque éstos se encuentren por encima de la media de la población. Por otro lado, si se intenta predecir qué tan bien se desempeñará en la Preparatoria un estudiante de habilidad y logro promedios, puntajes de T60 en So y en Li serían "altos" en el sentido estadístico y predictivo.

LAS ESCALAS Sb, Bi Y Cm. (15)

La simulación deliberada, tergiversación o falseamiento de los tests psicológicos probablemente es un fenómeno mucho menos común de lo que se piensa, aunque este fenómeno no puede ser ignorado en un test diseñado para usarse en diferentes situaciones, algunas de las cuales pueden dar lugar a motivaciones-

para presentar una impresión particular sea buena o mala.

El IPP, como otros instrumentos empíricamente desarrollados, no es particularmente fácil de falsear, debido a que el método de la construcción de las escalas produce muchos ítems sutiles; ítems cuyas relaciones con cualquier rasgo particular o conducta no son obvios a la inspección (V.gr.: "La educación es más importante que lo que la mayoría de la gente piensa"). Indudablemente, muchos ítems son menos sutiles y han sido desarrolladas tres escalas para ayudar a detectar a aquellos sujetos que deliberadamente o de otro modo exageran o distorsionan sus respuestas al inventario: Bi (Buena impresión) Sb (Sentido de Bienestar) y Cm (Sentido común).

Los métodos de construcción empleados para estas escalas son descritos - más adelante, pero su uso para identificar protocolos invalidados, pueden ser resumidos aquí:

ESCALA Bi (Buena impresión).- Los resultados obtenidos en esta escala indican, según el autor, su papel para ayudar a identificar los intentos exagerados del sujeto por colocarse en un ángulo favorable. Los puntajes ligeramente por encima del promedio (T 50) en Bi son indicadores de actitudes favorables y de esfuerzos concientes por ajustarse y adaptarse. Puntajes muy altos suscitan la posibilidad de "falseamiento" de la prueba, o de un interés excesivo por causar buena impresión.

Escala Bi (Buena Impresión): Composición, Dirección Crítica y Superposición con otras escalas.

Bi, Cierto

127.- Siempre trato de considerar el sentir de los demás antes de hacer algo.

133.- Me siento tan bien ahora como siempre lo he estado.

140.- Me gustan las conferencias sobre acontecimientos mundiales.

165.- No me molesta que me den órdenes y que me digan lo que tengo que hacer.

195.- Las cosas más importantes para mí son mis obligaciones para con mis compañeros y mi trabajo.

222.- Me gustaría pertenecer a un círculo o grupo de estudios y debates.

254.- Nunca he dicho una mentira deliberadamente.

Bi, Falso

34.- Hay algunas personas en quienes no se puede confiar.

- 56.- Detesto ser interrumpido cuando estoy trabajando en algo.
- 109.- Algunas veces me siento desalentado.
- 153.- Algunas veces cuando no me siento bien, estoy malhumorado.
- 159.- Me pongo nervioso si tengo que reunirme con mucha gente.
- 203.- Cuando las cosas salen mal, a veces culpo a otros.
- 238.- Algunas veces simplemente siento que no puedo seguir adelante.
- 262.- Ha habido ocasiones en las que he sido muy duro con otra persona.
- 268.- A veces he estado muy ansioso por zafarme de mi familia.
- 293.- De vez en cuando estoy de mal humor y nadie puede hacer nada para - complacerme.

Bi, Ac, Falso

- 44.- Algunas veces siento como si destrozara las cosas.
- 48.- La mayoría de la gente mentiría si pudiera ganar algo con ello.
- 66.- A veces siento deseos de maldecir.
(MMPI.- Escalas L, K, 2, 3: falso).
- 81.- Debo admitir que a menudo he tratado de salirme con la mía, inde---
pendientemente de lo que otros puedan querer.
- 91.- Algunas veces pienso cosas demasiado malas para hablar de ellas.
- 120.- No siempre digo la verdad.
(MMPI.- Escala L: falso).
- 178.- La mayoría de la gente se alegra por dentro cuando alguien se mete-
en aprietos.
- 248.- Tengo que admitir que tengo mal genio, una vez que me enojo.

Bi, Falso; Ps, Cierto

- 10.- Algunas personas exageran sus desdichas para que se condelan de -
ellas.

Bi, Ps, Falso

- 150.- La crítica o el regaño me hieren profundamente.
(MMPI.- Escalas 2, 0: cierto; K: falso).

Bi, Cierto; Ps, Falso

14.- Siempre sigo esta regla: primero el deber, después el placer.

Bi, Sb, Falso

70.- A veces cruzo la calle para evitar encontrarme con alguien.

Bi, Cs, Cierto

103.- Me gusta escuchar conciertos sinfónicos en la radio.

Bi, Falso; Aa, Cierto

101.- Si pasa desapercibido, hago tan poco trabajo como puedo.

Bi, Pi, Falso

289.- Ha habido ocasiones en las que he estado muy preocupado por algo -
que realmente no era importante.

Bi, Falso; Ps, Ei, Cierto

30.- A veces chismeo o murmuro un poco de la gente.
(MMPI.- Escala L: falso).

Bi, Aa, Fe, Falso

38.- Es difícil para mí iniciar una conversación con extraños.

Bi, Ac, Falso; Aa, Cierto

44.- Algunas veces aparento saber más de lo que realmente sé.

Bi, Ac, Fe, Falso

78.- Me gusta vanagloriarme de mis triunfos.

Bi, So, Ac, Falso; Ps, Cierto

170.- A menudo actúo impulsivamente, sin detenerme a pensar.

Bi, Falso; Do, Pi, Cierto

207.- Algunas veces voto en las elecciones por gente que casi no conozco.
(MMPI.- Escala L: falso).

Bi, Ac, Falso; Ps, Cierto

231.- Soy capaz de lucirme de alguna manera si tengo la oportunidad.

Bi, Falso; Ps, Aa, Cierto

275.- Algunas veces me gusta ir contra las reglas y hacer cosas que nunca imaginé.

ESCALA Sb (Sentido de Bienestar).- Uno de los propósitos de esta escala es el de identificar personas que tienden a exagerar excesivamente sus preocupaciones y problemas, minimizando su bienestar personal. Las personas psiquiátricamente enfermas califican, por lo general, debajo del promedio en Sb, pero los puntajes excepcionalmente bajos (T40 ó menos), se encuentran entre sujetos que intentan falsear la prueba (Gough, 1969). Por lo tanto, los puntajes en el rango moderadamente bajo (T41 a T49) pueden ser considerados como indicadores de un sentido de bienestar reducido, en tanto que los puntajes por debajo de estos límites, dan lugar a considerar que el sujeto está exagerando indebidamente sus problemas personales y fingiendo francamente una posición irreal.

Escala Sb (Sentido de Bienestar): Composición, Dirección Crítica y Superposición con otras Escalas.

Sb, Cierto

312.- Cualquier persona capacitada y dispuesta a trabajar fuerte, tiene - buenas posibilidades de obtener éxito.
(MMPI.- Escala F: falso).

313.- Muy raras veces siento dolor en la nuca.
(MMPI.- Escala 1: falso).

Sb, Falso

299.- No parece importarme lo que me pase.
(MMPI.- Escalas 2 y 8: cierto).

- 301.- Me dá miedo estar solo en la oscuridad.
(MMPI.- No codificada: falso *).
- 306.- Me dan pesadillas con mucha frecuencia.
(MMPI.- Escala F: cierto).
- 308.- Sufro mucho de trastornos estomacales.
(MMPI.- Escala 1: cierto).
- 309.- He tenido miedo a cosas y a personas que sabía que no me podían hacer daño.
(MMPI.- Escalas 7 y 8: cierto).
- 318.- Cuando era niño nunca me interesó pertenecer a un grupo o pandilla.
(MMPI.- No codificada: falso).
- 325.- Cuando me siento muy feliz y activo, alguien que esté deprimido me desanima por completo.
(MMPI.- No codificada: falso).
- 330.- Todo me sabe igual.
(MMPI.- Escalas F y 8: cierto).
- 337.- La mayor parte del tiempo parece dolerme toda la cabeza.
(MMPI.- Escala 3: falso).
- 341.- Mi familia me trata más como un niño que como adulto.
(MMPI.- No codificada: falso).
- 344.- Ciertos animales me ponen nervioso.
(MMPI.- No codificada: falso).
- 351.- Algunos de mis familiares tienen hábitos que me molestan y perturban mucho.
(MMPI.- Escalas 5 y 9: cierto).
- 353.- Nadie parece comprenderme.
(MMPI.- Escalas 4, 6 y 8: cierto).
- 358.- Sueño frecuentemente acerca de cosas que preferible mantener en secreto.
(MMPI.- Escala 8: cierto; escala 2: falso).
- 372.- Tengo motivos para sentirme celoso de uno o más miembros de mi familia.
(MMPI.- Escala F: cierto).

* La anotación "no codificada", se refiere a que el ítem en cuestión aparece en el MMPI, pero no se califica para ninguna de las escalas básicas. La anotación "cierto" o "falso" indica la dirección en que el reactivo es contestado por la mayoría de los sujetos (según el Manual).

- 375.- Hay ciertas personas que me disgustan tanto que me alegro interiormente cuando están pagando las consecuencias por algo que han hecho.
(MMPI.- No codificada: falso).
- 381.- Siento la boca seca casi todo el tiempo.
(MMPI.- No codificada: falso).
- 388.- Cuando estoy en una situación difícil digo sólo aquella parte de la verdad que no me perjudique.
(MMPI.- No codificada: falso).
- 406.- Tengo uno o varios malos hábitos tan arraigados que es inútil luchar contra ellos.
(MMPI.- No codificada: falso.)
- 411.- Padezco de acidez estomacal varias veces a la semana.
(MMPI.- Escala 1: cierto).
- 415.- Me he sentido avergonzado por la clase de trabajo que alguien de mi familia ha hecho.
(MMPI.- No codificada: falso).
- 425.- A menudo me he sentido culpable porque he fingido mayor pesar del que realmente sentía.
(MMPI.- No codificada: falso).
- 430.- Ciertas cosas que han hecho algunos de mis familiares me han asustado.
(MMPI.- Escala 8: cierto).
- 437.- Sufro de ataques de náusea y vómito.
(MMPI.- Escalas F, 1, 2 y 3: cierto).
- 438.- Yo hubiera tenido mucho más éxito si la gente me hubiera dado una buena oportunidad.
- 449.- Casi todos los días sucede algo que me asusta.
(MMPI.- Escalas 7 y 8: cierto).
- 454.- A mi familia no le gusta el trabajo que he escogido (o el trabajo que pienso escoger para el resto de mi vida).
(MMPI.- Escalas F y 4: cierto).
- 455.- Casi siempre me parece que tengo un nudo en la garganta.
(MMPI.- Escalas 3 y 7: cierto).

Sb, To, Falso

- 15.- Varias veces por semana siento como si algo terrible fuera a suceder.
(MMPI.- No codificada: falso).
- 89.- Una vez a la semana, o más, repentinamente me siento caliente en todo el cuerpo, sin razón aparente.
- 266.- Creo que la mayoría de la gente mentiría para salir adelante.
(MMPI - Escalas 3 y 6: falso)

Sb, Ac, Falso

191.- Recuerdo haberme fingido enfermo para zafarme de algo.
(MMPI.- Escala 0: falso).

297.- A veces tengo un fuerte impulso de hacer algo dañino o escandaloso.
(MMPI.- Escalas 8 y 9: cierto).

Sb, Bi, Falso

70.- A veces cruzo la calle para evitar encontrarme con alguien.

Sb, Ps, Cierto

224.- Por lo general espero tener éxito en las cosas que hago.
(MMPI.- Escala F: falso).

259.- Generalmente siento que la vida vale la pena.
(MMPI.- Escala 2: falso).

Sb, Ei, Falso

434.- Mi piel parece ser muy sensible al tacto.
(MMPI.- No codificada: falso).

Sb, So, Falso

398.- La vida sólo me trata injustamente.

Sb, La, Falso

236.- Soy tan sensible acerca de algunos asuntos que ni siquiera puedo hablar de ellos.
(MMPI.- Escala 8: cierto).

Sb, Ac, Cierto

276.- Tengo muy pocos disgustos con mi familia.
(MMPI.- Escala K: cierto; escala 4: falso).

Sb, Do, Ps, Ac, Pi, Falso

54.- Encuentro difícil concentrarme en una tarea o trabajo.
(MMPI.- Escalas 2, 3, 4, 7, 8 y 0: cierto).

ESCALA Cm (Sentido Común).- Dentro del rango usual de los puntajes, la magnitud de la calificación en Cm, indica el cuidado y escrupulosidad con la que el individuo se aproxima a la prueba. Cuando el puntaje cae muy bajo, se crea la fuerte posibilidad de que sus respuestas hayan sido, en cierta forma, dadas al azar o sin significado.

Escala Cm (Sentido Común): Composición, Dirección Crítica y superposición con otras escalas.

Cm, Cierto

- 316.- Mis padres querían hacer de mí alguien de éxito en la vida.
- 322.- No me gusta prestar mis cosas a gente descuidada.
- 333.- La educación es más importante de lo que la mayoría de la gente piensa.
- 342.- Algunas personas exageran sus desdichas sólo para obtener simpatía.
- 343.- La mayoría de los maestros fueron imparciales y justos conmigo.
- 348.- Generalmente trato de hacer aquello que se espera de mí y evitar así la crítica.
- 371.- Preferiría ser un trabajador confiable y estable a uno brillante pero inestable.
- 410.- Considero como cosa muy importante el derecho de decir lo que uno piensa.
- 426.- A veces he sido de carácter muy violento.
- 427.- Hay algunas cuantas personas en las que sencillamente no se puede confiar.
- 440.- A veces he estado desanimado.
- 445.- Lucharía si alguien tratara de arrebatarme mis derechos.
- 446.- Debo admitir que algunas veces la gente me desilusiona.
- 447.- Si viera algunos niños lastimando a otro, seguramente trataría de detenerlos.

Cm, Falso

- 307.- Cuando manejo, trato de evitar que otros me rebasen.
- 311.- No puedo hacer nada bien.
(MMPI.- No codificada: falso).

- 321.- Casi nunca sueño.
(MMPI.- Escala 7: falso).
- 324.- Las campañas de la Cruz Roja no son más que un fastidio.
- 332.- Podría ser perfectamente feliz sin ninguna amistad.
- 349.- Si una persona es lo suficientemente hábil como para estafarle a alguien una gran cantidad de dinero, debe permitírsele quedarse con él.
- 350.- No debe esperarse que una persona haga algo por su comunidad a me--nos que se le pague.
- 360.- Es imposible que un hombre honesto progrese en la vida.
- 366.- Parece que nunca me enojo.
- 374.- Nunca me molestaría ayudar a otra persona, aunque ello signifique -renunciar a alguna comodidad personal.
- 378.- Dudo de si alguien es realmente feliz.
- 384.- La mayoría de la gente estaría en mejores condiciones si nunca hubiera estado en la escuela.
- 401.- La mayoría de los jóvenes tienen suficientes oportunidades de estu-diar.
- 421.- Realmente me tiene sin cuidado si le gusto o no a la gente.

El Manual de la prueba señala algunos de los estudios de validez realiza-dos con estas escalas. La tabla siguiente muestra los puntajes medios obteni-dos de un estudio en el que se pidió a los sujetos que tomaran la prueba falsi-ficando en sentido "positivo" o "negativo" .

PUNTAJES PROMEDIO Y DESVIACIONES ESTANDAR DE ESTUDIANTES
DE BACHILLERATO QUE INTENTARON FALSEAR LOS RESULTADOS.

Escala	H O M B R E S				M U J E R E S			
	Distorsión "positiva"		Distorsión "negativa"		Distorsión "positiva"		Distorsión "negativa"	
	N = 8		N = 6		N = 14		N = 13	
	Prom.	DE	Prom	DE	Prom	DE	Prom	DE
Do	36.8	4.7	20.8	7.7	34.1	4.6	19.7	8.0
Cs	24.9	2.3	12.2	4.7	25.4	3.3	13.6	4.3
Sd	28.9	3.0	15.8	6.0	30.7	3.8	15.3	6.8
Ps	37.6	3.0	27.7	6.4	37.8	3.5	25.7	9.2
Aa	22.4	2.1	20.2	3.2	22.9	3.1	18.8	5.4
Sb	41.8	1.6	11.7	7.4	42.4	1.5	17.2	10.4
Re	35.3	3.4	14.2	7.2	37.4	4.1	17.5	10.9
So	41.0	3.5	18.7	4.6	43.4	2.4	24.4	7.8
Ac	38.1	4.8	11.8	5.0	40.3	5.2	18.7	7.5
To	26.4	2.1	6.2	2.9	28.6	2.3	11.2	4.2
Bi	30.6	7.8	8.0	2.8	31.6	6.3	13.8	7.5
Cm	25.8	1.2	17.2	3.7	25.7	1.7	16.3	7.0
La	32.6	4.2	11.8	3.4	34.3	3.4	18.6	6.4
Li	19.6	2.9	11.2	3.6	21.4	2.8	13.3	2.6
Ei	41.9	3.1	20.2	7.6	43.6	4.8	23.3	7.0
Pi	13.3	1.5	6.8	2.6	13.4	2.1	9.0	2.1
Fx	7.9	3.1	6.3	1.7	8.5	3.7	8.8	3.7
Fe	16.9	2.8	18.2	5.0	21.9	2.7	21.3	6.1

(Todos los datos en puntajes crudos).

Los protocolos falsificados "positivamente" se encontraron por encima del promedio virtualmente en todas las escalas; su punto más alto, como cabría esperar, se encuentra en la escala Bi (T68 para los varones y T70 para las mujeres). Este hallazgo coloca al perfil simulado aparte de la mayoría de los perfiles obtenidos bajo situaciones de prueba normal.

Los perfiles falsificados "negativamente", están abiertamente distorsionados; los puntajes en Sb se encuentran por debajo de los niveles observados en situaciones ordinarias de prueba (T-5 para los hombres y T-4 para las mujeres), en tanto que otros puntajes se encuentran también disminuidos por debajo de los límites normales: Ei (hombres, T 8; mujeres, T17), Cm (hombres, T13; mujeres, T 5), La (hombres, T16; mujeres, T29) y To (hombres, T14; mujeres, T22).

Otro problema que se presenta en algunas situaciones, es la identificación de perfiles de personas cuyas respuestas han sido esencialmente dadas al azar. Confusiones de este tipo pueden deberse a falta de cuidado para aparear los números de los ítems en el cuadernillo con el número correspondiente de la hoja de respuestas, a inhabilidad para leer los reactivos o a deliberada falta de cooperación.

Para obtener evidencia en este sentido, 30 protocolos del IPP fueron llenados sobre la base de una tabla de números aleatorios; cada número par se clasificó como una respuesta cierta y cada número non como una falsa. Los 30 protocolos fueron calificados obteniéndose como resultados más relevantes los siguientes:

Tanto el perfil de hombres como el de mujeres abatieron significativamente los puntajes de las escalas Cm (T cero) Sb (T10) y Ei (T20). El pico en Cm, resulta extremadamente bajo y el mayor puntaje individual en dicha escala fue de 20 (T25); este valor está en el 1% inferior de puntajes en la escala Cm. El valor medio de 13.8 está por debajo de cualquier puntaje individual que hayamos observado en más de 3,000 registros de prueba. Esto parece sugerir que cualquier protocolo producto de respuestas al azar, resulta detectable por los bajos puntajes de la escala Cm, máxime si se consideran otras características distintivas, como son el abatimiento en las escalas Sb y Ei.

Como último comentario de este inciso, debemos mencionar el hecho de que los perfiles de la prueba pueden verse influidos no solamente por la distorsión deliberada, sino también por la tendencia de algunos sujetos a exagerar o minimizar las propias virtudes. Esta tendencia puede entenderse mejor como la forma característica que un individuo tiene para percibirse a sí mismo y a los demás y, como tal, es igualmente una parte de su personalidad como cualquier otra tendencia conductual. El método empírico de construcción de las escalas utilizado para desarrollar la mayoría de ellas, reconoce y utiliza este hecho. Por ejemplo, si los sujetos denominados "responsables", también tienden a ser más autocríticos que otras personas, un análisis de sus respuestas deberá revelar esta tendencia. Si los sujetos de inteligencia superior tienden a hablar de sí mismos en forma socialmente deseable, una escala empíricamente desarrollada deberá incluir una cierta proporción de tales actitudes. De acuerdo con la estructura de la prueba no se requieren escalas especiales ni métodos muy elaborados para tratar con este tipo de tendencia de respuestas.

DESARROLLO DE LAS ESCALAS. (15)

El método fundamental empleado en la construcción de las escalas del IPP, ha sido denominado "técnica empírica". Este método busca, en primer lugar, - definir el criterio o dimensión que se pretende medir. En segundo término, - se estructura una escala preliminar a partir de aquéllos ítems que parecen tener relevancia psicológica con respecto a la dimensión evaluada; estos reactivos son administrados a sujetos que demuestran, por algún otro procedimiento - totalmente independiente de la prueba, tener características significativamente similares a las de la dimensión en cuestión, analizando, por último, la forma en que tales sujetos responden a los ítems preliminares.

El objetivo de este análisis de reactivos es descubrir aquellas cuestiones que son respondidas en un sentido diferencial por los sujetos empleados para responder a la escala; es decir, con una alta frecuencia de respuestas para "falso" o para "cierto" en determinado ítem. Los reactivos desarrollados por esta "técnica empírica", establecen ciertas discriminaciones que se van seleccionando para posteriores estudios o ajustes. La respuesta "correcta" a un ítem específico, es aquella que tiene una alta frecuencia de respuestas dada por los sujetos que se considera que poseen en alto grado la característica evaluada.

El resultado de esta serie de pasos para cualquiera de los criterios que se trata de evaluar, es la identificación de 30 o 40 ítems, cada uno de los cuales posea una relación demostrable con la conducta que está siendo medida. Once de las escalas del IPP fueron desarrolladas por esta técnica (Do, Cs, Sd, Re, So, To, La, Li, Ei, Pi y Fe).

Cuatro escalas adicionales fueron desarrolladas por medio de un análisis de consistencia interna (Ps, Aa, Ac y Fx). Este método se inicia por la configuración de una escala previa, integrada por reactivos que tienen relevancia teórica o predictiva con respecto al criterio que se trata de medir o a un patrón conductual específico. En este caso el experimentador asigna un peso específico a cada ítem. El conjunto de ítems se administra entonces a una muestra de sujetos y se obtiene un registro del puntaje obtenido por cada sujeto. Se identifica a los sujetos con calificaciones elevadas en el conjunto preliminar de ítems y se hace lo mismo con aquéllos que tienen puntuaciones bajas. Como siguiente paso, se enfoca el análisis a identificar aquellos reactivos que diferencian más significativamente entre ambas sub-muestras. Los ítems con mayor capacidad diferencial se conservan en la versión definitiva de la escala.

El mayor peligro de este método de construcción de escalas es que el experimentador puede incurrir, con relativa facilidad, en errores al decidir cuál debe ser el peso de un ítem o bien al decidir qué reactivos incluir (o excluir) en la escala preliminar; por ejemplo, el constructor del test pudiera pensar -

que el reactivo "Una vez que he tomado una decisión, difícilmente cambio de opinión" indica "flexibilidad" si es contestado como "falso". Si un estudio con personas consideradas "muy flexibles" mostrará que éstas (a) tienden a contestar el ítem como "cierto", o (b) no muestran una tendencia particular al contestar al ítem, entonces la decisión inicial del examinador hubiera sido incorrecta.

Debido a las limitaciones inherentes a la técnica de consistencia interna, ésta ha sido empleada en la construcción de escalas del I. P. P., únicamente en aquellos casos en donde es posible derivar este análisis a partir de muestras de gran tamaño y donde un análisis empírico de los reactivos resultaría impracticable. En las cuatro escalas en las que se empleó este método, se practicaron cuidadosas verificaciones de las escalas completas para determinar su validez empírica. Esto es, los sujetos para formar la muestra se obtuvieron de listas de nombres y por otros métodos independientes de la prueba y se estudió la capacidad de la escala total para identificar a estos sujetos. En las 4 escalas discutidas se obtuvieron niveles aceptables de validez.

En el desarrollo de la escala Sb (Sentido de Bienestar) se siguieron los procedimientos empíricos mencionados antes, pero debido a sus particulares características, requiere discusión especial. La principal función de la escala es la de identificar a aquellas personas que menosprecian su bienestar y exageran sus preocupaciones y desgracias y distinguir, por otra parte, a aquellos sujetos que presentan una imagen objetiva y precisa de sus problemas e inquietudes. Con el fin de obtener un fundamento para la tendencia a responder en este sentido, se pidió a los sujetos del grupo experimental, que respondieran al inventario simulando sufrir severos conflictos internos. El problema consistió en identificar, bajo tales circunstancias, los malestares reales de los que no lo eran. También se examinó a pacientes psiconeuróticos hospitalizados y sus respuestas fueron consideradas como un índice de problemática real. Las respuestas elegidas para ser incluidas en la escala fueron aquéllas que diferenciaron entre los sujetos del grupo experimental y las personas con trastornos reales. La escala resultante reveló una elevada validez, como más adelante se indica. Las muestras con casos psiquiátricos califican ligeramente por debajo del promedio en Sb, como es de esperarse, pero puntajes muy bajos se encuentran entre sujetos que deliberadamente pretenden dar una imagen de preocupación, dudas y baja moral.

La escala Bi (Buena impresión) también intenta evaluar algunos aspectos --tendenciosos al responder al inventario. En la construcción de esta escala, los sujetos que integraron las muestras, respondieron los ítems de la escala en dos ocasiones: la primera en condiciones normales de prueba y después bajo instrucciones de presentar la mejor imagen posible de sí mismos. Aquellos ítems que mostraron una mayor y más significativa alteración entre la primera y la segunda aplicación, se seleccionaron para ser incluidos en la escala. Puntajes muy elevados en la escala Bi, tal como marcados descensos en Sb, indican algún tipo de exageración consciente al respecto de sí mismo. Sin embargo, en ambas escalas, las desviaciones menores con respecto a la media no tienen implicaciones de distorsión; por ejemplo, sujetos cuyo puntaje se encuentra ligeramente -

por encima de T50 en Bi, tienden a ser más cooperadores, adaptables y abiertos que aquellos con bajos puntajes; aunque a partir de los datos observados, estos sujetos tienden, de hecho, a crear una impresión favorable de sí mismos.

Por último, la escala Cm (Sentido común), fue estructurada con 28 ítems - en el inventario, en los cuales se observó una gran concentración de respuestas que indican concordancia con el patrón modal del inventario; es decir, de la postura de determinados sujetos a aparentar estar de acuerdo con ciertas - premisas del inventario, cuando sus puntos de vista personales al respecto de tales enunciados son, en realidad diferentes (quedan excluidos de esto los simuladores detectados experimentalmente). Cada uno de estos ítems un tipo modal de concordancia o aceptación y el conjunto total de enunciados es un "común denominador" de las actitudes y convicciones de un sujeto. La designación de esta escala con el rubro de "sentido común" fue elegida para reflejar esta propiedad de la escala.

Los estudios al respecto de la confiabilidad de las escalas de validez señalan que en la situación test-retest, en períodos que oscilan entre 7 y 21 días, la escala Sb mantiene coeficientes entre .72 y .75; la escala Bi de .68 a .81 y Cm entre .38 y .58. Gough (1969) señala que los bajos coeficientes - arrojados por Cm, pueden atribuirse a la reducida longitud de la escala (28 reactivos), así como a que su distribución es extremadamente asimétrica (izquierda).

El problema de la confiabilidad de las escalas, aún cuando presenta algunas dificultades para estimarse, tiene la ventaja de poder representarse por un índice numérico simple. La validez de una escala es, en cambio, mucho más difícil de expresar y resumirse. Para una escala como Re (responsabilidad), puede estimarse su validez correlacionándola con algún criterio subjetivo de responsabilidad; pero estas estimaciones pueden resultar inexactas y falibles por sí mismas; sin embargo, estas correlaciones pueden tener algún grado de "verdad" con respecto a la responsabilidad social. Las escalas de validez del IPP han sido objeto de algunos estudios para verificar su validez, de los cuales reportamos, aquí, los más importantes. Debe indicarse que, en todos los casos, la evidencia que se presenta adelante fue obtenida por medio de estudios de validación cruzada del inventario.

SENTIDO DE BIENESTAR (Sb).

(a).- En la evaluación de una muestra de 100 oficiales militares estudiados en la Universidad de California, Sb correlacionó .26 con el promedio personal de "salud y vitalidad" y .27 con su propio promedio en "estado físico general".

(b).- 354 estudiantes de bachillerato fueron instruídos para responder al inventario, distorsionando los ítems de manera de presentar serios problemas personales y sentimientos de ansiedad. Sus puntajes fueron comparados con -

los de pacientes psiquiátricos y con las calificaciones de otros estudiantes.

