

PUNTO DE VISTA EN EVALUACIÓN (QUE VAMOS A TOMAR)

1.- Montado sobre un modelo de psicología científica: se atiene al método científico: trabaja con hipótesis. La experiencia es considerada como aquella operación empírica que sirve para contrastar hipótesis científicas. Trabajamos hipotéticamente, con hipótesis de trabajo.

2.- La función principal de la evaluación psicológica no es reproducir exactamente la realidad: la evaluación psicológica describe para algo, no es la descripción en sí misma. Lo que le interesa a la evaluación psicológica es predecir la conducta y se ayuda de la descripción en la medida en que ésta le sirve para predecir. Interesa elaborar modelos predictivos de la conducta no solamente modelos descriptivos de la conducta.

3.- Psicodiagnóstico es distinto de tests psicológicos: la evaluación psicológica es más amplia que los tests psicológicos, al menos en dos sentidos:

- Longitudinalmente: la evaluación psicológica es esencialmente un proceso que parte del planteamiento de un problema y que termina en la recomendación de la toma de acción. En este proceso pueden entrar tests como elementos que ayudan a esa toma de decisión pero también pueden no entrar.
- Tenemos mucha más instrumentación que los tests, por ejemplo la entrevista, técnicas psicofisiológicas, cuestionarios, observación de la conducta, etc.

En base a los tests se desarrolló la teoría de tests que ha producido unos criterios de calidad de los instrumentos de evaluación y éstos sí que son comunes a todos los instrumentos de evaluación, los deben cumplir y eso se lo debemos a los tests.

4.- En evaluación psicológica se circunscribe el campo de trabajo a sujetos humanos: se abandona a los animales, se trabaja con humanos.

5.- La evaluación psicológica tiene dos orientaciones básicas: una teórica o de investigación y una práctica o de aplicación.

6.- El interés de la evaluación psicológica es fundamentalmente aplicado: se trata de conocer a la gente para ayudarla. El criterio de calidad fundamental dentro de la evaluación psicológica es la utilidad, ser útil significa ayudar a tomar decisiones correctas; en pro de la utilidad, no de un mero activismo. El concepto de utilidad implica también el concepto de economía; si una evaluación es muy costosa y da poca ganancia no es útil.

7.- La evaluación psicológica es acción efectiva sobre los demás: con la evaluación intervenimos a veces con mucha más fuerza que con lo que se llama intervención o tratamientos. Crea problemas deontológicos y legales porque los evaluadores se olvidaron de los problemas deontológicos y legales durante décadas, ahora ya no. Esto hace que la comunicación de los resultados de la evaluación sea tan importante (el informe psicológico).

8.- La evaluación psicológica es siempre imperfecta: toda formulación científica es siempre imperfecta, pero a la vez es perceptible, se puede mejorar. Por eso el diagnóstico psicológico es algo continuo, se hace día a día, nunca está acabado, hacemos cortes en el momento en que el diagnóstico se considera útil.

Hay que ser firmes en los planteamientos porque están científicamente contrastados pero también hay que ser modestos porque sabemos que nuestro diagnóstico es imperfecto. Teniendo en cuenta que ayudamos a tomar decisiones pero pocas veces somos quienes tomamos las decisiones. Al tomar decisiones hay que recordar que

los datos son sólo una perspectiva del problema pero también hay otras: política, ética, religiosa, etc.; los problemas no son únicamente psicológicos, son humanos y por tanto tienen muchas perspectivas.

9.- Esta disciplina es de un primer ciclo de estudios: no es una disciplina especializada, es general.

10.- Esta disciplina tiene prerequisites de conocimiento: estadística y muy buen manejo del concepto de correlación, materias psicométricas (escalas de medición).

Tema 1: HISTORIA

1.- Interés primario en la clasificación de personas.

2.- Interés primario en la clasificación de personas.

3.- Interés primario en la predicción de la conducta o de acontecimientos.

No son fenómenos distintos.

1.- Interés primario en la clasificación de personas:

Ha sobresalido. El esfuerzo por clasificar los sujetos según su temperamento pero fuertemente arraigados en variables de tipo biológico.

Escuela hipocrática: primeros esfuerzos de la clasificación. Tiene su fruto más maduro en Teofrasto: discípulo de Aristóteles: libro sobre los caracteres: está mejor dibujada esta escuela temperamentalista.

La figura más importante es la del médico latino Galeno: siglo I-II d.C., formula la teoría de los temperamentos postulando los 4 temperamentos: sanguíneo, flemático, melancólico y colérico; según dominara uno de los cuatro humores. La clasificación de los temperamentos sigue persistiendo, Eysenck fundamentó la teoría de la personalidad en los temperamentos de Galeno, aunque ahora se denominen de distinta manera. Eysenck ha revitalizado esta tradición galénica.

Platón: La República: divide a los hombres en tres clases:

- Los que domina la cabeza: filósofos destinados a gobernar.
- Los que domina el tórax: el corazón como valentía: los soldados.
- Los que domina el abdomen: destinados a servir a los filósofos y a los soldados.

Esta división no tuvo tanta influencia como la de Galeno.

Juan Huarte de San Juan: Examen de los ingenios: ; se adelantó a su época, tenía clara la idea de que cada sujeto tenía unos ingenios en los cuales sobresalía y según los ingenios se debía orientar al sujeto en sus actividades futuras. Fue una excepción en su época y durante siglos pasó olvidado, hasta casi el siglo XX.

2.- Interés primario en la selección de personas:

- Biblia: libro de los jueces.
- Mitología griega: Edipo y la Esfinge.
- Edad Media: juicios de Dios: se recurría a él cuando la justicia era incapaz de dar la razón a una u otra de las partes.
- Caza de brujas: para dirimir si una mujer era bruja o no.

El ejemplo más notable de la prehistoria de la selección esta en el Imperio Chino, ya a partir del segundo milenio a.C. se ven atisbos de construir procedimientos de selección para elegir a los funcionarios del imperio. Llegó a su madurez en los tiempos de Cristo. En 1905 cuando Binet crea el primer test occidental se abole en China este sistema.

3.- Interés primario en la predicción de la conducta o de acontecimientos:

- Oráculos: había personas especializadas. Se le daba el oráculo a cambio de fuertes cantidades de dinero siempre un lenguaje ambiguo.
- Sueños: se interpretaban con fines predictivos.
- Horóscopo: basado en la astrología (pseudociencia).
- Quiromancia: lectura de manos como predictora.

INFLUENCIAS HISTÓRICAS MÁS INMEDIATAS (S. XIX):

Teoría evolucionista: Darwin:

Influye porque en ella se habla de la supervivencia del más fuerte e intentaron estudiar las diferencias individuales y esto lo hacían con fines eugenésicos, de mejorar la raza. Estimuló a la evaluación psicológica a estudiar diferencias individuales y ese es el punto clave en el nacimiento de la evaluación psicológica.

Asociacionismo filosófico:

Plantea que todo nuestro psiquismo se construye en función de unos átomos psíquicos (sensaciones). En la evaluación psicológica influyó por dos razones:

- Dio un campo de estudio a la evaluación psicológica: las sensaciones, los elementos simples del psiquismo.
- Dio la idea de que podía construir lo complejo como una combinación de lo simple.

El ítem es el átomo de la evaluación, y la combinación de ítems constituye el test: atomismo constructivista presente en los tests actuales.

Matemáticas: estadística:

Quetelet: aplicó conocimientos estadísticos en variables humanas, aunque no en variables psicológicas, más bien de tipo sociológico. Descubrió cómo la curva normal se podía aplicar a un grupo heterogéneo y significativo de variables.

Desarrollo de la física:

Weber, Fechner,...., influyó de distintas formas:

- Significó la cuantificación de afectos psíquicos, sensaciones.
- Pone en relación la experiencia subjetiva y los estímulos objetivos; con esta relación la psicofísica prefigura la relación test – respuesta al test.
- Desarrolla una serie de métodos de medición que han inspirado métodos de medición propiamente evaluativa.

Psicología experimental naciente:

Es un fruto de la segunda mitad del XIX, influencia ambigua: por un lado positivamente y por otro

negativamente:

- Positivamente: se dedica al estudio de fenómenos sensomotrices que van a ser el caldo de cultivo de los primeros tests.
- Tiende a inhibir su desarrollo porque está decididamente dedicada a estudiar las leyes generales del psiquismo y en cambio la evaluación psicológica se va a desarrollar en la medida en que se potencie el estudio de las diferencias individuales.

Desarrollo de la educación y de la psicología educativa:

- Porque comienza un interés por fomentar la igualdad de oportunidades: que puedan sobresalir los que más valen, para conocer cuáles son hay que estudiar sus características: evaluación psicoeducativa de los alumnos.
- Porque en el siglo XIX comienza la preocupación por la evaluación escolar que es similar a la evaluación psicológica.

PSIQUIATRIA: Nace en el siglo XIX. Hace esfuerzos por encontrar un sistema clasificatorio objetivo de las enfermedades mentales así como para su pronóstico.

Pinel: psiquiatra de la revolución francesa. Distinguió la deficiencia mental de la enfermedad mental.

Esquirol: aporte taxonómico. Distinción entre la ilusión y la alucinación (fenómenos normales o neuróticos y fenómenos psicóticos).

Kraepelin: psiquiatra alemán, que trabajó en las últimas décadas del siglo XIX y principios del XX. Propone una taxonomía con gran repercusión en la psiquiatría. Abogó por el examen mental: historiales clínicos.

Gall (frenología): proponía un isomorfismo entre la superficie del cerebro y la craneal. Tuvo aportes en la evaluación psicológica:

- Fomentó el estudio de las localizaciones de las funciones cerebrales.
- Aporte a la evaluación psicológica por ser defensores de las diferencias individuales justificando las funciones mentales.
- Esfuerzo por construir instrumentos de evaluación.

Galton, J.M. Cattell y Binet hicieron posible la aparición de la disciplina de evaluación.

Galton: interesado por la teoría evolucionista de Darwin. Objetivo: mejorar el potencial humano (interés eugenésico). Estudio de las diferencias individuales:

- De tipo biológico (antropométrico).
- De tipo psicológico: sensaciones y motricidad.

Prefirió estudiar grandes grupos de población llevándole a una formulación estadística de la información (correlación).

J.M. Cattell: interés por la psicología: doctorado junto a Wundt. Se interesó por las diferencias individuales oponiéndose a sus compañeros: diferencias individuales en tiempos de reacción. Ya en Estados Unidos se refuerza en la diferencia individual: creación de instrumentos tipificados de medición aplicándolos a estudiantes (tests mentales). Expansión e institucionalización de los intereses diagnósticos y evaluativos. En 1892 la Asociación Psicológica Americana creó divisiones de medición mental (psicológica). Creó la primera casa constructora y distribuidora de test: Corporación Psicológica.

Binet: hace de su programa de psicología una psicología diferencial; él la denomina individual. Con ciertas características peculiares:

- Le interesa el estudio de las funciones superiores.
- Estudiarlas en un medio natural, no en laboratorio.
- La psicología diferencial debe responder a requisitos prácticos, exigidos por la sociedad y no solamente de forma teórica tanto para la selección, predicción, etc.

Binet construye con Simon un test dedicado a la praxis profesional: test de inteligencia (1905).

ETAPAS HISTÓRICAS DE LA EVALUACIÓN PSICOLÓGICA:

Algunos autores por etapas cronológicas y otros por etapas lógicas que pueden coincidir en el tiempo.

ETAPAS LÓGICAS:

Propuestas por el profesor **Pelechano** , según él existen tres aproximaciones desde un punto de vista lógico a la evaluación psicológica:

- Aproximación constitucionalista: aquella que toma al concepto de constitución o al concepto de temperamento como el concepto fundamental para constituir una tipología temperamental. Las categorías que sirven para distinguir a los sujetos tienen una base biológica, en la biología hay que buscar el marco estructural que integra a la información que se pueda recoger.
- Acercamiento funcionalista: tipologías de tipo funcional. Son las funciones psicológicas (memoria, pensamiento...) las que constituyen el marco integrador de la información, son categorías comportamentales y funcionales las que se usan como ejes de referencia. Es una aproximación psicologista.
- Acercamiento situacionista: tipologías situacionales. Lo que delimita las categorías van a ser variables estimulares o contextuales, esos van a ser los ejes de referencia. De ahí la importancia por ejemplo de escenarios de conducta. Viene de la mano de aproximaciones de tipo conductistas.

La cuarta etapa la agrega el profesor:

- Aproximación interaccionista: Endler, Magnusson. Lo esencial en psicología está en el concepto de interacción entre estímulos y respuestas. Categorías interaccionistas que ordenaran el comportamiento. Puente entre aproximaciones endógenas (constitucionalista y funcionalista) y exógenas (situacionista).

No hay sucesión cronológica, lo que se ha dado es un fenómeno de aproximación: primero aparece la aproximación constitucionalista, luego sin desaparecer ésta aparece la aproximación funcionalista, sin desaparecer las dos primeras aparece la aproximación situacionista y finalmente sin desaparecer ninguna de las tres ha aparecido la aproximación interaccionista. En la actualidad coexisten las cuatro aproximaciones.

ETAPAS CRONOLÓGICAS:

Ha habido dos intentos. Uno de parte de Thorndike y Hagen y el otro de parte de Drenth.

Thorndike y Hagen: según estos autores habría cuatro etapas, divididas en periodos de 15 años entre 1900 y 1960:

- 1900 – 1915: búsqueda y desarrollos iniciales de los métodos de medición en psicología y educación; lugar privilegiado entre estos métodos es el test de Binet – Simon de 1905.

- 1915 – 1930: auge en el desarrollo de los tests. Epoca Dorada.
- 1930 – 1945: delatan contradicciones como por ejemplo la que existe entre una construcción, evaluación e interpretación de tests muy refinada y por otro lado una práctica más bien descuidada donde dominan también las técnicas proyectivas (que tienen un estudio y desarrollo psicométrico nulo).
- 1945–1960: etapa más compleja. Etapa de crítica sociales a los tests, por otro lado los tests continúan desarrollándose. Desarrollo de baterías de tests y de programas de valuación. Baterías de tests: se entienden dos cosas:

- Un test que tiene varios subtests independientes entre sí.
- Conjunto de tests distintos.

Las baterías múltiples se desarrollan de la mano del análisis factorial y abren la posibilidad de dar respuestas a tareas diagnósticas más específicas por ejemplo de inteligencia mecánica, de aptitud artística, etc.

A partir de los ´60 lo consideran actualidad al igual que Drenth.

