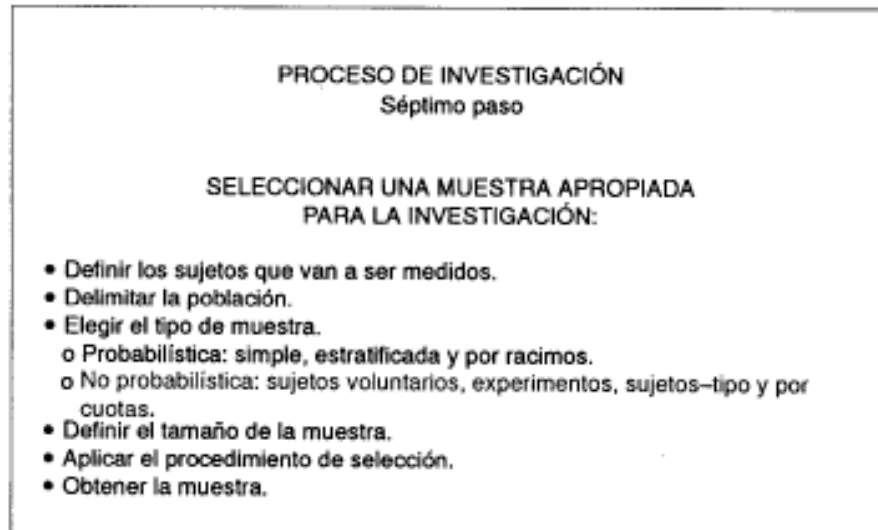


¿Cómo seleccionar una muestra?



OBJETIVOS DE APRENDIZAJE

Que el alumno:

- 1) Comprenda los conceptos de muestra, población y procedimiento de selección de la muestra.
- 2) Conozca los diferentes tipos de muestras, sus características, las situaciones en que es conveniente utilizar cada uno y sus aplicaciones.
- 3) Esté capacitado para determinar el tamaño adecuado de muestra en distintas situaciones de investigación.
- 4) Pueda obtener muestras adecuadas desde el punto de vista científico, aplicando diferentes métodos de selección.

SÍNTESIS

El capítulo discute los conceptos de muestra, población o universo, tamaño de muestra, representatividad de la muestra y procedimiento de selección. También presenta una tipología de muestras: probabilísticas y no probabilísticas. Explica cómo definir los sujetos que van a ser medidos, cómo determinar el tamaño adecuado de muestra y cómo proceder a obtener la muestra dependiendo del tipo de selección elegido.

8.1. ¿QUIÉNES VAN A SER MEDIDOS?

Aquí el interés se centra en quienes”, es decir, en los *sujetos u objetos de estudio*. Esto desde luego, depende del planteamiento inicial de la investigación. Así, si el objetivo es por ejemplo, describir el uso que hacen los niños de la televisión, lo más factible es que tendremos que interrogar a una muestra de niños. Desde luego, también sería posible entrevistar a las mamás de los niños. Escoger entre los niños o sus mamás, o ambos,

dependería no sólo del objetivo de la investigación sino del diseño de la misma. El caso —ya citado en el libro— de la investigación de Fernández Collado, Baptista y Elkes (1986) en donde el objetivo básico del estudio es el de describir la relación niño-televisión, determinó que los sujetos seleccionados para el estudio fueron niños que respondieron sobre sus conductas y percepciones relacionadas con este medio de comunicación. En otro estudio de Greenberg, Ericson y Vlahos (1972) el objetivo de análisis era investigar las discrepancias o semejanzas en las opiniones de madres e hijos con respecto al uso de la televisión. Aquí el objetivo del estudio supuso la selección de mamás y niños, para entrevistarlos cada uno por su lado, correlacionando posteriormente la respuesta de cada par madre-hijo.

Puede lo anterior ser muy obvio, pues los objetivos de los dos ejemplos mencionados son claros. En la práctica esto no parece ser tan simple para muchos estudiantes que en propuestas de investigación y de tesis no logran una coherencia entre los objetivos de la investigación y la unidad de análisis de la misma. Algunos errores comunes se encuentran en la tabla 8.1.

Para seleccionar una *muestra*, lo primero entonces es definir nuestra *unidad de análisis* —personas, organizaciones, periódicos, etc.— *El ‘quiénes van a ser medidos’, depende de precisar claramente el problema a investigar y los objetivos de la investigación.* Estas acciones nos llevarán al siguiente paso, que es el de delimitar una población.

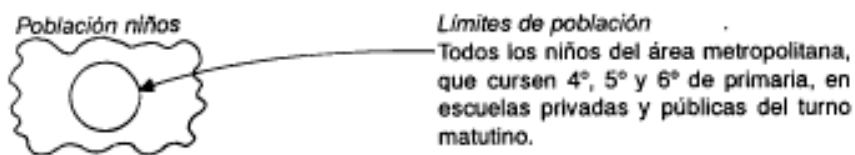
TABLA 8.1
¿QUIÉNES VAN A SER MEDIDOS?: ERRORES Y SOLUCIONES

Pregunta de investigación	Unidad de análisis errónea	Unidad de análisis correcta
¿Discriminan a las mujeres en los anuncios de la televisión?	Mujeres que aparecen en los anuncios de televisión Error no hay grupo de comparación	Mujeres y hombres que aparecen en los anuncios de televisión para comparar si categorías de análisis difieren entre los dos grupos.
¿Están los obreros del área metropolitana satisfechos con su trabajo?	Computar el número de conflictos sindicales registrados en Conciliación y Arbitraje durante los últimos 5 años, Error: la pregunta propone indagar sobre actitudes individuales y esta unidad de análisis denota datos agregados en una estadística laboral y macrosocial	Muestra de obreros que trabajan en el área metropolitana cada uno de los cuales contestará a las preguntas de un cuestionario.
¿Hay problemas de comunicación entre padres e hijos?	Grupo de adolescentes, aplicarles cuestionario. Error: se procedería a describir únicamente cómo perciben los adolescentes la relación con sus padres	Grupo de padres e hijos. A ambas partes se le aplicará el cuestionario.

8.2. ¿CÓMO SE DELIMITA UNA POBLACIÓN?

Una vez que se ha definido cuál será nuestra *unidad de análisis*, se procede a *delimitar la población* que va a ser estudiada y sobre la cual se pretende generalizar los resultados. Así, una población es el conjunto de todos los casos que concuerdan con una serie de especificaciones (Selítiz, 1974). La *muestra* suele ser definida como un *subgrupo de la población* (Sudman, 1976). Para seleccionar la muestra deben delimitarse las características de la población. Muchos investigadores no describen lo suficiente las características de la población o asumen que la muestra representa automáticamente a la población. Es frecuente que muchos estudios que únicamente se basan en muestras de estudiantes universitarios —porque ‘es fácil aplicarles el instrumento de medición, pues están a la mano’— hagan generalizaciones temerarias sobre jóvenes que probablemente posean otras características sociales. Es preferible entonces, establecer claramente las características de la población, a fin de delimitar cuáles serán los *parámetros muestrales*. Lo anterior puede ilustrarse con el ejemplo de la investigación sobre el uso de la televisión por los niños. Está claro que en dicha investigación la unidad de análisis son los niños. Pero, ¿de qué población se trata?, de ¿todos los niños del mundo?, de ¿todos los niños de la República Mexicana? Sería muy ambicioso y prácticamente imposible referirnos a poblaciones tan grandes. Así tenemos que en nuestro ejemplo la población fue delimitada de la siguiente manera:

FIGURA 8.1



Esta definición eliminó entonces a niños mexicanos que no vivieran en el área metropolitana del D.F, a los que no van al colegio y a los menores de 9 años. Pero por otra parte permitió hacer una investigación costeable, con cuestionarios contestados por niños que ya sabían escribir y un control sobre la inclusión de niños de todas las zonas de la metrópolis, al usar la ubicación de las escuelas como puntos de referencia y de selección. En este y otros casos, la delimitación de las características de la población no sólo depende de los objetivos del estudio, sino de otras razones prácticas. No será un mejor estudio, por tener una población más grande, sino la calidad de un trabajo estriba en delimitar claramente la población con base en los objetivos del estudio. *Las poblaciones deben situarse claramente en torno a sus características de contenido, lugar y en el tiempo.* Por ejemplo, en un estudio sobre los directivos de empresa en México (Baptista, 1983) y con base en las consideraciones teóricas del estudio que describe el comportamiento gerencial de los individuos y la relación de éste con otras variables de tipo organizacional se procedió a definir la población de la siguiente manera:

Nuestra población comprende a todos aquellos directores generales de empresas industriales y comerciales que en 1983 tienen un capital social superior a 30 millones de pesos, con ventas superiores a los 100 millones de pesos y/o con más de 300 personas empleadas.

Vemos que en este ejemplo se delimita claramente la población, excluyendo a personas que no son los directores generales, a empresas que no pertenezcan al giro industrial y comercial, como por ejemplo bancos, hoteles, casas de bolsa. Se establece también claramente que se trata de empresas medianas y grandes con base en criterios de capital y de recursos humanos. Finalmente se indica que estos criterios operaron en el año 1983.

⁵ Algunos investigadores usan el término universo, pero los autores preferimos utilizar el término población, ya que como Kisch (1974), consideramos que universo es más bien un término descriptivo de un Conjunto infinito de datos, lo que no se aplica a la población.

Los criterios que cada investigador cumpla dependen de sus objetivos de estudio, lo que es importante es establecerlos claramente. Toda investigación debe ser transparente, sujeta a crítica y a réplica, y este ejercicio no es posible si al examinar los resultados, el lector no puede referirlos a la población utilizada en un estudio.

8.3. ¿CÓMO SELECCIONAR LA MUESTRA?

Hasta este momento hemos visto que se tiene que *definir cuál será la unidad de análisis y cuáles son las características de la población*. En este inciso hablaremos de la *muestra* o mejor dicho de los *tipos de muestra* que existen, a fin de poder elegir la más conveniente para un estudio.



La muestra es, en esencia, un subgrupo de la población. Digamos que es un subconjunto de elementos que pertenecen a ese conjunto definido en sus características al que llamamos población. Esto se representa en la figura 8.2. Con frecuencia leemos y oímos hablar de “*muestra representativa*”, “*muestra al azar*” “*muestra aleatoria*” como si con los simples términos se pudiera dar más seriedad a los resultados. En realidad, pocas veces se puede medir a toda la población, por lo que obtenemos o seleccionamos una muestra y se pretende —desde luego— que este subconjunto sea un reflejo fiel del conjunto de la población. Todas las muestras deben ser representativas, por tanto el uso de este término es por demás inútil. Los términos *al azar* y *aleatorio* denotan un tipo de procedimiento mecánico relacionado con la probabilidad y con la selección de elementos, pero no logra esclarecer tampoco el tipo de muestra y el procedimiento de muestreo. Hablemos entonces de esto en los próximos incisos.

8.3.1. Tipos de muestra

Básicamente categorizamos a las muestras en dos grandes ramas: *las muestras no probabilísticas* y *las muestras probabilísticas*. En estas últimas todos los elementos de la población tienen la misma posibilidad de ser escogidos. Esto se obtiene definiendo las características de la población, el tamaño de la muestra y a través de una selección aleatoria y/o mecánica de las unidades de análisis. Imagínense el procedimiento para obtener el número premiado en un sorteo de lotería. Este número se va formando en el momento del sorteo, a partir de las bolitas (con un dígito) que se van sacando después de revolverlas mecánicamente hasta formar el número, de manera que todos los números tienen la misma probabilidad de ser elegidos.

En las *muestras no probabilísticas*, la elección de los elementos no depende de la probabilidad, sino de causas relacionadas con las características del investigador o del que hace la muestra. Aquí el procedimiento no es mecánico, ni en base a fórmulas de probabilidad, sino que depende del proceso de toma de decisiones de una persona o grupo de personas, y desde luego, las muestras seleccionadas por decisiones subjetivas tienden a estar sesgadas. *El elegir entre una muestra probabilística o una no probabilística, depende —sí, otra vez— de los objetivos del estudio, del esquema de investigación y de la contribución que se piensa hacer con dicho estudio*. Para ilustrar lo anterior mencionaremos varios ejemplos que toman en cuenta dichas consideraciones.

EJEMPLO 1

En un primer ejemplo tenemos una investigación sobre inmigrantes extranjeros en México, (Baptista et al. 1988). El objetivo de la investigación es documentar las experiencias de viaje, de vida y de trabajo. Para cumplir dicho propósito se seleccionó una muestra no probabilística de personas extranjeras que por diversas razones —económicas, políticas, fortuitas— hubieran llegado a México entre 1900 y 1960. Las personas se seleccionaron a través de conocidos, de asilos, de referencias. De esta manera se entrevistaron a 40 inmigrantes con entrevistas semiestructuradas que permitieron al sujeto hablar libremente sobre sus experiencias.