M u e s t r a	N	M	DE
1.- Estudiantes que fingieron ansiedad	354	11.81	6.95
2.- Pacientes psiquiátricos	915	34.84	5.94
3.- Estudiantes de bachillerato	2800	37.46	4.51

Cada una de las tres muestras difiere de las otras dos de manera significativa (P .01). El valor de 37.46 es típico entre adultos de muestras no psiquiátricas. El puntaje menor encontrado entre sujetos psiquiátricos mantiene la validez de la escala. El extremadamente bajo puntaje obtenido por los sujetos que fingieron disturbios psiquiátricos corrobora la utilidad de Sb para -- identificar perfiles disimulados o distorsionados intencionalmente.

BUENA IMPRESION (Bi).

(a).- En una muestra de 152 adultos del sexo masculino, Bi correlacionó -- .60 con la corrección K del MMPI.

(b).- A un grupo de 179 estudiantes de secundaria se les pidió que respondieran a los reactivos del inventario intentando presentar "la mejor impresión posible" de sí mismos. Los resultados obtenidos en la escala Bi por este grupo y por una muestra no seleccionada de alumnos de secundaria fueron los siguientes:

M u e s t r a	N	M	DE
1.- Estudiantes que fingieron	179	23.87	8.48
2.- Estudiantes de secundaria	7628	15.37	6.20

dif.= 8.50

Raz. Crítica = 13.33

P .01

El mayor puntaje obtenido por los simuladores tiende a confirmar la validez de la escala y a corroborar su capacidad discriminativa para identificar -- protocolos distorsionados.

SENTIDO COMUN (Cm).

(a).- Al valorar una muestra de 100 oficiales militares, Cm correlacionó .28 con la frase: "Es confiable y práctico; tiene sentido común y buen juicio" y -.32 con la frase: "Está en conflicto consigo mismo, tiene problemas inter---nos."

C A P I T U L O I I I

ANALISIS DEL FUNCIONAMIENTO DE LAS ESCALAS L, F, K, Sb,
Bi Y Cm EN UNA SITUACION DE SELECCION DE PERSONAL.

Como ya se mencionó en la introducción de este trabajo, la observación del funcionamiento de las escalas señaladas en el rubro de este capítulo, muestran, al ser aplicados los cuestionarios respectivos, una configuración específica que parece mantener cierta estabilidad intersujetos, como ha podido observarse al través de aproximadamente 3500 pares de inventarios aplicados.

Dado que la fundamentación de un trabajo de esta naturaleza no puede conformarse, para ser llevado a cabo, con apreciaciones tan imprecisas, se procedió inicialmente a tomar una submuestra al azar de 100 sujetos (profesionales de las áreas de ingeniería), con el fin de verificar si tales configuraciones aparecían de hecho cuando eran cuantificadas. Se calcularon asimismo las correlaciones con este grupo, observándose que mostraban asociaciones significativas para la mayoría de los casos (mediana en 0.369). Visto lo anterior, además de que en conjunto los reactivos de las escalas estudiadas del MMPI sólo se encuentran superpuestos sobre las del IPP en una proporción de 1 a 8 (de los 111 reactivos que componen las escalas Sb, Bi y Cm, sólo 14 son idénticos a los otros 109 incluidos en las escalas L, F y K, aunque no siempre en la misma dirección crítica, tal como se señala en el capítulo anterior), se consideró la utilidad de realizar este estudio, en virtud de que los datos iniciales muestran cierta concatenación lógica entre las escalas de ambos inventarios, aun cuando las escalas incluyen reactivos diferentes.

A partir de estos antecedentes, se planteó la hipótesis de que en la situación de selección de personal estudiada, el número de características que condicionan la distorsión de los inventarios de personalidad empleados, puede ser explicada en función de un número de factores que debe ser menor que el número de escalas que, con distinta denominación, pretenden medir este efecto; con el objeto de poder sentar las bases para la construcción de escalas de validez con altas cargas factoriales, las cuales en un momento dado proporcionen aproximaciones métricas más precisas y permitan fundamentar, al menos, la hipótesis de que la consideración de las medidas arrojadas por cada una de las escalas, no puede, en todos los casos, considerarse como un criterio independiente o no contaminado con los aspectos presuntamente evaluados por otras escalas.

Ahora bien, la comprobación de esta hipótesis, presupone, asimismo, la verificación previa de algunos hechos y supuestos iniciales, que permitan apoyar los aspectos técnicos sobre los que descansa este estudio. (28)

El primero de estos supuestos, se refiere a que en la población la configuración ya señalada de las escalas se mantiene con cierta estabilidad; para poder demostrar este hecho, se calculan las distribuciones de frecuencias, las medias aritméticas, las desviaciones estándar, la curtosis y la asimetría de cada una de las escalas, sobre una muestra representativa de tal población. La segunda premisa que debe ser comprobada para realizar el análisis factorial de las escalas en relación a la hipótesis de trabajo, consiste en que las escalas deben mantener intercorrelaciones significativamente mayores de cero, que-

permitan prever que los rasgos fundamentales que explican la varianza común de las escalas están dados en un número menor que el número de escalas. (4, 28)

Si estos supuestos pueden ser demostrados objetivamente, el análisis factorial y el análisis de varianza pueden arrojar datos, debidamente fundamentados, para explicar adecuadamente la composición factorial de la varianza y sentar, de esta manera, las bases para la construcción de nuevas escalas de validez para estos inventarios, lo que de hecho, constituye la tesis sustentada por este trabajo.

Como un paréntesis en el análisis de los datos, se presenta un análisis de las diferencias entre las medias aritméticas de las escalas, por el método de clasificación múltiple de Duncan, en donde se señalan los niveles de significancia de tales diferencias, como un complemento de nuestro primer supuesto, referido a la configuración de las escalas estudiadas.

1).- LA MUESTRA.

La población de sujetos de las áreas profesionales de ingeniería que han tomado el MMPI y el IPP como parte del conjunto de 14 tests que les son aplicados en la evaluación psicométrica, del proceso de selección de personal que atañe a este trabajo, se componía, a la fecha de inicio de este estudio, de aproximadamente 3500 sujetos. Para la realización de esta tesis se integró la muestra con 337 sujetos, elegidos al azar por medio de una tabla de números aleatorios, la cual de acuerdo con Arkin y Colton (1963) (3) y con las características de nuestra población, opera con una representatividad probable de .99 y con un margen de error que no excede de $\pm 3\%$, quedando integrada de la manera siguiente:

COMPOSICION DE LA MUESTRA

PROFESION	NUMERO DE SUJETOS
Ingeniería Química	78
Ingeniería en Comunicaciones y Electrónica	51
Ingeniería Mecánica	42
Ingeniería Química Industrial	34
Ingeniería Eléctrica	28
Ingeniería Mecánica Electricista	26
Ingeniería Civil	14
Ingeniería Petrolera	13
Ingeniería Geológica	11
Ingeniería Química Petrolera	8
Ingeniería Industrial op. Química	7

PROFESION	NUMERO DE SUJETOS
Ingeniería Industrial op. Mecánica	6
Ingeniería Mecánica Electricista op. Industrial	3
Ingeniería Industrial op. Electricidad	3
Ingeniería Química Administrativa	3
Ingeniería Mecánica Administrativa	3
Ingeniería Mecánica Naval	1
Ingeniería Agrónoma	1
Ingeniería Mecánica op. Industrial	1
Ingeniería Mecánica Electricista op. Electricidad	1
Ingeniería Mecánica Electricista op. Control y Comunicaciones Electrónicas	1
Ingeniería de Transmisiones	1
Ingeniería Mecánica Electricista op. Mecánica	1

2).- DISTRIBUCION DE FRECUENCIAS, MEDIAS, DESVIACIONES ESTANDAR, CURTOSIS Y ASIMETRIA. (2, 30)

En esta sección se consignan los datos referidos al encabezado del inciso, los cuales fueron obtenidos a partir de las calificaciones crudas obtenidas por los sujetos de la muestra en las escalas en cuestión. En primer término se muestran los polígonos de frecuencias y las siguientes medidas, para cada ejemplo.

MEDIA ARITMETICA.- Se obtuvo por medio de:

$$\bar{X} = \frac{\sum_{j=1}^N X_j}{N}$$

donde: X_j = Puntajes crudos individuales

N = Número total de casos

DESVIACION ESTANDAR.- Esta medida de variabilidad se calculó por:

$$S = \sqrt{\frac{\sum_{j=1}^N X_j^2}{N} - \left(\frac{\sum_{j=1}^N X_j}{N}\right)^2}$$

CURTOSIS.- Se calculó el nivel de elevación de la distribución en relación con la curva normal, llamando leptocúrtica a una distribución alta, platicúrtica a la distribución baja o aplanada y mesocúrtica a la distribución normal. El procedimiento de cómputo fue:

$$S_2 = (m_4 / m_2^2)$$

donde m_4 es el cuarto momento a partir de la media aritmética, y m_2 el segundo. Los momentos a partir de la media se calcularon por:

$$m_r = \frac{\sum_{j=1}^N (x_j - \bar{x})^r}{N}$$

ASIMETRIA.- La asimetría de las distribuciones de las escalas corresponde al momento de tercer orden con respecto a la media, expresado en forma adimensional, y se computa por:

$$g_1 = \frac{m_3}{\sqrt{m_2^3}}$$

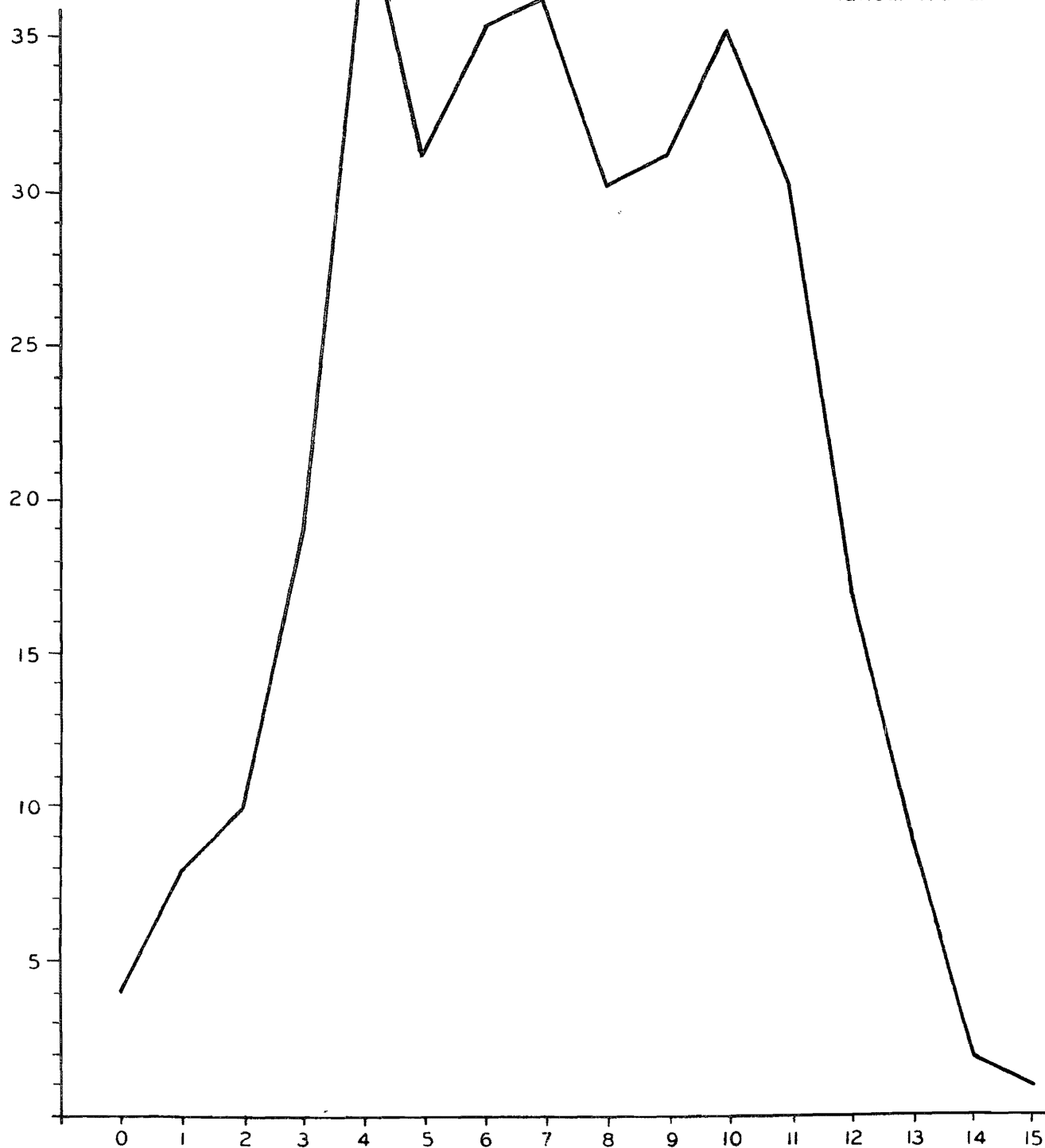
ESCALA L

Media = 7.184

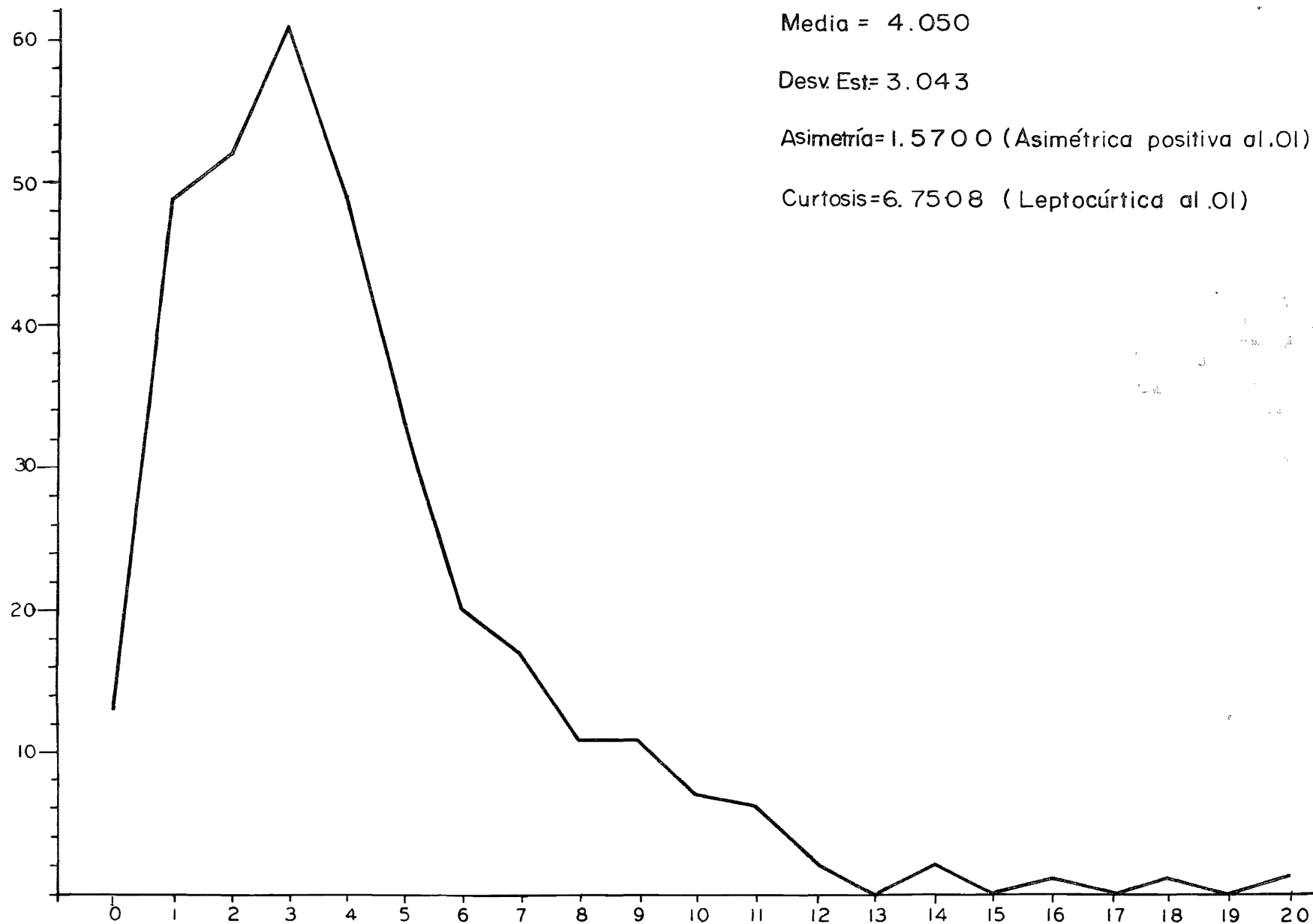
Desv. Est. = 3.175

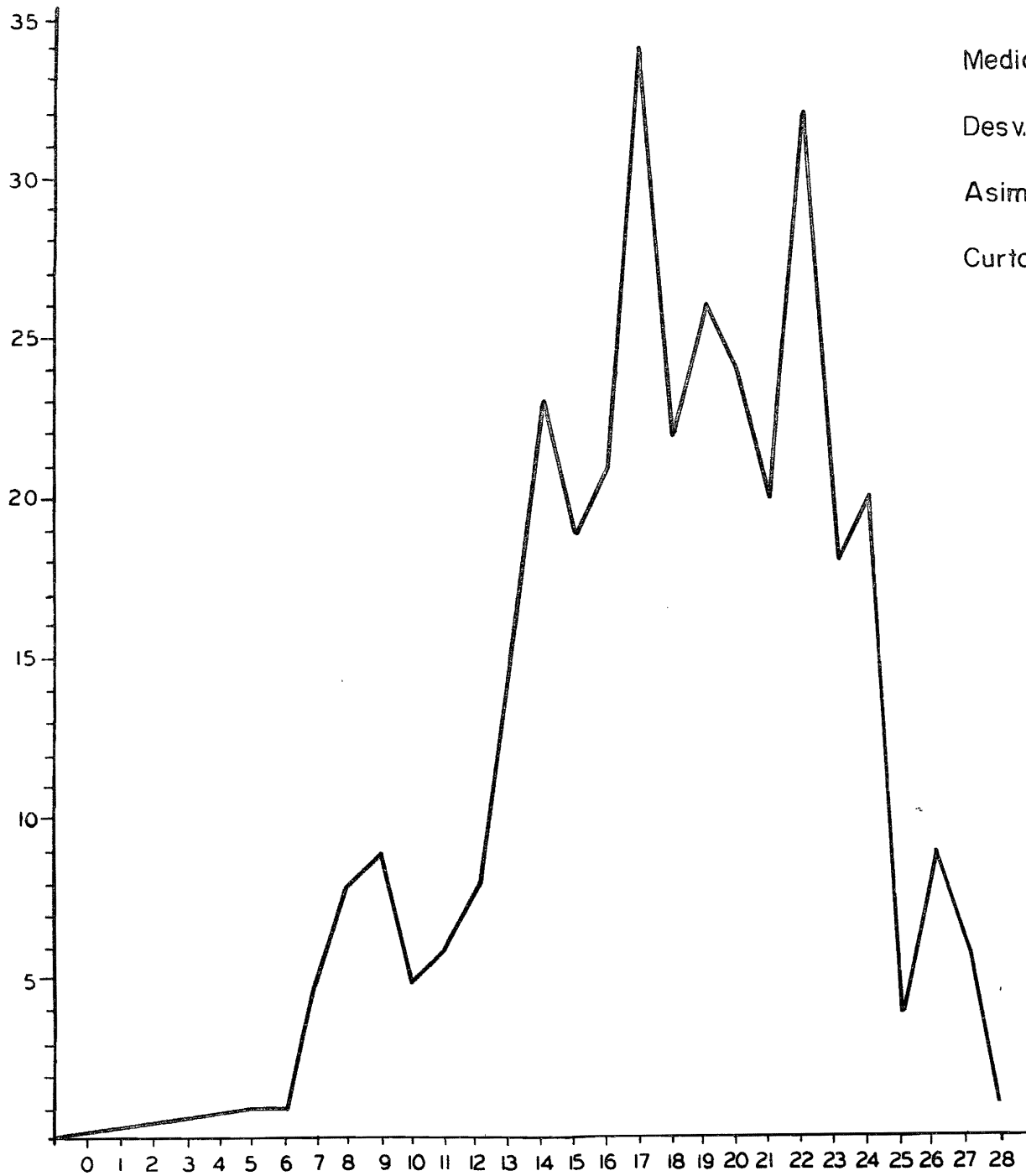
Asimetría = - 0.0988 (Simétrica)

Curtosis = 2.2371 (Platicúrtica al .01)



ESCALA F





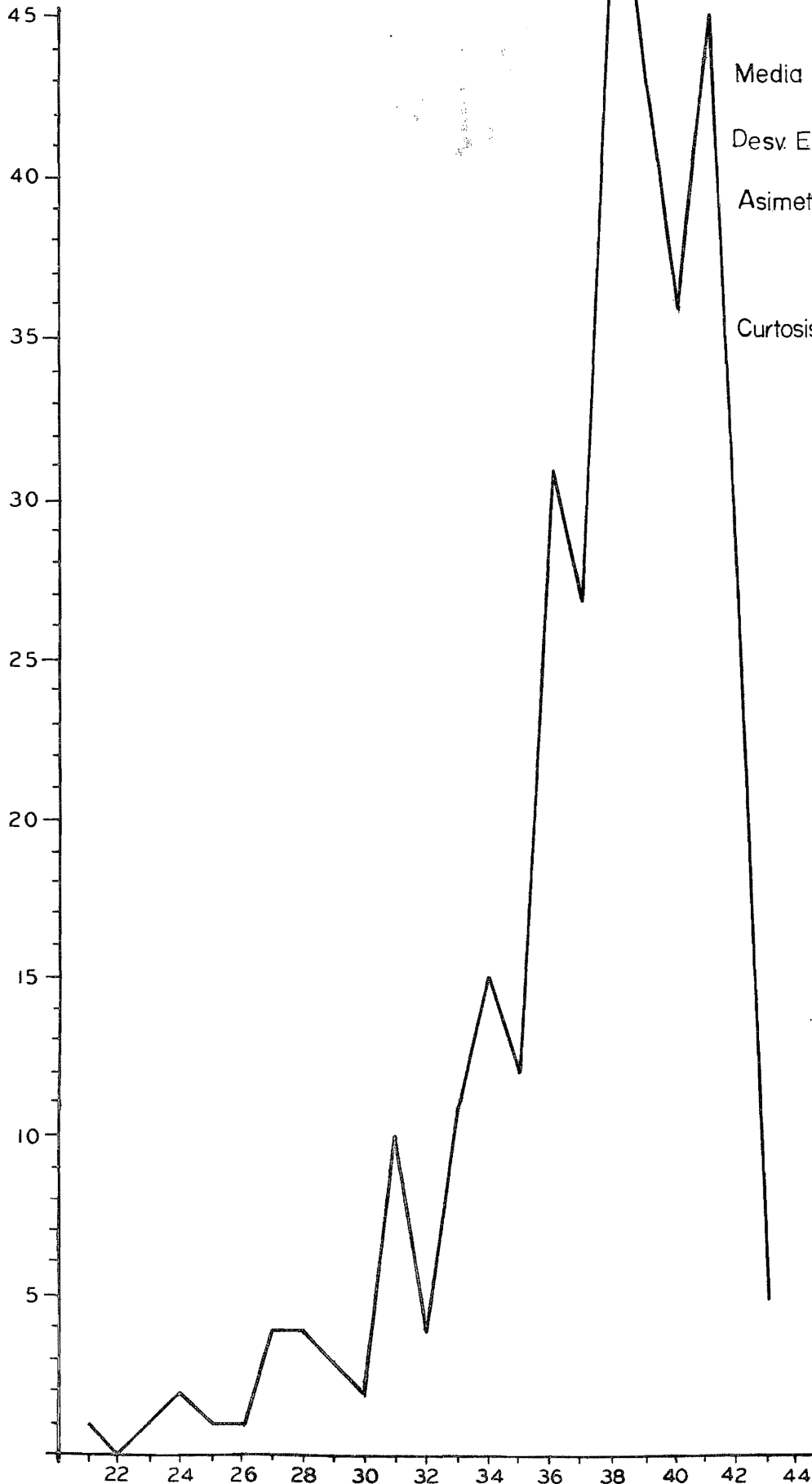
Media = 17.896

Desv. Est. = 4.786

Asimetría = -.3359 (Asimétrica negativa al .01)

Curtosis = 2.6103 (Platicúrtica al .05)

ESCALA Sb



Media = 37.525

Desv. Est. = 3.838

Asimetría = -1.3872 (Asimétrica
negativa al .01)

Curtosis = 5.2729 (Leptocúrtica al .01)

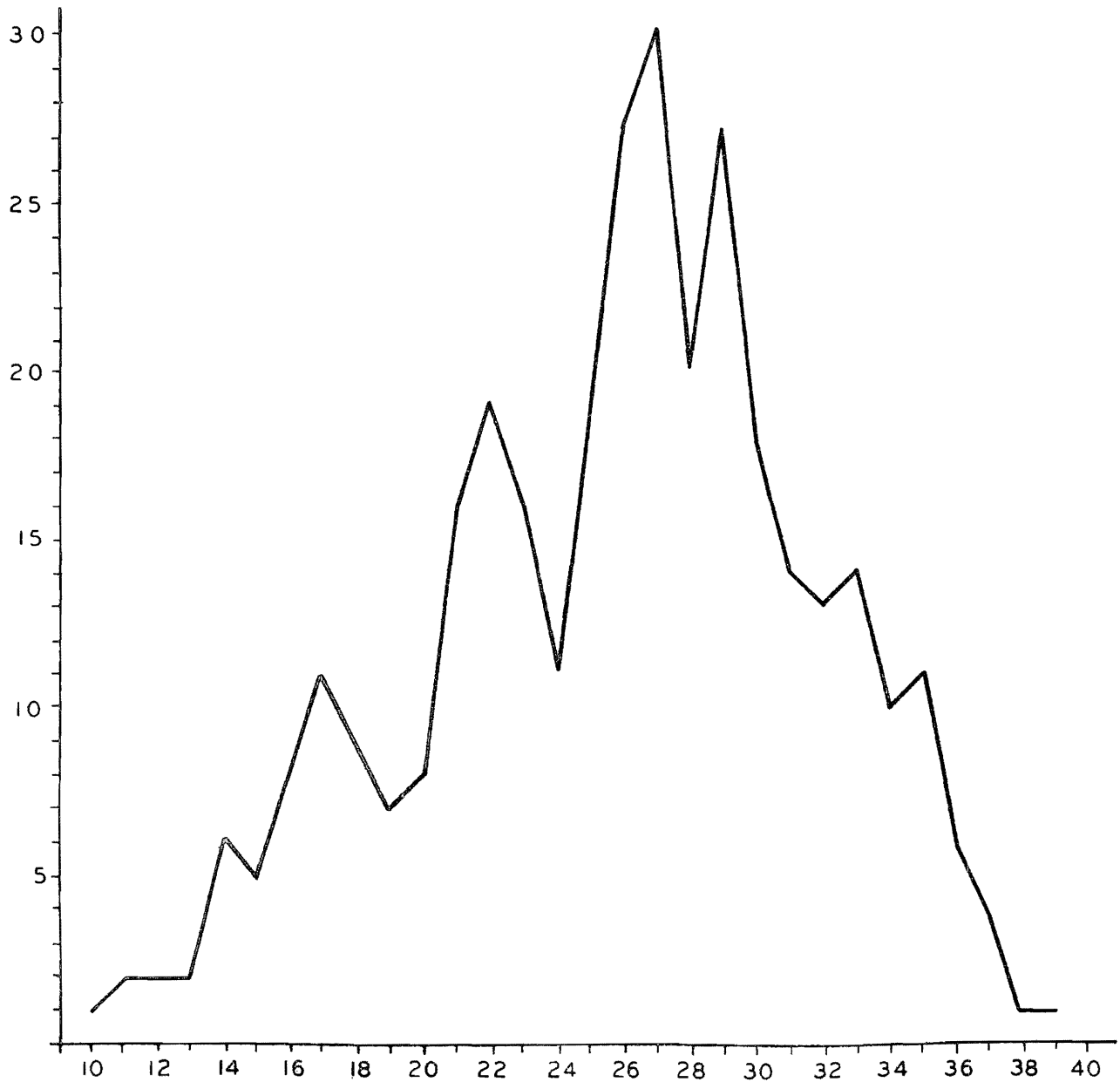
ESCALA B_i

Media = 25.875

Desv.Est= 7.723

Asimetría= -0.0277 (Simétrica)

Curtosis= 2.5888 (Platicúrtica al.05)



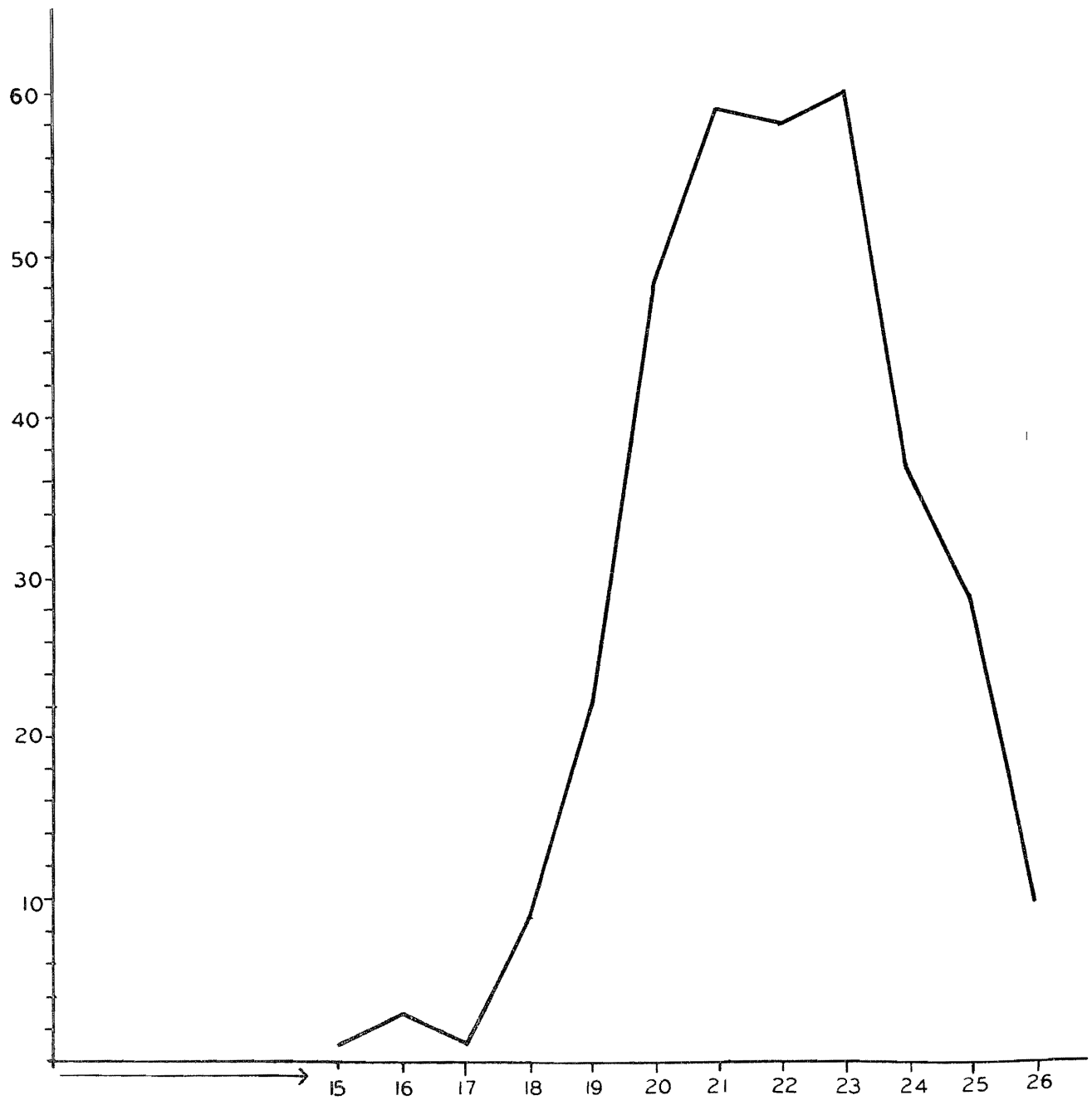
ESCALA Cm

Media = 21.923

Desv. Est. = 2.040

Asimetría = -0.2190 (Asimétrica negativa al .05)

Curtosis = 2.9399 (Mesocúrtica)



3).- DIFERENCIAS ENTRE MEDIAS.