Drenth: psicólogo holandés que distingue:

- Antes de la creación del test de Binet – Simon: prehistoria del psicodiagnóstico.
- Entre la aparición del test de Binet y Simon y la Primera Guerra Mundial: la primera y segunda Guerra Mundial han sido muy importantes en el desarrollo de la evaluación psicológica. Características de esta etapa:
 - Aparecen los primeros intereses por definir, por medir y por estructurar la inteligencia. **Spearman**: a comienzos de siglo plantea su teoría de la inteligencia, es el que produce la teoría de mayor calidad; papel fundamental en la historia de la teoría de la inteligencia, por lo menos por haber planteado la existencia del factor G (factor general de inteligencia).
 - Gran expansión del test de **Binet – Simon**: se encontró con una sociedad sensibilizada a su aparición y en pocos años había más de 200 publicaciones sobre el test de Binet. La adaptación norteamericana del test fue muy importante, la llevó a cabo **Terman** en la universidad de Stanford, no fue una simple adaptación sino que fue con perfeccionamiento del instrumento. Luego la historia de este test sigue por esta adaptación, no del original y se le denomina: test de Stanford – Binet o el test de Terman (esta definición no es buena porque olvida a sus creadores, sobre todo a Binet).
 - Evaluación individual fundamentalmente: se evalúan individuos aislados.
- Primera Guerra Mundial – Segunda Guerra Mundial: se caracteriza por:
 - Aparecen y se expanden las pruebas de aplicación colectiva. Una comisión desarrolló dos tests para el ejercicio: army y army , el primero está hecho para sujetos que dominan el idioma inglés escrito, el segundo para los que no lo dominan (analfabetos, extranjeros...).
 - Se desarrollan tanto la evaluación psicométrica como la evaluación proyectiva de la personalidad.
 - Se formula una teoría de tests: la teoría de los tests psicológicos.
 - Etapa de gran expansión y optimismo psicométrico, es un optimismo un poco ingenuo.
- Segunda Guerra Mundial hasta la actualidad (´60 – ´70): características:
 - Continúa la expansión del material diagnóstico con un interés creciente por aspectos más específicos de la conducta.
 - Etapa crítica: hay una reflexión sobre los fundamentos metodológicos de la evaluación psicológica y hay una fuerte crítica social: revuelta antitests: movimiento contra los tests por parte de la sociedad y de la prensa en Estados Unidos, fundamentalmente aplicado al campo educativo. Llegaron incluso a quemar literalmente tests en una hoguera. Había parte de razón en su crítica que la ha asimilado la

evaluación psicológica para progresar.

- Interés creciente por las estrategias y los problemas de decisión diagnóstica.

A finales de los 60 y principios de los 70 la evaluación psicológica se caracteriza por fuertes contradicciones o disociaciones: por ejemplo **Cronbach** en 1957 en su trabajo sobre las dos disciplinas de la psicología científica (la psicología experimental o manipulativa y la psicología correlacional o descriptiva) delata que el psicodiagnóstico se ha desarrollado exclusivamente dentro de la psicología correlacional, no existe en esa época una evaluación experimental.

Anastasi: ha delatado también la disociación que existe entre la teoría de tests y la teoría psicológica; se van alejando los tecnólogos o metodólogos y los psicólogos y esto es muy nocivo para la evaluación psicológica. Delata también una disociación entre la teoría de tests y el uso de los tests: la primera ha alcanzado un alto grado de satisfacción, el segundo sigue todavía a un nivel muy primitivo.

Algunos autores defienden respuestas a tests con características personales, aunque se haya demostrado que no es así. A este tipo de relaciones se las llama correlaciones insólitas, y el psicodiagnóstico está lleno de ellas (correlaciones ilusorias es por ejemplo cuando en una lámina del test de Rochar un sujeto ve flores y el evaluador dice que el sujeto es gay: se ha demostrado que no es así, pero no se hace caso y se sigue convencido de que es verdad y se sigue diciendo).

La década de los 50 ha sido la época que ha marcado el desarrollo posterior de la psicología de la evaluación, tanto positivo como negativo (críticas).

Teoría de los tests mentales (libro) que cerró una época dorada.

Año 50 Sobin.

Años 52 y 53, Eysenck, publica unos trabajos muy críticos en dos vertientes:

- Baja fiabilidad y validez de los diagnósticos psicopatológicos taxonomía psiquiátrica usada.
- Evaluación al éxito de las terapias en uso (generó un movimiento: eficacia de las psicoterapias).

Meehl (54) publica Predicción clínica versus estadística: enfrenta dos maneras de hacer predicción.

Cronbach y Meehl (55) Validez de constructo en los tests psicológicos delatan una deficiencia por estudiar las interpretaciones psicológicas de los tests (qué miden los tests: validez conceptual).

Polémica:

- Posiciones científicas

Evaluación psicológica

- Posiciones no científicas

Crisis dentro del diagnóstico científico.

Técnicas proyectivas: Rorschach:

A partir de los 40 se criticó a las técnicas proyectivas, por no tener en cuenta los criterios científicos de calidad.

(65) Zubin, Eron y Schumer, revisan 500 trabajos sobre la calidad métrica de la técnica proyectiva: las técnicas proyectivas no dan nada más de lo que aporta una buena entrevista.

(Meehl) el método clínico basa su diagnóstico en la experiencia y en la intuición del experimentador o examinador. Se basa en una fórmula matemática derivada de estudios empíricos.

Polémica entre lo idiográfico y lo nomotético:

Aproximación idiográfica. Se introdujo en las ciencias sociales por Windelband a comienzos de siglo. A mediados se introdujo en psicología por medio de Allport. Para Allport es aquella que da cuenta de la unicidad del individuo, que lo distingue de todos los demás lo que es lo más propio, esencial y único de él.

Nomotético: se trabaja nomotéticamente cuando el individuo es visto como un perfil de variables que son universales, el perfil es único pero las variables con las que se describe son universales para todos los individuos.

A mediados de siglo hubo una gran polémica entre las dos posturas. El enfoque idiográfico se hizo extremista y esa unicidad del individuo era tal que no se podía ni siquiera describir y con esa postura no se puede avanzar en la ciencia. El enfoque nomotético avanzaba a pasos agigantados; terminó ganando y los mismos idiógrafos cuando tenían que trabajar con sujetos trabajaban nomotéticamente. Siempre quedó una nostalgia por estudiar lo propio del individuo; esto ha sido posible con una definición menos estricta de lo idiográfico: estudia el individuo según las variables que extrae del mismo individuo, entonces sí podemos trabajar.

Queda claro que las dos posturas no son incompatibles: podemos trabajar con un individuo idiográficamente por ejemplo con una entrevista en profundidad y luego podemos aplicarle tests para ver la posición de ese individuo en ciertas variables universales (cuál es su posición en neuroticismo, extraversión, etc.).

Hoy por hoy toma nuevamente fuerza el enfoque idiográfico pero con una definición menos estricta de la de Allport y seguidores y también se sigue trabajando nomotéticamente.

(Fin de la polémica entre posiciones científicas y no).

El problema también se extiende al diagnóstico científico mismo:

En los años 50 Cronbach y Meehl habían protestado pro los conceptos psicológicos que usaban en los tests. Su trabajo nos introduce la validez teórica de los instrumentos de medición.

Ghiselli: haciendo trabajos de revisión sobre la validez predictiva de los instrumentos psicológicos concluyó que su poder predictivo era en la mayoría de las veces entre bajo y muy bajo.

¿Cuáles son las razones de que la predicción haya sido tan deficiente? :

- Dicotomización o separación creciente entre el test y el criterio: entendemos por test cualquier instrumento de predicción; por criterio entendemos toda variable que deba ser predicha. Los psicólogos de los 30 a los 50 no aprendieron de Binet: Binet para construir su test pasó mucho tiempo estudiando la situación de criterio, de rendimiento académico y de ahí extraía sus ítems para su test. Los psicólogos empezaron a construir tests que no tenían que ver con las situaciones de criterio y lógicamente el valor predictivo de los tests fue bajando. Los psicólogos se dieron cuenta, y de los años 60 en adelante hay un movimiento de regreso a los criterios, de vuelta al estudio de los criterios, de esas variables que deben ser predichas. Potenciación de la observación natural de conductas: siempre la observación ha estado presente en la evaluación psicológica pero en un segundo plano, pero ahora vuelve a estar en primer plano.
- carácter indiferenciado de la predicción tan típica de la época clásica (30-50-60): la predicción se hacía

típicamente para grandes grupos sin diferenciar entre grupos donde podría haber una predicción así mismo diferencial. Se descubrieron en los 50 las variables moduladoras que han echado abajo ese carácter indiferenciado de la predicción. Una variable moduladora o moderadora es una variable que no predice al criterio pero que sirve para diferenciar, o sea para modular, la predicción de otras variables sobre ese criterio. Ejemplo: si nuestra tarea es predecir rendimiento académico y tenemos una prueba de inteligencia general y otra de motivación para el estudio y la predicción es mediocre (alrededor de 0.30); en un momento suponemos que el género de los sujetos puede modular la predicción del rendimiento académico, volvemos a aplicar las dos variables predictoras pero separadas según el sexo. En varones el cuestionario sobre motivación para el estudio no predice rendimiento académico pero el de inteligencia es un buen predictor del rendimiento académico. En mujeres la inteligencia es un nulo predictor de rendimiento académico pero la motivación es un muy buen predictor de rendimiento académico. El valor predictivo de la inteligencia general y de la motivación para estudiar es muy distinta según si se trata de un varón o una mujer. Se restringe el campo de la predicción, ahora tiene que tomar el sexo del sujeto. El diagnóstico se especifica pero la predicción sube.

- presupuestos infundados de la teoría clásica de los tests, más concretamente del concepto de fiabilidad: Spearman ha identificado precisión con consistencia interna y estabilidad temporal. Mientras menos consistente es el test menos fiable es. Mientras más estable temporalmente sea un test mejor es el test. Lo importante es que el test tenga aquella consistencia y estabilidad que tiene la conducta que mide ya que hay conductas poco consistentes, la conducta varía de situación a situación y de momento a momento y que sin embargo pueden ser conductas muy importantes. Han aparecido nuevas áreas de estudio como la de los tests situacionales, que toman en cuenta las distintas situaciones que están muestreando, son los tests S-R. Relativizan el concepto de consistencia comportamental. Evaluación de cambios de conducta: los tests seleccionaban sólo los ítems estables temporalmente, sin darse cuenta de la sensibilidad de los ítems para detectar cambios allí donde los haya, por ejemplo para evaluar los efectos de las terapias los tests eran muy malos, tuvieron que buscar otros conjuntos de ítems que fueran sensibles al cambio de conducta.
- Diagnóstico como descripción de productos: por mucho tiempo sólo ha habido una preocupación por la respuesta al test pero no por cómo era dada esa respuesta al test, interesaba el producto final y nada más: sólo nos acercaremos a una predicción de 1 cuando sepamos qué ocurre entre las respuestas de test y de criterio: ese qué ocurre puede ser conceptualmente interpretado: cuáles son los procesos comunes en la respuesta a test y al criterio que nos explican su correlación o su falta de ella, cuáles son los procesos que subyacen a las respuestas a los tests.

El diagnóstico científico se está perfeccionando en una serie de áreas como por ejemplo estas 4.

Terminología: el aspecto semántico de la evaluación psicológica.

Roscharch: publicó su test en forma de libro en 1921 **Psicodiagnóstico** y creó así esa expresión que aparece por primera vez. Esto hace que la expresión psicodiagnóstico se tiña de una serie de características; Roscharch es un médico y el libro responde al modelo médico: se diferencian síntomas (lo que aparece) y la enfermedad (lo que está detrás de lo que aparece: lo importante). Se le llama también enfoque de signos porque los síntomas actúan como signos que hay detrás o puede haber detrás:

- Un rasgo psicológico: una respuesta a un test puede estar explicada por un rasgo por ejemplo la agresividad.
- La dinámica inconsciente: aplicación preferida por los psicólogos psicodinámicos o psicoanalíticos, explican la respuesta a los tests en función de conflictos inconscientes.
- Una base biológica.

No son incompatibles entre sí al menos completamente.

Psicodiagnóstico, se centra en lo negativo, en lo patológico. Esta significación muy clínica del término psicodiagnóstico fue la que pasó a Estados Unidos junto con el test de Roscharch (años 30). En Europa el

término se ha liberado de estas connotaciones y tiene un espectro muy amplio como es el de evaluación psicológica sin más.

Hay una segunda expresión más antigua que la de psicodiagnóstico; es la expresión de **Psychological testing**: no tiene una traducción exacta en Español, es más o menos examen psicológico en base a pruebas es de raíz americana (Cattell), las características están en que se centra en la construcción, aplicación e interpretación empírica de pruebas psicológicas, se le ha llamado peyorativamente testología (mala interpretación). Tareas del Psychological testing:

- Selección de personas
- Clasificación de personas

Evaluación psicológica: traducción de Psychological assessment. En la segunda Guerra Mundial las potencias en conflicto dieron especial cuidado en la selección de espías, en Estados Unidos esta tarea se la encomendaron a la Organización de servicios estratégicos (en los 50 se fusionó con otras organizaciones y crearon la CIA) que es donde crearon un método especial para la selección de espías que llamaron assessment meted (método de evaluación), a la cabeza de los psicólogos que trabajaban allí estaba Murria; esto se publicó en 1948: Assessment of men

Este método de evaluación era bastante peculiar, específico: era una aproximación muy intensiva (al candidato/s se le/s decía/n que viniesen al centro de evaluación y allí se quedaban internos una semana donde las 24 horas del día eran evaluados), era una aproximación multimétodo (se le aplicaban cuestionarios, tests, entrevistas, observación de conducta, métodos de panel, etc.) y había un método que los autores consideraban necesario para que se pudiera hablar de evaluación que eran los tests situacionales (imitan una situación real para observar cómo se comportan los sujetos ante esa situación). El enfoque teórico era flojo, un funcionalismo en sentido amplio pero no se dedicaron mucho a la teoría, estaban mas avocados a la práctica y en esto hicieron también otro aporte fundamental: vieron la evaluación como un proceso de toma de decisiones que nacía del planteamiento de un problema al equipo de evaluadores y la sugerencia de medidas de acción; era por tanto también una aproximación multijueces, multievaluadores que terminaban en una mesa redonda donde se discutía toda la información sobre el sujeto.

No se centra sólo en lo negativo de los sujetos sino más en lo positivo de los sujetos, de lo que los sujetos tienen por repertorios conductuales para poder vérselas con el ambiente. Esta es la acepción original. Hoy por hoy esta acepción se ha perdido, evaluación psicológica exige ahora mucho menos: no exige tests situacionales, que haya muchos jueces ni centrarse en lo positivo, la única exigencia que se mantiene es el ser un proceso de toma de decisiones diagnósticas.

Evaluación conductual (Behavioral assessment): nació en los años 60. La psicología experimental del aprendizaje venía pasando décadas, pero o había salido a hacer una oferta social de forma sistemática, era una psicología experimental que trabajaba con modelos paulovianos y skinerianos. En los años 50 esta psicología experimental del aprendizaje se encontró capaz de hacer una oferta social en el campo de la terapia y nace la terapia de conducta o modificación de conducta, que desde su nacimiento ha tenido necesidades diagnósticas, necesidades de evaluación, en dos sentidos:

- Para que le provea información útil para montar el plan terapéutico.
- Para medir los efectos de la psicoterapia o de la terapia de conducta.

Los terapeutas conductuales fueron a la evaluación de esa época a buscar respuesta a sus necesidades y la evaluación de esa época fue incapaz ya que estaba especialmente mal empleada para evaluar cambios.

Al no encontrar respuestas tuvieron que crear una evaluación ellos mismos y la llamaron evaluación conductual. Es difícil caracterizarla porque ha evolucionado desde su nacimiento:

En sus comienzos:

- Veía la conducta en función de variables ambientales, había un marcado situacionismo, la conducta era enteramente determinada por el estímulo
- Había un énfasis en lo que se pudiera observar externamente (la conducta motora) y también de conducta fisiológica (la preferida era la primera). La conducta motora era recogida por los métodos de registro directos (métodos de observación externa).
- Había un énfasis en conductas muy concretas, era un enfoque muy molecular, y esas conductas eran sometidas a evaluación y a terapia.
- Había un énfasis idiográfico en el sentido de que las variables objeto de estudio y modificación se sacaban del sujeto mismo, no se imponían desde fuera, se exploraba al sujeto y de allí salían las conductas de interés. Pero también tenía su momento nomotético.