Comentado. En este caso una muestra no probabilística es adecuada pues se trata de un estudio con un diseño de investigación exploratorio, es decir, no es concluyente, sino su objetivo es documentar ciertas experiencias. Este tipo de estudio pretende generar datos e hipótesis que constituyan la materia prima para investigaciones más precisas.

EJEMPLO 2

Como segundo caso mencionaremos el caso de una investigación para saber cuántos niños han sido vacunados y cuántos no, y variables asociadas <nivel socioeconómico, lugar donde se vive, educación) con esta conducta y sus motivaciones. En este caso se hizo una muestra probabilística nacional de 1600 personas y de los datos se tomaron decisiones para formular estrategias de vacunación y mensajes dirigidos a persuadir la pronta y oportuna vacunación de los niños.

Comentario. Este tipo de estudio, en donde se hace una asociación entre variables, cuyos resultados servirán de información para tomar decisiones políticas que afectarán a una población, se logran por medio de una investigación por encuestas y definitivamente a través de una *muestra probabilística*, diseñada de tal manera que los datos pueden ser generalizados a la población con una estimación precisa del error que pudiera cometerse al hacer tales generalizaciones.

EJEMPLO 3

Se diseña un experimento para medir si contenidos violentos en la televisión generan conductas antisociales en los niños. Para lograr tal objetivo se seleccionan en un colegio 60 niños de 5 años de edad de Igual nivel socioeconómico e igual inteligencia y se asignan aleatoriamente a 2 grupos o condiciones. 30 niños verán caricaturas pro-sociales (ej. Heidi) y otros 30 verán caricaturas muy violentas. Inmediatamente después de la exposición a dichos contenidos violentos, los niños serán observados en un contexto de juego y se medirán sus conductas violentas y pro-sociales.

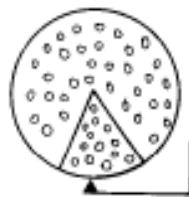
Comentario. Esta es una muestra *no probabilística*. Aunque se asignen los niños de manera aleatoria a las dos condiciones experimentales, para generalizar a la población se necesitarían repetidos experimentos. Un estudio así es valioso en cuanto a que el nivel causa-efecto es más preciso al aislar otras variables, sin embargo los datos no pueden generalizarse a todos los niños, sino a un grupo de niños con las mencionadas características. Se trata de una muestra dirigida y “clásica” de un estudio de este tipo. La selección de la muestra no es al azar, aunque la asignación de los niños a los grupos si lo es.

8.4. ¿CÓMO SE HACE UNA MUESTRA PROBABILÍSTICA?

Resumiremos diciendo que *la elección entre la muestra probabilística y una no probabilística se determina con base en los objetivos del estudio, el esquema de la investigación y el alcance de sus contribuciones*. Las muestras probabilísticas tienen muchas ventajas, quizás la principal es que puede medirse el tamaño de error en nuestras predicciones. Puede decirse incluso que el principal objetivo en el diseño de una muestra probabilística es el de reducir al mínimo este error al que se le llama error estándar (Kish, 1965).

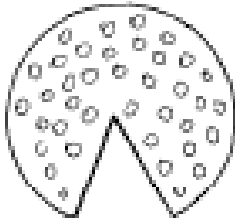
Las *muestras probabilísticas* son esenciales en los diseños de investigación por encuestas en donde se pretende hacer estimaciones de variables en la población, estas variables se miden con instrumentos de medición (capítulo 9) y se analizan con pruebas estadísticas para el análisis de datos en donde se presupone que la muestra es probabilística, donde todos los elementos de la población tienen una misma probabilidad de ser elegidos. Los elementos muestrales tendrán valores muy parecidos a los de la población, de manera que las mediciones en el subconjunto, nos darán estimados precisos del conjunto mayor. Que tan preciso son dichos estimados depende del error en el muestreo, el que se puede calcular, pues hay errores que dependen de la medición y estos errores no pueden ser calculados matemáticamente.

Para hacer una muestra probabilística es necesario entender los siguientes términos y sus definiciones:



La población, a la que llamaremos N , es un conjunto de elementos.

La muestra, a la que denominaremos n , es un subconjunto de la población N .



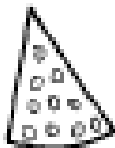
En una población N —previamente delimitada por los objetivos de la investigación— nos interesa establecer expresiones numéricas de las características de los elementos de N .

Nos interesa conocer valores promedio en la población, el cual se expresa como:

\bar{Y} = es decir se refiere al valor de una variable determinada (\bar{Y}) que nos interesa conocer.

Nos interesa conocer también:

V = es decir la varianza de la población con respecto a determinadas variables.



Como los valores de la población no se conocen, seleccionamos una *muestra* n y a través de estimados en la muestra, inferimos valores en la población. \bar{y} será el valor de \bar{Y} el cual desconocemos. \bar{y} es un estimado promedio en la muestra el cual podemos determinar. Sabemos que en nuestra estimación habrá una diferencia ($\bar{Y} - \bar{y} = ?$) es decir, habrá un error, el cual dependerá

del número de elementos muestreados. A dicho error le llamaremos estándar

=Se

Se = es la desviación *estándar de la distribución muestral* y representa la fluctuación de \bar{y} .

$(se)^2$ = el *error estándar al cuadrado*, es la fórmula que nos servirá para calcular la varianza (V) de la población (N). Y la varianza de la muestra (n) será la expresión S^2

S^2 = *varianza de la muestra*, la cual podrá determinarse en términos de probabilidad donde $S^2 = p(1-p)$

Para una *muestra probabilística* necesitamos principalmente dos cosas: determinar el *tamaño de la muestra* (n) y seleccionar los elementos muestrales, de manera que todos tengan la misma posibilidad de ser elegidos. Para lo primero, daremos una fórmula que contiene las expresiones ya descritas. Para lo segundo,

necesitamos de un marco de selección adecuado y de un procedimiento que permita la aleatoriedad en la selección. Hablaremos de ambas cosas en los siguientes incisos.

8.4.1. El tamaño de la muestra

Cuando se hace una muestra probabilística, uno debe preguntarse ¿Cuál es el número mínimo de *unidades de análisis* (personas, organizaciones, capítulos de telenovelas, etc.), que necesito para conformar una muestra (n) que me asegure un error estándar menor de .01 (fijado por nosotros), dado que la población N es aproximadamente de tantos elementos? En esta pregunta se inquiriere cuál será la probabilidad de ocurrencia de \bar{y} , y de que el valor de \bar{y} —basado en n observaciones— se sitúe en un intervalo que comprenda al verdadero valor de la población. Es decir que mi estimado \bar{y} se acerque a \bar{Y} , al valor real. Si nosotros establecemos el *error estándar* y fijamos .01, sugerimos que esta fluctuación promedio de nuestro estimado y con respecto a los valores reales de la población \bar{Y} , no sea $>.01$, es decir que de 100 casos, 99 veces mi predicción sea correcta y que el valor de \bar{y} se sitúe en un intervalo de confianza que comprenda el valor de \bar{Y} . La *fórmula para determinar el tamaño de n* es la siguiente:

$$n' = \frac{S^2 \text{ varianzadelamuestra}}{V^2 \text{ varianzadelapoblación}}$$

lo cual se ajusta si se conoce el tamaño de la población N . Entonces tendremos que:

$$n' = \frac{n'}{1 - n'/n}$$

Pongamos el siguiente ejemplo. En el ejemplo que ya habíamos dado en el inciso 8.2 de este capítulo, delimitamos a una población diciendo que para un estudio de directores generales consideramos a “todos aquellos directores generales de empresas industriales y comerciales que en 1983 tienen un capital social superior a 30 millones de pesos, con ventas superiores a los 100 millones de pesos y/o con más de 300 personas empleadas”. Con estas características se precisó que la población era de $N = 1\ 176$ directores generales ya que 1 176 empresas conformaban las mencionadas características. ¿Cuál es entonces el número de directores generales n que se tiene que entrevistar, para tener un error estándar menor de .015, y dado que la población total es del 176?

N = población de 1176 empresas.

\bar{y} = valor promedio de una variable = 1, un director general en cada empresa.

Se = error estándar - .015, lo determinamos. Es aceptable pues es muy pequeño.

V = varianza de la población. Su definición $(Se)^2$ el cuadrado del error estándar.

S^2 = varianza de la muestra expresada como la probabilidad de ocurrencia de \bar{y}

Sustituyendo tenemos que:

$$n' = \frac{S^2}{V^2}$$

$$S^2 = p(1-p) = .9(1-.9) = .09$$

$$V = (.015)^2 = .000225$$

$$n' = \frac{.09}{.000225} = 400$$

y ajustando tenemos que:

$$n' = \frac{n'}{1 + n/N} = \frac{400}{1 + 400/1176} = 298$$

Es decir que, para nuestra investigación, necesitaremos una muestra de 298 directores generales.

Esto (como habíamos dicho) es el primer procedimiento para obtener la muestra probabilística: el determinar el tamaño de la misma, con base en estimados de la población. El segundo procedimiento estriba en cómo y de dónde seleccionar a esos 298 sujetos.

8.4.2. Muestra probabilística estratificada

El pasado ejemplo corresponde a una muestra probabilística simple. Determinamos en este caso que el tamaño de la muestra sería de $n = 298$ directivos de empresa. Pero supongamos que la situación se complica y que esta que tendremos que estratificar a fin de que los elementos muestrales o unidad de análisis posean un determinado atributo. En nuestro ejemplo este atributo es el giro de la empresa. Es decir, cuando no basta que cada uno de los elementos muestrales tengan la misma probabilidad de ser escogidos, sino que además es necesario *estratificar la muestra* en relación a estratos o categorías que se presentan en la población y que aparte son relevantes para los objetivos del estudio, se diseña una muestra probabilística estratificada. Lo que aquí se hace es dividir a la población en subpoblaciones o estratos y se selecciona una muestra para cada estrato. La estratificación aumenta la precisión de la muestra e implica el uso deliberado de diferentes tamaños de muestra para cada estrato, “a fin de lograr reducir la varianza de cada unidad de la media muestral” (Kish, 1965). Dice Kish (p. 92) en su libro de muestreo que en un número determinado de elementos muestrales $n = Z n h$ la varianza de la media muestral σ^2 puede reducirse al mínimo si el tamaño de la muestra para cada estrato es proporcional a la desviación estándar dentro del estrato.

Esto es,

$$fh = \frac{n}{N} = KSh$$

En donde fh es la fracción del estrato, n el tamaño de la muestra, N el tamaño de la población, sh es la desviación estándar de cada elemento en el estrato h , y K es una proporción constante que nos dará como resultado una q óptima para cada estrato.

Siguiendo nuestro ejemplo de los directores de empresa tenemos que la población es de 1 176 directores de empresa y que el tamaño de muestra es $n = 298$. La fracción para cada estrato fh será:

$$fh = \frac{n}{N} = \frac{298}{1176} = .2534$$

De manera que el total de la subpoblación se multiplicará por esta fracción constante a fin de obtener el tamaño de muestra para el estrato. Sustituyendo tenemos que:

$$N_h \times fh = nh$$

TABLA 8.2

MUESTRA PROBABILÍSTICA ESTRATIFICADA DE DIRECTORES DE EMPRESA

Estrato por giro	Directores generales de empresa del giro	Total población* (fh) = .2534 Nh (fn) = nh	Muestra
1	Extractivo y Siderúrgico	53	13
2	Metal mecánicas	109	28
3	Alimentos, bebidas, tabaco	215	55
4	Papel y artes gráficas	87	22
5	Textiles	98	25
6	Eléctricas y electrónicas	110	28
7	Automotriz	81	20
8	Químico-farmacéutica	221	56
9	Otras empresas transformación	151	38
10	Comerciales	<u>51</u>	<u>13</u>
		N=1176	n=298

por ejemplo:

Nh = 53 directores de empresas extractivas corresponde a la población total de este giro
fh = .2534 es la fracción constante.

nh = 13 es el número redondeado de directores de empresa del giro extractivo que tendrán que entrevistarse.

*Fuente de Industridata, 1982.

8.4.3. Muestreo probabilístico por racimos

En algunos casos en donde el investigador se ve limitado por recursos financieros, por tiempo, por distancias geográficas o por una combinación de éstos y otros obstáculos, se recurre a otra modalidad de *muestreo* llamado *por racimos*. En este tipo de muestreo se reducen costos, tiempo y energía al considerar que muchas veces nuestras unidades de análisis se encuentran encapsuladas o encerradas en determinados lugares físicos o geográficos a los que denominamos racimos. Para dar algunos ejemplos tenemos la tabla 8.3., en donde en la primera columna se encuentran unidades de análisis que frecuentemente vamos a estudiar en ciencias sociales. En la segunda columna, sugerimos posibles racimos en donde se encuentran dichos elementos.