(Clasificación Múltiple de Duncan). (22)

Duncan (1955) ha desarrollado un procedimiento que permite realizar comparaciones, por pares, entre medias. El error de medición para esta prueba es establecido por α * para cada grado de libertad. El nivel de confiabilidad para k medias en esta prueba es igual a $(1-\alpha)^{k-1}$. Para seis medias ordenadas jerárquicamente, cuando $\alpha = .01$, el nivel de confiabilidad es $100(1-.01)^{6-1} = 95\%$, lo cual es la probabilidad máxima de encontrar diferencias significativas entre medias apareadas. El nivel de confiabilidad decrece a medida que aumenta el número de medias; sin embargo, Duncan arguye que si $K > 2$ resulta más razonable esperar que existan algunas diferencias realmente significativas entre las medias que si K es solamente igual a 2. Consecuentemente, al incrementar K la prueba debe hacerse más poderosa, esto es, más adecuada para detectar diferencias reales. Este incremento en el poder de la prueba es obtenido a expensas del nivel de confiabilidad de la misma. Esta prueba, como las de Student (1927) y Newman-Keuls (1952), está basada en aproximaciones estratificadas (es decir, dependiendo del número de medias consideradas) a los tests de significancia.

La diferencia R_p que una comparación debe exceder para poder ser declarada significativa de acuerdo con el procedimiento de Duncan, es:

$$R_p = S_e \cdot r_p \alpha \sqrt{1/N}$$

donde r_p es obtenido de las tablas especialmente preparadas por Duncan para esta prueba.

El primer paso para el cálculo es ordenar las medias en orden creciente de magnitud, para lo cual los valores obtenidos en las escalas de validez por los sujetos de la muestra fueron convertidos a puntajes T (media aritmética = 50, desviación estándar = 10) para contar con una base homogénea para la comparación, de donde se tiene:

	Cm	Sb	F	Bi	L	K
N	337	337	337	337	337	337
ΣX	11 811	16 821	17 764	20 153	20 321	20 346
ΣX^2	443 537	870 643	955 234	1237 975	1263 059	1254 990
SS	29591.24	31070.51	18855.08	32801.68	37708.73	26622.89
\bar{X}	35.047	49.514	52.712	59.801	60.300	60.374

La media C_m se define como encontrándose a seis pasos de la media K , a cinco pasos de la media L , y así sucesivamente.

* α Nivel de significancia. Probabilidad de rechazar la Hipótesis Nula cuando ésta es verdadera.

De los datos de la tabla anterior, la suma de los cuadrados (SS) se obtiene por medio de:

$$SS = \frac{\sum X^2 - (\sum X)^2}{N}$$

Estos valores sirven de base para obtener la raíz cuadrada del error de variación (Se), como sigue:

$$Se = \sqrt{SS_1 + SS_2 + SS_3 + SS_4 + SS_5 + SS_6 / k (N-1)}$$

lo que para este caso resulta:

$$Se = \sqrt{176\ 650.13 / 6(336)} = 9.361$$

Se calculan los grados de libertad por medio de:

$$gl = N_{tot} - k = 2016$$

Se obtienen, de la tabla correspondiente, los valores r_p para $\alpha .01$ y $.05$ y se calculan las distancias significativas mínimas (Rp).

Número de Pasos	Valores r_p		Valores R_p	
	$\alpha = .05$	$\alpha = .01$	$\alpha = .05$	$\alpha = .01$
2	2.77	3.64	1.412	1.856
3	2.92	3.80	1.489	1.938
4	3.02	3.90	1.540	1.989
5	3.09	3.98	1.576	2.030
6	3.15	4.04	1.606	2.060

Con estos datos se formula el siguiente cuadro completo de comparaciones.

	Cm	Sb	F	Bi	L	K
\bar{Y}	35.047	49.914	52.712	59.801	60.300	60.374
Cm		14.867	17.665	24.754	25.253	25.327
Sb	.01 (2)		2.798	9.887	10.386	10.460
F	.01 (3)	.01 (2)		7.089	7.588	8.022
Bi	.01 (4)	.01 (3)	.01 (2)		.499	.573
L	.01 (5)	.01 (4)	.01 (3)	DNS (2)		.074
K	.01 (6)	.01 (5)	.01 (4)	DNS (3)	DNS (2)	

En donde:

- a) En la parte superior de la diagonal aparecen las diferencias entre las medias (en puntajes T).
- b) En la parte inferior de la diagonal se muestran los niveles de significancia α ; de la diferencia entre las medias apareadas.

DNS= Diferencia no significativa.

El número entre paréntesis indica el número de pasos entre cada par de medias.

4).- CORRELACIONES ENTRE LAS ESCALAS. (2, 23, 30)

En este inciso se calculan los coeficientes de correlación por medio de una derivación de la fórmula producto-momento de Pearson, utilizando los puntajes originales:

$$(1) \quad r = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{[N \sum X^2 - (\sum X)^2][N \sum Y^2 - (\sum Y)^2]}}$$

Se utiliza esta fórmula y no la que utiliza el cociente de la covarianza entre las varianzas de X y Y, por ser el cálculo de estos términos más complejo.

Todos los coeficientes de correlación reportados en este trabajo son lineales, y las líneas de regresión se obtuvieron por el método de mínimos cuadrados. La recta de mínimos cuadrados, se calcula en relación a las distancias perpendiculares de la línea de regresión predicha con respecto al eje de la variable predictora.

El tipo más sencillo de curva de aproximación al diagrama de dispersión, es la línea recta, cuya ecuación es:

$$(2) \quad Y = a + bX$$

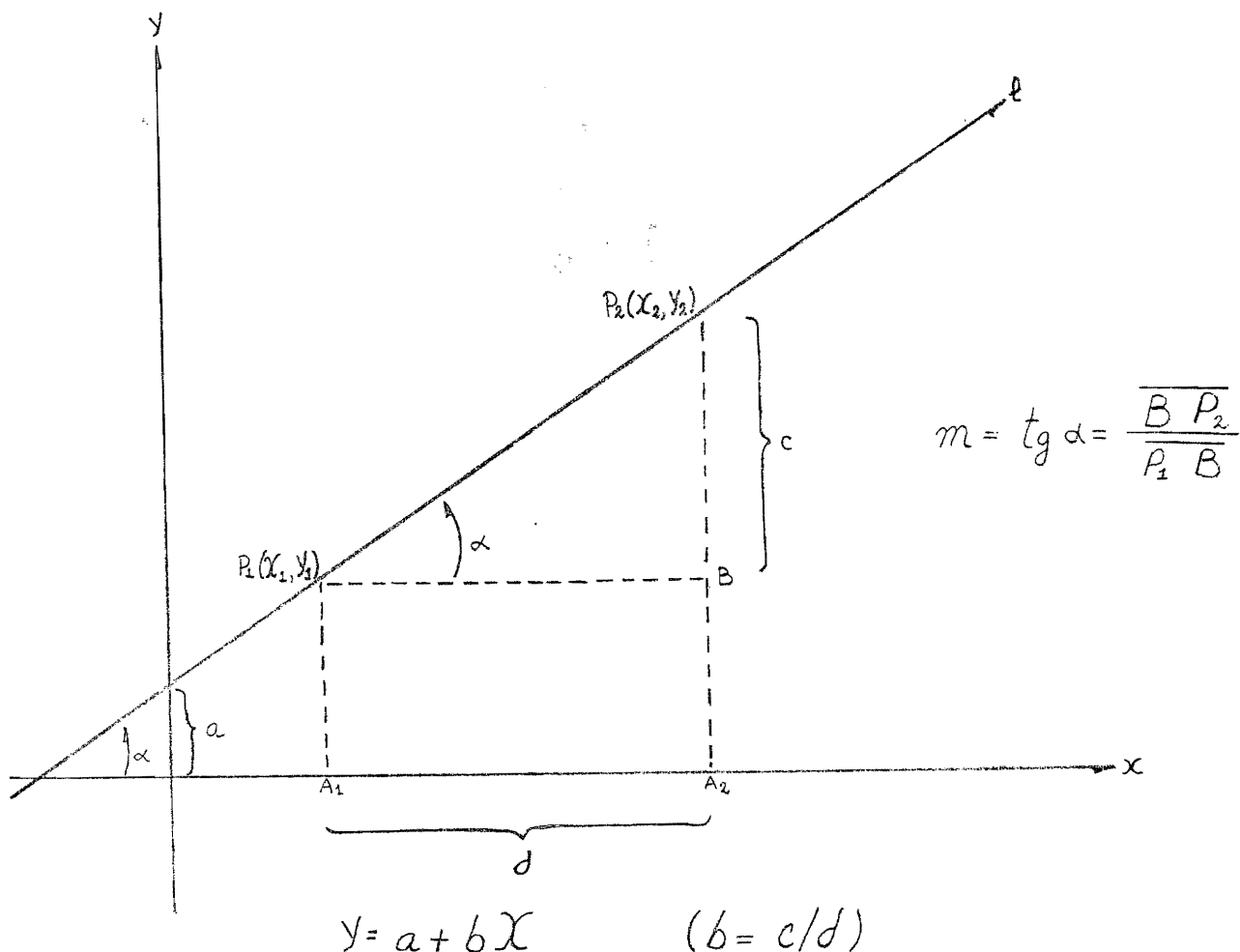
La ecuación de un lugar geométrico cualquiera, puede obtenerse a partir de un número suficiente de las propiedades únicas que lo definen. Las líneas de regresión para la correlación serán aquí siempre rectas, por lo que si tomamos dos puntos diferentes cualesquiera $P_1 (X_1, Y_1)$ y $P_2 (X_2, Y_2)$ de la línea, el valor de la pendiente (m) resulta siempre constante y puede ser calculado por medio de:

$$(3) \quad m = \frac{Y_2 - Y_1}{X_2 - X_1}, \quad \text{donde} \quad X_1 \neq X_2$$

Geoméricamente, una recta queda perfectamente determinada por uno de sus puntos y su pendiente, o sea por las coordenadas de uno de sus puntos y su ángulo de inclinación. Así, la recta que pasara por el punto dado $P_1 (X_1, Y_1)$ y tuviera la pendiente dada m, tendría por ecuación:

$$(4) \quad Y_2 - Y_1 = m (X_2 - X_1)$$

$m = \text{tg } \alpha$ (tangente del ángulo formado por dos rectas dirigidas, donde si α es agudo la pendiente es positiva y si es obtuso la pendiente es negativa.



Las coordenadas del punto A_1 son $(x_1, 0)$ las de $A_2: (x_2, 0)$ y las de $B: (x_2, y_1)$ de donde se tiene que:

$$\begin{aligned} \overline{BP_2} &= y_2 - y_1 \\ \overline{P_1B} &= x_2 - x_1 \end{aligned}$$

con lo que se demuestra que $m = \operatorname{tg} \alpha$

Consideremos una recta l , cuya pendiente es m y cuya ordenada en el origen, es decir, su intercepción con el eje Y , es a . Como se conoce a , el punto cuyas coordenadas son $(0, a)$ está sobre la recta. Por lo tanto, el problema se reduce a encontrar la ecuación de la recta, que pasa por el punto $(0, a)$ y tiene una pendiente dada m .

De acuerdo con estos datos se tiene que la ecuación buscada es:

$$(5) \quad \begin{aligned} y - a &= m(x - 0) \\ y &= mx + a \end{aligned}$$

De todas las curvas de aproximación a una serie de datos, la curva que tiene la propiedad de que la suma de los cuadrados de las desviaciones de la misma sea mínimo se conoce como la mejor curva de ajuste. Una curva que presente

esta propiedad, se dice que se ajusta a los datos por mínimos cuadrados y se llama curva de mínimos cuadrados.

La recta de aproximación por mínimos cuadrados del conjunto de puntos $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$, tiene por ecuación:

$$(6) \quad Y = a + bX$$

donde la constante "a" señala la intercepción de Y (variable dependiente) y b indica la pendiente de la línea de regresión y es una forma equivalente de (5). Estas constantes quedan determinadas mediante el sistema de ecuaciones normales para la recta de mínimos cuadrados (6), como sigue:

$$(7) \quad \begin{aligned} \sum Y &= aN + b\sum X \\ \sum XY &= a\sum X + b\sum X^2 \end{aligned}$$

El cálculo de tales constantes puede simplificarse por las fórmulas siguientes, las que son una derivación de (7):

$$a = \frac{(\sum Y)(\sum X^2) - (\sum X)(\sum XY)}{N\sum X^2 - (\sum X)^2}$$

$$b = \frac{N\sum XY - (\sum X)(\sum Y)}{N\sum X^2 - (\sum X)^2}$$

Las líneas de regresión, tienen por objeto estimar el valor de una variable a partir de otra. En general, la línea de regresión de Y sobre X no es la misma que la de X sobre Y, excepto cuando la correlación es 1 o -1, en donde ambas rectas se superponen. Si la correlación es cero, ambas líneas forman ángulo recto, coincidiendo con sus respectivas medias aritméticas, lo que indicaría que la mayor predicción posible de una variable a partir de otra, sería el promedio de la distribución predicha.

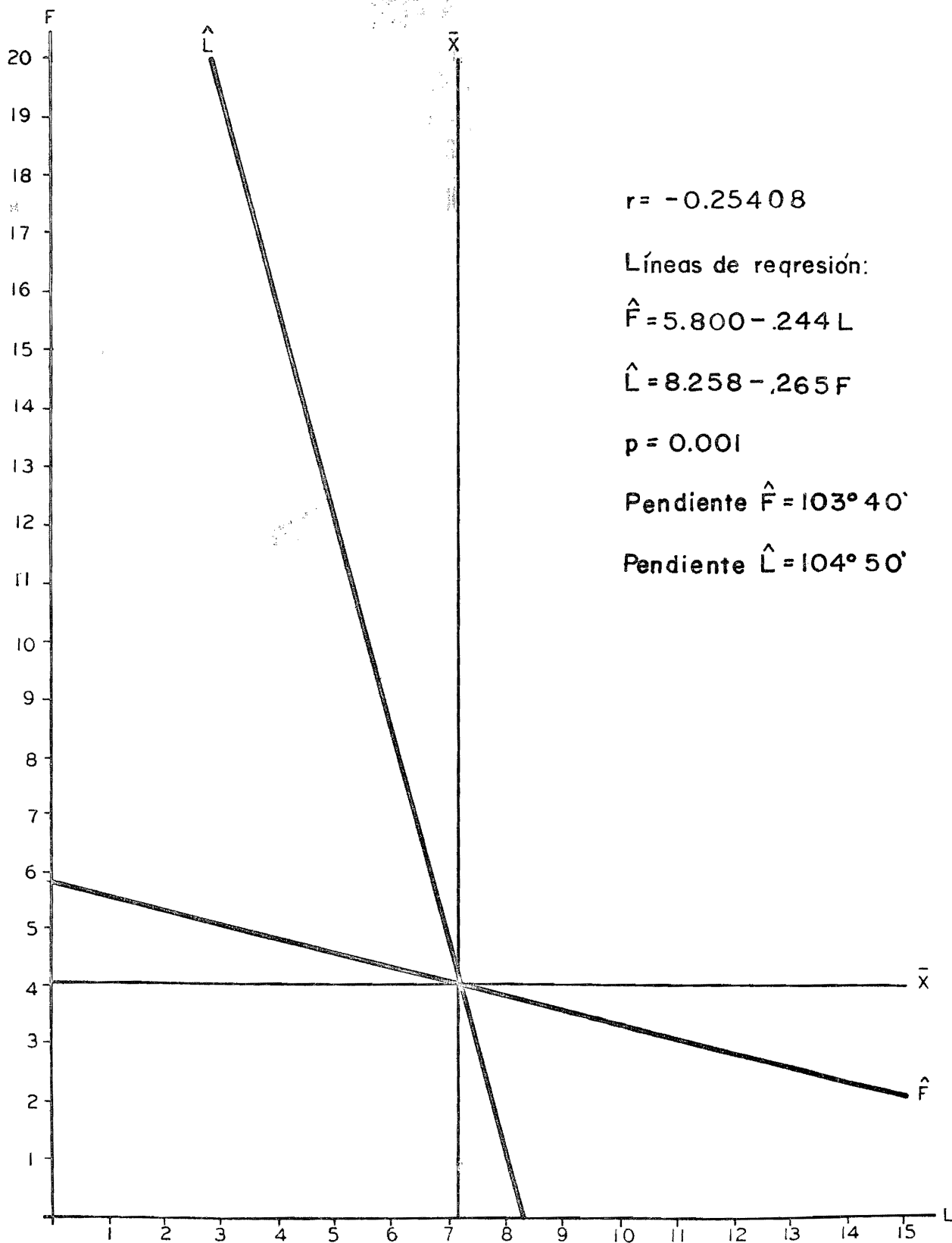
Las gráficas que aparecen en las páginas siguientes, muestran las líneas de regresión de X a partir de Y (\hat{X}) y de Y a partir de X (\hat{Y}) para cada uno de los 15 pares de variables correlacionadas; los coeficientes de correlación son calculados por medio del producto de los momentos* de las variables; para las ecuaciones de las líneas de regresión, la primera constante indica el punto donde la línea intercepta el eje de la variable dependiente y la segunda la pendiente de la línea de regresión (tangente del ángulo formado por la línea de regresión con su propio eje**) y el nivel de probabilidad de que la relación entre ambas variables sea atribuible a fluctuaciones del azar (H_0), por medio de la relación:

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \quad \text{donde} \quad gl = n-2$$

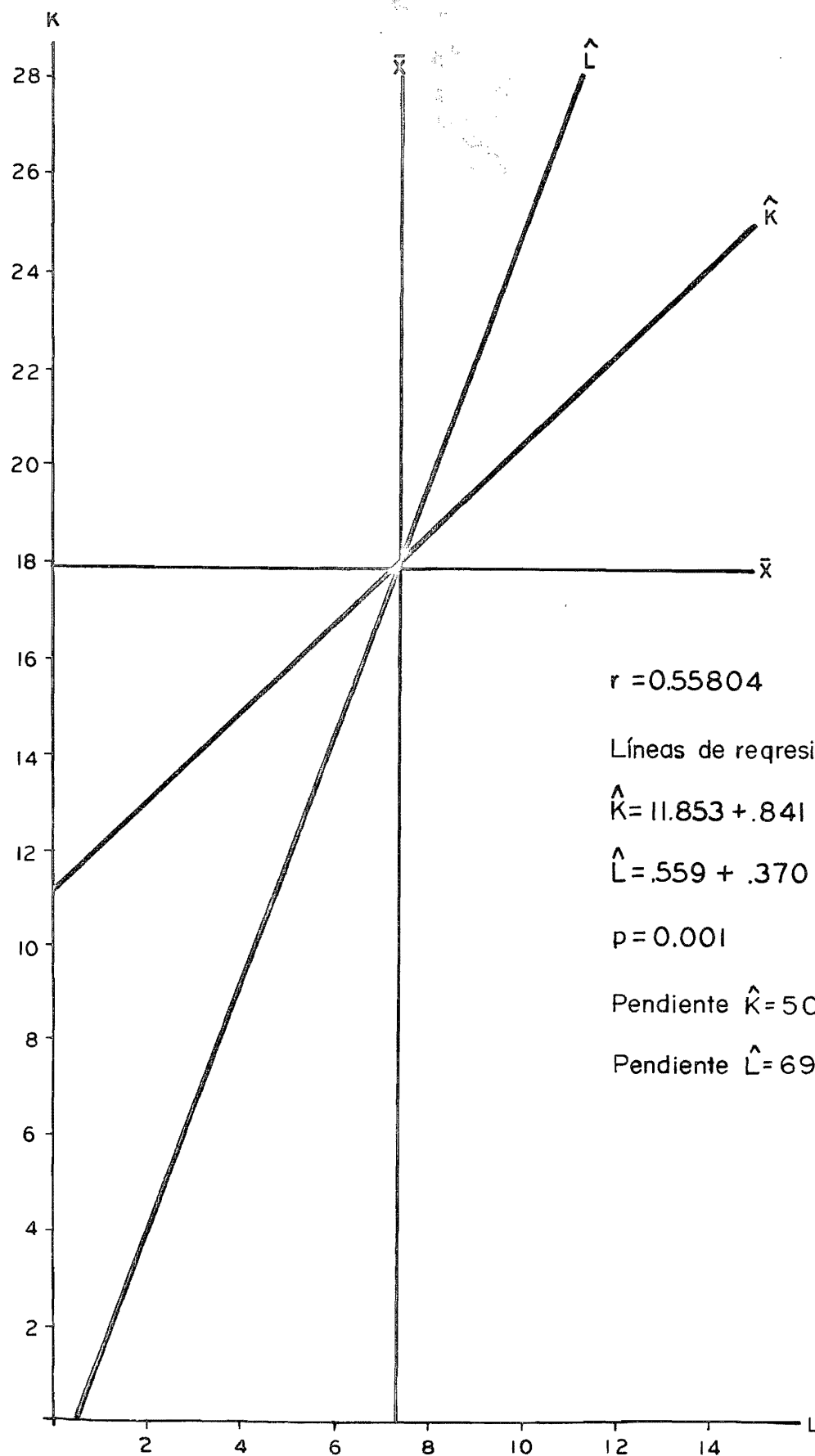
* Segundo momento a partir de la media aritmética.