Ahora: ya no es situacionista, es interaccionista porque ve al sujeto como un sujeto activo por lo menos en que selecciona ambientes. Sigue importando lo motor y lo fisiológico pero también importa y mucho lo cognitivo – verbal y se echa de mano a procesos cognitivos, cosa impensable en los comienzos. Los intereses no son más moleculares, interesan conductas más globales, conjuntos de conductas, el enfoque es sindromático, de síndromes, por ejemplo interesa la depresión que es un conjunto de conductas muy complejas. Hay una ampliación de los modelos explicativos, se echa mano de modelos explicativos cognitivos, es una evaluación cognitivo – conductual más que puramente conductual.

Hay cosas específicas de la evaluación conductual:

- Análisis funcional de la conducta: búsqueda de modelos explicativos de la conducta dentro de paradigmas experimentales.
- Gran acento que pone en la operacionalización y el control experimental de las variables (este es un enfoque diagnóstico que viene de la psicología experimental)
- El fundamento teórico está en los avances de la psicología experimental del aprendizaje (acepta modelos que hayan pasado por un rasero de una psicología estrictamente experimental).

Valoración: hablamos de valoración cuando recolectamos información psicológica y de otro tipo y además cuando hacemos un juicio de valor acerca del curso y sobre todo de los resultados e una intervención psicológica. Decimos si ha valido la pena o no, y hasta qué punto, la intervención.

Se puede hacer desde un individuo hasta grandes grupos, comunidades enteras.

Va siempre sobre el programa de intervención; para ver si ha valido la pena o no.

Término original: evaluation: evaluación de programas; assessment: evaluación (no de programas). Evaluación y valoración son sinónimos.

La evaluación se extiende no sólo a pretratamiento sino a una evaluación del tratamiento y evaluación post – tratamiento.

¿QUÉ ES UN TEST?

No es un instrumento de adivinación, tampoco es un artículo de fe (no se cree o no se cree en ellos). Es una técnica , un instrumento de medición.

DEFINICIONES DE TEST PSICOLÓGICO:

Lienert: instrumento científico de rutina (un test no se construye para una sola aplicación, sino para una

rutinaria) para la investigación de una o más características de personalidad (usado en sentido amplio, abarca los aspectos intelectuales, cognitivos), con el objeto de dar en lo posible una expresión cuantitativa sobre el grado relativo de pregnancia individual del rasgo (hasta qué punto las personas poseemos determinadas características) a través de comparaciones tanto inter como intraindividuales.

Cronbach: técnica sistemática (planificada) que compara la conducta de dos o más personas.

Drenth: exploración sistemática de una región definida de comportamientos que debe obtener, en comparación con otros, una evaluación de aspectos definidos del examinado.

Instrumento sistematizado de categorización o medición (incluyendo cualitativa y cuantitativa) que posibilita hacer una aseveración sobre una o más características empírico – teóricamente fundadas o sobre conductas específicas fuera de la situación de prueba (en las situaciones reales de vida: familia, colegio, etc.).

Michel: un test psicodiagnóstico puede caracterizarse como un experimento psicológico específico que sirve para explorar diferencias psíquicas individuales. Debe primar el control de las variables (tal como se controlan en un experimento psicológico).

Anastasi: un test psicológico constituye esencialmente una medida objetiva y tipificada de una muestra de conducta. Surge el concepto de muestra de conducta: ningún test puede evaluar todas las conductas del ámbito que pretende medir.

Freeman: instrumento estandarizado designado para medir objetivamente uno o más aspectos de la personalidad total a través de una muestra de respuestas verbales o no verbales.

Definición:

Instrumento de rutina que elicitaba una muestra definida de conductas en forma sistemática y estandarizada (tipificada) arrojando resultados categoriales o cuantitativos los cuales, en base a investigaciones teóricas y empíricas (no puede existir un test sólo en base a investigaciones teóricas, ni tampoco un test sólo en base a investigaciones empíricas) y en función de comparaciones interindividuales y/o intraindividuales, permiten sacar conclusiones más o menos probables (siempre las conclusiones de los tests son probabilísticas, nunca son dados en términos absolutos) acerca de conceptos psicológicos y/o acerca de comportamientos de los sujetos en contextos más amplios que la situación de prueba (un test que sólo nos da información sobre la situación de prueba es irrelevante).

CRITERIOS DE CALIDAD DE LOS TESTS PSICOLÓGICOS:

OBJETIVIDAD:

No va por las respuestas sino por el sistema de corrección que es estándar para todos los sujetos, protocolos.

El concepto de objetividad es polivalente, tiene distintas significaciones:

- significaciones especiales para distintos autores
- objetividad en sentido estricto
- objetividad en sentido más amplio

Significados especiales para distintos autores:

Cattell: distingue un tipo de test, como tests objetivos: son aquellos tests en los que los sujetos no pueden manipular su respuesta en una dirección deseada. También llamados tests opacos, no son tests transparentes

en su significación. Si a un sujeto le presentan un test que dice: a menudo estoy triste lloro con facilidad, aunque el sujeto no sea psicólogo sabe que si responde a todo que sí le van a catalogar como depresivo: ese es un test no objetivo. Según su acepción el test de Roscharch es mucho más objetivo ya que no sabes lo que significa por ejemplo que tú veas pájaros.

Esta no es la acepción general.

Campbell: objetividad de los ítems: un ítem es objetivo si el examinado puede elegir una respuesta unívoca por verdadera, que tiene alternativas unívocamente falsas. Es una definición que se limita a los tests de aptitud, conocimiento, que es donde se puede hablar de respuestas verdaderas y falsas, pero no por ejemplo a los de personalidad. Nuestros exámenes objetivos serían ítems no objetivos en el sentido de Campbell porque a veces una es medio verdadera y la otra es totalmente verdadera.

Objetividad en sentido estricto:

Cronbach: un test íntegramente objetivo es aquél en que todos los observadores asignan el mismo valor a una ejecución determinada, cuanto más subjetivo sea el test más diferirán en su juicio los observadores y viceversa.

Drenth: la objetividad debe ser interpretada como intersubjetividad, acuerdo entre los distintos evaluadores

Los resultados son independientes del examinador e idénticamente valorados por una serie de correctores.

Este es un criterio de calidad, es una característica deseable de los tests; pero existen muchos otros criterios de calidad.

Objetividad en sentido más amplio:

Lienert: existe objetividad en distintos momentos del proceso de aplicación del test: cuatro momentos:

- la objetividad de la situación de test: un test es objetivo situacionalmente cuando la situación de test es para todos los examinados iguales.
- Objetividad en la aplicación del test: un test es objetivo en su aplicación cuando se aplica a todos los examinados de la misma manera. Es importante que las instrucciones sean rigurosamente las mismas, los tiempos, etc.
- Objetividad en la corrección de un test: objetividad en sentido estricto, distintos evaluadores dan la misma puntuación a la misma respuesta.
- Objetividad de la interpretación del test: un test es objetivo en su interpretación cuando distintos evaluadores dan una misma interpretación a un mismo protocolo de tests. Se suele apoyar en datos de fiabilidad, validez y en los baremos que tiene la prueba.

La objetividad se puede medir, hay coeficientes en el sentido de coeficientes de concordancia entre evaluadores o coeficientes de convergencia entre evaluadores. Una objetividad se puede transformar en un coeficiente cuantitativo.

Diferencias entre el concepto de concordancia entre evaluadores y el concepto de exactitud:

Exactitud: se refiere al grado en que los registros de un observador se superponen con un registro patrón sobre los mismos datos, este registro patrón actúa como criterio para evaluar la calidad de los datos del observador.

En el concepto de exactitud se habla de un desnivel en la calidad de los datos de las dos evaluaciones:

- Evaluación del observador
- Evaluación patrón: la hecha por un experto en la materia y por tanto sabemos que sus datos son de una alta calidad, por eso actúa como patrón frente al nuevo observador no entrenado, hay un desnivel de calidad, en qué punto se aproxima el nivel de calidad inferior al superior.

Toda evaluación donde presumamos razonablemente que es de mejor calidad que la de otro observador puede ser usada como registro patrón y entonces ya no se evalúa la simple concordancia entre evaluadores sino que se evalúa la exactitud de la observación de nuestro observador, si se aproxima a la observación del patrón será más exacta.

FIABILIDAD:

Definición: es una proposición de toda medición. Indica la precisión (consistencia y estabilidad) de la medida de un test.

Comparación entre por lo menos dos medidas

Consistencia y estabilidad: fiabilidad entre:

- Forma diversa
- Tiempos diversos

Anastasi: fiabilidad: consistencia de puntuaciones de los mismos individuos en los mismos tests, en distintas situaciones.

Pervin: fiabilidad: consistencia de las medidas de un test y libre de varianza error.

Concepto de varianza error: en términos técnicos : porcentaje de varianza verdadera en la observación en el test.

La fiabilidad se indica siempre por coeficiente de fiabilidad: ¿a qué tipo de fiabilidad se refiere y cómo se obtuvo ese coeficiente de fiabilidad?

La mayor información que nos da la fiabilidad es la descomposición de la varianza error en sus distintos componentes.

Formas test

Una dos+ sesiones de

Bipartición formas paralelas 1 aplicación

KR-20 (inmediatas) del test

Coeficiente

Fiabilidad del punt.

Test- retest formas paralelas 2+

Con intervalo

Tipos de coeficiente de fiabilidad:

Bipartición: aplicándole la corrección Spearman – Brown

K–R: Kuder – Richardson

Cronbach demostró:

KR20 es igual a la media de todos los coeficientes de bipartición; es mucho más informativo.

KR20 sólo se puede calcular cuando sólo hay dos opciones de respuestas (p,q). Cuando hay varias opciones de respuesta: coeficiente alfa: de más amplio espectro de aplicación. Es el que más se usa.

Los tres coeficientes son de consistencia interna del test.

Para obtener el coeficiente de fiabilidad necesitamos al menos dos mediciones para hallarlo, ¿dónde están? En bipartición se divide el test, vale. Pero en KR20 y alfa como se trabaja con cada ítem por separado, cada ítem es como una forma.

Fiabilidad del puntuador: objetividad. Convergencia entre puntuaciones. Es de distinto tipo que los coeficientes de consistencia interna, se trabaja a través de puntuadores, la fuente de error son los puntuadores (no como antes que eran los ítems: fuente de error interna)

Una forma de test y dos sesiones del test: test – retest: esencial el intervalo entre el test y el retest.

Dos formas del test y una sesión: formas (inmediatas) paralelas

Dos formas del test y dos sesiones: formas paralelas con intervalo.

La mayor cantidad de información que nos puede dar un coeficiente de fiabilidad es la descomposición de la varianza error.

Ejemplo de Anastasi:

Fiabilidad de formas paralelas con intervalo: 0,70

División de mitades (S–B)=0,80

Fiabilidad del evaluador=0,92

Varianza error: contenidos del test (formas paralelas)+tiempo (intervalo)=0,30

Varianza error:

Formas paralelas = contenido + tiempo = 0,30

División en mitades = contenido = 0,20

Debido al tiempo: $0,30 - 0,20 = 0,10$

Debido al puntuador = 0,08

Porcentaje de varianza error= $0,20 + 0,10 + 0,08 = 0,38$ (0,30: contenido + tiempo: no lo ponemos porque se descompone en 0,20 y 0,10)

Varianza verdadera: $1 - 0,38 = 0,62$

V. Verdadera 62% V. Error

Estable a lo largo del tiempo 20% 10% 8%

Consistente en las distintas muestreo de muestreo de diferencias

formas contenido tiempo entre

Libre de la diferencia entre puntuaciones

puntuaciones

Un coeficiente de fiabilidad no puede ser utilizado inespecíficamente, siempre debe ser utilizado en función de la varianza de error que detecta.

¿para qué sirve un coeficiente de fiabilidad? Sirve para fijar un intervalo dentro del cual estará la puntuación verdadera. Se denomina intervalo de confianza y se mide en función del error de medida del test. Este error está en directa relación con el coeficiente de fiabilidad.

La puntuación observada no debe confundirse con la puntuación verdadera.

Error típico de medida:

$$\text{med} = \text{test} \sqrt{1 - r_{xx}}$$

r_{xx} = coeficiente de correlación intratest.

Ejemplo: un sujeto responde al WIS y podemos calcular su CI y su puntuación es de 90. podemos calcular en qué intervalo de puntuaciones se encuentra la puntuación verdadera del sujeto con cierta probabilidad; podemos calcular una distribución, los errores de medida se distribuye normalmente, entre qué puntuaciones está el CI verdadero del sujeto. Para ello calculamos el error típico de medida implica:

test = 15 (en este caso)

$r_{xx} = 0.91$

$$\text{med} = 15 \sqrt{1 - 0.91} = 15 \sqrt{0.09} = 15 \times 0.3 = 4.5$$

I.C. $\pm$ un error típico de medida

$90 \pm 4.5 = (85.5, 94.5)$: I.C. entre el cual se encuentra el valor verdadero del CI del sujeto, es un CI de 9 puntos de CI con un 68% de probabilidad (2/3 de probabilidad).

34% 34%

89.5 90 94.5

Se suele trabajar con un I.C. de más o menos un error de medida porque aunque nos da menos probabilidad nos da un I.C. más estrecho que es lo que nos interesa.

*Nunca debe informarse solamente del valor observado en un test porque entonces probablemente va a llevar al usuario a confundir el valor observado con el verdadero, lo que en principio no es cierto.

Siempre debemos informar además del valor observado, del error típico de medida y del I.C. en el cual se encuentra la puntuación verdadera con una cierta probabilidad.*

Segundo uso del coeficiente de fiabilidad: podemos calcular el error típico de medida de las diferencias entre puntuaciones. Queremos saber si dos puntuaciones distintas de un mismo sujeto son verdaderas diferencias o son diferencias que pueden ser debidas al azar, el procedimiento más correcto es hacerlo a través del cálculo del error típico de medida de las diferencias:

$$dif = med1 + med2$$

1,2 : test 1 y 2

En el caso en que las desviaciones típicas de los dos tests o de los dos subtests sea la misma (caso bastante frecuente con subtests), la fórmula es:

$$dif = DT_{común}^2 - r_{11} - r_{22}$$

DT =desviación típica

Ejemplo: un sujeto obtiene en el WIS una puntuación de 90. observamos que en el test de semejanzas (pensamiento abstracto) ha obtenido 12, en el de cubos (inteligencia práctica) ha obtenido 8. los subtests del WIS tienen una media y una desviación típica común y la desviación típica es de 2. ¿responde esta diferencia entre 8 y 12 a una diferencia verdadera o no? Hay que calcular para ello el error típico de las diferencias:

$$dif = 2^2 - 0.65 - 0.75 = 2 \cdot 0.6 = 1.54$$

I.C. para +/- un error típico de medida de las diferencias:

$$\text{Intervalo} = 1.54 \times 2 = 3.08$$

La diferencia es de 4 y el I.C. de 3,08, la diferencia es mayor y por lo tanto la diferencia se supone verdadera por lo menos con el 68% de la probabilidad. Si estuvieran dentro del I.C. las dos puntuaciones no habría una diferencia real; al estar fuera del I.C. podemos decir que las diferencias son verdaderas con cierta probabilidad (68%).

