

## TEMA 1: INTRODUCCIÓN AL DISEÑO DE SERIES TEMPORALES

Asignatura: *Diseños longitudinales*  
Profesor: *Salvador Chacón Moscoso*

- I. Diseños de investigación.
  1. Diseños experimentales y no experimentales.
  2. Características de los diseños cuasi-experimentales.
  3. Diferentes estrategias dentro de los diseños cuasi-experimentales.
  
- II. Evolución de los diseños cuasi-experimentales.
  1. Diseños antecedentes.
  2. El análisis experimental de Sidman.
  
- III. Tipos de diseños cuasi-experimentales.
  
- IV. Los diseños longitudinales.
  1. Perspectiva histórica.
  2. Características de los diseños longitudinales.
  3. Su representación: el cubo de datos.
  4. Clasificación de los diseños longitudinales.
  5. Series temporales v. medidas repetidas.
  6. Tipos de diseños de series temporales. DSTI intra-series e inter-series temporales.
  
- V. Bibliografía.

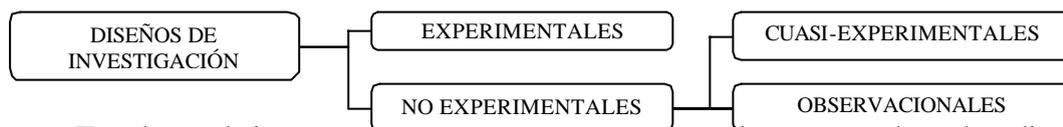
Realizado por:  
*Esther Bernal Guzmán*  
*M<sup>a</sup> José Díaz Domínguez*

## **I) Diseños de investigación:**

### **1) Estructura del diseño de investigación: Diferencia entre diseños experimentales y no experimentales.**

Podemos considerar que un diseño “*es una estrategia o secuencia de decisiones acerca de cómo recoger, ordenar y analizar los datos*” (<http://www.efdeportes.com>).

Cuando nos enfrentamos a una investigación, una de las decisiones más importantes que se han de tomar y también de las primeras, es, que tipo de diseño vamos a utilizar; evidentemente esto estará en función de varios factores (del objetivo de la investigación, de la muestra, etc.). De acuerdo con Arnau (1995) podemos distinguir dos grandes estrategias de investigación: “la experimental *versus* la no experimental”:



En la práctica surgen, por tanto, tres amplias categorías de diseños: los experimentales, los cuasi-experimentales y los observacionales.

La clave de la estrategia experimental frente a la no experimental consiste en la aleatorización, el control y la manipulación (o intervención). Por tanto, si adoptamos la estrategia experimental, debemos asignar aleatoriamente a los sujetos a las distintas condiciones o tratamientos, al tiempo que manipulamos de forma activa los factores que producen el fenómeno que estamos estudiando. La asignación aleatoria de los sujetos a los grupos persigue garantizar la validez interna, favoreciendo la comparabilidad de los grupos, y la validez de constructo, al conseguir la representatividad de la muestra y, por tanto, aumentar la generalización de los resultados de la evaluación.

Cuando un investigador utiliza la estrategia no-experimental puede optar por un diseño observacional o por uno cuasi-experimental (sobre el que nos centraremos el resto del tema). En los cuasi-experimentos, el investigador decide qué valores de la variable independiente recibirá cada uno de los sujetos y en qué momento, al igual que en los experimentos; pero a diferencia de ellos, no interviene sobre la situación ni sobre la respuesta. La diferencia con los diseños observacionales radica en que en estos últimos no se manipula la variable independiente.

Los cuasi-experimentos, por tanto, son aquellos sistemas de investigación donde el criterio de asignación de los sujetos o unidades a las condiciones de tratamiento no se rige por

las leyes del azar; ya sea por pertenecer de forma natural al grupo objeto de intervención o por circunstancias del contexto. Como afirma Helena Ramírez (1999), “Aunque los cuasi-experimentos son más vulnerables a las amenazas a la validez que las pruebas aleatorias, los cuasi-experimentos no requieren asignaciones aleatorias a los grupos experimentales y por eso son generalmente más factibles que las pruebas aleatorias”.

Ejemplo: Estamos interesados en aplicar un programa de disminución del consumo de tabaco en chicos/as de la clase de tercero de E.S.O de un colegio de Sevilla. Nos encontraríamos ante un cuasi-experimento si sólo podemos aplicar el tratamiento a la clase de tercero A, pero no en la de tercero B, donde sólo podremos medir. Por tanto, no podremos asignar a los sujetos a un grupo experimental de forma aleatoria, sino que cada chico pertenecerá al grupo-clase natural.

## ***2) Características de los diseños cuasi-experimentales.***

Como ya hemos comentado, en los cuasi-experimentos la asignación a las condiciones de intervención no es aleatoria. Esa falta de aleatorización provoca que los grupos formados no tengan equivalencia probabilística, lo que dificulta el logro de validez interna; se intenta compensar esa falta de equivalencia con otras estrategias, como el uso de grupos o medidas semejantes que puedan compararse antes y después del tratamiento. Otra mejora se da cuando se asigna a los sujetos en base a una regla conocida, ofreciendo ventajas respecto a los diseños en que no conocemos el criterio de asignación, porque permite analizar qué proporción de variabilidad observada se debe a la regla de asignación utilizada.

Por tanto, aunque el cuasi-experimento es más vulnerable a los sesgos de selección (o sea, que el grupo de tratamiento puede diferir del grupo control en características que están correlacionadas con los resultados estudiados, distorsionando los resultados del impacto) y aunque es sensible al uso de modelos estadísticos apropiados y al tratamiento correcto de los problemas de estimación estadística, cuando se diseñan, controlan y analizan apropiadamente, los cuasi-experimentos “*pueden ofrecer una evidencia casi tan fuerte del impacto del programa como la de las pruebas aleatorias y más fuerte que la mayoría de los estudios no experimentales*”, Ramírez, H. (1999).

## ***3) Diferentes estrategias dentro de los diseños cuasi-experimentales: estrategia transversal, longitudinal y mixta.***

Existen diversas clasificaciones de los diseños cuasi-experimentales; nosotros utilizaremos una de las más aceptadas, basada en el criterio temporal. Así, según Arnau, (1995), dentro del marco cuasi-experimental, los diseños pueden diversificarse de acuerdo con una determinada estrategia de recogida de datos:

A. La estrategia transversal, también denominada por comparación de grupos no equivalentes y por estrategia de grupos paralelos, consiste en la comparación de grupos de sujetos o comparación entre-sujetos: las distintas respuestas de las unidades o registros son obtenidas en un punto de corte en el tiempo, siendo el carácter básico de esta estrategia el sentido estático de la comparación. Es decir, se caracteriza por tomar un solo registro o medida de respuesta de grupos de sujetos, siendo un enfoque de corte puntual en el tiempo, de modo que se pierde la posibilidad de captar los posibles cambios que se operan en las respuestas de los individuos en función del tiempo.

Ejemplo: Siguiendo el ejemplo anterior, usaríamos una estrategia transversal si sólo medimos sus ideas sobre el tabaquismo una vez a cada grupo-clase, y comparamos las diferencias entre los dos grupos.

B. La estrategia longitudinal, puede entenderse como una alternativa al enfoque estático: la recogida suele hacerse a lo largo de una serie de puntos u ocasiones de observación en el tiempo. Es decir, consiste en registrar la misma respuesta en diferentes puntos del tiempo. Enfatiza, por tanto, el carácter dinámico del proceso, permitiendo estudiar el cambio, la evolución operada en las respuestas de un mismo grupo de sujetos. Se basa, fundamentalmente, en comparaciones de tipo intra-sujeto. La estrategia longitudinal es la mejor alternativa cuando nuestro diseño es cuasi-experimental, ya que conseguiremos mayor validez que con la estrategia transversal.

