



**UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO**

Escuela Nacional de Estudios Profesionales "IZTACALA"



U.N.A.M. CAMPUS  
IZTACALA

**METODOS CUANTITATIVOS EN  
DISEÑOS DE GRUPO**

COI  
31921  
C9  
1984-1

**TESIS PROFESIONAL**

QUE PARA OBTENER EL TITULO DE  
LICENCIADO EN PSICOLOGIA  
P R E S E N T A N

MARGARITA CHAVEZ BECERRA  
ANTONIO E. CASTILLO MERCADO

SN. JUAN IZTACALA, MEXICO

SEPTIEMBRE 1984



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



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

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

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

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

A NUESTROS PADRES .

INDICE GENERAL

Introducción..... 1

I.- Diseños de Grupo. **IZT: 1000378**

- a) Origen y Fundamentación Técnica..... 7
- b) Tipos de Diseños de Grupo..... 29-
- c) Aplicación de los Diseños de Grupo en psicología..... 80

II.- Análisis Numérico.

- A) Pruebas Paramétricas..... 97
- B) Pruebas No Paramétricas..... 99
- C) Cómo elegir una Prueba Estadística.....100-
- D) Lógica General de las Pruebas Estadísticas....101
- E) Su Aplicación a los Diseños de Grupo en Investigación Psicológica.....105-

III.- Correspondencia entre Pruebas de Significancia Estadística y los Diversos Tipos de Diseño.

- 1.- Diseños Aleatorizados.....108-
- 2.- Diseño de Bloques Homogéneos.....120
- 3.- Diseño Intra-Sujeto o de Medidas Repetidas...128

Conclusiones.....130

Bibliografía.....133

## INTRODUCCION.

Los procedimientos o métodos de investigación utilizados en psicología han sido prácticamente el método experimental y los métodos no experimentales.

Los métodos no experimentales se caracterizan esencialmente por la no intervención del investigador en la situación bajo estudio a fin de no provocar cambios, es decir, en ningún momento el investigador tiene un control directo de las variables. Kerlinger 1975, se refiere a este tipo de estudios como investigaciones ex post facto o estudios de campo. Arnau 1978, incluye dentro de los estudios de campo el llamado método natural o cuasiexperimental (Campbell y Stanley 1973), a los métodos diferenciales o selectivos y al método correlacional.

Las investigaciones que utilizan el método experimental se centran en el estudio de la relación entre eventos, planteando un control de todos los posibles factores ajenos a la situación experimental y un control directo por parte del experimentador de la(s) variable(s) que manipulará para observar su(s) efecto(s) sobre el evento en cuestión (VD), teniendo siempre un punto de comparación o contrastación. En el caso concreto de la psicología se trata de observar los efectos que pueden tener los eventos medio ambientales (VI) sobre la conducta de los organismos (VD).

La aplicación de la metodología experimental en el estudio de los fenómenos psicológicos ha seguido básicamente dos estrategias de investigación: Los diseños de grupo o tradicionales y los diseños N=1, operantes o de replicación intrasujetos.

Al seguir o aplicar determinado método de estudio, se plantean procedimientos o técnicas de analizar o evaluar sus resultados. La evaluación o análisis de los resultados de las investigaciones experimentales en psicología se ha efectuado a través de dos procedimientos, los métodos cuantitativos y los no cuantitativos o visuales. Girando en forma general. - La función de ambos, en determinar la presencia de cambios o efectos en la variable dependiente, atribuibles a la manipulación experimental.

La forma más común de una evaluación de tipo no cuantitativo es el análisis visual y la replicabilidad de los resultados, donde, su uso o aplicación ha sido asociado con los diseños intrasujeto o N=1.

El análisis visual se apoya en elementos de la estadística-descriptiva, tales como las tablas y gráficas para el registro y presentación en forma comprensible de los datos, a -- partir de los cuales se procede a determinar si existe una evidencia indicativa de una relación funcional entre los aspectos estudiados.

En cuanto a la replicabilidad, nos referimos a la repetición de la investigación bajo las mismas condiciones con el objeto de verificar los resultados obtenidos en la primera.

Se han establecido dos criterios de significancia para la validez de la intervención en los métodos no cuantitativos:

- 1) Criterio de Significancia Clínico o Terapéutico. - Tiene un fin terapéutico y el cambio conductual va orientado a una validación social, su aplicación se enfoca en los efectos de la variable independien--

te y los cambios que se dan entre la línea base y la intervención.

- 2) Criterio de Significancia Experimental; incluye la comparación de la variable dependiente antes y después de la intervención (VI). Kazdin, Craighed y Mahoney 1981.

En lo que se refiere al empleo de métodos cuantitativos en la evaluación de los resultados, tenemos que este tipo de análisis nos da bases numéricas para la interpretación de los datos, cuantificando las observaciones a los diferentes niveles de medición e interpretándolos por medio de pruebas estadísticas. El análisis cuantitativo se ha ligado en forma exclusiva a los diseños de grupo o tradicionales, sin embargo, en los últimos años se ha incrementado el interés por aplicar los métodos cuantitativos a los diseños N=I ( e.g. Kratochwill, 1978; Kazdin, 1976), perdiendo exclusividad su asociación con los diseños de grupo.

Se han presentado ciertas divergencias en cuanto a la utilización de ambos métodos de análisis; algunos autores; Sidman, 1978; Skinner 1956; en Kazdin 1976, no están de acuerdo en la aplicación de los métodos cuantitativos para el análisis de sus resultados, argumentando que es suficiente el análisis visual y el criterio terapéutico o experimental; además de su fácil aplicación e interpretación y de que permite seguir la secuencia de los datos en las diferentes fases, registrando paso a paso los cambios en la conducta. (Kazdin, Craighed y Mahoney 1981). Sin embargo también, hay que considerar ciertas desventajas en el sen-

tido de que no permiten la evaluación de los cambios cuando éstos no son muy evidentes, cuando los resultados son ambiguos y cuando el investigador no está familiarizado con la intensidad de los efectos de las diferentes variables -- que se están manejando (Kazdin 1976).

Por otra parte, los investigadores que trabajan con estudios grupales han recurrido a los métodos cuantitativos como herramienta primordial en el análisis e interpretación de sus resultados. Kazdin 1976, menciona que los métodos cuantitativos ayudan a percibir los cambios poco notables a través de un análisis visual, por ejemplo, cuando los efectos de la intervención son ambiguos (los datos presentan inestabilidad), los métodos cuantitativos permiten establecer mayor significancia a los datos; además señala -- que el análisis cuantitativo resulta útil cuando:

- a) La intervención no alcanza significancia clínica; el criterio estadístico ayuda a enfocar la orientación que debe dar el investigador en estas intervenciones.
- b) Hay poco control en la investigación, los efectos de la intervención pueden ser reducidos o menos obvios. La variabilidad dentro de las condiciones experimentales debida a factores extraños puede -- ser inicialmente muy notable y después verse reducida, en tales casos la evaluación cuantitativa puede incrementar la precisión para determinar la significancia del efecto.

Kazdin, Craighed y Mahoney (1981), consideran que un análisis de tipo cuantitativo presenta algunas desventajas, tales como:

- a) Obscurece los estados de transición entre una y otra condición experimental, ya que, mientras el análisis visual sigue paso a paso los cambios durante todo el experimento; el análisis cuantitativo se centra en medidas estáticas, sin tomar en cuenta que entre esos períodos también pueden ocurrir cambios de gran importancia.
- b) Los resultados del análisis cuantitativo puede ser irrelevante para los resultados esperados en una situación aplicada, pues en una situación dada puede ser más importante la validación social.

A las dos desventajas anteriores hay que suma la creencia de que no existen reglas claras para seguirlo y que un mal manejo del análisis cuantitativo pueda confundir la significancia de la investigación, así como hacer tediosa y lenta su aplicación.

De este último punto, hay que considerar que más que ser una desventaja lo que ha ocurrido es que algunos investigadores desconocen el manejo y aplicación de los métodos cuantitativos, lo que los conduce a su indebida utilización o a un encasillamiento en cierto número de pruebas, en muchas ocasiones también mal utilizadas que finalmente a lo que conllevan es a un rechazo de tales métodos y a optar por un análisis no numérico.

La mala aplicación de los métodos cuantitativos, no sólo --

se presenta por la carencia de bases y principios tanto -- estadísticos como probabilísticos, sino por la falta de -- vinculación que se ha dado entre los diferentes diseños de investigación y las técnicas de evaluación cuantitativa. -- Definitivamente no depende únicamente de la habilidad en -- cuanto a operatividad de los análisis numéricos por sí mis-- mos, es más importante su utilización partiendo siempre -- del objetivo de investigación y del conocimiento de las ca-- racterísticas del diseño utilizado.

Considerando los comentarios hechos hasta estos momentos -- en torno al análisis cuantitativo, en el presente trabajo -- pretendemos:

- a) Despertar el interés acerca del estudio y aplica-- ción de los métodos cuantitativos en estrecha -- vinculación con los diseños experimentales de -- grupo, partiendo de las necesidades de estos úl-- timos.
- b) Mostrar las implicaciones del tipo de diseño uti-- lizado, tipo de medición, tamaño de la muestra; -- en la selección de la prueba estadística a utili-- zar.

Cabe aclarar que la tesis sólo está dirigida al uso de los métodos cuantitativos en diseños de grupo, no por restarle importancia a los diseños de un sólo sujeto, sino por ser parte de un trabajo más amplio dentro del área de métodos cuantitativos que abarca otros puntos complementarios a -- éste.

## I. DISEÑOS DE GRUPO

### A.- Origen y Fundamentación teórica.

Hacia el siglo XIX la psicología deja de ser estudio filosófico de los fenómenos mentales para centrarse en el estudio científico de esos fenómenos.

Los principios de la psicología experimental comienzan con Fechner y se fijan en 1860, año en el cual se publicó su libro *Elemente der Psychophysik* (Boring, 1979). En este trabajo Fechner determinó los umbrales (sensaciones) y sus diferencias a través de efectuar una serie de medidas repetidas de una respuesta ante diferentes intensidades o diferentes localizaciones de un estímulo presentado a un organismo.

La investigación experimental en los principios de la psicología se llevó a cabo con organismos individuales. Se puede afirmar que todos los pioneros de la psicología experimental (Wundt, James, Watson, Pavlov y otros) utilizaron un solo sujeto en sus investigaciones. Sin embargo, Hersen y Barlow (1976) señalan que "quizá el caso más claro de la utilización de esta metodología, basada esencialmente en el estudio de un solo sujeto es la investigación de Ebbinghaus (1885) sobre la memoria". En su estudio Ebbinghaus adoptó el método expuesto por Fechner (medidas repetidas) al problema de la medición de la memoria; toma la frecuencia de repetición con sílabas sin sentido como condición esencial de la asociación y muestra cómo la repetición o repeticiones pueden usarse como medidas de la memoria (Boring, 1979).

Fue a principios del siglo XX cuando apareció una nueva metodología en la investigación básica y aplicadas en psicología que se basaba en el descubrimiento y medida de las diferencias individuales y que consistió en el desarrollo de diseños con grandes grupos de sujetos, anteponiéndose de esta manera a la metodología establecida hasta esos momentos en la investigación psicológica.

Boring (1979) señala a Maskely, astrónomo real del observatorio de Greenwich como el primero en registrar las diferencias individuales en 1795, más el verdadero descubridor de las diferencias personales fue Bassel, astrónomo de Konigsberg quien se dedicó a investigarlo detalladamente, tratando de encontrar una solución para eliminar ese error. De esta forma desarrollo la ecuación personal la cual en sus orígenes se refería a la diferencia de segundos entre las estimaciones de dos observadores (Anne Anastasi, 1979). Bassel demostró que no había forma de calibrar al observador para eliminar ese error. El problema de dichas diferencias fue heredado a la psicología, puesto que se decía eran de origen fisiológico tema que sería tomado posteriormente por lo que se conoce como Psicología Psicofisiológica.

Es necesario señalar que algunos estudios en torno a la ecuación personal y los métodos utilizados en las mismas propiciaron u originaron dos procedimientos de investigación al surgimiento de la psicología experimental: El experimento de compilación y el experimento de reacción. Este último de bastante relevancia ya que fue muy utilizado en la determinación de los tiempos de varios procesos mentales y fue un método muy utilizado por Wundt en su laboratorio (Boring, 1979).

Adolphe Quetelet (1796-1874) astrónomo Belga estudioso de las diferencias individuales, fue el primero en aplicar la Ley Normal del Error a la distribución de datos humanos, biológicos y sociales, descubrió esos rasgos humanos siguiendo la curva normal (Stilson, 1966) citado en Hersen y Barlow. Se le puede considerar como el primero en aplicar métodos estadísticos a datos humanos, la importancia de estos descubrimientos reside en el énfasis puesto en las diferencias individuales. Este hecho se vió reforzado por Darwin, ambos contribuyeron al desarrollo de los diseños con grupos en la investigación psicológica. Además Quetelet proporcionó elementos para el desarrollo posterior de los métodos estadísticos aplicados a problemas psicológicos y antropológicos inicialmente implementados por Galton y posteriormente desarrollados por Pearson y Fisher.

Los trabajos de Darwin sobre su teoría de la evolución tienen una gran influencia en el surgimiento no sólo de la psicología Británica, sino de la ciencia en general y más específicamente reforzó e incrementó el interés por el estudio de las diferencias individuales. En su libro "El Origen de las Especies" (1859), propone que los diferentes tipos de animales cambian a través de varias series de generaciones y que las similitudes de las estructuras y de las funciones podían interpretarse en términos psicológicos, afirmando con ésto que la evolución ocurre y propone tres mecanismos que implican una variación natural; una selección natural y una herencia.

En suma el Darwinismo hizo incapié en la evolución y adaptabilidad de los organismos, subrayó las variaciones individuales dentro de las especies y la transición de estas variaciones a través de la herencia.

Los trabajos de Darwin seguían como método de investigación la observación naturalista y el experimento simple. Es importante señalar que las observaciones de Darwin de los grupos de investigación, dieran origen a computar errores y a la estimación de medias.

A partir de los postulados de Darwin la teoría de la evolución y el planteamiento de si existía una continuidad entre la mente de los animales y el hombre, Galton se dedicó a -- estudiar las diferencias individuales de las capacidades -- mentales, inventando una serie de pruebas que permitieron -- medir algunas habilidades tales como: La determinación del tono más audible, utilizando un silbato; la estimación de la extensión manual por medio de una barra con una distancia variable, etc. Todos estos estudios orientados a buscar, analizar y mostrar las diferencias individuales.

Galton es considerado el pionero de la Psicología Experimental en Gran Bretaña, intentó aplicar los principios evolutivos de la variación, selección y adaptación al estudio de -- los individuos. Realizó estudios sobre la herencia humana, incluyendo que los individuos emparentados o no, deben medirse objetivamente y en grandes números, a fin de descubrir los grados de parecido que entre ellos existen, con este -- propósito ideó numerosos tests y medidas, y en 1882 estableció su famoso laboratorio antropométrico, con esto dió inicio al uso de test de asociación libre (Anne Anastasi, 1979), y es el primero en emplear métodos estadísticos en la medición de fenómenos mentales basándose en los estudios de Que- telet. Galton percibe la necesidad de técnicas estadísticas

especializadas para tratar los datos sobre las diferencias individuales que reunía estudios sobre la curva de distribución normal (Anne Anastasi). Desarrollo el método de correlación al cual Pearson dió su base matemática.

Se considera a estos dos autores como los que establecieron la investigación estadística a los problemas psicológicos como uno de los métodos fundamentales (Boring, 1979).

Durante los años de 1930's Fisher tuvo una gran influencia en la investigación psicológica, fue quien inventó la mayoría de los procedimientos estadísticos sofisticados usado en la comparación de grupos. Fisher, matemático interesado en la genética, cambió el lenguaje del diseño experimental en investigación psicológica introduciendo términos de agricultura. Contribuyó en la consideración de los problemas de inducción o inferencia, todo esto concerniente a la generalidad, trabajó sobre las propiedades de las pruebas estadísticas haciendo posible estimar la relevancia de los datos de un grupo pequeño con ciertas características del universo. Esto hizo posible los principios psicológicos de generalidad y aplicabilidad, basándose en la estadística apropiada al caso, promedios y variabilidad inter sujetos en la muestra (Herseny Barlow, 1976). Estos principios reforzaron en gran medida la utilización de la aproximación de comparación de grupos en investigación psicológica y dieron origen a la Psicología Diferencial.

La Psicología Diferencial tiene por objetivo la investigación cuantitativa de las diferencias individuales en la --

conducta, se interesa en el análisis, la naturaleza y características de las principales agrupaciones. (Anne Anastasi, -- 1979).

El estudio de tales diferencias tiene un triple propósito: a) ciertos grupos son reconocidos y se reacciona ante ellos como tales en la sociedad contemporánea; b) ayuda a aclarar los problemas básicos de las diferencias individuales en general; c) la comparación de un fenómeno psicológico cuando se produce en diferentes grupos puede contribuir a una más clara comprensión del fenómeno mismo (Anne Anastasi, 1979).

De acuerdo a los propósitos de estudio de las diferencias individuales se puede ver el por qué las investigaciones en psicología diferencial se llevan a cabo sólo con grupos de sujetos, donde el análisis estadísticos resulta su principal instrumento de análisis y evaluación de sus resultados.

La nueva metodología de los diseños de grupo invadió la investigación psicológica durante un largo tiempo sin dejar de lado que hubo excepciones en las cuales se siguió la investigación con un solo sujeto, tal es el caso de los problemas relacionados con la psicofísica y la investigación clínica. Los estudios clínicos durante la primera mitad del siglo -- veinte se efectuaron a través del método de estudio de un caso o casos únicos, metodología considerada como la base para el estudio experimental de casos únicos (Hersen y Barlow, 1976). Finalmente, sólo nos resta señalar la utilización de un solo sujeto en los estudios de la naciente aproximación de la modificación de conducta que se fue desarrollando paralelamente y en forma separada de la metodología dominante en la investigación psicológica (diseños de grupos).

A pesar del gran uso de los estudios de grupo en los estudios psicológicos, este tipo de metodología ha recibido un sinnúmero de críticas sobre todo desde el punto de vista de la investigación clínica aplicada. Las objeciones o limitaciones al empleo de diseños de grupo en la investigación aplicada son señaladas por Bergin y Strupp, 1972 en Hersen y Barlow (1976), quienes consideran que: a) no se toma en cuenta a los individuos que integran el grupo control; b) la dificultad para crear grupos con pacientes homogéneos; c) el promedio grupal no refleja los efectos individuales; d) los diseños de comparación entre grupos no consideran la variabilidad intersujeto.

Se puede considerar que Campbell y Stanley (1973) son los que dan las bases para el desarrollo de los diseños de grupo que se conocen en la actualidad poniendo énfasis en el control de la validez del experimento.

A continuación detallamos los diferentes intentos que se han dado para clasificar a los diseños experimentales, así como los planteamientos teóricos que se han presentado acerca de los mismos.

La aplicación de la metodología experimental en la investigación psicológica ha seguido dos estrategias básicas; los diseños de grupo y los diseños  $N=1$ . Es importante señalar que estas dos estrategias se han desarrollado paralelamente, y cada una ha logrado establecer una línea de investigación que ha conformado a la psicología experimental actual. Así tenemos que en algunos centros de investigación hay una tendencia a trabajar con un solo sujeto o un grupo único, mientras que en otros prefieren trabajar con grupos de sujetos.

Ante esta situación podríamos plantear las siguientes preguntas: ¿En qué criterios se basa el investigador para utilizar uno u otro diseño?, ¿Qué tipo de diseño tiene más validez -- dentro de la investigación psicológica?.

Creemos que el fin de este trabajo no es tomar preferencia -- por uno u otro diseño, ni tampoco discutir el carácter científico de los mismos, ya que esto implicaría cambiar el objetivo de nuestro trabajo, además debemos ser honestos y reconocer que no somos especialistas en el estudio de las ciencias, por lo cual trataremos solamente de señalar algunos puntos -- acerca de lo que se entiende por diseño experimental, así como los argumentos que se han dado para la elección de la estrategia a seguir.

Como lo indica Bunge (1977), el conocimiento científico se da a través de la investigación científica la cual tiene principios generales básicos (leyes, teorías, modelos), pero desarrolla métodos y técnicas muy diversos para los diferentes -- sectores de la ciencia (ver anexo 1, al final de este apartado). Debemos recordar que la ciencia es una, pero debido a -- la extensión y diversidad de los fenómenos científicos se hace necesario estudiarla en diferentes áreas de conocimiento.

El método en la ciencia es impersonal, se corrigió asimismo, es objetivo, los hombres de ciencia usan sistemática y conscientemente el aspecto autocorrectivo del enfoque científico (Kerlinger, 1975).

El método científico formula las reglas que nos permiten pasar del fenómeno natural, de la realidad del mundo exterior a

los símbolos con los cuales se expresan las características - que hemos juzgado importantes, esto es, "el conjunto de procedimientos por los cuales se plantean: a) Los problemas científicos; b) Se ponen a prueba las hipótesis científicas" (Bunge, 1977) o según Sidman, (1978), las preguntas experimentales.

Bernard citado en Hersen y Barlow (1976), en su "Introducción al estudio de la medicina experimental", da pautas para dividir a la investigación científica en dos grupos, el de observación y el de experimentación, dando preferencia al segundo pero sin descartar que ambos se complementan, esto se reafirma cuando dice que "La observación muestra en tanto que la experimentación instruye y que "la experimentación es una observación razonada".

Hasta aquí nos hemos centrado en la metodología experimental, no obstante, es pertinente que aclaremos que la investigación en psicología no sólo se ha hecho por medio de esta metodología, sino también se han seguido procedimientos no experimentales; caracterizados principalmente por la no intervención en forma activa del investigador en la situación bajo estudio.

Dentro de estos métodos no experimentales se encuentran: -- 1) el método de observación natural (Underwood, 1973; Plutchik, 1975; y Arnau, 1978) o estudios cuasiexperimentales - (Campbell y Stanley, 1973); 2) el método correlacional (Plutchik, 1975 y Arnau, 1978) o investigaciones ex post facto -- (Kerlinger, 1975).