Tercer uso del coeficiente de fiabilidad: validez criterial: con ayuda del coeficiente de fiabilidad podemos corregir por atenuación : ¿cuál sería la corrección entre el test y el criterio si sus fiabilidades fueran perfectas, es decir, de 1? Si la diferencia entre el coeficiente de correlación observado y el atenuado es grande, entonces vale la pena seguir trabajando con esas variables intentando subirles su fiabilidad; en cambio si la diferencia entre los dos coeficientes no es grande decimos que no vale la pena seguir trabajando en la fiabilidad de estas variables porque la ganancia que vamos a obtener, aunque aproximamos la fiabilidad a 1, la diferencia no va a ser grande.

Corrección por atenuación:

r_{xy} (observado)

$r_{xy}(\text{optima}) =$

$r_{xx} r_{yy}$

Ejemplo: predecir rendimiento académico con un test que hemos construido y $r_{xy}(\text{observado}) = 0.30$ (es una correlación modesta, a primera vista nos indica abandonar nuestro test) ¿si sus fiabilidades fueran de 1?

- 0.30

$r_{xy}(\text{optima}) = = = 0.43$

$0.75 \times 0.65 = 0.49$

Si las fiabilidades fuesen de 1 en el test y en el criterio mi coeficiente mejoraría de 0.30 a 0.43.

Los coeficientes de correlación entre variables distintas se interpretan como el coeficiente de determinación que es su cuadrado.

Sí valdría la pena. Hay una mejora del doble: $0.30^2 = 0.09$: 9% de varianza común test / criterio. $0.43^2 = 0.18$: 18% de varianza común

VALIDEZ:

A.P.A.: grado en que un test es capaz de lograr sus objetivos. Con esto no distingue entre la validez y la utilidad por ejemplo.

Validez es medir lo que queremos que el test mida, todo ello y sólo ello: esta definición es más correcta.

La validez es el grado en el que un instrumento de medición mide lo que pretende medir.

Es importante subrayar que existen distintas vertientes de la validez pero éstas se interpretan entre sí, cosa que no pasaba con las distintas formas de fiabilidad que son independientes entre sí.

Cronbach: en sentido estricto no se valida un test sino interpretaciones que hacemos del test, deducciones que sacamos de él, por eso un test puede ser válido en ciertas interpretaciones y no serlo en otras.

Estos distintos tipos de validez están interrelacionados entre sí: sólo conceptualmente son independientes (APA):

Validez aparente: face validity : es el grado en que un test parece medir a los ojos de quien lo cumplimenta (no del psicólogo) lo que se dice o supone que pretende medir. Es la transparencia del test a los ojos del usuario. Un test con validez aparente sería subjetivo en el sentido de Cattell. Los tests transparentes son por lo general mejores que los opacos. No tiene ninguna relevancia científica, teórica pero el práctico sí que la valora porque le sirve:

- para crear un buen rapport, una buena relación con el examinado porque no tiene sensación de que le engañan.
- Eleva la motivación del sujeto para responder (es muy importante)

El psicólogo práctico se lo toma en serio y prefiere test con validez aparente.

Tipos de validez:

Validez de contenido: APA : es el grado en que el contenido del test refleja las situaciones o materias acerca de las cuales van a sacarse conclusiones. Hay tres etapas:

- Está en determinar el dominio completo de contenidos que es relevante para el objeto de la medición.
- Muestrear contenidos específicos a partir de ese dominio total.
- Poner estos contenidos en una forma adecuada para su contestación, en forma de ítems.

Estrategias útiles para la validación de contenido:

- Análisis de tareas: análisis cuidadoso de lo que el sujeto está ejecutando en lo que nos interesa.
- Análisis racional de los contenidos por parte de expertos: es muy usual echar mano de algún experto en la materia para que examine mi prueba.
- Análisis de los ítems: por un lado se analiza el contenido y el modo de presentación de los ítems, por otro lado también se examinan algunos datos estadísticos, por ejemplo el nivel de dificultad.

Son muy útiles unas tablas de doble entrada para hacer un buen muestreo de contenidos. Ejemplo: prueba de ansiedad:

Situación tipo de respuestas

C – v M F

Familia 2 1 1

Trabajo 2 1 1

Comunidad 2 1 1

¿cuál va a ser el peso relativo que le demos a cada parámetro?

En el tipo de respuesta por ejemplo nuestra teoría puede señalar lo C–V, entonces le damos coeficiente doble (2) y a lo motor y fisiológico un coeficiente simple (1). En las situaciones les damos el mismo peso (1).

Para dar respuestas cognitivo – verbal construyo 2 ítems, para F y M sólo 1.

Está subsumida, incluida dentro de la validez de constructo, conceptual.

Validez criterial: (de criterio u orientada al criterio) APA: se determina comparando las puntuaciones del test con una o más variables externas que se considera que ofrecen una medida directa de las características o conducta en cuestión.

El problema es decirnos algo sobre una variable de importancia social que está más allá del test, que es externo, independiente del test. A esto es a lo que se llama predecir en sentido estadístico de la palabra que es compartir varianza común de dos VI, que esas varianzas están separadas en el tiempo estadísticamente es irrelevante. Se mide con un coeficiente de correlación rxy (x: predictor, y: criterio).

En la práctica sí es importante si un test predice el presente, pasado o futuro (teórica y estadísticamente no).

Indica la correlación entre nuestro predictor y una variable externa e independiente, que llamamos criterio. Implica la comparación de las puntuaciones del test con una o más variables externas o independientes.

Criterio: toda variable que debe ser predicha.

Antes a la validez criterial se la llamaba predictiva, pero esta expresión se ha dejado de lado porque en la práctica el tiempo sí es importante. En sentido estadístico predecir es compartir varianza común y por eso la palabra predicción era ambigua.

Tipos de validez criterial:

- Validez concurrente
- Validez predictiva en sentido estricto.

La validez concurrente a su vez tiene dos sentidos:

- Se habla de validez concurrente cuando comparamos dos predictores con el fin de sustituir uno por otro. Normalmente se quieren sustituir uno por otro por razones de economía. Ejemplo: tenemos un predictor muy largo: ejemplo: Binet: intentamos hacer un test más breve y que tenga una correlación alta con el test de Binet, eso es un estudio de validez concurrente. Ejemplo: para sustituir un test de aplicación individual por un test de aplicación colectiva. Trabajo con dos predictores los cuales correlacionan con el fin de sustituir uno por el otro.
- La comparación entre un predictor y una medición contemporánea de un criterio. Los dos están disponibles a la vez. Aquí entran todos los tests que llamamos de diagnóstico. Ejemplo: test para evaluar esquizofrenia, lo construimos y comparamos con la esquizofrenia que el sujeto tiene en ese momento que es medida por algún criterio y quedarnos con nuestro test que es mucho más fácil de aplicar.

Validez predictiva en sentido estricto: la correlación entre un predictor y un criterio que se medirá en el futuro. El fin es pronóstico, ya no es diagnóstico; usar nuestro test como un pronóstico de algo. Ejemplo: usamos las calificaciones de secundaria para pronosticar el rendimiento en la universidad. Hay siempre un desfase temporal. Se habla de predecir en sentido estricto de la palabra.

Se dice que el gran problema de la validez criterial es la medición de los criterios: durante décadas los psicólogos intentaron construir los mejores predictores posibles y descuidaron la medición de los criterios porque normalmente no la hace el psicólogo sino otros profesionales. Entonces se construían tests muy buenos para predecir criterios que se medían muy mal y al pasar eso sufre la correlación.

Desde hace décadas los psicólogos vienen trabajando en la medición de los criterios, colaborando con otros profesionales para mejorar la medición de los criterios, así por ejemplo en la mejora de la medición de los diagnósticos psiquiátricos, los psicólogos han tenido un papel esencial, la primera escala criterial de evaluación psiquiátrica la hizo un psicólogo, también se ha colaborado con los educadores. Tendremos buenas predicciones sólo si tanto el predictor como el criterio están bien mediados.

Cualidades de un buen criterio de validación:

- Adecuación racional: que sea racionalmente adecuado, que el criterio responda realmente al concepto que se está queriendo medir.
- Que tenga ausencia de sesgos: de tal manera que todas las personas tengan en principio las mismas oportunidades de ser evaluados por el criterio.
- Debe tener fiabilidad: su medición debe ser fiable, con la fiabilidad que le sea más idónea (consistencia interna, fiabilidad del evaluador, fiabilidad temporal, etc.).
- Tiene que tener disponibilidad: hay criterios que pueden ser muy buenos pero son difícilmente disponibles. Ejemplo: estudio sobre predicción de rendimiento académico y los profesores se niegan a dar las notas. Hay que asegurarse previamente.

La mayor estimación del criterio es el desempeño continuado en la tarea. Ejemplo: si pudiera calificar

nuestros conocimientos cada clase tendría una calificación óptima.

Algunos conceptos incluidos en la validez criterial:

- **Validez diferencial:** un predictor tiene validez diferencial cuando es capaz de diferenciar entre dos o más criterios. Por ejemplo: si tenemos una prueba que distingue muy claro entre el diagnóstico de trastorno depresivo y de ansioso, tenemos una buena prueba en este sentido para diferenciar esos dos trastornos.
- **Validez de incremento:** es la contribución del test o predictor a un incremento en la predicción del criterio en relación con otras informaciones. Nuestro objetivo no es obtener un coeficiente de precisión mayor de cero, sino uno mayor del que ya está funcionando. Hay que averiguar cómo está siendo la predicción que ya se está llevando a cabo, y si es alto hay que decir que no se puede contribuir mucho, pero si la predicción es deficiente nos motiva a hacer el estudio y a hacer un incremento de la validez.
- **Validez cruzada:** cross validity, se habla de ella cuando replicamos o repetimos los estudios de validez criterial en una muestra distinta de la original pero proveniente de la misma población. Es repetir el estudio de validez con otra muestra. Esto es importante porque los estudios de validez criterial suelen hacerse con muestras que no son muy representativas de la población y es posible que el coeficiente que obtengamos esté sesgado de alguna manera, normalmente es mayor que el coeficiente de predicción real para esa población. Se aconseja fervientemente realizar un estudio de validación igual al anterior pero con otra muestra y vemos si el coeficiente de validación se mantiene o baja y normalmente baja un poco, si es así hemos tenido éxito, pero hay ocasiones en que baja dramáticamente y nos ha hecho fracasar en nuestro estudio de validez criterial. Menor será el descenso cuanto más representativa sea la muestra de la población en cuestión.

Los distintos tipos de validación se interrelacionan entre sí. Validez de contenido: incluida dentro de la validez de constructo. También entre la validez criterial y la validez de constructo: esto queda claro cuando uno no se conforma con el mero r_{xy} , sino que se pregunta por qué hay una relación entre esas dos variables, por qué comparten varianza común, en este caso cuando toma la postura del científico ese por qué hay que responderlo en términos de compartir procesos psicológicos, conceptos psicológicos, constructos, y entonces el estudio que comenzó como un estudio de validez criterial se continúa en un estudio propiamente de validez de constructo.

- **Validez de constructo:** Cronbach: análisis de la significación de las puntuaciones de un test, expresadas en conceptos psicológicos. Drenth: es la explicación del rendimiento del test, según conceptos que están en el marco de la teoría psicológica. APA: la validez de constructo es evaluada para investigar qué cualidades mide un test, en qué grado ciertas hipótesis o conceptos explicativos están en la base de la ejecución del test.

Un constructo se define como un concepto propuesto formalmente, cuya definición y límites están explícitamente relacionados con datos empíricos. Concepto operacional.

Postulado básico de Cronbach y Mill: la validación de constructo significa lo mismo y tiene los mismos pasos que la comprobación de teorías científicas.

Esto tiene dos consecuencias:

- La validación de constructo trabaja por aproximaciones sucesivas, no es de todo o nada. Nos aproximamos a la validez de constructo poco a poco, en cuanto que acumulamos evidencias a favor de la validez de constructo. Un test siempre tiene más o menos validez de constructo, si hay evidencia acumulada.
- Para examinar la validez de constructo, se requiere combinar procedimientos lógicos y empíricos. No

es un ejercicio puramente lógico ni racional. Es también un ejercicio empírico, datos.

Etapas en la validación de constructo: las mismas que las del método científico de la vía hipotético – deductivo:

1.– construir hipótesis acerca de comportamiento del constructo, respecto a sus más variadas características, o el comportamiento de una red de constructos con sus correspondientes correlatos (datos) empíricos: denominados REDES NOMOLÓGICAS

2.– deducción de consecuencias contrastables: si mi hipótesis es tal, deduzco que mi test se comportará así. Son en forma de predicciones:

- Confirmativas. Aquella en que una R+ confirme la hipótesis
- Infirmitivas: aquella en que una R+ rechaza la hipótesis.

3.– comprobación empírica de las consecuencias: aplico mi test.

4.– aceptación (si los datos responden a mis expectativas)/rechazo o modificación (si los datos no responden a mis expectativas).

La validez de constructo significa que mi test se comporta como debe. Si mi test no lo hace: mala validez de constructo.

Estrategias más usadas en la validación de constructo: cualquiera de lo referente a mi test, parece ser interesante respecto a su validación de constructo.

Relacionado con la investigación de la estructura:

- Contenidos del test: deben responder al constructo. Todo lo relacionado con la validación de contenido.
- Estudios de fiabilidad: tengo que tener cierta consistencia interna y estabilidad.
- Intercorrelación de los ítems y los análisis factoriales: una prueba básica de validez de constructo es que el constructo aparezca reflejado en un determinado factor. Si no fuera así, el constructo no tendría entidad.

Validez convergente y divergente: la medición de mi constructo debe converger con otras variables y debe divergir con otras que yo predigo (típico estudio de redes nomológicas).

Campbell y Fiske plantean una estrategia para estudiar hasta qué punto la varianza de mis instrumentos y mis constructos se deben al método con que yo la mido o se deben al constructo que yo digo medir. Para ello Campbell y Fiske trabajan con la matriz multirasgo y multimétodo: miden distintos rasgos y métodos.

Hay que probar que nuestro constructo tenga correlación más alta entre los métodos que intra – métodos. Para esto usaríamos la estrategia de Campbell – Fiske: matriz multirasgo – multimétodo (rasgo ! constructo). Debemos tener distintos métodos que midan distintos rasgos (ver matriz multirasgo – multimétodo).

La diagonal principal mide la correlación de un rasgo consigo mismo y medida por el mismo método, se obtiene el coeficiente de fiabilidad llamado de consistencia interna (correlación monorasgo – monométodo).

La primera línea subordinada es un mismo rasgo correlacional con distintos métodos, correlación monorasgo – heterométodo.

Correlación heterorasgo – monométodo: segunda línea subordinada, distintos rasgos medidos con el mismo método; nos da la varianza del método (de las 3).

Correlación heterorasgo – heterométodo.

Los coeficientes que reflejan mayor correlación son los monorasgos – monométodo : coeficiente de fiabilidad.

Las correlaciones siguientes más altas son las monorasgo – heterométodo. Las siguientes son heterorasgo – monométodo y las de menos correlación las heterométodo – heterorasgo.

IMPORTANTE: lo crucial en esta matriz es que las correlaciones monorasgo – heterométodo sean significativamente superiores a las correlaciones heterorasgo – monométodo. Entonces vemos surgir los rasgos por encima de los métodos. Es difícil pasar esta prueba.