Ejemplo: Si sólo tomamos el grupo-clase al que le aplicamos el tratamiento, y medimos sus ideas sobre el tabaquismo antes de la intervención, justo después de realizarla y dos meses después, estaremos utilizando una estrategia longitudinal.

C. La estrategia mixta, registra la conducta de diferentes sujetos o grupos de sujetos en diferentes momentos temporales.

Ejemplo: Si medimos las ideas sobre el tabaquismo de los dos grupos/clases, antes de realizar la intervención y después de ésta, estaríamos usando una estrategia mixta, que combina las dos anteriores.

## **II) Evolución de los diseños cuasi-experimentales:**

### ***1. Antecedentes de los diseños cuasi-experimentales: Campbell***

Es en los años 60/70 cuando Campbell y Stanley (1963, cit. en Barlow y Hersen, 1988) y Cook y Campbell (1979, cit. en Barlow y Hersen, 1988), propusieron una serie de importantes innovaciones metodológicas que denominaron diseños cuasi-experimentales, en el área de investigación que trata el cambio social o educacional. En

éste área interesan más los efectos de los programas en amplias muestras de sujetos que el cambio conductual individual. Estos diseños, muchos de los cuales son aplicables tanto a grupos como a individuos, son relevantes en nuestro contexto.

Coincidiendo con lo mencionado en el artículo recogido en <http://www.efdeportes.com>, creemos que uno de los autores que más ha impulsado la utilización de la metodología experimental en los contextos sociales ha sido Campbell (1969), por ello, la consulta de sus trabajos es condición indispensable para la aplicación de la experimentación en contextos aplicados. En relación a los antecedentes de los diseños cuasi-experimentales, nos interesa especialmente su obra de 1966 (Campbell y Stanley, cit en <http://www.efdeportes.com>), donde se tratan los diseños “ex post facto” o “pre-experimentales”, también denominados “sin grupo control” o “sin pretest”. Estos diseños representan los principios a partir de los cuales se han desarrollado diseños más complejos.

Ya Shapiro (1959, cit. en Barlow y Hersen, 1988), había planteado un diseño donde observaba no sólo durante el tratamiento, sino también después de éste (B). Pero al no existir información sobre la situación antes del tratamiento, las mediciones carecían de validez. Este diseño también se denomina: ***Diseño de un solo grupo con prueba posterior:***

$$\boxed{X \quad O_1}$$

Se refiere a un diseño en el que se miden los resultados de un grupo de sujetos una vez que se ha aplicado la intervención, en el que no existen ni grupo control ni medidas previas a la intervención.

Un ejemplo de este diseño sería un programa diseñado para disminuir el consumo de tabaco de los alumnos de la clase de 3º de ESO. Se trata del caso en el que el programa sea implantado en dicha clase y la medida de la conducta de fumar se realice únicamente después de la aplicación del programa. Por lo que, con este tipo de diseño no se dispondrían de datos suficientes para conocer si el programa ha tenido efectos ya que no existe ningún referente con el que comparar los posibles efectos que tiene el programa.

Este tipo de diseño tiene fundamentalmente un objetivo exploratorio o de descripción de una situación, más que la validación de relaciones causa-efecto.

Surgen después los diseños A-B, donde la inclusión de la línea base (A) permite observar cambios como resultado de una intervención. Aunque la inclusión de una línea base constituye una mejora metodológica importante, este diseño es de naturaleza básicamente correlacional, y no permite establecer relaciones causa-efecto. Puede denominarse también como ***Diseño de un solo grupo con prueba previa y posterior:***

$O_1 \quad X \quad O_2$

En este tipo de diseño, el único grupo de sujetos existente es medido en la variable/s objeto del programa antes y después de la aplicación de la intervención.

Por ejemplo, siguiendo con el programa anterior, se podría realizar el registro de la conducta de fumar (nº de cigarrillos fumados) de los alumnos antes de aplicar el programa y posteriormente, una vez implementado el programa, se volvería a medir la conducta de fumar. Este tipo de diseño tiene muchas amenazas de validez, por ejemplo, el nº de cigarrillos fumados en una semana por los alumnos puede disminuir significativamente entre la medida previa y posterior a la implementación del programa. Esto podría deberse a la historia.

Este diseño de un solo grupo con prueba previa y posterior tiene poca potencia dada la gran cantidad de problemas de validez que plantea, por lo que se usa poco para generar inferencias causales. En general, sólo podría ser adecuado en aquellos casos en los que la variable medida sólo puede ser afectada por la intervención implementada y que a la vez transcurra poco tiempo entre ambas medidas.

Por otra parte, podrían utilizarse algunos procedimientos para mejorar este diseño:

1. Añadir otra medida antes de aplicar el programa. ( $O_1 \quad O_2 \quad X \quad O_3$ ).
2. Añadir la medida de una VD no equivalente antes y después de aplicar el programa. ( $O_{1A}, O_{1B} \quad X \quad O_{2A}; O_{2B}$ ).

A partir de este tipo de diseño se pueden desarrollar también variantes, como las que veremos a continuación. Estaríamos ante un diseño de series temporales equivalentes (A-B-A), que implica la manipulación experimental de variables independientes mediante el cambio de tratamientos, o mediante la introducción y la retirada de un mismo tratamiento siguiendo la secuencia A-B-A:

*A) Diseño de supresión del tratamiento ( $O_1 \quad X \quad O_2 \quad O_3 \quad X \quad O_4$ )*

Una vez aplicado el programa y medido su efecto se añade un tercer posttest y luego se suprime el tratamiento antes de realizar una medida final.

El objetivo principal de este diseño es el análisis de la validez interna. Se pretende demostrar que los resultados mejoran o empeoran con la presencia o ausencia de intervención. Este diseño presenta distintos problemas: a) en relación con la validez de conclusión estadística; b) muchos programas suponen efectos beneficiosos desde un punto de vista ético; c) suelen presentar problemas de mortalidad elevados; d) las medidas han de ser tomadas en iguales intervalos temporales.

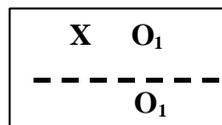
*B) Diseño de tratamiento repetido ( $O_1 \quad X \quad O_2 \quad X \quad O_3 \quad X \quad O_4$ )*

Se trata de introducir, eliminar y después volver a introducir el programa en distintos momentos en una única población cuando el efecto del programa es transitorio. Este diseño presenta los mismos problemas del diseño anterior, pero más acentuados. Entre los problemas de validez interna destacan las amenazas de maduración e historias cíclicas. También puede

darse problemas de validez de constructo de la causa cuando los usuarios del programa detectan la introducción y retirada de las intervenciones. Este diseño es más efectivo cuando las reintroducciones de las intervenciones son frecuentes y se distribuyen aleatoriamente en el tiempo.

En general, los diseños que no disponen de grupo de comparación pueden experimentar algunas mejoras mediante la "construcción de contrastes". Por otra parte, también se pueden realizar comparaciones normativas y fuentes secundarias de comparación. Aunque, estas tres posibilidades de construcción de contrastes tienen muchas debilidades, por lo que sería mejor no utilizarlos.

Por último, debemos nombrar los *Diseños con grupo no equivalente y sólo prueba posterior*:



En este caso nos referimos a un diseño en el que la medida de la conducta de fumar en los alumnos sería tomada después de implementar el programa tanto en la clase de 3° A como en otra, la clase de 3° B, donde no se aplicó. En este ejemplo, el grupo control o de comparación, serían los alumnos de 3° B, que no han participado en el programa. El objeto de utilizar este grupo de comparación no equivalente es comprobar los efectos del programa mediante la comparación de las medidas posteriores a la intervención de la conducta de fumar en ambos grupos.