TABLA 8.3

EJEMPLOS DE RACIMOS

UNIDAD DE ANÁLISIS	POSIBLES RACIMOS
Adolescentes	Preparatorias
Obreros	Industrias
Amas de casa	Mercados
Niños	Colegios
Personajes de televisión	Programas de televisión

El muestrear por racimos implica diferenciar entre la unidad de análisis y la unidad muestral. La unidad de análisis —como lo indicamos al principio de este capítulo— se refiere a quiénes van a ser medidos, o sea, el

sujeto o sujetos a quienes en última instancia vamos a aplicar el instrumento de medición. La unidad muestral —en este tipo de muestra— se refiere al racimo a través del cual se logra el acceso a la unidad de análisis. El *muestreo por racimos* supone una selección en dos etapas, ambas con procedimientos probabilísticos. En la primera, se seleccionan los racimos, siguiendo los ya reseñados pasos de una muestra probabilística simple o estratificada. En la segunda, y dentro de estos racimos se seleccionan a los sujetos u objetos que van a ser medidos. Para ello se hace una selección que asegure que todos los elementos del racimo tienen la misma probabilidad de ser elegidos. A continuación daremos un ejemplo que comprenda varios de los procedimientos descritos hasta ahora y que ilustra la manera como frecuentemente se hace una muestra probabilística en varias etapas.

EJEMPLO

¿COMO HACER UNA MUESTRA PROBABILÍSTICA ESTRATIFICADA Y POR RACIMOS?

- *Problema de investigación:* Una estación de radio local necesita saber con precisión —a fin de planear sus estrategias— cómo usan la radio los adultos de una ciudad de 2 500 000 habitantes. Es decir, qué tanto radio escuchan, a qué horas, qué contenidos prefieren y sus opiniones con respecto a los programas noticiosos.
- *Procedimientos:* Se diseñará un cuestionario que indague estas áreas sobre uso del radio. Los cuestionarios se aplicarán por entrevistadores a una muestra de sujetos adultos.
- *Población:* Todos aquellos sujetos —hombres o mujeres— de más de 21 años de edad, y que vivan en una casa o departamento propio o rentado de la ciudad X.
- *Diseño por racimos:* Los directivos de la estación de radio desconocen el número total de sujetos con las características arriba señaladas. Sin embargo, nos piden que diseñemos una muestra que abarque a todos los sujetos adultos de la ciudad, adultos por edad cronológica y por ser jefes de familia, es decir, excluye a los adultos dependientes. Se recurre entonces a la estrategia de seleccionar racimos y se considera el uso de un mapa actualizado de la ciudad y que demuestra que en dicha ciudad hay 5 000 cuadras. Las cuadras se utilizarán como racimos, es decir como unidades muestrales a partir de las cuales obtendremos en última instancia a nuestros sujetos adultos. Lo primero entonces es determinar ¿Cuántas cuadras necesitaremos muestrear, de una población total de 5000 cuadras, si queremos que nuestro error estándar sea no mayor de 0.15 y con una probabilidad de ocurrencia del 50%?

Tenemos entonces que $n' = \frac{S^2}{V^2}$ para una muestra probabilística simple.

$$S^2 = p(1-p) = .5 = .25$$

$$V^2 = (\text{error estándar})^2 = (.015)^2 = .00025$$

$$n' = \frac{S^2}{V^2} = \frac{.25}{.00025} = 1111.11$$

$$n' = \frac{n}{1 + n/N} = \frac{1111.11}{1 + 1111.11/5000} = 909.0902 = 909$$

Necesitaremos una muestra de 909 cuadras de ciudad X para estimar los valores de la población con una probabilidad de error menor a .01.

* Sabemos que la población $N = 5\ 000$ cuadras de la ciudad está dividida por previos estudios de acuerdo a 4 *estratos socioeconómicos*, que categorizan las 5 000 cuadras según el ingreso mensual promedio de sus habitantes, de manera que se distribuyen como sigue:

Estrato	No. de cuadras
1	270
2	1940
3	2000
4	790
	T = 5 000

* Estratificación de la muestra:

$$fh = \frac{n}{N} = KSh$$

$$fh = \frac{909}{5000} = .1818$$

¿Cómo distribuiremos los 909 elementos muestrales de η_1 para optimizar nuestra muestra, de acuerdo a la distribución de la población en los 4 estratos socioeconómicos?

Estrato	No. de cuadras	fh = .1 818	η_h
1	270	(.1818)	50
2	1 940	(.1818)	353
3	2000	(.1818)	363
4	790	(.1818)	143
	N=5000		n=909

Tenemos que en principio, de 5 000 cuadras de la ciudad se seleccionarán 50 del estrato 1, 353 del estrato 2, 363 del estrato 3 y 143 del estrato 4. Esta selección comprende la selección de los racimos, los cuales se pueden numerar y elegir aleatoriamente hasta completar el número de cada estrato <ver sección 8.4.2). En una última etapa se seleccionan a los sujetos dentro de cada racimo. Este procedimiento también se hace de manera aleatoria, hasta lograr un número de sujetos determinados en cada racimo. En el próximo inciso describiremos dicho procedimiento.

Estrato	Nh cuadras	nh	Número de hogares-sujeto en cada cuadra	Total de hogares por estrato
1	270	50	20	1000
2	1940	353	20	7060
3	2000	363	20	7220
4	790	143	20	2860
	N=5000	n=909		11840

8.5. ¿CÓMO SE LLEVA A CABO EL PROCEDIMIENTO DE SELECCIÓN?

Cuando iniciamos nuestra discusión sobre *muestra probabilística*, señalamos que dichos tipos de muestra dependen de dos cosas:

1) del *tamaño de la muestra*; 2) del *procedimiento de selección*. De lo primero, hemos hablado con todo

detalle, de lo segundo hablaremos ahora. Se determina el *tamaño de la muestra* n , pero ¿cómo seleccionar los elementos muestrales? Se precisa el número de racimos necesario ¿cómo se seleccionan a los sujetos dentro de cada racimo? Hasta el momento sólo hemos dicho que los elementos se eligen aleatoriamente, pero ¿cómo se hace esto?

Las unidades de análisis o los elementos muestrales se eligen siempre aleatoriamente para asegurarnos que cada elemento tenga la misma probabilidad de ser elegidos. Pueden usarse 3 procedimientos de selección:

8.5.1. Tómbola

Muy simple y no muy rápido, consiste en numerar todos los elementos muestrales del 1.. al n . Hacer unas fichas, una por cada elemento, revolverías en una caja, e ir sacando n fichas, según el tamaño de la muestra. Los números elegidos —al azar— conformarán la muestra.

Así en la tabla 8.2., tenemos que de una población $N = 53$ empresas extractivas y siderúrgicas, se necesita una muestra $n = 13$ de directivos generales de dichas empresas. En una lista se puede numerar cada una de estas empresas. En fichas aparte

se sortean cada uno de los 53 números. Los números obtenidos se checan con los nombres y direcciones de nuestra lista, para precisar los que serán sujetos de análisis.

8.5.2. Números random o números aleatorios

El uso de *números random* no significa la selección azarosa o fortuita, sino la utilización de una tabla de números que implica un mecanismo de probabilidad muy bien diseñado. Los números random de la Corporación Rand, fueron generados con una especie de ruleta electrónica. Existe una tabla de un millón de dígitos, publicada por esta corporación; partes de dicha tabla se encuentran en los apéndices de muchos libros de estadística. Estas tablas son como lo muestra la tabla 8.4 y el apéndice 5.

Siguiendo el ejemplo del inciso anterior, determinamos una muestra de 909 manzanas o cuadras, y a partir de este número se determinó una submuestra para cada estrato. Véase que para el estrato 1, la población es de 270, manzanas. Numeramos entonces en nuestro listado o mapa las 270 cuadras y seleccionamos —a partir de la tabla de números random— los 50 casos que constituirán nuestra muestra.

TABLA 8.4
NÚMEROS RANDOM

26804	29273	79811	45610	22879	72538	70157	17683	67942	52846
90720	96215	48537	94756	18124	89051	27999	88513	35943	67290
85027	59207	76180	41416	48521	15720	90258	95598	10822	93074
09362	49674	65953	96702	20772	12069	49901	08913	12510	64899
64590	04104	16770	79237	82158	04553	93000	18585	72279	01916
06432	08525	66864	20507	92817	39800	98820	18120	81860	68065
02101	60119	95836	88949	89312	82716	34705	12795	58424	69700
19337	96983	60321	62194	08574	81896	00390	75024	66220	16494
75277	47880	07952	35832	41655	27155	95189	00400	06649	53040
59535	75885	31648	88202	63899	40911	78138	26376	06641	97291
76310	79385	84639	27804	48889	80070	64889	99310	04232	84008
12805	65754	96887	67060	88413	31883	79233	99603	68989	80233
32242	73807	48321	67123	40637	14102	55550	89992	80593	64642
16212	84706	69274	13252	78974	10781	43629	36223	36042	75492

75362	83633	25620	24828	59345	40653	85639	42613	40242	43160
34703	93445	82051	53437	53717	48719	71858	11230	26079	44018
01556	58563	36828	85053	39025	16688	69524	81885	31911	13098
22211	86468	76295	16663	39489	18400	53155	92087	63942	99827
01534	70128	14111	77065	99358	28443	68135	61696	55241	61867
09647	32348	56909	40951	00440	10305	58160	62235	89455	73095
97021	23763	18491	65056	95283	98232	86695	78699	79666	88574
25469	63708	78718	35014	40387	15921	58080	03936	15953	59658
40337	48522	11418	00090	41779	54499	08623	49092	654.31	11390
33491	98685	92536	51626	85787	47641	95787	70139	42383	44187
44764	14986	16642	19429	01960	22833	80055	39851	47350	70337

Fuente: Rand Corporation.

Se eligen aquellos casos que se dictaminen en la *tabla de números random*, hasta completar el tamaño de la muestra. Los números pueden recorrerse hacia arriba, hacia abajo, horizontalmente. Al fin siempre se logra que cada elemento muestral tenga la misma probabilidad de ser elegido. Se eligen aquellos números que contenga el listado. Así si en nuestro ejemplo la población es de 270, se escogen los 3 últimos dígitos y se procede de la siguiente manera a seleccionar los casos hasta completar el número de elementos muestrales.



TABLA 8.5
SELECCIÓN MUESTRA BASADOS EN LA TABLA DE NÚMEROS
RANDOM

78986	45691	28281	82933	24786	55586
83 830	59 025	40379	99 989	63 822	99 974
(1)30 226	19863	(5)95039	08909	(7)48 197	(8)23 270
(2)02 073	(4)59 042	26440	(6)16 161	14496	24786
(3)05 250	47 552	95659	92 356	13 334	23471

8.5.3. Selección sistemática de elementos muestrales

Este procedimiento de selección es muy útil y fácil de aplicar e implica el seleccionar dentro de una población N a un número n de elementos a partir de un intervalo K.

K es un intervalo que va a estar determinado por el tamaño de la población y el tamaño de la muestra. De manera que tenemos que $K = N/n$, en donde K = es un intervalo de selección sistemática N = es la población a = es la muestra

Ilustramos los anteriores conceptos con un ejemplo. Supongamos que se quiere hacer un estudio sobre varios aspectos de la publicidad en México. Específicamente se pretende medir qué número de mensajes informativos y qué número de mensajes motivacionales tienen los comerciales en la televisión mexicana. Para tal efecto supongamos que los investigadores consiguen videocasetes con todos los comerciales que han pasado al aire —en los diferentes canales de televisión— durante un periodo de tres años. Quitando los comerciales repetidos, se tiene una población de $N = 1\,548$ comerciales. Se procede con este dato a determinar qué número de comerciales necesitamos analizar para generalizar a toda la población nuestros resultados con un error estándar no mayor de .015. Con la fórmula que ya hemos dado en la lección 8.4.3 de este capítulo tenemos que si $p = .5$ $s^2 = p(1-p) = .5(.5) = .25$

$$n = \frac{S^2}{V^2} = \frac{.25}{.00025} = 1111.11, \quad n = \frac{1111.11}{1 + 1111.11/1548} = 647$$

Si necesitamos una muestra de $n = 647$ comerciales, podemos utilizar para la selección al intervalo K en donde:

$$K = \frac{N}{n} = \frac{1548}{647} = 2.39 = 3^* \text{ (redondeando)}$$

El intervalo $1/K = 3$ indica que cada tercer comercial $1/K$ será seleccionado hasta completar $n = 647$.

La *selección sistemática de elementos muestrales* $1/K$ se puede utilizar para elegir los elementos de n para cada estrato y/o para cada racimo. La regla de probabilidad que dice que cada elemento de la población tiene que tener la misma probabilidad de ser elegido, se mantiene empezando la selección de $1/K$ al azar. Siguiendo nuestro ejemplo, *no* empezamos a elegir de los 1 548 comerciales grabados, el 1,3,6,9... sino que procuramos que el comienzo sea determinado por el azar. Así, en este caso, podemos tirar unos dados y si en sus caras muestran 1, 6, 9, empezaremos en el comercial 169 y seguiremos: 169, 172, 175 $1/K$ volver a empezar por los primeros si es necesario. Este procedimiento de selección es poco complicado y tiene varias ventajas: cualquier tipo de estratos en una población X, se verán reflejados en la muestra. Asimismo, la selección sistemática logra una muestra proporcionada, pues por ejemplo tenemos que el procedimiento de selección $1/K$ nos dará una muestra con nombres que inician con las letras del abecedario en forma proporcional a la letra inicial de los nombres de la población.