Estrategias que tienen que ver con las correlaciones del test, con otras variables: aquí se incluyen estudios de diferencias entre grupos. Correlación con distintas edades, también correlación con criterio. Ejemplo. Un test que mide ansiedad debe correlacionar con estudio de ansiedad; necesidad de detectar si un test no está midiendo otras variables que comúnmente se filtran en los tests. Ejemplo: muchos tests de rendimiento académico, correlaciona a lo mejor, tan alto con inteligencia que consigo mismo. Por ello hubo muchos problemas en determinar si un test era de inteligencia o de creatividad.

Puede influir también en la respuesta (produciendo sesgos), ejemplo: deseabilidad social, debe ser menor a la correlación del test consigo mismo.

Manipulación experimental de las variables: ejemplo: al hacer pensar en voz alta a los sujetos cuando elaboren la respuesta. Así veremos los procesos que utiliza el sujeto para llegar a las respuestas del test, vemos si es deductivo, inductivo... veremos los constructos que hay detrás de las respuestas.

Cualquier dato es interesante en la validación del constructo, pero debe estar engarzado a alguna hipótesis precisa respecto al comportamiento de mi constructo o test.

Fiabilidad, validez y generalizabilidad: introducción de la teoría de la generalizabilidad de Cronbach, en un principio se aplica a la fiabilidad, y no se centra en el concepto de valor verdadero, sino en si podemos generalizar en algo más allá a través de los ítems del test.

Si hago un estudio de estabilidad temporal, no busco el valor verdadero, sino hasta qué punto puedo generalizar resultados a través del tiempo (si el retest es alto podré generalizar, sino, no).

Cronbach subraya el hecho de que se puede generalizar a través de distintos universos a través de evaluadores, a través de los ítems, de la forma del test.

Cronbach llama a la puntuación verdadera, puntuación universo.

Los universos de generalización que distingue Cronbach se dividen en dos áreas:

- Personas
- Condiciones

Personas: la inferencia que se hace de muestra a población. La recomendación es que la muestra en que se hace el estudio de fiabilidad, tiene que ser lo más representativo posible para la población.

Condiciones. Cuestiones de fiabilidad:

- Los ítems que componen un test
- Las formas de un test
- Los estímulos que acompañan a un test (por ejemplo las instrucciones)
- Los observadores
- Las ocasiones en que se aplica el test.
- Las situaciones de medición: cuestión de validez concurrente.

Mezcla en los problemas de generalización, la fiabilidad y validez. Uno de los mayores aportes de esta teoría, es diluir los límites entre fiabilidad y validez. Ambos problemas son de generalización (hasta qué punto puedo generalizar las puntuaciones). La diferencia es que las puntuaciones de fiabilidad son de generalización intrainstrumento, por ejemplo: la consistencia interna a través de los ítems del test, o la estabilidad temporal a través de distintas aplicaciones del test.

Las puntuaciones de validez son de generalización interinstrumento, por ejemplo en la validez predictiva nos preguntamos hasta qué punto se puede generalizar de un test a un criterio, esto es una variable distinta.

La generalización va en aumento, mientras pasamos de un criterio de calidad a otro.

En la consistencia interna : generalización mínima ítem a ítem.

En la estabilidad temporal (test – retest): generalización de distintas aplicaciones del test, va aumentando.

Deformas paralelas: ya no es una sola forma del test, sino varias distintas y se suelen aplicar también en momentos distintos. Mayor generalización.

Validez criterial de tipo convergente: intenta generalizar entre un test y otro; o entre un test y un criterio medido contemporáneamente.

Validez criterial predictiva: intenta generalizar de un test a un criterio mediante un espacio temporal. Generalización mayor.

Validez conceptual o constructo: es la más pretenciosa de todas, porque a través del constructo pretende predecir una serie indeterminada de criterios. Es el máximo intento de generalizabilidad.

Visión unitaria de la fiabilidad y validez, las une la pretensión de generalización.

Teorías de la utilidad:

La utilidad de un test : criterio de calidad que ha llamado la atención de los investigadores.

La teoría de la decisión nació en un campo distinto a la evaluación psicológica.

La utilidad tiene que ver con la cantidad de decisiones correctas e incorrectas que se hacen en un proceso de evaluación, y con los valores de utilidad relacionados en estas decisiones. Este concepto de utilidad está relacionado con el concepto de validez, pero el de utilidad es más complejo porque introduce nuevos parámetros.

Problema de selección = aceptar determinados sujetos y rechazar a otros, según las puntuaciones obtenidas en un test o un criterio cuantificable.

Y (ordenada)

FN VP

TB

VN FP

X (abcisas)

Se calcula a partir de la nube de puntos el coeficiente de correlación (de validez criterial).

Cuando esto se transforma en un problema de selección, tenemos que introducir dos parámetros:

- La cantidad de sujetos con las características necesarias para ser seleccionado (que deberían ser seleccionados): tasa de base (TB) más o menos en el 50% en nuestro ejemplo. Tasa base: cantidad de sujetos bien seleccionados que hacemos al azar (+/- 50% en este caso).
- Razón de selección (RS) (selection ratio): aquí +/- 50%, cuántos sujetos tengo que seleccionar con mi test.

Con estos parámetros obtendremos cuatro resultados posibles:

- Cuadrante superior derecho : VP: sujetos positivos según el criterio (por encima de la tasa de base) y a la vez son seleccionados por mi predictor, son válidos positivos: están bien seleccionados.
- Cuadrante inferior izquierdo: VN : porque lo predice bien como negativo. Son negativos, no dan la talla o no tienen las características determinadas. Y mi proceso de selección lo rechaza, porque están a la izquierda del ratio. Son negativos y válidamente negativos.
- Cuadrante inferior derecho : FP: decisión errónea: sujetos negativos en relación con el criterio (debajo de la tasa de base), sin embargo son predictivos como positivos porque están a la derecha de la ratio: predichos como exitosos o que poseen las características buscadas. Falsos positivos.
- Cuadrante superior izquierdo: FN: decisión errónea: aquellos sujetos que según el criterio son positivos (son los que tienen las características buscadas y dan la talla) sin embargo la ratio dice que son negativos, que no dan la talla. Falsos negativos.

La utilidad es distinta de la validez, con un mismo coeficiente de validez podemos tener utilidades distintas.

Con una hipotética validez perfecta (todos los sujetos sobre la diagonal) la correlación predictor – criterio = 1

TB=0.5

RS=0.5

¿Existen sujetos válidos positivos? Sí, los del cuadrante superior derecho.

¿Existen sujetos falsos positivos? No.

¿Existen sujetos válidos negativos? Sí.

¿Existen sujetos falsos negativos? No.

Todas mis predicciones son correctas.

Con la validez perfecta, pero con tasa base = 0.60, y mi razón de selección (ratio) es muy estricta = 0.05.

VN FN VP

Conclusión: siendo la validez la misma en los dos casos (=1). Sin embargo en un caso tenemos un 100% de decisiones correctas y en el otro apenas llegamos a un 50%.

La utilidad !decisiones correctas.

Validez !correlación test – criterio.

FN VP

VN FP

En el cuadro existen dos estrategias posibles:

- Sólo considerar a los sujetos aceptados en la selección.
- Que me interesen los 4 tipos de resultados posibles.

Solo considerar a los sujetos aceptados en la selección: típico de las selecciones de personal: sólo me interesa el personal que yo selecciono, las dos de la derecha, sacar la mayor cantidad posible de VP y la menor posible de FP.

Lo que hay que lograr es que $VP / (FP + VP)$ se acerque a la unidad.

Para lograr esto se debe hacer la razón de selección lo más estricta posible.

0.80 0.30 5%

Si ponemos una razón de selección muy baja (0.80), los FP son prácticamente igual o más que los VP. Si ponemos de 0.30 la proporción de VP frente a la proporción de FP es ya mucho mayor. Si sólo seleccionamos a un 5% de los candidatos, sólo hay VP, han desaparecido los FP.

La recomendación para acercar la fórmula a 1 es hacer la razón de selección lo más pequeña posible. Por eso en selección de personal se seleccionan tan pocos candidatos para aumentar todo lo posible los VP y reducir lo más posible los FP.

Que me interesen los 4 tipos de resultados posibles: FN y VN pasan a ser de alguna forma de importancia. Ejemplo: selección de unos determinados sujetos para recibir un tratamiento médico: dejar fuera a sujetos que tienen la enfermedad pero no van a recibir el tratamiento (FN) es una decisión grave; a veces puede ser mortal; se intentaría que fueran los más posibles de VN en problemas sanitarios.

VP + VN

VP + FP + VN + FN

Tenemos que pretender que esta fórmula se acerque lo más posible a 1.

La recomendación es igualar lo más posible la razón de selección a la tasa de base, y ojalá ambas lo más cerca de la línea media, de 0.50: $RS = TB = 0.50$. Si se tiende a eso se dan las mayores posibilidades de decisiones correctas y las menores posibilidades de decisiones erróneas.

En este caso serían todas las decisiones correctas.

Todo desnivel entre TB y RS se paga en más decisiones erróneas.

Asignación de valores de utilidad a cada uno de los resultados: ¿cuáles son más importantes, o son igual /desigualmente importantes?

La asignación de valores de utilidad es el Talón de Aquiles de la teoría de la decisión, es muy difícil asignar los valores de utilidad. Lo tienen más fácil las empresas privadas que lo convierten a valor económico (ganancias o pérdidas). En decisiones que tocan a la salud o enseñanza pública las cosas son más complejas, a veces no se puede reducir a valores puramente económicos, aunque se intenta hacer. Hay autores que dicen que se tome en cuenta la subjetividad del sujeto, ejemplo: la frustración de un FN. Esto hace que a veces se renuncie a dar valores de utilidad, pero eso es renunciar a decidir sobre una base racional. La forma de la utilidad es:

$$U = U1 \times p(VP) + U2 \times p(VN) + U3 \times p(FP) + U4 \times p(FN) + U_t$$

U1: valor de utilidad de un VP

p (VP): proporción de verdaderos positivos.

U_t = valor de utilidad del proceso de selección: normalmente es negativo (al igual que FP y FN) porque es costo.

Si sólo tengo en cuenta VP y FP el segundo y cuarto sumando sería cero y la fórmula me quedaría:

$$U = U1 \times p(VP) + U3 \times p(FP) + U_t$$

¿CÓMO HACER DIAGNÓSTICO DE PRINCIPIO A FIN?

PROCESO DE EVALUACIÓN DESDE EL ENFOQUE CONDUCTUAL (APLICABLE A CUALQUIER MODELO)

Teorías de la evaluación:

- Aportar información útil para montar programas de entrenamiento: tarea de indicación de tratamientos, significa una configuración individualizada de programa de tratamiento.
- Evaluar los efectos del tratamiento: contrastar la eficacia obtenida en el tratamiento con los objetivos que nosotros mismos nos hemos planteado: Tarea de control de resultados.

Características del enfoque conductual:

- Afán de científicidad: sobre todo contrastabilidad, las hipótesis deben ser empíricamente contrastables.
- Análisis funcional de la conducta: cuando de lo que se trata es de cambiar conductas es inevitable proponer un modelo explicativo de la conducta en cuestión. Tenemos que proponer un modelo explicativo de la conducta. Si nuestro trabajo no es cambiar conducta, podemos prescindir de modelos explicativos pero no en el caso contrario porque sobre esos modelos explicativos se va a apoyar el tratamiento. Al proceso de construcción de un modelo explicativo de la conducta en cuestión se le llama análisis funcional de la conducta.
- Trabaja en un proceso hipotético – deductivo: construye hipótesis, los modelos funcionales son modelos hipotéticos, de ellos se deducen consecuencias contrastables y se contrastan empíricamente.
- Aproximación experimental: no quiere decir que necesariamente se tengan que hacer experimentos, a veces se hacen; lo que quiere decir es que necesariamente se tienen que manipular variables, porque de lo que se trata es de lograr cambios de conducta, sino no sería necesaria la manipulación de variables. Experimental

se refiere a la fase de contrastación de hipótesis.

- Es necesario un enfoque longitudinal del proceso de evaluación: porque se pretenden cambios, existen procesos transversales de evaluación (un proceso de selección). Hay una evaluación antes del tratamiento, junto al tratamiento y una evaluación que sigue al tratamiento y todo esto es necesariamente un transcurso temporal.

Proceso de evaluación conductual:

Tareas: indicación control

Fases

Diagnóstico: modelo

Funcional ! tratamiento ! seguimiento

Indicación de tratamientos y control de éstos, y sus efectos.

Primera fase: tarea de indicación: diagnóstico o fase de evaluación de pre – tratamiento. El objetivo es estar volcado hacia el tratamiento, ofrecer toda la información útil para el tratamiento. Desde el punto de vista conductual, la tarea central es ofrecer un modelo explicativo de las conductas problemáticas o las alternativas a ésta, algunos autores (no conductistas) reducen toda la fase de diagnóstico a la tarea de análisis funcional: esto no es bueno. Hay otras tareas muy importantes.

Segunda fase: de tratamiento, vemos que entra la evaluación de control de resultados. Este control se da en dos momentos:

- Durante el tratamiento: hay un control de resultados. Se hace evaluación.
- Al finalizar el tratamiento: aquí no termina el control de resultados, una vez que el tratamiento ha terminado sigue pasando el tiempo, y en la fase de seguimiento: sabemos si se mantienen o no los resultados.

Relación entre diagnóstico y tratamiento: la flecha en sentido izquierdo: ha tenido una interpretación errónea: el diagnóstico. Se rehace a un diagnóstico hipotético, y la fase de tratamiento valida al diagnóstico. De tal manera que si el tratamiento resulta exitoso, el diagnóstico no validado! este sentido es erróneo, porque inmuniza el tratamiento, ya que éste, según este razonamiento, no será el culpable del fallo.

El tratamiento puede fallar, y es erróneo que la fase de tratamiento es la fase de validación del diagnóstico. Antes de comenzar el tratamiento el diagnóstico tiene que contrastar sus propias hipótesis.

El movimiento hipotético – deductivo experimental debe darse completo durante la fase de tratamiento.

La flecha en sentido inverso: durante la fase de tratamiento sigue apareciendo información diagnóstica de interés. Ocurre un fenómeno feed – back.

Relación tratamiento y seguimiento: también se ha argumentado mal; se ha dicho que el seguimiento valida al tratamiento. Si el seguimiento OK: el tratamiento OK. Si el seguimiento fracasa el tratamiento es malo. Esto no es verdad, durante el seguimiento pueden ocurrir situaciones adversas, que hagan que la persona recaiga, sin querer por esto decir, que el tratamiento no ha sido un éxito. Durante la fase de seguimiento obtenemos información de interés sobre el trato mismo. De tal manera que se produce feed – back.

TAREA DE INDICACIÓN DE TRATAMIENTO:

- exploración de conducta problemática y sus circunstancias
- exploración de los enlaces funcionales y postulación de un método
- fijación de la línea basal
- fijación de objetivos
- diseño del plan de intervención.

Exploración de conducta problemática y sus circunstancias.