** Se indica también, en grados, la pendiente de la línea de regresión.

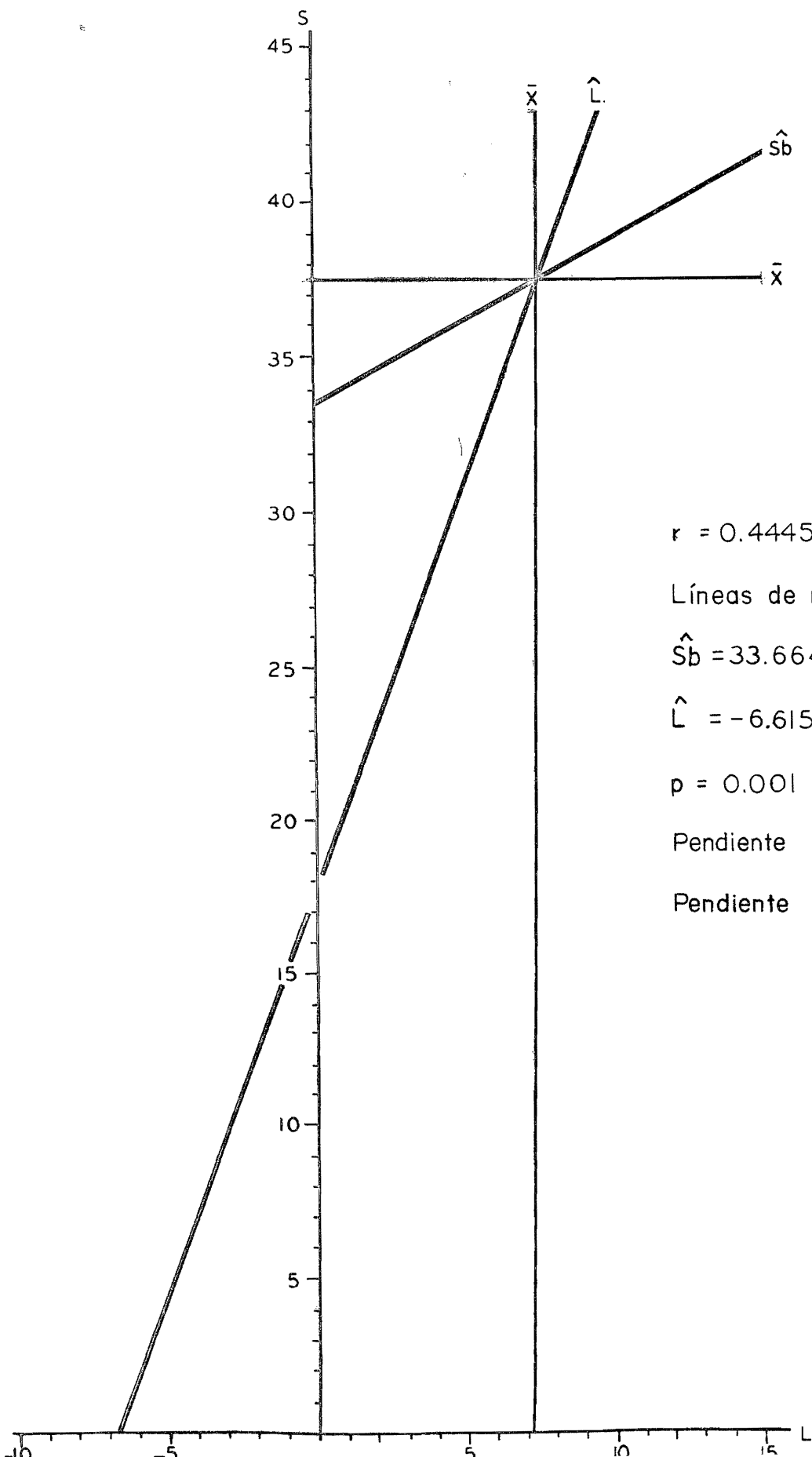
ESCALAS L y F



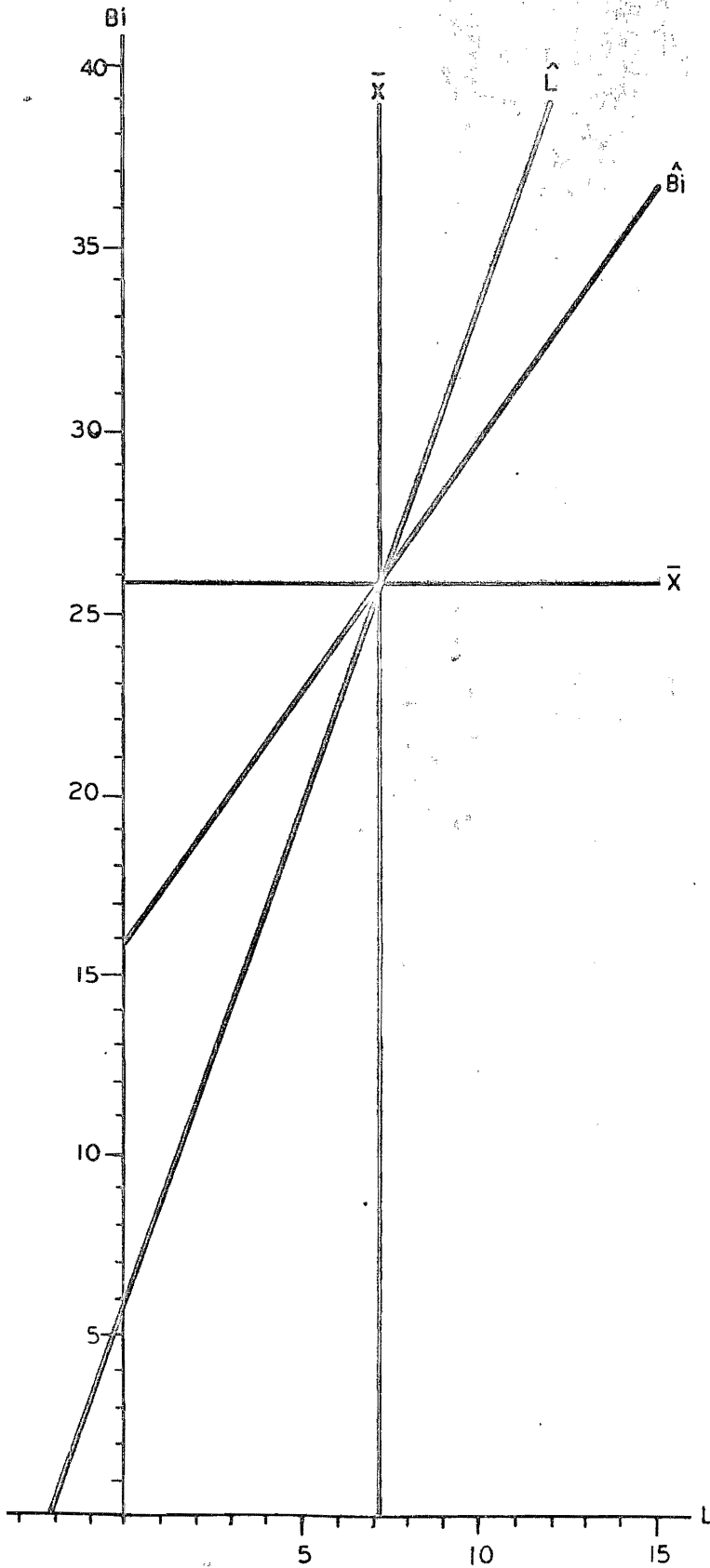
ESCALAS L y K



ESCALAS L y Sb



ESCALAS L y Bi



$$r = 0.51759$$

Líneas de regresión

$$\hat{B}_i = 16.830 + 1.259 L$$

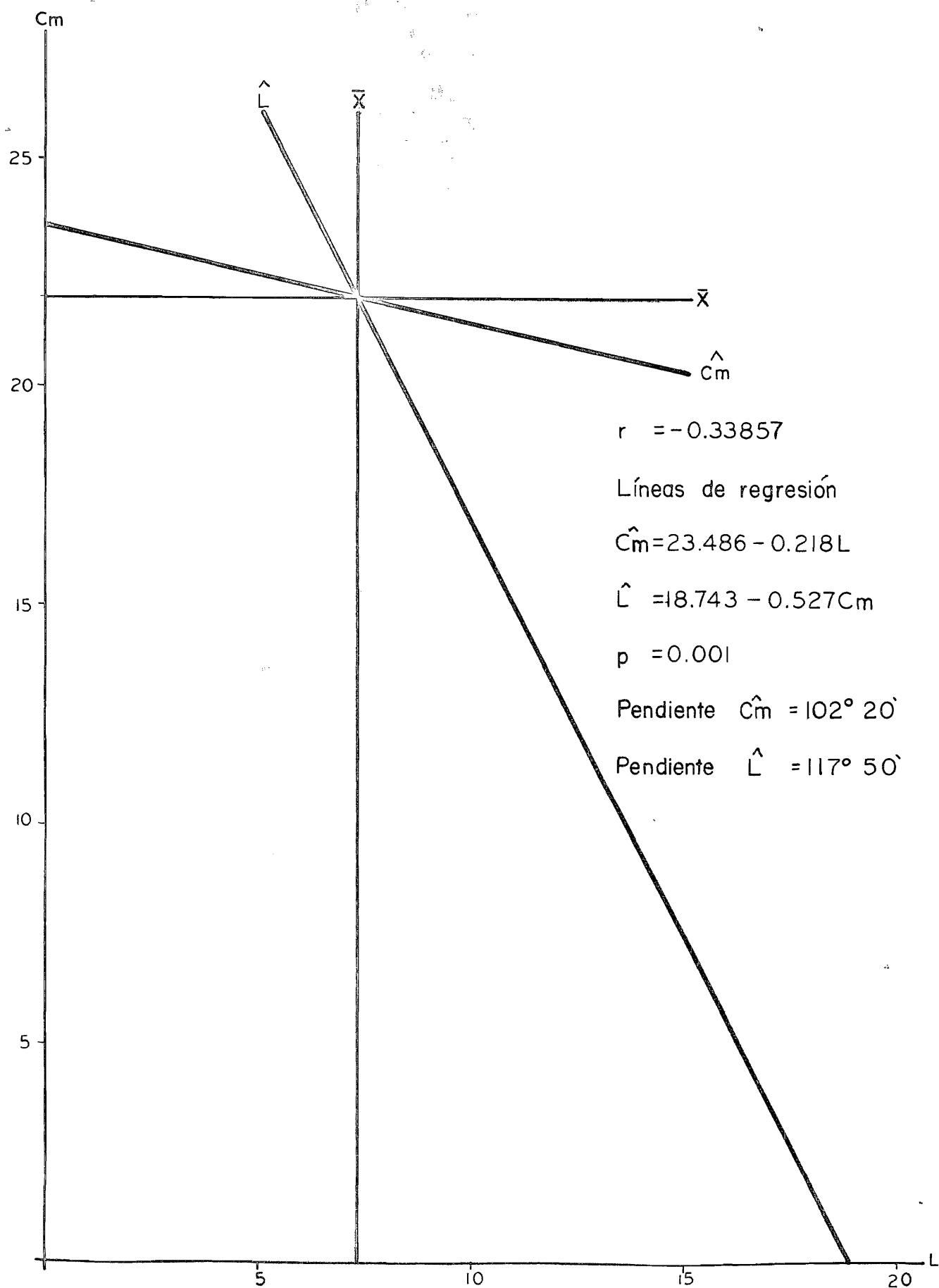
$$\hat{L} = -2.209 + .363 B_i$$

$$p = 0.001$$

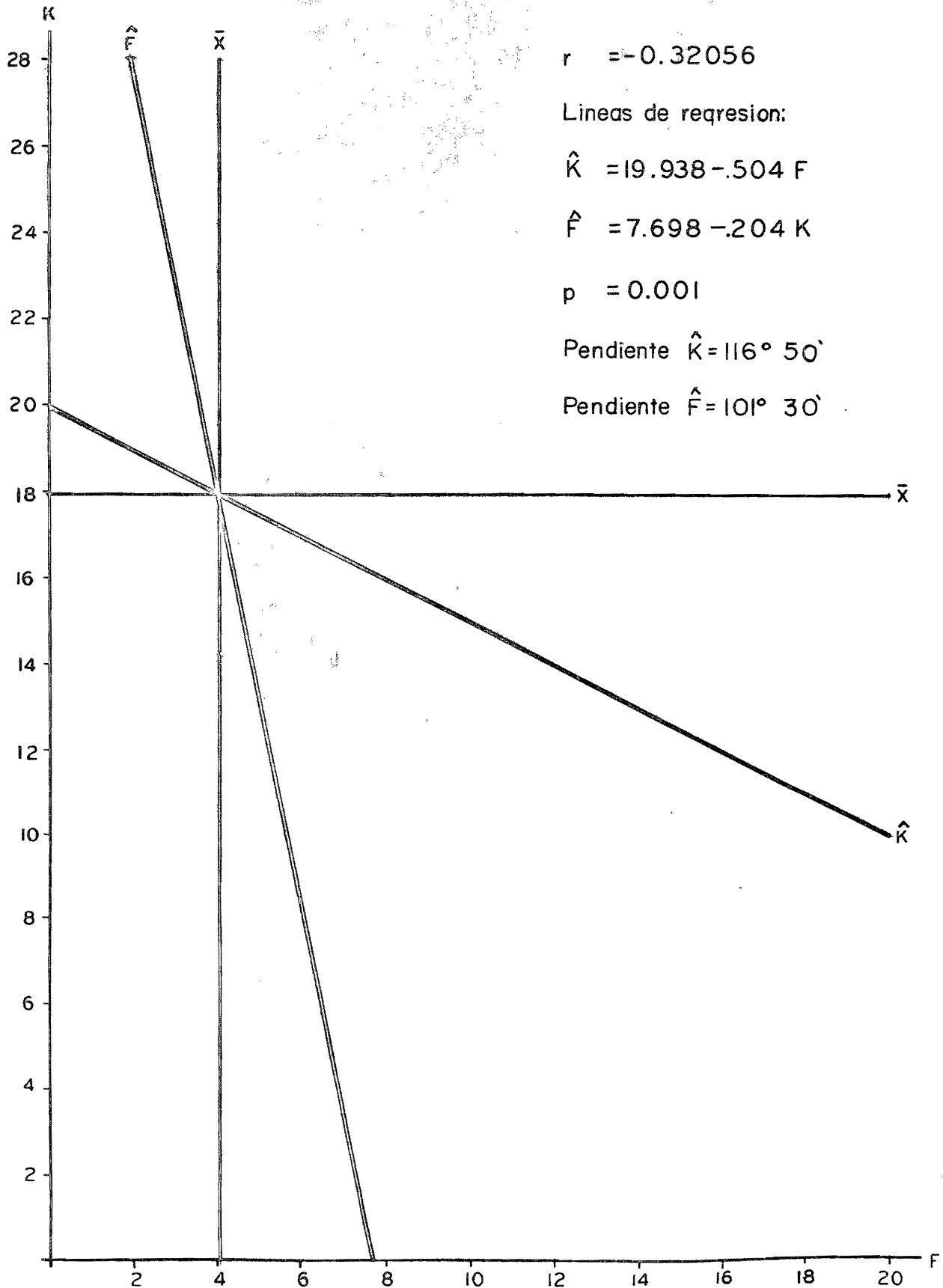
$$\text{Pendiente } \hat{B}_i = 70^{\circ} 00'$$

$$\text{Pendiente } \hat{L} = 38^{\circ} 30'$$

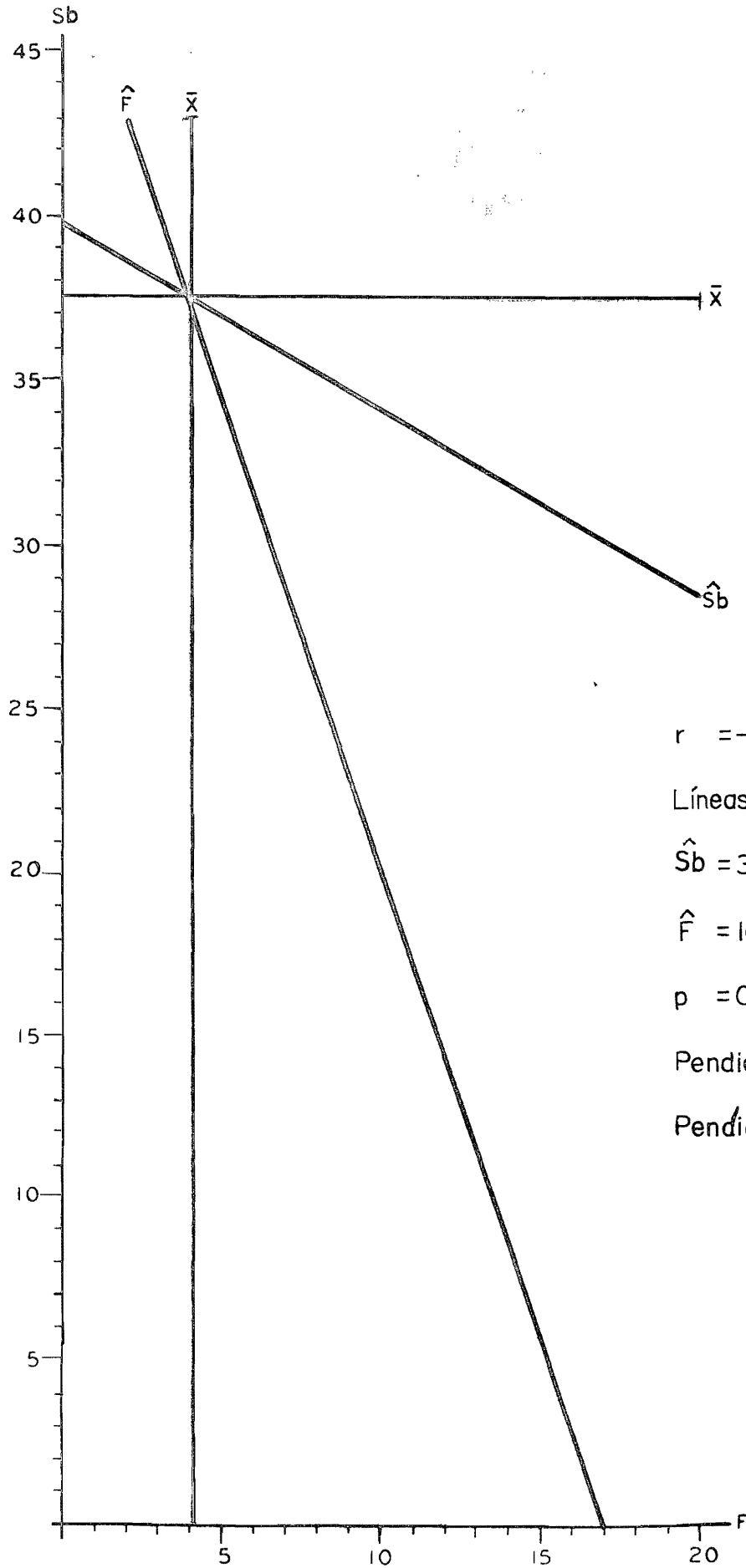
ESCALAS L y Cm



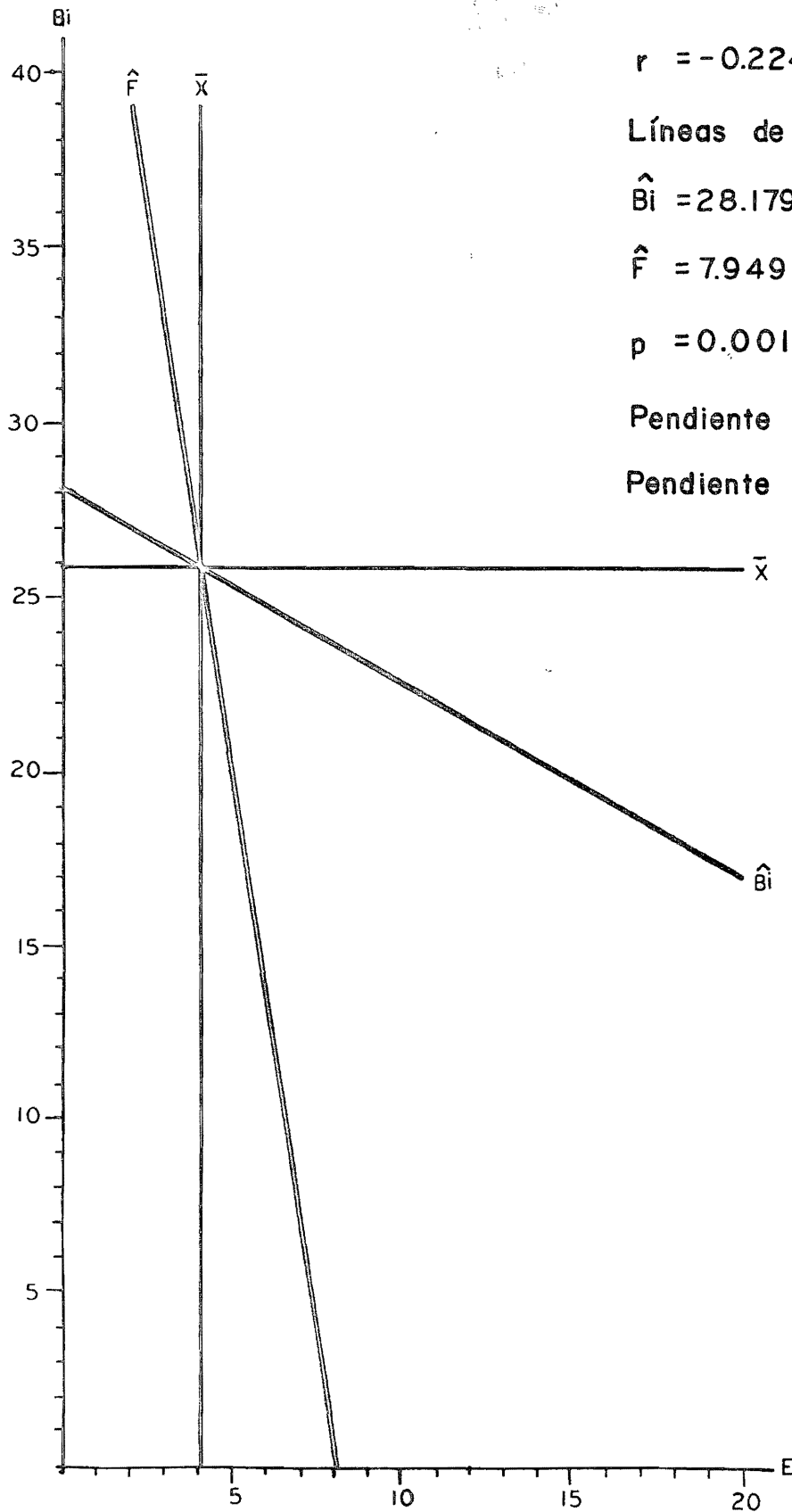
ESCALAS F y K



ESCALAS F y Sb



ESCALAS F y Bi



ESCALAS F y Cm

$$r = -0.19779$$

Lineas de regresion:

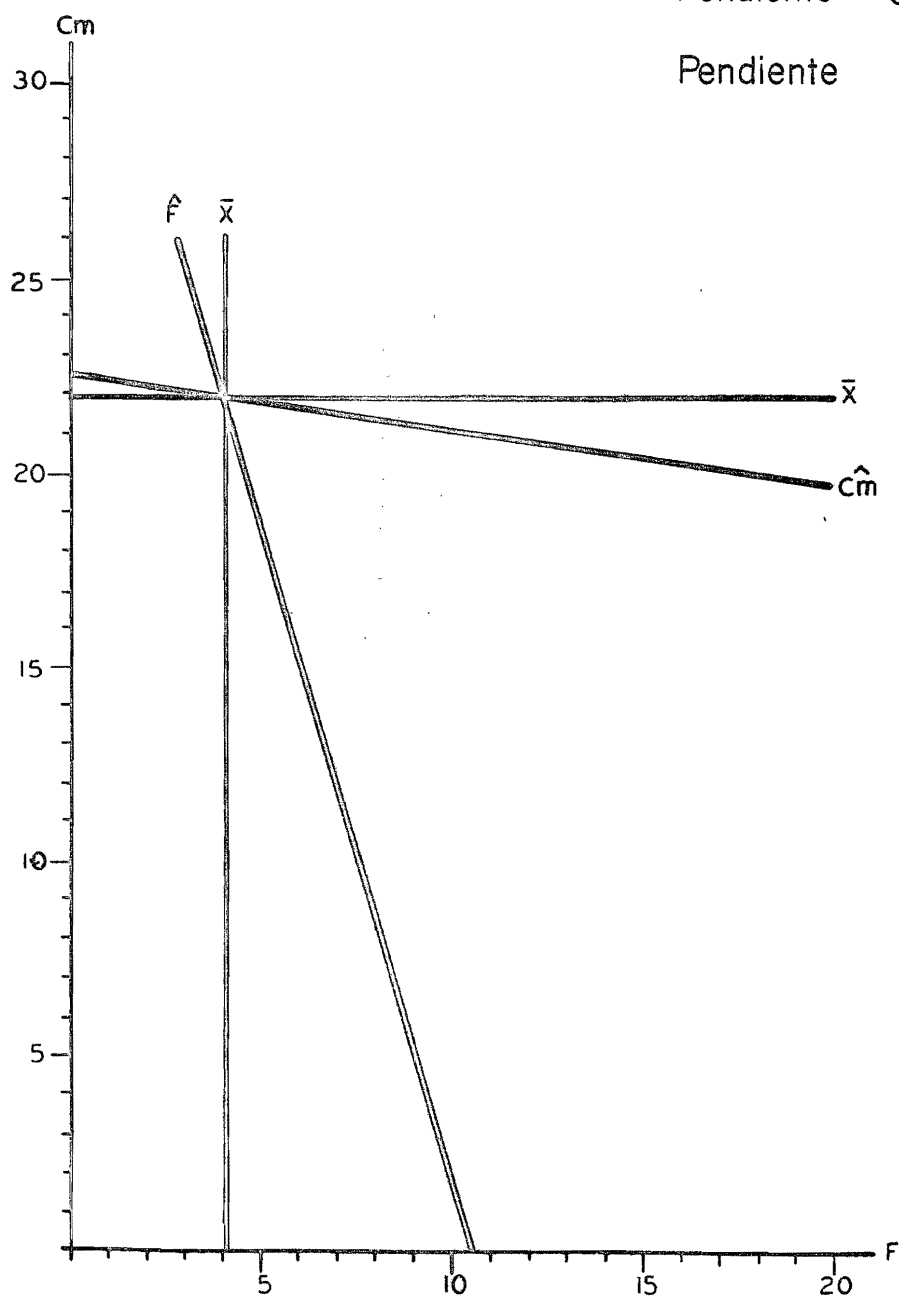
$$\hat{C}_m = 22.460 - .133 F$$

$$\hat{F} = 10.519 - .295 C_m$$

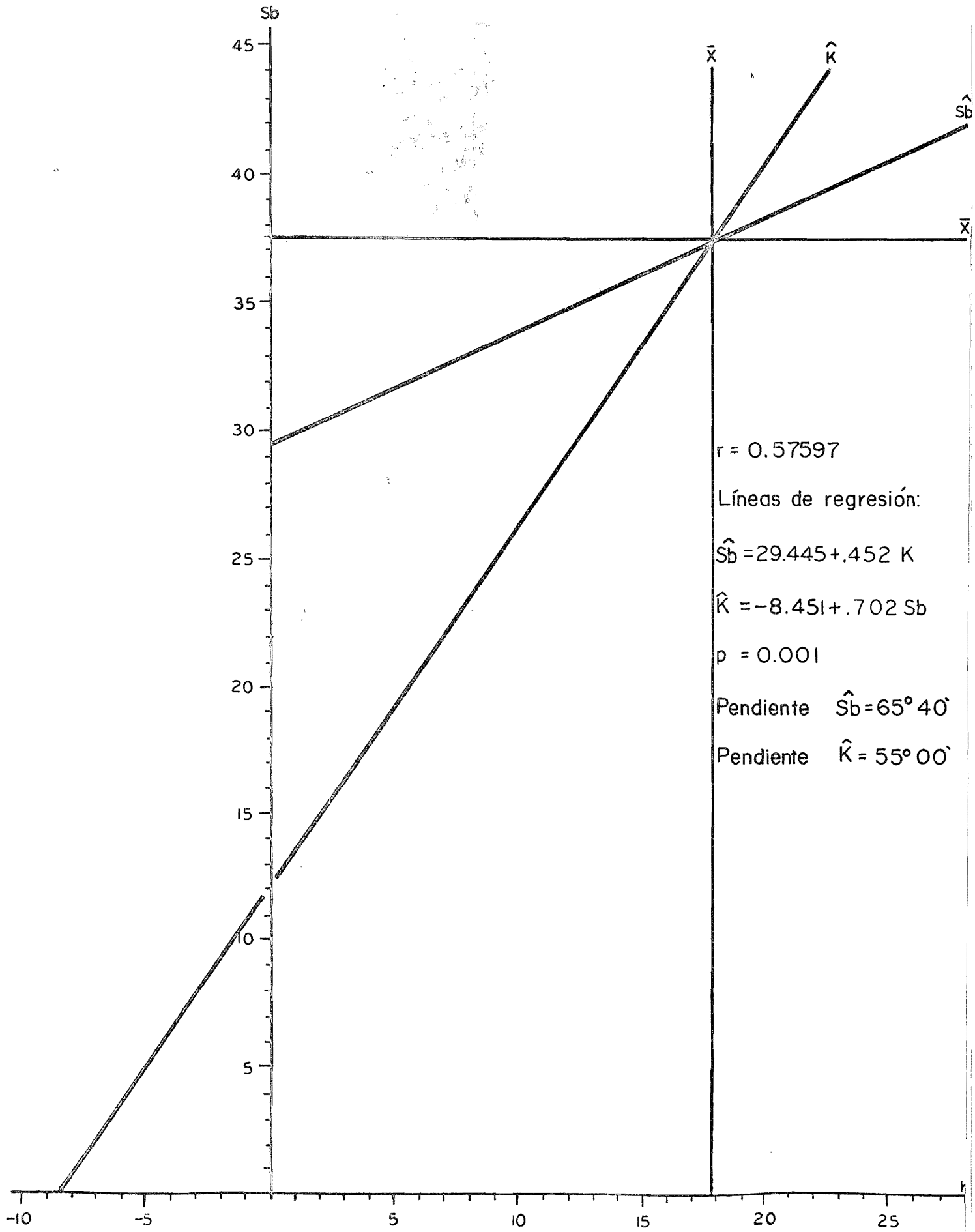
$$p = 0.001$$

$$\text{Pendiente } \hat{C}_m = 97^\circ 40$$

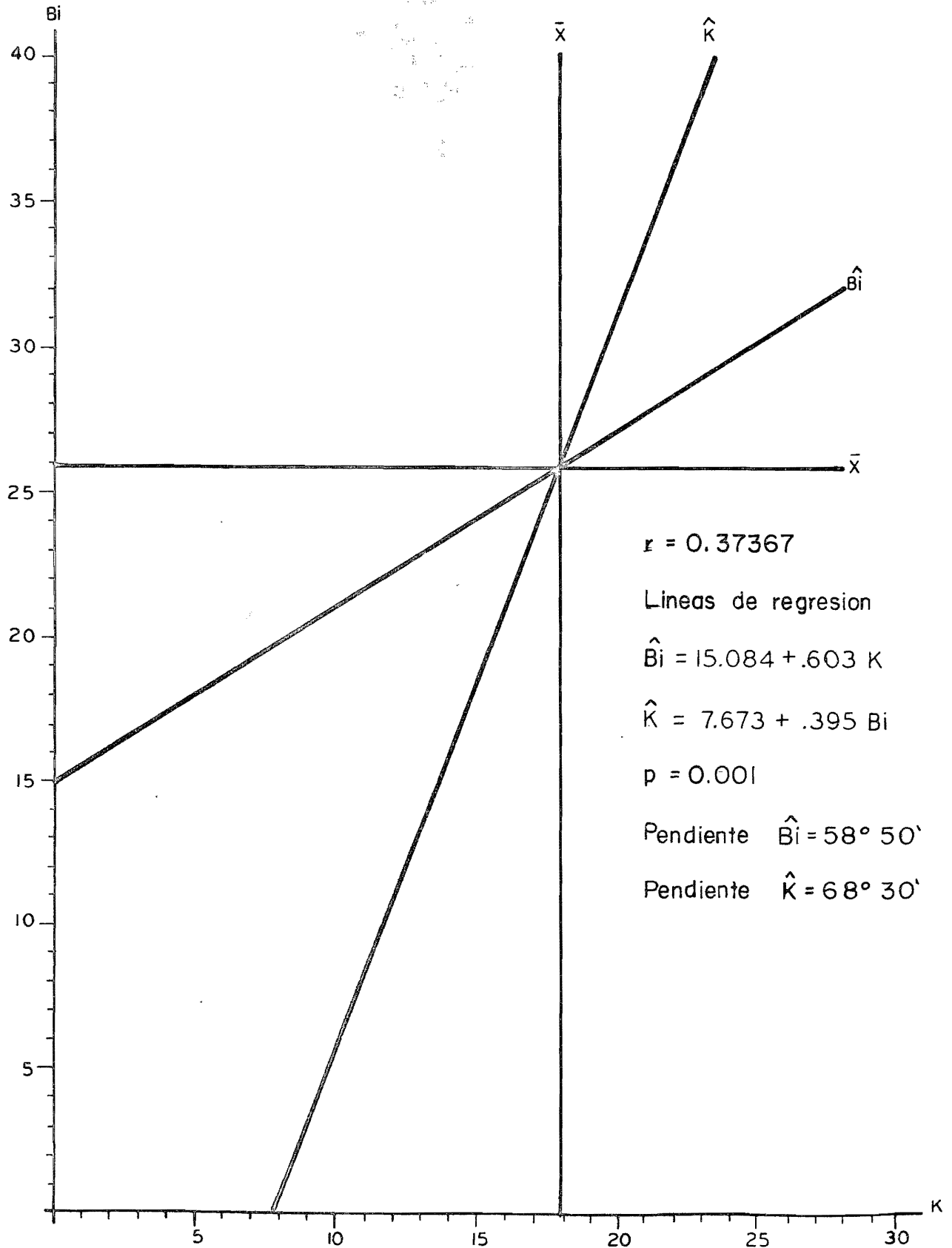
$$\text{Pendiente } \hat{F} = 106^\circ 30$$