El problema de este tipo de diseños puede ser de selección diferencial, es decir, los alumnos de la clase A en comparación con los otros pueden tener características diferenciales...

También se pueden introducir mejoras en el diseño para intentar atenuar algunas de su amenaza a la validez. Algunas de estas mejoras pueden ser:

a) Pruebas retrospectivas en las que se les pregunta a los sujetos acerca de las variables medidas antes de implementar el programa.

b) Pruebas antes de la implementación del programa a muestras independientes de sujetos.

c) Pruebas previas aproximadas, donde se toman medidas de variables que estén conceptualmente relacionadas y presenten una correlación con la variable sobre la que actúa el programa.

d) Uso de controles o mediante el emparejamiento de los grupos, de tal forma que se logre conformar grupos lo más similares posibles entre sí, lo que a su vez se puede combinar con el uso de múltiples grupos de comparación.

En definitiva, aproximándose al estudio del individuo desde una perspectiva diferente a la de Shapiro, Campbell y Stanley llegaron a conclusiones parecidas acerca de la posibilidad de manipulación de variables independientes y el establecimiento de relaciones causa-efecto en el estudio de un caso único.

Quizás la contribución más importante de estos metodólogos fue la descripción de varias limitaciones de estos diseños y su habilidad para descartar hipótesis alternativas plausibles (validez interna), o hasta qué punto se pueden generalizar las conclusiones obtenidas a partir de dichos diseños (validez externa).

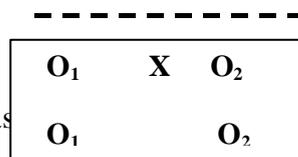
## 2. *El análisis experimental de la conducta: Sidman*

A partir de los años 70, adquiere auge el enfoque del análisis experimental de la conducta, que aporta a la investigación aplicada el desarrollo de una metodología sofisticada que permite el estudio intensivo de sujetos individuales.

Fueron Skinner y sus colaboradores quienes desarrollaron y refinaron la metodología de caso único. El punto álgido lo constituyó la publicación de un tratado metodológico definitivo por Sidman (1960, “*Técnicas de investigación científica*”). Esta obra subraya los presupuestos y condiciones de un verdadero análisis experimental de la conducta, lo que sin duda influyó en los posteriores desarrollos de los diseños cuasi-experimentales, que veremos a continuación.

### **III) Tipos de diseños cuasi-experimentales:**

#### ***1. Diseño con grupo control no equivalente y prueba previa y posterior:***



Este diseño es el más común, ya que además de las ventajas que presenta frente a los casos anteriormente descritos, es relativamente fácil de implementar. El diseño implica realizar registros antes y después de la intervención en ambos grupos (grupo objeto de la intervención y grupo no equivalente).

Siguiendo con el ejemplo anterior, nos referimos al caso en el que antes de aplicar el programa en una clase en concreto se tomarían medidas previas de la conducta de fumar tanto en los adolescentes de la clase objeto de intervención, como en los adolescentes de otra clase tomados como referente de comparación.

Su principal problema radica en la configuración de grupos equivalentes. En el caso en que no fuese posible el uso de grupos naturales equivalentes debido a las circunstancias del

contexto de intervención, la otra vía podría ser mediante el emparejamiento de sujetos en variables relevantes para el diseño. Debemos de partir de sujetos pertenecientes a una misma población.

Por otra parte, los diseños de grupo control no equivalente pueden presentar problemas éticos ligados al hecho de que un grupo de sujetos no está recibiendo los efectos de una intervención potencialmente beneficiosa. Presenta otras amenazas contra la validez:

a) La interacción entre selección y maduración. No es posible controlar las circunstancias diferenciales del ciclo vital de los sujetos.

b) Problemas de instrumentación, utilizar escalas de medidas no comparables. Problemas de sensibilidad que surgen al aplicar los instrumentos.

c) El problema de la regresión estadística.

d) Historia local diferencial de los distintos grupos tampoco es controlable.

A pesar de la existencia de importantes amenazas a la validez en este diseño, éstas son amenazas específicas y contingentes. A partir de las amenazas planteadas se han desarrollado otras opciones para mejorar las condiciones en la inferencia causal perseguida. Éstas son:

a) Doble medida previa con objeto de estudiar la estabilidad de las medidas antes de la implementación del programa.

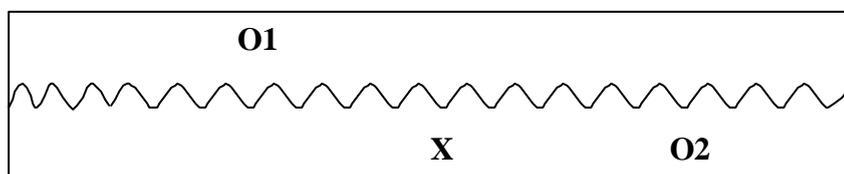
b) La medida de otra V.D. no equivalente.

c) Diseño de replicación intercambiado.

d) Grupo tratamiento y control invertidos.

Todas estas variantes se pueden resumir en tres principios: 1) Aumentar en lo posible la comparabilidad de los grupos. 2) Aumentar el nº de medidas previas a la implementación del programa. 3) Usar medidas similares a la VD en el grupo control o VD no equivalentes con las que estudiar posibles interpretaciones alternativas a la influencia del programa.

## 2. *Diseño de cohortes:*



Las cohortes son grupos de sujetos que se van sucediendo por los distintos niveles en los que se pueden organizar las poblaciones. Un ejemplo serían los colegios, centros de salud, asociaciones... Las cohortes pueden ser consideradas como grupos de comparación (control) en evaluaciones cuasi experimentales.

A modo de ejemplo de un diseño de cohorte básico tendríamos un programa de disminución del consumo de tabaco dirigido a los distintos niveles educativos de un instituto

de la E.S.O., en el que se puede plantear si ha disminuido el consumo de tabaco en la cohorte de niños de 3º E.S.O. que han recibido un tipo de intervención diferente frente a sus compañeros de 3º E.S.O. del año anterior, que no recibieron ese tipo de intervención.

La lógica de este tipo de diseño es la misma que presenta el diseño de grupo control no equivalente, su utilidad radica en la cuasicomparabilidad entre los grupos cohortes.

A pesar que este diseño supone mayores garantías de validez respecto a amenazas como maduración de los sujetos, regresión estadística y mortalidad, todavía presenta dos amenazas principales contra la validez, la selección de los sujetos y la historia.

A partir del diseño básico de cohortes se han planteado algunos desarrollos con el objeto de aumentar la validez de los resultados obtenidos. Se pueden aumentar las medidas previas y posteriores a la implantación del programa en las distintas cohortes o generar subdivisiones de grupos de tratamiento.

Por otra parte, podemos encontrar tres tipos de diseños de cohortes: 1) Diseños de cohorte con pretests. 2) Diseños con tres cohortes. 3) Extensiones de este tipo de diseños a su vez.