Se trata de definir las conductas problemáticas y delimitar la situación en que aparece. De tres formas:

- Precisa: definición precisa de conductas y circunstancias: definidas con claridad y que las distingue de otras conductas y circunstancias. Claridad y distinción.
- Objetiva: independiente de cualquier examinador en particular. La definición debe ser común a los evaluados.
- Cuantificable: en escalas de razón, y también en sentido más suave con escalas de estimación (nunca / alguna vez / frecuentemente /siempre). Los parámetros en cuantificación son:
 - Frecuencia (el más usado).
 - Intensidad
 - Duración

En la frecuencia de conducta, podemos utilizar escalas de razón. La teoría skineriana constituida por el parámetro frecuencia. En la duración de conducta (con reloj) también podemos utilizar escalas de razón. Los intervalos son iguales y hay cero absoluto.

Esta fase tiene problemas:

- El cliente o paciente (que no tiene por qué ser psicólogo) nos puede dar definiciones de sus circunstancias demasiado vagas o generales.
- De lo general tenemos que hacer algo particular, de lo subjetivo algo objetivo y de lo vago algo preciso.
- Los clientes hacen referencia a situaciones, dicen que SON de tal forma, esto no vale, porque le da un carácter de continuidad a un comportamiento, que no lo es. De ser a estar.

Análisis funcional de la conducta:

- ◆ **Momento descriptivo**
- ◆ **Momento interpretativo**
 - ◇ **Control situacional**
 - ◇ **Reforzamiento**
 - **Positivo**
 - **negativo**
 - ◇ **Conducta respondente**
- ◆ **[aspectos funcionales de conductas alternativas]**

Control situacional: control por estímulos discriminativos ya sean facilitadores o inhibidores de la conducta.

Reforzamiento: indagar sobre un posible reforzamiento, es decir, buscar si hay estímulos consecuentes que sigan con cierta regularidad, no azarosamente, a la respuesta y que aumentan su probabilidad de aparición.

Ro = respuesta operante.

Dos posibilidades: ambas son importantes en la clínica:

- ◆ Control por refuerzo positivo (C+): consecuencia positiva, se identifica porque actúa por presencia del estímulo y comúnmente se observan conductas de aproximación del sujeto al estímulo.
- ◆ Refuerzo negativo: es un refuerzo, aumenta la probabilidad de aparición de la respuesta pero por reducción o evitación de consecuencias negativas. La sigla es C-. Se habla de dos paradigmas:
 - ◇ El paradigma de huida: se reduce la estimulación aversiva presente por retirada del sujeto, el sujeto huye y la conducta se refuerza.
 - ◇ Paradigma de evitación: a través de la conducta se evita la presencia del estímulo aversivo. Un sujeto que tiene fobia a los perros, ve uno y pasa a la cera de enfrente y sigue caminando, una vez que ha pasado vuelve a la cera anterior.

Existen también autoreforzamientos: se los da el propio sujeto: pueden ser orgánicos (estados orgánicos placenteros, ejemplo: adicciones a drogas) o subjetivos (autoverbalizaciones: qué bien lo he hecho).

Conducta respondente: la conducta en vez de ser un operante sea una conducta respondente (Rr). Se identifica porque hay un control reflejo, forzoso, determinista al 100% de la conducta en cuestión por parte de estímulos antecedentes. EC, RC que se constituyen sobre EI y RI (terminología más al uso).

Aspectos funcionales de conductas alternativas: son enormemente importantes. Indagar en ello significa: dar una definición topográfica de la conducta, tenemos que preguntarnos si existe la conducta en el repertorio de conductas del sujeto o si no existen y por lo tanto hay que crearla porque los mecanismos terapéuticos para uno y otro caso son muy distintos. Hay que preguntarse sobre las circunstancias antecedentes y consecuentes y hay que formular la secuenciación E-R.

Hay que intentar una interpretación de por qué la conducta alternativa o incompatible no está suficientemente desarrollada. Comúnmente lo que se encuentra son dos posibilidades: responsables de que la conducta alternativa no esté suficientemente desarrollada:

- Que la conducta alternativa está seguida por un refuerzo insuficiente: hay consecuencias positivas truncadas (), no suficientemente desarrollada.
- Que la conducta alternativa, siendo una conducta adaptativa, es castigada (C-)

Ambas posibilidades se encuentran con relativa frecuencia en las respuestas alternativas.

Ejemplo de conductas adaptativas no suficientemente reforzadas son todas esas conductas que los padres a los niños no les refuerzan porque dicen que no lo merecen porque lo único que hacen es cumplir con su deber. Carencia de reforzamiento positiva ().

Ejemplo de C-: niños de primeros cursos de EGB que no se integraba con sus compañeros, en los recreos se quedaba con los profesores, éstos le incitaban a ir con sus compañeros, a veces les hacía caso pero cometía torpezas y era regañada por sus profesores y rápidamente volvía con ellos porque cuando estaba con ellos no la reñían.

Hacer un análisis funcional de conducta no es tarea fácil, por ello ha elaborado unas preguntas guía, no tomarlo como algo que da la verdad definitivamente sino como meros indicadores porque puede tener muchas excepciones:

- ¿Puede describirse la conducta problemática como innata o refleja? En el caso que la respuesta sea positiva abusaremos hacia una Rr. Si es negativa guiaremos nuestra indagación por conducta operante.
- ¿Sigue la conducta problemática con un 100% de probabilidad y de forma más o menos inmediata a la

aparición de un estímulo? Si ocurre así toma la forma de un reflejo: indagar si no se trata de Rr. Si no ocurre así indagaremos sobre conducta operante más bien.

- ¿Puede suponerse que los estímulos que siguen de forma inmediata a la conducta sintomática o problemática tienen un valor reforzante para ella? SI: conducta operante. Esto se refiere a que el refuerzo tiene que ser refuerzo para la persona en cuestión. El refuerzo más universal es el dinero pero puede no ser refuerzo para una determinada persona en determinadas condiciones. Si hay carácter autopropiativo hay refuerzo, sino no.
- ¿Se logra aproximación o alejamiento? Si se logra aproximación del estímulo nuestra indagación irá en el sentido de R+, alejamiento R-.

(Esto es en márgenes estrictamente conductistas).

Ejemplos:

Cuando Jaimito se va por la noche a la cama empieza a llorar con fuerza, la madre va entonces inmediatamente a su cama y espera hasta que el niño se duerme.

Jaimito ha sido sorprendido (mintiendo) por su madre, ésta que a menudo notó que su hijo no siempre dice la verdad, le suelta un largo discurso acerca de lo feo que es que él mienta. Escucha con la cabeza baja y promete luego no lo voy a hacer nunca más, la madre lo coge en brazos y le dice que él sigue siendo su preferido. Este proceso se repite en los días siguientes. La madre se asombra de que Jaimito no deje de mentir, pese a sus promesas.

Muy típico esta secuencia en la vida cotidiana y se va elevando la promesa y la conducta problemática.

La madre al lanzarle el discurso *le presta atención* y es un poderoso reforzador positivo, aunque aparentemente parecería negativo, entonces mejor es hacer un buen reforzamiento negativo, hay que ver el *valor funcional del E o R.*

Entrevista: para ver modelos funcional para tratar problemas, dos razones son importantes:

- Método hipotético deductivo experimental: en la entrevista monta hipótesis funcional, deduce la consecuencia contrastable y las pone a prueba, siempre en el diálogo de preguntas y respuestas.
- Al contrastar las hipótesis se siguen dos caminos: buscar evidencias confirmatorias (corroborar nuestras hipótesis), infirmativas (rechazarlas: mejor camino).

En general los examinadores siguen el camino confirmatorio y Schult en cambio intenta rechazarla y ser si la hipótesis resiste a la información.

Ya se ha recogido información: chica que comenzó hace dos meses sintomatología con carácter físico por hipótesis: el estudio médico y el examen médico salió negativo. Hija única que vive con sus padres. Sufre nauseas, y ha vomitado en casa. Se da cuenta antes de que se vaya a poner mala y sale de la clase porque está temblorosa, se sale todos los días del cole porque se siente mal en clase, sólo se siente mal en el cole y no en casa o en el camino a casa. Sintomatología: florida de carácter físico: nauseas, escalofríos. E: si tuviera fobia escolar tendría miedo a ir y ella dice que se siente bien. Frecuencia: todos los días: muy importante para ver la gravedad de la situación, es grave porque afecta a la vida e intereses de la niña porque perderá el curso.

¿Se puede montar un modelo funcional? Todavía no, sintomatología por la vertiente semántica: se da sólo en las clases, todos los días desde hace dos meses.

Entrevista:

Estudiante de COU, hija única. Naturaleza del síntoma. Escalofríos, náuseas, mareos, temblores, etc.: vertiente somática de lo que podría ser un ataque de pánico. A las primeras manifestaciones del síntoma se va a su casa y con ello el síntoma remite. El síntoma se da sólo en las clases. Todos los días desde hace dos meses.

Recurso al ejemplo real y reciente.

Hipótesis funcional:

La respuesta está precedida por una situación de evaluación y esa respuesta ¿va seguida de alguna estimulación? Ella sale de clase y ya no sigue siendo evaluada, desaparece la situación de evaluación.

¿Qué podríamos indagar para rechazar nuestra hipótesis? Si se da en otras situaciones que son las situaciones de evaluación, si es así nuestra hipótesis se viene abajo.

La sucede también en otras ocasiones cuando no está siendo evaluada.

Nuestra hipótesis no sirve, ¿por dónde podríamos seguir nuestra indagación? Por ejemplo por los profesores: *se siente mal en matemáticas y ciencias, y es el mismo profesor en ambas ocasiones, el profesor no la cree cuando dice que se encuentra mal.*

Para tratar de rechazarla : ¿pasa con otros profesores?: *Sí. La hipótesis sigue sin ser la correcta; también la pasa con profesor preferido en su asignatura preferida.*

Hasta ahora no hemos hablado de los compañeros: *al principio estaban preocupados, son muy amables, cariñosos y la consuelan.*

¿Antes de que se produjeran los síntomas los compañeros se comportaban igual o de una forma distinta? Si es igual no existe relación de contingencia y deberíamos rechazar la hipótesis. *Siempre se ha entendido bien con los compañeros, es la delegada desde hace tres años; nota que están pendientes y eso es lo que no querría; con la profesora de lengua siempre ha tenido una buena relación.*

Parece que en la clase no vamos a encontrar una solución y nos vamos al ámbito familiar: *los padres están bastante preocupados, la relación con ellos es muy normal, están ahora más pendientes, antes nunca llamó la atención; la madre está casi siempre en casa cuando llega. Antes la dejaba sola normalmente, no la hacía ningún caso especial.*

¿Hay relación de contingencia? Sí, ha habido un cambio grande en la conducta de la madre.

Carácter autopropiativo del refuerzo: puede no ser para ella un refuerzo positivo.

La chica manifiesta su disgusto por estar enferma y que se estén preocupando por uno. A Schultz le hace dudar sobre le carácter autopropiativo del refuerzo, sobre que la conducta de atención sea un refuerzo para la chica.

La molesta que su madre se preocupe y la cuide tanto: no hay carácter autopropiativo del refuerzo, se le hace insoportable. Nuestra hipótesis vuelve a ser rechazada.

Podíamos seguir investigando por el camino entre la casa y el colegio. *Intenta llegar lo más pronto a casa, cuando está un rato fuera ya está mejor. Antes de esto se juntaban todos a la salida y se iban al bar, ahora ya no lo puede hacer, le da rabia. Antes la cosa no le gustaba tanto, un amigo con el que no se llevaba bien, la perseguía. Ahora ya no sabe cuándo se va a casa y ya no lo tiene que ver.*

Para rechazar esta hipótesis: aunque no le vea al salir de clase ¿le ve de otra manera? Teléfono, portal... Ya no lo ve desde hace dos meses. Nuestra hipótesis sería correcta. Schult se queda con esa hipótesis como hipótesis funcional de trabajo.

Tratamiento: que se enfrente al chico; se la entrenó en juego de roles para enfrentarse al chico y para decirle que no le quiere ver más. La chica se enfrentó a él y el síntoma remitió, seguimiento de dos años con éxito pleno.

Todas las preguntas han tenido un objetivo claro.

Fijación de la línea base:

Línea base: aquél registro de las conductas problemáticas que se hace previo a toda intervención sistemática; previo a la fase de tratamiento.

Línea base: dato de gran importancia: ventajas:

- es el dato esencial para juzgar la eficacia de un tratamiento. Lo juzgamos en la medida que los datos de tratamiento y al finalizarlo se alejen lo más posible de la línea base, entonces habremos tenido éxito: lo hacemos juzgando los datos del tratamiento en relación con los datos de la línea base.
- durante la fase de línea base se obtiene normalmente una cierta cantidad de información que es de gran importancia para las decisiones terapéuticas, puede hacernos decidir no intervenir, dejar un período de observación o cambiarnos el punto de vista de nuestro enfoque.

Problemas comunes en la fijación de objetivos:

- ◆ *motivación del evaluador*
- ◆ *efectos de reactividad*
- ◆ *fiabilidad del evaluador*
- ◆ *variabilidad de las tasas de conducta*
- ◆ *duración de la línea base*

Motivación del evaluador: Evaluador: sujeto que realmente registra la conducta, rara vez o nunca es el psicólogo, la mayoría de las veces es el mismo cliente o personas allegadas al cliente. Necesita una motivación especial en la fase de línea base porque esa fase es odiosa para el cliente porque el cliente no quiere que le registren sus conductas, quiere que le quiten sus problemas, no que le observen sin hacer nada. Esta fase la ve como una fase contranatura. Le motivamos:

- explicándole la importancia que tiene esta fase de línea base, que es la única manera de tener una base objetiva para juzgar el éxito terapéutico
- reforzando su rendimiento, le damos palabras de apoyo y si tiene errores los minimizamos.

Observadores motivados: mejores observaciones.

Efectos de reactividad: sesgo de la observación. Efecto de la reactividad: un sujeto que es observado cambia su manera de actuar habitual al saberse observado y nosotros queremos observamos la manera habitual de actuar del sujeto. Va contra la deontología profesional ocultarle que le observamos, tiene que dar su consentimiento.

Es un efecto efímero, las tasas de conducta remiten a las habituales a condiciones de que el observador se comporte correctamente: esto es que no se convierta en un reforzador diferencial de las conductas del sujeto. Debe convertirse en un mueble, del cual el sujeto no puede obtener claves de ningún tipo.

Cuando el observador participa activamente junto con el sujeto observado: recomendación: que el observador no cambie su modo habitual de actuar frente a los sujetos observados (ejemplo: profesor que observa a sus alumnos mientras da clase) sobre todo frente a las conductas que son observadas y registradas.

Si se comporta correctamente la reactividad remite rápidamente.

Fiabilidad del evaluador: se calcula mediante la concordancia entre observadores. Hay muchos tipos de coeficientes, se dividen dependiendo de si los datos son dicotómicos o no dicotómicos.

No dicotómicos: cuyos valores sean de 0,1 o de más de 1. se usa el coeficiente de correlación de Pearson: se correlacionan los datos del observador 1 con los del observador 2.

Dicotómicos: presencia – ausencia de conducta: 0–1, cálculos de frecuencia de conducta. Es el registro más utilizado. Se usan otros coeficientes: se usan coeficientes de dos tipos: porcentuales (%de acuerdos) y coeficientes de tipo correlacional o cuasicorrelacional.

Si hay que elegir entre uno de los dos, son de mejor calidad los coeficientes correlacionales, no hay excepción, son de mejor calidad porque:

- Controlan los acuerdos posiblemente debidos al azar.
- Actúa como un coeficiente de fiabilidad y se puede operar con él como con cualquier coeficiente de fiabilidad.