8.6. LOS LISTADOS Y OTROS MARCOS MUESTRALES

Como se ha visto a lo largo de este capítulo, las *muestras probabilísticas* requieren de la determinación del tamaño de la muestra y de un *proceso de selección aleatoria que asegure que todos los elementos de la población tengan la misma probabilidad de ser elegidos*. Todo esto lo hemos visto, sin embargo nos falta discutir sobre algo esencial que precede a la selección de una muestra: *el listado, el marco muestra*. El listado se refiere a una lista existente o a una lista que se tiene que confeccionar ‘ad hoc’, de los elementos de la población, y a partir de la cual se seleccionarán los elementos muestrales. El segundo término se refiere a un marco de referencia que nos permita identificar físicamente a los elementos de la población, la posibilidad de enumerarlos y por ende, proceder a la selección de los elementos muestrales.

* 2.39 se redondea para que sea un entero. Véase Kish (1969) p. 115-117.

Los listados basados en listas existentes sobre una población pueden ser variados: el directorio telefónico, la lista de miembros de una asociación, directorios especializados, las listas oficiales de escuelas de la zona, las listas de las canciones de éxito publicadas por una revista, la lista de alumnos de una universidad, etc. En todo caso hay que tener en cuenta lo completo de una determinada lista, su exactitud, veracidad, su calidad, y qué tanta cobertura tiene en relación con el problema a investigar y la población que va a medirse, ya que todos estos aspectos influyen en la selección de la muestra. Por ejemplo, para algunas encuestas se considera que el directorio telefónico es muy útil. Sin embargo hay que tomar en cuenta que muchos teléfonos no aparecerán porque son privados o que hay hogares que no tienen teléfono. La lista de socios de una asociación como Canacintra (Cámara Nacional de la Industria de la Transformación) puede servir si el propósito del estudio es —por ejemplo— conocer la opinión de los asociados con respecto a una medida gubernamental. Más si el objetivo de la investigación es el análisis de opinión del sector patronal del país, el listado de una asociación no será adecuado por varias razones, entre otras: hay otras asociaciones patronales, la Canacintra representa solamente el sector de la Industria de Transformación, las asociaciones son voluntarias y no todo patrón o empresa pertenece a éstas. Lo correcto en este caso, sería construir una nueva lista, con base en los listados existentes de las asociaciones patronales, eliminando de dicha lista los casos duplicados, suponiendo que una o más empresas pudieran pertenecer a dos asociaciones al mismo tiempo, como director a la COPARMEX (Confederación Patronal de la República Mexicana) y como empresa a la ANIQ (Asociación Nacional de Ingenieros Químicos).

Hay listas que proporcionan una gran ayuda al investigador. Pensamos en directorios especializados como el Industriadata que enlista a las empresas mexicanas medianas y grandes, el directorio de la Ciudad de México por calles, el directorio de medios, que enlista casa productoras, estaciones de radio y televisión, periódicos y revistas. Este tipo de directorios realizados por profesionales son útiles al investigador pues representan una compilación (sujetos, empresas, instituciones), resultado de horas de trabajo e inversión de recursos. Recomendamos pues utilizarlos cuando sea pertinente, tomando en cuenta las consideraciones que estos directorios hacen en su introducción y que revelan a qué año pertenecen los datos, cómo los obtuvieron, (exhaustivamente, por cuestionarios, por voluntarios) y *muy importante*, quiénes y por qué quedan excluidos del directorio.

En México se cuenta también con directorios de anunciantes en publicidad y mercadotecnia como el publicado por Mercamétrica Ediciones, 5. A. Frecuentemente es necesario construir listas ad hoc, a partir de las cuales se seleccionarán los elementos que constituirán las unidades de análisis en una determinada investigación. Por ejemplo en la investigación de La Televisión y el Niño (Fernández Collado, et. al., 1986) se hizo una muestra probabilística estratificada por racimo, en donde en una primera etapa se relacionaron escuelas para en última instancia llegar a los niños. Pues bien, para tal efecto se consiguió una lista de las escuelas primarias del Distrito Federal. Cada escuela tenía un código identificable por medio del cual se eliminaron, las escuelas para niños atípicos. Este listado contenía además información sobre cada escuela, sobre su ubicación —calle y colonia—, sobre su propiedad —pública o privada—.

Con ayuda de otro estudio que catalogaba en diferentes estratos socioeconómicos a las colonias del Distrito Federal con base al ingreso promedio de la zona, se hicieron 8 listas:

- 1 escuelas públicas clase A
- 2 escuelas privadas clase A
- 3 escuelas públicas clase B
- 4 escuelas privadas clase B
- 5 escuelas públicas clase C
- 6 escuelas privadas clase C
- 7 escuelas públicas clase D
- 8 escuelas privadas clase D

Cada lista representaba un estrato de la población y de cada una de ellas se seleccionó una muestra de

escuelas.

No siempre existen listas que permitan identificar a nuestra población. Será necesario pues recurrir a otros marcos de referencia que contengan descripciones del material, organizaciones o sujetos que serán seleccionados como unidades de análisis.

Algunos de estos marcos de referencia son los archivos, los mapas, volúmenes de periódicos empastados en una biblioteca o las horas de transmisión de varios canales de televisión. De cada una de estas instancias daremos ejemplos con más detalles.

8.6.1. Archivos

Un jefe de reclutamiento y selección de una institución quiere precisar si algunos datos que se dan en una solicitud de trabajo están correlacionados con el ausentismo del empleado. Es decir, si a partir de datos como edad, sexo, estado civil y duración en otro trabajo, puede predecirse que alguien tenderá a ser faltista. Para establecer correlaciones se considerarán como población a todos los sujetos contratados durante 10 años. Se relacionan sus datos en la solicitud de empleo con los registros de faltas.

Como no hay una lista elaborada de estos sujetos, el investigador decide acudir a los archivos de las solicitudes de empleo. Estos archivos constituyen su marco muestral a partir del cual obtendrá la muestra. Calcula el tamaño de la población, obtiene el tamaño de la muestra y selecciona sistemáticamente cada elemento $1/K$ (ver sección 8.5.3) cada solicitud que será analizada. Aquí el problema que surge es que en el archivo hay solicitudes de gente que no fue contratada, y por tanto, no pueden ser consideradas en el estudio. En este caso y en otros en donde no todos los elementos del marco de referencia o de una lista (por ejemplo nombres en el directorio que no corresponden a una persona física) los especialistas en muestreo (Kish, 1965; Sutman, 1976) *no* aconsejan el reemplazo, con el siguiente elemento, sino simplemente no tomar en cuenta ese elemento, es decir como si no existiera, continuándose con el intervalo de selección sistemática.

8.6.2. Mapas

Los mapas son muy útiles como marco de referencia en muestras por racimo. Por ejemplo, un investigador quiere saber qué motiva a los compradores en una determinada tienda de autoservicio. Sobre un mapa de la ciudad y a partir de la lista de tiendas que de cada cadena competidora, marca todas las tiendas de autoservicios, las cuales constituyen una población de racimos, pues en cada tienda seleccionada, entrevistará a un número n de clientes. El mapa le permite ver la población (tiendas autoservicio) y su situación geográfica, de manera que eligió zonas donde coexistan tiendas de la competencia, como para asegurarse que el consumidor de la zona tenga todas las posibles alternativas.

8.6.3. Volúmenes

En este ejemplo supongamos que un estudioso del periodismo quiere hacer un análisis de contenido de los editoriales de los tres principales diarios de la ciudad durante el porfiriato. El investigador va a la Hemeroteca Nacional y encuentra en los volúmenes que encuadernan a los diarios por trimestre y año un marco de referencia ideal a partir del cual se seleccionará n volúmenes para su análisis. Supongamos, sin embargo, con que se encuentra que el volumen X que contiene el periódico "el Hijo del Ahuizote" (Enero-Marzo 1899), falta en la Hemeroteca. ¿Qué hace? Pues redefine la población, manifestando explícitamente que de N volúmenes tiene 99% de los elementos y a partir de este nuevo número de N calculó su muestra n y la seleccionó.

8.6.4. Horas de transmisión

En un estudio de Portilla y Solórzano (1982), los investigadores querían hacer un análisis de anuncios en la TV mexicana. Las emisoras no proporcionan una lista de anuncios ni sus horas de transmisión. Por otra parte sería muy caro grabar todos los anuncios a todas horas e imposible estar frente al televisor para hacerlo. Ante

la imposibilidad de tener un listado de comerciales, se hicieron listados que identificaron cada media hora de transmisión televisiva en cada canal 2, 4, 5, 9, y 13 de las 7 a las 24 horas durante siete días de una semana de octubre de 1982. La población estaba constituida del número total de medias horas de transmisión televisiva, $N = 1190$ horas. Esta población se dividió en estratos —mañana, mediodía, tarde y noche— y se procedió a calcular el tamaño de la muestra tomando en consideración que por cada media hora de transmisión hay 6 minutos de comerciales (De Noriega, 1979).

Se calculó el número de medias horas que se seleccionarían para obtener una muestra n . Una vez obtenido el tamaño de la muestra, se seleccionaron aleatoriamente n medias horas y por último se grabaron y analizaron únicamente aquellos comerciales contenidos en las medias horas seleccionadas al azar y que representaron diferentes canales y segmentos del día. El punto en este ejemplo es la construcción concreta de un marco muestral que permitiera el análisis de una muestra probabilística de comerciales.

8.7. TAMAÑO ÓPTIMO DE UNA MUESTRA Y EL TEOREMA DEL LÍMITE CENTRAL

Las muestras probabilísticas, como lo hemos visto en incisos anteriores, requieren dos procedimientos básicos: 1) la determinación del tamaño de la muestra y 2) la selección aleatoria de los elementos muestrales. El primer procedimiento, lo hemos descrito en su modalidad más simple, en la sección 8.4.1 de este capítulo. El precisar adecuadamente el tamaño de la muestra puede tornarse en algo muy complejo dependiendo del problema de investigación y la población a estudiar. Se nos ocurre que para el alumno y el lector en general, pueda resultar muy útil el comparar qué tamaño de muestra han utilizado otros investigadores en ciencias sociales. Para tal efecto reproducimos las siguientes tablas preparadas por Sudman (1976) y que indican el tamaño de la muestra más utilizada por los investigadores según sus poblaciones (nacionales o regionales) y según los subgrupos que quieren estudiarse en ellas.

TABLA 8.6
MUESTRAS FRECUENTEMENTE UTILIZADAS EN
INVESTIGACIONES NACIONALES Y REGIONALES SEGÚN ÁREA
DE ESTUDIO

<i>Tipo de estudio</i>	<i>Nacionales</i>	<i>Regionales</i>
Económicos	1000+	100
Médicos	1000+	500
Conductas	1000+	700 — 300
Actitudes	1000 +	700 — 400
Experimentos de Laboratorio	— — —	100

En esta tabla vemos que el tipo de estudio poco determina el tamaño de la muestra, sino más bien el hecho de que sean muestras nacionales o regionales. Las muestras nacionales, es decir, muestras que representan a la población de un país son típicamente de más de 1 000 sujetos. La muestra del estudio “¿cómo somos los mexicanos?” (Hernández Medina, Harro, et. al., 1987) consta de 1837 sujetos repartidos de la siguiente manera:

Frontera y norte	696
Centro (sin D.F.)	426
Sur-sureste	316
Distrito Federal	<u>299</u>
	1 837

Las muestras regionales (por ejemplo las que representen al área metropolitana) algún estado del país o algún municipio o región son típicamente más pequeñas con rangos de 700 a 400 sujetos.

El tamaño de una muestra tiende más a depender del número de subgrupos que nos interesan en una población. Por ejemplo, podemos subdividirla aún más en hombres y mujeres de 4 grupos de edad; o aún más en hombres y mujeres de 4 grupos de edad en cada uno de 5 niveles socioeconómicos. Si este fuera el caso estaríamos hablando de 40 subgrupos y por ende de una muestra mayor. En la siguiente tabla se describen típicas muestras según los subgrupos bajo estudio, según su alcance,

—estudios nacionales o estudios especiales o regionales— y según su unidad de análisis, es decir se trata de sujetos o de organizaciones, en esta última instancia el número de la muestra se reduce, ya que éstas representan casi siempre una gran fracción de la población total.