El método de observación natural se refiere al proceso de búsqueda, observación y descripción del fenómeno que se estudia, sin que el observador intervenga en la situación, ni modifique el curso natural de los procesos. Este método -- constituyó el instrumento de trabajo propio de los estadios iniciales de las ciencias. En psicología, fue adoptado a -- partir de los estudios de Bernard y Darwin.

Arnau (1978a) nos señala que el método de observación natural ha permitido a los etólogos conocer conductas específicas de los animales, referentes a los hábitos de comida, -- conducta de apareamiento, maternidad y agresión. Por otra parte el autor nos recuerda que Jean Piaget ha dedicado su obra al estudio y desarrollo del pensamiento y lenguaje del niño, empleando para esto el método de observación definido como "método clínico", el cual se considera una variante del método natural.

El método de investigación denominado ex post facto, recibe ese nombre precisamente porque es adecuado para aquellas -- situaciones en que el investigador se propone explicar o es tudiar un fenómeno pasado, esto es, que ya ha ocurrido.

Kerlinger (1975), define a la investigación ex post facto -- como "una búsqueda sistemática empírica, en la cual el científico no tiene control directo sobre las variables independientes porque ya acontecieron sus manifestaciones o por -- ser intrínsecamente no manipulables. Se hacen inferencias sobre las relaciones de ellas, sin intervención directa, a partir de la variación concomitante de las variables"(p.268).

A los interesados en el tema les sugerimos revisar directamente a Kerlinger (ver, bibliografía).

Como último punto al respecto de los estudios *ex post facto* hay que añadir que no se puede tener la veracidad de una relación causal entre las variables estudiadas, sino que regularmente se habla de una relación o correlación entre dos o más fenómenos ya observados o medidos, esta acepción es asumida por Plutchik (1975) y Arnau (1978 a); razón por la cual consideramos los autores lo denominan correlacional.

El método más utilizado en la investigación psicológica ha sido el experimental, ya que este a diferencia de los no experimentales, no se queda en la mera observación, pues implica la manipulación de ciertas variables y la observación de sus efectos sobre otras, de tal forma que hay un control de la situación en la cual se lleva a cabo el experimento. Para Sandler y Davidson, (1977), el método experimental analiza la relación entre eventos e intenta controlar todas -- las posibles condiciones ajenas a la situación, mientras -- se está manipulando una variable determinada (variable independiente) y observándose sus efectos sobre el evento en -- cuestión (variable dependiente).

Plutchik (1975) ,menciona que el método experimental busca a través de diferentes estrategias de planeación, eliminar o minimizar las fuentes de error o sesgo a fin de que puedan establecerse relaciones causales inequívocas. Por otra parte Plutchik, señala al diseño experimental como el instrumento del método experimental en el cual se consideran formas de disponer las condiciones de un experimento de tal manera que se logren las respuestas a las preguntas del caso.

Con el fin de tener una idea más amplia de lo que es un diseño experimental enunciaremos algunas de sus definiciones y caracterizaciones que con mayor frecuencia se han utilizado en la investigación psicológica.

Para Kirk (1972) en Castro (1975), el diseño experimental es un plan de acuerdo con el cual se asignan los sujetos a los diferentes grupos o condiciones experimentales.

Robinson (1976), en Arnau, 1981; lo define en función de dos aspectos: 1) como un procedimiento de asignación de sujetos a las condiciones experimentales, así como a la selección de las técnicas de análisis adecuadas, y 2) como sinónimo de seis actividades básicas: a) la formulación de hipótesis o pregunta experimental, esta decisión depende de la tradición experimental del investigador (tradición hipotético deductiva o inductiva), b) selección de las variables dependientes o independientes adecuadas, c) control de las variables extrañas, d) manipulación de variables independientes y registro de la variable dependiente, e) análisis de datos, f) inferencia de las relaciones entre variable dependiente y variable independiente (conclusiones).

Otra definición de diseño es la que nos presenta Craighead, Kazdin y Mahoney, (1981) la cual considera al mismo como "la manera según se disponga la situación a fin de valorar el efecto de la intervención del tratamiento o de alguna variable independiente". Estructura la situación de tal forma que la causa del cambio de conducta puede ser demostrada sin ambigüedad.

A pesar de la gran diversidad de definiciones del diseño - creemos que la mayoría de estas incluye, tal como lo indica Arnau (1981), a los procedimientos requeridos en una investigación experimental que van desde la formulación de - una hipótesis o pregunta experimental, la asignación de los sujetos a las condiciones experimentales, la selección de - técnicas de control, etc.; hasta las conclusiones.

Finalmente al respecto de la definición del diseño, Arnau (1981), nos dice que el objetivo primordial del diseño experimental es el control de las fuentes de variación extrañas o secundarias a fin de poner de manifiesto de manera - clara la acción de las variables manipuladas, sobre la o las variables observadas y eliminar los posibles cambios no deseados causados por factores extraños.

El control de las variables o factores extraños es sumamente importante, ya que atentan contra la validez de un diseño.

Campbell y Stanley, 1973 plantearon la validez del diseño experimental a través de dos fuentes: Validez Interna y -- Validez Externa. La validez interna es la mínima imprescindible, sin la cual es imposible interpretar el modelo, en otras palabras, es el control que se ejerce sobre las variables externas en los diseños (Castro, 1975), esta -- fuente se relaciona con el nivel del control del mismo.

Hay ocho clases de variables externas que pueden afectar -- la variable interna del diseño: Historia, Maduración, Administración de test, Mortalidad experimental e Intracción entre la selección y la maduración.

La Validez Externa plantea el interrogante de la posibilidad de generalización a otras situaciones, poblaciones, variables de tratamiento y variables de medición (Campbell y Stanley, 1973), se trata de estar seguro de que X variable dependiente ocurrió por la manipulación de determinada variable independiente, o qué efectos fueron causados por la aplicación o intervención de la variable independiente. Los factores que amenazan este tipo de validez son: El efecto reactivo o de interacción de las pruebas, los efectos de interacción de los sesgos de selección y la variable experimental, los efectos reactivos de los dispositivos experimentales y la interferencia de los tratamientos múltiples. Aunque nunca se puede esperar un pleno control de los factores que atentan contra la validez, es importante que se seleccionen diseños ricos en una y otra validez. Para una mejor comprensión de los factores que atentan contra la validez recomendamos consultar a Campbell y Stanley, 1973.

Como lo mencionamos en la introducción, el objeto de este trabajo es presentar en forma vinculada el uso de métodos numéricos en los diseños de grupo, pero tomando en cuenta que hay diferentes diseños experimentales que se utilizan en la investigación psicológica, creemos importante señalar algunos intentos que se han dado para clasificarlos.

Existen dos enfoques metodológicos que han sido la guía de la diversificación de los diseños experimentales: De acuerdo con el contexto en que los diseños experimentales son aplicados, así como también la línea de investigación seguida por el experimentador (tradicción hipotética deductiva y tradición inductiva).

La más conocida de estas clasificaciones es la que predomina hasta estos momentos y divide a los diseños experimentales en diseños tradicionales o de grupo, y diseños de caso único, operantes o de replicación intrasujeto. Como ya se había visto anteriormente los diseños de grupo aparecen en la investigación psicológica a principios de este siglo, y parten básicamente del descubrimiento y desarrollo del estudio de las diferencias individuales el cual utilizó la curva normal para obtener el error promedio de los individuos, basándose en que las distribuciones de frecuencia en la mayoría de los rasgos presentan una agrupación de individuos cerca del centro del campo de variación y en un decenso gradual en el número de casos que se acercan a ambos extremos. Una distribución de este tipo puede representar se gráficamente mediante un polígono de frecuencia o un histograma, así como también mediante el cálculo de medidas de tendencia central y medidas de variabilidad (Anastasi, 1979).

La variación entre un individuo y otro, tanto en sus habilidades como en sus medidas antropométricas dió origen a un gran interés por los estudios de grupo de tal forma que un número grande de observaciones permitiera acercarse al fenómeno real. Las diferencias entre individuos así como la distribución de estas diferencias exigía la comparación de individuos y la descripción de un grupo o población de un todo (Hersen y Barlow, 1976).

De esta forma la estadística fue desarrollando técnicas apropiadas para este tipo de diseño experimental, los cu

les requerían que se cumplieran ciertas suposiciones (normalidad de la distribución, homogeneidad de las varianzas, etc.) estos se optimizaban con el uso de un mayor número de sujetos (Castro, 1975).

Debido a la estrecha vinculación que hay entre los diseños de grupo y la estadística, a estos diseños experimentales -- también se les conoce como diseños estadísticos, más adelante mencionaremos las características de este diseño, así como sus diferentes tipos.

El segundo grupo de diseños experimentales, también se le conoce como no estadísticos o diseños operantes, se caracterizan por utilizar un solo sujeto o un número muy reducido en su experimentación y en oposición a los diseños de grupo no necesariamente han recurrido a procedimientos estadísticos -- como técnica de análisis; el análisis de sus resultados se hace en forma más frecuente a través del análisis no estadístico, el cual implica un simple registro visual de los datos -- (según Sidman, 1978, "criterio por registro"), el registro implica normalmente la comparación de datos gráficos en la forma de líneas (oblicuas) a la altura de columnas de datos bajo diferentes condiciones, (Craighead, Kazdin y Mahoney, 1981).

A pesar de que los diseños de un solo sujeto no han recurrido al uso de pruebas estadísticas, en los últimos tiempos, se han hecho intentos por adaptar métodos estadísticos a estos diseños tales como la teoría de análisis series de tiempo (Arnau Grass, 1981). Esta técnica de análisis apenas comienza a aplicarse en psicología y está dando resultados positivos.

Existen dos tipos de análisis de los datos los cuales se vinculan estrechamente con las dos clases de diseños vistos anteriormente. El análisis numérico de los datos se vincula con los diseños de grupo aunque también puede aplicarse a los diseños de un solo sujeto y consiste en el empleo de procedimientos matemáticos en la comparación de los valores de la variable dependiente, implican probabilidad, se establece una hipótesis, se elige un nivel de significancia, se recogen los datos esenciales y se analizan los datos como soportes neutrales o no soportes de la hipótesis.

El análisis no numérico o no estadístico de datos se vincula con los diseños de un solo sujeto,  $N=1$ . Existen 2 criterios para la evaluación no estadística: el criterio experimental y el criterio terapéutico. El primero se utiliza tanto en análisis estadístico de datos como en no estadístico, se refiere a una comparación entre como es la conducta (cuando se revela en L.B.) y como debería haber sido la conducta si los procedimientos experimentales no se hubieran introducido, o sea la comparación de la variable dependiente antes y después de la intervención de la variable independiente; hay mayor rigidez en la interpretación de los datos.

El criterio terapéutico como su nombre lo indica tiene un fin terapéutico y el cambio conductual va orientado a una validación social, este criterio se emplea y se enfoca básicamente en los efectos de la variable independiente y los cambios que se dan entre la línea base y la intervención.

Como se puede apreciar, este tipo de clasificación de diseños de grupo y diseños de un solo caso se basa en la dicotomía

mía del diseño y la estadística, dando a entender que el uso de la estadística es exclusivo de los diseños de grupo, y - que los diseños de un solo caso no lo pueden utilizar. El uso reciente de técnicas estadísticas en los diseños de un solo caso le quita la exclusividad a los diseños de grupo en el uso de la misma. Debe quedar claro que la estadística no es un diseño experimental, sino solo una técnica de análisis al servicio de cualquier planteo experimental. (Arnau Grass, 1981). "Los diseños experimentales, tanto los de grupo, como los de un solo caso existen independientemente de sus usuarios y de la ausencia, presencia o abuso de los análisis cuantitativos" (Castro, 1976). El uso de las mismas las determina el investigador.

Otro criterio para clasificar a los diseños es el número de variables que intervienen en el mismo, de este modo hay autores que hacen referencia a diseños univariables y multivariables (Castro, 1976); otros mencionan a los diseños bivalentes y multivalentes (Plutchik, 1975), etc., todos basándose en el número de variables manipuladas; no obstante no ampliaremos aquí esta cuestión, pues será detallada en el siguiente inciso.

Por otra parte, Arnau Grass (1981), considera que la clasificación de los diseños debe darse de acuerdo a los criterios: a) el control, y b) la cantidad de variables manipuladas y niveles con que dichas variables actúen (condición experimental). Con estos criterios tenemos que a partir del grado de control que permita el diseño se obtendrá una mayor precisión y efectividad de los efectos de los tratamientos,

esto por un lado, y por el otro la cantidad de variables independientes utilizadas nos va a dar una aproximación más -- real al fenómeno conductual.

Arnau (1981) pone énfasis en los aspectos del propio diseño experimental, independientemente de que se pueda asignar al mismo uno o más sujetos; lo cual quedaría determinado por -- el grado de control y uso de variables independientes que -- crea conveniente usar el investigador, así como también los recursos con que cuenta para hechar a andar su experimento.

Creemos que de estas clasificaciones del diseño la que más -- se conoce en la actualidad, es la primera la cual considera que los diseños se clasifican en diseños de grupo y diseños de un sólo caso o  $N=1$ . Aunque en un principio la investigación psicológica utilizó básicamente un solo organismo, a -- partir de la aparición de los diseños de grupo en el escenario psicológico y a pesar de que este último dominó durante varias décadas la investigación psicológica, ambos enfoques metodológicos fueron desarrollados paralelamente dando lu-- gar a que cada uno por su lado buscara y desarrollara procedimientos y técnicas de control, así como diferentes tipos de análisis de los datos.

Así tenemos que algunos campos de la psicología tales como -- el aprendizaje, la percepción, la memoria, psicología social, etc., la mayor parte de la investigación incluso en la década de los 60's y principios de los 70's se ha llevado a cabo con grupos de sujetos. Por otra parte, la psicología y el -- condicionamiento (análisis experimental de la conducta), ha utilizado en sus experimentos el diseño de un solo sujeto -- (Castro, 1975).

El uso de uno u otro diseño parte de tres puntos; la tradición, el uso de la estadística y los requerimientos de la investigación. En cuanto a la primera tenemos que el número de sujetos se decide en base a la tradición en que se ha desarrollado determinada área de investigación, y la experiencia de la misma. De acuerdo con el contexto en que los diseños experimentales son aplicados, así como también la línea de investigación seguida (tradición hipotético-deductiva y tradición inductiva). La tradición no se centra únicamente en el número de sujetos, sino también en el uso de la estadística, esto ha provocado que algunos autores opongan a los diseños de grupo con los de  $N=1$ , llamándolos también diseños estadísticos y no estadísticos, como podemos ver, la estadística juega un papel muy importante en los diseños de grupo, mientras que en el otro tipo de diseños no se ha utilizado, ya que los investigadores prefieren hacer uso del análisis descriptivo visual y consideran que la estadística sale sobrando.

En cuanto al tercer punto debemos considerar que algunos campos de la psicología requieren de cierto número de sujetos para la ejecución del experimento, así tenemos que en las investigaciones relacionadas con el aprendizaje, motivación, formación de conceptos, percepción, memoria, etc., es posible encontrar un número sustancial de informes llevados a cabo con grupos de sujetos. Por otro lado tenemos la psicofísica y el condicionamiento operante ejecutan sus experimentos con organismos individuales y la toma de múltiples medidas, periódicas y continuas sobre el mismo sujeto. Debemos aclarar que en ambos casos no es tan tajante la cuestión, nos

referimos también a la mayoría de los casos, más no a la totalidad. Este último punto es muy importante porque la validez del experimento depende básicamente de que se tomen en cuenta la mayoría de los factores de validez, esto da lugar a que según el fenómeno que vamos a investigar, utilicemos el diseño experimental que más se apegue a nuestras necesidades, sin plantearnos de antemano la utilización de una u otra clase de diseño, esto también incluye el tipo -- de análisis de datos que se vaya a explicar.

Algunos autores, entre ellos Plutchik, 1975; Castro, 1975; Craighhead et al, 1981; han mencionado algunas razones con respecto al uso de uno u otro diseño.

Razones e inconvenientes.

Razones por las cuales se utilizan diseño de grupo.

1.- Por seguir la indicación derivada de la teoría de la medida en el sentido de que mientras mayor sea el número de observaciones, más se acercarán los datos obtenidos a -- los valores verdaderos del fenómeno medido.

2.- Por la influencia de la estadística en la experimentación, la cual requiere que se cumplan ciertas suposiciones (normalidad de la distribución, homogeneidad de las varianzas, etc.). Estas suposiciones, se cumplen mejor entre mayor sea el número de sujetos utilizados.

3.- Por el tiempo limitado de acceso a los sujetos.

4.- Porque el tipo de diseño lo requiere (esto es para ambos diseños).

5.- Por un intento de que los datos sean representativos de la población estudiada.

6.- Para producir una replicación intergrupos.

7.- Cuando se desarrolla investigación aplicada cuyo --  
objetivo es resolver problemas de grupo de sujetos.

Inconvenientes de los diseños de grupo.

1.- Obscurecen el comportamiento individual.

2.- Oculta los estados de transición.

3.- Distinción entre significación estadística y signi-  
ficación clínica. Es posible que un resultado estadística--  
mente significativo sea inoportuno para fines aplicados.

4.- La consagración de las técnicas estadísticas y de la  
significación, en gran medida se abusa de las mismas conside-  
rándolos como una evidencia indiscutible.

5.- Los supuestos aplicados en el enfoque estadístico, -  
tratan de la naturaleza del material estudiado.

Razones por las cuales se utilizan diseños de  $N=1$  u  
operantes.

1.- Por necesidad de control experimental.

2.- Cuando el experimentador decide hacer una replicación  
intrasujeto.

3.- Por limitaciones de equipo.

4.- Por limitaciones de espacio.

5.- Cuando se estudia un sujeto con características parti-  
culares.

6.- Fenómenos de ocurrencia escasa.

Inconvenientes de los diseños N=1 u operantes.

1.- Los cambios considerados como fiables deben ser obvios y dramáticos, no hay reglas claras para establecer cuando un descubrimiento es fiable.

2.- No siempre es posible establecer tasas constantes - de línea base de respuestas especialmente en investigación - realizada con humanos. Cuando se usan animales se requieren periodos de tiempo más largo para establecer tasas constantes.

3.- Buena parte de la conducta no es reversible, lo cual significa esencialmente que se presentan cambios en la línea base por razones que se relacionan con los antecedentes del individuo en particular.

4.- A muchos de los investigadores no les interesa en - calidad de medida, la tasa de respuesta de presión de palancas, sino que se ocupan de otra clase de conductas (más complejas).

5.- La generalización es limitada.

B) Tipos de Diseños de Grupo.

Como ya se ha mencionado, los diseños también han sido clasificados por el grado de control que logran de los factores que atentan contra la validez de un diseño; diseños pre-experimentales, experimentales y cuasiexperimentales (Campbell y Stanley, 1973). La preocupación en este sentido obedece a que si bien a través de los diseños se busca establecer y conocer la relación o relaciones entre variables, es necesario contar con un mínimo de confiabilidad y validez que asegure que los cambios

observados en el fenómeno estudiado (VD) se deben efectiva-- mente a la variable o variables introducidas (VI) y no a fac-- tores extraños (variables extrañas) que afectan a la VD; ha-- ciéndola variar. Si esto último sucediera los resultados es-- tarían contaminados, es decir no habría ninguna certeza para afirmar que la variable manipulada es responsable de los cam-- bios en la VD.

Campbell y Stanley (1973), explican y resumen una serie de -- factores, ya mencionados en el apartado anterior, que aten-- tan tanto contra la validez interna como externa de un dise-- ño. Considerando a la validez interna como la exigencia míni-- ma que garantice que la VI manipulada es la que ha originado los cambios observados; y a la validez externa como la que -- está asociada directamente con la generalización: Cuando ha-- blamos de ella es para preguntarnos a que poblaciones, situa-- ciones, variables de tratamiento y variable de medición pue-- de generalizarse este aspecto.

Es pues tomando en cuenta estos aspectos que dentro de su -- mismo trabajo, Campbell y Stanley (1973), indican que aque-- llos diseños que logren un mayor control de los factores de invalidez, particularmente interna, son los que deben consi-- derarse como propiamente experimentales, así mismo, nos men-- cionan que un diseño experimental se caracteriza por:

- a) la introducción de un grupo control.
- b) La utilización de grupos independientes y equivalentes logrados por aleatorización.
- c) Un buen control de los siete factores de validez interna.
- d) Su presentación en forma de comparaciones diversas -- de una VI con ninguna (o con otros valores de VI, co-- mo posteriormente se hará notar). A partir de este momen--

to emplearemos mucho los términos valores o niveles de una variable independiente, por lo que es conveniente especificar qué entendemos y se debe entender cuando los utilizamos, a fin de esclarecerlo daremos algunos ejemplos:

1) En psicología experimental es ya clásico el estudio de los efectos de la privación de las respuestas de sujetos experimentales, así se puede privar desde 6 hrs., 12 hrs., 24 hrs., etc. En estos estudios la VI es la privación y las diferentes horas son lo que llamamos valores o niveles.

2) Tenemos que en psicología clínica, se cuenta con algunos métodos o técnicas desarrolladas con objeto de disminuir o eliminar la ansiedad, entre uno de estos se encuentra la desensibilización sistemática a la que Joseph Wolpe (1977) señala como uno de los métodos para contrarrestar gradualmente los hábitos neuróticos de la respuesta de ansiedad, la cual tiene variaciones en su aplicación, que se dividen en tres categorías: a) Variaciones Técnicas del procedimiento estandar de consulta en cuarto; b) Respuestas alternativas contra-ansiedad que se emplean con estímulos imaginarios; y c) Métodos que implican la utilización de fuentes exteroceptivas de respuesta de ansiedad. Si en un caso particular un terapeuta elige la desensibilización como tratamiento de la ansiedad, su variable independiente sería precisamente la desensibilización, y si estuviese interesado en la comparación de una de las variaciones técnicas del procedimiento estandar (en especial la de grupo), con el modelamiento por parte de la desensibilización empleando estímulos exteroceptivos; estos dos últimos constituyen los valores o niveles de la VI.

Las características del diseño experimental que acabamos de enunciar y con las que Campbell y Stanley (1973) diferencian un diseño experimental de uno no experimental, son características propias de los diseños de grupo.