Hay 4 resultados posibles en la relación entre la evaluación de dos observadores:

A: el observador 1 ha registrado presencia y el observador 2 ausencia. Celdilla de desacuerdo.

B: celdilla de acuerdos por presencia de conductas.

C: celdilla de acuerdos por ausencia de conductas.

D: celdilla de desacuerdo.

De esta matriz de datos van a salir los distintos coeficientes, ya sean porcentuales o correlacionales.

Coefficientes porcentuales: % de acuerdos:

Coefficientes correlacionales: se actúa con los valores proporcionales de A, B, C, D, es decir, con base 1.

Ejercicio: dos observadores, período de observación de 16 minutos, divididos en 8 intervalos de 2 minutos cada uno. Que puntúen presencia / ausencia de conducta al final de cada intervalo.

Los valores de cada celdilla esperados al azar son el producto de sus márgenes, para B es el producto de 0.50 x 0.62 = 0.31.

" y K se asemejan mucho, normalmente K es un poco más bajo porque es un coeficiente un poco más estricto.

Debe haber un mínimo de 12 puntuaciones, normalmente de 24 a 30 observaciones.

Ejercicio:

Ejercicio: 30 observaciones.

La dificultad está en conseguir dos evaluadores, por eso hay que trabajar en equipos profesionales.

Variabilidad de las tasas de conducta:

Una línea base puede tener una variabilidad de tasas de conducta muy distintas. La línea base no se da en la realidad.

Normalmente tiene una cierta variabilidad, toleramos la variabilidad que nos deje ver el sentido de la línea base.

El problema es cuando la variabilidad de la línea base es extrema. Ejemplo:

No se puede pensar nada porque

no sabemos qué dirección está

tomando

Cuando una línea base aparece con una extrema variabilidad: qué hacer y qué no hacer:

Qué no hacer: promediar los datos contiguos: se va construyendo una línea base que por supuesto va a la horizontal y si seguimos se llega a la horizontal. El problema no se soluciona, se enmascara.

Qué hacer: hay que averiguar qué es lo que está ocurriendo: revisar la calidad de nuestro procedimiento: una baja fiabilidad puede dar una alta variabilidad o una pobre definición de las conductas puede dar también una alta variabilidad de las tasas de conducta. Si una vez revisada la calidad del procedimiento es adecuada, los datos son objetivos, tenemos que intentar averiguar a qué realidad están respondiendo: se recomienda coger las puntuaciones extremas y preguntarnos si tienen algún denominador común por un lado las altas y por otro las bajas, si es así hemos ganado información de valor funcional, entonces podemos eliminar las situaciones bajas, podemos quedarnos en las tasas medias y dejar las altas para un momento posterior.

Duración de la línea base: debe durar lo menos posible, no existe un número, la recomendación es la línea base ha de durar tanto cuanto se encuentre una dirección clara de la curva y esta dirección sea en sentido no terapéutico. Ejemplo:

Línea base en tres sentidos:

- Horizontal o estacionario: si aparece con suficiente claridad y si la tasa de conducta es suficientemente problemática entonces está indicado terminar con la línea base y empezar con la fase de tratamiento.
- Ascendente: significa que cada vez hay más conducta problemática, entonces decimos que es contraterapéutico y terminamos la línea base e intentamos con el tratamiento de invertir el sentido de la curva.
- Descendente: ocurre con frecuencia, el sentido es terapéutico, se va reduciendo la conducta problemática, en este caso no está indicado terminar con la línea base porque se está produciendo un efecto terapéutico, pero no sabemos por qué, debemos continuar con la línea base tratando de indagar por qué se está produciendo ese efecto terapéutico.

Las curvas circulares son extremadamente raras.

Si la línea base se tomase sobre una conducta adaptativa la cuestión sería al revés, cortaríamos con la línea base si desciende de la conducta adaptativa y continuaríamos si está ascendiendo la conducta adaptativa.

Sólo existe una excepción a esto y es cuando la aceleración del incremento o descenso de conducta nos parece insuficiente, aunque descienda introducimos el tratamiento para aumentar la aceleración, pero este es un caso raro.

Fijación de objetivos de tratamiento:

Preocupación tardía dentro de la psicología conductual. Antes se dejaba a la intuición del terapeuta.

Principios a tener en cuenta en el planteamiento conductual de objetivos: dos principios fundamentales:

- La conducta es reversible: puede aprenderse y desaprenderse y volverse a aprender. Existen muy pocos aprendizajes que queden para toda la vida (ejemplo: andar en bicicleta) y no se dan en la clínica, en ésta reina la reversibilidad, un trastorno puede mejorar pero si las condiciones de vida posteriores facilitan la aparición de ese trastorno, vuelve a aparecer y no significa que haya sido mal tratado anteriormente.
- No existe la inmunización en psicología

Estos principios tienen unas consecuencias importantes:

- El paciente debe participar en forma muy activa en su tratamiento, como su propio terapeuta; tiene que comprender qué es lo que le sucede, tiene que saber acerca de su enfermedad.
- Debemos potenciar los recursos del sujeto para evitar las recaídas. Tenemos fortalecedores de la capacidad del sujeto para enfrentarse a un medio adverso, estos fortalecedores se convierten en objetivos inespecíficos del programa de intervención, son comunes a los distintos programas terapéuticos. Estos objetivos inespecíficos son básicamente de tres tipos:
 - ◊ Entrenamiento en habilidades sociales: la mayoría de los trastornos psicológicos se dan en una matriz social.
 - ◊ Entrenamiento en solución de problemas
 - ◊ Entrenamiento en autocontrol: técnica que nos ayuda a distanciarnos un poco del ambiente y por eso podemos prepararnos para enfrentarnos mejor al medio ambiente.

Tratamiento:

- Objetivo específico: problema (alucinaciones, delirio....)
- Objetivos inespecíficos: fortalecedores.

Ventajas de la fijación de objetivos:

- Ayuda a una estimación objetiva del éxito terapéutico de una determinada intervención: el éxito terapéutico, se juzga en relación con la línea base, cuanto más nos alejemos más cerca estaremos de lograr el éxito terapéutico. Mientras más nos acerquemos al objetivo propuesto mejor será nuestro juicio sobre el logro del éxito terapéutico.
- Posibilita una comunicación y una colaboración óptimas entre terapeuta y cliente. El cliente debe saber hacia dónde a de dirigir sus esfuerzos, eso no es algo impuesto por el terapeuta sino que el algo negociado, consensuado.
- ayuda a detectar posibles conflictos en cuanto a los objetivos, por ejemplo puede haber en el programa terapéutico objetivos contradictorios entre sí (ejemplo: máxima autonomía personal y optimización de las relaciones de pareja), pueden ser también contradictorias las consecuencias que se deriven de los objetivos o los objetivos que se planteen el cliente y los objetivos que se plantee el terapeuta.

Dimensiones que debe poseer un objetivo de tratamiento.

Debe poseer tres dimensiones:

- Dimensión tecnológico – instrumental: tiene que ver con la definición operacional del objetivo, con su medición, con el desarrollo de instrumentos para medir el logro de objetivos
- Dimensión conceptual del objetivo: su definición cualitativa, es previa a la dimensión tecnológico – instrumental
- Dimensión ético – valorativa del objetivo: todo objetivo tiene que pasar por un rasero deontológico, hay que valorarlo éticamente. Relevancia personal que tenga el objetivo para el cliente, la relevancia social, qué valores están implicados en el objetivo.

Características deseables de un objetivo de tratamiento:

- Que actúe como criterio, como norma hacia donde debemos dirigir nuestros esfuerzos terapéuticos, debe traducir el criterio de rendimiento aceptable y esto varía según los casos y también puede haber subobjetivos, y cada uno deberá ser definido adecuadamente.
- Que esté referido a rendimiento: un objetivo debe ser capaz de decirnos lo que se espera que el cliente sea capaz de hacer.
- Que esté referido a determinadas condiciones, un objetivo debe describir las condiciones en las cuales el rendimiento debe ocurrir, no debe ocurrir en cualquier condición.

Preguntas a responder en la fijación de objetivos:

¿Quién y qué debe cambiar? A veces es obvio, otras es difícil (ejemplo: relación padre – hijo).

¿Qué expectativas tiene el cliente en relación con las posibilidades de cambio? ¿son realistas los objetivos en relación con las circunstancias concretas con las que nos toca trabajar? Mucho mejor es fijarse objetivos más humildes pero más realistas, donde el éxito terapéutico esté más garantizado.

¿Qué representaciones se hace el cliente y las personas que le rodean acerca de los objetivos? No podemos dar por descontado que es la misma que se hace el terapeuta.

¿Qué disposición tienen las personas cercanas al sujeto para participar en el programa? Debemos contar con colaboradores que en muchas ocasiones no tienen el 100% de toda su disposición para colaborar.

El colaborador tienen que llegar a saber el papel que va a tener en el proceso, lo que se espera de él en la terapia.

¿Qué influencia tienen eventuales experiencias terapéuticas anteriores? (sobre las expectativas de cambio, la relación terapéutica, el curso terapéutico, los objetivos del tratamiento). Los sujetos que vienen a la consulta no vienen por primera vez a la misma, vienen con prejuicios que hay que sacar a la luz, para saber las expectativas que tienen el paciente.

¿Existe en alguno de los puntos anteriores discrepancias o contradicciones en el triángulo formado por el terapeuta, cliente y personas cercanas? Y si existe ¿cómo deben superarse?.

No podemos empezar un programa terapéutico habiendo discrepancias entre este triángulo.

¿Qué consecuencias se esperan de la posible modificación de la conducta? El cambio de conducta va a alcanzar una serie de cambios en su vida, debemos ser conscientes de que puede haber consecuencias negativas y positivas. Hay que preverlas y tomar las medidas necesarias para minimizar las negativas.

La tarea de fijación de objetivos es compleja: hay que consensuar.

Pasos para fijar objetivos:

- Definición cualitativa del objetivo: se determinan los componentes esenciales del objetivo de tratamiento.
- Definición cuantitativa: elegir un instrumento de medición que nos de cuantitativamente la consecución de objetivos.
- Determinar las circunstancias en las que debe aparecer la conducta en cuestión, y viceversa, en qué circunstancias no debe: la conducta es siempre situacional, no debe aparecer siempre. Hay que especificar las situaciones en las que debe aparecer y no.
- Determinar la cantidad mínima de evaluaciones (logro del objetivo conductual) hasta el final del programa.
- Formular las evaluaciones: determinar qué vamos a evaluar, cómo y cuándo vamos a evaluar.
- Examinar la evaluación en cuanto a su validez de contenido, ver hasta qué punto hay adecuación ante los contenidos de la prueba y los objetivos terapéuticos. Hay que dar mucha importancia a la representatividad de la prueba, con respecto al objetivo.
- Fijar la puntuación crítica en la instrumentación usada, para cumplir con la definición del objetivo.

Diseño del plan de intervención:

¿En qué dirección debe producirse el cambio? ¿Se trata de crear una mera conducta? O ¿de aumentar o reducir la frecuencia de una conducta existente? ¿De disminuirla? ¿Se trata de un cambio topográfico (de las condiciones bajo las cuales aparece la conducta)?

¿En qué lugar preciso de los enlaces funcionales se ha de intervenir? De esta respuesta dependerán las medidas terapéuticas.

¿Qué principios y programas de modificación deben emplearse?

Consejos:

- Hacer una exploración de rutina de posibles refuerzos a utilizar. Un programa terapéutico se apoya casi siempre en refuerzos.
- Relación con el diseño del plan de intervención. Hay que estructurarlo. Debemos dividir los objetivos en subobjetivos, los pasos en subpasos, los plazos a cumplir en subplazos. Hay que diferenciar al máximo posible los programas de intervención, esto nos obliga a mejorarlo clasificándolo, objetivándolo, ya que si lo dejamos globalmente existen pasos subjetivos y esto no es bueno; la estructuración de un programa ayuda a la evaluación durante el tratamiento, esto es enormemente importante.

Criterios para seleccionar las conductas a tratar:

Adecuación de objetivos: se le debe dar prioridad a aquellas conductas que están en adecuación con los objetivos básicos propuestos.

Severidad del impacto: severidad del impacto del síntoma sobre la persona o sobre las personas que le rodean. Pueden ser muy severas o no. Si hay una amenaza para la integridad física del sujeto o de los que le rodean, si el síntoma impide el funcionamiento mínimo del sujeto, etc., y si el síntoma supone graves consecuencias (divorcio, pérdida del puesto en trabajo...) hasta la severidad.

Significado personal y social: la conducta más molesta para el cliente o para las personas que le rodean, se le da prioridad a las conductas más molestas.

- La valoración social de la conducta a tratar, se le da prioridad a la conducta con una valoración social más negativa.
- El grado de armonía que existe entre el cambio propuesto y objetivos y valores del cliente y de las

personas alrededor suyo, se le da prioridad a aquellas circunstancias.

- Donde existe mayor armonía entre el cambio propuesto y los objetivos y valores de la persona.
- Ayudan a motivar, comprender y cooperar al cliente, y a las personas que cooperan en el tratamiento y están cercanas.

Relación con el éxito terapéutico: trátase la conducta que sea más fácil de tratar: se tratan las conductas fáciles no tanto por la conducta misma que se trata, su contenido es lo de menos, lo importante es darle al sujeto experiencia de éxito, probarle que puede cambiar su propia conducta porque eso es un motivador extraordinario para que el sujeto siga progresando (nada tiene más éxito que el éxito).

Da prioridad a aquellas conductas cuyo éxito sea más factible o viable: los tratamientos están llenos de limitaciones, en cuanto a las técnicas terapéuticas, ala habilidad del terapeuta, a los recursos del cliente, a la tolerancia que pueda darnos el entorno social, eso hace que podamos tener conductas a tratar cuyo éxito es más viable y factible y conductas cuyo éxito sea menos factible.

Hay que dar prioridad a aquellas conductas donde se suponga un mejor mantenimiento post-terapéutico del cambio: hay conductas que podemos tratar pero que sabemos que con las condiciones adversas del medio es muy posible que se de una recaída, esas las posponemos un poco, en otras conductas ocurre lo contrario y a ellas se las debe dar prioridad.

Relación con otras personas: consecuencias colaterales: tenemos que pronosticar qué consecuencias colaterales va a tener un cambio conductual, sobre sus relaciones sociales, vida personal, valores. Las consecuencias colaterales normalmente son positivas porque la conducta es adaptativa, pero las puede haber de dudosa valoración o que valoremos negativamente, por ejemplo podemos lograr un éxito terapéutico en autonomía personal y poner con ello en problemas las relaciones de pareja. Da prioridad a aquellas conductas que tengan los mejores efectos colaterales posibles.

Todas las conductas tratadas tienen más o menos probabilidad de dar efectos de generalización, que se generalicen a otras conductas o situaciones no tratadas, como los efectos de generalización son beneficiosos dé prioridad a aquellas conductas donde supongamos que se darán los mayores efectos de generalización.

Cuando hay una valoración funcional, causal entre conductas problemáticas por ejemplo una ansiedad social como conducta sintomática que está causada por un déficit en habilidades sociales, tenemos probada la relación causal entre ambos (no siempre es así), ¿a qué conducta le damos prioridad? Al déficit en habilidades sociales porque es la conducta genéticamente determinante, luego va a ser infinitamente más fácil tratar la ansiedad social. de prioridad a la conducta genéticamente determinante y postérguese el tratamiento a la conducta genéticamente determinada.