Tabla 8.7
MUESTRAS TÍPICAS DE ESTUDIOS SOBRE POBLACIONES HUMANAS Y ORGANIZACIONALES

Número de subgrupos	Población de sujetos u hogares		Poblaciones de organizaciones	
	Nacionales	Regionales	Nacionales	Regionales
Ninguno-pocos	1000-1500	200-500	200-500	50-200
Promedio	1500-2500	500-1000	500-1000	200-500
Muchos	2500-+	1000+	1000+	500+

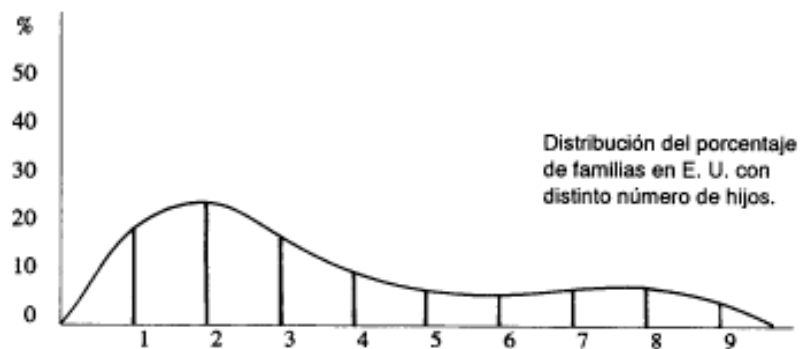
Estas tablas (Sudman 1976: 86-87) fueron construidas en base a artículos de investigación publicados en revistas especializadas y nos dan una idea de las muestras que utilizan otros investigadores, de manera que pueden ayudar al investigador a precisar el tamaño de su muestra. Recordemos que lo óptimo de una muestra depende en qué tanto su distribución se aproxima a la distribución de las características de la población. Esta aproximación mejora al incrementarse el tamaño de la muestra. La “normalidad” de la distribución en muestras grandes, no obedece a la normalidad de la distribución de una población. Al contrario, la distribución de las variables en estudio de ciencias sociales están lejos de ser normales. Sin embargo, la distribución de muestras de 100 o más elementos tienden a ser normales y esto sirve para el propósito de hacer estadística inferencial sobre los valores de una población. A esto se le llama teorema de límite central.

Distribución normal: esta distribución en forma de campana se logra generalmente con muestras de 100 o + unidades muestrales y es útil y necesaria cuando se hacen inferencias de tipo estadístico.

FIGURA. 8.3



FIGURA. 8.4



Esta es la distribución de una población; es anormal y, sin embargo, la distribución de una muestra de esta población de más de 100 casos tenderá a distribuirse normalmente. Esta tendencia —teorema del límite central— permite estimar los valores de la población, a partir de la inferencia estadística.

8.8. ¿CÓMO SON LAS MUESTRAS NO PROBABILÍSTICAS?

Las *muestras no probabilísticas*, las cuales llamamos también *muestras dirigidas* suponen un procedimiento de selección informal y un poco arbitrario. Aún así estas se utilizan en muchas investigaciones y a partir de ellas se hacen inferencias sobre la población. Es como si juzgásemos el sabor de un cargamento de limones, solamente probando alguno, como si para “muestra bastase un botón”. La muestra dirigida selecciona sujetos “típicos” con la vaga esperanza de que serán casos representativos de una población determinada. La verdad es que las muestras dirigidas tienen muchas desventajas. La primera es que, al no ser probabilísticas, no podemos calcular con precisión el error estándar, es decir, no podemos calcular con qué nivel de confianza hacemos una estimación. Esto es un grave inconveniente si consideramos que la estadística inferencial se basa en teoría de la probabilidad, por lo que pruebas estadísticas (X^2 , correlación, regresión, etc.), en muestras no probabilísticas tienen un valor limitado y relativo a la muestra en sí, mas no a la población. Es decir, los datos no pueden generalizarse a una población, que no se consideró ni en sus parámetros, ni en sus elementos para obtener la muestra. Recordemos que, *en las muestras de este tipo, la elección de los sujetos no depende de que todos tienen la misma probabilidad de ser elegidos, sino de la decisión de un investigador o grupo de encuestadores.*

La ventaja de una *muestra no probabilística* es su utilidad para un determinado diseño de estudio, que requiere no tanto de una “representatividad de elementos de una población, sino de una cuidadosa y controlada elección de sujetos con ciertas características especificadas previamente en el planteamiento del problema”. Hay varias clases de muestras dirigidas y éstas se definirán a continuación.

8.8.1. La muestra de sujetos voluntarios

Las muestras de sujetos voluntarios son frecuentes en ciencias sociales y ciencias de la conducta. Se trata de muestras fortuitas, utilizadas también en la Medicina y la Arqueología en donde el investigador elabora conclusiones sobre especímenes que llegan a sus manos de manera casual. Pensemos por ejemplo en los sujetos que voluntariamente acceden a participar en un estudio que monitorea los efectos de un medicamento o en el investigador que anuncia en una clase que está haciendo un estudio sobre motivación en el universitario e invita a aquellos que acepten someterse a una prueba proyectiva TA.T. En estos casos la elección de los individuos que serán sujetos a análisis depende de circunstancias fortuitas. Este tipo de muestra se usa en estudios de laboratorio donde se procura que los sujetos sean homogéneos en variables tales como edad, sexo, inteligencia, de manera que los resultados o efectos no obedezcan a diferencias individuales, sino a las condiciones a las que fueron sometidos.

8.8.2. La muestra de expertos

En ciertos estudios es necesaria la opinión de sujetos expertos en un tema. Estas muestras son frecuentes en estudios cualitativos y exploratorios que para generar hipótesis más precisas o para generar materia prima para diseño de cuestionarios. Por ejemplo en un estudio sobre el perfil de la mujer periodista en México (Barrera, et. al., 1989) se recurrió a una muestra de $n = 227$ mujeres periodistas pues se consideró que estos eran los sujetos idóneos para hablar de contratación, sueldos y desempeño de las mujeres periodistas. Estas son muestras válidas y útiles cuando los objetivos del estudio así lo requieren.

8.8.3. Los sujetos-tipos

Al igual que las muestras anteriores, ésta también se utiliza en estudios exploratorios y en investigaciones de tipo cualitativo, donde el objetivo es la riqueza, profundidad y calidad de la información, y no la cantidad, y estandarización. En estudios de perspectiva fenomenológica donde el objetivo es analizar los valores, ritos y significados de un determinado grupo social, el uso tanto de expertos como de sujetos-tipo es frecuente. Por ejemplo pensamos en los trabajos de Howard Becker (“El músico de jazz”, “Los muchachos de blanco”) en donde se basa en grupos de típicos músicos de jazz y típicos estudiantes de medicina para adentrarse en el análisis de los patrones de identificación y socialización de estas dos profesiones: la de músico, la de médico.

Los estudios motivacionales, los cuales se hacen para el análisis de las actitudes y conductas del consumidor, también utilizan muestras de sujeto-tipo. Aquí se definen los grupos a los que va dirigido un determinado producto —por ejemplo jóvenes clase socioeconómica A y B, amas de casa, clase B, ejecutivos clase A-B— y se construyen grupos de 8 ó 10 personas, cuyos integrantes tengan las características sociales y demográficas de dicho subgrupo.

Con dicho grupo se efectúa una sesión, en que un facilitador o moderador dirigirá una conversación donde los miembros del grupo expresen sus actitudes, valores, medios, expectativas, motivaciones hacia las características de un determinado producto o servicio.

8.8.4. La muestra por cuotas

Este tipo de muestra se utiliza mucho en estudios de opinión y de mercadotecnia. Los encuestadores reciben instrucciones de administrar cuestionarios a sujetos en la calle, y que al hacer esto vayan conformando o llenando cuotas de acuerdo a la proporción de ciertas variables demográficas en la población. Así, por ejemplo, para un estudio sobre la actitud de la población hacia un candidato político, le dice a los encuestadores “van a tal colonia y me entrevistan a 150 sujetos. Que el 25% sean hombres mayores de 30 años, 25% mujeres mayores de 30 años; 25% hombres menores de 25 años y 25% mujeres menores de 25 años”. Así se construyen estas muestras, que como vemos dependen en cierta medida del juicio del entrevistador.

Hemos terminado este capítulo de muestra y, a manera de conclusión, resumiremos en una tabla que esquematice los diferentes tipos de muestra, y los estudios en donde se usan con mayor frecuencia.

TABLA 8.8

TIPOS DE MUESTRA

Muestras probabilísticas

Muestras dirigidas

(Estudios descriptivos, diseños de investigación por encuestas, censos, raitings, estudios para toma de decisiones).

Muestra probabilística simple

Sujetos voluntarios (diseños experimentales, situación de laboratorio).

Muestra probabilística estratificada.

Muestras de experimentos,

Muestra probabilística estratificada y por racimos.	Muestras de sujetos-tipo estudios cualitativos, investigación motivacional.
	Muestras por cuotas. Estudios de opinión y de mercado.
Resultados. Las conclusiones se generalizan a la población, y se conoce el error estándar de nuestros estimados.	Las conclusiones difícilmente pueden generalizarse a la población. Si esto se hace debe ser con mucha cautela.

RESUMEN

1. En este capítulo describimos el cómo seleccionar una muestra. Lo primero que se tiene que plantear es el quiénes van a ser medidos, lo que corresponde a definir la unidad de análisis —personas, organizaciones o periódicos—. Se procede después a delimitar claramente la población con base en los objetivos del estudio y en cuanto a características de contenido, de lugar y en el tiempo.
2. La muestra es un subgrupo de la población —previamente delimitada— y puede ser probabilística o no probabilística.
3. El elegir qué tipo de muestra se requiere depende de los objetivos del estudio y del esquema de investigación.
4. Las muestras probabilísticas son esenciales en los diseños de investigación por encuestas donde se pretenden generalizar los resultados a una población. La característica de este tipo de muestra, es que todos los elementos de la población tienen al inicio la misma probabilidad de ser elegidos, de esta manera los elementos muestrales tendrán valores muy aproximados a los valores de la población, ya que las mediciones del subconjunto, serán estimaciones muy precisas del conjunto mayor. Esta precisión depende del error de muestreo, llamado también error estándar.
5. Para una muestra probabilística necesitamos dos cosas: determinar el tamaño de la muestra y seleccionar los elementos muestrales en forma aleatoria.
6. El tamaño de la muestra se calcula con base a la varianza de la población y la varianza de la muestra. Esta última expresada en términos de probabilidad de ocurrencia. La varianza de la población se calcula con el cuadrado del error estándar, el cual determinamos. Entre menor sea el error estándar, mayor será el tamaño de la muestra.
7. Las muestras probabilísticas pueden ser: Simples, estratificadas y por racimos. La estratificación aumenta la precisión de la muestra e implica el uso deliberado de submuestras para cada estrato o categoría que sea relevante en la población. El muestrear por racimos implica diferencias entre la unidad de análisis y la unidad muestral. En este tipo de muestreo hay una selección en dos etapas, ambas con procedimientos probabilísticos. En la primera se seleccionan los racimos —escuelas,

- organizaciones, salones de clase— en la segunda y dentro de los racimos a los sujetos que van a ser medidos.
8. Los elementos muestrales de una muestra probabilística *siempre* se eligen aleatoriamente para asegurarnos de que cada elemento tenga la misma probabilidad de ser elegido. Pueden usarse tres procedimientos de selección: 1. Tómbola, 2. Tabla de números random y 3. Selección sistemática. Todo procedimiento de selección depende de listados, ya sea existentes o contruidos ad hoc. Listados pueden ser: el directorio telefónico, listas de asociaciones, listas de escuelas oficiales, etc. Cuando no existen listas de elementos de la población se recurren a otros marcos de referencia que contengan descripciones del material, organizaciones o sujetos seleccionados como unidades de análisis. Algunos de éstos pueden ser los archivos, hemerotecas y los mapas.
 9. Las muestras no-probabilísticas, pueden también llamarse muestras dirigidas, pues la elección de sujetos u objetos de estudio depende del criterio del investigador.
 10. Las muestras dirigidas pueden ser de varias clases: (1) Muestra de sujetos voluntarios —frecuentemente utilizados con diseños experimentales y situaciones de laboratorio. (2) Muestra de expertos —frecuentemente— utilizados en estudios exploratorios. (3) Muestra de sujetos tipo —o estudios de casos—, utilizados en estudios cualitativos y motivacionales y (4) muestreo por cuotas —frecuentes— en estudios de opinión y de mercadotecnia. Las muestras dirigidas son válidas en cuanto a que un determinado diseño de investigación así los requiere, sin embargo los resultados son generalizables a la muestra en sí o a muestras similares. No son generalizables a una población.
 11. En el teorema de límite central se señala que una muestra de más de cien casos, será una muestra con una distribución normal en sus características, sin embargo la normalidad no debe conjuntarse con probabilidad. Mientras lo primero es necesario para efectuar pruebas estadísticas, lo segundo es requisito indispensable para hacer inferencias correctas sobre una población.