- ✦ Los diseños de grupo han sido definidos como aquellos en los que existe un grupo experimental y un grupo control, ambos equivalentes; así como variaciones de medidas (antes y después o sólo después) cuyo objetivo es recoger información que permita evaluar los cambios observados en ambos grupos. Esperando que se presenten cambios notables en el grupo experimental como producto de la intervención a la cual fue expuesto, y que en el grupo control (condición idéntica a la experimental, excepto porque no es sometida a ningún tipo de manipulación) no se presenten cambios o que estos sean poco significativos con respecto a la condición experimental.

Cabe aclarar que el referirse a un grupo experimental y a un grupo control, se hace para efectos de simplificación y convención gráfica en su presentación, más esto no implica en ningún momento que en diseños de grupo exista exclusivamente una comparación entre una única condición experimental y un grupo control; puesto que hay diferentes tipos de diseño que comprenden más de dos grupos de acuerdo al número de valores o niveles en que se esté interesado a estudiar en una determinada variable independiente. Por otra parte, también dentro de esta estrategia hay tipos de diseños que permiten el manejo de más de una VI al igual que una o más VD; diseños que fueron ampliamente desarrollados a partir de la influencia de Fisher sobre la psicología y a los que haremos referencia más adelante.

La introducción del grupo control tiene gran relevancia, ya que un primer nivel cumple la función de grupo de contrastación; - con el que se comparan los resultados del grupo experimental y se evalúan los efectos de la intervención; de tal manera que - si los resultados de ambos grupos al finalizar el estudio son iguales podemos afirmar que la variable manipulada no tuvo ningún efecto. Si por el contrario el grupo experimental muestra gran diferencia con el control, tendremos una amplia seguridad de afirmar que la variable introducida es la responsable de los cambios. Un tercer caso sería cuando la diferencia es muy pequeña, por lo que no habría la plena confianza de señalar a la VI como la causa, y sugerir que los cambios se deben a otros - factores o que la variable o sus valores no son los adecuados. Las situaciones señaladas no son exclusivamente las que pueden presentarse; de acuerdo a las características intrínsecas de - cada estudio pueden presentarse otras más, sin embargo esto - requeriría ya un análisis específico para cada una de ellas. Los aspectos mencionados sólo pretenden ejemplificar lo que - de una u otra manera, un tanto idealizada, mostrarían los resultados experimentales.

En segundo lugar, el grupo control juega un papel importante - en el control de los factores que atentan contra la validez interna.

Con la introducción del grupo control encontramos en Campbell y Stanley (1973) que los factores de invalidez; maduración, - historia, etc., se controlan en el sentido en que los posibles acontecimientos referentes a la historia del individuo, procesos de maduración, etc. que intervienen de alguna manera alterando las medidas finales, serían iguales en ambos grupos, --

puesto que ambos grupos son sometidos a las mismas situaciones, exceptuando la experimental, lo que indicaría que si -- algún proceso de maduración interviniese, sería tanto en el grupo control como experimental, situación que también se -- presentaría con los factores de invalidez. Finalmente, si -- recordamos que la obtención de datos científicos requiere por lo menos una comparación. Los diseños con grupo control incluyen minimamente un punto de comparación.

En diseños de grupo la utilización de grupos equivalentes es fundamental, dicha equivalencia significa partir de grupos -- homogéneos con respecto a los posibles factores que puedan -- interferir, con la intervención. Un procedimiento que se ha considerado muy importante para la formación de grupos equivalentes u homogéneos preexperimentalmente es la aleatorización (1). "Procedimiento que en experimentación se refiere a dos aspectos: 1) La extracción azarosa de una muestra de -- una población, y 2) La asignación azarosa de sujetos a las -- condiciones experimentales" Kazdin (1980).

El primero punto señalado de la aleatorización tiene que ver con la representatividad de la muestra de la población con -- la que se va a trabajar, para lo que es necesario la extracción azarosa; la representatividad está relacionada también con la generalización de los resultados, dado de que si se -- desea inferir de una muestra a una población, o conocer las características de una población a través de una muestra, es necesario dicha representatividad para que esto se pueda hacer -- sin problemas, de no ser así se pueden presentar serios problemas tales como: La heterogeneidad de los participantes,

(1) En Campbell y Stanley (1973) encontramos que la introducción de la aleatorización es uno de los elementos que permite pasar de los diseños pre-experimentales a los propiamente experimentales.

El sesgo de la muestra hacia una característica específica,-- (ejemplo: al escoger una población entre 12 y 22 años, el 80% de los sujetos varían entre 18 y 22 años).

Existen dos formas para determinar la muestra de la población: El muestreo probabilístico y el muestreo no probabilístico. -- Los os tipos de muestreo probabilístico son: a) El aleatorio simple, b) Estratificado, c) Por racimos, d) Sistemático. Las ventajas de este tipo de muestreo reside en que las unidades de observación son seleccionados en forma aleatoria, esto es, cada elemento tiene la misma probabilidad de ser elegido y es posible conocer el error de muestreo.

Los tipos de muestreo no probabilístico son: a) muestreo de -- cuotas, b) muestreo intencional o selectivo. Aunque su utilización se justifica por comodidad y economía, tiene la desventaja de que los resultados de la muestra no pueden generali-- zarse a toda la población.

En lo que respecta a cómo determinar el tamaño de la muestra es conveniente tener en cuenta algunos puntos, tales como:

1) Mientras mayor sea la muestra, se puede alcanzar mayor presentatividad.

2) Antes de decidir el tamaño de la muestra tener bien -- claro los objetivos de la investigación, así como la forma en que se analizarán sus datos.

3) Con la revisión bibliográfica se puede tener una muestra similar a la utilizada en otros trabajos que representan condiciones semejantes o bien calcular el tamaño de la mues--

tra, usando los procedimientos específicos de muestras para estudios sencillos y muestras para estudios complejos.

4) Al establecer el tamaño de la muestra, tener en cuenta el tiempo y los recursos con que se disponen.

5) Tomar medidas sobre problemas que puedan afectar el tamaño de la muestra como: La no representatividad de algunos de los integrantes de la misma (2).

Una vez seleccionada la muestra se asignan los sujetos en forma azarosa a cualquiera de las condiciones experimentales (formaciones de grupos independientes); lo que conlleva a presuponer que todos los factores extraños, posibles (VE), que de una u otra manera podrían influir en los resultados finales, se distribuyen por igual a todas las condiciones, por lo que si afectan o intervienen en una condición, también se manifestarán en las otras. De aquí el significado de equivalencia, expresado inicialmente al abordar este punto.

La equivalencia y el grupo control son dos elementos básicos de los diseños de grupo para el logro de una buena validez interna, por consecuencia, la eliminación de las posibilidades de obtener un experimento contaminado.

Hemos desglosado dos de las características principales de los diseños de grupo: El grupo control y la aleatorización.

(2) Para una mayor extensión sobre el tema "Selección de la Muestra y Cómo Determinar su Tamaño", ver a Rojas Soriano. Guía para realizar investigaciones Sociales. UNAM Méx., 1981. ó Padua J. Técnicas de investigación aplicadas a las C.S.F.C.E. México, 1979.

ción; elementos que como ya se había mencionado, permiten - considerar a los diseños como propiamente experimentales; - los cuales difieren con otros en cuanto al grado y tipo de control que logre, dado que sacrifican la validez externa - por alcanzar una alta confiabilidad interna, caso contra- rio a lo que sucede con los diseños cuasiexperimentales.

Tenemos pues que la aleatorización como selección de una -- muestra de una población y como asignación azarosa a una u otra condición, es vista como un camino confiable para pro- ducir grupos equivalentes preexperimentalmente. No obstan- te es importante conocer que existen procedimientos diferen- tes, que han sido desarrollados con la misma finalidad que la aleatorización; el control de factores extraños y por -- tanto el logro de grupos equivalentes. Es común referirse a estos procedimientos como técnicas de control de varia-- bles extrañas.

Estas técnicas parten de tres principios fundamentales: la eliminación, la constancia y el sujeto como técnica de con- trol o medidas repetidas. La elección de cada una de es-- tas técnicas se hace en base a objetivos y características de la investigación. Es también en base a ellas que los - diferentes tipos de diseños de grupo toman su nombre de - tal forma que encontramos: diseños aleatorios o al azar, - diseños de bloques homogéneos y diseños intra-sujetos o de medidas repetidas.

La eliminación, propone como su nombre lo indica la elimi- nación del factor extraño como tal, de la situación expe--

rimental y resulta muy útil cuando es posible emplearla. Su aplicación resulta más propia y notable en la estrategia de diseños N=1. No obstante en diseños de grupo también es viable solo que en el sentido de seleccionar sujetos con la mayor homogeneidad posible en relación con el factor extraño identificado. Por la forma en que acabamos de exponer se sigue en diseños de grupo, dicho procedimiento queda contemplado dentro de la constancia.

La constancia dice que puesto que en ocasiones resulta dificil y aún en otras más, casi imposible eliminar los factores extraños que interfieren en la situación experimental, - si es posible mantenerlos constantes, esto significa esencialmente que: a) sea cual sea la variable o variables extrañas, un mismo valor o nivel de ella se usa en la selección de los sujetos (Mc Guigan, 1971); que la variable extraña sea introducida dentro del diseño en forma de variable atributiva. (3)

Siguiendo este principio de constancia se propone el bloqueo como una técnica de control de fuentes de invalidez. -- El bloqueo consiste en la agrupación de sujetos que presenten valores iguales en relación con algún o algunos aspectos psicológicos, biológicos o sociales (variables atributiva

- (3) Kerlinger (1975), considera una variable atributiva a - cualquier variable que no se puede manipular o que es - difícil hacerlo. Señala que podemos considerar variables atributivas a todas las que representan características humanas: inteligencia, aptitud, sexo, posición socio-económica, actitudes, etc.



U.N.A.M. CAMPUS  
IZTACALA

vas) que se crea o detecte, mantengan una estrecha relación con la variable dependiente; con el fin de formar grupos lo más homogéneos posibles antes de la aplicación de los tratamientos. El criterio de selección que se establezca es conocido como variable de bloqueo, y para el caso en que se utilicen diseños con sólo dos condiciones, dicho criterio se denomina variable de apareamiento. Para la elección del criterio es pertinente considerar a Edwards (1968), "no se debe seleccionar en base a variables impertinentes o no relacionadas con las medidas que van a tomarse", en Arnau (1978).

**IZT.** 1000378

Una vez determinado el criterio de selección se procede a la agrupación de los sujetos, dando pauta a la formación de lo que se conoce como bloques. Se puede formar más de un bloque de acuerdo a los criterios establecidos (variables extrañas) y a su división o fragmentación. Si se considera que sólo un factor está en estrecha vinculación con la variable dependiente y que pueda influir en ella haciéndola variar, tendremos un sólo bloque dentro del cual habrá una mayor homogeneidad intra-bloque, por tanto entre-sujetos; y dentro del cual se aplicarán los diferentes valores de la variable independiente, dando origen a dos o más condiciones experimentales dentro del mismo bloque, según los niveles de VI de interés. Por otra parte, si la variable de acuerdo a la cual se formaron los bloques se divide en dos o más partes tendremos 2, 3, 4 ó más bloques que responden a la división que se haya hecho de la variable extraña; y dentro de cada bloque se introducirán todas las condiciones, de tal forma que podemos considerar a un bloque como una réplica de los otros, en ---

cuanto a la introducción de las mismas condiciones en todos los bloques. Finalmente cuando se requiere del control de dos variables extrañas, se habla de un doble bloqueo, - y sin dejar de lado que cada una de ellas puede dividirse en dos o más partes, se procede a la formación de los bloques pertinentes.

Una vez formados los bloques sea cualquiera de los tres casos descritos, la asignación de los sujetos a las condiciones experimentales se realiza en forma azarosa.

La manera a través de la cual se aplica el bloqueo como técnica de control, diferencia y rotula a los diferentes diseños dentro de esta categoría de diseños de bloque, así cuando se usa un solo bloque, los diseños se conocen como de -- dos grupos apareados (un sólo criterio de selección). En el caso de que se formen dos o más bloques (una sola VE, pero dividida o fragmentada en dos o más) es común llamarlos diseños de bloques al azar. Finalmente, cuando se cree pertinente o necesario un doble bloqueo (dos variables extrañas) se señalan los diseños de cuadrado latino.

El procedimiento seguido particularmente por cada uno de -- estos diseños de bloqueo, lo ampliaremos posteriormente, en un inciso dedicado explícitamente a éstos.

Antes de abordar la otra técnica de control de factores extraños, queremos una vez más enfatizar que tanto la aleatorización como el bloqueo, persiguen la homogeneización de -

los grupos de sujetos participantes. En el caso de la aleatorización, controlando todos aquellos factores que puedan interferir; en cuanto al bloqueo, controlando una o más variables, determinadas previamente, asegurando así la equivalencia o igualdad pre-experimental. Ahora si consideramos que muchos de estos factores o variables extrañas tienen su origen en las diferencias individuales, afirmamos que la -- aleatorización y el bloqueo intentan mantener constantes o equilibrar dichas diferencias. Es considerado este aspecto e intentando de alguna manera ofrecer, algún procedimiento que permita un máximo de homogeneidad y que evite la contaminación de los resultados experimentales, debido a las diferencias individuales, como se desarrolló la técnica de -- control: el sujeto como su propio control o medidas repetidas; denominándose a los diseños que ocupan dicha técnica -- diseños intra-sujetos o de medidas repetidas.

El procedimiento seguido por dicha técnica, consiste en considerar a un sólo sujeto como un bloque, eliminándose de esta forma la cuestión de las diferencias individuales, dado que evidentemente el bloque tendrá una alta homogeneización al estar formado por un sólo sujeto. De acuerdo a esta técnica habrá tantos bloques como sujetos se determinen necesarios, y cada sujeto será sometido a todas las condiciones experimentales, recogiendo o registrándose datos para cada una de las condiciones experimentales, por las que se ha sometido a cada uno de los sujetos. Un ejemplo de la aplicación de este principio lo constituyen aquellos estudios -- en los que la información proporcionada por un grupo de suje

tos, expuestos independientemente a varias condiciones experimentales diferentes, es más amplia y útil que el sólo observar la conducta ante una sola condición. Tal es el caso de algunas investigaciones con drogas donde en diferentes sesiones y en el mismo grupo, es posible estudiar toda una gama de dosificaciones de una droga. A manera de ilustración, supongamos que se quieren ver los efectos de 3 diferentes dosificaciones de x droga (.5, 1, y 1.5 mg) en la presión sanguínea de 6 ratas. Con esta disposición de que el sujeto sea su propio control tendríamos la siguiente forma de disponer el experimento.

Sujeto	Tratamientos		
	.5	1	1.5
1			
2			
3			
4			
5			
6			

Como posteriormente señalaremos, existen varios procedimientos para determinar el orden de presentación de las condiciones experimentales, lo cual abordaremos al referirnos explícitamente a estos diseños.

Del ejemplo podemos notar que existen una serie de casillas para cada sujeto ante cada condición, estas también indican que hay una medición para cada condición, lo que permite hacer comparaciones entre las diferentes medidas generadas por un mismo sujeto ante cada una de las condiciones a que fue expuesto (intra sujetos) y entre sujetos, lo que hace posible hacer contrastaciones entre uno y otro sujeto, ya-

sea en forma general o particular para cada condición en -- que se tenga interés. Sin embargo el interés de la evaluación se centra en el grupo.

Es fundamental dejar claro que el hacer N número de medidas a un mismo sujeto, correspondientes a cada diferente condición por la que pasó, es por lo que en ambas estrategias de diseños (conductual y de grupos) se habla del uso del sujeto como su propio control, más con una diferencia que cambia totalmente el sentido del concepto en una y otra estrategia. Mientras en los diseños conductuales las mediciones o registros son periódicas y continuas, y en los diseños pertenecientes a la estrategia de grupo las mediciones solo son -- finales para cada una de las condiciones, es decir en los -- primeros las mediciones permiten observar en forma continua la evolución del fenómeno en el tiempo y en los segundos, -- las mediciones son estáticas. Además como ya mencionábamos en los diseños intra sujetos de la estrategia de grupo, el objeto de evaluación final reside en la ejecución por grupo. Las diferencias expuestas son básicas para no confundir la -- aplicación del concepto en una y otra estrategia.

Ya mencionábamos que suele denominarse a los diseños de --- acuerdo a la técnica de control de factores extraños que -- utilicen de tal manera que señalábamos: los diseños al azar; diseños de bloques homogéneos y diseños intra-sujetos. Los diseños que acabamos de enunciar son los que podemos llamar tipos de diseños de la estrategia diseños de grupo, y dentro de cada uno de estos tipos existen variantes, variantes

que obedecen tanto al arreglo de medidas como a la utilización de 2 o más condiciones, lo segundo de acuerdo al número de variables a manipular, así como a los distintos niveles en que se esté interesado.

La introducción o manejo de un número determinado de variables y sus niveles ha desencadenado que se clasifique a los diseños como bivalentes, multivalentes y paramétricos (Plutchik, 1975); univariables: bicondicionales y multicondicionales; multivariables (Castro, 1975); bivalentes, multivalentes o funcionales y factoriales (Arnau, 1978 y 1981).

Partiendo entonces de dos aspectos: la técnica de control y el número y niveles de la VI a manipular, se ha hecho una clasificación de los tipos de diseño, lo cual contemplamos, es producto de tratar de proporcionar una forma clara y ordenada que conlleve a facilitar su presentación, su estudio y su adecuada elección de acuerdo a los objetivos de las diferentes investigaciones.

El término clasificación lo utilizamos en este apartado exclusivamente como se acaba de exponer, en ningún momento como se vio en el punto origen de los diseños, en el que recalcábamos que la clasificación se ha orientado a marcar la separación del diseño en dos estrategias.

Brevemente explicaremos a lo que cada uno de los tres autores citados se refiere con dichas clasificaciones. El motivo por el que se mencionan sólo tres de un extenso número de estudiosos del tema responde a que simplemente se quiere ejemplificar, como se han efectuado las clasifica--

ciones en cuanto a las variables y niveles que intervienen - y considerándolas definir la forma en que se manejará la -- clasificación en el presente trabajo.

Para Plutchik (1975), tanto los experimentos que siguen la estructura de los diseños bivalentes y multivalentes, buscan determinar la relación entre dos variables (una VI y - una VD); más en los bivalentes únicamente existe una comparación entre dos condiciones o dos puntos; grupo experimental con tratamiento y grupo control (ausencia de tratamiento) o el grupo experimental I con un valor de la VI ( $x_1$ ) y el grupo experimental II con otro valor diferente de la VI ( $x_2$ ). En lo que respecta a los multivalentes, requieren - mínimamente el manejo de una VI pero con tres niveles o -- más. Paramétricos, aquellos que inter-relacionan tres variables (2 VI, 1 VD).

Castro (1975) centra su clasificación en diseños univariables y multivariables. Con univariables se refiere a aquellos en los que se manipula una VI y los subdivide en dos tipos: bicondicionales y multicondicionales. Bicondicionales, en los que exclusivamente hay dos condiciones. Multicondicionalmente, en los que se presentan más de dos condiciones, tomando en cuenta que se determinan tres o más niveles de la variable independiente, por lo que cada nivel origina una condición experimental diferente.

Multivariables, diseños en los que intervienen dos o más - variables independientes.

Arnau (1978 y 1981), presenta a los diseños como: bivalentes, multivalentes o funcionales y factoriales. Al igual que Plutchik (1973), señala que los bivalentes son aquellos en los que se comparan dos condiciones. Multivalentes o funcionales, en las que se manipula una VI pero con más de dos niveles (3 o más condiciones). Factoriales -- las disposiciones en las que se estudian simultáneamente la acción conjunta de dos o más VI.

Las clasificaciones en cuanto al número de variables y sus niveles, tomando como referencia a las anteriores ejemplificaciones, inciden en que se deben de marcar una diferencia cuando se trabaja sólo con una VI y dos condiciones, -- una VI y más de dos condiciones; así como cuando intervienen dos o más VI. Sin embargo varían en cuanto a la forma de denominar a tales clasificaciones, dichas variables -- consideramos se deben a que cada investigador persigue -- presentar una clasificación más accesible y entendible para su estudio. No obstante pese a ese interés, se cae en la confusión en términos de manejo, ¿Cuál es cuál?, situación que al parecer tiene su origen en lo que respecta a los valores o niveles de la VI a manipular; y puesto que comúnmente tiende a asociarse el término bivalente, a la determinación de relaciones entre dos variables (esto es una VI y una VD); así como por otro lado, también se asocia multivalente, a la búsqueda de relaciones entre más -- de dos variables lo que implica dos o más VI, así como -- una o más VD; es que nos resulta pertinente englobar dentro de los diseños bivalentes a todos los diseños en los

que sólo se manipula una VI; ya sea que solo se presente - dos condiciones, o que existan más de dos. Para el primer caso, los denominaremos diseños bivalentes de dos condiciones (comparaciones entre x valor de una VI y la ausencia - de tratamiento o entre dos valores diferentes de una VI). Para el caso en el que se maneje una VI con dos o más niveles, serán diseños bivalentes multigrupo o de más de dos condiciones (comparaciones entre dos valores diferentes de una VI y una condición con ausencia de tratamiento o comparaciones entre tres o más valores de una VI). En lo que - respecta a la situación en la que se decida o se crea conveniente la manipulación conjunta de dos o más VI, así como una o más VD, nos referimos a diseños multivalentes, es decir dentro de esta categoría caerán los denominados diseños factoriales.

Una vez especificada la clasificación que haremos aquí de los tipos de diseños dentro de la estrategia diseños de -- grupo, en cuanto al número de variables y niveles, se presentarán haciendo tres divisiones que responden a la técnica de control; 1) Diseños aleatorizados, 2) Diseños de bloques, y 3) Diseños intra-sujeto. Dentro de cada uno de estos tipos señalaremos cuáles son bivalentes (de dos grupos y multigrupo) y multivalentes. (4)

(4) Arnau 1981; dada ya la conveniencia o facilidad de estudio de los diseños, clasificados con respecto a la - técnica de control y al número de variables, propone - que esta clasificación se haga en forma general para - todos los diseños, pertenezcan a una u otra de las ya clásicas estrategias, con el fin de romper precisamente con esa división.