ESCALAS K y Sb



ESCALAS K y Bi



ESCALAS K y Cm

$r = -0.24947$

Líneas de regresión

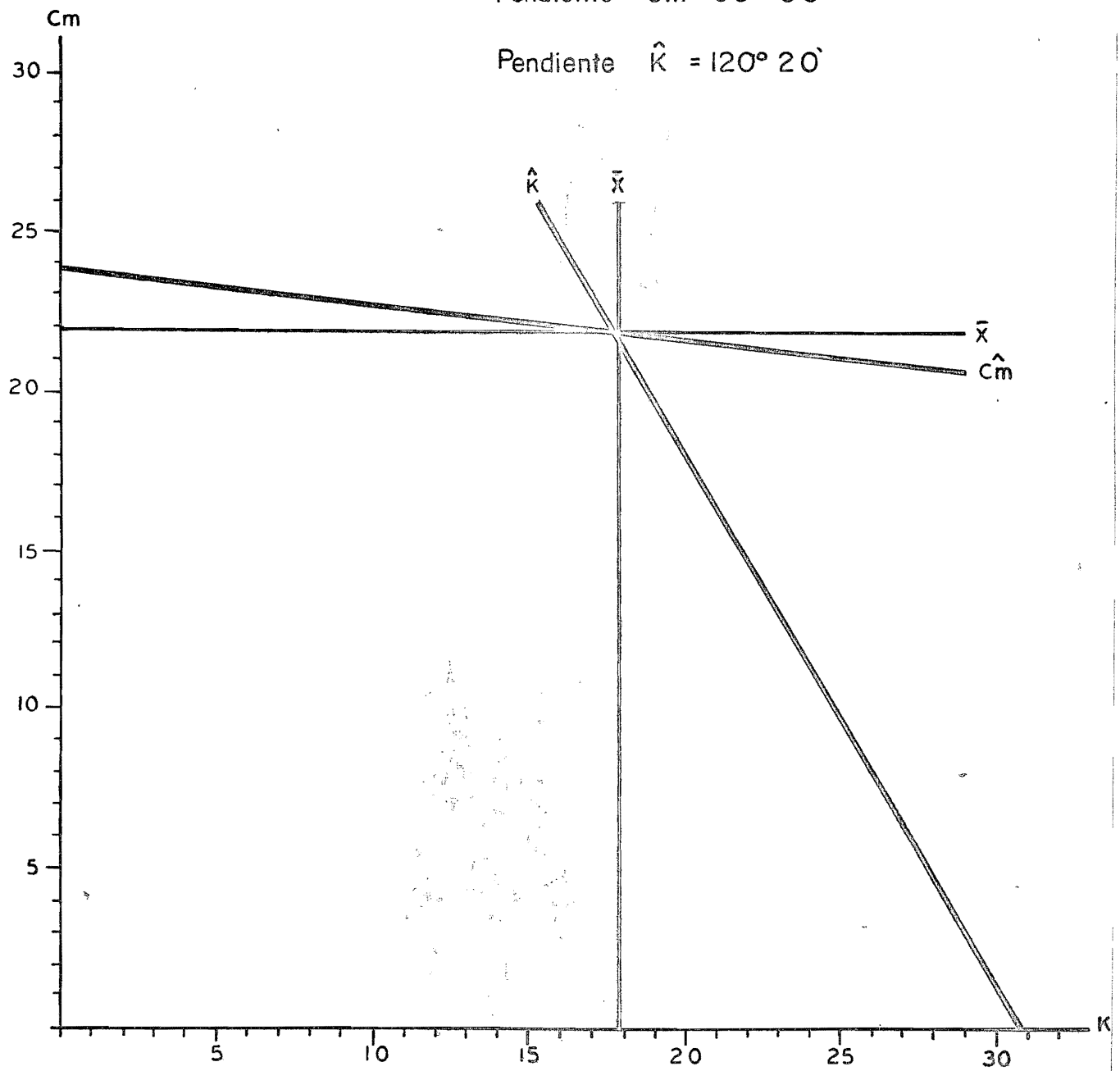
$\hat{C}_m = 23.826 - .106 K$

$\hat{K} = 30.729 - .585 C_m$

$p = 0.001$

Pendiente $\hat{C}_m = 96^\circ 00'$

Pendiente $\hat{K} = 120^\circ 20'$



ESCALAS Sb y Bi

$$r = 0.45250$$

Líneas de regresión

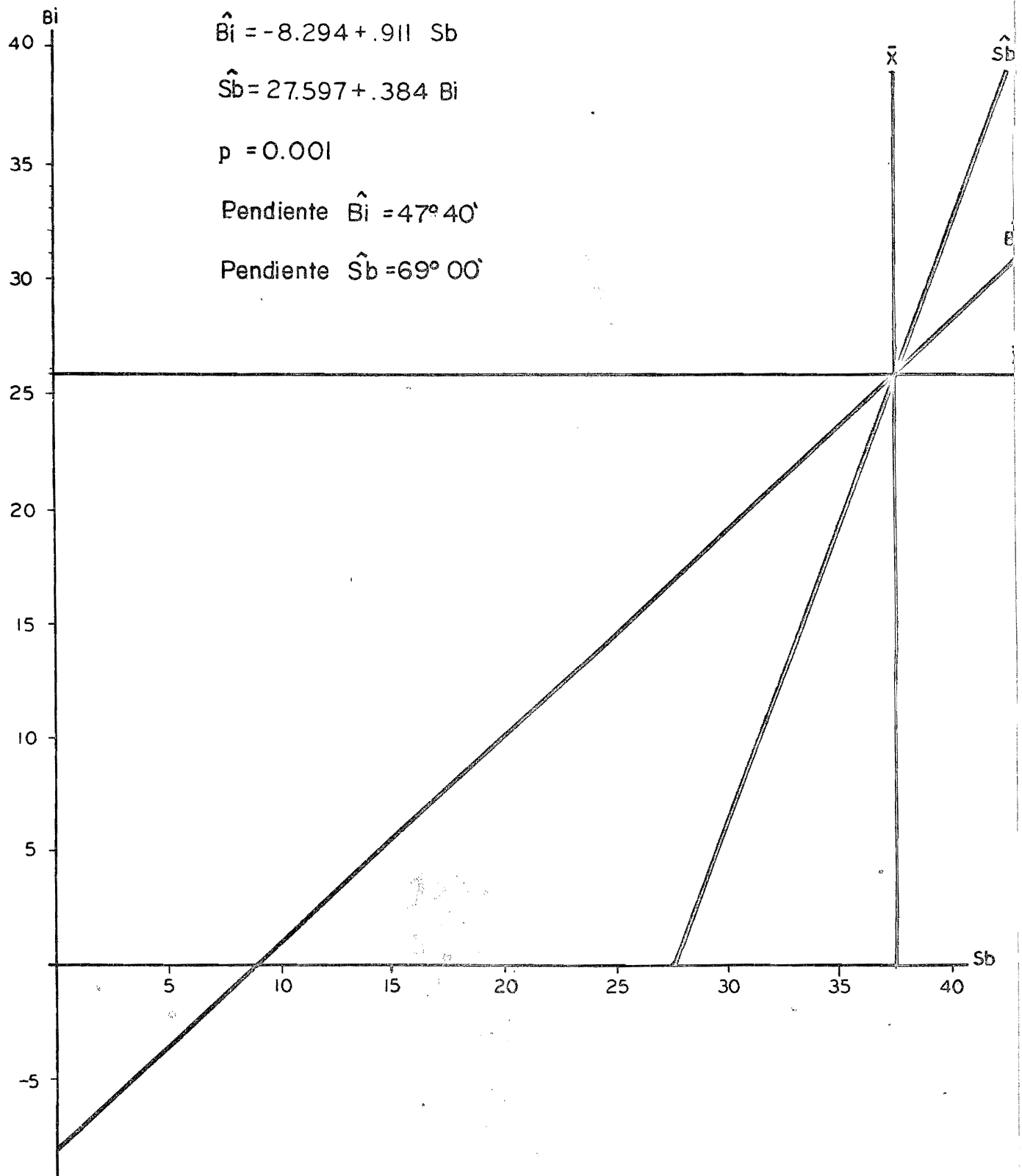
$$\hat{B}_i = -8.294 + .911 S_b$$

$$\hat{S}_b = 27.597 + .384 B_i$$

$$p = 0.001$$

$$\text{Pendiente } \hat{B}_i = 47^\circ 40'$$

$$\text{Pendiente } \hat{S}_b = 69^\circ 00'$$



ESCALAS Sb y Cm

$r = -0.42695$

Líneas de regresión:

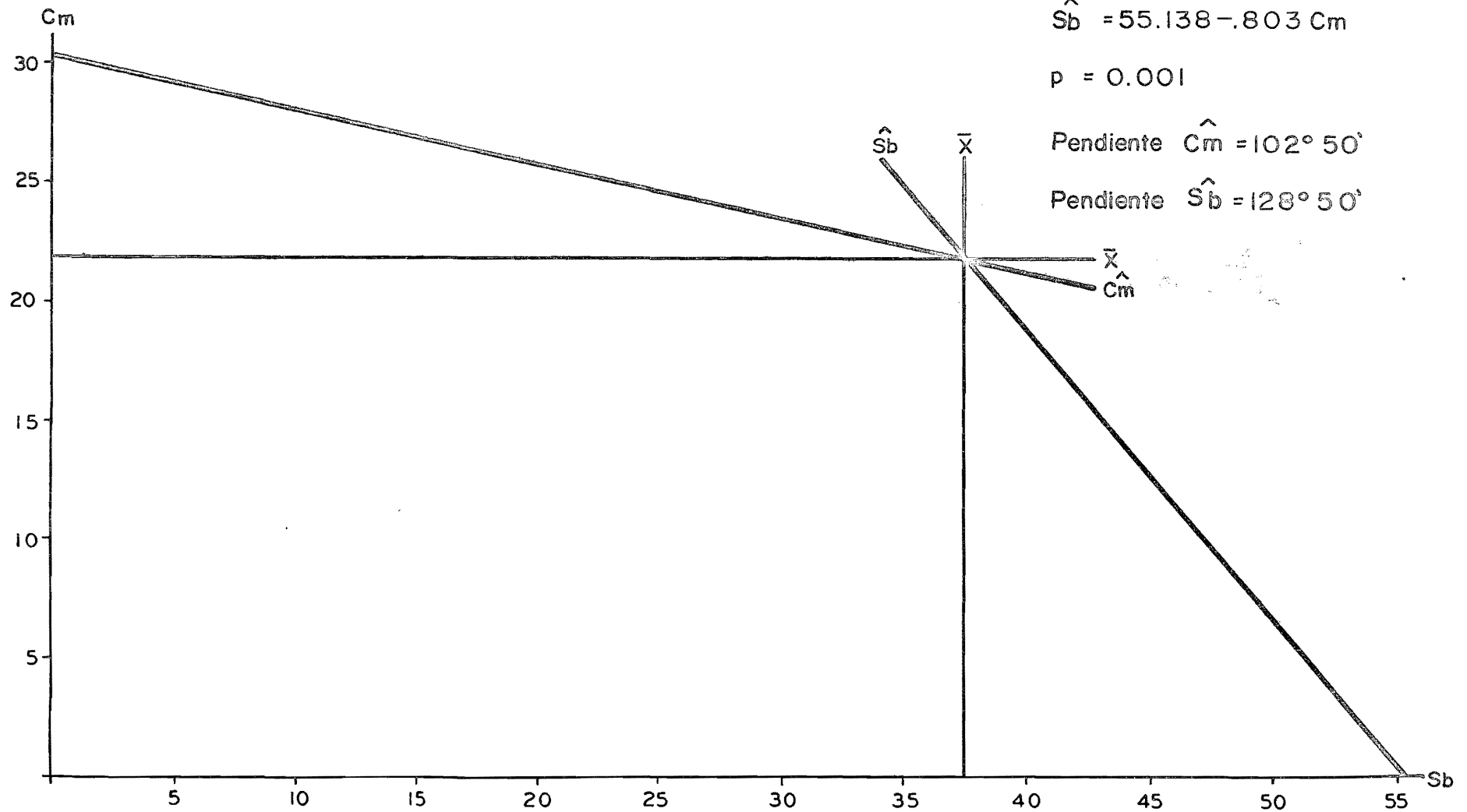
$\hat{C}_m = 30.437 - .227 S_b$

$\hat{S}_b = 55.138 - .803 C_m$

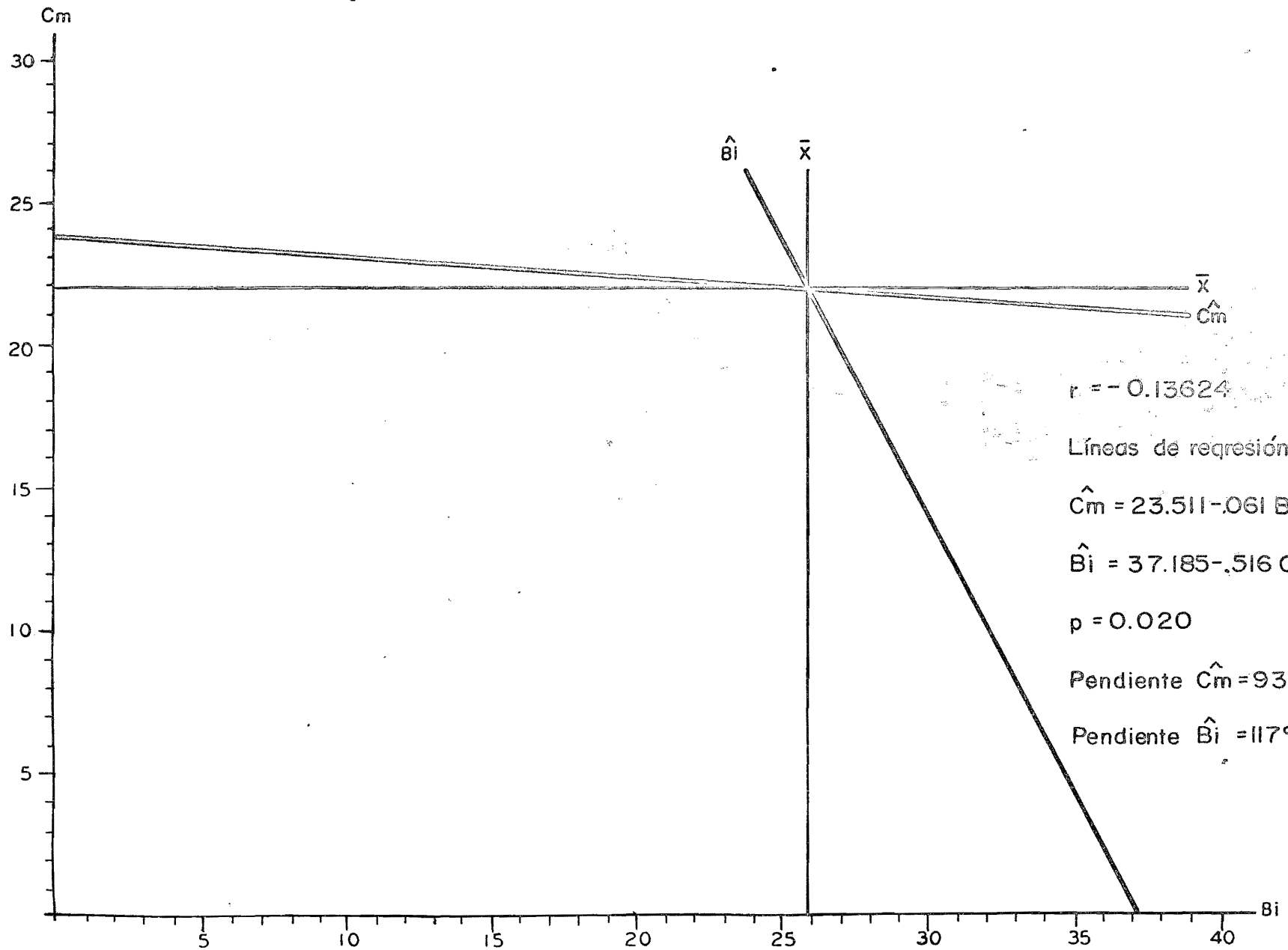
$p = 0.001$

Pendiente $\hat{C}_m = 102^\circ 50'$

Pendiente $\hat{S}_b = 128^\circ 50'$



ESCALAS Bi y Cm



5) ANALISIS DE VARIANZA. (2, 3, 22, 24 y 30)

Se han considerado ya las diferencias entre las medias de las escalas de validez incluídas en este estudio; como puede apreciarse, las medias (apareadas) LBi, LK y KBi, no difieren significativamente entre sí y muestran tener elevaciones similares entre sí.

Un procedimiento alternativo para poner a prueba la hipótesis nula en relación al funcionamiento de la variabilidad entre las escalas, está dado por el análisis de varianza, método desarrollado por Fisher hace aproximadamente medio siglo; este método cumple con dos funciones:

1.- Permite computar la varianza conjunta de las escalas, viniendo a ser un procedimiento relativamente simple para determinar la varianza de error, y

2.- Proporciona otra prueba, la razón F, para la hipótesis nula de que las medias de L, F, K, Sb, Bi y Cm son idénticas entre sí (Snedecor y Cochran, 1969), permitiendo la inspección preliminar de los resultados, antes de pasar al análisis factorial.

Supongamos, momentáneamente, que no existen diferencias entre los promedios obtenidos por los sujetos de la muestra en las seis escalas consideradas. En esta situación todos los 2022 puntajes observados estarían distribuídos al rededor de una media aritmética común, con una varianza (σ^2) determinada.

El análisis de varianza fue desarrollado por Fisher, partiendo del hecho de que pueden efectuarse tres estimaciones de la varianza, a partir de los datos de la siguiente tabla:*

	L	F	K	Sb	Bi	Cm
$\sum X$	20 321	17 764	20 346	16 821	20 153	11 811
\bar{X}	60.300	52.712	60.374	49.914	59.801	35.047
$\sum X^2$	1263 059	955 234	1254 990	870 673	1237 975	443 537
$(\sum X)^2/n$	1225 350	936 379	1228 367	839 602	1205 173	413 946
$\sum x^2$	37 709	18 855	26 623	31 071	32 802	29 591
gl	336	336	336	336	336	336

T O T A L		
$\sum X$	107 216 (G)	Varianza conjunta = 176 651/2016 = 87.625
\bar{X}	318.148	Error estándar de la diferencia entre me--
$\sum X^2$	6 025 468	dias (σ_0) = $\sqrt{2\sigma^2/n} = \sqrt{(2)(87.625)/337} =$
$(\sum X)^2/N$	5 848 817	.721
$\sum x^2$	176 651	
gl	2016	

* Puntajes T.

Puesto que asumimos que los 2022 puntajes provienen de la misma distribución, podemos calcular la suma de cuadrados total de las desviaciones para los 2022 calificaciones, como sigue:

$$(1) \quad SS_{TOT} = (\sum X_1^2 + \sum X_2^2 + \sum X_3^2 + \sum X_4^2 + \sum X_5^2 + \sum X_6^2) - (\sum X_1 + \sum X_2 + \sum X_3 + \sum X_4 + \sum X_5 + \sum X_6)^2 / N = 340\,368.764$$

El número de grados de libertad para esta suma de cuadrados es:

$$gl = N_{TOT} - 1 = 2021$$

El cuadrado medio es entonces, $SS_{TOT} / gl = 340\,368.764 / 2021 = 168.416$ que es la primera estimación de la varianza.

La segunda aproximación es la varianza conjunta ($\sum X^2$) que se obtuvo en la tabla, y a la que se denomina suma de cuadrados intraescalas; su fórmula de cálculo es:

$$SS \text{ Intraescalas} = SS \text{ Total} - SS \text{ Interescalas} = 176\,650.125 \quad (\sum \sum x^2)$$

La suma de cuadrados es dividida por sus grados de libertad ($gl = N_{TOT} - nv^*$), 2016, por lo que la media de cuadrados es entonces 87.625.

La tercera estimación, considera las medias de las seis escalas cuyas varianzas son $\sigma^2/337$, puesto que son medias de muestras de 337 puntajes. La suma de cuadrados de las desviaciones es:

$$X_1^2 + X_2^2 + X_3^2 + X_4^2 + X_5^2 + X_6^2 - (\sum X)^2 / 6 = 485.832$$

con $(ng-1)$ grados de libertad. El cuadrado medio, $485.832/5=97.166$, es una estimación de $\sigma^2/337$. En consecuencia, si se multiplica por 337, tendremos la tercera estimación de la varianza, esto es, $163\,725.384$, en donde el cuadrado-medio resulta $163\,725.384/5 = 32\,745.077$.

Puesto que la suma de cuadrados total es 337 veces la media de cada escala, esta suma de cuadrados puede ser calculada por los totales de cada escala como sigue:

$$(3) \quad SS \text{ interescalas} = (\sum X_1)^2 / N_1 + (\sum X_2)^2 / N_2 + (\sum X_3)^2 / N_3 + (\sum X_4)^2 / N_4 + (\sum X_5)^2 / N_5 + (\sum X_6)^2 / N_6 - (\sum X_1 + \sum X_2 + \sum X_3 + \sum X_4 + \sum X_5 + \sum X_6)^2 / N_{TOT} = 5848\,817.875 - 5685\,099.236 = 163\,718.639$$

A partir de los datos anteriores, se formula el cuadro siguiente, consiguiendo las sumas de cuadrados y los grados de libertad.

* nv = número de variables.

Fuente de Variación	Grados de libertad	Suma de Cuadrados
Interescalas	5	163 718.639
Intraescalas	2016	176 650.25
T o t a l	2021	340 368.764

Debe señalarse que la suma de cuadrados y los grados de libertad para ambos componentes de la varianza (Interescalas e intraescalas) se suman para obtener los totales. El resultado para los grados de libertad, no resulta difícil de verificar; con nv variables y n puntajes por variable, los grados de libertad son $(nv-1)$ para la varianza interescalas, $nv(n-1)$ para la varianza intraescalas, y $(nv.n)-1$ para la varianza total; en donde:

$$(nv-1) + nv(n-1) = nv-1 + nv.n - nv = (nv.n)-1$$

La tabla siguiente, muestra los aspectos más frecuentemente incluidos en el análisis de varianza, para seis escalas con 337 observaciones en cada una de ellas. El símbolo T denota el total de una variable, en tanto que $G = \sum T = \sum \sum X$ es el gran total.

El primer paso para la formulación de la tabla, es el cálculo de la corrección para la media de cuadrados:

$$C = G^2 / nv \cdot n = (107\ 216)^2 / (6)(337) = 5685\ 099.236$$

Este cálculo se efectúa en razón de que C se presenta tanto en el cálculo de (1) como de (3); los pasos restantes se especifican en la tabla:

FORMULAS PARA EL CALCULO DE LA TABLA DE ANALISIS DE VARIANZA.

Fuente de Variación	Grados de Libertad	Suma de Cuadrados	Media de Cuadrados
Interescalas	$nv-1 = 5$	$(\sum T^2/n) - C = 163\ 718.639$	32 743.728
Intraescalas	$nv(n-1) = 2016$	Resta = 176 650.25	87.625
T o t a l	$(nv.n)-1 = 2021$	$\sum \sum X^2 - C = 340\ 368.764$	

Se calcula después la proporción de la varianza aportada por la interacción de las escalas por medio de:

$$(4) \quad SS_{interacción} = \frac{(\sum X_1 + \sum X_2)^2}{n_1 + n_2} + \frac{(\sum X_3 + \sum X_4 + \sum X_5 + \sum X_6)^2}{n_3 + n_4 + n_5 + n_6} - \frac{(\sum X_1 + \sum X_2 + \sum X_3 + \sum X_4 + \sum X_5 + \sum X_6)^2}{N_{TOT}}$$

El número de grados de libertad es $df = nv-1$; por lo que la media de cuadrados para cada fuente de variación por interacción entre las escalas es igual a la suma de cuadrados de cada condición entre sus grados de libertad, tal como se muestra en el siguiente cuadro.

FUENTE DE VARIACION	SUMA DE CUADRADOS	GRADOS DE LIBERTAD	MEDIA DE CUADRADOS.
Interacción LF	12 252.108	5	2450.422
Interacción LK	54 054.458	5	10 810.892
Interacción LSb	4 382.814	5	876.563
Interacción LBi	49 903.669	5	9 980.734
Interacción LCM	28 949.654	5	5 789.931
Interacción FK	12 514.589	5	2 502.918
Interacción FSb	2 962.048	5	592.410
Interacción FBi	10 560.392	5	2 112.078
Interacción FCM	14 549.229	5	2 909.846
Interacción KSb	4 540.363	5	908.073
Interacción KBi	50 431.989	5	10 086.398
Interacción KCM	28 549.710	5	5 709.942
Interacción SbBi	3 396.251	5	679.250
Interacción SbCM	52 399.209	5	10 479.842
Interacción BiCM	31 709.441	5	6 341.888

La proporción de la varianza aportada por cada una de las escalas en relación a la varianza total se computa por medio de:

$$\begin{aligned} \text{SS variación} = & \text{SS interacción 1, 2} + \text{SS interacción 1, 3} + \\ & \text{SS interacción 1, 4} + \text{SS interacción 1, 5} + \\ & \text{SS interacción 1, 6} \end{aligned}$$

Donde el número de grados de libertad es: $df = (nv-1) (nv-1) = 25$

Por lo que para las seis escalas de validez consideradas esta fuente de variación resulta:

FUENTE DE VARIACION	SUMA DE CUADRADOS	GRADOS DE LIBERTAD	MEDIA DE CUADRADOS
Variación L	149 542.703	25	5981.708
Variación F	52 838.366	25	2113.535
Variación K	150 091.109	25	6003.644
Variación Sb	67 680.685	25	2707.227
Variación Bi	146 001.742	25	5840.070
Variación Cm	156 157.243	25	6246.289

Por último se calcula la varianza de error para cada una de las escalas - por medio de:

$$\text{SS error de variación} = \text{SS intraescalas} - \text{SS variación},$$

donde el número de grados de libertad es igual a $(nv-1) = 5.$, de donde se tiene:

FUENTE DE VARIACION	SUMA DE CUADRADOS	GRADOS DE LIBERTAD	MEDIA DE CUADRADOS
Error L	27 107.421	5	5 421.484
Error F	123 811.758	5	24 762.352
Error K	26 559.015	5	5 311.803
Error Sb	108 969.439	5	21 793.888
Error Bi	30 648.382	5	6 129.676
Error Cm	20 492.811	5	4 098.562

El criterio utilizado para poner a prueba la hipótesis nula de que las medias de la población son las mismas para todas las escalas, es proporcionada por la razón de variación F, la que es el cociente de la media de cuadrados de cada condición, entre la media de cuadrados intraescalas. El valor de F debe ser de alrededor de 1 cuando la hipótesis nula es aceptada; en tanto que su valor se hará mayor cuando las medias difieran sustancialmente. La distribución fue tabulada por Fisher en la forma: $Z = \log_e \sqrt{F}$ y en su honor, la razón fue nominada como F por Snedecor (1934). Fisher y Yates (1938) habían denominado a F como la "razón de variación".