### 3. *Diseño de discontinuidad en la regresión:*

<b>C</b>	<b>O<sub>1</sub></b>	<b>X</b>	<b>O<sub>2</sub></b>
<b>C</b>	<b>O<sub>1</sub></b>		<b>O<sub>2</sub></b>

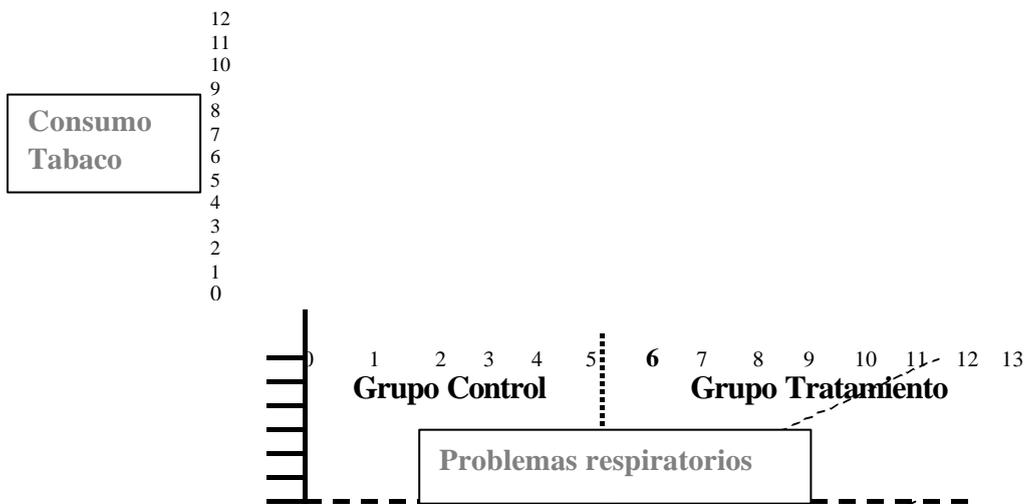
Este diseño, en su forma básica, se conforma por dos grupos de sujetos, un grupo recibe la intervención y el otro, al no recibir esta intervención, se toma como grupo de comparación. La característica principal es que la regla de asignación de los sujetos aunque no es aleatoria, sí es conocida. Los sujetos serán asignados a un grupo u otro dependiendo del nivel que presenten con respecto a una variable cuantitativa continua.

La lógica de este diseño se basa en que cada sujeto que participa en el programa puede ser clasificado de acuerdo con su puntuación en esa variable cuantitativa, y que a su vez existe un punto de corte tal que los sujetos que puntúen por encima de ese nivel recibirán la intervención y los que no lo sobrepasen no la recibirán, o viceversa.

Una vez realizada la asignación de los sujetos supongamos que conectásemos las dos rectas de regresión obtenidas en ambos grupos. En este caso, en tanto que sólo se ha realizado una medida previa se obtendría una línea continua a ambos lados del punto de corte. En cambio en una medida a posteriori, una vez implantada la intervención a los sujetos que estaban por encima del punto de corte, es de esperar que se produzca un salto o discontinuidad entre ambas líneas de regresión si dicha intervención ha tenido algún efecto. Por lo que, el modelo puede analizar la gran mayoría de las amenazas a la validez.

Un ejemplo de este diseño sería un programa de disminución del consumo de tabaco en alumnos con problemas respiratorios; suponiendo que estas dos variables guarden una relación lineal creciente. En primer lugar, debemos decidir cuál será el punto de corte a partir del que consideraremos que los pacientes necesitan el tratamiento, por ejemplo, los que tengan problemas respiratorios serios (punto 6 en nuestra escala). Así, clasificamos a todos los pacientes como del grupo con problemas moderados o más, (al que aplicaremos el tratamiento), frente al grupo con poco problemas, (que medimos para usarlo como control).

### Tratamiento



Así, el efecto de un programa efectivo se traduce en detectar una discontinuidad en la recta de regresión a partir del punto de corte y un cambio en la pendiente de la recta.

El diseño de discontinuidad en la regresión tiene potencia cuando nos referimos a relaciones lineales entre variables, ya que en el momento en que la relación sea curvilínea se estará sesgando la interpretación. Es decir se rompería el principio por el que el cambio entre el pretest y el postest sigue un patrón lineal desde la distribución del pretest, por lo que no se podría asegurar si el cambio de nivel en el postest se debe a la implementación del programa.

Las restricciones de este diseño, punto de corte y relación lineal, hace que los resultados encontrados sean difíciles de generalizar. La aplicabilidad de este diseño a su vez se complejiza debido a la dificultad de disponer de un punto claro de corte a partir del cual realizar la asignación de los sujetos. Una solución posible a este problema es el uso de área o intervalos de corte en vez de puntos de corte. El problema de esta solución es que se pierde potencia estadística y la posibilidad de descartar interpretaciones alternativas ligadas al punto de corte exacto.

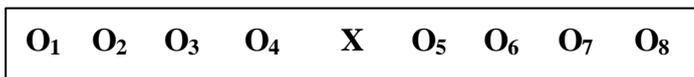
El hecho de presentar unas condiciones de aplicación tan precisas hace que se posibilite una estimación insesgada del efecto del programa; de ahí que algunos autores consideran el diseño a medio camino entre el experimento auténtico y el diseño cuasi-experimental sin conocimiento de regla de asignación.

El diseño de discontinuidad en la regresión se hace muy útil en el marco de evaluación de programas, ya que los resultados obtenidos controlarán en gran medida las amenazas a la validez, y presenta una alta potencia para detectar el efecto causal del programa.

#### **4. Diseños de grupo único o diseños de series temporales interrumpidas:**

Son diseños que no presentan grupo de comparación; tratan de constatar con exactitud cuál es el nivel que presenta la/s variables sobre las que se va a intervenir antes de implementar el programa, de tal forma que una vez realizada la intervención se pueda apreciar si ha habido cambio en ese nivel inicialmente detectado. En este caso, la comparabilidad no se busca entre grupos, sino entre series de mediciones de las variables objeto de la evaluación en un mismo grupo de sujetos antes y después de implementar el programa.

##### **a) Diseño de series temporales interrumpidas**



La estrategia de este diseño se reduce a tomar una serie de medidas de un grupo de sujetos antes y después de la implantación del programa.

El principio de este diseño se fundamenta en que las medidas previas y posteriores a la implantación del programa deberían presentar una tendencia o estructura regular. La validez del análisis del efecto del programa se basa en la detección de dichos patrones regulares de datos registrados o saber hasta que punto la estructura de la serie ha quedado alterada como consecuencia del programa.

Los efectos de la implementación del programa pueden apreciarse en las diferencias entre las series obtenidas antes y después de la implantación con respecto a distintos aspectos: a) un cambio de nivel, b) un cambio de tendencia una vez aplicado el programa.

Este tipo de diseños es útil cuando la implementación del programa o intervención se realiza en un momento temporal específico y se espera que provoque un efecto inmediato.

A modo de ejemplo, supongamos que en el centro de enseñanza anterior se realizan una serie de medidas durante un mes en las que se detecta una frecuencia muy alta de consumo de tabaco en los alumnos. Una vez que son verificadas la estabilidad y regularidad de estas conductas, los profesionales del centro instauran un programa para disminuir las conductas relacionadas con el consumo de tabaco. Una vez implementado el programa se produce una disminución significativa de estas conductas.

En este tipo de diseños hay dos vías para potenciar la validez: 1. Priorizar una perspectiva de análisis, donde a partir de los datos registrados mediante procedimientos analíticos se estiman las tendencias aportando criterios desde los cuales tomar la decisión de cuando realizar la intervención.

A pesar de ello, los problemas más frecuentes que presenta este tipo de diseños están ligados al estudio de la estabilidad en las series obtenidas. Otro problema que puede ocurrir es que antes de aplicar el programa se haya producido un cambio extremo en las variables objeto de evaluación.

En resumen, las principales amenazas a la validez interna se pueden minimizar utilizando estrategias de diseño, que pueden estar ligadas tanto al estudio de la estabilidad de las series como al análisis de todas las variables que pueden incidir al mismo tiempo que la implementación del programa. Por otra parte, con respecto a la validez externa, al tratarse de un único grupo de sujetos, la posibilidad de extrapolar resultados es escasa y sólo se sustenta en disponer de datos de archivo de las características de los sujetos que han participado en el programa. Pese a los problemas de validez que hemos comentado, supera a los diseños anteriores en el logro de la misma; ahí radica su importancia.