## 1. Diseños aleatorios.

Al utilizar estos diseños la aleatorización o el azar como técnica de control de factores extraños, se presupone que las posibles variables extrañas que intervengan, afectan -- por igual a un grupo como al otro. "Con base en el principio del azar cabe esperar que la influencia de tales variables extrañas se distribuya por igual a todos los grupos y que las medidas de las observaciones sean en promedio equivalentes" (Arnau, 1978).

Resumiendo el procedimiento general de aleatorización a seguir es: a) la extracción de una muestra representativa de la población (ver los procedimientos ya señalados en cuanto al tema), b) la formación de grupos en forma azarosa, esto es asignado a cada sujeto en forma aleatoria a las diferentes condiciones, c) la aplicación de los tratamientos en -- forma independiente a cada uno de los grupos.

Los diseños de este tipo alcanzan una buena validez interna, gracias a dos elementos fundamentales: La aleatoriza-- ción y el grupo control.

### 1.2.- Bivalentes

#### 1.2.1. De dos condiciones.

La estructura básica de estos diseños, implica la formación de un grupo experimental, al que se expondrá a una manipulación experimental y un grupo control que estará bajo las -- mismas condiciones que el experimental, excepto de que no

se someterá a ninguna intervención y con el cual se contrastará el experimental, para evaluar los resultados.

Los diseños de dos condiciones bivalentes aleatorizados incluyen algunas variantes cuya diferencia entre uno y otro estriba en el procedimiento de la introducción de las observaciones o mediciones.

#### 1.2.1.1. Diseño antes y después o Pretest-Posttest.

Su característica esencial es la evaluación de los sujetos antes y después de la intervención.



Las observaciones antes de iniciar el experimento (O<sub>1</sub> y O<sub>3</sub>) - tienen por objeto alcanzar una mayor certeza de que existe -- equivalencia en ambos grupos. Campbell y Stanley (1973) señalan que la medida antes es opcional puesto que la aleatorización presupone la equivalencia de grupos y es suficiente, más si se desea una mayor confiabilidad en dicho aspecto, la medida antes resulta conveniente.

Arnau (1978), indica que la medida antes resulta útil cuando existe interés en conocer, si alguna variable es o no pertinente, esto es, si los grupos han tenido alguna interacción con esa variable y si la hay o hubo, conocer en qué medida.

\* Representación gráfica propuesta por Campbell y Stanley (1973); donde:

X.-representa la exposición del grupo a una variable o acontecimiento experimental, cuyo efecto se medirá.

O.-proceso particular de observación o medición.

R.-indica el procedimiento de aleatorización.

La evaluación de los efectos causados por la introducción de la VI, se realiza de dos maneras; la primera permitirá observar los cambios comparando  $O_2$  con  $O_1$ . La segunda pondría en marcha la función del grupo control, como nivel de contrastación, al confrontar  $O_2$  con  $O_4$ .

Un factor que atenta contra la validez interna de este diseño es la interacción de la medida antes con la final (pretest-postest), que siguiendo el planteamiento general del control de los siete factores de invalidez interna es controlable. No obstante se ha planteado la necesidad de eliminar dicha interacción totalmente por considerarlo negativo para determinados estudios, en los que una interacción de este tipo contaminaría las observaciones finales por consiguiente es necesario proponer otro diseño que ofrezca una alternativa a tal situación, siendo este el caso del siguiente diseño.

#### 1.2.1.2. Diseño solamente después o con Postest únicamente.

El diseño al utilizar una sola medida después, descarta el factor de interacción pretest-postest, originado por el diseño que utiliza ambas medidas.

R	$X_1$	$O_1$
R		$O_2$

La comparación  $O_1$  Vs  $O_2$  permite la evaluación de los efectos experimentales.

### 1.2.1. De más de dos condiciones.

Permiten estudiar una gama más amplia de la variable independiente a través de la formación de tantos grupos como niveles de la VI se manipulen (más de dos). Cada grupo es expuesto independientemente a los diferentes niveles de la VI. También se les considera como diseños de dos grupos extendidos, puesto que permiten el estudio de una VI y toman como base la estructura del de dos grupos. Su representación gráfica varía de acuerdo a los niveles a estudiar, así como en la forma de efectuar las mediciones.

Partiendo entonces de que toman como base a los diseños de dos grupos, los diseños de más de dos condiciones aleatorizados pueden ser con medidas solo después o antes y después; son útiles cuando existe interés de conocer u observar los efectos de diferentes valores o niveles de la VI. Si se desea estudiar tres diferentes niveles de la variable, su representación será:

R		$O_1$
R	$X_1$	$O_2$
R	$X_2$	$O_3$

donde  $X_1$  y  $X_2$  representan los diferentes niveles de la variable y la ausencia de X indica la ausencia de tratamiento. Nótese que aunque solo existe un grupo control de acuerdo a su definición (ausencia de tratamiento), las mediciones de los otros grupos sirven también como niveles de contrastación; --

$O_2$  funge como nivel de comparación con  $O_1$ ;  $O_1$  Vs  $O_3$  y  $O_3$  Vs  $O_2$ .

Recordar que de acuerdo al número de valores a estudiar -- cambiará el diseño y representación.

### 1.2.2.1. Diseño de Cuatro Grupos de Solomon.

Como lo indica su nombre se caracteriza por el empleo de -- cuatro grupos y la combinación de medidas antes y después:

R	$O_1$	X	$O_2$
R	$O_3$		$O_4$
R		X	$O_5$
R			$O_6$

De acuerdo a su representación gráfica, podemos observar -- que integra al diseño antes y después, así como al diseño -- solamente después. Dicha integración permite determinar la interacción pretest-postest al igual que hace posible eva-- luar o medir los efectos de la variable manipulada de cua-- tro maneras diferentes:  $O_2$   $O_1$ ,  $O_2$   $O_4$ ,  $O_5$   $O_6$ ,  $O_5$   $O_3$ , es-- tas comparaciones dan al investigador un margen mayor para afirmar que la VI manipulada es la responsable de los cam-- bios.

Hay que dejar claro que en las dos condiciones experimenta-- les se maneja el mismo nivel de la variable o la misma va--

riable.

Una de las principales ventajas de este diseño es que permite el control del efecto pretest-postest.

### 1.3. Multivalentes.

Los diseños multivalentes estudian la acción conjunta de dos o más variables independientes sobre una o más variables dependientes; además también es posible observar los efectos de cada VI por separado cuando no están relacionadas y se denominan efectos principales. Cada VI, puede tomar dos o más valores de tal manera que cada condición experimental estará constituida por la combinación de los respectivos valores de una VI con los de la otra.

Los diseños que cumplen con estas características son los denominados Diseños Factoriales, cuyo precursor es Fisher quién da ampliamente sus bases y propone de la misma manera sus procedimientos estadísticos de análisis.

En la planeación de un diseño factorial es común elegir una serie de valores, para cada una de las variables independientes elegidas y esto se hace de tres maneras:

a) seleccionando los valores según algún criterio apriorístico (el investigador selecciona los valores de acuerdo a un criterio). Cuando los valores de las variables son elegidos así, el diseño recibe el nombre de modelo fijo o de constantes fijas.

b) cuando los tratamientos, se han elegido al azar, el

diseño se conoce por modelo al azar o de efectos aleatorios.

c) cuando se da el caso en el que algunos valores de la VI, se selecciona en base a "a" y otros, azarosamente, se habla de un modelo mixto.

El optar por valores fijos o aleatorios tiene implicaciones que deben considerarse y serán señaladas más adelante.

Una vez seleccionada la muestra azarosamente, donde su tamaño obedece a los objetivos de la investigación y considerando los puntos sobre cómo determinar el tamaño de la muestra; una cantidad igual de sujetos se asigna a todas las condiciones. Si se quieren observar los efectos de dos VI, y cada una tiene 2 niveles, tendremos un diseño factorial 2x2 que origina 4 condiciones experimentales. El número de números de la operación indica la cantidad de VI manipuladas y el valor de cada número representa los niveles o valores de cada variable independiente.

Un ejemplo de un diseño factorial 2 x 2 sería un estudio en el que estuviese interesado en observar los efectos de la experiencia o inexperiencia del terapeuta, así como la utilización de dos variantes de la desensibilización sistemática en la reducción de la ansiedad.

La variable independiente sería la experiencia de los terapeutas. La VI dos, la desensibilización sistemática, sus niveles las variantes: a) variaciones técnicas del procedimiento estandarizado; b) la utilización de fuentes exterocepti-

vas. Representándose de la siguiente manera:

Variantes de la desensibilización

	2	1
experiencia		
inexperiencia		

De igual forma que se expresó con una operación de produc--  
to el diseño factorial  $2 \times 2$ , es común expresar todos los -  
diseños factoriales, con la variante de que la cantidad de  
números indican las variables que se manipularán y con el -  
cambio de valores de los números, representan sus niveles,  
así encontramos diseños:

- 2 x 3      2 VI, la primera con 2 niveles, la segunda con 3
- 3 x 3 x 3      3 VI, con tres niveles cada una
- etc.

el producto de la operación nos da el número de condiciones  
experimentales. A mayor número de variables, así como ni-  
veles de cada VI, es necesario más condiciones, lo que impli-  
ca mayor cantidad de sujetos, más energía y esfuerzo, más -  
condiciones económicas, etc., razones por las que no es muy  
frecuente la utilización de diseños factoriales que emplean  
muchas VI o que en el caso en que se plantee necesario el -  
uso de muchas VI, se de la alternativa de aplicar técnicas  
que reduzcan el número de condiciones a una fracción de la  
totalidad de los tratamientos. Esta técnica es conocida -  
como replicación fraccionada y "es una reducción del tamaño  
del experimento de forma que se sacrifican algunas combina-  
ciones de tratamientos, con los subsecuentes inconvenientes  
que ello implica desde el punto de vista de la interpreta--

ción de los resultados", Arnau (1981). De lo expuesto anteriormente diremos que en ocasiones se habla de diseños factoriales completos e incompletos. (5)

Para concluir lo referente a los diseños de tipo aleatorizado, es elemental tener presente que ya que se utilizan sujetos diferentes para cada grupo, se habla de grupos independientes logrados por aleatorización. Conocer ésto es fundamental para poder aplicar un correcto análisis estadístico.

## 2. Diseños de Bloques Homogéneos.

Los diseños pertenecientes a este tipo, permiten o buscan la formación de grupos de sujetos homogéneos, persiguiendo con ésto reducir la variabilidad causada por la intervención de variables extrañas.

Si en base a estudios anteriores o a las propias apreciaciones de un investigador, se descubre o cree que existe una alta correlación entre su variable a observar (VD) y alguna otra u otras VE; que influyen en forma decisiva; los diseños de bloques resultan convenientes, puesto que plantean la equivalencia de grupos a través de la medición de la variable o variables extrañas, agrupando a los sujetos que tienen un mismo valor de tal variable.

(5) Para una mayor profundidad en el tema de "Técnicas de Reducción de Condiciones en Diseños Factoriales", ver a Arnau (1981). Diseños experimentales en psicología y educación, Cochran y Cox (1965) diseños experimentales.

## 2.1.- Bivalentes.

### 2.1.1. Diseño de dos grupos apareados.

El diseño permite la comparación o evaluación de dos condiciones experimentales ( $x$  valor de una VI y la ausencia de tratamiento o dos niveles de una VI), donde los grupos han sido formados en base a un criterio, variable de apareamiento, y de acuerdo a la cual se espera que exista una homogeneidad inicial entre los grupos, lo que constituye la base para inferir que cualquier cambio ulterior después de la introducción de las condiciones experimentales se deba a estas últimas. Cuantitativamente una media y desviación típica lo más idéntica posible, indican un buen apareamiento y una igualdad pre-experimental de grupos. Regularmente al seguir adecuadamente el apareo la igualdad de medios ( $x$ ) y desviaciones ( $s$ ) se suponen. En ocasiones cuando existe una estrecha relación de la variable de apareamiento con la VD, se dice que es mejor el uso de estos diseños apareados pues reducen al máximo los cambios que puede producir las variables extrañas y en consecuencia aumenta la validez interna.

Los pasos a seguir para lograr un buen apareo son los siguientes:

- a) la selección azarosa de una muestra de una población.
- b) la selección de sujetos de acuerdo a una variable que mantenga una alta vinculación con la VD. Si el apareamiento se hace con otras variables que no tenga relación con la VD, el procedimiento carecerá de eficiencia.
- c) la asignación azarosa de cada elemento seleccionado a una y otra condición. Un elemento a uno, el segundo a la otra.

- d) asignación o aplicación azarosa de los tratamientos (opcional).

Existen cuatro técnicas por las que se puede efectuar el apareo:

- 1) apareo mediante variable de correlación o de variable - correlacionada.

Se forman los dos grupos con características similares en relación con una variable que esté muy correlacionada con la VD. - Es posible aparear con variables tales como; edad, sexo, educación, etc., siempre y cuando se verifique que tienen una influencia directa en la VD.

- 2) apareo por pares o parejas similares.

Se seleccionan pares de sujetos con igual coeficiente intelectual, edad, rendimiento escolar, u otras variables relacionadas con la VD. Una vez elegido el par se asigna azarosamente un miembro del par a una condición y el otro a la segunda, -- formando así dos grupos equivalentes.

- 3) apareo según un rendimiento previo o criterio de ejecución.

Siguiendo esta técnica los sujetos se entrenan hasta alcanzar el criterio establecido, una vez logrado este criterio se inicia el experimento. Esperar a que todos los sujetos de la muestra alcancen el criterio, implica en muchas ocasiones pérdida - de tiempo, por lo que con frecuencia se opta por una segunda alternativa.

Se estipula cierto número de sesiones pre-experimentales, con -

el objeto de alcanzar un criterio, si al finalizar dicho periodo de entrenamiento algunos sujetos no logran alcanzar el criterio, se descartan del experimento. Una vez que de acuerdo a las dos alternativas planteadas los sujetos alcanzan el criterio, se asignan azarosamente a las dos condiciones.

#### 4) Apareamiento por control acoplado.

Este procedimiento surgió a partir de un experimento específico en el que se planteaba, si la causa de la formación de úlceras en animales era por la presencia de una situación conflictiva (evitar un choque) o por recibir en sí el estímulo - choque. Para poder aclarar la situación se acopló un sujeto - al experimento, es decir había un sujeto experimental y una - control, ambas en idénticas condiciones, excepto que el experimental podía controlar la aplicación de los choques, presionando en un periodo determinado, si la presionaba evitaba el choque sino lo recibía; en cambio el sujeto control no se --- veía ante tal situación, pues el que recibiera los choque dependía totalmente del sujeto experimental, así si el experimental evitaba el choque, no recibía el choque ni él ni el -- control. En resumen el apareamiento por control acoplado con siste en la igualación de condiciones, sólo que en la experimental se da el control de algún estímulo al sujeto.

La formación de los bloques para los diseños de este tipo se puede efectuar siguiendo cualquiera de las cuatro técnicas de apareo. En el diseño de dos grupos sólo se forma un bloque - en base a una VE, y dentro del bloque se forman las dos condiciones experimentales. Al igual que los diseños de dos -- grupos aleatorizados, en el diseño de dos grupos apareados -

pueden variarse las observaciones antes y después o solo después y su esquematización gráfica sería igual exceptuando que en lugar de colocar una R que indica aleatorización, se hace notar que los grupos son apareados.

Diseños de dos grupos apareados con medidas antes y después.

Apareo  $O_1$  X  $O_2$

Apareo  $O_3$   $O_4$

Con medida sólo después

Apareo X  $O_1$

Apareo  $O_2$

2.2.- De más de Dos Condiciones.

2.2.1. Diseño de Bloques al Azar.

El diseño es útil para aquellos experimentos en que es necesario mantener constantes dos o más aspectos (niveles) de un sólo factor extraño que influya en la VD. Cada valor especificado como criterio para la selección de sujetos, origina un bloqueo diferente. De tal manera que si son dos criterios elegidos, habrá dos bloques, si son tres, serán tres bloques, etc. Ejemplo: Si se contempla que la edad es un factor importante a controlar para observar la retención de palabras, "la edad" es la VE a mantener constante, más si se cree que son edades específicas las relacionadas, 10, 12 y 14 años, - estas serán niveles de la variable extraña y en base a esos

niveles se harán tres bloques diferentes. Resumiendo el diseño de bloques al azar mantiene controlado un sólo factor, pero dividido o fragmentado en niveles, a lo que corresponde igual número de bloques.

El diseño es bivalente, porque sólo estudia la relación entre una VI y una VD, a diferencia del de dos grupos, en éste la VI toma más de un valor de donde resultan igual número de condiciones experimentales.

Todas las condiciones experimentales se deben introducir en cada uno de los bloques en un orden azaroso, constituyendo cada bloque una replica del otro.

Se debe asignar azarosamente un número igual de sujetos a cada condición experimental.

La esquematización de un diseño con tres bloques (10, 12 y 14), y tres diferentes tratamientos (niveles de una VI) es:

Bloques	A	B	C
10 años			
12 años			
14 años			

- A.- lista de palabras con poco significado.
- B.- palabras sin significado.
- C.- palabras significativas.

Las filas indican los diferentes bloques, las columnas las condiciones. Observar como en cada bloque se presentan las tres condiciones, pasando a ser cada bloque una réplica del otro. La representación es la forma final en que se acomodarían los datos, pues recordar que el orden de presentación no será siempre

pre así.

Para efectuar un análisis cuantitativo en cada una de las casillas se colocan los resultados de las mediciones finales de cada condición y en cada bloque; acomodadas de tal manera es posible anotar al margen y al pié del cuadro los totales pero -- cada bloque (margen) y para cada condición (pié). Notar en -- base al cuadro, que un diseño con tres bloques y tres niveles de una VI origina nueve subgrupos y por tanto nueve condiciones experimentales. Finalmente hay que señalar que con fines de evaluación es posible hacer comparaciones intra-bloques y entre-- bloques.

### 2.3.- Diseños de Doble Bloqueo.

Permiten el control de dos factores extraños, se caracterizan -- por formar una matriz cuadrada y porque el número de niveles de la variable extraña, es igual al número de tratamientos. El -- cuadrado latino es un ejemplo de estos diseños. A fin de dejar claro el procedimiento a seguir con este diseño, trataremos de ejemplificar su uso en una área de investigación.

Dentro del área de aprendizaje verbal existen tres variables -- fundamentales de investigación, una de ellas es la semejanza y en particular la formal que puede ser ortográfica o acústica -- (6). Tomando esta información; supongamos que en un estudio --

(6) En Dominousky R, y otros (1975) encontramos que la semejanza formal se define mediante algunos tipos de semejanzas físicas reales entre las unidades: a) Ortográficas: cuando la semejanza se refiere a letras idénticas repetidas en las diferentes unidades (ejemplo JZH, ZJQ, QHZ, etc.); b) acústica: cuando la semejanza se define en términos de letras o unidades de sonido semejante (Ejemplo b v, ca, gf, etc.).

con un grupo de estudiantes; 8 de 12 años y 8 de 14, entre los cuales el 50% son varones y el otro 50% mujeres. Para cada grupo hay cuatro mujeres y cuatro hombres y se observará que tipo de semejanza facilita más el aprendizaje verbal.

De acuerdo con este planteamiento la primera variable extraña que se está controlando es la edad, con dos diferentes niveles (12 y 14 años) en base a las cuales se forman los dos primeros bloques colocándolos en las filas de una matriz.

12 años

14 años

La segunda variable que se considera importante mantener constante es el sexo (f y m), por lo que se procede a formar los segundos, dos bloques colocándolos en las columnas.

	F	M
12 años		
14 años		

Determinada de esta manera la matriz, cada columna y fila constituyen bloques diferentes. Como ya se había mencionado las dimensiones de los bloques corresponden a los de la VI, es decir tenemos dos niveles de las variables sexo y edad; y también dos niveles de nuestra VI: semejanza ortográfica y acústica.

El orden de presentación de las condiciones para cada fila y columna se hace dentro de lo posible en forma azarosa, pues se debe cuidar que cada condición sólo aparezca una vez en cada

columna y fila (sobre todo cuando existe un número mayor, de dimensiones tanto de las variables extrañas como de la VI).

La razón por la que el diseño recibe el nombre de cuadrado latino, responde a que antiguamente existía un pasatiempo que — consistía en buscar diferentes colocaciones de unas letras latinas en una tabla cuadrada, donde una letra sólo debía aparecer una vez por columna y fila. La matriz del ejemplo que venimos haciendo, finalmente quedaría representada de esta manera:

	F	M	
12 años	S.A	S.O.	S.A.—semejanza acústica.
14 años	S.O	S.A.	S.O.—semejanza ortográfica.

El número de sujetos para cada condición debe ser igual; si se creyera conveniente seis sujetos por condición se tendría un total de veinticuatro sujetos. Existe la posibilidad de trabajar con un sujeto por condición, no obstante cuando así se hace, el diseño pasa a formar parte de los diseños intra-sujeto o de medidas repetidas.

La evaluación del diseño se puede hacer haciendo comparaciones intra columna o fila (intra-bloques) y entre-columna y fila (entre bloques).

#### 2.4.- Multivalentes.

El bloqueo también es posible pero aquellos diseños en los que se manipulan dos o más variables indiferentes.

### 2.4.1.- Diseños factoriales de Bloques al Azar.

Su estructura básica es idéntica al diseño de bloques al azar, con la diferencia de que es posible estudiar dos o más variables independientes conjuntamente.

Volviendo nuevamente al ejemplo de la retención de palabras - como parte de las investigaciones de aprendizaje verbal, en - que señalábamos la edad como un factor importante a controlar y se manipulaba la significancia de palabras (sin significado, poco significado y muy significativas), y amplificámos el diseño que nos ocupa, utilizando el mismo ejemplo sólo que se -- manipulará una variable más con tres diferentes niveles: el -- tiempo (5 min., 10 min., 15 min.). De esta manera tendremos -- dos VI: a) La significancia y b) El tiempo.

Lo primero que procedería es la selección de sujetos que nos - permitan formar tres bloques, esto es, un grupo de sujetos de 8 años, otro idéntico en tamaño de 10 y 12 años.