El cuadro siguiente muestra el análisis de varianza completo:

ANÁLISIS DE VARIANZA COMPLETO

FUENTE DE VARIACION	SUMA DE CUADRADOS	df	MEDIA DE CUADRADOS	F	P	PROPORCION DE LA VARIANZA APORTADA POR CADA CONDICION
Interacción LF	12 252.108	5	2 450.422	27.97	.005	.036
Interacción LK	54 054.458	5	10 810.892	123.38	.005	.159
Interacción LSb	4 382.814	5	876.563	10.00	.005	.013
Interacción LBi	49 903.669	5	9 980.734	113.90	.005	.147
Interacción LCm	28 949.654	5	5 789.931	66.08	.005	.085
Interacción FK	12 514.589	5	2 502.918	28.56	.005	.037
Interacción FSb	2 962.048	5	592.410	6.76	.005	.009
Interacción FBi	10 560.392	5	2 112.078	24.10	.005	.031
Interacción FCm	14 549.229	5	2 909.846	33.21	.005	.043
Interacción K Sb	4 540.363	5	908.073	10.36	.005	.013
Interacción K Bi	50 431.989	5	10 086.398	115.11	.005	.148
Interacción K Cm	28 549.710	5	5 709.942	65.16	.005	.084
Interacción Sb Bi	3 396.251	5	679.250	7.75	.005	.010
Interacción Sb Cm	52 399.209	5	10 479.842	119.60	.005	.154
Interacción Bi Cm	31 709.441	5	6 341.888	72.38	.005	.093
Intraescalas	176 650.124	2016	87.624			.519
Interescalas	163 718.639	5	32 743.728	373.684	.005	.481
Total	340 368.764	2021	168.416			1.000
Error L	27 107.421	5	5 421.484	61.87	.005	.080
Error F	123 811.758	5	24 762.352	282.60	.005	.364
Error K	26 559.015	5	5 311.803	60.62	.005	.078
Error Sb	108 969.439	5	21 793.888	248.72	.005	.320
Error Bi	30 648.382	5	6 129.676	69.95	.005	.090
Error Cm	20 492.881	5	4 098.562	46.77	.005	.060
Variación L	149 542.703	25	5 981.708	68.27	.005	.439
Variación F	52 838.366	25	2 113.535	24.12	.005	.155
Variación K	150 091.109	25	6 003.644	68.52	.005	.441
Variación Sb	67 680.685	25	2 707.227	30.90	.005	.199
Variación Bi	146 001.742	25	5 840.070	66.65	.005	.429
Variación Cm	156 157.243	25	6 246.289	71.29	.005	.459

Para todas las condiciones, se rechaza la hipótesis nula (al 99.5% de confiabilidad*) de que las medias de los cuadrados sean las mismas para todas las fuentes de variación entre las escalas.

* Los niveles de probabilidad para rechazar la H_0 , fueron obtenidos de Snedecor y Cochran (30); los cálculos de los niveles de probabilidad por medio de $z = \log_e \sqrt{F}$ no fueron factibles dada la carencia de tablas con valores Z más allá de 5 desviaciones estándar; aunque para el valor F más pequeño de la tabla (6.76) la H_0 es rechazada con una confiabilidad del 99.53% ($p=0.0047$).

6).- ANALISIS FACTORIAL. (13, 14, 16, 18, 29)

El análisis factorial, así como las correlaciones, son métodos multivariantes (en contraposición con los métodos bivariantes más tradicionales de la Psicología Experimental) que se han constituido en aspectos metodológicos-necesarios para la identificación de habilidades y rasgos básicos diferentes, susceptibles de ser medidos por medio de tests. Habitualmente estos procedimientos se aplican a muestras formadas por N individuos, cada uno de los cuales ha tomado las pruebas; los datos empíricos se presentan en forma de una matriz de puntajes. A partir de dicha matriz de puntajes, trata de encontrarse un orden o sistema subyacente, algún principio que tenga significado psicológico en relación con las diferencias individuales en los puntajes de los tests.

La capitalización de las diferencias individuales como la fuente de varianza de los datos, es la característica más obvia. En los experimentos bivariantes de tipo tradicional, el experimentador aplica "tratamientos" en forma de variaciones de las condiciones estímulo, o condiciones de tiempo, número de exposiciones, etc.; es decir está interesado en la dependencia de la respuesta como una función del estímulo (dependencia E-R), en tanto que el análisis factorial trabaja con dependencias R-R, en las cuales los puntajes en distintas escalas son considerados como variables de respuesta.

El método de correlación no enfoca su atención sobre las medias y desviaciones estándar de las variables involucradas en la medición, en tanto que los experimentos bivariantes fundamentan sus hallazgos en las medias y en las varianzas. Estas se ven afectadas por un gran número de condiciones anteriores al experimento, lo que significa que el experimentador debe considerar los antecedentes de los sujetos, si es que quiere asegurarse de que es sólo su tratamiento el que produce determinados resultados. El ideal de conocer substancialmente la historia de los sujetos, sólo puede lograrse, de hecho, en grupos de animales bajo control.

Lo anterior no quiere, de ninguna manera, decir que quienes trabajan con métodos multivariantes pueden ignorar ciertas condiciones experimentales. Desafortunadamente, en muchas ocasiones se utilizan estos procedimientos como si no fuera necesario ningún control experimental, dado que puede someterse al análisis factorial a cualquier matriz de correlaciones que resulte. Thurstone ha advertido contra esta práctica, la cual, aparentemente, ha surtido poco efecto.

El análisis factorial debe prestar atención a dos importantes fuentes que determinan los resultados: la selección de la muestra de individuos y la elección de las variables experimentales. Es importante que los individuos que conforman la muestra en un estudio de rasgos de la personalidad sea lo más uniforme posible por lo que respecta a características tales como edad, escolaridad, sexo, como condiciones que puedan influir en los coeficientes de correlación; estas variables no pueden dejar de considerarse, a menos que se demuestre que no influyen considerablemente en las correlaciones.

Por otra parte la selección de las variables es importante, ya que para poder saturar un factor, comúnmente se requiere de un mínimo de tres escalas, un análisis factorial con un número restringido de variables puede ocasionar que algunos factores que aparentemente se reflejan fuera del contexto central de la matriz, no puedan ser interpretados.

El descubrimiento de factores diferenciados de rasgos de personalidad, implica por sí mismo la existencia de diferencias intraindividuales, con la posibilidad de describir a cada persona en términos de un perfil, más que de una sola calificación, permitiendo explicar el funcionamiento de distintas escalas en términos de un número mínimo de rasgos o factores.

El término factor se usa cuando menos, con un doble sentido; existe el factor matemático que se extrae de las intercorrelaciones y que permite representar al factor de manera taquigráfica. Después de realizar el análisis, se obtienen las correlaciones de n escalas con k factores, donde k es menor que n , en tanto que previamente se contaba una matriz de intercorrelaciones mucho mayor, de n tests, uno con otro. A partir de las propiedades conocidas de las escalas que están estrechamente relacionadas con un factor matemático, a diferencia de las propiedades que no están correlacionadas con el factor, se puede obtener, por aprehensión de los datos inherentes a la prueba, como una síntesis perceptual, una idea de qué variable psicológica puede estar siendo representada por el factor matemático. Dicha variable psicológica equivale al factor psicológico o rasgo. Un investigador puede equivocarse al denominar o interpretar un factor matemático; aunque esto no es más cierto para quien utiliza el análisis factorial que lo que es para quien utiliza cualquier otro procedimiento, máxime si la interpretación se apoya en la evidencia empírica para concebir la existencia de un factor.

Los resultados obtenidos del análisis factorial tienen dos funciones fundamentales: la primera es que tales resultados permiten demostrar que algo, que no es azar, opera para producirlos, y la segunda es que un concepto que se deriva de resultados empíricos y se mantiene en forma estable merece confianza en su utilidad potencial.

Cae fuera de los límites de este trabajo el analizar la estabilidad de los factores al través de diferentes procedimientos, aunque un gran número de estudios en este sentido corroboran la afirmación anterior (Guilford, 1948 y 1961; Zacherter y Friedman, 1953; Petersen, Hoepfner y Merrifield, 1963; French, 1951; etc.)

Al hablar de factores surge, casi automáticamente, la pregunta: ¿qué es un factor psicológico? Aunque no existe una respuesta unívoca a esta cuestión, puede decirse que es una forma de expresar la existencia de una variable subyacente, a lo largo de la cual los individuos difieren entre sí, del mismo modo que difieren en la escala de una prueba en la cual hay diferentes valores numéricos. Las escalas de los tests al través de las cuales se aprecian diferencias interindividuales, son variables manifiestas. Al estudiar

la prueba se observa el conjunto de puntajes obtenidos por los diferentes sujetos, pero no podemos observar ni el factor, ni las posiciones de los individuos a lo largo de la escala del factor, pero si pueden calcularse las cargas factoriales y, en algunos casos, construir escalas o pruebas con altas saturaciones factoriales. En otras palabras, los factores permiten demostrar que una escala tiene alguna significancia psicológica general que es algo más que la mera colección sumatoria de variables de actitud o aptitud.

Existen otros métodos que no son los del análisis factorial para llegar a concepciones sobre funciones psicológicas particulares. La clasificación de funciones ha sido reconocida siempre como una tarea legítima y necesaria del psicólogo teórico; pero probablemente ninguno de los métodos de investigación diseñados hasta la fecha resulte más indicado que el análisis factorial, adecuadamente utilizado, para la tarea de responder a las interrogantes taxonómicas de la Psicología.

E L M E T O D O C E N T R O I D E

El empleo del método centroide para efectuar el análisis de las escalas de validez del MMPI e IPP, en este trabajo, obedece básicamente a que es razonablemente factible de realizar con una calculadora de escritorio, cuando no se tiene acceso a computadoras de alta velocidad, que pueden efectuar este análisis haciendo uso de los complejos procedimientos desarrollados actualmente para esta técnica, a más de que este método permite identificar los factores comunes a la matriz de correlaciones original, lo que se adecúa perfectamente al objetivo de esta tesis, la cual pretende establecer qué factores comunes de las escalas estudiadas permiten explicar la postura asumida por los sujetos al responder los inventarios.

El método de factorialización, a partir de matrices de correlación, que probablemente ha sido más utilizado, es el método centroide, desarrollado por Thurstone (1947); que, aunque no proporciona una solución matemática única por mínimos cuadrados como puede ser el caso del método de ejes principales, está basado en totales y resulta relativamente sencillo de computar.

El nombre del método denota una estrecha relación con el concepto mecánico de un centroide o centro de gravedad; por esta razón, la forma centroide de análisis factorial encuentra su mejor descripción en términos geométricos. Las variables pueden considerarse como representadas por un grupo de n vectores los cuales se encuentran contenidos en un espacio de m dimensiones, donde m es el número de factores comunes; y el producto escalar de cualquier par de vectores es la correlación entre ellos. Las variables pueden considerarse también representadas por las m coordenadas de los puntos finales de estos vectores con respecto a los m ejes ortogonales, arbitrarios de referencia. A partir de la configuración de los vectores que representan a las variables puede determinarse la matriz de correlaciones original, el sistema de referencia puede rotarse sin afectar a dichas correlaciones. El sistema arbitrario de coordenadas puede ser rotado de manera tal que el centroide de un grupo de n puntos junto al origen, determina el primer eje de referencia; de donde es posible obtener la proyección de cada uno de los vectores, o la coordenada de cada variable sobre el primer eje de referencia por medio del centroide. Asumiendo que los residuos desaparecen, las correlaciones pueden obtenerse por medio de la siguiente fórmula:

$$(1) \pi_{jk} = r_{jk} = a_{j1} a_{k1} + a_{j2} a_{k2} + \dots + a_{jm} a_{km} \quad (j, k = 1, 2, \dots, n)$$

Donde m es el número de factores comunes. Los valores numéricos de a_{jp} ($p = 1, 2, \dots, m$) son determinados por la posición de los ejes ortogonales de referencia en tanto que a_{jp} es la p coordenada de la variable Z_j . En el sistema arbitrario de referencia ortogonal, las m coordenadas de cada uno de los n puntos son las siguientes:

$$(2) \begin{aligned} P_1 &: (a_{11}, a_{12}, \dots, a_{1p}, \dots, a_{1m}) \\ P_k &: (a_{k1}, a_{k2}, \dots, a_{kp}, \dots, a_{km}) \\ P_n &: (a_{n1}, a_{n2}, \dots, a_{np}, \dots, a_{nm}) \end{aligned}$$

Cualquiera de las m coordenadas del centroide es el promedio de las coordenadas correspondientes a estos n puntos, por lo tanto:

$$(3) \text{ Centroide: } \left(\frac{1}{n} \sum_K a_{k1}, \dots, \frac{1}{n} \sum_K a_{kp}, \dots, \frac{1}{n} \sum_K a_{km} \right),$$

donde la suma desde 1 hasta n se indica por medio del índice k cuando los límites no están dados específicamente.

Al trazar el marco de referencia del espacio factorial, se tiene que el primer eje F_1 pasa a través del centroide; luego entonces, todas las coordenadas del centroide serán cero, excepto la primera:

$$(4) \quad \sum_K a_{k2} = \sum_K a_{k3} = \dots = \sum_K a_{km} = 0$$

Los m valores se reducen entonces a:

$$(5) \quad \frac{1}{n} \sum_K a_{k1}, 0, 0, \dots, 0 \quad \text{donde hay } (m-1) \text{ ceros.}$$

Entonces el centroide cae sobre F_1 por lo tanto, la primera coordenada es también la distancia del centroide al origen.

Resulta ahora posible determinar los coeficientes, del primer factor centroide, por ejemplo las coordenadas a_j , en términos de las correlaciones observadas. Entonces, sumando todas las variables k en una determinada columna j de la matriz de correlaciones, resulta:

$$(6) \quad \sum_K r_{jk} = a_{j1} \left(\sum_K a_{k1} \right) + a_{j2} \left(\sum_K a_{k2} \right) + \dots + a_{jm} \left(\sum_K a_{km} \right)$$

La cual puede simplificarse de la manera siguiente:

$$(7) \quad \sum_K r_{jk} = a_{j1} \left(\sum_K a_{k1} \right)$$

Entonces la suma de todas las celdillas de la matriz de correlación es simplemente:

$$(8) \quad \sum_j \sum_K r_{jk} = \left(\sum_j a_{j1} \right) \left(\sum_K a_{k1} \right) = \left(\sum_K a_{k1} \right)^2$$

Ahora, la raíz cuadrada del lado izquierdo de la fórmula anterior puede ser sustituida por el término entre paréntesis de (7) obteniéndose la siguiente fórmula centroide básica:

$$(9) \quad a_{j1} = \frac{\sum_K r_{jk}}{\sqrt{\sum_j \sum_K r_{jk}}} = \frac{S_j}{\sqrt{T}} \quad (j=1, 2, \dots, n)$$

Donde S_j es la suma de todas las correlaciones en la columna j de la matriz, y T es el total de todas las correlaciones en la matriz, inclusive, los valores de la diagonal. En la fórmula de arriba, la raíz positiva fue obtenida arbitrariamente. Por supuesto, si aparece un signo negativo en el radical, el coeficiente del factor deberá cambiar de signo obteniéndose un factor

igualmente aceptable. La fórmula (9) da el coeficiente del primer factor cen-
troide \bar{F}_1 para cada variable Z_j o la primera coordenada para cada punto que
represente a una variable.

El siguiente paso es obtener los residuos del primer factor de donde se -
obtendrá la segunda coordenada. En tanto que los residuos de las correlacio-
nes con 1, 2, ... (m-1) factores que se mueven en el espacio factorial son uti-
lizados en sucesivas fases del método cenroide. Los residuos del primer fac-
tor están dados entonces por:

$$(10) \quad 1\pi_{jk} = \pi_{jk} - a_{j1}a_{k1} = a_{j2}a_{k2} + a_{j3}a_{k3} \dots a_{jm}a_{km}$$

Las correlaciones residuales pueden ser consideradas como productos esca-
lares de pares de vectores residuales en un espacio de (m-1) dimensiones (la di-
mensión del espacio residual es igual al número de términos del miembro dere-
cho de (10), o del rango de la matriz residual de correlaciones).

En este espacio residual, las (m-1) coordenadas de cada uno de los n pun-
tos puede ser representada por:

$$(11) \quad \begin{aligned} 1P_1 &: (a_{12}, a_{13}, \dots, a_{1p}, \dots, a_{1m}) \\ 1P_k &: (a_{k2}, a_{k3}, \dots, a_{kp}, \dots, a_{km}) \\ 1P_n &: (a_{n2}, a_{n3}, \dots, a_{np}, \dots, a_{nm}) \end{aligned}$$

Las (m-1) coordenadas del cenroide de estos n puntos son:

$$(12) \quad \left(\frac{1}{n} \sum_k a_{k2}, \dots, \frac{1}{n} \sum_k a_{kp}, \dots, \frac{1}{n} \sum_k a_{km} \right)$$

en donde, de acuerdo con (4), desaparecen todas; por consecuencia, el cenroide
se encuentra en el origen de este (m-1) espacio, y la fórmula (9) no puede ser
utilizada de manera directa para calcular los valores de los coeficientes del-
segundo factor. Debe señalarse que en la obtención de (9) la expresión (7) -
fue dividida por $\sum_k a_{k1}$ o n veces la distancia del cenroide al origen.

Se asume tácitamente que el cenroide no se encuentra en el origen, ya -
que si esto no fuera así, esta división no sería posible.

El problema inmediato es, entonces, mover el cenroide del origen en el -
espacio (m-1), de manera tal que el procedimiento señalado pueda aplicarse de-
nuevo.

Por medio de una rotación 180° de algunos vectores, alrededor del origen-
(a lo que se llama "reflexiones en el origen") el cenroide puede ser traslada-
do desde el origen, hasta un punto distinto del espacio factorial. Si las -
coordenadas de un punto P_j que representan una variable Z_j son:

$$(a_{j1}, a_{j2}, \dots, a_{jm}),$$

entonces, el punto reflejado $-P_j$ con coordenadas

$$(-a_{j1}, -a_{j2}, \dots, -a_{jm})$$

representará a la variable $-Z_j$, que corresponde a la variable original con dirección opuesta.

Ahora, partiendo de (1), resulta evidente que la inversión de signos de las coordenadas de P_j tienen el efecto de invertir los signos de todas las correlaciones de la variable Z_j , con los mismos residuos; esto quiere decir que la reflexión en el origen de una variable se realiza simplemente invirtiendo los signos de las correlaciones de esta variable en la matriz de correlaciones.

Por supuesto, el mismo argumento se mantiene, tanto para el espacio residual $(m-1)$, como para el espacio original de m dimensiones. Por lo tanto, la reflexión de una variable en el espacio residual se efectúa cambiando los signos de las correlaciones residuales para cada variable. En un intento para determinar qué variables deberían ser reflejadas, Thurstone (1935), sugiere que "es deseable explicar tanto como sea posible de la varianza residual por medio de cada uno de los factores sucesivos". Mientras que una aplicación rigurosa de este principio matemático nos llevaría a una solución de factor principal, el procedimiento centroide es solamente una aproximación gruesa. Para poder explicar toda la varianza residual posible por medio del segundo factor, el eje de referencia que lo representa, debe pasar a través de un grupo de vectores residuales. Si hay una agrupación de vectores (como podría ser el caso de un conjunto de variables que tuvieran grandes correlaciones residuales positivas) que permita equilibrar la dispersión de vectores en un punto diferente del origen, puesto que el centroide se encuentra en el origen; entonces el segundo eje de referencia deberá estructurarse de manera que pase a través de tal agrupación; de esta forma, parecería que los vectores que se dispersan, en vez de agruparse, podrían ser reflejados de manera que caigan fuera de la agrupación. Las coordenadas del segundo factor pueden entonces calcularse como en el caso del primero. En la práctica, aquellas variables que tienen un gran número de correlaciones negativas, deben ser reflejadas primero, trasladándolas al hemisferio de la agrupación. En situaciones prácticas Thurstone (1935) sugiere "invertir todos los signos de una escala en cada reflexión, hasta que el número de coeficientes negativos de la matriz residual sea menor que $N/2$ esto es, que para cada variable el número de signos negativos sea menor que $N/2$ (no los del total de la matriz). Sin embargo, no todo termina aquí, dado que, si se desea por añadidura "maximizar" la varianza moviendo cada uno de los factores sucesivos, la reflexión de variables debe continuarse hasta que la suma de las correlaciones para cada variable sea tan grande (positivamente) como sea posible.

Para designar al residuo de este trabajo de análisis resulta conveniente usar un símbolo que designe si un punto, que represente a una variable, ha sido reflejado en el origen.

Designando con ϵ_j la dirección del signo algebraico de P_j esto es, el punto $\epsilon_j P_j$ es $+P_j$ o $-P_j$. Si P_j no ha sido reflejado, ϵ_j es más, pero si P_j ya fue reflejado, entonces ϵ_j es menos. Además, ϵ_j puede considerarse como un operador algebraico definido como sigue:

$$\varepsilon_j = \begin{cases} +1 & \text{si } Z_j \text{ no ha sido reflejado} \\ -1 & \text{si } Z_j \text{ ya fue reflejado} \end{cases}$$

En consecuencia, ε_j puede ser asociado a las coordenadas de P_j y ser tratado como cualquier otra cantidad algebraica. Así, si los residuos de las correlaciones del primer factor después de la reflexión de ciertas variables se indican por γ_{ijk} (a diferencia de $|\gamma_{ijk}$ antes de la reflexión), estos términos pueden escribirse de la manera siguiente:

$$(13) \quad \gamma_{ijk} = \varepsilon_j \varepsilon_k (a_{j2} a_{k2} + a_{j3} a_{k3} + \dots + a_{jm} a_{km})$$

Este resultado sigue inmediatamente después de (10), donde cada a_{jp} y a_{kp} es reemplazada por $\varepsilon_j a_{jp}$ y $\varepsilon_k a_{kp}$ respectivamente. Si ninguna de las variables Z_j ni Z_k fueron reflejadas, o si ambas fueron reflejadas, entonces $\gamma_{ijk} = -|\gamma_{ijk}$. En otras palabras,

$$\gamma_{ijk} = \varepsilon_j \varepsilon_k (|\gamma_{ijk}|)$$

Las $(m-1)$ coordenadas del centroide de n puntos, dadas originalmente en (12), después de la reflexión de variables se convierte en:

$$\frac{1}{n} \sum_k \varepsilon_k a_{kp} \quad (p=2, 3, \dots, m)$$

Ahora el sistema de referencia puede ser rotado sobre el primer eje de referencia F_1 de manera tal que el segundo eje F_2 pase a través de este centroide*. Asumiendo que esta operación ha sido realizada, se encuentra que no resulta necesario cambiar la notación para las coordenadas del centroide, las que son:

$$\frac{1}{n} \sum_k \varepsilon_k a_{k2}, 0, 0, \dots, 0$$

dado que el centroide cae sobre el eje F_2 . A partir de los valores de las últimas $(m-2)$ coordenadas, puede derivarse la siguiente expresión:

$$(14) \quad \sum_k \varepsilon_k a_{k3} = \sum_k \varepsilon_k a_{k4} = \dots = \sum_k \varepsilon_k a_{km} = 0,$$

que corresponde a (4) en el caso del primer centroide.

Ahora las proyecciones de los vectores sobre el segundo eje centroide -- pueden ser expresadas en términos de las correlaciones residuales. En consecuencia, sumando (13) a todas las variables k en una determinada columna j de la matriz de correlaciones residuales (después de la reflexión de variables) -- y aplicando (14), resulta:

$$\sum_k \gamma_{ijk} = \varepsilon_j a_{j2} \sum_k \varepsilon_k a_{k2}$$

* El espacio residual $(m-1)$ es ortogonal al primer eje de referencia F_1 . El segundo eje F_2 puede entonces rotarse a cualquier posición en el espacio residual, y formará un ángulo recto con respecto a F_1 .

Entonces, sumando para todas las columnas:

$$\sum_j \sum_k \gamma_{jkk} = \sum_j \epsilon_j a_{j2} \sum_k \epsilon_k a_{k2} = \left(\sum_k \epsilon_k a_{k2} \right)^2$$

De donde se obtiene que:

$$\epsilon_j a_{j2} = \frac{\sum_k \gamma_{jkk}}{\sqrt{\sum_j \sum_k \gamma_{jkk}}} = \frac{S_{j1}}{\sqrt{T_1}}$$

o, multiplicando ambos lados por ϵ_j ,

$$(15) \quad a_{j2} = \frac{\epsilon_j S_{j1}}{\sqrt{T_1}} \quad (j=1, 2, \dots, n)$$

donde S_{j1} es la suma de todas las celdillas en la columna j de la matriz residual de las correlaciones del primer factor y T_1 es el total de todas las correlaciones en esta matriz, los signos de todos los coeficientes son los que aparecen después de la reflexión. Las ϵ_j indican que, si una variable Z_j fue reflejada, debe haber cambiado su signo algebraico, pero si esta variable no fue reflejada, entonces ϵ_j es simplemente +1. En otras palabras, $\epsilon_j S_{j1}$ es la suma de todas las correlaciones residuales. Entonces, por definición:

$$(16) \quad |S_j = \epsilon_j S_{j1}$$

y la fórmula (15) puede transformarse en:

$$(17) \quad a_{j2} = \frac{|S_j}{\sqrt{T_1}} \quad (j=1, 2, \dots, n)$$

En esta fórmula el numerador representa la suma de las correlaciones residuales para la variable j , en tanto que el denominador T_1 indica la suma de todas las correlaciones residuales después del cambio de signo. La fórmula (17) da los coeficientes del segundo factor centroide F_2 para cada variable.

Las restantes cargas factoriales pueden obtenerse de forma similar. Si guiendo el principio básico de explicar tanto como sea posible de la varianza de las variables para cada factor, éstas aparecen reflejadas en los espacios residuales producto de su agrupación. Cuando se efectúa el cambio de signo, el centroide del conjunto de puntos en el espacio residual cae en algún lugar de la agrupación de variables, y el siguiente eje de referencia pasa a través de este centroide. Cada uno de los sucesivos ejes centroides forma un ángulo recto con el eje precedente, dado que el espacio residual es ortogonal al espacio de los ejes centroides previamente establecidos. Después de la extracción de cada factor centroide, las correlaciones residuales decrecen en magnitud, y el rango de cada matriz residual se reduce, teóricamente, en uno. El desarrollo precedente fue hecho sin ninguna restricción en los elementos diagonales; el número de factores que pueden ser obtenidos de la matriz, depende de estos valores en la diagonal.

En la más simple de las estimaciones arbitrarias de la comunalidad, la mayor de las correlaciones de cada columna de la matriz*, se elige como ele--

* mento para la diagonal; en cada una de las matrices residuales subsecuentes, el término de la diagonal no se mantiene, sino que se va fijando para cada columna como los demás residuos de la matriz, (otra alternativa puede ser colocar en cada celdilla de la matriz residual, el coeficiente residual más grande) sin atender a su signo algebraico, en cada columna. Este procedimiento no proporciona ninguna indicación al respecto del número de factores comunes en una matriz.

Si en vez de modificar los valores de la diagonal en cada fase del procedimiento, el análisis se aplica directamente, se mantienen los términos de la diagonal con los mismos valores calculados, entonces el número de factores se determina cuando los valores originales de la diagonal han sido resueltos. Si se analizan los valores de la diagonal de manera directa, solo pueden obtenerse los factores comunes de la matriz, es decir no pueden extraerse factores específicos. En la práctica, es recomendable emplear, no una aproximación arbitraria, sino una estimación completa de las comunales; cuando se usa tal estimación, y se aplica directamente el análisis centroide (cual es el presente caso), la solución resultante, culmina en un número definido de factores comunes.

Por último, y antes de proceder al análisis factorial de los datos, señalaremos que el cálculo de las comunales para este trabajo se hizo por medio de una aproximación centroide completa, en donde las comunales involucran, en esencia, el cálculo del primer factor centroide. Como punto de arranque, se inserta en la diagonal principal de la matriz, el mayor coeficiente de correlación de cada variable. Entonces, la comunalidad estimada se obtiene como la razón del cuadrado de la suma de cada columna sobre la suma total de todas las correlaciones en la matriz, como sigue:

$$(18) \quad h_j^2 = \left(\sum_{k=1}^n r_{jk} \right)^2 / \sum_{k=1}^n \sum_{l=1}^n r_{kl}$$

* Este procedimiento simple fue usado por Thurstone y sus seguidores en un gran número de análisis por solución centroide, (se encontró, sin embargo, que esta aproximación arbitraria de estimar las comunales puede ser empleada en matrices grandes de correlaciones, pero no resulta recomendable cuando se emplea un número restringido de variables); aunque existen otras aproximaciones arbitrarias para estimar la comunalidad, como pueden ser el promedio de los coeficientes $(h_j^2 = \sum_{k=1}^n r_{jk} / (n-1))$ o el empleo de triadas con las dos correlaciones más altas (l y k) de una columna determinada $(h_j^2 = r_{jk} r_{jl} / r_{kl})$. Otros métodos que implican un cálculo previo del rango de la matriz, por lo general no son factibles sin una computadora de alta velocidad, dada la complejidad de las fórmulas y de la labor de cálculo para determinar la comunalidad.

PROCEDIMIENTO DE CALCULO.

El cálculo de la presente solución centroide se realizó usando las inter correlaciones de seis escalas de validez: L, F y K del MMPI y Sb, Bi y Cm del IPP. Las correlaciones fueron realizadas con 337 sujetos del sexo masculino y con estudios de ingeniería (ver la composición de la muestra en el inciso 1 de este capítulo) y aparecen en la tabla III-1.