A partir del diseño básico de series temporales interrumpidas se han desarrollado los siguientes tipos de diseños:

**5. Desarrollos a partir del diseño DSTI:**

- a) Diseño de serie temporal interrumpido con grupo control no equivalente;

O <sub>1</sub>	O <sub>2</sub>	O <sub>3</sub>	O <sub>4</sub>	X	O <sub>5</sub>	O <sub>6</sub>	O <sub>7</sub>	O <sub>8</sub>
O <sub>1</sub>	O <sub>2</sub>	O <sub>3</sub>	O <sub>4</sub>	-	O <sub>5</sub>	O <sub>6</sub>	O <sub>7</sub>	O <sub>8</sub>

Este tipo de diseño tiene como objetivo obtener más evidencias sobre el posible efecto de la historia. Siguiendo con el ejemplo anterior, supongamos que las series tomadas a los niños del centro de enseñanza antes de aplicar el programa también fueran tomadas en otro centro de características similares tomado como serie de comparación.

- b) Diseño de series temporales interrumpidas con variables dependientes no equivalentes;

O <sub>1A</sub>	O <sub>2A</sub>	O <sub>3A</sub>	O <sub>4A</sub>	X	O <sub>5A</sub>	O <sub>6A</sub>	O <sub>7A</sub>	O <sub>8A</sub>
O <sub>1B</sub>	O <sub>2B</sub>	O <sub>3B</sub>	O <sub>4B</sub>	(X)	O <sub>5B</sub>	O <sub>6B</sub>	O <sub>7B</sub>	O <sub>8B</sub>

Otro modo de amortiguar el posible efecto de la historia sería mediante el uso de este tipo de diseño, por lo que los datos serían recogidos de distintas variables dependientes no equivalentes (rendimiento escolar o estabilidad emocional) que estarían sometidos a los mismos efectos de la historia, pero el programa sólo se implantaría en el grupo de escolares que participan en la condición experimental.

- c) Diseño de series con replicaciones múltiples

O <sub>1</sub>	O <sub>2</sub>	O <sub>3</sub>	O <sub>4</sub>	O <sub>5</sub>	O <sub>6</sub>	O <sub>7</sub>	O <sub>8</sub>	X	O <sub>9</sub>	O <sub>10</sub>	O <sub>11</sub>	O <sub>12</sub>	O <sub>13</sub>	O <sub>14</sub>	O <sub>15</sub>	O <sub>16</sub>
----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	---	----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------

En algunos contextos de intervención es posible introducir el programa, eliminarlo posteriormente, volverlo a introducir y eliminarlo de nuevo sucesivamente. Este diseño se puede usar si la variable medida responde de manera similar en cada ocasión que se introduzca el programa aunque de modo distinto cuando se elimina, es el denominado efecto reversión. Aunque este diseño es muy útil para obtener inferencias válidas, presenta muchos problemas de aplicabilidad.

d) Diseño de series temporales interrumpidas con replicaciones cambiadas.

O <sub>2</sub>	O <sub>3</sub>	O <sub>4</sub>	O <sub>5</sub>	O <sub>6</sub>	X	O <sub>7</sub>	O <sub>8</sub>	
O <sub>2</sub>	O <sub>3</sub>	O <sub>4</sub>	O <sub>5</sub>	O <sub>6</sub>		X	O <sub>7</sub>	O <sub>8</sub>

En este tipo de diseños se utiliza dos muestras de sujetos, de tal forma que cuando una de ellas recibe el programa la otra no la recibe y viceversa. La principal ventaja de este diseño, además de aportar más evidencias para el estudio de la validez interna, es que potencia la validez externa ya que el efecto del programa puede ser encontrado en dos muestras de sujetos en momentos y contextos diferentes. Otra ventaja es que pueden encontrarse efectos retardados del programa. Su principal problema sería que dos eventos diferentes (efecto historia) actuaran de manera diferencial cada vez que se aplicara el programa en cada una de las muestras.

**IV) Los diseños longitudinales:**

Hemos incluido éste último punto para clarificar y ampliar el tipo de diseños que estudiaremos el resto del curso. Como ya hemos comentado anteriormente, dentro de los diseños cuasi-experimentales podemos utilizar diferentes tipos de estrategias: longitudinal, transversal y mixta; siendo la longitudinal la que más nos interesa, ya que permite solventar algunos de los problemas que la estrategia transversal presenta en los diseños no experimentales. De ahí, que la mayoría de los diseños cuasi-experimentales vistos hasta ahora utilicen una estrategia longitudinal.

**1. *Perspectiva histórica:***

Dentro de un contexto netamente social, donde el interés se centraba en evaluar las actitudes de los ciudadanos en torno a diferentes problemas de carácter sociopolítico, surge la necesidad de un tratamiento estadístico de los datos longitudinales, a partir del año 1944. Es decisiva la aportación de Campbell y Stanley (1963/1966, cit en Arnau, 1995), quienes dentro del ámbito cuasi-experimental, proponen el Análisis del panel transretardo y los diseños de

series temporales interrumpidas, así como otra clase de esquemas de investigación, y cuya influencia va a ser decisiva para el posterior desarrollo de los Diseños Longitudinales.

De forma simultánea, dentro del ámbito social y educativo, aumenta la importancia de estos diseños al incrementarse los estudios sobre el desarrollo y el cambio: autores del movimiento del “ciclo vital”, como Baltes, destacaron la necesidad de una metodología rigurosa y de carácter longitudinal en diferentes publicaciones.

A lo largo de los años 60 se incrementa de forma espectacular la metodología de análisis de datos de series temporales. Diseños alternativos fueron descritos por Campbell y Stanley (1963). Box y Tiao proponen en 1965 los modelos ARIMA, los métodos estadísticos para el análisis adecuado de las series temporales.

A finales de los 70, Cook y Campbell publican un trabajo sobre los principales Diseños cuasi-experimentales, con especial énfasis en los longitudinales. En la misma fecha, Nesselroade y Baltes recogen aportaciones sobre la metodología y el tratamiento de datos longitudinales.

En los últimos años, destacan los trabajos con marcado carácter estadístico, que aportan mejoras en el tratamiento de los datos longitudinales.

## ***2. Características de los diseños longitudinales:***

Al introducir la dimensión temporal se abren nuevas perspectivas explicativas, puesto que se puede aprovechar la información obtenida del “sentido del cambio” que presentan los datos u observaciones a lo largo de esa dimensión. Como ha señalado Kandel (cit. en Arnau, 1995), los diseños longitudinales constituyen los instrumentos más potentes disponibles a los científicos sociales, dado que proporcionan un conjunto óptimo de datos para la prueba de asunciones causales. Por ello, en los diseños no experimentales, la estrategia longitudinal es la mejor alternativa a la investigación transversal; entre otras ventajas, permite medir los cambios para cada caso o sujeto, sin que para ello sea necesario inferirlo a partir de la diferencia entre los individuos. Esto evita algún tipo de incertidumbre y permite establecer, de forma más clara, la acción causal de una variable.

Por tanto, se plantea el diseño longitudinal como una alternativa a la investigación transversal. Esto es debido no sólo a las mejores posibilidades que ofrecen estos esquemas para la inferencia de relaciones válidas, sino también a que dentro del contexto de las ciencias sociales y del comportamiento, interesa adoptar un enfoque dinámico, que implica la obtención de datos a lo largo del tiempo más bien que a partir de un punto de corte.