Una vez formados los bloques, se decidirá sobre el orden de - presentación de las diferentes combinaciones de tratamientos que constituirán cada condición (recordar que es aleatoria). La representación de la matriz podría quedar así:

$A_3B_3$	$A_1B_1$	$A_2B_2$
$A_2B_2$	$A_3B_3$	$A_1B_1$
$A_1B_1$	$A_2B_2$	$A_3B_3$

donde

$A_1$  sin significado

$B_1$  15 min.

A<sub>2</sub> poco significativo                      B<sub>2</sub> 10 min.

A<sub>3</sub> significativa                              B<sub>3</sub> 5 min.

y para cada condición un número igual de sujetos, por ejemplo cuatro que haría un total de treinta y seis sujetos. La matriz para que sea adecuadamente analizada una vez obtenidos los datos, es conveniente colocarla así:

	B <sub>1</sub>	B <sub>2</sub>	B <sub>3</sub>
A <sub>1</sub>			
A <sub>2</sub>			
A <sub>3</sub>			

Al margen de la columna se podrá obtener el total de cada fila (significancia de las palabras) y al pie el total para cada columna (tiempo).

### 3. Diseños Intra-Sujetos o de Medidas Repetidas.

Los diseños intra-sujetos son todos aquellos en que el sujeto sirve como su propio control en el sentido de que cada uno de los diferentes tratamientos bajo estudio son presentados al mismo sujeto, "Con el fin de generalizar los resultados a toda la población suelen emplear más de un sujeto" (Arnau 1978, p 438). De este modo se tiene un grupo de sujetos, donde cada uno de ellos en forma independiente será sometido a todas las condiciones experimentales, obteniéndose una medición fi-

nal ante cada condición por sujeto, lo que posibilita hacer - comparaciones intra-sujetos y entre-sujetos. No obstante hay que remarcar, que aunque con estos diseños es posible hacer - observaciones intra y entre-sujetos, el objetivo final de la evaluación se centra en el grupo y no en forma individual.

Existen variaciones de estos diseños que dependen del número de tratamientos y de la manera en que se presenten (arreglos u ordenaciones de los tratamientos).

Para evitar confusiones de estos diseños con los conductuales o de replicación intrasujeto, Castro 1975, las denomina diseños de medidas repetidas dentro del mismo sujeto. La razón - por la que es importante no confundirlos ya fue tratada en el momento que expusimos la técnica del sujeto como su propio -- control.

La propia estructura de estos diseños, origina algunos problemas, causados por la exposición temporal de dos o más tratamientos experimentales a un mismo sujeto. Uno de ellos es el llamado "error progresivo" (Underwood, 1973), que incluye tanto los efectos positivos o negativos, originados por la exposición de una variable independiente o un nivel de ella antes de otra, haciendo variar la respuesta, variaciones que harían confusa la acción de la VI. Dentro de los efectos positivos se incluye principalmente el aprendizaje (entrenamiento) y en los efectos negativos se podría tropezar con frecuencia la -- presencia del cansancio, fatiga, etc. Un segundo problema es el llamado efecto residual o persistencia de la acción de los

tratamientos más allá de su aplicación (Arnaud, 1981); quien nos indica que este efecto tiene más probabilidades de ocurrir en aquellos experimentos en que está implicado algún tipo de aprendizaje. El mismo autor, nos indica que en general podemos referirnos a estos dos problemas como "efectos de orden o secuenciales".

A fin de controlar tales factores y evitar que los resultados sean un tanto reflejo de la intervención de estos factores, se han propuesto algunas alternativas.

Habíamos mencionado que cada sujeto, separadamente tiene que ser expuesto a un orden o secuencia diferente de las condiciones, la manera como se puede decidir dicha secuencia es a través del: balanceo o contrabalanceo; determinación azarosa y; pre-estableciendo orden de secuencia y asignando azarosamente a los sujetos a esos órdenes.

Kazdin (1980) hace referencia a que cuando los diferentes tratamientos pueden ser presentados en orden diferente, se está balanceando o contrabalanceando los tratamientos. Es decir se está intentando equilibrar su presentación para así mantener constantes los posibles efectos de orden o secuencia.

El contrabalanceo plantea, la aplicación de los tratamientos en todas las posibles combinaciones y requiere por cada combinación a un sujeto; es decir el número de combinaciones -- posibles de los diferentes tratamientos determina el número de sujetos necesarios. Debido a que la aplicación del prin-

cipio de contrabalanceo, en casos en que se manipula un número extenso de variables o niveles de estas determinan un número grande de sujetos, resulta inconveniente su aplicación; razón por la que se ha planteado un contrabalanceo incompleto, que siguiendo el mismo principio, consiste en presentar las combinaciones de los tratamientos, pero sólo aquellos necesarios de acuerdo a la cantidad de sujetos mar cada por el investigador y sus objetivos.

La determinación de la secuencia de los tratamientos azarosamente es particularmente útil, cuando hay un gran número de sujetos, o cuando el número es pequeño pero el número de tratamientos es grande y no es posible o conveniente balancearlos. Es así como se asignan azarosamente los diferentes tratamientos a cada sujeto, cuidando que no se repita ninguna y que se presenten todas a cada sujeto.

Finalmente, el preestablecer un orden de secuencia y asignar azarosamente a los sujetos a cada diferentes orden, indica que el investigador decida de antemano los diferentes órdenes y una vez que los establezca, asigne azarosamente a los sujetos, independientemente, a cada orden preestablecido.

Una segunda medida para el control especialmente de los efectos residuales es la utilización de intervalos de tiempo entre los tratamientos. Intervalos relativamente grandes, con el fin de que al pasar al siguiente tratamiento los efectos del anterior se hayan eliminado o disipado.

### 3.1.- Diseño Experimental Tratamientos por Sujetos.

Conocido también como diseño experimental simple de medidas repetidas. Su estructura responde a aquellas situaciones en las que sólo se desea manipular una VI con  $x$  niveles, de aquí que de acuerdo a la clasificación que hemos manejado, el diseño se puede considerar bivalente de dos condiciones (comparación entre dos niveles de la VI o un nivel de la VI con una condición control); o bivalente de más de dos condiciones -- (comparaciones entre dos o más niveles de la VI).

El diseño es el modelo de lo que constituyen los diseños de medidas repetidas. Básicamente consiste en la aplicación de todos los tratamientos (niveles de la VI) a cada uno de los sujetos. El orden de presentación de las condiciones regularmente se hace siguiendo el balanceo, si es que el número de condiciones es pequeño. Cuando hay un número mayor de tratamientos se procede a la aplicación del contrabalanceo incompleto o a la asignación azarosa. Un ejemplo de este diseño, se dio en el momento en que presentamos la técnica del sujeto como su propio control.

### 3.2.- Diseño Cuadrado Latino intra-sujeto o de Medidas Repetidas.

Inicialmente señalábamos que uno de los aspectos por los que se presentan variaciones en estos diseños es la manera en que se efectúan los arreglos de presentación de las condiciones. Pues bien el presente diseño obedece a ese aspecto y determina el -- arreglo u orden de presentación, preestableciendo un orden de secuencia y posteriormente asignando azarosamente a los sujetos

a esos órdenes; de tal forma que el sujeto no sea expuesto a una misma condición 2 veces.

Es pertinente su utilización cuando se manipulan más de 2 niveles de una VI o cuando se manipulan dos o más VI, retomando nuestra clasificación pueden ser tanto bivalentes de más de dos condiciones o multivalentes según sea el caso.

La forma de representar el diseño es a través de un cuadrado, donde en las filas se colocan los sujetos y en las columnas los diferentes órdenes de presentación de los tratamientos.

A fin de ejemplificar el diseño, supongamos que se quieren comparar los efectos de 3 procedimientos (3 VI): desensibilización sistemática, implosión y práctica reforzada; en la reducción de la ansiedad de pacientes fóbicos a situaciones sociales, (un grupo de tres).

Si a los diferentes procedimientos de terapia conductual los denotamos:  $X_1$ ) desensibilización;  $X_2$ ) implosión;  $X_3$ ) práctica reforzada; y determinamos 3 órdenes o secuencias de presentación: 1)  $X_2X_1X_3$ ; 2)  $X_3X_2X_1$ ; 3)  $X_1X_3X_2$ , colocados en las columnas; que se presentarán a cada sujeto independientemente, tendremos la siguiente esquematización:

Suje- tos	Ordenes de Presen- tación		
	1	2	3
1	$X_2$	$X_3$	$X_1$
2	$X_1$	$X_2$	$X_3$
3	$X_3$	$X_1$	$X_2$

donde cada sujeto se asigna azarosamente a las diferentes secuencias y pasa por todas las condiciones.

Para la evaluación del diseño se considera la ejecución de -- los 3 sujetos pero como grupo; sin que por esto se niegue que es posible hacer comparaciones intra-sujetos y entre-sujetos.

Posiblemente se pregunta cuál es la diferencia de este cuadrado latino, con el de bloqueo, y diremos que es precisamente la técnica de control de variables extrañas, pues en el cuadrado latino de bloques, grupos diferentes de sujetos son expuestos a condiciones diferentes; en cambio en el cuadrado latino intra-sujetos el número de sujetos se reduce, dado que sólo hay un -- grupo reducido y cada integrante de este, en forma independiente pase por todas las condiciones.

Siguiendo la estructura de los diseños aquí presentados es posible que en diferentes obras especializadas en diseños de grupo señalen otras variantes en cuanto al número de variables y orden de presentación de las condiciones, tales como los diseños contrabalanceados de tratamientos múltiples, diseños alternativos, etc. Debido a que en este apartado dedicado a los -- diseños intra-sujetos cubrimos sus principios y estructura básica y considerando que en esto se fundan todas las variantes que puedan existir de este específico tipo de diseño, podemos sólo dejar mencionado que existen otras variantes cuya diferencia fundamental con los dos aquí presentados sería su denominación.

A manera de conclusión es importante dejar presente que su --

utilización causa efectos de orden y secuencia, que aunque -- existan procedimientos alternativos para mantenerlos constantes, balancearlos, etc.; en ciertas áreas de investigación, - ésto no resulta pues los efectos son fuertes y los intentos - por balancearlos son nulos.

#### 4.- Comentarios sobre los Tipos de Diseño.

En este punto queremos exponer algunas cuestiones que resulta conveniente tener presentes en el momento de la elección de - alguno de ellos, considerando:

4.1.- Grado de control.- Si partimos de que los tres tipos de diseño (aleatorizado, de bloque e intra-sujetos) permiten el control de factores extraños, logrando una buena validez interna, sería válido el uso indistinto de cualquiera de ellos; no obstante, hay que contemplar que aún dentro de esa garantía de validez existen variantes en lo referente a que unos ofrecen mayor control que otros de uno de los factores que - atentan contra la validez interna de un diseño, "las diferencias individuales"; y es la técnica de control empleada por cada tipo de diseño lo que da pauta a ciertas consideraciones en cuanto a grado de control, las que hay que tener presentes en el momento de la elección de un diseño. Aclarando que la elección del diseño no se hace de acuerdo a la técnica de control, sino por los objetivos y situación de investigación.

En primer término hay que recordar que los diseños aleato-- rizados son los que menos control ofrecen, específicamente

de las diferencias individuales. En segundo, que los diseños de bloques alcanzan un mayor control que los anteriores, más las variables de apareamiento deben estar fuertemente -- correlacionadas con la VD. Finalmente que los diseños, señalados como los de mejor control en el sentido de que minimizan los efectos causados por las diferencias individuales entre los diferentes tratamientos, más no entre los sujetos.

Si el diseño elegido corresponde a cualquiera de los dos tipos de diseño que aseguran un mayor control, hay que recordar algunos aspectos que enseguida mencionamos.

4.1.1.- Si se elige un diseño de bloques se debe:

a) Tener la seguridad de que las variables sobre las que se hará el bloqueo, esten realmente relacionadas con la VD. -- En caso contrario, que se aparee con variables no relacionadas el diseño no garantiza el control.

b) Una vez determinada la variable o variables relacionadas, hay que enfatizar qué procedimiento de apareamiento es -- posible llevar con más éxito, teniendo presente que el apareamiento por criterio de ejecución requiere del establecimiento de un periodo pre-experimental, para que los sujetos alcancen el criterio establecido y así todos inicien con el mismo ante tal variable. Definitivamente el apareamiento por criterio de ejecución, implica la planeación de un tiempo más amplio para todo estudio, incrementándose las probabilidades o posibilidades de la pérdida de sujetos, aumento de la historia intra-sesión y el aumento de costos.

c) En el caso que el apareamiento se haga a través de variables correlacionadas o pares de sujetos, hay que señalar -- que por lo general al apareamiento sólo se hace de variables -- atributivas (edad, sexo, coeficiente intelectual, etc.).

d) Un último punto que es importante tener presente, so-- bre todo para el análisis cuantitativo de los resultados en -- que los grupos son correlacionados y que las técnicas de análi-- sis cuantitativo difieren para grupos correlacionados e inde-- pendientes. Se dice que los grupos son correlacionados pues -- existe una relación entre unos y otros sujetos de las condicio-- nes, relación que se propicia con el apareamiento.

e) En el caso en el que se crea más conveniente la utili-- zación de un diseño intra-sujetos, los puntos en los que hay -- que tener cuidado son los efectos de orden y secuencia cuya -- influencia positiva o negativa, podría ocasionar cambios en la VD. Lindquist (1952) --en Arnau, 1978; aconseja que estos di-- seños sólo se apliquen en aquellas situaciones en las que los efectos de los tratamientos son temporales.

f) En lo que se refiere a la elección de un tipo de dise-- ño de acuerdo al grado de control tiene que ver la generaliza-- ción.

Sabemos que al optar por un diseño propiamente experimental se garantiza la validez interna, pero se sacrifica la externa; ca-- racterística de los diseños experimentales que los distingue -- de otros ejemplo: las cuasiexperimentales. Campbell y Stanley (1973) señalan que el cumplimiento de ambos criterios es impor-- tante, sin embargo se contraponen, pues al emplear procedimien-- tos que proporcionan un buen control de los factores extraños,

se repercute negativamente en la validez externa y lo mismo -  
acontecerá si se invierte la situación.

El tema en torno a las implicaciones de priorizar uno u otro tipo de validez, sale de nuestros objetivos; más lo que si es necesario marcar, es que dentro de las limitaciones de generalización que presentan los diseños propriamente experimentales en el sentido de extrapolación de los resultados o situaciones no representadas en las condiciones experimentales o de estudio; creemos que en unos tipos de diseños éstas se agudizan, tal es el caso de los diseños de bloque e intra-sujetos, que permiten un buen control de los factores extraños pero alteran más las situaciones. En los diseños de bloques, tenemos que se busca la homogenización de los grupos a través del apareamiento de algunas variables extrañas específicas, en lo que respecta a los diseños intra-sujetos, los participantes son sometidos a condiciones diferentes y secuenciales; ambos diseños pues, alteran en mayor medida la situación, reduciendo las posibilidades de generalización a situaciones con una máxima similitud a la experimental y por supuesto compatibilidad con los factores de invalidez que se controlaron. Lo anterior también funciona con los diseños aleatorizados, sólo que en éstos en menor medida, pues hay que recordar que en ellos se define una población, se toma una muestra y se forman azarosamente los grupos, presuponiéndose la distribución equitativa de los factores de invalidez y por lo tanto dando más margen a la extrapolación en toda la población de la cual se extrajo la muestra.

4.2.- Número de Variables.- El número de variables queda totalmente al criterio del investigador y de los objetivos que persigue. En este inciso simplemente señalaremos algunos -- comentarios que se han hecho alrededor de cómo se consideran los diseños bivalentes y multivalentes por la información -- que es posible conseguir con su utilización.

Plutchik (1975) comenta que los diseños bivalentes de dos -- condiciones han fungido como: a) Métodos de aproximación al estudio de fenómenos; b) Procedimiento para la identificación de variables y c) Una herramienta para iniciar el estudio de áreas desconocidas; ésto responde principalmente a que al manipular un sólo valor de una VI, la relación que se encuentra se limita a ese valor por lo tanto no es determinante pues obedece a una situación en concreto, además las posibilidades de -- que al repetir el experimento con otro valor diferente o el mismo de la VI se encuentren resultados opuestos a los primeros -- son considerables, de aquí que al utilizarlos hay que recordar que la información obtenida es muy limitada.

Se dice (Arnau, 1978), que los diseños bivalentes con más de -- dos condiciones o multigrupo, permiten especificar más claramente la relación funcional entre las variables al posibilitar el estudio de un rango más amplio de los valores intermedios entre los elegidos de la VI. Su principal desventaja estriba en el -- mayor requerimiento de sujetos al incrementar el número de con-- diciones.

Los diseños factoriales (multivalentes) dan una información más amplia al permitir la evaluación de los efectos de las varia--

bles separadas en un sólo experimento, y única acerca de los efectos combinados de las variables independientes; sin embargo no son útiles en áreas relativamente nuevas de investigación (Kazdin, 1980).

4.3.- Número de Sujetos.- El número de sujetos a utilizar en un experimento es una decisión necesaria y difícil que debe cumplir un investigador. Al iniciar el capítulo señalábamos algunos factores que son importantes y se se deben tener presentes en el momento de tomar la decisión, resumidamente se dijo que la cantidad de sujetos dependerá del fenómeno a estudiar, de los objetivos, de las tradiciones seguidas por determinadas áreas de investigación, etc. No perseguimos en este inciso repetir la información que como ya expresamos se expuso en otro momento; el mencionarlo es para aclarar que no es este el aspecto al que haremos referencia aquí, sino el simplemente marcar la cantidad de sujetos necesarios planteado para cada tipo de diseño según su estructura.

Los diseños completamente al azar que parten tanto de la selección azarosa de una muestra, como de la asignación de los sujetos a las diferentes condiciones de la misma forma, requieren regularmente de muchos sujetos, a fin de asegurar la representatividad de la muestra y garantizar la igualdad o equivalencia de los grupos pre-experimentales, la cual se presupone se logra con la aleatorización al distribuir equitativamente las variables que tienen probabilidad de afectar a la VD. Si los grupos son pequeños se dice que existe la posibilidad de que no sean equivalentes, pues a mayor número de su-



jetos mejor distribución de los factores **extraños**. De acuerdo a esta presuposición teórica, que desde el punto de vista práctico no se logra en un sentido estricto, los diseños completamente aleatorizados se caracterizan por un gran número de sujetos.

**IZT** 1000378

Los sujetos disminuyen en los diseños de bloques en comparación con los aleatorizados y varía de acuerdo a la forma seguida de apareamiento. Si el apareo se hizo mediante un criterio de ejecución hay una gran tendencia a eliminar los sujetos que no alcanzan dicho criterio y en consecuencia sólo se integran azarosamente a los bloques los que le cumplieron, reduciéndose la muestra inicialmente seleccionada. El apareamiento a través de pares similares produce grupos pequeños, - debido a que en muchas ocasiones es difícil localizar o encontrar sujetos con idénticas características. Si el apareo se hace por variables correlacionadas, por lo regular el número de sujetos se determina de antemano en base a las oportunidades que se tengan de conseguir sujetos que cubran el requisito de la variable correlacionada; lo que definitivamente se - facilita si dichas variables son atributivas. En diseños que utilizan el apareamiento es posible obtener buenos resultados con números tan pequeños como cinco o diez (Plutchik, 1975).

Los diseños intra-sujetos caracterizados por la reducción de variabilidad entre las condiciones experimentales, disminuyen al máximo el número de sujetos en comparación con los otros tipos, dado que plantean un número igual de sujetos para cada ordenación diferente de las condiciones.

4.4.- La Forma de Elección de los Valores de la VI.- La -- manera de elección de los valores de la VI, da origen a -- que se hable de un modelo fijo o aleatorio y tienen implicaciones importantes que repercuten en la generalización.

Si se opta por un modelo fijo, se dice que presenta limitaciones en el sentido de que sólo se pueden hacer afirmaciones acerca de los valores de la VI que se manipularon y no a otros. Lo que significa que no se deben hacer inferencias fuera de los valores manipulados, es decir ni interpolaciones (inferir sobre valores intermedios, comprendidos dentro de los presentados) ni extrapolaciones (inferir fuera del rango de valores seleccionados).

Si la elección de los valores se realizó, aleatoriamente, se dice que los valores seleccionados son una muestra representativa de todo un rango marcado por el valor más pequeño y el mayor de los elegidos, por lo que la generalización a valores no manipulados, pero comprendidos en el rango (interpolación) no ofrece limitaciones.

C) Aplicación de los Diseños de Grupo en Psicología.

Después de haber hecho una revisión de los diseños de grupo, creemos conveniente investigar cuál es su influencia actual dentro de la investigación psicológica, ya que según dijimos al principio este tipo de diseños dominaron -- casi en su totalidad la investigación en esta área durante los últimos 50 años.

Para lograr nuestro propósito numeramos títulos de 20 revistas de investigación psicológica de los últimos 3 años (1980, 81 y 82) a excepción de la Revista Mexicana de Análisis de la Conducta del que se revisó el volumen de 1977. Se tomó este criterio porque en las bibliotecas por lo general, todavía no se encuentran todas las revistas de 1983 o en algunos casos no han llegado éstas.

A continuación presentamos los datos que se obtuvieron en los volúmenes mencionados. Para evitar repeticiones agrupamos a las revistas que se encontraron resultados semejantes.

En unas de las revistas que se mencionaron dominaron en su totalidad las investigaciones que utilizaron diseños de grupo, es importante anotar que el análisis de los datos siempre se hizo por medio de métodos numéricos (estadística). +

En cinco de las revistas se observó que en la mayoría de los casos se utilizan diseños de grupos, pero también se incluyen diseños operantes y en algunos casos presentan artículos relacionados con el estudio de los diseños operantes. En el caso de los diseños de grupo, éstos siempre son acompañados de un análisis estadístico de los datos. En cuanto a los diseños operantes, éstos en su mayoría utilizan sólo un análisis visual y la replicación en algunos casos llegan a usar las medidas de tendencia central y de dispersión. ++

En el Psychological Bolletín V.91, 1982, se observó un equilibrio entre los diseños que utilizan grupos y los que utilizan uno o pocos sujetos. Al igual que en el caso anterior, los diseños de grupo utilizan análisis estadístico.

Creemos importante señalar, que esta revista trae una sección dedicada expresamente al estudio de los métodos cuantitativos y su aplicación a la psicología, estos artículos se orientan tanto a diseños de grupo como para diseños N=1 u operantes (análisis de series de tiempo).