Las estimaciones de comunalidad en la diagonal, se obtuvieron por medio de una aproximación centroide completa de acuerdo con (18).

TABLA III-1

MATRIZ DE CORRELACIONES (R)

	1	2	3	4	5	6
	L	F	K	Sb	Bi	Cm
1 L	(.586)	-.254	.558	.445	.518	-.339
2 F	-.254	(.264)	-.321	-.434	-.224	-.198
3 K	.558	-.321	(.608)	.576	.374	-.249
4 Sb	.445	-.434	.576	(.374)	.452	-.427
5 Bi	.518	-.224	.374	.452	(.598)	-.136
6 Cm	-.339	-.198	-.249	-.427	-.136	(.225)
Sj	1.514	-1.167	1.546	.986	1.582	=1.124

En la mayoría de los casos, los totales de las columnas de la tabla de intercorrelaciones, son todos positivos; sin embargo si alguno de los totales de las columnas resulta negativo, como en el caso de la tabla III-1, es necesario realizar el proceso de reflexión para la matriz de correlaciones. Este procedimiento es igual al que se efectúa para con las matrices residuales, y queda explicado adelante.

TABLA III-2

Reflexión de signos para la matriz de correlación.

Variable	Variables Reflejadas	Antes de la Reflexión	Después de la reflexión de variables sucesivas	
			F	Cm
L		2	1	0
F	-	5	0	1
K		2	1	0
Sb		2	1	0
Bi		2	1	0
Cm	-	5	4	1
Total		18	8	2
Diferencia			10	6

En consecuencia, la matriz de correlaciones R después de la reflexión queda de la manera siguiente:

TABLA III-3
Matriz de Correlaciones (R)
(Después de la Reflexión)

	L	F*	K	Sb	Bi	Cm*
L	(.586)	.254	.558	.445	.518	.339
F*	.254	(.264)	.321	.434	.224	-.198
K	.558	.321	(.608)	.576	.374	.249
Sb	.445	.434	.576	(.374)	.452	.427
Bi	.518	.224	.374	.452	(.598)	.136
Cm*	.339	-.198	.249	.427	.136	(.225)
Sj	2.700	1.299	2.686	2.708	2.302	1.178
Aj1	.753	.362	.749	.755	.642	.328

$$T = 12.873 = 3.588^2$$

(*)Variables reflejadas.

Los coeficientes del primer factor se obtienen por S_j/\sqrt{T} por lo que es necesario calcular las sumas de las columnas (S_j) y el total (T) de todas las correlaciones. El total (T) se obtiene simplemente sumando todas las S_j . Entonces, dividiendo S_j por \sqrt{T} se obtienen los coeficientes del primer factor centroide (a_{j1}) que aparece en la última línea de la matriz R.

Un procedimiento para verificar este cálculo, consiste en sumar las cargas de cada factor (a_{j1}) el que se denota por D_1 como sigue:

$$(19) \quad D_1 = \sum a_{j1} = \sum S_j / \sqrt{T} = T / \sqrt{T} = \sqrt{T}$$

En este caso $D_1 = 3.588$

Se calcula la matriz de saturación (Tabla III-4), multiplicando las cargas factoriales (a_{j1}) entre sí, tal como se indica:

TABLA III-4
Matriz de Saturación: $R_1 = (a_{ji}a_{ki})$

	L	F	K	Sb	Bi	Cm	
	.753	.362	.749	.755	.642	.328	
L	.753	(.567)	.273	.564	.569	.483	.247
F	.362	.273	(.131)	.271	.273	.232	.119
K	.749	.564	.271	(.561)	.565	.481	.246
Sb	.755	.569	.273	.565	(.570)	.485	.248
Bi	.642	.483	.232	.481	.485	(.412)	.211
Cm	.328	.247	.119	.246	.248	.211	(.108)
E _{j1}	2.703	1.299	2.688	2.710	2.304	1.179	
A _{j1} D ₁	2.702	1.299	2.687	2.709	2.303	1.177	

Los elementos de la matriz se verifican por medio de la siguiente relación entre las sumas de las columnas E_{j1} y D_1

$$(20) \quad E_{j1} = \sum_k a_{ji} a_{ki} = a_{ji} \sum_k a_{ki} = a_{ji} D_1$$

El procedimiento para obtener la matriz residual del primer factor en el método centroide es el siguiente:

Se sustraen los coeficientes de la matriz de saturación de los coeficientes de la matriz R que ocupen el mismo lugar en las matrices respectivas, tal como se muestra a continuación.

TABLA III-5
Residuos del Primer Factor (γ_{jk})

	L	F	K	Sb	Bi	Cm
L	(.019)	-.019	-.006	-.124	.035	.092
F	-.019	(.133)	.050	.161	-.008	-.317
K	-.006	.050	(.047)	.011	-.107	.003
Sb	-.124	.161	.011	(-.196)	-.033	.179
Bi	.035	-.008	-.107	-.033	(.186)	-.075
Cm	.092	-.317	.003	.179	-.075	(.117)
$\sum_k \gamma_{jk}$	-.003	.000	-.002	-.002	-.002	-.001

En la localización espacial de los factores residuales, el centroide debe encontrarse en el origen y por consiguiente, la suma de las columnas debe ser cero, excepto por errores de aproximación. Al trasladar el centroide, desde el origen en el espacio de los factores residuales, para incrementar la contribución del segundo factor a la varianza residual, se reflejan ciertas variables en el origen, como se muestra en la tabla III-6.

TABLA III-6

Número de Signos Negativos para los residuos del primer factor después de sucesivas reflexiones de las variables.

Variable	Variables Reflejadas	Antes de la Reflexión	Después de la Reflexión de las variables sucesivas		
			Bi	L	
L	-	3	4	1	
F		3	2	1	
K		2	1	0	
Sb		2	1	0	
Bi	-	4	1	0	
Cm		2	1	2	
Total		16	10	4	
Diferencia			6	6	

En la columna encabezada por "variables reflejadas" se coloca un signo de menos frente a cada una de las variables que son reflejadas. Las columnas restantes son introducidas sucesivamente a medida que las variables van siendo reflejadas. El proceso se inicia contando el número de signos negativos para cada variable de la matriz residual, y registrándolo en la columna "Antes de la Reflexión". Si varias variables tienen el mismo número de máximo de signos negativos, cualquiera de ellas puede ser seleccionada, arbitrariamente, para la reflexión.

En el presente caso la variable Bi tiene cuatro signos negativos, por lo que se la selecciona para la reflexión. Frente a esta variable, en la primera columna de la Tabla III-6, se escribe un signo negativo para indicar que es la que se utilizará para la reflexión. Se hace un ajuste en el número de signos negativos para cada variable, como si todos los signos de la variable Bi hubieran sido cambiados. Tales resultados se registran en la columna encabezada por "Bi" para indicar que el cómputo de signos negativos para cada variable es el que resulta después de que la variable Bi es reflejada.

Después de la reflexión de una variable dada, todos los residuos que eran positivos se transforman en negativos y cada residuo negativo se vuelve positivo, excepto los valores de la diagonal de la matriz residual, los que permanecen sin cambios. Por lo tanto, para la variable reflejada, el ajuste de signos negativos es $(n-1)$ menos el número de signos negativos que tenía antes de la reflexión. En este caso, $n-1 = 5$ y la anotación para la variable Bi después de la reflexión es $5 - 4 = 1$.

No es necesario cambiar todos los signos de los residuos para la variable que está siendo reflejada para poder contar el número de signos negativos para

las otras variables después de la reflexión. En vez de ello considérese el signo de cada coeficiente, excepto los de la diagonal de la matriz residual, para la variable que está siendo reflejada, y se prosigue como sigue:

a).- Si la anotación para una variable particular, que no haya sido previamente reflejada, es positivo, se incrementa en uno el número de signos para la variable y se registra el nuevo valor para esta variable en la columna "Bi". Por ejemplo, el coeficiente de la variable Cm en la columna Bi de la Tabla III-5 es positivo, y puesto que Cm no ha sido reflejada previamente, el número de signos negativos para Cm se incrementa en uno, de 1 a 2, en la Tabla III-6, después de que la variable Bi ha sido reflejada.

b).- Si la anotación de una variable particular, la cual no haya sido previamente reflejada, es negativa, se resta uno al número de signos negativos para esta variable y se anota el nuevo valor en la columna encabezada por Bi. Por ejemplo, el coeficiente para la variable Sb en la columna Bi de la matriz residual es negativo, y puesto que Sb no ha sido reflejada, el número de signos negativos es decrementado en uno, produciendo un coeficiente negativo después de que Bi es reflejada. Las reglas generales para el cambio de signo son:

REGLAS PARA AJUSTAR EL CAMBIO DE SIGNO.

Reflexiones Previas	El coeficiente en la hilera (o columna) de la variable reflejada es positivo	El coeficiente en la hilera (o columna) de la variable reflejada es negativo.
No reflejada previamente (o reflejadas un número igual de veces).	AUMENTAR UNO	DISMINUIR UNO
Reflejadas previamente una vez (o cualquier número impar de veces)	DISMINUIR UNO	AUMENTAR UNO

Después de que una variable ha sido reflejada, se prosigue con la siguiente columna de la Tabla III-6 y se selecciona la variable con mayor número de signos negativos para la reflexión, siguiendo el mismo procedimiento hasta que el número de residuos negativos para cada variable sea menor que la mitad del número total de variables. En este caso, $n = 6$, de modo que las reflexiones son llevadas hasta el punto en que haya 2 ó menos signos negativos para cada variable

Deben señalarse varias situaciones excepcionales:

Si aparecieran ceros en cualquier correlación o residuo deberán ser tratados como números positivos al efectuar ajustes de signos para la reflexión de variables. Los valores de la diagonal de las matrices residuales no se considerarán en el cómputo de signos negativos, así, si una variable es reflejada, esta "autocorrelación" permanece sin cambio. Puede acontecer que una variable, la cual hubiera sido reflejada anteriormente, pueda volver a aparecer como la variable con un máximo de signos negativos, después de que varias otras variables hayan sido reflejadas. En este caso, la variable se refleja de nuevo, cambiando de menos a más el signo de la primera columna de la Tabla III-6, y el número de signos negativos se ajusta para cada una de las variables. Para formular la tabla de residuos del primer factor después de la reflexión (γ_{ijk}) y el cálculo de los coeficientes del segundo factor, se pone en primer lugar en los encabezados de la tabla III-7, un signo negativo antes de cada una de las variables que han sido reflejadas, en este caso Bi y L como sigue:

TABLA III-7

Residuos después de la Reflexión

	-L	F	K	Sb	-Bi	Cm
-L	(.019)	.019	.006	.124	.035	-.092
F	-.019	(.133)	.050	.161	.008	-.317
K	.006	.050	(.047)	.011	.107	.003
Sb	.124	.161	.011	(-.196)	.033	.179
-Bi	.035	.008	.107	.033	(.186)	.075
Cm	-.092	-.317	.003	.179	.075	(.117)
S_{j1}	.111	.054	.224	.312	.444	-.035
$\epsilon_j S_{j1}$	-.111	.054	.224	.312	-.444	-.035
Z_{j2}	-.105	.051	.213	.296	-.421	-.033

$$T=1.110 = 1.054^2$$

Los signos de los residuos originales del primer factor (γ_{ijk}) pueden ser cambiados de manera de obtener los signos algebraicos de los residuos después de la reflexión de variables, de la manera siguiente:

$$(21) \quad \gamma_{ijk} = \epsilon_j \epsilon_k (\gamma_{ijk})$$

puesto que las épsilons son meramente símbolos algebraicos para los signos positivos o negativos, si ninguna de las variables Z_j o Z_k es reflejada, o si ambas aparecen reflejadas, entonces $\gamma_{ijk} = |\gamma_{ijk}|$; pero, si solamente una de ellas aparece reflejada, entonces $\gamma_{ijk} = -|\gamma_{ijk}|$. El número de signos negativos después de la reflexión de variables debe coincidir con los indicados en la última columna de la Tabla III-6.

Los coeficientes del segundo factor se calculan por medio de (15), utilizando los residuos después de la reflexión de la misma manera que si fuera la -

matriz de correlaciones original, repitiendo el procedimiento hasta que la matriz se agote.

Después de obtener la suma de las columnas, se cambian los signos algebraicos de aquellas variables que han sido reflejadas. Los coeficientes que aparecen en la última línea de la Tabla III-7 se refieren a las variables observadas, no a las reflejadas. Una forma de verificar el cómputo de los coeficientes para todos los factores, es proporcionada por el hecho de que la suma de los coeficientes debe ser aproximadamente de cero (en este caso .000). Los coeficientes del segundo factor suman precisamente cero, excepto por errores de aproximación (aquí, la $\sum a_{j2} = .001$).

El procedimiento señalado se repite hasta obtener los factores suficientes, de manera que la comunalidad total con que se inició el cálculo en la matriz R, pueda ser explicada.

TABLA III-8
Matriz de Saturación: $R_2 = (a_{j2} a_{k2})$

	L	F	K	Sb	Bi	Cm
	.105	.051	.213	.296	.421	-.033
L	.105 (.011)	.005	.022	.031	.044	-.003
F	.051	.005 (.003)	.011	.015	.021	-.002
K	.213	.022	.011 (.045)	.063	.090	-.007
Sb	.296	.031	.015	.063 (.088)	.125	-.010
Bi	.421	.044	.021	.090	.125 (.177)	-.014
Cm	-.033	-.003	-.002	-.007	-.010	-.014 (.001)
E_{j2}	.110	.053	.224	.312	.443	-.035
$a_{j2} D_2$.111	.054	.225	.312	.444	-.035

TABLA III-9
Residuos del Segundo Factor ($2\gamma_{jk}$)

	L	F	K	Sb	Bi	Cm
L	(.008)	.014	-.016	.093	-.009	-.089
F	.014	(.130)	.039	.146	-.013	-.315
K	-.016	.039	(.002)	-.052	.017	.010
Sb	.093	.146	-.052	(-.284)	-.092	.189
Bi	-.009	-.013	.017	-.092	(.009)	.089
Cm	-.089	-.315	.010	.189	.089	(.116)
$\sum_k 2\gamma_{jk}$.001	.001	.000	.000	.001	.000

TABLA III-10

Número de Signos Negativos para los Residuos del segundo factor después de sucesivas reflexiones de las variables.

Variable	Variables Reflejadas	Antes de la Reflexión	Después de la reflexión de variables sucesivas		
			Bi	K	Cm
L		3	2	1	0
F		2	1	2	1
K	-	2	3	2	1
Sb		2	1	0	1
Bi	-	3	2	1	0
Cm	-	2	3	4	1
Total		14	12	10	4
Diferencia			2	2	6

TABLA III-11

Residuos después de la Reflexión (γ_{2jk})

	L	F	-K	Sb	-Bi	-Cm
L	(.008)	.014	.016	.093	.009	.089
F	.014	(.130)	-.039	.146	.013	.315
-K	.016	-.039	(.002)	.052	.017	.010
Sb	.093	.146	.052	(-.284)	.092	-.189
-Bi	.009	.013	.017	.092	(.009)	.089
-Cm	.089	.315	.010	-.189	.089	(.116)
S_{j2}	.229	.579	.058	-.090	.229	.430
$\epsilon_j S_{j2}$.229	.579	-.058	-.090	-.229	-.430
a_{j3}	.191	.483	-.048	-.075	-.191	-.359

$$T = 1.435 = 1.198^2$$

TABLA III-12

RESULTADO CENTROIDE PARA SEIS ESCALAS DE VALIDEZ

	Coeficientes de los Factores			Comunalidad		
	C ₁	C ₂	C ₃	(1) Original	(2) Calculada	(1) - (2)
L	.753	-.105	.191	.586	.614	-.028
F	-.362	.051	.483	.264	.367	-.103
K	.749	.213	-.048	.608	.608	.000
Sb	.755	.296	-.075	.374	.664	-.290
Bi	.642	-.421	-.191	.598	.625	-.027
Cm	-.328	-.033	-.359	.225	.238	-.013
Total	—	—	—	2.655	3.116	-.461
Contribución del Factor (Vp*)	2.349	.325	.442	—	—	—
Porcentaje del - total de la comu- nalidad original	88.475	12.241	16.648	—	117.363	-17.363

El número máximo de factores (r) que pueden ser obtenidos de una matriz de correlaciones con n variables queda determinado por:

$$r = \frac{2n + 1 - \sqrt{8n + 1}}{2}$$

por lo que, en nuestro caso, el número máximo de factores es:

$$r = \frac{2(6) + 1 - \sqrt{8(6) + 1}}{2} = 3$$

De acuerdo con los lineamientos señalados antes, el análisis factorial debe continuarse hasta un nivel que permita explicar la comunalidad inicial de la matriz R; los datos de la Tabla III-12, muestran que la extensión del procedimiento de cálculo hasta su máximo teórico (tres factores) producen excedente de más de 17% de la comunalidad original, en tanto que si la solución centroide se da al nivel de dos factores, este excedente es sólo del 0.716%. En consecuencia la solución centroide definitiva para este trabajo queda explicada, de acuerdo con el procedimiento de cálculo, en dos factores, tal como se muestra en la siguiente tabla.

$$* V_p = \sum_{j=1}^n a_{jp}^2 \quad (p = 1, 2, \dots, m)$$

TABLA III-13

SOLUCION CENTROIDE PARA SEIS ESCALAS
DE VALIDEZ

	Coeficientes de los Factores		Comunalidad		
	C_1	C_2	(1) Original	(2) Calculada	(1) - (2)
L	.753	-.105	.586	.578	.008
F	-.362	.051	.264	.134	.130
K	.749	.213	.608	.606	.002
Sb	.755	.296	.374	.658	-.284
Bi	.642	-.421	.598	.589	.009
Cm	-.328	-.033	.225	.109	.116
Total	—	—	2.655	2.674	-.019
Vp	2.349	.325	—	—	—
Porcentaje - del total de la comunali- dad original	88.475	12.241	—	100.716	-.716

ROTACION DE LOS EJES DE REFERENCIA.

Una crítica que se ha hecho siempre al análisis factorial ha sido que no existe una localización única para el marco de referencia de los vectores. Una tabla dada de correlaciones tiene sin embargo, una configuración simple de vectores, y tal configuración puede ser descrita con respecto a un número infinito de localizaciones dentro del marco de referencia de un sistema de coordenadas. Pueden obtenerse un gran número de rotaciones de los vectores por rotación de los ejes de referencia alrededor del origen.

En las páginas siguientes se muestra la localización espacial de las seis variables utilizadas para este estudio, tanto para su posición original en el espacio factorial, como después de haber rotado los ejes hasta posiciones que faciliten su interpretación.

Cuando como en este caso, los vectores se encuentran representados en un plano puede emplearse cualquier grado de rotación de los ejes sin alterar la configuración de los vectores; sin embargo, con cada nueva localización de los ejes las cargas de las escalas en los factores serán diferentes, aunque la comunalidad no se altera. Los ejes pueden ser rotados en forma dextrógira o sinestrógira en el grado deseado sin alterar las correlaciones entre las variables. Para efectuar rotaciones dextrógiras de dos ejes se utilizan las siguientes ecuaciones:

$$a'_{j1} = a_{j1} \cos \phi + a_{j2} \sin \phi$$

$$a'_{j2} = -a_{j1} \sin \phi + a_{j2} \cos \phi$$

En tanto que para rotaciones sinestrogiras las ecuaciones son:

$$a'_{j1} = a_{j1} \cos \phi - a_{j2} \sin \phi$$

$$a'_{j2} = a_{j1} \sin \phi + a_{j2} \cos \phi$$

Donde:

ϕ = Angulo al través del cual se rotan los ejes.

a'_{j1} = La carga rotada del test j en el factor I.

a'_{j2} = La carga rotada del test j en el factor II.

a_{j1} = La carga del test j en el factor I antes de la rotación.

a_{j2} = La carga del test j en el factor II antes de la rotación.

Ejes C_1 y C_2

Rotación Sinestrogira de $33^\circ 10'$

$$a'_{j1} = a_{j1} \cos \phi - a_{j2} \sin \phi$$

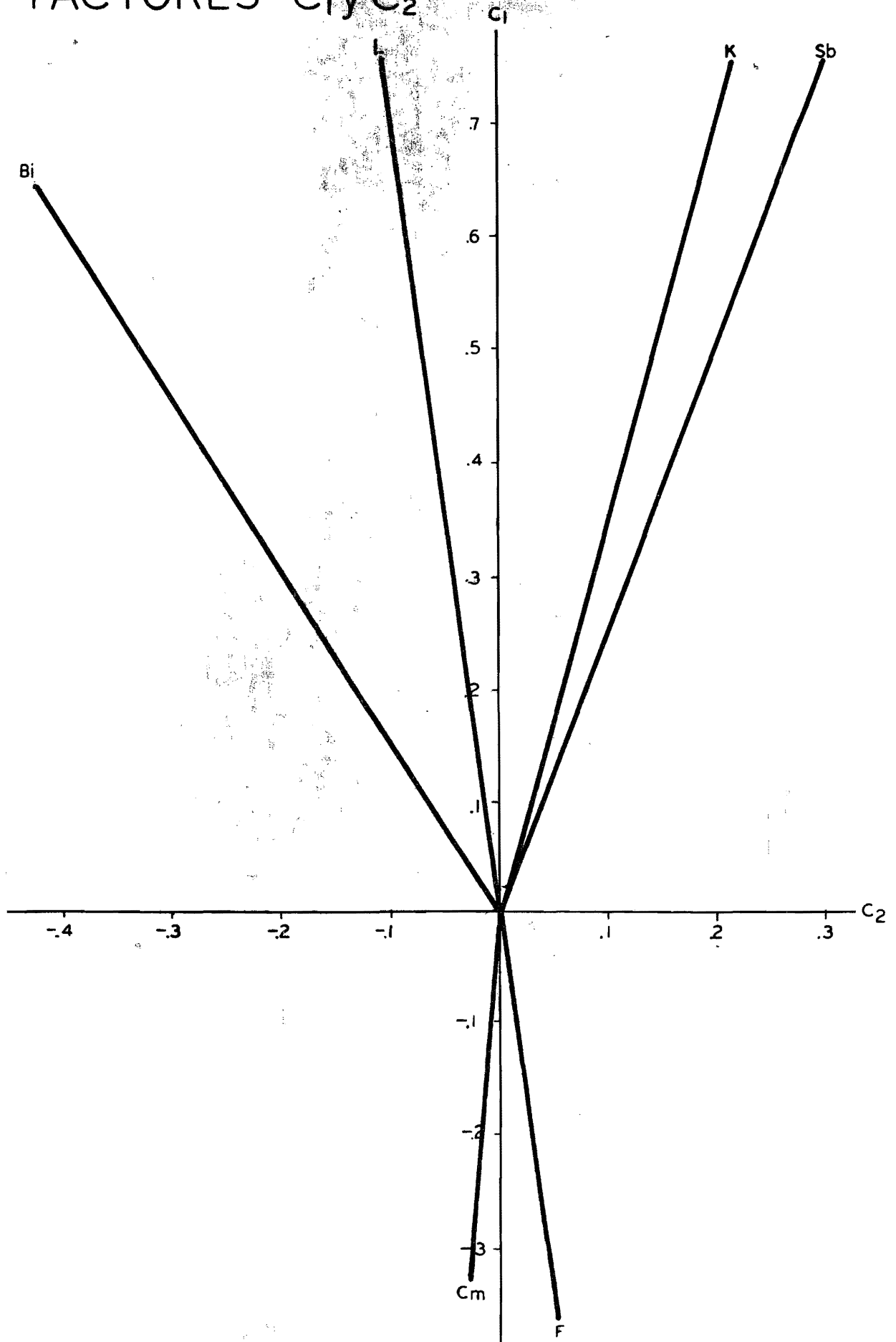
$$a'_{j2} = a_{j1} \sin \phi + a_{j2} \cos \phi$$

$$a'_{j1} = a_{j1} (.8371) - a_{j2} (.5471)$$

$$a'_{j2} = a_{j1} (.5471) + a_{j2} (.8371)$$

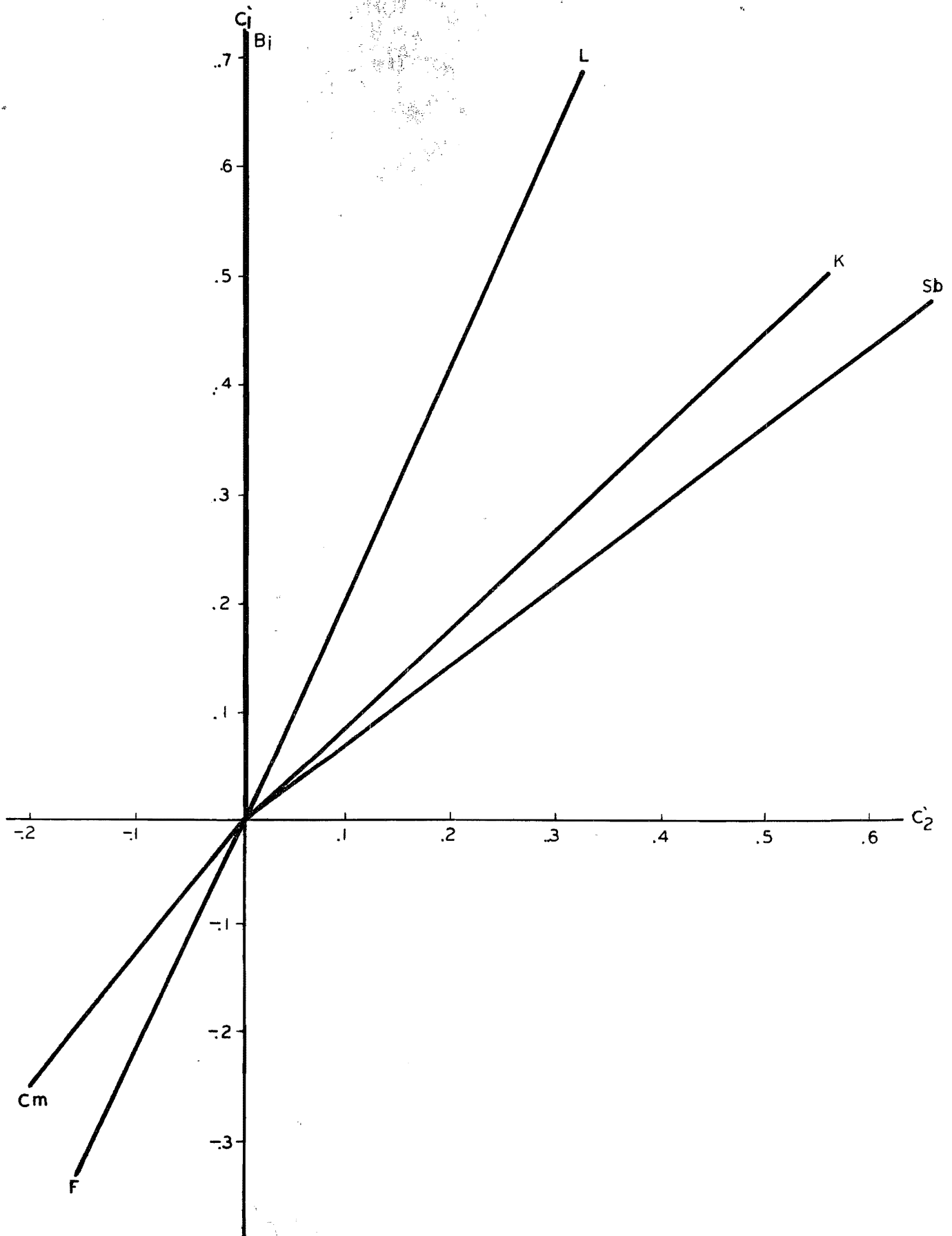
	C	C	C	C	h	h
L	.753	-.105	.687	.324	.578	.577
F	-.362	.051	-.331	-.155	.134	.134
K	.749	.213	.510	.588	.606	.606
Sb	.755	.296	.470	.661	.658	.658
Bi	.642	-.421	.767	-.001	.589	.588
Cm	-.328	-.033	-.257	-.207	.109	.109

FACTORES C_1 y C_2



FACTORES C_1 y C_2

Rotacion Sinestrógrica: $33^\circ 10'$



COMENTARIOS

Al hacer una revisión de los resultados arrojados por el análisis factorial, saltan a la vista tres grandes aspectos que merecen algunos comentarios e intentos de interpretación o explicación.

El primero de ellos se refiere a las cargas factoriales del primer factor, después de haber realizado una rotación ortogonal de los ejes de referencia, para facilitar su estudio. Como ha podido verse, de las seis escalas analizadas, cuatro de ellas tienen comunales elevadas, a saber: Sb (.658), K (.606), Bi (.589) y L (.578), lo que señala que una gran parte de la varianza total está determinada por aspectos que son comunes a las cuatro escalas. Cuando revisamos las cargas factoriales de las escalas en el primer factor (C_1) vemos que las escalas Bi y L correlacionan .77 y .69 con C_1 y constituyen las cargas más elevadas; se observan cargas moderadamente elevadas en K (.51) y en Sb (.47); en tanto que las escalas F y Cm correlacionan negativamente con el primer factor (-.33 y -.26 respectivamente).