GLOSARIO

Elementos muestrales; Casos o unidades que conforman una muestra.

Error estándar: Error en el muestreo, definido como la desviación promedio de un estimado de los valores reales de la población.

Listados: Lista o marco de referencia del cual se obtienen los elementos muestrales.

Muestra: Subconjunto de elementos de la población.

Muestra probabilística: Subconjunto donde todos los elementos de la población tienen la misma probabilidad de ser escogidos.

Muestra no probabilística: Muestra dirigida, en donde la selección de elementos dependen del criterio del investigador.

Población: Conjunto de todos los casos que concuerdan con una serie de especificaciones.

Selección aleatoria: Selección probabilística de los elementos de una población. Selección sistemática: Selección de elementos de una población a partir de un intervalo.

Teorema Límite Central: Proposición de que aun en muestras de tamaño moderado —más de 100 casos-, la distribución será aproximadamente normal.

Unidad de análisis: Quienes van a ser medidos en una investigación.

Unidad muestral: El racimo a través del cual se logra el acceso a la unidad de análisis.

Varianza: Fluctuación o variabilidad promedio de un determinado valor de la población.

EJERCICIOS

1. Se forman grupos de 3 o 4 personas. Cada grupo dispone de 15 minutos, para formular una pregunta de investigación. El problema puede ser de cualquier área de estudio. Lo que conviene aquí, es que sea algo que realmente inquiete a los estudiantes, algo que ellos consideren un fenómeno social importante. Las preguntas de investigación se van anotando en el pizarrón. Después y junto a cada de éstas preguntas se define ¿quiénes van a ser medidos? Discutir por qué y por qué no son correctas las respuestas de los estudiantes.

2. Como secuencia del ejercicio anterior se proponen los siguientes temas de investigación. Supongamos que en otro curso, estudiantes de un taller de investigación sugirieron los siguientes temas para investigar. Decir en cada caso quiénes van a ser medidos, para lograr resultados en las investigaciones propuestas.
- Tema 1. ¿Cuál es el impacto que sobre los jóvenes tienen los anuncios de bebidas alcohólicas?
- Tema 2. Hace tres meses que se implantó en una fábrica de motores un programa de círculos de calidad. ¿Ha tenido éxito dicho programa?
- Tema 3. Los niños que asistieron en la primaria a escuelas laicas y mixtas, ¿tienen un mejor desempeño académico en la universidad que los que provienen de escuelas religiosas de un solo sexo?
- Tema 4. ¿Qué diferencias significativas existen entre los comerciales de la televisión mexicana, la norteamericana y la venezolana?
3. Seleccione 2 estudios de alguna publicación científica <ver apéndice 1> y 2 tesis de licenciatura (que curse o que ya cursó). Obviamente los 4 estudios tienen que entrar dentro de la categoría de estudios exploratorios, descriptivos y/o experimentales: Analice los siguientes aspectos: a) ¿Cuál es el problema de investigación? b) ¿Cuál es la muestra? c) ¿Cómo fue elegida? d) ¿Es adecuada la muestra y el procedimiento de muestreo para el problema que se investigó? e) ¿Cuáles son los principales resultados o conclusiones? f) ¿Dichos resultados son generalizables a una población mayor? g) Con base en la muestra, ¿pueden tomarse como serias dichas generalizaciones? Evalúe la solidez de los 4 estudios, tomando como criterios los aspectos a, b, c, d, e, f y g.
4. Supongamos que trabaja en un despacho que hace investigaciones sociales y que diversos clientes le preguntan que los asesore en estudios de diferente índole. ¿Qué tipo de muestra sugeriría para cada uno? Fundamente su sugerencia

Cliente Necesidad Tipo de muestra

4.1.	Clínica de terapias psicoemocionales.	Pacientes con cáncer que siguen la terapia reaccionan mejor a los tratamientos médicos usuales que los enfermos de cáncer que no toman la terapia
4.2.	Empresa en el giro químico.	Definir cuáles son nuestros empleados y obreros, anteriores y presentes, que tienen menos ausentismo Es decir, ¿hay un perfil del ausentista?
4.3.	Empresa de cosmología.	¿Qué nociones tienen las jóvenes (de 15 a 20 años) sobre su arreglo personal y cuidado de su cutis. Funcionaria crear una línea de productos exclusivamente para ellas?
4.4.	Grupo que defiende los derechos del consumidor.	¿Qué quejas tienen los niños sobre los juguetes del mercado?, ¿se rompen?, ¿son peligrosos?, ¿aburridos?, durabilidad, etcétera.
4.5.	Partido político.	¿Por cuál candidato a gobernador votarán los ciudadanos de determinado Estado?

5. Supongamos que una asociación iberoamericana de profesionales cuenta con 5 000 miembros. La junta directiva ha decidido hacer una encuesta (por teléfono o por fax) a los suscritos para indagar - entre otras cosas— lugar de trabajo, puesto que ocupa, salario aproximado, carrera cursada, generación, estudios posteriores, oportunidades de avance percibidas, etc. En resumen, se piensa publicar un perfil profesional actualizado con propósito de retroalimentar a los asociados. Como sería muy costoso llegar a los 5 000 miembros repartidos en España, Iberoamérica y Estados Unidos, ¿qué tamaño de muestra se necesita, si queremos un error estándar no mayor de .015?

Una vez definido el tamaño de la muestra, ¿cómo sería el proceso de selección, de manera que los resultados obtenidos con base en la muestra puedan ser generalizados a toda la población? Es decir, se pretende reportar un perfil certero de los 5 000 socios de dicha asociación profesional.

6. Una institución quiere lanzar por televisión mensajes de prevención de uso de sustancias dañinas (alcohol y drogas). Los productores no saben realmente el grado de realismo que deben contener estos mensajes ni su tono, es decir si deben apelar al miedo, a la salud o a los problemas morales que se desencadenan en las familias. Se sabe con certeza que hay que hacer esta campaña, pero no se tiene idea clara de cómo estructurar el mensaje para que sea más efectivo. En resumen, para conceptualizar y poner en imágenes dichos mensajes, se necesita información previa sobre la relación sujeto-sustancia. ¿Qué se aconsejaría aquí? ¿Qué tipo de muestra se necesitaría para recabar dicha información?

BIBLIOGRAFÍA SUGERIDA

- COSMOS, S. (1982). Lite style and consumption patterns. *Journal of Consumer Research*. March, p. 453.
- DOUGLAS, JACK D. (1980). *Introduction to the sociology of everyday life*. New York, N.Y: Allyn and Bacon.
- GLASS, GENE V. y Julian C. Stanley (1970). *Statistical methods in education and psychology*. New Jersey: Prentice-Hall.
- HANUSHEK, ERIC A. y JE. Jackson (1977). *Statistical methods for social scientists*. New York, N.Y.: Academic Press~ capítulo 3.
- HARRISON, DANIEL R (1976). *Social forecasting methodology: Suggestions for research*. New York, N.Y: Ruselí Sage Foundation.
- HOLGUÍN QUIÑONES, FERNANDO (1972). *Estadística descriptiva (aplicada a las ciencias sociales)*. México, D.F.: Universidad Nacional Autónoma de México.
- KISH, LESLIE (1975). *Survey Sampling*. New York, N.Y: John Wilen & Sons.
- KREYSZIQ, ERWIN (1974). *Introducción a la estadística matemática*. México, D.F.: Ed. LIMUSA. Parte II.
- LOFLAND, JOHN & L.H. LOFLAND (1984). *Analyzing social setting*. Belmont, CA: Wadsworth Publishing Co., Inc. University of California.
- PADUA, JORGE (1979). *Técnicas de investigación aplicadas a las ciencias sociales*. México, D.F.: El Colegio de México/Fondo de Cultura Económica. Capítulo III.



Recolección de los datos

PROCESO DE INVESTIGACIÓN

Octavo paso

RECOLECTAR LOS DATOS

- Definir la forma idónea de recolectar los datos de acuerdo al contexto de la investigación.
- Elaborar el instrumento de medición.
- Aplicar el instrumento de medición.
- Obtener los datos.
- Codificar los datos.
- Archivar los datos y prepararlos para el análisis.

OBJETIVOS

Que el alumno:

- 1) Comprenda el significado de “medir” en ciencias sociales.
- 2) Comprenda los requisitos que toda medición debe cumplir: confiabilidad y validez.
- 3) Conozca los métodos para determinar la confiabilidad y validez de un instrumento de medición.
- 4) Comprenda los niveles de medición en que pueden ubicarse las variables.
- 5) Conozca los principales instrumentos de medición disponibles en ciencias sociales.
- 6) Esté capacitado para elaborar y aplicar diferentes instrumentos de medición.
- 7) Se encuentre habilitado en la preparación de datos para su análisis.

SÍNTESIS

El capítulo presenta una definición de medición en el contexto de las ciencias sociales, así como los requisitos que todo instrumento de medición debe reunir: confiabilidad y validez. Diversos métodos para determinar la confiabilidad y validez son revisados.

Además, el capítulo analiza y ejemplifica las principales maneras de medir en ciencias sociales: escalas de actitudes, cuestionarios, análisis de contenido, observación, pruebas estandarizadas, sesiones en profundidad y utilización de archivos.

Finalmente en el capítulo se presenta el procedimiento de codificación de los datos obtenidos y la forma de

prepararlos para el análisis.

9.1. ¿QUÉ IMPLICA LA ETAPA DE RECOLECCIÓN DE LOS DATOS?

Una vez que seleccionamos el diseño de investigación apropiado y la muestra adecuada de acuerdo con nuestro problema de estudio e hipótesis, *la siguiente etapa* consiste en *recolectar los datos* pertinentes sobre las variables involucradas en la investigación.

Recolectar los datos implica tres actividades estrechamente vinculadas entre sí:

- a) *Seleccionar un instrumento de medición* de los disponibles en el estudio del comportamiento o desarrollar uno (el instrumento de recolección de los datos). Este instrumento debe ser válido y confiable, -de lo contrario no podemos basarnos en sus resultados.
- b) *Aplicar ese instrumento de medición*. Es decir, obtener las observaciones y mediciones de las variables que son de interés para nuestro estudio (medir variables).
- c) *Preparar las mediciones obtenidas* para que puedan analizarse correctamente (a esta actividad se le denomina *codificación de los datos*).

9.2. ¿QUÉ SIGNIFICA MEDIR?

De acuerdo con la definición clásica del término —ampliamente difundida— *medir* significa “asignar números a objetos y eventos de acuerdo a reglas” (Stevens, 1951). Sin embargo, como señalan Carmines y Zeller (1979), esta definición es más apropiada para las ciencias físicas que para las ciencias sociales, ya que varios de los fenómenos que son medidos en éstas no pueden caracterizarse como objetos o eventos, puesto que son demasiado abstractos para ello. La disonancia cognitiva, la alienación, el producto nacional bruto y la credibilidad son conceptos tan abstractos para ser considerados cosas que pueden verse o tocarse” (definición de objeto) o solamente como “resultado, consecuencia o producto” (definición de evento) (Carmines y Zeller, 1979, p. 10).

Este razonamiento nos hace sugerir que es más adecuado definir la *medición* como “*el proceso de vincular conceptos abstractos con indicadores empíricos* proceso que se realiza mediante un plan explícito y organizado para clasificar (y frecuentemente cuantificar) los datos disponibles —los indicadores— en términos del concepto que el investigador tiene en mente (Carmines y Zeller, 1979, p. 10). Y en este proceso, el *instrumento de medición o de recolección de los datos* juega un papel central. Sin él no hay observaciones clasificadas.

La definición sugerida incluye dos consideraciones: La primera es desde el punto de vista empírico y se resume en que *el centro de atención es la respuesta observable* (sea una alternativa de respuesta marcada en un cuestionario, una conducta grabada vía observación o una respuesta dada a un entrevistador). La segunda es desde una perspectiva teórica y se refiere a que *el interés se sitúa en el concepto subyacente no observable* que es representado por la respuesta (Carmines y Zeller, 1979). Así, los registros del instrumento de medición representan valores observables de conceptos abstractos. *Un instrumento de medición adecuado es aquel que registra datos observables que representan verdaderamente a los conceptos o variables que el investigador tiene en mente.*

En toda investigación aplicamos un instrumento para medir las variables contenidas en las hipótesis (y cuando no hay hipótesis, simplemente para medir las variables de interés). Esa medición es efectiva cuando el instrumento de recolección de los datos realmente representa a las variables que tenemos en mente. Si no es así nuestra medición es deficiente y por lo tanto la investigación no es digna de tomarse en cuenta. Desde luego, no hay medición perfecta, es prácticamente imposible que representemos fielmente variables tales como la inteligencia, la motivación, el nivel socioeconómico, el liderazgo democrático, la actitud hacia el sexo y otras más; pero sí debemos de acercarnos lo más posible a la representación fiel de las variables a

observar, mediante el instrumento de medición que desarrollemos.