En el caso de JABA Y RMAC,<sup>+++</sup> se observó que en su mayoría las investigaciones se llevan a cabo con diseños experimentales N=1 a operantes, y en forma muy reducida utilizan diseños de grupo.

De esta revisión podemos hacer dos consideraciones, la primera de ellas referente al binomio que hay entre los diseños de grupo y el análisis estadístico; y el segundo en cuanto al "predominio" actual de los diseños de grupo en la investigación psicológica.

En cuanto a la primera consideración, ya anteriormente se había mencionado que este tipo de diseño tuvo su origen junto con la estadística, pero posteriormente se targiversó su uso a tal grado que para algunos investigadores tenía más importancia la aplicación de pruebas estadísticas que el propio diseño, abusando de ellas y

desvirtuando su fin.

Debe quedar claro que por las características de los diseños de grupo señalados en los puntos anteriores, el análisis estadístico es un instrumento que le proporciona al investigador los elementos necesarios para darle significación a su investigación.

En cuanto a los diseños operantes se manejan otros índices de medición (análisis visual, replicación), pero en algunos casos también el análisis estadístico inferencial que en la actualidad se usa con poca frecuencia.

En cuanto a la segunda consideración, antes que nada queremos aclarar que el hecho de haber revisado 20 revistas de investigación psicológica no quiere decir que los resultados obtenidos sean del todo completo ya que hay mucho más revistas que no se revisaron por no tener acceso a ellas y porque implicaría mucho tiempo en hacer esta revisión. A pesar de esto creemos que las revistas que se revisaron son las más representativas dentro de la investigación psicológica. Es to en base a las bibliotecas que se visitaron en las cuales se encontró, son las revistas que se consultan con mayor frecuencia y las que hay en existencia.

Podemos concluir sin temer a equivocarnos, que todavía hay predominación de los diseños de grupo en la investigación psicológica, pero ésta no se antepone a otro tipo de diseño, creemos que hay cierta madurez en los investigadores para -

hacer uso adecuado de los diseños, sin tratar de antagoni--  
zar o negar la existencia de uno u otro tipo de diseño.

A N E X O 1

Definición y Clasificación de la Ciencia.

Para comprender mejor este punto creemos que es importante presentar algunas definiciones de ciencia así como su clasificación en forma muy breve.

Bunge (1977), define a la ciencia como un cuerpo de ideas-que puede caracterizarse como un conocimiento racional, -- sistemático, exácto, verificable, perfectible y por consiguiente falible.

Rosenblueth (1981), "La ciencia es el conocimiento ordenado de los fenómenos naturales y de sus relaciones mutuas", "El conocimiento científico se ocupa de los aspectos dinámicos reproducibles de los fenómenos naturales que son invariables en el tiempo y en el espacio, y que son fundamentales para el desarrollo de una sociedad".

Bunge (1977), clasifica a la ciencia a partir del objeto - de estudio en ciencia formal y ciencia fáctica. "La ciencia formal no se ocupa de hechos objetivos, más bien trata de entes abstractos, pero éstos se desarrollan a partir de objetos reales (naturales o sociales)".

"La ciencia fáctica no emplea símbolos vacíos, para confirmar sus conjeturas, necesita la observación y/o el experimento, exige que las conjeturas sean verificables en la -- realidad".

La lógica y la matemática por ocuparse de entes abstractos o formales se consideran dentro de la ciencia formal. Debe quedar claro que esta clasificación es sólo con fines explicativos, más en la realidad vemos que se puede establecer o que existe una relación entre los objetos formales y procesos pertenecientes a la realidad, empleando los entes formales como herramienta para la explicación de los segundos.

"Las ciencias formales demuestran o prueban; las ciencias fácticas verifican, confirman o desconfirman".

Nota :

- + Journal of Experimental Child Psychology. Vol.29,1980.  
Journal of Abnormal Psychology. Vol. 91, 1982.  
International Journal of Psychology. Vol.15, 1980.  
Journal of Counseling Psychology. Vol. 28, 1981.  
Journal of Educational Psychology. Vol. 72,1980.  
Social Psychology Quarterly. Vol.45, 1982.  
Research in Higher Education. Vol. 17, 1982.  
Psychological Review. Vol. 89, 1982.  
Psychology in the Schools. Vol. XIX, 1982.  
The Journal of Psychology (The General Field of Psychology) Vol.112, 1982.  
Journal of Personality and Social Psychology.
  
- ++ Revista de Psicología General y Aplicada. Vol.37,1982.  
Psychological Record. Vol. 32, 1982.  
Journal of Consulting and Clinical Psychology.Vol.49,1981.  
Journal of Educational Research. Vol. 75,1982.  
Revista Latinoamericana de Psicología. Vol.13,1981.
  
- +++ JABA. Vol. 14 y 15, 1981 y 1982 respectivamente.  
Revista Mexicana de Análisis de la Conducta. Vol.3,1977.

C A P I T U L O    I I

ANALISIS    NUMERICO

Como se había mencionado con anterioridad, en la investigación psicológica, hay dos tendencias claras en cuanto al análisis de los datos que van ligadas estrechamente con el tipo de diseño experimental que se utilizó. Hay quienes prefieren utilizar el análisis numérico a el análisis no numérico, (análisis visual, réplica) y viceversa. En cuanto al análisis no numérico, debemos aclarar que utiliza procedimientos de estadística descriptiva, básicamente con el fin de describir y/o resumir los datos, más la interpretación de los mismos se queda en el análisis visual, esto es, sin emplear nunca procedimientos inferenciales. En este apartado nos centraremos en lo que se entiende por análisis numérico en psicología. Teniendo presente que **este tipo de análisis es el que, se relaciona estrechamente con los diseños revisados en el punto anterior (diseños experimentales de grupo).**

Para que los resultados de una investigación tengan sentido y sean accesibles al investigador y posteriormente a los lectores, es necesario hacer un tratamiento de los resultados, estableciendo categorías, ordenando los datos, resumiéndolos, - etc.; todo esto orientado a extraer una respuesta de los resultados obtenidos. Cuando llevamos a cabo todos estos procedimientos, se puede decir que estamos analizando los datos (expresiones cualitativas o cuantitativas de nuestras variables), y en base a ellos interpretamos los resultados para -

hacer inferencias pertinentes y extraer conclusiones sobre las relaciones bajo estudio. Ahora cuando el análisis emplea números, se cuantifican las observaciones (índices de X variable o conducta a estudiar) a los diferentes niveles de medición, se da lugar a un análisis numérico.

Antes de abordar las escalas de medición hablaremos un poco acerca de lo que se entiende y los problemas a los que se enfrenta la medición en psicología. Si se desea profundizar en el tema, sugerimos ver a Siegel, 1976 o Magnuseon, 1978.

Por medición se entiende la asignación de números a objetos o individuos según determinadas reglas; por medio de la medición, los atributos de nuestras percepciones se transforman en entidades conocidas y manejables (Pick y López Velásco, 1979). Cuando medimos damos ciertas propiedades a uno o más objetos con ayuda del sistema numérico.

Las medidas que se obtienen con instrumentos físicos, casi nunca presentan problemas prácticos, ni al interpolar los resultados; con ésto nos referimos a instrumentos que nos dan resultados fijos y precisos en forma de puntajes (longitud, peso, tiempo, etc.). La relación entre los objetos que se están observando y los números es tan directa que -- mediante la manipulación de los números el investigador obtiene información acerca de los mismos (Siegel, 1976). El problema comienza cuando queremos medir variables psicológicas, las cuales se definen como una propiedad o caracte-

rística que poseen diferentes individuos en cantidades distintas (Magnusson, 1978).

La posición de un individuo sobre un continuo no es dada -- como un puntaje absoluto, sino como un puntaje relativo puesto que necesita un nivel de comparación y éste no se puede medir si se tiene solamente un individuo, pues sólo lo podemos comparar con otros individuos, es por esto que los instrumentos para medir variables psicológicas sobre una escala de intervalo parte de la suposición de una distribución normal de los puntajes. Lo anterior sólo desde el punto de vista de la psicometría clásica, pues desde los principios psicométricos de las técnicas de evaluación conductual, las puntuaciones de cada sujeto se comparan con algún nivel o criterio previamente establecido en lugar de con las puntuaciones de otros sujetos, "test orientados al criterio", Popham y Husek (1969) en Fernández Ballesteros y Carrobles (1981).

En psicología todavía no se ha desarrollado un sistema de medición que se apegue realmente a las necesidades del fenómeno psicológico, esto se debe a que es muy complejo plantearse un sistema de este tipo que se ha validado para cualquier individuo, y lo más importante, que logre medir objetivamente lo que se quiere medir.

No siempre existe una correspondencia exacta, isomórfica, entre el sistema numeral y las propiedades psicológicas, por esto algunos autores han iniciado la labor de desarrollar una axiomatología para cada uno de los diversos niveles de medi--

ción. Los axiomas especifican las relaciones que deben mantenerse en los datos así como en las propiedades de las relaciones, es pues de suma importancia el conocimiento del modelo de cada escala de medida, ya que de él depende el grado en que nuestros datos, reunidos mediante una de estas escalas, reflejen adecuadamente, las propiedades de los objetos (Arnau, 1978).

La clasificación de las escalas de medida según Stevens(1946), en Arnau, 1978; es la siguiente:

Nivel Nominal o Clasificadorio. Simplemente se involucra el proceso de denominar, etiquetar, clasificar, o agrupar objetos en clases. Se pueden asociar números en cada categoría, no se pueden aplicar operaciones aritméticas.

Escala Ordinal o de Rango. Busca ordenar en términos del grado en que poseen una determinada característica. No representa el tamaño de las diferencias que hay entre las categorías o entre los grupos.

Escala de Intervalo. Indica el orden y la distancia exacta que existe entre una y otra categoría, por lo tanto se necesita una unidad de medición que pueda ser utilizada una y otra vez (grados, tiempo, etc.).

Escala de Proporción o de Razón. Se diferencia de la escala anterior únicamente en que aquí se puede localizar un punto - cero absoluto. La proporción de un punto a otro cualquiera -

de la escala es independiente de la unidad de medida (Siegel, - 1976), por ejemplo: altura y peso.

Las diferentes escalas de medición conforman entre sí una - escala acumulativa. Es importante ver las suposiciones que tiene una técnica estadística, antes de aplicarla a nues-- tros datos, estas suposiciones están ligadas a la escala - de medición, al tamaño y al tipo de muestra con que se es-- tá trabajando (Pick y López Velásco, 1979).

En la mayoría de las investigaciones educativas, psicológi-- cas o de las ciencias del comportamiento; las mediciones se efectúan principalmente en escalas nominales y ordinales; - sin embargo también existen variables que pueden medirse en escala de intervalo o razón, sólo que en menor grado, ejem-- plos de variables en estas escalas son: los tiempos entre - respuestas y la duración de una conducta.

En psicología la mayor parte de las escalas que se utilizan son ordinales y algunas de intervalo, siendo su origen arbi-- trario; las unidades de medida que se emplean parten de la suposición de que son Constantes (Magnusson, 1978).

Para que los datos obtenidos con diferentes tipos de instru-- mentos de medición puedan usarse en situaciones prácticas, de-- ben satisfacer ciertas condiciones: a) El instrumento de me-- dida que se usa en un caso y con un propósito dado, debe me-- dir realmente el rasgo que se intenta medir; b) El instrumen-- to debe dar medidas confiables de manera que se obtengan los mismos resultados al volver a medir el rasgo bajo condiciones similares (Magnusson, 1978).

Cuando se tienen una serie de datos numéricos obtenidos de la o las relaciones bajo estudio, se procede a su interpretación y análisis, apoyados en la estadística.

La estadística es una rama de las matemáticas aplicadas -- (Glass y Stanley, 1980, Plutchik, 1975) de la que hay una -- diversidad de definiciones, coincidiendo la mayoría de -- ellas en que es un método (Haber y Runyon 1973, Cortada -- 1975, Meza, Morales y Magaña 1980) un instrumento o un conjunto de técnicas (Levin 1979, Yamane 1979) para la organización, presentación, interpretación y análisis de datos -- numéricos que conllevan a conclusiones acerca de los fenómenos de estudio.

A diferencia de los autores antes citados, Daniels W.(1981), no sólo menciona que la estadística sea un conjunto de técnicas, sino también un conjunto de conceptos. El agregado que hace Daniels a la definición de la estadística es válido, pues si revisamos las diferentes obras especializadas en estadística, notaremos que se nos presentan una serie de conceptos que fundamentan las técnicas propuestas y sin las cuales no serían comprensibles.

En suma, considerando las definiciones que se dan de la estadística, que tiene bases probabilísticos (Plutchik, 1975), y que como rama de las matemáticas se recurre a la aritmética simple y a álgebra elemental (Glass y Stanley, 1980), parece apropiado afirmar, tal como indica Kerlinger (1975) que la estadística es la teoría y el método de analizar -

datos cuantitativos obtenidos de nuestras observaciones.

Se dice que la estadística tiene varias finalidades:

a) Ordenar y resumir una gran cantidad de datos numéricos en forma útil y comprensible (Kerlinger, 1975; Plutchik, 1975; Daniels, 1977; y Levin, 1979).

b) Facilitar el estudio de poblaciones a partir de muestras, y conociendo o estudiando las propiedades de la muestra inferir, las de la población (Glass y Stanley, 1980; Haber y Runyor, 1973; Plutchik, 1975; Kerlinger, 1975; Pick y López Velásco, 1979).

c) Tomar decisiones sobre la aceptación o rechazo de las relaciones hipotéticas entre los fenómenos -Probar hipótesis. (Kerlinger 1975, Levin 1979).

De acuerdo a las finalidades anteriores, la estadística se ha dividido en dos ramas generales.

La primera finalidad, corresponde a lo que se conoce como -- estadística descriptiva, integrada por conceptos y técnicas que hacen posible, se cumpla su finalidad y sobre todo la - extracción de conclusiones que no rebasen al conjunto de co nocimientos de los datos estudiados (Cortada, 1975).

Las técnicas proporcionadas por la estadística descriptiva - para la presentación de los datos en forma comprensibles incluyen:

a) La elaboración de tablas, gráficas y figuras que -- permiten la visualización de los resultados; b) la trans-- formación de los resultados en medidas descriptivas útiles y convenientes para resumir las características del conjunto de datos estudiados (medidas de tendencia central y dispersión; c) la obtención de algún tipo de información específica (porcentajes, razones, correlaciones).

Con respecto a las finalidades de la estadística, marcadas anteriormente pertenecen a lo que se conoce como estadística inferencial, que al igual que la descriptiva comprende, conceptos (matemáticos y probabilísticos) y técnicas - (diferentes modelos) que a diferencia de la descriptiva hacen posible: 1) La extracción de conclusiones que rebasan los límites del conjunto de datos estudiados y permiten -- inferir, dentro de los límites siempre probables, valor -- para un individuo o un conjunto mayor (Nuria Cortada, 1975). Teniendo claro que inferir significa "derivar o deducir una conclusión de premisas o de hechos concretos" Kerlinger, - 1975; 2) Probar hipótesis, esto es, poner a prueba relaciones hipótéticas.

Esta diferencia entre la estadística descriptiva y la inferencia es por lo que frecuentemente se dice que el empleo - de una técnica inferencial implica ir más allá de la simple descripción de los datos. Sin que ese ir más allá, separe a la estadística descriptiva de la inferencia, de manera - tajante, pues es indispensable estudiar primero la estadística descriptiva para comprender mejor, con mayor facilidad,

con mayor rigor a la estadística inferencial (Holguin - Quiñones, 1975); aparte de que es preciso manejar y entender diversos conceptos probabilísticos, básicos, en los que se apoyan las técnicas o modelos inferenciales de análisis "para la estimación probabilística de inferencias que se obtienen a partir de las observaciones" (Arnau, 1978, pág. 278).

Algunos de los conceptos importantes de la estadística inferencial comprenden: a) Concepto de población y muestra; b) Qué es un parámetro y un estadístico; c) Distribuciones muestrales; d) Teorema del límite central; e) Ley de los grandes números; f) Intervalos de confianza y niveles de significancia; g) Hipótesis nula y alternativa; sin los que no serían comprensibles y entendibles las técnicas, procedimientos o modelos de la estadística inferencial.

Antes de continuar, queremos señalar que no desarrollaremos en ningún momento de manera técnica los conceptos y pruebas o modelos de análisis estadístico, puesto que existe en este momento un gran conjunto de autores que ya lo han hecho. Existe un extenso número de obras especializadas (ver bibliografía) elaboradas por expertos en la estadística que pueden consultarse.

En la investigación psicológica con diseños de grupo, la utilización de la estadística inferencial en el análisis de datos expresados cuantitativamente, se ha remitido principalmente a las pruebas de significancia, que permu

ten probar hipótesis sobre relaciones hipotéticas, planeadas estadísticamente y concluir si existe o no una relación en términos probabilísticos. "Una hipótesis es -- una suposición sobre una población estadística que en -- virtud de la información obtenida a partir de datos obser vados decidiremos aceptar o rechazar" (Arnau 1978, pág.266).

Dichas pruebas se han clasificado en dos tipos, de acuerdo a suposiciones o supuestos que mantienen, así como por el nivel de medición que exigen para su correcta aplicación.

#### A) Pruebas Paramétricas.

Las pruebas paramétricas fueron las primeras técnicas inferenciales que aparecieron en el desarrollo de los métodos estadísticos y se caracterizan por hacer un buen número de suposiciones acerca de la población de la que se obtuvo la muestra (Siegel 1976), y porque sólo son útiles cuando se alcanzan niveles de medición verdaderamente numéricos.

Siegel (1976); Pick y López Velásco (1979) enuncian y presentan claramente los supuestos de estas pruebas:

1) Las observaciones deben ser independientes entre sí; el puntaje que se asigne a un caso cualquiera no debe influir en el puntaje que se asigne a cualquier otro. Esto es, dado el evento A la probabilidad del evento B es igual a la probabilidad de B.

2) Las muestras deben tomarse de poblaciones que se encuentran distribuidas normalmente.

La suposición de normalidad es la más famosa en que se fundamentan las pruebas paramétricas (Kerlinger 1975). Por una distribución normal se entiende aquella que toma la forma de una campana (curva normal) con un punto máximo que la divide en dos mitades idénticas (simétricas).

3) Las muestras deben tener la misma varianza.

4) Las variables deberán estar medidas mínimamente en escala de intervalo.

Una aplicación incorrecta de las pruebas paramétricas será su utilización con datos nominales u ordinales; así como con observaciones no independientes y con una distribución diferente a la normal.

Aunado con el punto de escala de medida, se desprende otra característica de estas pruebas que reside en que los datos (observaciones cuantificadas) de la muestra pueden sumarse, multiplicarse, dividirse, debido a que su nivel de medición es realmente cuantitativa; operaciones que no se efectuarían si los datos estuviesen medidos en escalas más débiles.

En suma "las pruebas estadísticas clásicas o paramétricas se basan en supuestos que implican una serie de condiciones altamente restrictivas acerca de los parámetros de la población. Estas condiciones no son por lo general, verificadas, partiendo del supuesto de su cumplimiento, a excepción de las varianzas iguales" (Arnau, 1981).

La significancia de los resultados de una prueba paramétrica depende de la validez de sus suposiciones.

Las diferentes pruebas paramétricas, comunmente empleadas las señalaremos posteriormente en un cuadro, indicando -- cuáles son las más apropiadas según el número de variables manipuladas, etc.

#### B) Pruebas No Paramétricas.

Son procedimientos desarrollados en forma más reciente, -- que no dependen para su validez de la forma, naturaleza o distribución de la población (Daniels 1981); es decir no suponen que las distribuciones deben ser necesariamente -- normales (Plutchik, 1975). Razón por la que también se -- les conoce como pruebas de distribución libre.

"No especifican las condiciones de los parámetros de la población de la que se obtuvo la muestra. Hay algunas condiciones que se asocian con la mayoría de las pruebas estadísticas no paramétricas: Observaciones independientes y variable de continuidad básica, pero dichas suposiciones son pocas y mucho más débiles que las asociadas con las paramétricas. Además las pruebas no paramétricas no requieren de mediciones tan fuertes, la mayoría de las pruebas se aplican a datos de una escala ordinal y algunos de alguna escala -- nominal." (Siegel, 1976, pág. 52).

Las pruebas no paramétricas dada la escasez de sus condiciones y sobre todo los débiles niveles de medida con los que se pueden trabajar, han sido más adoptados por la investigación

ción en ciencias conductuales, pues como ya antes menciona bamos es difícil en muchos casos lograr medidas realmente - cuantitativas de las variables a estudiar.

Resumiendo las pruebas no paramétricas resultan muy útiles, y deben aplicarse sólo cuando: 1) Las observaciones hayan sido medidas en una escala nominal u ordinal; 2) Cuando no es factible hacer suposiciones acerca de la forma de dis- tribución de la población, ni de que las poblaciones de -- donde proceden las muestras, tienen varianzas iguales; -- 3) Si las muestras son muy pequeñas (3, 4, 5 ó 6) y no se conoce la distribución de la población; 4) Las observacio- nes sean independientes. Esta última condición que compar ten con las paramétricas, pero en un sentido menos fuerte y estricto.

Las ventajas de las no paramétricas en contraste con las - paramétricas estriban en que los métodos no paramétricos - son más rápidos y fáciles de aplicar y sobre todo que es - posible trabajar con datos que sólo clasifican u ordenan.

### C) Como Elegir una Prueba Estadística.

Identificar la escala de medida en que se encuentran los - datos numéricos de nuestro experimento; así como si es per tinente una prueba paramétrica o no paramétrica en base a sus características y supuestos, es importante tener pre-- sentes algunos otros factores para la correcta elección y aplicación de una prueba apropiada, que nos permita un --

buen análisis de nuestros datos y más que nada, que sea -- el instrumento pertinente para tomar la decisión de si -- existe o no una influencia efectiva de la variable o va-- riables independientes manipuladas, sobre la VD. Estos -- factores a los que nos referimos, y que son válidos para ambos tipos de pruebas son:

1) Número de variables independientes manipuladas.

2) Número de condiciones experimentales determinadas por los niveles de la VI o variables manipuladas y expresada por la cantidad de grupos.