Ahora bien, la mayor parte de la varianza de las escalas Bi y L puede explicarse en función de C_1 como factor común. Al hacer una revisión de los reactivos que aparecen en ambas escalas, parece bastante coherente pensar que dicho factor puede ser explicado como "deseabilidad social"; es decir, una disposición de respuesta a los ítems de las escalas que no parece estar necesariamente dada porque las respuestas atiendan estrictamente al contenido mismo de los enunciados, si no más bien por el deseo del sujeto de presentar un cuadro favorable de sí mismo, que considera como socialmente deseable. Al hacer estas afirmaciones, no hemos perdido de vista el hecho de que los datos aquí manejados fueron producto de una situación en la que los sujetos realizaban las pruebas como un requisito para conseguir empleo.

Si revisamos algunos de los reactivos de estas escalas (la composición total de las escalas se muestra en el Capítulo II) resulta bastante clara la forma en que los ítems pueden reflejar aspectos socialmente deseables:

MMPI - De vez en cuando dejo para mañana lo que debiera hacer hoy. (Falso: -.53*).

- Mis modales en la mesa no son tan correctos en casa como cuando salgo a comer fuera en compañía de otros. (Falso: .52)

* Las proporciones de respuesta entre paréntesis fueron obtenidas de un estudio factorial en preparación sobre 33 escalas del MMPI en una muestra de 2000 profesionales mexicanos que tomaron la prueba como parte del proceso de selección de personal (Gudiño, ... 1973). Todavía no se dispone de datos al respecto para el IPP.

- Si pudiera entrar a un cine sin pagar, y estuviera seguro de no ser visto, seguramente lo haría. (Falso: .63)
- Me gusta conocer gente de importancia porque eso me hace sentir importante. (Falso: .58)
- A veces murmuro o chismeo un poco de la gente. (Falso: .59)
- En las elecciones, algunas veces voto por candidatos acerca de quienes conozco muy poco. (Falso: .56)

- IPP: - Siempre sigo esta regla: primero el deber después el placer. (Cierto).
- Siempre trato de considerar el sentir de los demás antes de hacer algo. (Cierto).
 - Las cosas más importantes para mí son mis obligaciones para con mis compañeros y mi trabajo. (Cierto).
 - Algunas veces aparento saber más de lo que realmente sé. (Falso).
 - Me gustan las conferencias sobre acontecimientos mundiales. (Cierto).
 - Algunas veces me gusta ir contra las reglas y hacer cosas que nunca imaginé. (Falso).

Pudiera pensarse que la forma de responder a estos reactivos reflejara solamente una manera de distorsionar el inventario, actitud que pudiera reflejarse también en las demás escalas de ambas pruebas; pero no puede descartarse la alternativa de que la forma de responder a los reactivos refleje también el nivel de ajuste de los sujetos a los valores de la cultura a que pertenecen.

Como un comentario en apoyo de esta explicación, podemos citar un estudio realizado por Gough en 1952 ("On making a good impression", Journal of Educational Research, XLVI, 1952) (7) en donde se pidió a un grupo de sujetos que respondieran a 115 ítems escritos en primer lugar con las instrucciones habituales, y en segunda con instrucciones de "crear una impresión excepcionalmente favorable en una persona importante que proporcione el mejor cuadro de usted". Cuarenta de tales ítems sufrieron cambios significativos en la manera de contestarse con ambas versiones del instructivo, y fueron utilizados para estructurar la escala Bi. Un estudio posterior de Gough (29), indica que la escala Bi resulta altamente efectiva para identificar registros del I. P.-P. distorsionados deliberadamente en una dirección socialmente favorable, tal como ya se mencionó en el Capítulo II.

El segundo de tales aspectos se encuentra relacionado con las cargas factoriales de las escalas en el segundo factor (C_2). Puede apreciarse que tanto para la escala Sb como para la escala K, la mayor parte de su varianza puede explicarse en función de sus cargas en este segundo factor (.66 y .59 respectivamente); en tanto que para L (.32) y para Bi (.00) tales cargas quedan abatidas. Las escalas Cm y F mantienen correlaciones negativas que solo alcanzan un nivel de -.21 para Cm y -.16 para F.

Este factor parece indicar la capacidad del sujeto para mostrar a los demás conductas que son el resultado de un adecuado manejo de la economía psicológica. Parece incluir, asimismo, ciertos aspectos de la autoestima, y de la forma de manejar los recursos personales para entretarse a la realidad y pudiera designarse con el nombre de "manejo de los recursos psicológicos". Cuando hablamos de manejo de los recursos psicológicos, nos estamos refiriendo a la forma en que el sujeto produce conductas que le permiten adaptarse de manera adecuada al medio ambiente, conciliar los estímulos que recibe del exterior con sus impulsiones internas, de manera tal que la conducta general que puede ser apreciada por los demás se encuentre adecuadamente enmarcada dentro del contexto social en que el sujeto se mueve. Lo cual, dicho de otra manera, parece mostrar la capacidad del sujeto para hacer uso de sus recursos mentales y emocionales sin que tal manejo implique un elevado gasto psicológico.

Como ejemplo de esta situación mostramos algunos de los ítems de estas escalas que parecen reflejar con mayor claridad los aspectos arriba señalados:

- MMPI.- - Se necesita discutir mucho para convencer a la mayor parte de la gente de la verdad. (Falso: .57)
- Encuentro difícil entablar conversación con alguien que conozco por primera vez. (Falso: .82)
 - Cuando estoy en un grupo de gente, tengo dificultad pensando en las cosas apropiadas de qué hablar. (Falso: .82)
 - Creo que casi todo el mundo mentiría para evitarse problemas. (Falso: .52)
 - La gente me desilusiona con frecuencia. (Falso: .79)
 - Algunas veces he sentido que las dificultades se acumulaban de tal modo que no podía vencerlas. (Falso: .74)
- IPP.- - Soy tan sensible acerca de algunos asuntos que ni siquiera puedo hablar de ellos. (Falso)
- Cuando me siento muy feliz y activo, alguien que esté deprimido me desanima por completo. (Falso)

- Hay ciertas personas que me disgustan tanto que me alegro interiormente cuando están pagando las consecuencias por algo que hanhecho. (Falso)
- Tengo uno o varios malos hábitos tan arraigados que me es inútil luchar contra ellos. (Falso)
- A menudo me he sentido culpable porque he fingido mayor pesar del -- que realmente sentía. (Falso)
- No parece importarme lo que me pase. (Falso)

Independientemente de lo que hemos dicho aquí al respecto de la escala K, no debe perderse de vista el hecho de que esta escala fue estructurada con el objeto de identificar los casos de perfiles falsos-negativos. De los 30 ítems que integran la escala, 24 de ellos tienen altas correlaciones con el factor de Deseabilidad Social de Edwards (Carson, 1969*) (32). La escala, en su forma actual, es el producto de un gran número de esfuerzos de los autores para desarrollar una medida de las actitudes defensivas al tomar la prueba, y en este sentido K es utilizada como una variable supresora. La escala K se aproxima bastante a lo que mide la escala L; por lo cual tampoco debemos perder de vista que, aunque estos comentarios se han enfocado fundamentalmente hacia dos factores diferentes en los que tanto L como K tienen cargas diferentes, ambas escalas correlacionan entre sí .56 y K correlaciona .51 con el primero de los factores aquí descritos; lo cual quiere decir que si bien K muestra una alta saturación en el segundo factor, no por ello está libre de contaminaciones en el primero de ellos, al que aquí hemos denominado como "deseabilidad social".

Por último, pasaremos a referirnos a la posición que guardan las escalas F y Cm con respecto a los factores encontrados.

La escala F muestra cargas negativas tanto para el primero (-.331) como para el segundo (-.155) de estos factores; parece lógico pensar que estas cargas negativas obedecen fundamentalmente a que la escala fue estructurada atendiendo exclusivamente a la proporción de respuesta de los reactivos, sin atender a su contenido. Teóricamente los ítems de esta escala fueron incluidos cuando su proporción de respuesta en la dirección crítica no excediera de .10; consecuentemente la escala incluye ítems cuyo contenido queda francamente fuera de lo que pudiéramos llamar características normales de la población.

De ahí que resulte difícil tratar de definir un conjunto de características psicológicas unitarias a partir de una escala que fue estructurada con reactivos que poseen contenidos tan heterogéneos.

* Robert C. Carson: Interpretative Manual to the MMPI.

Adicionalmente cabe señalar que 13 de los reactivos de esta escala no funcionan adecuadamente para la población de la que fue extraída la muestra para este trabajo, inclusive hay alguno cuya proporción excede a .50, cuestión que probablemente determina la elevación del promedio de respuestas en F, de T50 a T55 para los ítems contestados en dirección crítica.

Los ítems cuya proporción de respuesta en dirección crítica funciona inadecuadamente en relación con la muestra original son:

- 115.- Creo en otra vida después de esta. (Falso: .552)
- 56.- Cuando muchacho (a) me suspendieron de la escuela una o más veces - por hacer travesuras. (Cierto: .305)
- 215.- He bebido alcohol con exceso. (Cierto: .230)
- 218.- No me molesta mucho el ver sufrir a los animales. (Cierto: .229)
- 252.- A nadie le importa mucho lo que le suceda a usted. (Cierto: .170)
- 258.- Creo que hay un Dios. (Falso: .149)
- 156.- He tenido épocas durante las cuales he hecho cosas que luego no he recordado haber hecho. (Cierto: .146)
- 272.- A veces estoy lleno de energía. (Falso: .146)
- 40.- La mayor parte del tiempo preferiría soñar despierto antes que hacer cualquier otra cosa. (Cierto: .121)
- 65.- Yo quise a mi padre. (Falso: .119)
- 199.- Se debe enseñar a los niños la información básica sobre la vida sexual. (Falso: .113)
- 185.- Aparentemente oigo tan bien como la mayoría de las personas. (Falso: .112)
- 54.- Le agrado a la mayor parte de la gente que me conoce. (Falso: .110)

Por otra parte encontramos que esta escala incluye varios de los reactivos cuya dirección crítica de respuesta, se encuentra entre las más bajas del inventario y que, en gran medida parecen reflejar alteraciones más o menos severas de los sujetos. De entre éstos pueden anotarse los siguientes:

- 23.- Sufro de ataques de náusea y de vómito. (Cierto: .005)
- 151.- Alguien ha estado tratando de envenenarme. (Cierto: .007)

246.- Con frecuencia me salen manchas rojas en el cuello. (Cierto: .009)

123.- Creo que me están siguiendo. (Cierto:.016)

184.- Frecuentemente oigo voces sin saber de donde vienen. (Cierto:.017)

202.- Creo que estoy condenado o que no tengo salvación. (Cierto: .017)

Por lo que se refiere a la escala Cm, aún cuando las correlaciones que mantiene con los dos factores encontrados son negativas (-.257 para C_1 y -.207 para C_2) cabe pensar que la escala refleja algunas características de los sujetos que toman el I. P. P. Entre los aspectos que parecen ser evaluados por la escala, puede pensarse en cierta capacidad del sujeto para tolerar la presión ejercida por la prueba, manteniendo una postura lógica al responder a los reactivos sin mostrarse aquiescente para con los lugares comunes que aparentemente pudieran ubicarlo en una situación más ventajosa ante quien evalúa los resultados de la prueba.

Como ejemplo podemos citar los siguientes reactivos:

- No me gusta prestar mis cosas a gente descuidada. (Cierto)
- La educación es más importante de lo que la mayoría de la gente piensa. (Cierto)
- Si una persona es lo suficientemente hábil como para estafar a alguien una gran cantidad de dinero, debe permitírsele quedarse con él. (Falso).
- Preferiría ser un trabajador confiable y estable a uno brillante pero inestable. (Cierto)
- Hay algunas cuantas personas en las que sencillamente no se puede confiar. (Cierto)
- Debo admitir que algunas veces la gente me desilusiona. (Cierto)

De acuerdo con Gough (29) los ítems de la escala Cm representan puntos modales de confiabilidad de los sujetos en relación al contexto social en que se desenvuelven, y el conjunto total de ítems vendría a ser una especie de "común denominador" de sus opiniones y actitudes.

7) COMPOSICION FACTORIAL DE LA VARIANZA

	L	F	K	Sb	Bi	Cm
CONFIABILIDAD	.920	.636	.922	.680	.910	.940
COMUNALIDAD	.578	.134	.606	.658	.589	.109
VARIANZA UNICA	.422	.866	.394	.342	.411	.891
VARIANZA ESPECIFICA	.342	.502	.316	.022	.499	.831
VARIANZA ERROR	.080	.364	.078	.320	.090	.060
TOTALIZACION DE FACTORIALIZACION	62.83	21.07	65.73	96.76	54.14	11.60

ASPECTOS DE LA COMPOSICION DE LA VARIANZA. (18, 24)

CONFIABILIDAD.- Mide el grado de certeza con que los ítems de la escala miden lo que miden, en función del número de ítems de cada escala y de la homogeneidad del grupo de ítems, es decir, el grado en que los ítems miden la misma variable; en este sentido, funciona como coeficiente de consistencia interna. Esta proporción de la varianza total, se encuentra formada por componentes verdaderos.

COMUNALIDAD.- Proporción de la varianza total que una escala tiene en común con las otras escalas de la matriz de correlación analizada. Es la parte de la varianza que determina la correlación entre una escala dada y las otras. Está determinada por factores que son comunes (en este caso, dos) a la escala en cuestión y a las demás, y se basa enteramente en componentes verdaderos.

VARIANZA UNICA.- Es la proporción de la varianza, única a la escala y que está formada por la varianza específica y la varianza de error.

VARIANZA ESPECIFICA.- Es la parte de la varianza verdadera que no aparece sistemáticamente en las otras escalas de la matriz de correlaciones entre una escala dada y las otras que se incluyen en tal matriz. No contribuye a la correlación entre las escalas.

VARIANZA ERROR.- Es la parte de la varianza que está compuesta por factores no correlacionados con las escalas incluidas en la matriz; se caracteriza por la aparición de factores fortuitos y es provocada por la sensibilidad del instrumento, por factores cuyo efecto varía de una ocasión a otra o por factores diferentes a aquéllos que determinan los puntajes verdaderos del sujeto.

TOTALIZACION DE FACTORIALIZACION.- Es un índice que mide el porcentaje de la confiabilidad de una variable explicada por factores comunes. Este índice es siempre menor que 100 y se aproxima a esta cantidad solo cuando la varianza

específica desaparece. Obviamente el análisis para determinar las cargas factoriales, no puede llevarse hasta un punto donde la varianza específica no se presente, cuando se trabaja con un grupo finito de variables.

COMPOSICION DE LA VARIANZA EN TERMINOS FACTORIALES:

$$\text{VARIANZA TOTAL } (1): h^2 + b^2 + e^2 = h^2 + d^2$$

$$\text{CONFIABILIDAD } (r_j): h^2 + b^2 = 1 - e^2$$

$$\text{COMUNALIDAD } (h^2): 1 - d^2$$

$$\text{VARIANZA UNICA } (d^2): b^2 + e^2 = 1 - h^2$$

$$\text{VARIANZA ESPECIFICA } (b^2): d^2 - e^2$$

$$\text{VARIANZA DE ERROR } (e^2): 1 - r_j$$

$$\text{TOTALIZACION DE FACTORIALIZACION } (C_j): 100 h^2 / (h^2 + b^2)$$



C O N C L U S I O N E S .

CONCLUSIONES

Al plantear la hipótesis sobre la que se fundamenta este trabajo se estableció el supuesto de que dada la configuración observada en las seis escalas que fueron analizadas, así como la existencia de ciertos niveles significativos de correlación entre ellas, mismos que aparecen sistemáticamente en la muestra utilizada, permiten establecer conceptualmente que el número de variables-subyacentes que expliquen la varianza conjunta de estas escalas debe ser menor que el número de escalas. Ahora bien, el análisis realizado ha permitido -- identificar dos factores comunes a estas escalas, a las cuales se ha denominado como "deseabilidad social" y "manejo de los recursos psicológicos"; desafortunadamente, y aún cuando se conocía este riesgo, el limitado número de variables incluido en el estudio ocasionó de hecho que dos de las variables, F y Cm, se reflejaran fuera del marco central de la matriz por lo que su interpretación resulta impracticable en el contexto de las otras cuatro variables. Si bien en los comentarios del Capítulo III se hacen algunas anotaciones al respecto de estas variables, esto no quiere decir que sean ubicables dentro del patrón configurativo de los vectores que representan a las escalas L, K, Sb y Bi.

La nominación de los dos factores encontrados se hizo en función del contenido de las escalas, tratando de sintetizar las variables psicológicas que están siendo representadas por el factor matemático; como ya se dijo antes, la realización del análisis factorial, no es garantía de certeza para denominar o interpretar el o los factores encontrados, aunque el empleo de otros procedimientos no escapa tampoco de esta dificultad, y aún pudieran ser más deficientes si se considera que el análisis factorial se encuentra apoyado en la evidencia empírica de los datos para demostrar que los resultados son producto de algo que no es azar y que si los datos mantienen estabilidad interindividualmente, merecen confianza en su utilidad potencial.

Esto nos lleva, necesariamente, a hacer algunos comentarios al respecto de los supuestos que fueron planteados al inicio del Capítulo III.

El primero de tales supuestos se refería a que la configuración de las seis escalas estudiadas mantenía un patrón configurativo estable; los datos obtenidos de la muestra empleada indican que esta suposición queda comprobada por los cálculos realizados, tal como se muestra en los gráficos iniciales del Capítulo III; esto no quiere decir que las escalas arrojen distribuciones que sean idénticas entre sí, ya que como puede apreciarse, aunque el patrón configurativo se conserva, las escalas L, K y Bi muestran una mayor dispersión en los datos (platicúrticas al .01, .05 y .05 respectivamente) que las escalas Cm (mesocúrtica) y las escalas F y Sb que son leptocúrticas al .01. Asimismo se aprecia que las escalas K, Sb y Cm son asimétricas izquierdas, las dos primeras al .01 y la última al .05; las escalas L y Bi muestran una distribución simétrica y la escala F asimétrica derecha al .01; datos que son todos coherentes con las premisas planteadas en el cuerpo del trabajo.

En el inciso referido a las diferencias entre medias, se ve que la mayor parte de las medias difieren significativamente entre sí, por lo que se refiere al promedio del puntaje T alcanzado por los sujetos en las escalas; excepto para los pares L-Bi, K-Bi y L-K, donde las escalas alcanzan elevaciones similares entre sí.

El segundo de tales supuestos en relación a la hipótesis de trabajo, se refería a la existencia de intercorrelaciones significativamente mayores de cero para las escalas estudiadas. Todas las correlaciones encontradas fueron lo suficientemente significativas como para permitir rechazar la H_0 , siendo la media de la distribución de correlaciones de 0.374.

Una vez comprobadas las premisas de apoyo para la realización del trabajo, se procedió a analizar la varianza conjunta de las seis escalas con lo cual se concluyó que la razón F para cada una de las fuentes de variación es, en todos los casos, significativamente mayor de 1, por lo que puede afirmarse que la variación de cada condición, en relación a la varianza intraescalas, no es producto del azar y, como aspecto no menos importante, permitió determinar la varianza de error por un procedimiento relativamente sencillo.

A partir de los datos de la varianza de error, así como los de la comunalidad (aportados por el análisis factorial) pudo llegarse al cuadro completo de la composición factorial de la varianza que aparece al final del tercer capítulo; mismo que presenta un análisis de la varianza que no puede lograrse cuando se utiliza sólo uno de ambos procedimientos.

Si se observan los datos de confiabilidad del cuadro de composición factorial de la varianza, éstos pudieran parecer espuriamente altos, por lo que cabe aquí señalar que los coeficientes reportados deben ser vistos como coeficientes de consistencia interna y no como coeficientes de equivalencia, en donde la varianza específica es tratada como varianza de error por no contribuir a la correlación entre las escalas. Como coeficiente de equivalencia, la confiabilidad correspondería a la comunalidad reportada en el mismo cuadro.

Con respecto a la validez de las escalas, es posible derivar a partir de los datos obtenidos de este estudio los coeficientes de validez de construcción, ya que, aunque no se incluye en este trabajo ningún criterio externo que permita establecer la validez predictiva o concurrente, los coeficientes de validez de construcción pueden derivarse de los resultados del análisis factorial, mismo que puede detectar el grado en que las escalas miden factores comunes.

Los coeficientes de validez de construcción cuando la confiabilidad es tomada como coeficiente de consistencia interna son los siguientes:

L	=	.636
F	=	.274
K	=	.745
Sb	=	.659
Bi	=	.518
Cm	=	.318

Si estos coeficientes de validez se apoyan en la confiabilidad como coeficiente de equivalencia, entonces la validez de construcción (factorial) de las escalas es:

L	=	.544
F	=	.126
K	=	.604
Sb	=	.649
Bi	=	.417
Cm	=	.108

Tales coeficientes fueron calculados como la suma del cuadrado de la mitad de la confiabilidad por la carga factorial de cada una de las escalas en los dos factores encontrados.

Independientemente de los coeficientes reportados arriba, cuando se habla de escalas de validez para el MMPI y para el IPP, se hace un paréntesis en relación a los tipos de validez formulados por la American Psychological Association en su "Technical recommendations for psychological tests and diagnostic techniques" (26), para referirnos específicamente a la aceptabilidad de los perfiles obtenidos de la administración de ambos inventarios; o sea que las escalas funcionan como indicadores de la utilidad de la prueba a partir del resultado de tales escalas y para determinar la aceptabilidad de un conjunto de datos para un sujeto particular, por medio del inventario.

Con respecto al análisis factorial de las escalas por medio de la solución centroide debe hacerse un último comentario:

Esta solución ha sido una de las más populares en el análisis factorial, aunque con frecuencia ha sido equívocamente comprendida y utilizada. El método centroide de análisis intenta obtener una solución del mismo tipo que la proporcionada por el método de factor principal desarrollado por Pearson a

principios de siglo y estuvo muy en boga durante las décadas de los treinta y los cuarenta y abrevia un gran número de complicaciones de cálculo y de labor de procedimiento; intenta asimismo evaluar, tanto como sea posible, la varianza con cada factor sucesivo; sin embargo la solución centroide no es de ninguna manera la única posible para un conjunto de variables y no proporciona algunas de las propiedades matemáticas de la solución por factor principal. Thurstone (1947) (18) señala que "el método centroide de factorialización y la solución centroide para la ubicación de los ejes de referencia han de ser vistos como una alternativa de cálculo, de la que se ha encontrado que involucra una labor mucho más sencilla que la solución de factor principal". Por supuesto, en la actualidad existen un gran número de programas que permiten obtener una solución de factor principal en cuestión de minutos por medio de una computadora. Cuando no se encuentra con medios de este tipo, el método centroide resulta bastante adecuado y efectivo en la solución preliminar de problemas sencillos.

Cuando una matriz está integrada exclusivamente por correlaciones positivas, esta solución proporciona un factor general y los factores restantes generalmente resultan ser bipolares. Una solución centroide puede ser obtenida también de una matriz que contenga correlaciones negativas, pero, como se vio en nuestro caso, el procedimiento se vuelve ligeramente más complicado.

Finalmente, aún cuando este trabajo se ha limitado a reportar el planteamiento inicial del funcionamiento de las escalas estudiadas, cabe señalar que la labor de análisis de las mismas se continúa; encontrándose a la fecha en preparación los datos correspondientes a las proporciones de respuesta de los reactivos del IPP para toda la población de la que se extrajo la presente muestra. Posteriormente se pretende establecer las correlaciones biserials entre los reactivos, a fin de estructurar, a partir de las escalas L, K, Sb y Bi, dos escalas factoriales para evaluar los factores a los que aquí se ha denominado tentativamente como "deseabilidad social" y "manejo de los recursos psicológicos". Se espera, asimismo, que al eliminar los reactivos cuya varianza no contribuye a la correlación entre las escalas, existan mayores posibilidades de corroborar los supuestos conceptuales de nominación e interpretación de los factores encontrados y que la información adicional permita delimitarlos en un contexto más preciso.

B I B L I O G R A F I A

1. - ANASTASI, ANNE.
Tests Psicológicos.
Aguilar, S. A. de Ediciones.
Madrid, 1966.
2. - ARKIN, H. y COLTON, R.
Statistical Methods.
Barnes & Noble, Inc.
New York, 1964.
3. - ARKIN, H. y COLTON, R.
Tables for Statisticians.
Barnes & Noble, Inc.
New York, 1963.
4. - BEST, J. W.
Research in Education.
Prentice - Hall, Inc.
New Jersey, 1959.
5. - BUROS, O. K.
The Sixth Mental Measurements Yearbook.
The Gryphon Press.
Highland Park, N. J. 1965.
6. - BUTCHER, J. N.
MMPI: Research Developments and Clinical Applications.
Mc Graw - Hill Book Company.
New York, 1969.

7. - CATELL, R. B. y WARBURTON, F. W.
Objetivo, Personality & Motivation Tests.
University of Illinois Press.
Urbana, Ill., 1967.
8. - CRONBACH, L. J.
Fundamentos de la Exploración Psicológica.
Biblioteca Nueva.
Madrid, 1963.
9. - CRONBACH, L. J. Y GOLDINE C. G.
Psychological Tests and Personnel Decisions.
University of Illinois Press.
Urbana, Illinois, 1965.
10. - DAHLSTROM W. G. Y WELSH G. S.
An MMPI Handbook, a Guide To Use In Clinical.
Practice and Research.
University of Minnesota Press.
Minneapolis, 1962.
11. - DAHLSTROM, W. G.; WELSH, G. S. Y DAHLSTROM L. E.
An MMPI Handbook, Clinical Interpretation.
University of Minnesota Press.
Minneapolis, 1972.
12. - DUNNETE, M. D. y KIRCHNER, W. K.
Psicología Industrial.
Ed. F. Trillas, S. A.
México, 1972.
13. - FISCHER, H.
Estadística Aplicada a la Psicología.
Ed. Paidós.
Buenos Aires, 1968.

14. - FRUCHTER , B.
Introduction to Factor Analysis.
D. Van Nostrand Company, Inc.
Princeton , N. J. 1954.
15. - GOUGH , H. G.
California Psychological Inventory.
Consulting Psychologists Press , Inc.
Palo Alto , Calif. , 1969.
16. - GUILFORD , J. P.
The Nature of Human Inteligence.
McGraw - Hill Book Company.
New York , 1967.
17. - GUION , R. M.
Personnel Testing.
McGraw - Hill Book Company.
New York , 1965.
18. - HARMAN , H. H.
Modern Factor Analysis.
The University of Chicago Press.
Chicago , 1967.
19. - HATHAWAY , S. R. Y MCKINLEY.
Minnesota Multiphasic Personality Inventory.
The Psychological Corporation.
New York , 1967.
20. - HATHAWAY , S. R. y MONACHESI , E. D.
An Atlas of Juvenile MMPI Profiles.
The University of Minnesota Press.
Minneapolis , 1961.

21. - HAWK, R. H.
Reclutamiento y Selección de Personal.
Editora Técnica, S. A.
México, 1968.
22. - KIRK, R. E.
Experimental Design: Procedures for the Behavioral Sciences.
Brooks - Cole Publishing Company.
Belmont, Calif., 1968.
23. - LEHMAN, C.
Geometría Analítica.
U. T. H. E. A.
México, 1959.
24. - MAGNUSSON, D.
Teoría de los Tests.
Ed. F. Trillas, S. A.
México, 1969.
25. - MARKS, P. A. y SEEMAN, W.
An Actuarial Description of Abnormal Personality.
(An Atlas for Use with the MMPI).
The Williams & Wilkins Company.
Baltimore, Md., 1963.
26. - MEGARGEЕ, E. I.
Métrica de la Personalidad.
Ed. F. Trillas.
México, 1971.

27. - NUÑEZ, R.
Aplicación del Inventario Multifásico de la Personalidad (MMPI) a la Psicopatología.
El Manual Moderno, S.A.
México, 1968.
28. - PARDINAS, F.
Metodología y Técnicas de Investigación en Ciencias Sociales.
Siglo XXI Editores.
México, 1970.
29. - REUCHLIN, M.
Methodes D'Analyse Factorielle a L' Usage des Psychologues.
Presses Universitaires de France.
Paris, 1964.
30. - SNEDECOR, G. W. y COCHRAN, W. G.
Statistical Methods.
The Iowa University Press.
Ames, Iowa, 1969.
31. - STRAUSS, G. y SAYLES, L. R.
Personal. Los Problemas Humanos de la Dirección.
Herrero Hnos., Sucs., S. A.
México, 1964.
32. - TIFFIN, J. y McCORMICK, E. J.
Psicología Industrial.
Ed. Diana, S. A.
México, 1970.
33. - WELSH, G. S. y DAHLSTROM, W. G.
Basic Readings on the MMPI in Psychology and Medicine.
University of Minnesota Press.
Minneapolis, 1956.