9.3. ¿QUÉ REQUISITOS DEBE CUBRIR UN INSTRUMENTO DE MEDICIÓN?

Toda medición o instrumento de recolección de los datos debe reunir dos requisitos esenciales: *confiabilidad* y *validez*. La *confiabilidad* de un instrumento de medición se refiere al grado en que su aplicación repetida al mismo sujeto u objeto, produce iguales resultados. Por ejemplo, si yo midiera en este momento la temperatura ambiental mediante un termómetro y me indicara que hay 22°C. Un minuto más tarde consultara otra vez y el termómetro me indicara que hay 5°C. Tres minutos después observara el termómetro y ahora me indicara que hay 40°C. Este termómetro no sería confiable (su aplicación repetida produce resultados distintos). Igualmente, si una prueba de inteligencia la aplico hoy a un grupo de personas y me proporciona ciertos valores de inteligencia; la aplico un mes después y me proporciona valores diferentes, al igual que en subsecuentes mediciones. Esa prueba no es confiable (analícese los valores de la figura 9.1, suponiendo que los coeficientes de inteligencia puedan oscilar entre 95 y 150). Los resultados no son consistentes; no se puede “confiar” en ellos.

FIGURA 9.1
EJEMPLO DE RESULTADOS PROPORCIONADOS POR
UN INSTRUMENTO DE MEDICIÓN SIN CONFIABILIDAD

PRIMERA APLICACIÓN		SEGUNDA APLICACIÓN		TERCERA APLICACIÓN	
Martha	130	Laura	131	Luis	140
Laura	125	Luis	130	Teresa	129
Arturo	118	Marco	127	Martha	124
Luis	112	Arturo	120	Rosa María	120
Marco	110	Chester	118	Laura	109
Rosa Maria	110	Teresa	118	Chester	108
Chester	108	Martha	115	Arturo	103
Teresa	107	Rosa María	107	Marco	101

La *confiabilidad* de un instrumento de medición se determina mediante diversas técnicas, las cuales se comentarán brevemente después de revisar el concepto de *validez*.

La *validez*, en términos generales, se refiere al grado en que un instrumento realmente mide la variable que pretende medir. Por ejemplo, un instrumento para medir la inteligencia válido debe medir la inteligencia y no la memoria. Una prueba sobre conocimientos de Historia debe medir esto y no conocimientos de literatura histórica. Aparentemente es sencillo lograr la *validez*. Después de todo —como dijo un estudiante— “pensamos en la variable y vemos cómo hacer preguntas sobre esa variable”. Esto sería factible en unos cuantos casos (como lo sería el “sexo” de una persona). Sin embargo, la situación no es tan simple cuando se trata de variables como la motivación, la calidad de servicio a los clientes, la actitud hacia un candidato político y menos aun con sentimientos y emociones, así como diversas variables con las que trabajamos en ciencias sociales. La *validez* es una cuestión más compleja que debe alcanzarse en todo instrumento de medición que se aplica. Kerlinger (1979, p. 138) plantea la siguiente pregunta respecto a la *validez*: ¿Está usted midiendo lo que usted cree que está midiendo? Si es así, su medida es válida; si no, no lo es.

La *validez* es un concepto del cual pueden tenerse diferentes tipos de evidencia (Wiersma, 1986; Gronlund, 1985): 1) *evidencia relacionada con el contenido*, 2) *evidencia relacionada con el criterio* y 3) *evidencia relacionada con el constructo*. Hablemos de cada una de ellas.

1) Evidencia relacionada con el contenido

La *validez de contenido* se refiere al grado en que un instrumento refleja un dominio específico de contenido de lo que se mide. Es el grado en que la medición representa al concepto medido (Bohrnstedt, 1976). Por ejemplo, una prueba de operaciones aritméticas no tendrá validez de contenido si incluye sólo problemas de resta y excluye problemas de suma, multiplicación o división (Carmines y Zeller, 1979). O bien, una prueba de conocimientos sobre las canciones de “Los Beatles” no deberá basarse solamente en sus álbumes Tet it Be”y “Abbey Road”, sino que debe incluir canciones de todos sus discos.

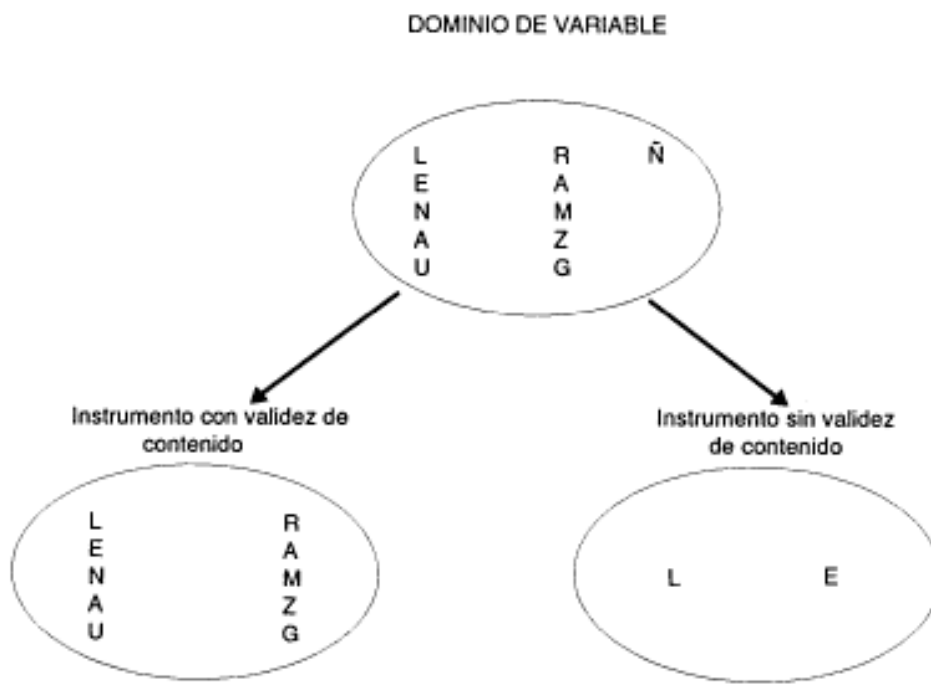
Un instrumento de medición debe *contener* representados a todos los items del dominio de contenido de las variables a medir. Este hecho se ilustra en la figura 9.2.

2) Evidencia relacionada con el criterio

La *validez de criterio* establece la validez de un instrumento de medición comparándola con algún criterio externo. Este criterio es un estándar con el que se juzga la validez del instrumento (Wiersma, 1986). Entre los resultados del instrumento de medición se relacionen más al criterio, la validez del criterio será mayor. Por ejemplo, un investigador valida un examen sobre manejo de aviones, mostrando la exactitud con que el examen predice qué tan bien Un grupo de pilotos puede operar un aeroplano.

FIGURA 9.2

ILUSTRACIÓN DE UN INSTRUMENTO DE MEDICIÓN CON VALIDEZ DE CONTENIDO VERSUS CON UNO QUE CARECE DE ÉSTA.



Si el criterio se fija en el *presente*, se habla de *validez concurrente* (los resultados del instrumento se correlacionan con el criterio en el mismo momento o punto del tiempo). Por ejemplo, un cuestionario para detectar las preferencias del electorado por los distintos partidos contendientes, puede validarse aplicándolo tres o cuatro días antes de la elección y sus resultados compararlos con los resultados finales de la elección (si no hay fraude —desde luego—).

Si el criterio se fija en el *futuro*, se habla de *validez predictiva*. Por ejemplo, una prueba para determinar la capacidad administrativa de altos ejecutivos se puede validar comparando sus resultados con el futuro

desempeño de los ejecutivos medidos.

3) Evidencia relacionada con el constructo

La *validez de constructo* es probablemente ³⁵ la más importante sobre todo desde una perspectiva científica y se refiere al grado en que una medición se relaciona consistentemente con otras mediciones de acuerdo con hipótesis derivadas teóricamente y que conciernen a los conceptos (o constructos) que están siendo medidos. Un *constructo* es una variable medida y que tiene lugar dentro de una teoría o esquema teórico.

Por ejemplo, supongamos que un investigador desea evaluar la *validez de constructo* de una medición particular, digamos una escala de motivación intrínseca: “el Cuestionario de Reacción a Tareas”, versión mexicana (Hernández-Sampieri y Cortés, 1982). Estos autores sostienen que el nivel de motivación intrínseca hacia una tarea está relacionado positivamente con el grado de persistencia adicional en el desarrollo de la tarea (v.g., los empleados con mayor motivación intrínseca son los que suelen quedarse más tiempo adicional una vez que concluye su jornada). Consecuentemente, la predicción teórica es que a mayor motivación intrínseca, mayor persistencia adicional en la tarea. El investigador administra dicho cuestionario de motivación intrínseca a un grupo de trabajadores y también determina su persistencia adicional en el trabajo. Ambas mediciones son correlacionadas. Si la correlación es positiva y sustancial, se aporta evidencia para la validez de constructo del Cuestionario de Reacción a Tareas, versión mexicana (a la validez para medir la motivación intrínseca).

La *validez de constructo* incluye tres etapas:

- 1) *Se establece y especifica la relación teórica entre los conceptos* (sobre la base del marco teórico).
- 2) *Se correlacionan ambos conceptos y se analiza cuidadosamente la correlación.*
- 3) *Se interpreta la evidencia empírica de acuerdo a qué tanto clarifica la validez de constructo de una medición en particular.*

El proceso de *validación de un constructo* está vinculado con la teoría. No es posible llevar a cabo la validación de constructo, a menos que exista un marco teórico que soporte a la variable en relación con otras variables. Desde luego, no es necesaria una teoría sumamente desarrollada, pero si investigaciones que hayan demostrado que los conceptos están relacionados. Entre más elaborado y comprobado se encuentre el marco teórico que apoya la hipótesis, la validación de constructo puede arrojar mayor luz sobre la validez de un instrumento de medición. Y mayor confianza tenemos en la validez de constructo de una medición, cuando sus resultados se correlacionan significativamente con un mayor número de mediciones de variables que teóricamente y de acuerdo con estudios antecedentes están relacionadas. Esto se representa en la figura 9.3.

Para analizar las posibles interpretaciones de evidencia negativa en la *validez de constructo*, se sugiere consultar a Cronbach y Meehí (1955) y Cronbach (1984).

VALIDEZ TOTAL = VALIDEZ DE CONTENIDO + VALIDEZ DE CRITERIO + VALIDEZ DE CONSTRUCTO

Así, la validez de un instrumento de medición se evalúa sobre la base de tres tipos de evidencia. Entre mayor evidencia de validez de contenido, validez de criterio y validez de constructo tenga un instrumento de medición; éste se acerca más a representar la variable o variables que pretende medir.

Cabe agregar que *un instrumento de medición puede ser confiable pero no necesariamente válido* (un aparato —por ejemplo— puede ser consistente en los resultados que produce, pero no medir lo que pretende). Por ello es requisito que el instrumento de medición demuestre ser *confiable* y *válido*. De no ser así, los resultados de la investigación no los podemos tomar en serio.

³⁵ La explicación se basa en Carmines y Zeller (1979).

FIGURA 9.3

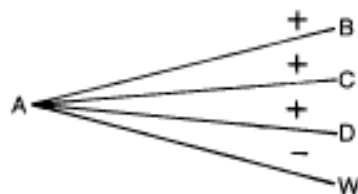
PRESENTACIÓN GRÁFICA DE UN INSTRUMENTO CON VALIDEZ DE CONSTRUCTO

Instrumento mide constructo "A"

Teoría

(Investigaciones hechas) encontraron que "A" se relaciona positivamente con "B", "C" y "D". Y negativamente con "W".

Si el instrumento mide realmente "A" sus resultados deben relacionarse positivamente con los resultados obtenidos en las mediciones de "B", "C" y "D"; y relacionarse negativamente con los resultados de "W".



El instrumento parece realmente medir "A".

FACTORES QUE PUEDEN AFECTAR LA CONFIABILIDAD Y VALIDEZ

Hay diversos factores que pueden afectar la confiabilidad y la validez de los instrumentos de medición.

El primero de ellos es la improvisación. Algunas personas creen que elegir un instrumento de medición o desarrollar uno es algo que puede tomarse a la ligera. Incluso algunos profesores piden a los alumnos que construyan instrumentos de medición de un día para otro, o lo que es casi lo mismo, de una semana a otra. Lo cual habla del poco o nulo conocimiento del proceso de elaboración de instrumentos de medición. Esta improvisación genera —casi siempre— instrumentos poco válidos o confiables y no debe existir en la investigación social (menos aún en ambientes académicos). Aun a los investigadores experimentados les toma tiempo desarrollar un instrumento de medición. Es por ello que los construyen con cuidado y frecuentemente están desarrollándolos, para que cuando los necesiten con premura se encuentren preparados para aplicarlos, pero no los improvisan. Además, para poder construir un instrumento de medición se requiere conocer muy bien a la variable que se pretende medir y la teoría que la sustenta. Por ejemplo, generar —o simplemente seleccionar— un instrumento que mida la inteligencia, la personalidad o los usos y gratificaciones de la televisión para el niño, requiere amplios conocimientos en la materia, estar actualizados al respecto y revisar cuidadosamente la literatura correspondiente.