3) Grupos independientes o relacionados; lo cual depende de la manera en que los sujetos sean asignados a las condiciones experimentales (aleatorización, o por medio del -- apareo, el sujeto como su propio control).

Grupos correlacionados.- Aquellos en los que se utilizan el apareo o se utiliza al sujeto como su propio control.

Grupos independientes.- los que se forman **aleatoriamente**.

A continuación presentamos un cuadro de las pruebas que resultan propias según la situación de investigación, considerando los aspectos arriba citados. (Ver la siguiente hoja)

D) **Lógica General de las Pruebas Estadísticas.**

La utilización de las diferentes pruebas estadísticas en el análisis de datos numéricos arrojados por una investigación con diseños de grupo, se pasan en la prueba de hipótesis; -

Prueba	Número de VI	Número de Condiciones	Pruebas Específicas	
			Grupos correlacionados	Grupos independientes
Paramétrica (distribución normal)*	una VI	dos	t para grupos correlacionados.	t
		tres o más	Análisis de varianza (AVAR) para grupos relacionados.	AVAR, Tukey, - Neuman, Duncan.
	dos o más	dos o más - por variable	Análisis factorial de la varianza. AVAR de dos factores correlacionados.	Análisis factorial de la varianza. AVAR de dos factores independientes.
	una VI	dos	T de Wilcoxon(ordinal).	$\chi^2$ Ji cuadrada (nominal).
tres o más		Prueba de los signos (ordinal) Q de Cochran(nominal).	U. de Mann Whitney - (ordinal) AVAR de Kruskal Wallis (ordinal).	
No Paramétricas(distribución libre)**	dos o más	dos o más	AVAR de Friedman (ordinal).	Ji cuadrada para k -- muestras(nominal u ordinal).
		dos o más	dos o más	diferentes formas de - Ji cuadrada (nominal u ordinal).

\*\* Observaciones en escala nominal u ordinal.

\* Observaciones (datos) mínimamente en escala de intervalo.

que consiste en formular una hipótesis estadística llamada nula ( $H_0$ ) y que expresa una igualdad supuesta en la ejecución de los diferentes grupos (condiciones experimentales), lo que indica que la variable dependiente no ha sido afectada por la manipulación de la variable o variables manipuladas o que si hay cambios se deben a otros factores; -- esta hipótesis puesta a prueba en forma estadística contrastándose con otra hipótesis, conocida como alterna ( $H_1$ ), también planteada en términos estadísticos y que coincide con la hipótesis de investigación. Es decir la hipótesis alterna contempla que existen diferencias entre las condiciones a causa de la introducción experimental de un evento.

La hipótesis nula, es la única que explícitamente es enunciada para ser aceptada o rechazada; su aceptación indica que la VI no altera a la VD. Por otra parte, su rechazo implica que la VI si tiene efectos sobre la variable dependiente y evidentemente confirma el enunciado de la hipótesis alterna (hipótesis de causalidad).

Sin embargo, tomar una decisión de aceptación de una hipótesis, no es tan simple como pudiera parecer. La estadística inferencial proporciona un conjunto de reglas o criterios (\*) para probar la hipótesis y tomar una decisión, -- siempre con bases probabilísticas (Arnau, 1978), y en virtud de la información obtenida a partir de los datos observados.

(\*) Intervalos de confianza, niveles de significancia, zonas de aceptación y de rechazo, etc.

Al aplicar las pruebas de significancia estadística "nos hacemos siempre la misma pregunta como condición fundamental para la inferencia de hipótesis de causalidad: ¿Cuál es la probabilidad de que los resultados observados hayan ocurrido al azar?" (pág. 17, Arnau, 1981).

La cita anterior marca claramente como se determina una probabilidad de ocurrencia para cada hipótesis; además de hacer referencia a otra forma de plantear la hipótesis nula ( $H_0$ ), muy importante y que en psicología es adoptada: La probabilidad de que los resultados observados ocurran a causa del azar. El plantear la hipótesis de esta manera contempla que las variaciones entre una condición experimental y otra, no se deben a la acción de la variable independiente, sino a otros factores, tales como las diferencias individuales, errores de muestreo, instrumentos de medición, posibles influencias de un factor extraño; estos entre otros comprendidos dentro del concepto del azar. El azar representa el efecto conjunto de muchas pequeñas causas que van a influir en el fenómeno y cuyos efectos por separado van a ser despreciables (en algunas ocasiones) pero que al sumarse pueden tener un efecto conjunto importante (Ibarrola, 1978).

En suma, las pruebas de significancia estadística constituyen instrumentos que permiten contrastar dos hipótesis: La de si los resultados de nuestras observaciones se deben al azar o a la intervención de la variable o variables manipuladas, siempre en términos probabilísticos.

E) Su Aplicación a los Diseños de Grupo en Investigación Psicológica.

La investigación psicológica desde sus orígenes y aún en la actualidad ha utilizado los métodos cuantitativos para el análisis de sus datos, esto, a pesar de que hay investigadores que prefieren no utilizarlos, la anterior afirmación la hacemos en base a una revisión que se realizó en las revistas dedicadas a la investigación psicológica (ver anexo 2, al final de este capítulo). Lo que se pudo observar, es que en todos los reportes de investigación con diseños de grupo, se aplicó un análisis estadístico, en el caso de los reportes de investigación con diseños operantes, se observó que aunque en muchos casos no se utiliza el análisis **numérico** de los datos, en un número importante (considerable) de estos reportes, sí se recurre a los métodos cuantitativos. Es importante señalar, que casi en la totalidad de los reportes de investigación operantes se emplearon básicamente técnicas de estadística descriptiva para la interpretación de los datos, debemos recordar que en estos casos sólo se describen y se resumen; y a partir de ellos se da la interpolación de los resultados.

El uso de la estadística en los diseños de grupo se basa principalmente en que desde un principio, lo concerniente al estudio de las diferencias individuales se cerró en el trabajo con grupos y en el promedio, por lo que se comienza a adoptar y desarrollar pruebas estadísticas, que se utilizan en la actualidad como el principal instrumento de

análisis de datos que nos proporciona la diferencia entre los individuos o la variabilidad intersujeto (Hersen y -- Barlow, 1976). En la actualidad se están desarrollando -- técnicas apropiadas para el análisis de datos en diseños experimentales operantes, con ésto nos referimos a los -- diseños de series de tiempo que enfatizan la medida repe tida, continua y periódica, bajo condiciones de línea -- base e intervención (Kratochwill, 1978).

Consideramos que el análisis estadístico ha ido a la par con el desarrollo de la investigación, como un instrumen- to de análisis y que en el caso de la psicología nos ha -- proporcionado gran ayuda en tres aspectos: a) En la des-- cripción y resumen de los resultados; b) Como el único -- índice de obtener información cuando no resulta prudente o no proporcione información otro tipo de análisis; c) Co mo otro índice de información para reforzar la ya obteni- da a través de otro tipo de análisis y a la vez no falsear los datos obtenidos en la descripción y resumen de los mis mos.

Estos puntos son válidos para cualquier tipo de diseño ex- perimental, no hay que olvidar que los métodos cuantitati- vos son un instrumento de análisis de los datos y no el -- diseño en sí mismo, por lo tanto no debemos anteponer el -- análisis cuantitativo al diseño.

A N E X O 2

- **Behavior Therapy.** Vol. 11, 1980; vol. 12, 1981.
- Journal of Applied Psychology. Vol.66,1981; vol.67,1982.
- Journal of Verbal Learning and Verbal Behavior. Vol.21,1982.
- Journal of Educational Psychology. Vol.72,1980; vol.73,1981.  
y; vol. 74, 1982.
- Journal of Experimental Psychology Human Learning and Memory.  
Vol. 7, 1981.
- Journal of Experimental Psychology. Human Perception and ---  
Performance. Vol. 7, 1981.
- The Journal of General Psychology. Vol. 106.y 107, 1982.
- The Journal of Genetic Psychology. Vol. 140 y 141, 1982.
- The Psychological Record. Vol. 31, 1981.
- Personality Social Psychology. Vol. 42 y 43, 1982.

Además de las ya citadas en el capítulo II.

C A P I T U L O

III

C A P I T U L O     I I I

CORRESPONDENCIA ENTRE PRUEBAS DE SIGNIFICANCIA  
ESTADISTICA Y LOS DIVERSOS TIPOS DE DISEÑO

El objetivo de este capítulo es hacer una serie de recomendaciones y sugerencias en torno al tipo de análisis numérico -- más adecuado entre determinado tipo de diseños de grupo. Para ésto, nos basaremos en la estructura del apartado II, referente al inciso: Tipo de Diseños de Grupo.

Cabe aclarar que las recomendaciones y sugerencias que hare-- mos en este trabajo, se deben tomar como tales, y no como un hecho en sí, hay que tener presente que el análisis numérico sólo es una herramienta del diseño experimental y el inves-- tiguador debe elegir la estadística de prueba a partir del -- tipo de diseño experimental que utilice, así como del tipo - y número de variables que maneje.

1.- Diseños Aleatorizados.

Como ya expusimos en el capítulo II, una característica fundamental de este tipo de diseños es el que la selección de - los sujetos que participarán en un experimento debe hacerse azarosamente, al igual que la formación de los grupos que -- constituirán las diferentes condiciones. Remarcar aquí nue-- vamente está característica obedece a que para una buena elec-- ción de la prueba estadística que conlleva a un análisis ade-- cuado de los resultados, es vital conocer el procedimiento --

mediante el cual se seleccionaron los participantes y la manera en que se asignan a las diferentes condiciones, pues — ésto último es lo que determina que se hable de grupos independientes o correlacionados, existiendo pruebas estadísticas propias para cada caso.

Los diseños aleatorizados que siguen tanto la selección de — sujetos de manera azarosa, como su asignación del mismo modo a las condiciones, se habla de grupos independientes, es decir que no existe ninguna relación específica entre los sujetos de un grupo y otro.

Dejando entonces claro, que siempre que se utilicen diseños aleatorizados, nuestros grupos serán independientes y por — consiguiente tendremos que recurrir a la elección de un procedimiento de análisis cuantitativo para grupos independientes. Aquí presentaremos algunas pruebas específicas paramétricas y no paramétricas para cada uno de los diseños alea—torizados, tratados en el apartado II. Los comentarios que haremos a los diseños serán desde el punto de vista del análisis cuantitativo.

## 1.1 Bivalentes.

### 1.2.1 De Dos Condiciones.

Para los diseños cuya estructura sólo contempla dos grupos, ya sea uno experimental y uno control, o dos experimentales, resulta pertinente la aplicación de las pruebas: i) Paramé—

tricas, "z" o "t"; ii) Las no paramétricas, Ji cuadrada o la U de Mann Whitney.

Tener presente que la elección de una prueba paramétrica se determina por el cumplimiento que se haga de los supuestos asociados a cada una de ellas (ver Cap. III). En lo que — respecta a si optar por "z" o "t", depende del tamaño de — las muestras (grupos) con los que se trabaja; "z" es regu— larmente útil para grupos grandes, mientras que "t" es para grupos pequeños (menores de 25). Ambas pruebas requieren — que se trabaje con variables numéricas, y que impliquen mí— nimamente un nivel de medición de intervalo.

Por otro lado, la decisión entre si optar por Ji cuadrada o la U de Mann Whitney se determina por el nivel de medición logrado en la variable de estudio. Ji cuadrada ( $x^2$ ) se — suele aplicar cuando se tiene una variable categórica y las observaciones que se recogen para obtener información acerca de su comportamiento, consisten en simples frecuencias (con— teos) de las ocurrencias en las diferentes categorías discre— tas previamente definidas. La escala de medición implicada para este tipo de observaciones es la nominal.

La aplicación de la U de Mann Whitney, requiere que, la medi— ción de la variable de interés se haga en una escala ordinal y que ésta sea continua.

Tanto  $x$  y  $t$ , así como Ji cuadrada y la U de Mann Whitney son pruebas de significancia estadística, que siguiendo la lógica

general de las pruebas ya sean paramétricas o no paramétricas (ver Cap. III) ayudan a determinar si las diferencias entre los grupos (muestras) son un índice convincente de una diferencia realmente atribuible a las diversas condiciones a que fueron expuestos. En el caso particular de las pruebas aquí mencionadas para diseños de dos grupos independientes, permiten probar la hipótesis nula ( $H_0$ ) de que no existe diferencias entre la ejecución de los dos grupos bajo estudio, es decir se plantea que tanto la ejecución en el grupo experimental como en el control o experimental dos es igual. Otro planteamiento de  $H_0$ , es el que las diferencias entre una y otra condición se deben al azar. Cualquiera de las dos formas de enunciar la hipótesis nula, implica llegar a afirmar, si es aceptada, que la variable independiente manipulada no causa ningún efecto, en el caso de que se compare un grupo experimental con una condición control; o que ambos niveles de la VI tienen efectos iguales sobre la variable dependiente, cuando se comparan dos condiciones experimentales.

El valor numérico que usan las pruebas en la comparación de los grupos, es una medida descriptiva de la ejecución de interés del grupo (conjunto de observaciones). Los valores de estas medidas descriptivas son conocidos como estadísticos o estadígrafos y varían para cada prueba. Glass y Stanley 1974, señalan que los estadígrafos son abreviaciones estadísticas, cuyo valor numérico describe una característica de un conjunto de observaciones.

La estadística o el estadígrafo empleado por "t" o "z", en-

la comparación de los dos grupos es la media ( $\bar{x}$ ); En el caso de Ji cuadrada " $\chi^2$ ", se habla de la proporción de casos para una y otra categoría; en cuanto a la U de Mann Whitney la estadística utilizada es el valor de U.

Continuando con la lógica de las pruebas estadísticas, tenemos que existe una segunda hipótesis llamada alterna ( $H_1$ ), la que regularmente coincide con la hipótesis de investigación y que es favorecida (apoyada) si  $H_0$  se rechaza. El enunciado de dicha hipótesis depende de lo que un investigador pretenda. Antes de pasar a revisar las formas específicas en que  $H_1$  se puede plantear, hay que tener claro, que de manera general contempla que si existen diferencias entre las ejecuciones de los dos grupos y que éstas no se deben al azar, sino a la manipulación de la variable independiente.

Una de las formas de plantear  $H_1$ , responde a aquellos casos en que sólo existe interés por observar o mostrar que si hay diferencias entre los efectos de un grupo y otro, sin importar si éstos son mayores o menores; para tales casos se dice que la hipótesis no tiene dirección y se expresa por una desigualdad entre las ejecuciones de los dos grupos.

Por otro lado, si un investigador de antemano señala que al exponer a un grupo a una cierta variable su ejecución incrementará o decrementará con respecto a la condición control o a otro valor de la VI, su hipótesis tiene dirección y se expresa por una relación de mayor que " $>$ ", o menor que " $<$ "

entre las dos condiciones, según en la que se esté interesado.

El desarrollo de cada una de las pruebas ya sugeridas y que seguiremos sugiriendo para cada tipo de diseño, así como — sus respectivos cálculos numéricos que conllevan a la aceptación o rechazo de una hipótesis nula, no los señalaremos aquí, pues insistimos en que se recurra a otras especializadas en esos aspectos, por lo que recomendamos ver la bibliografía.

Resumiendo, tenemos que la aplicación de las pruebas para — diseños bivalentes de dos condiciones son útiles concretamente para los diseños pretest-postest y sólo postest. En lo que se refiere al diseño sólo postest comparando el valor de las observaciones finales (los postest, de las dos condiciones). Para el diseño pretest-postest, Campbell y Stanley (1973) nos mencionan que hay que tener cuidado en no caer en un error muy peculiar que suele hacerse al aplicar "t" o "z" a la diferencia entre el pretest-postest para el grupo — experimental y calcular otra para el pretest-postest de la — condición control; si ésto se hace señalá, se llega a conclusiones erróneas, pues no se está haciendo una comparación estadística directa entre las dos condiciones experimentales. — La alternativa para evitar ese error consiste en calcular para cada grupo los puntajes de ganancia entre el pretest y el postest, y entonces si calcular una "t" o "z" sobre esos puntajes entre los dos grupos.

### 1.2.2 De Más de Dos Condiciones.

Diseños que permiten el estudio, de una sola VI pero con dos o más niveles; lo que genera que se manejen tres o más grupos.

Las pruebas estadísticas indicadas para la evaluación de los resultados (datos), en estos diseños, consideramos una extensión de las dos condiciones es el análisis de varianza (ANOVA) de un sólo factor, por las paramétricas; y el análisis de varianza (AVAR) de una clasificación por rangos de Krus Kal-Wallis, o Ji cuadrada para K muestras independientes, por las no paramétricas.

El hecho de que en la denominación de los análisis de varianza se aclare que son de un sólo factor y una clasificación, obedece precisamente a su tuidad únicamente cuando se manipula una variable independiente, evidentemente con sus respectivos niveles.

Si la variable de estudio es numérica (nivel de medición minimamente de intervalo) y las observaciones cumplen con los supuestos de las paramétricas en definitiva se elige el análisis de varianza. Si la variable es categórica y se registran frecuencias para categorías discretas (sea en escala nominal u ordinal), la prueba indicada será  $\chi^2$  para K muestras. Por último para aplicar el AVAR de Krus Kal-Wallis es necesario que la variable en estudio tenga una distribución continua y el nivel de medición implicado sea ordinal.

Los tres tipos de análisis, siguen la lógica ya conocida con la variante de que la formulación de  $H_0$  no se limita a plantear una igualdad entre dos condiciones, sino que contempla la igualdad de ejecución entre tres o más grupos expuestos a diferentes condiciones. En consecuencia la hipótesis alternativa enunciará que sí existen diferencias entre las distintas condiciones.

La hipótesis puesta a prueba contemplará que existe una igualdad: a) entre las medidas de ejecución de cada grupo, si se trata del ANOVA de un factor; b) entre la proporción de casos para todas las condiciones, esto para  $\chi^2$  cuadrada; y c) entre el puntaje promedio de los diferentes grupos, cuando se emplee el AVAR de una clasificación. La otra forma de plantear dicha hipótesis, afirmará que las diferencias son debidas al azar. En el caso de que los resultados conlleven el rechazo de la hipótesis nula, lo único que nos indicará el encontrar diferencias significativas es que los cambios se pueden atribuir a la intervención de los niveles de la variable independiente, manipulados, más no especificará cuál de los tratamientos es más efectivo.

Para el ANOVA de un factor, pueden emplearse a posteriori de su aplicación, una serie de pruebas: Tukey, Duncan, Newman-Keuls, Soheffe y otras, a fin de precisar la dirección y fuerza de las diferencias, es decir que tratamiento es mejor que otro. Además de estas pruebas es posible aplicar una prueba "t", también a posteriori para comparar las dos condiciones.

Una vez señaladas agrosso modo las pruebas estadísticas pertenientes para los diseños bivalentes de más de dos condiciones, sólo nos resta especificar los diseños incluidos dentro de esta clasificación. De acuerdo al capítulo II, tenemos a aquellos diseños que toman como base a los de dos condiciones (pretest-postest, sólo postest) y que a diferencia de ellos, incluyen más niveles de la variable manipular, de tal forma que encontremos diseños bivalentes de más de dos condiciones que en su estructura contemplen el pretest y el postest, o sólo el postest.

Si se tiene un diseño de más de dos condiciones con sólo postest la aplicación del ANOVA considerará las medias de cada grupo, pero esencialmente su procedimiento trabaja con la variación o variabilidad: a) entre las medias de los grupos y b) entre la variación de la ejecución de un sujeto con respecto a la media de su grupo. De esto último que se hable de varianza entre grupos (entre las medias de los grupos) y varianza intra-grupos (cada sujeto con respecto a la media de su grupo). La decisión de aceptar o rechazar  $H_0$  depende del valor que se obtenga al aplicar una razón de variabilidad dada por la varianza entre grupos, sobre la varianza dentro de los grupos (razón  $F$ ). Sin necesidad de hacer cálculos numéricos se puede concluir de manera general, que cuando la variabilidad dentro de los grupos es mayor a la variabilidad entre los grupos, los efectos son atribuibles al azar. Caso contrario de lo que sucede si se encuentra una mayor variabilidad entre los grupos que intra-grupos, si esto último acontece  $H_0$  se rechaza y se dice que existen

diferencias significativas, a causa de los niveles manipulados.

Por lo que toca a la aplicación del AVAR de Krus Kal-Wallis en diseños sólo posttest, considera la medición final de la ejecución individual por sujeto. Agrupando y ordenando en un sólo conjunto todas las observaciones de cada uno de los sujetos, como si se tratara de un sólo grupo, sin importar si es experimental o control y finalmente para tomar la decisión se calcula el estadístico H.

Para un diseño de más de dos condiciones protest-postest, - el ANOVA es el análisis conveniente en la evaluación de sus resultados más hay que tener presente el señalamiento que - Campbell y Stanley, 1973, hacen para los diseños de dos condiciones. Haciendo caso de tal indicación marcada ya en -- los diseños de dos condiciones, creemos que el análisis de varianza de un sólo factor se debe aplicar sobre los puntajes de ganancia calculados entre el pretest- y el postest.

Para concluir con los diseños de más de dos condiciones, -- sólo nos queda mencionar la prueba estadística apropiada para el diseño cuatro grupos de Solomon. A este respecto nuevamente los autores arriba citados, nos dicen que no hay un procedimiento estadístico que utilice a un mismo tiempo los seis conjuntos de observaciones en tal disposición: dos observaciones pretest y cuatro postest; por lo que el ANOVA - se descarta. No obstante, si se dejan de lado los pretest es posible la aplicación del ANOVA de dos factores, el cual

será más ampliamente tratado en los diseños que permiten el análisis de dos variables independientes, conjuntamente.

### 1.3 Multivalentes.

El análisis numérico para los diseños factoriales ya sea que se manipulen dos o más variables independientes con determinados niveles cada uno es el análisis de varianza de dos factores. Si se manipulan tres variables el análisis de varianza adecuado es el de tres factores. El requisito para su aplicación, es que se cumplan con los supuestos de las pruebas paramétricas y que la variable de estudio sea continua en un nivel de medición fuerte.

En experimentos de este tipo, en que se manipula más de una variable independiente el interés del investigador se centra en observar de manera simultánea los efectos que causan sobre una variable dependiente, dos o más independientes.