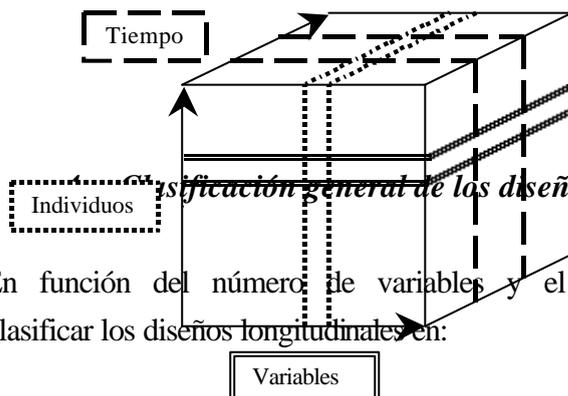
## ***3. La representación de los datos longitudinales: el cubo de datos***

Un modo de representar los datos longitudinales sería mediante "el cubo de datos". Las entradas de este cubo representa la triple dimensionalidad de los datos. Es decir, las filas

suelen representar a los individuos, las columnas a los instrumentos de observación (variables) y los estratos (sectores) a las ocasiones.

Así mismo, el hecho de que los datos pueden ser ordenados en el cubo de datos ilustra al menos dos nociones básicas sobre la naturaleza de los datos longitudinales: 1) El hecho de que posea tres dimensiones indica un mayor grado de complejidad en los datos; 2) El modo o la dimensión tiempo en el cubo de datos, posee una propiedad de orden de la que carecen los otros dos modos. Esto implica una restricción en el momento de reordenar los datos.

DISEÑOS LONGITUDINALES:  
CUBO DE DATOS



*Clasificación general de los diseños longitudinales:*

En función del número de variables y el número de unidades de observación, podemos clasificar los diseños longitudinales en:

A) **Diseños de medidas repetidas.** Son los diseños que incorporan más de una unidad de observación. Dichas unidades pueden formar parte de un solo grupo/muestra, o pueden estar distribuidas en diferentes grupos de acuerdo con alguna variable o criterio de clasificación. Cada una de las unidades de investigación es registrada a lo largo de una serie reducida de intervalos de tiempo u ocasiones. Por tanto, implican una cantidad restringida de datos y en intervalos de tiempo muy espaciados. Podemos diferenciar dos subtipos en función de:

- DE MEDIDAS REPETIDAS:**
  - DE PANELES:** Diseños que incorporan más de una unidad de observación.
  - DE SERIES:** Diseños que incorporan más de una unidad de observación.
- INTERMITENTES (DSIT):** Diseños que incorporan más de una unidad de observación.

Además, se mencionan **TIPOS DE DISEÑOS LONGITUDINALES** y **BIVARIABLES**.

de si usan una o más muestras de sujetos, las cuales pueden ser medidas en una o más variables.

Uno de los diseños más simples de medidas repetidas es conocido por Diseño antes y después: estructura formada por un solo grupo de sujetos registrados en dos momentos u ocasiones, antes y después de algún hecho, circunstancia o tratamiento. Dado que los sujetos son medidos dos veces, estos esquemas suelen denominarse *Diseño pretest-postest*; son especialmente apropiados para evaluar el cambio que se produce en las medidas tomadas en dos ocasiones distintas en el tiempo.

Ejemplo: siguiendo con el ejemplo anterior, nos encontraríamos ante un diseño de medidas repetidas si tomamos medidas del número de cigarrillos fumados por cada persona del grupo/clase de Tercero A y del grupo de Tercero B, en varias ocasiones antes de implementar el programa, un día después de implementarlo, y dos meses después.

Una de las grandes ventajas de la estrategia intrasujeto radica en que las comparaciones se hacen dentro de los sujetos (estrategia intrasujeto), y no entre los sujetos (propio de los diseños clásicos), lo que evita que la comparación quede afectada por las diferencias individuales. Desde una perspectiva longitudinal, los datos de respuesta o medidas de la variable objeto de estudio de cada uno de los sujetos son función del tiempo, y en consecuencia, el Diseño de medidas repetidas se convierte en un instrumento apropiado para la modelización de curvas de crecimiento y evaluación de los procesos de cambio (perfiles) en contextos sociales, evolutivos y educativos. Constituye, por tanto, un enfoque de estudio idóneo, cuando el investigador se propone estudiar las tendencias que presentan los datos, en función del tiempo.

Existen enfoques alternativos en el estudio del cambio, que ya hemos analizado con más detalle: diseño de cohortes y diseño de discontinuidad en la regresión.

B) Diseños de panel: El acento se pone en la cantidad de variables que se registran, así como en la cantidad de tandas de observaciones (o tomas de datos) que se distribuyen, de forma muy restringida, a lo largo del tiempo. Se trabaja con un solo grupo de sujetos: lo importante es la dirección de los efectos de las variables a lo largo de las diferentes tandas. Datos de una sola muestra en una o más variables. Por tanto, la diferencia con los diseños de medidas repetidas radica en que en los de panel se toman muy pocas medidas, y de un solo grupo de sujetos, mientras que en los de medidas repetidas se toma un mayor número de medidas y de un grupo o de varios grupos de sujetos.

Ejemplo: Siguiendo con el ejemplo anterior, estaríamos ante un diseño de panel si observáramos el número de cigarrillos fumados por cada persona de uno sólo de los grupos/clases en dos ocasiones, antes y después de aplicar el tratamiento.

Es una de las modalidades de Diseño más ampliamente utilizadas en investigación social, de carácter longitudinal. Requiere, en su formato más elemental, que de una misma

muestra de sujetos (o casos) se tomen registros o medidas de dos o más variables, en intervalos diferentes a lo largo del tiempo. Como destacan algunos autores, los diseños en panel constituyen una combinación de los diseños de series temporales y los transversales, donde las medidas son obtenidas en cortes o tandas en cada punto de observación. La estructura del Diseño en panel admite diferentes modalidades según se combinen las variables y las tandas (o puntos de observación): 2 tandas y 2 variables (el más simple: 2W2V); 3 tandas y 2 variables: 3W2V; etc.

C) Diseños de series temporales: Una serie temporal es una secuencia cronológicamente ordenada de valores de medición, a lo largo de cincuenta o más puntos equidistantes de tiempo. Se trabaja con un solo dato por punto del tiempo y por variable medida, es decir, datos de un solo sujeto en una o más variables. Lo que se pretende es la búsqueda de regularidades del comportamiento del usuario del programa en el tiempo, así como el pronóstico o predicción de su comportamiento más probable en el futuro. Se suelen plantear en procesos de duración prolongada.

Ejemplo: En el programa antes citado vamos a medir el número de cigarrillos fumados por la clase, cada día en dos ocasiones, durante tres meses, antes y después de aplicar nuestro programa.

Por tanto, estaríamos hablando de una clase de diseños donde un solo sujeto o unidad genera un conjunto o más de un conjunto de datos/observaciones, que son registradas a lo largo de una serie de puntos en el tiempo. El diseño puede ser caracterizado como de  $N = 1$ , tanto en su vertiente uni como bivariable (es decir, con registros de una o más variables de medida o dependientes).

### ***5. Diferenciación entre los diseños de series temporales y los de medidas repetidas:***

Como ya se ha comentado, la principal diferencia reside en que en los diseños de series temporales sólo hay un sujeto o unidad observacional, mientras que en los de medidas repetidas hay varias unidades observacionales, pertenecientes a una o más muestras o grupos de sujetos.

Por tanto, según Arnau (1995) “el paso del Diseño de series temporales al de medidas repetidas, va paralelo al cambio entre la utilización de una sola unidad a un grupo de unidades”.

En consecuencia, en los diseños de series temporales se trabaja con un solo dato por punto del tiempo y variable medida, lo que implica que se necesiten más mediciones: da lugar a una gran cantidad de registros, de forma sucesiva y secuencial. Los diseños de medida repetida implican más de un dato por punto del tiempo y obtienen una cantidad más restringida de datos y en intervalos de tiempo mucho más espaciados.