¿Qué criterios son más importantes? E de severidad del impacto es muy importante y es difícil que otro criterio pase por encima de él, exceptuando éste no está claro que un criterio sea superior a otro, habrá que juzgarlo clínicamente frente a un caso particular, hay que repasar todos los criterios para determinar cuál será más importante en el caso de tratar.

(fin de las tareas en el proceso de indicación)

PROCESO DE CONTROL DE RESULTADOS:

- **Evaluación de los efectos de la intervención**
 - ◊ **Durante la fase de tratamiento**
 - ◊ **Al finalizar el tratamiento**
 - ◊ **Durante el seguimiento**
- **Perfeccionamiento del plan de intervención**

Evaluación de los efectos de la intervención:

Durante la fase de tratamiento: la evaluación durante la fase de tratamiento preocupó tardíamente a los evaluadores, las hacían sólo al finalizar el tratamiento, sin embargo tiene evidentes ventajas:

- Podemos hacer una demostración objetiva al cliente pronta y rápida del éxito que se va logrando. Es muy importante porque nada tiene más éxito que el éxito, es extraordinariamente motivador para el cliente el ver objetivamente cómo su conducta va cambiando en la dirección deseada.
- Demostración pronta de desviaciones de la curva: no tenemos que esperar hasta el final para decir que hemos fracasado, si se desvía significativamente de los objetivos funcionales habrá que tomar las medidas necesarias.
- Con la evaluación durante el tratamiento se detecta con más facilidad los factores responsables de las desviaciones: si demoramos la evaluación es más difícil, no nos acordamos de lo que ocurría. Es más fácil porque en ese mismo momento está actuando el factor perturbador.

Aspectos a considerar durante esta fase:

- Hay que poner los datos a trabajar, se recogen para que den fruto, no para archivarlos; los datos están para servir, hay que realizar curvas, configurar esa información.
- Durante la fase de tratamiento hay que hacer una eventual evaluación de los efectos de generalización, se pueden empezar a dar durante la fase de tratamiento.

Sesgos en la fase de evaluación durante el tratamiento: típicos de la habituación de la observación:

- Sesgo de expectativas del observador: el observador sabe hacia dónde debe dirigirse la conducta y le hace registrar en ese mismo sentido, a veces no corresponde con la realidad. No cabe el recurso a ocultar el objetivo conductual, si ocultamos el observador se hace sus propias expectativas y desfigura la información con expectativas que no conocemos, lo que es peor. ¿Cómo detectarlo? Introduciendo a un observador nuevo y haciendo un registro de fiabilidad, si la diferencia es significativa entre los dos registros y el del observador habitual va en dirección de las expectativas, podemos afirmar que hay sesgo de expectativas del observador. Se ha demostrado que daña poco en situaciones clínicas, hace mucho daño en diseños experimentales, hay que cumplir una condición: que la conducta que se está observando esté clara y precisamente definida: si es así es muy difícil que se de un sesgo por expectativas del observador.
- Sesgo de desplazamiento del criterio: se desplaza el criterio de registro. Por ejemplo: conocemos a un sujeto y vemos que realiza una serie de actos agresivos, nos hacemos más indulgentes a medida que nos habituamos a una persona determinada, como va en sentido terapéutico parece que el sujeto se va mejorando pero no es así. Hay dos maneras de detectarlo: introduciendo un observador nuevo o una segunda manera que requiere que hallamos grabado en video las sesiones de entrenamiento del observador, a veces al ver el video el mismo observador se da cuenta de su desplazamiento y por tanto están en condiciones de corregir.

Cuánto debe durar un tratamiento: debe durar lo que dure lograr los objetivos conductuales; pero el problema surge en el hecho de que existen distintos criterios para juzgar el éxito terapéutico y por lo menos existen dos tipos de criterios:

- Criterio científico: es el criterio que nosotros nos hemos fijado en la fase de determinación de objetivos, cuando se llegue a una determinada tasa de conducta habremos logrado el objetivo.
- Criterio social. Cómo se siente el cliente y cómo lo ven los que le rodean, es subjetivo, pero no por subjetivo despreciable.

A veces los dos criterios van en paralelo (se logra uno y el otro), entonces no hay problema, hay veces que se

cumple uno y no el otro, ejemplo: se logra el criterio científico pero el sujeto sigue sintiéndose mal, o al revés: el sujeto se siente sano y sin embargo falta para llegar al criterio científico. ¿a cuál hay que darle prioridad? A ninguno de los dos; hay que revisar a cada uno de ellos porque podemos habernos equivocado en el criterio científico (no lo pusimos en su justa medida por ejemplo porque lo pusimos demasiado exigente), se puede equivocar el sujeto, sentir que está bien y si lo examinamos ver que tiene unos puntos débiles y eso es frecuente observarlo.

Fase de evaluación al finalizar el tratamiento: se ha convertido en una evaluación más de tantas, se le quita esa gran importancia que tenía antes cuando sólo se evaluaba al final. Es especialmente importante en el sentido de que es una evaluación más global que abarca todo el sujeto y entre los aspectos más notables una evaluación de efectos de generalización, ha perdido su carácter único.

Estrategias para la evaluación del éxito terapéutico al finalizar el tratamiento:

- Juicio del terapeuta: es muy importante
- Autoevaluación del cliente: junto con el juicio del terapeuta se refieren tanto al estado general del sujeto como a los aspectos específicos que han sido objeto de tratamiento.
- Juicio profesional independiente: ejemplo: juicio de otro psicólogo, aparece nuevamente la ventaja de trabajar en equipo. Condición: que haya examinado al cliente antes de iniciar la terapia y lo examine de nuevo al finalizar la terapia (el profesional independiente).
- Juicio no profesional independiente: debe conocer al cliente desde antes de empezado el tratamiento y tener trato con él una vez terminado el tratamiento, puede ser un pariente, pareja.
- Instrumentos objetivos generales: cuestiones de personalidad, también antes / después o listados de problemas.
- Instrumentos objetivos ad hoc: contruidos para el caso particular por ejemplo instrumentos observacionales, nos van dando de una manera objetiva el cambio del sujeto desde la línea base hasta la finalización del tratamiento, son de suma importancia.

Evaluación del seguimiento:

Problema general: la fase de seguimiento es una fase de evaluación de resultados terapéuticos una vez ha terminado el tratamiento. El problema general es un problema de generalización a través del tiempo.

Estrategias para mantener los resultados:

- Refuerzos que quedan manteniendo las respuestas adaptativas logradas, deben ser de fácil accesibilidad para el sujeto, que están en su medio natural y cuya adquisición sea fácil para él. Los refuerzos de más fácil adquisición son las actividades autoreforzadoras (ejemplo: qué bueno soy o actividades de ese estilo). Son las más fáciles porque dependen de uno mismo.
- Integrar la conducta adquirida en otras conductas más complejas, ejemplo: dejar al cliente en clubes donde puede entrenar esas habilidades sociales (en el caso de habilidades sociales).
- Habilidades para enfrentarse de mejor manera a un medio hostil, entrenarle en habilidades sociales, solución de problemas y autocontrol.

Tendencias posibles de los resultados:

Básicamente hay tres tendencias:

- Que se mantengan a través del tiempo. Los resultados del tratamiento se han mantenido:
- Tendencia a incrementarse los logros:

- Tendencia a descender, recaída: en conducta adaptativa es lo peor, en conducta problemática sería lo mejor.

Recaída: introducción de un nuevo bloque de sesiones de tratamiento, no siempre es necesario introducir un bloque completo de sesiones a veces basta con averiguar los aspectos responsables de la recaída e incidir sólo sobre ellos, y en un par de sesiones podemos cambiar la tendencia de la curva y volver a nuestros resultados satisfactorios.

¿Cuántos registros de seguimiento?: hay que tomar más de uno porque en el seguimiento se trata de detectar una curva del comportamiento y para ello necesitamos por lo menos un punto de apoyo.

Los distintos registros de seguimiento no deben estar demasiado espaciados entre sí, porque puede haber habido una recaída y cuando nosotros hemos cogido el segundo registro ya ha habido recuperación por lo que no nos damos cuenta de la recaída.

Registran distintos aspectos por ejemplo si se ha aplicado Desensibilización Sistemática cómo está haciendo el sujeto la relajación. Es importante hacer un registro de los efectos de generalización que puedan ocurrir.

Perfeccionamiento del plan de intervención:

Retrotrae a la fase de durante el tratamiento.

A veces tenemos que hacer una modificación del tratamiento sobre la marcha.

¿Qué puede fallar en el tratamiento?:

- Teoría: por ejemplo el modelo funcional que hemos construido no es el adecuado. Un refuerzo que hemos supuesto como tal no es tal refuerzo.
- Metodología: un índice de aplicación de refuerzos no resulta adecuado, aplicamos un índice de razón variable por ejemplo y no es capaz de mover la conducta. Otro por ejemplo es una baja objetividad de los datos.
- Práctica: por ejemplo tratando a un niño: el padre es el encargado de darle los refuerzos contingentes pero no hemos sido capaces de convencer al padre de la filosofía del refuerzo y a lo mejor le da el refuerzo pero con una mala actitud, negativa.

Hay que revisar estas posibilidades en sentido inverso: primero los defectos de la práctica porque suelen ser los más frecuentes, si están bien luego los de la metodología porque son menos frecuentes, y si está bien revisamos la teoría, porque son los menos frecuentes, entonces ganamos tiempo.

Requiere bastante sentido clínico, no hay que pasarse ni quedarse corto, muchas veces se precipitan a cambiar el programa y no hay que precipitarse porque la conducta tiene su tiempo, en algunas tardan más en cambiar de lo que pensábamos; pero si ya nos hemos convencido de que realmente las cosas no funcionan hay que actuar con decisión porque mantener un programa que está fracasando, está dando experiencia de fracaso al cliente, parientes y psicólogo, y las experiencias de fracaso son indeseables y cuanto antes se corrijan mejor.

Tener que cambiar un programa es siempre indeseable pero el grado puede estar desde leve hasta intenso, depende de distintos factores:

- De la imagen que se haya dado del psicólogo: mago, señor infalible: la experiencia de fracaso va a ser desastrosa, cuanto más se mitifique al psicólogo, peor va a ser la experiencia de fracaso.
- Programas de tratamiento: son científicos y por ello es susceptible de error, porque se trabaja con hipótesis y éstas pueden fallar, pero al ser el psicólogo un científico y el programa elaborado

científicamente también lleva sus herramientas de autocorrección. Hay que tomar el fracaso como una eventualidad pero no como algo que termina con todo. La imagen del psicólogo y programa científico es la que menos daño hace a la hora de posibles fracasos.

EVALUACIÓN INFANTIL Y ADOLESCENTE

ASPECTOS ESPECÍFICOS DE LA EVALUACIÓN PSICOLÓGICA INFANTIL:

El niño es traído a la consulta: es típico que los niños no acudan por sí solos a la consulta, sino que son llevados por una persona mayor: padres, parientes, profesores...

El adulto va por propia iniciativa.

Esta diferencia tiene una serie de consecuencias importantes: en la consulta del adulto está legitimada por sí misma, acude por propia iniciativa, se siente mal y pide ayuda; en el niño no pasa esto, la consulta del niño no está a priori legitimada, a lo mejor el adulto tiene razón o no la tiene y el niño está mal referido al psicólogo. Hay que plantearse si la consulta del niño es o no legítima. Hay veces que los padres llevan al niño al psicólogo porque son muy exigentes y le piden un desarrollo muy rápido, el psicólogo tiene que cuestionar la legitimidad de esa consulta. Si es legítima tiene una tarea muy importante que es intentar convencer al niño de que esa consulta es necesaria, que le conviene porque va a quedar mejor.

Si no es legítima tenemos que hacer la misma tarea pero con los padres, convencerles de que no tiene necesidad (eso no es tan fácil).

La perspectiva evolutiva: toda evaluación infantil debe hacerse desde una perspectiva evolutiva, en este periodo la evaluación es especialmente importante, debemos conocer los cursos vitales naturales de las conductas, tiene que conocer muy bien la psicología infantil porque el diagnóstico, la indicación y el pronóstico se hacen desde esa perspectiva evolutiva.

Nosotros determinamos el juicio sobre la intervención o no en unas conductas.

Acentuación de variables ambientales: en todo organismo humano las variables ambientales son importantes, sobre todo desde una óptica conductual, sin embargo en el adulto ya se ha formado la personalidad que mediatiza el ambiente, el ambiente influye sobre el adulto con menor intensidad de lo que influye en el niño, en el cual la personalidad no está todavía formada.

Exploración instrumental / intelectual: le tenemos que dar especial importancia a la exploración de los aspectos instrumentales e intelectuales del sujeto, (instrumentales: motricidad y percepción).

Se recomienda un examen de rutina, aunque sea breve, sobre todo si son niños más pequeños. Es recomendable hacerlo por dos razones:

- En los niños no está hecha todavía una selección en función de estos aspectos instrumentales e intelectuales.
- La niñez y adolescencia se pasan fundamentalmente en una matriz que se llama el colegio y ahí fundamentalmente se requieren aspectos instrumentales e intelectuales, estos aspectos toman una gran importancia, son el vehículo de la socialización de los niños; si tienen deficiencias en estos aspectos ello va a repercutir en la socialización.

MODELO TETRÁDICO:

Psicólogo profesores

Es un modelo que trata de envolver al niño, de cubrir todos los aspectos relevantes de su vida y conducta.

Hay un modelo anterior a éste: **modelo diádico**: sólo la relación entre el psicólogo y el niño: rechazan la información de los padres y profesores. Los partidarios critican a los partidarios del tetrádico de que ven muy poco al niño, porque el tiempo es el mismo y tienen que dedicarle tiempo a los padres y profesores.

El problema está en si el mejor modelo es el más eficiente o aquél en el que el psicólogo está más tiempo con el niño: sin duda el mejor es el más eficiente y en este sentido no cabe duda de que el modelo tetrádico tiene una clara ventaja sobre el diádico, es más eficaz porque en el diádico están el psicólogo y el niño en la consulta, en ese ambiente el psicólogo puede hacer grandes cambios, pero nada nos garantiza que se generalicen a otros ambientes, de hecho se generalizan muy poco; el modelo tetrádico diagnostica y trata la conducta en los medios ambientes propios donde la conducta se da, en el hogar, comunidad y colegio. El tetrádico demuestra mucha más validez ecológica, trabaja en los medios mismos donde la conducta problemática se da.

TAREAS BÁSICAS QUE UN PSICÓLOGO DEBE LLEVAR A CABO EN UNA ENTREVISTA CON LOS PADRES Y PROFESORES.

Informar y motivar: el objetivo es ganarse colaboradores y el principal enemigo es la ignorancia que tienen padres y profesores de los modos de actuación del psicólogo. El cliente trae una serie de expectativas que a veces están bastante descentradas porque no domina adecuadamente lo que es el rol del psicólogo y sus modos de actuación. Debemos combatir esa ignorancia.

Se lucha contra la ignorancia dándole información de nuestro modo de hacer psicología.

Se han elaborado pequeños manuales escritos de forma muy fácil para personas no profesionales (profesores y padres) que explican lo que es un psicólogo conductual, sus maneras de trabajar, etc. por ejemplo: Peine y Howarth: Padres e hijos: problemas cotidianos de conducta.

Ayudan a condición de que se discutan con los padres o profesores, es muy mala táctica decirles que se lo lean y no volver a hablar del libro; es condición indispensable someterlos a una discusión.