La disposición de los diseños factoriales, nos da pauta para observar tres tipos de efectos. El primero de ellos permite detectar en que medida el efecto de un nivel de la primera variable independiente (A) varía de acuerdo al nivel manipulado de una segunda variable (B); esto es lo que se conoce como efectos simples o efecto de cada variable independiente con sus respectivos niveles para cada uno de los valores de la otra variable.

Otro tipo de información que es posible adquirir es la acción

global de ambas variables en forma independiente, como si se hubieran realizado dos estudios por separado, conociéndose estos factores como efectos principales.

Hay otro efecto, además de los anteriores llamado efecto -- secundario o de interacción, que como su nombre lo indica, estima el efecto en la variable dependiente como consecuencia de la interacción o relación entre las variables manipuladas.

Los cálculos numéricos que se siguen para conocer los factores en forma también numérica, no los desarrollaremos, los motivos por los que se hará ya fueron señalados anteriormente. La importancia de haberlos mencionado es sólo para poder llegar a decir que con la aplicación del ANOVA a los resultados de un diseño factorial es posible probar una serie de hipótesis, que plantean: 1) que la ejecución ante cada nivel de la variable A, para cada diferente condición es igual, denotándose por una igualdad entre  $\alpha_1, \alpha_2 \dots$  etc. ( $\alpha_1 = \alpha_2 = \dots = \alpha_i$ ); 2) que la ejecución ante cada nivel de la variable B, para cada diferente condición es igual, señalándose por  $\beta_1 = \beta_2 \dots = \beta_i$  ( $\beta_i$  beta); y 3) que el efecto de interacción de los niveles de cada variable manipulada es igual  $\alpha\beta_1 = \alpha\beta_2 \dots = \alpha\beta_i$ .

Para poder probar estas hipótesis el análisis de varianza, comprende el cálculo de la variabilidad entre los niveles de la primera variable (A); la variabilidad entre los niveles de la segunda (B); la variación entre las diferentes --

interacciones de las dos variables, y finalmente la variabilidad de error.

## 2.- Diseños de Bloques Homogéneos.

En los diseños de bloques homogéneos los diferentes grupos se consideran correlacionados o relacionados.

### 2.1 Bivalentes.

#### 2.1.1 De Dos Condiciones.

Sólo se trabaja con dos grupos formados a través del apareamiento de los sujetos. Se dice que los grupos están relacionados porque los sujetos de ambos grupos comparten una o más características, o tienen el mismo valor respecto a una variable importante de controlar, ejemplo: Nivel de ansiedad, grado escolar, sexo, edad, etc.

Al igual que con los diseños aleatorizados las pruebas adecuadas en el análisis de los resultados con diseños apareados pueden ser paramétricas o no paramétricas, dependiendo del cumplimiento de los supuestos y del nivel de medición de la variable.

De las paramétricas, la prueba indicada es "t", también sugerida para los diseños aleatorizados de dos condiciones. La lógica es la misma, la única diferencia que existe entre la "t" para grupos independientes y relacionados es con respecto

a la manera de enunciar la hipótesis nula y el procedimiento de cálculo para obtener el estadístico. En lo que concierne a la hipótesis nula, ésta plantea la no diferencia significativa entre los puntajes de uno y otro grupo; mientras que -- con grupos independientes se enunciaba la igualdad entre medidas ( $\bar{x}$ ), del grupo experimental y control.

El referirnos a la diferencia entre los puntajes, nos está -- marcando en sí, que se sigue un procedimiento de cálculo dis -- tinto; consistente en comparar el puntaje de un sujeto del -- grupo experimental con el del otro grupo (control o experi -- mental II) al que fue apareado y sacar la diferencia entre -- sus ejecuciones. Lo anterior para cada uno de los sujetos -- en ambos grupos y sobre dicho puntaje de diferencia se pro -- cede a aplicar "t".

De las pruebas no paramétricas propias para estos diseños de dos grupos apareados, cabe señalar a la prueba de los signos y la T de Wilcoxon o prueba de los rangos señalados y pares igualados. La lógica de ambas pruebas es la ya conocida, y las dos requieren una continuidad de la variable.

La prueba de los signos emplea los símbolos + y - para seña -- lar la dirección de las diferencias entre las observaciones de los sujetos de un grupo y otro, siendo ésta la causa por la que precisamente recibe dicho nombre. Para conocer la -- dirección de esas diferencias, se comparan las ejecuciones, expresadas numéricamente, entre cada uno de los sujetos ex -- perimentales y los del control o experimental dos por pare --

jas. Es así, como si el puntaje de un sujeto experimental es mayor que el del sujeto del grupo control, se dice que la relación tiene una dirección positiva y se le marca con +. Si por el contrario el puntaje del sujeto experimental es menor que la del control, la dirección de esta relación se dice -- negativa y se signa con -.

La decisión de aceptar una hipótesis nula que plantea una igualdad entre la dirección de las diferencias de los pares - (igual número de relaciones marcadas con + y con -); se basa esencialmente en el conteo de la ocurrencia de relaciones positivas y negativas, en ningún momento se efectúan operaciones aritméticas. De acuerdo a esta hipótesis nula, la hipótesis alterna puede contemplar que: a) el número de parejas con una relación positiva o negativa no es igual; b) que las relaciones con dirección positiva son mayores que las marcadas con -; d) que los pares con una dirección negativa son - mayores que las positivas. Las diferentes formas de enunciar  $H_1$ , dependiendo del objetivo de investigación.

A diferencia de la prueba anterior la T de Wilcoxon además - de considerar la dirección de las diferencias, toma en cuenta su magnitud, dando un mayor paso al par que muestra una diferencia grande entre las dos condiciones que al par que exhibe una diferencia pequeña sin reparar en la dirección - (Siegel, 1970). Una vez que se han asignado los pesos a - las relaciones, se asignan rangos, conservando su respectiva dirección, se suman los rangos de todas las relaciones - positivas y todas las negativas por separado. De acuerdo - a lo expuesto,  $H_0$  enuncia que la suma de los rangos para --

las relaciones asignadas con + es igual a las denotadas con -, en caso contrario se rechaza  $H_0$ . Si resulta que la suma de los rangos para las diferencias con dirección positiva es mayor que para las diferencias negativas, se concluye que existe una diferencia significativa entre las condiciones y tal diferencia se atribuye a la variable manipulada. Lo mismo ocurrirá si  $H_1$  enuncia que la suma de los rangos negativos es mayor.

## 2.2 De más de Dos Condiciones.

### 2.2.1 Diseño de Bloques al Azar.

De acuerdo a las características descritas en el capítulo II, decíamos que es útil cuando es necesario controlar una serie de niveles de una misma variable que se cree puede contaminar los resultados al interferir con los diversos valores de la variable independiente manipulada.

El análisis de varianza (ANOVA) de dos factores es la prueba de significancia estadística idónea para la evaluación de los resultados de un diseño de bloques aleatorizados o al azar, siempre y cuando la variable a observar sea continua y numérica. El ANOVA es el mismo que se ocupa con los diseños factoriales aleatorizados, más hay que tener cuidado con las variantes que se presentan en su aplicación dependiendo del tipo de diseño.

Al referirnos al ANOVA de dos factores, en los diseños aleatorizados aclaramos que se deben manipular dos variables --

independientes, entonces porque aquí afirmamos que es idóneo para este diseño, cuya estructura sólo contempla una VI. La respuesta a esta cuestión es que aunque es definitiva sólo se manipula una VI, la variable sobre la que se realiza el bloqueo, es considerada como el segundo factor requerido para proceder a la aplicación de un ANOVA de dos factores.

Al analizar el ANOVA la atención se centra en evaluar si: -- a) las diferencias de ejecución entre las condiciones (grupos) se deben efectivamente a los niveles manipulados de la variable independiente; b) las diferencias entre los bloques son atribuidas a los niveles de la variable sobre la que se procedió a formar los bloques; c) las diferencias (variabilidad) son causadas por el azar (error). Todos los aspectos anteriores se consideran los factores que pueden contribuir a los cambios finales en la VD (varianza total) el término de un experimento. El modelo del análisis de varianza que incluye las anteriores fuentes de variación es conocido como aditivo, pues se asume que el efecto en potencia, causado para la interacción de la variable independiente y la variable de bloqueo se adhiere a la variabilidad de error.

Existe otro modelo de análisis de varianza de dos factores que además de contemplar como causas de la variación en la ejecución de una y otra condición experimental a la variable independiente, a la variable de bloqueo y al azar, incluye también la posibilidad de que los cambios se originan por la interacción de los niveles de la VI y los de la variable de bloqueo. Dicho modelo, se conoce como no adi-

tivo. Preutilización de este modelo del ANOVA se aplica una prueba para mostrar si existe aditividad o no; si hay aditividad al desarrollo del análisis de varianza desechará la hipótesis de que existe interacción entre la variable de bloqueo y la independiente. Por otra parte si esta prueba da muestras de no aditividad, la interacción de las variables mencionadas se considerará. La prueba a la que nos estamos refiriendo es la de Tukey,

La aceptación o rechazo de las hipótesis nulas, planteadas en términos de los factores a los que se atribuyen los cambios, depende de los resultados arrojados por la razón F, para cada una de las fuentes de variación.

### 2.3 Diseños de Doble Bloqueo.

El diseño cuadrado latino fue el que nos ocupó en el punto II, su característica distintiva es el control de dos potenciales variables extrañas.

Teniendo presente que a través de la estructura del diseño se controlan varios valores de dos variables extrañas y se manipulan igual número de niveles de la variable independiente, los cambios que operan en la variable dependiente pueden adjudicarse a cuatro aspectos: a) los niveles de la variable manipulada; b) la primera variable extraña sobre la que se procede a formar los bloques; c) la segunda variable de bloqueo; y d) el azar. Si nuestra variable es continua y numérica el análisis de varianza de tres factores será la herra-

mienta indicada para analizar la variación total entre - las condiciones experimentales y tener elementos para de cir a que factor o factores podemos atribuir los cambios observados en la variable dependiente.

El análisis de varianza se dice de tres factores, porque aunque el diseño sólo involucra una variable independien te, las dos variables extrañas pueden originar cambios en la variable de estudio y representan los otros dos facto res.

El planteamiento de hipótesis nulas para este modelo del ANOVA, contempla de acuerdo a los factores mencionados:

1) Que los niveles de la VI causan iguales efectos, - denotándose por  $\alpha_1 = \alpha_2; \dots \alpha_i$ .

2) Que la ejecución en tódas las condiciones es igual independientemente de que los sujetos de cada grupo difie ren en cuanto al valor de una variable extraña, (ejemplo - 12 años, 16 y 20) Representándose estadísticamente como -  $\beta_1 = \beta_2 = \dots \beta_i$

3) Que los cambios en la VD para todos los grupos es igual, aunque los sujetos de cada uno de ellos tengan ca racterísticas diferentes conforme a la segunda variable - de bloqueo (ejemplo grado escolar); representándose por  $\gamma_{k_1} = \gamma_{k_2} = \dots \gamma_{k_i}$

Si en base a los cálculos, respecto a la información ob-

tenida y cuantificada se concluye que no existe tal igualdad entre los factores ( $H_0$  se rechaza), o en alguno de ellos se dice que las diferencias son significativas, tales cambios se atribuyen a los niveles de la variable manipulada o a una de las dos variables de bloqueo, según los resultados proporcionados por las razones F, dadas por la razón entre la variabilidad de los tratamientos, variable de bloqueo uno, variable de bloqueo dos, sobre la variabilidad de error. De manera general se dice que si la variabilidad de error es mayor que la de las diferentes variables,  $H_0$  se acepta y por tanto las diferencias encontradas se deben a la acción del azar.

El modelo del análisis de varianza que acabamos de describir es el aditivo, ya que, suma las posibles variaciones debidas a la interacción de las variables a la variación del error. Cuando el modelo es no aditivo, dicha variación se contempla.

## 2.4 Multivalentes.

Diseños en los que se manipulan dos o más variables independientes.

### 2.4.1 Diseños de Bloques al Azar, Factoriales.

Nuevamente el análisis de varianza es el adecuado para la evaluación de los resultados proporcionados por estos diseños. La única variante de este modelado del ANOVA y los anteriores es en cuanto al procedimiento, pues al incrementar el número de variables incrementan las variaciones a estimar. Donde dichas variaciones se pueden atribuir a: 1) La acción de cada variable manipulada, consideradas independientemente (efectos

principales); 2) Interacción o interacciones entre las variables independientes; 3) La variable de bloques; y 5) El error.

### 3.- Diseños Intra-Sujetos o de Medidas Repetidas.

#### 3.1 Diseño Experimental Tratamientos por Sujetos.

De acuerdo a su disposición, tenemos que los cambios observados en la variable dependiente pueden obedecer a varios factores. En primer término y sin dudar, que aunque se eliminan las diferencias individuales entre los tratamientos, las ejecuciones entre los diferentes sujetos del mismo grupo, ante una misma condición difícilmente serán iguales; estas variaciones se denominan entre sujetos. Por otra parte una segunda fuente de variación y a la que esperamos se deban los cambios observados es la correspondiente a los diversos niveles manipulados y es conocida como variación entre tratamientos.

Si las características de nuestros datos, cumplen con los supuestos requeridos por el ANOVA, éste será el procedimiento que nos conduzca a el análisis de las fuentes de variación y concluir a cuál de ellas, se atribuyen los cambios en la VD.

Es posible tanto la utilización del modelo aditivo o no aditivo. Recordando que para utilizar el no aditivo es necesario antes probar que la interacción entre la VI y la va---

riable sujetos no existe, ésto a través de la aplicación de la prueba Tukey; cuyo resultado positivo indicará la conveniencia del modelo aditivo, y un rechazo de no interacción marcará como idóneo al modelo no aditivo, ambos modelos ya fueron mencionados con anterioridad.

El análisis de varianza es de dos factores, siendo el -- primero la VI y sus niveles, y el segundo la variable -- sujetos a la que Arnan (1981) dice que es más propio con siderarla como un pseudofactor.

### 3.2 Diseños Cuadrado Latino Intra-Sujeto o de Medidas - Repetidas.

La evaluación de los resultados de este diseño, también es a través del ANOVA, del que hasta aquí ya se ha vertido bastante información en lo referente a la lógica - que sigue, el planteamiento de hipótesis, etc. Es por ésto que sólo mencionaremos que la estructura del diseño, además de permitir el cálculo de la variación correspondiente a los sujetos, hace posible estimar lo que -- pueden provocar los diferentes órdenes de presentación y evidentemente la variación entre tratamientos.

## C O N C L U S I O N E S

Hay dos aspectos fundamentalmente que nos interesa dejar claros. En primer término que el diseño experimental; el cual hemos presentado como, un conjunto de pasos operativos con el fin de lograr un control de las fuentes de variación extrañas o secundarias con el objeto de evidenciar o resaltar los cambios en la variable dependiente como producto de las variaciones en la variable independiente; tiene un papel importante dentro de la Metodología experimental como un instrumento que permite abordar el objeto de estudio bajo una situación específica con diferentes grados de control. Lo anterior, considerando que la Metodología experimental comprende un conjunto de actividades tendientes a descubrir y a establecer un conocimiento científico de los fenómenos observados (Kazdin 1980 y Arnau 1978), y que se apoya para esto en diferentes mecanismos e instrumentos especiales, siendo uno de estos instrumentos el diseño experimental.

Por otra parte, hemos señalado que la estadística ofrece un conjunto de técnicas o procedimientos, para que un investigador pueda sistematizar, organizar y analizar la información proporcionada (expresiones cuantitativas) por los diseños de grupo o los conductuales, a fin de darles una interpretación y significado dentro de un contexto. La razón por la cual afirmamos que la estadística es una herramienta de análisis útil, tanto para los diseños de grupo como para los conductuales obedece a que como hicimos -

notar en ambos métodos de análisis; no numérico y numéricos, se emplean procedimientos estadísticos.

En el caso particular de los diseños de grupo la estadística ayuda a analizar sus resultados, separando y comparando básicamente la información, en la variabilidad como causa de la intervención (VI) y la adjudicada a factores extraños o secundarios con el objeto de poner a prueba -- las relaciones planteadas... "los métodos estadísticos -- permiten separar aspectos diferentes de la información -- aportada por una serie de observaciones y expresar tales aspectos en forma suficientemente condensada como para -- ser utilizable; son totalmente explícitos y por ende necesariamente conducen a los mismos resultados a investigadores diferentes". (Reuchlin N. y Hutea M., 1976. P.63).

En cuanto a la inquietud por hacer más accesible el estudio y aplicación de los métodos cuantitativos como una herramienta de análisis en la evaluación de los resultados obtenidos por un investigador, creemos haber dado un primer paso al: 1) marcar los aspectos que hay que considerar antes de la elección de una técnica, enfatizando que esto depende de las características del diseño y nunca al contrario. En cuanto a la elección del tipo de análisis se determine partiendo siempre de las características del -- diseño, este será más propio y en consecuencia se evitará la pérdida de tiempo, el tedio y la confusión de los resultados, causados por una mala elección y aplicación; --

2) Tratar de familiarizar al lector con la lógica de las pruebas, para que éste, no se centre o preocupe inmediatamente por las operaciones numéricas únicamente, sino que conociendo la lógica se le facilite el entendimiento de las operaciones. Hay que enfatizar que nunca un investigador debe esperar o encontrar en las pruebas estadísticas la solución y significación de sus resultados, sin tener presente que no son los resultados de la investigación los que se tienen que adecuar a una prueba específica, sino al contrario la estadística debe implementar o desarrollar pruebas que sistematicen y analicen los datos obtenidos.

Finalmente, nuestro trabajo no sólo hace la presentación y descripción de los diseños de grupo, sino que aparte de señalar o sugerir algunas pruebas a utilizar, dependiendo de sus características, muestra sus orígenes.

B I B L I O G R A F I A

Anastas: Anne; Psicología Diferencial. Ed. Aguilar, México, 1979, Cap. I y II.

Arnau Grass Jaime

1978. Psicología Experimental. Un enfoque Metodológico. Ed. Trillas, México.

1978 a. Métodos de Investigación en las Ciencias Humanas. Ediciones Omega, S.A. Barcelona.

1981. Diseños Experimentales en Psicología y Educación. Ed. Trillas, México.

Boring Edwin G.; Historia de la Psicología Experimental. Ed. Trillas, México, 1979.

Bunge Mario; La Ciencia : su método y su filosofía. Ed. Siglo Veinte, Buenos Aires, 1977, Cap. I, II, III.

Campbell y Stanley J.; Diseños Experimentales y Cuasi-experimentales en la Investigación Social. Ed. Amorrortu, Buenos Aires, 1973.

Castro L.; Diseño Experimental sin Estadística. Ed. Trillas, México, 1975.

Cortada de C.N.; Estadística Aplicada a las Ciencias Sociales. E.U.D.E.B.A., Argentina, 1975.

Craighead W. Edward, Kazdin Alan E.; Mahoney Michael J.; Modificación de Conducta: Principios, Técnicas y Aplicaciones. Ediciones Omega, S.A., Barcelona España, 1981.

Daniels W.; Estadística con Aplicaciones a las Ciencias Sociales y a la Educación. Ed. Mc Graw Hill, México, 1981.

Daniels W.; Bioestadística. Ed. Limusa, México, 1982.

Dominousky R. y otros; Psicología del Pensamiento. Ed. Trillas, México, 1975.

Fernández Ballesteros y Carrobbles.; Evaluación Conductual: Metodología y Aplicaciones. Ediciones Pirámide, Madrid,-- 1981, Cap. V.

Glass y Stanley.; Métodos Estadísticos Aplicados a las -- Ciencias Sociales. Ed. Prentice Hall International, Madrid 1980.

Haber Rungyon.; Estadística General. Ed. Fondo Educativo - Interamericano, S.A., México, 1973.

Hersen M. and Barlow D.H.; Single-case Experimental Desings: Strategies for studing behavior change. Ed. Pergamon Press. New York, 1976.

Holguín Quiñonez.; Estadística Descriptiva: Aplicada a las - Ciencias Sociales. Universidad Autónoma de México, México, - 1975.

Kazdin A.E.; Statistical Analyses for Single-case Experimental Desings. En Hersen y Barlow: Single-case Experimental Desing. 1976.

Kazdin Alan E.; Research Desing in Clinical Psychology. Harper y Row publishers, New York, 1980, Cap. III, V,VI y VII.

Kerlinger N.F.; Investigación del Comportamiento: Técnica y Metodología. Ed. Interamericana, México, 1975.

Kratochwill Thomas R.; Single-Subjet Research: Strategies for evaluating Change. Ed. Thomas R. Kratochwill, Academic Press Inc., New York, 1978.

Levin J.; Fundamentos de Estadísticas en la Investigación Social. Ed. Harla, México, 1979.

Magnusson.; Teoría de los Test. Ed. Trillas, México, 1978.

Meza,Morales y Magaña.; Introducción al Método Estadístico. Universidad Autónoma Metropolitana, México, 1980.

Mc Guigan.; Psicología Experimental: Un enfoque Metodológico. Ed. Trillas, México, 1971.

Pick de Weiss S. y López Velásco de Fubert A.L.; Cómo In-- vestigar en Ciencias Sociales. Ed. Trillas, México, 1979.

Plutchik Robert.; Fundamentos de Investigación Experimental. Ed. Harla, México, 1975.

Sandler J y Davidson R.; Psicopatología: Teoría del Aprendizaje, Investigación y Aplicación. Ed. Trillas, México,1977, Cap. I.

Sidman M.; Tácticas de Investigación Científica. Ed. Fontanella, Barcelona España, 1978.

Siegel Sidney.; Estadística no Paramétrica Aplicada a las Ciencias de la Conducta. Ed. Trillas, México, 1976.

Rosenblueth Arturo.; El Método Científico. CONACYT, -- México, 1981.

Reuchlin Maurice, Huteau Michael.; Introducción a la Psicología: Una guía para el Estudiante. Ed. Paidós, Buenos Aires, 1976.

Underwood Benton J.; Psicología Experimental. Ed. Trillas, México, 1973.

Wolpe Joseph.; Práctica de la Terapia de la Conducta. -- Ed. Trillas, México, 1977.

Yamane T.; Estadística. Ed. Harla, México, 1979.