Todo lo anterior conlleva que se diferencien en aspectos metodológicos y en el análisis de los datos. Para un correcto análisis de los datos de un diseño de series temporales lo más adecuado es utilizar un modelo ARIMA, que requiere al menos que contemos con 50 mediciones, mientras que los datos de un diseño de medidas repetidas, suelen analizarse mediante modelos condicionales (o de regresión) e incondicionales (Arnau, 2001).

- Ejemplos: estaríamos ante un diseño de series temporales si en un programa para reducir el absentismo escolar de la clase de primero de secundaria de un colegio, medimos el grado de absentismo del grupo en días alternos durante tres meses, y después de llevar a cabo nuestra intervención, volvemos a medir cada día durante tres meses. Sería de medidas repetidas si para llevar a cabo el programa anterior, medimos el grado de absentismo de cada chico/a que conforma la clase (y no sólo el del grupo clase), en varias ocasiones antes y después del tratamiento.

**6. Tipos de diseños de series temporales:** En función del número de variables podemos diferenciar entre:

**A) Diseños de series temporales bivariantes:** diseños con registro simultáneo de dos variables o dimensiones de variación. Se caracterizan por dos sistemas paralelos de observaciones o registros y el principal interés estriba en el estudio de la relación entre ellas.

Ejemplo: Si al aplicar el programa de intervención antes mencionado, además de medir el número de cigarrillos fumados por la clase, medimos las ideas sobre el tabaquismo, estaríamos ante un diseño de series temporales bivariantes.

**B) Diseños de series temporales interrumpidas (DSTI):** Este tipo de diseño que ya ha sido comentado al hablar de las tipologías de diseños cuasi-experimentales, se ampliará en este apartado, dada su importancia dentro de la asignatura. Los DSTI se caracterizan por el hecho de que se toman o registran medidas antes y después de alguna intervención. Las observaciones antes de la intervención reciben el nombre genérico de “línea base”, mientras que las de después se denominan “fase de intervención o tratamiento”.

Son una práctica común en las Ciencias Sociales y del comportamiento, donde se toman medidas o registros de una sola unidad observacional o sujeto (diseños de series temporales), habiendo ocurrido una circunstancia o intervención capaz de afectar las medidas o registros de la variable dependiente (interrumpidas):  $O_1 O_2 O_3 O_4 X O_5 O_6 O_7 O_8$ .

Ejemplo: En el programa antes citado vamos a medir el número de cigarrillos fumados por la clase, cada día en dos ocasiones, durante tres meses, antes y después de aplicar nuestro programa.

a) Técnicas de análisis: Como afirma Arnau, J. (2001), tradicionalmente, los datos generados por esta clase de diseños han sido analizados por medio de procedimientos gráficos, o por la aplicación de las técnicas estadísticas convencionales (como las pruebas *t* y *F*), a fin de comparar las medias de los períodos antes y después de la intervención. Ambos enfoques no son válidos porque no tienen en cuenta la posible estructura de dependencia que suele darse en los datos.

Un enfoque alternativo que parece bastante prometedor, consiste en la identificación de un modelo Autoregresivo Integrado de Medias Móviles (ARIMA). A pesar de las ventajas de este tipo de análisis, en la práctica los modelos ARIMA presentan dos grandes problemas:

- Aún cuando se disponga de una cantidad de datos suficiente y se dominen las técnicas estadísticas, con frecuencia se suele identificar incorrectamente el modelo.

- La necesidad de un mínimo de 50 observaciones antes y después de la intervención, lo que es poco frecuente en contextos aplicados, tales como diseños clínicos de un solo sujeto o en los estudios de evaluación de programas de una sola unidad.

La modelización estadística de las series temporales es el conjunto de procedimientos que permiten conocer la estructura latente de las series observadas, siendo en palabras de Arnau, J. (2001) “la única vía de acceso a los procesos o mecanismos ocultos responsables de la serie”. Por ello, el problema principal al que se enfrenta el analista de series temporales es la identificación del modelo, que es asumido como el auténtico mecanismo causante de la estructura de dependencia de la serie. Por tanto, la labor del analista consiste en detectar los componentes sistemáticos que son la causa de la dependencia serial, es decir, identificar la estructura latente del mecanismo generador, lo cual puede hacer identificando un modelo ARIMA. Desde el punto de vista de la investigación psicológica, estos modelos permiten evaluar el impacto de las intervenciones, pero para ello se requiere la construcción de un modelo global, donde se suman el modelo ARIMA y el modelo del impacto. Esto hace posible concluir si la variable tratamiento ha afectado a la serie, tras la previa transformación de la serie mediante el modelo adecuado. Esta técnica de evaluación del impacto es especialmente importante en aquellos ámbitos de aplicación en que usan diseños cuasi-experimentales, como es nuestro caso.

b) Por último, podemos diferenciar entre dos tipos de DSTI: los diseños intra-series e inter-series temporales. Los diseños intra-series se fijan en los datos dentro de la misma serie de punto de datos, frente a los diseños inter-series, en cuya estrategia se comparan los resultados entre dos series separadas de puntos de datos.

1). Los diseños intra-series temporales:

Si partimos del diseño de series temporales interrumpidas básico, nos encontramos con un diseño A-B. Este tipo de diseño tiene graves amenazas de validez interna, como recogió Campbell en un artículo de 1969 (cit. en Barlow y Hersen, 1988). Entre los problemas

metodológicos, destaca el hecho de que no podemos saber si las variaciones registradas entre la fase A (o línea base) y la fase B (tratamiento), se deben a la intervención, o a una tercera variable, como el paso del tiempo. Para solventarlo, podemos utilizar un grupo control equivalente, asegurándonos así de que las variaciones entre la fase A y la fase B en el grupo que recibe el tratamiento son debidas a éste. Otra estrategia con la que se mejora el control experimental es la de series interrumpidas con retirada de tratamiento o diseño A-B-A: los diseños de retirada son aquellas estrategias de análisis en los que la variable tratamiento es introducida y luego retirada; el más simple es el diseño A-B-A.

Frente al diseño A-B que tiene problemas con la certeza de la influencia del tratamiento, el A-B-A posibilita un análisis de los efectos de control de la introducción y retirada de la variable tratamiento. Si la aplicación del tratamiento supone una mejora respecto a la línea base, y su retirada, un empeoramiento, podemos concluir con bastante seguridad que la variable tratamiento es la responsable de los cambios producidos. A pesar de que la estrategia es aceptable desde un punto de vista experimental, tiene un problema ético en el contexto clínico: desafortunadamente para el paciente/sujeto, el diseño termina con la fase A (línea base), por lo que no le aporta los beneficios completos del tratamiento experimental. Otro problema es el de la confusión secuencial en los diseños A-B y sus variantes, que implica cierta limitación en la generalización al campo clínico. De manera parecida, la segunda introducción de la variable A puede afectar a la conducta de manera diferente que la primera introducción (en general, la conducta mejora más rápidamente en una segunda introducción del tratamiento).

- **Series interrumpidas con replicaciones o diseño A-B múltiple: Página 160/161**

## 2). Los diseños inter-series temporales y de series combinadas:

Como ya hemos comentado, la lógica básica de este diseño requiere la comparación de dos series separadas de puntos de datos.

En primer lugar, hablaremos de los diseños de tratamientos alternos (ATD). Consisten en la alternancia rápida de dos o más tratamientos o condiciones en un sujeto único. Esta estrategia para ver la eficacia relativa de dos tratamientos o condiciones, presenta mejoras respecto a la tradicional comparación entre grupos, que podía presentar problemas por las diferencias individuales. Por tanto, lo que interesa no es tanto la tendencia a mejorar en el tiempo, sino la comparación entre los dos tratamientos; de ahí que se incluya dentro de los diseños entre series.

Por otro lado, nos encontramos con los diseños de líneas de base múltiple. **Pág. 191.** En esta técnica se identifican y miden un cierto número de respuestas en el tiempo para proporcionar líneas de base en contraste con las cuales se pueden evaluar los cambios. Después se aplica una variable experimental a una de las conductas, provocando un cambio y quizá se observa poco o ningún cambio en las otras líneas de base. Después, el

experimentador aplica la misma variable experimental a una segunda conducta y observa los cambios de tasa en dicha conducta. A pesar de que los efectos de control de la variable de tratamiento es más débil en el diseño de línea base múltiple, que en el de retirada, tiene una gran ventaja: fomenta la medida simultánea de diversas conductas objetivo concurrentes. Esto tiene una gran importancia porque el registro de conductas concurrentes permite una mayor aproximación a las condiciones naturales, donde una variedad de respuestas se está produciendo al mismo tiempo.

### **Conclusión:**

A modo de resumen, hemos redactado este apartado en el que destacamos los diseños más importantes, sus ventajas e inconvenientes.

En primer lugar, hemos diferenciado entre diseños experimentales y cuasi-experimentales; nos hemos parado en estos últimos, por su relevancia en la investigación en el campo de las ciencias sociales y del comportamiento, donde no siempre es posible/deseable la asignación aleatoria de los sujetos a las muestras.

#### 1) Evolución de los diseños cuasi-experimentales:

Comenzamos con la aparición de los diseños de un solo grupo con prueba posterior, los cuales carecen de todo tipo de validez experimental. Ejemplo: si medimos el consumo de tabaco en la clase A, después de nuestra intervención. Al no existir referente de comparación, no sabremos si el tratamiento ha sido válido. Para ello necesitaríamos incluir una prueba previa: Diseño de un solo grupo con prueba previa y posterior: esta vez, medimos antes y después de aplicar el tratamiento. Pero la historia u otras variables pueden afectar a la validez, por lo que sería necesario usar otros diseños que mejoren ese aspecto. Surgen así los diseños con grupo no equivalente y sólo prueba posterior: Medimos el consumo de cigarrillos en la clase A, donde se aplicó el tratamiento, y en la clase B, donde no se aplicó. Pero también en este tipo de diseño se daban grandes problemas como los de selección diferencial, etc.

Por tanto, en estos primeros diseños, existían graves problemas de validez interna y externa, por lo que pronto fueron sustituidos por otros diseños que intentaron solventar esas dificultades, los denominados diseños cuasi-experimentales.

#### 2) Tipos de diseños cuasi-experimentales:

El diseño con grupo control no equivalente y prueba previa y posterior, es una derivación lógica de los anteriores diseños, e intenta superarlos; pero presenta algunos inconvenientes como: la dificultad de configuración de grupos equivalentes, problemas éticos, y otras amenazas contra la validez como la interacción entre maduración y selección y la historia local diferencial de los distintos grupos.

Surge también el diseño de cohortes, cuya utilidad radica en la cuasi-comparabilidad entre los grupos cohortes. Supone mayores garantías de validez respecto a amenazas como

maduración de los sujetos, regresión estadística y mortalidad, pero sigue presentando amenazas: selección de los sujetos e historia.

El diseño de discontinuidad en la regresión permite analizar la gran mayoría de amenazas a la validez, y además, el hecho de presentar unas condiciones de aplicación muy precisas, posibilita una estimación insesgada del efecto del programa. Pero su potencia decrece cuando la relación entre las variables no es lineal. Las restricciones de este diseño (punto de corte y relación lineal) hacen que los resultados sean difíciles de generalizar.

El diseño de grupo único DSTI: a diferencia de los anteriores diseños, la comparabilidad no se busca entre-grupos, sino entre series de mediciones de las variables objetos de evaluación en un mismo grupo de sujetos antes y después de implementar el programa. Por tanto, se minimizan problemas presentes en los diseños anteriores, y se evitan amenazas como la de la selección diferencial de la muestra, o la no comparabilidad de los grupos; etc. A pesar de ello, puede presentar ciertas amenazas ligadas al estudio de la estabilidad en las series obtenidas, que antes de aplicar el programa se haya producido un cambio extremo en las variables objeto de evaluación, o que al tener un solo grupo, la posibilidad de extrapolar los datos sea escasa. Estas amenazas a la validez interna y externa se pueden minimizar utilizando estrategias como el estudio de la estabilidad de las series, datos de archivo anteriores, etc. Por las grandes ventajas de este diseño, es uno de los que más interés ha despertado en investigación cuasi-experimental, lo que ha dado lugar a la aparición de diversos diseños desarrollados a partir de él.

- Desarrollos a partir del DSTI:

a) DSTI con grupo control no equivalente: respecto al DSTI, su mejora es que obtiene más evidencias sobre el posible efecto de la historia.

b) DSTI con variables dependientes no equivalentes: es otro modo de amortiguar el posible efecto de la historia.

c) Diseño de series con replicaciones múltiples: es muy útil para obtener inferencias válidas, pero presenta problemas de aplicabilidad.

d) DSTI con replicaciones cambiadas: aumenta la validez externa, ya que el efecto del programa puede ser encontrado en varias muestras de sujetos en momentos y contextos diferentes.

3) Estrategias: Dentro de los diseños cuasi-experimentales, hemos diferenciado entre la estrategia longitudinal, transversal y mixta; en los diseños no experimentales la longitudinal es la mejor alternativa, porque permite medir los cambios para cada sujeto sin que para ello sea necesario inferirlo a partir de la diferencia entre los individuos. Por tanto, la longitudinal ofrece mejores posibilidades para la inferencia de relaciones válidas y, además, dentro del contexto de las ciencias sociales y del comportamiento, interesa adoptar un enfoque dinámico, que implica la obtención de datos a lo largo del tiempo. Dentro de los diseños longitudinales,

los que más nos interesan son los de series temporales, cuyas ventajas acabamos de destacar, y que derivan de la comparación de una gran cantidad las medidas de la variable en un único grupo de sujetos, a lo largo del tiempo. Con los diseños de series temporales, frente a los de medidas repetidas y a los de panel, lo se que pretende es la búsqueda de regularidades del comportamiento del usuario del programa en el tiempo, así como el pronóstico de su comportamiento más probable en el futuro; por lo que es especialmente adecuado plantearlos en procesos de duración prolongada, como el caso/ejemplo que nos ocupa. Así, en el programa de disminución del consumo de tabaco, lo más adecuado sería usar este tipo de diseños, que recoge un mayor número de datos de un solo sujeto en una o más variables, favoreciendo así la validez.

#### V) **Bibliografía:**

- Arnau, J. (1995). **Diseños longitudinales aplicados a las Ciencias Sociales y Comportamentales**. México: Limusa.
- Arnau, J. (2001). **Diseños de series temporales: técnicas de análisis**. Barcelona: Universitat de Barcelona.
- Barlow, D. H. y Hersen, M. (1988). **Diseños experimentales de caso único**. Barcelona: Martínez Roca.
- Chacón, Blanco & Anguera (Eds.) (En preparación). **Evaluación de programas sociales y sanitarios**. Un abordaje metodológico. Madrid: Síntesis.
- **Lecturas: Educación Física y Deportes**. Año 8. Nº 46.  
<http://www.efdeportes.com/efd46/invest.htm>
- Ramírez, Helena (1999). ¿Qué son los cuasi-experimentos? <http://www.cpc.unc.edu/mirrors/ccp.ucr.ac.cr/cursoweb/242cuas.htm>
- Sidman, M. (1973). **Tácticas de investigación científica**. Barcelona: Fontanella.