El segundo factor es que a veces se utilizan instrumentos desarrollados en el extranjero que no han sido validados a nuestro contexto: cultura y tiempo. Traducir un instrumento —aun cuando adaptemos los términos a nuestro lenguaje y los contextualicemos— no es de ninguna manera (ni remotamente) validarlo. Es un primer y necesario paso, pero sólo es el principio. Por otra parte, hay instrumentos que fueron validados en nuestro contexto pero hace mucho tiempo. Hay instrumentos que hasta el lenguaje nos suena "arcaico". Las culturas, los grupos y las personas cambian; y esto debemos tomarlo en cuenta al elegir o desarrollar un instrumento de medición.

Un tercer factor es que en ocasiones el instrumento resulta inadecuado para las personas a las que se les aplica: no es empático. Utilizar un lenguaje muy elevado para el respondiente, no tomar en cuenta diferencias en cuanto a sexo, edad, conocimientos, capacidad de respuesta, memoria, nivel ocupacional y educativo, motivación para responder y otras diferencias en los respondientes; son errores que pueden afectar la validez y confiabilidad del instrumento de medición.

Un cuarto factor que puede influir esté constituido por las condiciones en las que se aplica el instrumento de medición. Si hay ruido, hace mucho frío (por ejemplo en una encuesta de casa en casa), el instrumento es demasiado largo o tedioso, son cuestiones que pueden afectar negativamente la validez y la confiabilidad. Normalmente en los experimentos se puede contar con instrumentos de medición más largos y complejos que en los diseños no experimentales. Por ejemplo, en una encuesta pública sería muy difícil poder aplicar una prueba larga o compleja.

Por otra parte, *aspectos mecánicos* tales como que si el instrumento es escrito, no se lean bien las instrucciones, falten páginas, no haya espacio adecuado para contestar, no se comprendan las instrucciones, también pueden influir de manera negativa.

9.4. ¿CÓMO SE SABE SI UN INSTRUMENTO DE MEDICIÓN ES CONFIABLE Y VALIDO?

En la práctica es casi imposible que una medición sea perfecta. *Generalmente se tiene un grado de error.* Desde luego, se trata de que este error sea el mínimo posible. Es por esto que la medición de cualquier fenómeno se conceptualiza con la siguiente fórmula básica:

$$X = t + e$$

Donde “X” *representa los valores observados* (resultados disponibles), “t” *son los valores verdaderos* y “e” *es el grado de error en la medición.* Si no hay error de medición (“e” es igual a cero), el valor observado y el verdadero son equivalentes. Esto puede verse claramente así:

$$X = t + 0$$

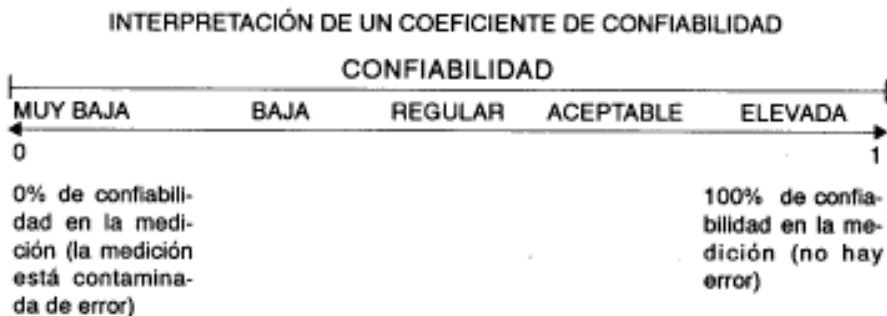
$$X = t$$

Esta situación representa el ideal de la medición. Entre mayor sea el error al medir, el valor que observamos (y que es en el que nos basamos) se aleja más del valor real o verdadero. Por ejemplo, si medimos la motivación de un individuo y esta medición está contaminada por un grado de error considerable, la motivación registrada por el instrumento será bastante diferente de la motivación real que tiene ese individuo. Por ello *es importante que el error sea reducido lo más posible.* Pero, ¿cómo sabemos el grado de error que tenemos en una medición? Calculando la confiabilidad y validez.

CÁLCULO DE LA CONFIABILIDAD

Existen diversos *procedimientos para calcular la confiabilidad* de un instrumento de medición. Todos utilizan fórmulas que producen *coeficientes de confiabilidad.* Estos coeficientes *pueden oscilar entre 0 y 1.* Donde un coeficiente de 0 significa nula confiabilidad y 1 representa un máximo de confiabilidad (confiabilidad total). Entre mas se acerque el coeficiente a cero (0), hay mayor error en la medición. Esto se ilustra en la figura 9.4.

FIGURA 9.4



Los procedimientos más utilizados para determinar la confiabilidad mediante un coeficiente son:

1. *Medida de estabilidad* (confiabilidad por test-retest). En este procedimiento un mismo instrumento de medición (o ítems o indicadores)³⁶ es aplicado dos o más veces a un mismo grupo de personas, después de un periodo de tiempo. Si la correlación entre los resultados de las diferentes aplicaciones es altamente positiva, el instrumento se considera confiable. Se trata de una especie de diseño panel. Desde luego, el periodo de tiempo entre las mediciones es un factor a considerar. Si el periodo es largo y la variable susceptible de cambios, ello puede confundir la interpretación del coeficiente de confiabilidad obtenido por este procedimiento. Y si el periodo es corto las personas pueden recordar cómo contestaron en la primera aplicación del instrumento, para aparecer como más consistentes de lo que son en realidad (Bohrnstedt, 1976).
2. *Método de formas alternativas o paralelas*. En este procedimiento no se administra el mismo instrumento de medición, sino dos o más versiones equivalentes de éste. Las versiones son similares en contenido, instrucciones, duración y otras características. Las versiones —generalmente dos— son administradas a un mismo grupo de personas dentro de un periodo de tiempo relativamente corto. El instrumento es confiable si la correlación entre los resultados de ambas administraciones es significativamente positiva. Los patrones de respuesta deben variar poco entre las aplicaciones.
3. *Método de mitades partidas* (split-halves). Los procedimientos anteriores (medida de estabilidad y método de formas alternas), requieren cuando menos dos administraciones de la medición en el mismo grupo de individuos. En cambio, el método de mitades-partidas requiere sólo una aplicación de la medición. Específicamente, el conjunto total de ítems (o componentes) es dividido en dos mitades y las puntuaciones o resultados de ambas son comparados. Si el instrumento es confiable, las puntuaciones de ambas mitades deben estar fuertemente correlacionadas. Un individuo con baja puntuación en una mitad, tenderá a tener también una baja puntuación en la otra mitad. El procedimiento se diagrama en la figura 9.5.

La confiabilidad varía de acuerdo al número de ítems que incluya el instrumento de medición. Cuantos más ítems la confiabilidad aumenta (desde luego, que se refieran a la misma variable). Esto resulta lógico, veámoslo con un ejemplo cotidiano: Si se desea probar qué tan confiable o consistente es la lealtad de un amigo hacia nuestra persona, cuantas más pruebas le pongamos, su confiabilidad será mayor. Claro está que demasiados ítems provocarán cansancio en el respondiente.

³⁶ Un ítem es la unidad mínima que compone a una medición; es un reactivo que estimula una respuesta en un sujeto (por ejemplo, una pregunta, una frase, una lámina, fotografía, un objeto de descripción).

FIGURA 9.5

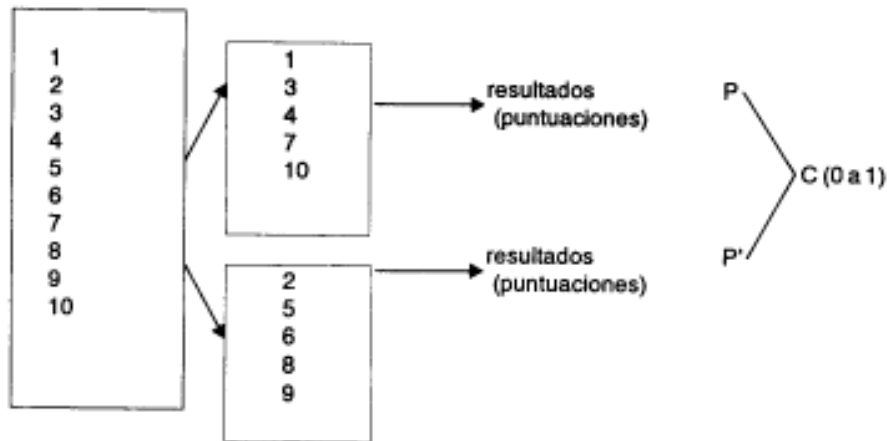
ESQUEMA DEL PROCEDIMIENTO DE MITADES-PARTIDAS

El instrumento de medición se aplica a un grupo

Los ítems se dividen en dos mitades (el instrumento se divide en dos)

Cada mitad se califica independientemente

Se correlacionan puntuaciones y se determina la confiabilidad



Al dividir los ítems, éstos se emparejan en contenido y dificultad.

4. *Coefficiente alfa de Cronbach.* Este coeficiente desarrollado por J. L. Cronbach requiere una sola administración del instrumento de medición y produce valores que oscilan entre 0 y 1. Su ventaja reside en que no es necesario dividir en dos mitades a los ítems del instrumento de medición, simplemente se aplica la medición y se calcula el coeficiente.

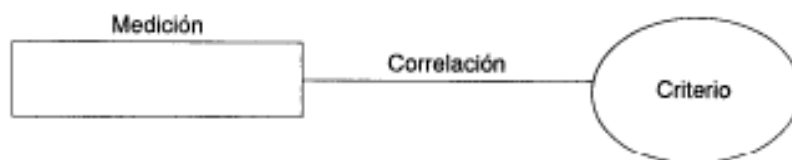
La manera de calcular este coeficiente se muestra en el siguiente capítulo.

5. *Coefficiente KR-20.* Kuder y Richardson (1937) desarrollaron un coeficiente para estimar la confiabilidad de una medición, su interpretación es la misma que la del coeficiente alfa.

CÁLCULO DE LA VALIDEZ

La validez de contenido es compleja de obtener. Primero, es necesario revisar cómo ha sido utilizada la variable por otros investigadores. Y en base a dicha revisión elaborar un universo de ítems posibles para medir la variable y sus dimensiones (el universo tiene que ser lo más exhaustivo que sea factible). Posteriormente, se consulta con investigadores familiarizados con la variable para ver si el universo es exhaustivo. Se seleccionan los ítems bajo una cuidadosa evaluación. Y si la variable tiene diversas dimensiones o facetas que la componen, se extrae una muestra probabilística de ítems (ya sea al azar o estratificada —cada dimensión constituiría un estrato—). Se administran los ítems, se correlacionan las puntuaciones de los ítems entre sí (debe haber correlaciones altas, especialmente entre ítems que miden una misma dimensión) (Bohrnstedt, 1976), y se hacen estimaciones estadísticas para ver si la muestra es representativa. Para calcular la validez de contenido son necesarios varios coeficientes.

La validez de criterio es más sencilla de estimar, lo único que hace el investigador es correlacionar su medición con el criterio, y este coeficiente es el que se toma como coeficiente de validez (Bohrnstedt, 1976). Esto podría representarse así:³⁷



La validez de constructo se suele determinar mediante un procedimiento denominado “Análisis de Factores”. Su aplicación requiere de sólidos conocimientos estadísticos y del uso de un programa estadístico apropiado en computadora. Para quien desee compenetrarse con esta técnica recomendamos consultar a Harman (1967), Gorsuch (1974), Nie et al. (1975), On-Kim y Mueller (1978a y 1978b) y Hunter (1980). Asimismo, para aplicarlos se sugiere revisar a Nieetal. (1975), Cooper y Curtis (197~) y —en español— Padua (1979). Aunque es requisito conocer el programa estadístico para computadora. Esta técnica se describe en la página 420.

9.5. ¿QUÉ PROCEDIMIENTO SE SIGUE PARA CONSTRUIR UN INSTRUMENTO DE MEDICIÓN?

Existen diversos tipos de instrumentos de medición, cada uno con características diferentes. Sin embargo, el procedimiento general para construirlos es semejante. Antes de comentar este procedimiento, es necesario aclarar que en una investigación hay dos opciones respecto al instrumento de medición:

- 1) Elegir un instrumento ya desarrollado y disponible, el cual se adapta a los requerimientos del estudio

en particular.

2) Construir un nuevo instrumento de medición de acuerdo con la técnica apropiada para ello.

En ambos casos es importante tener evidencia sobre la confiabilidad y validez del instrumento de medición.

El procedimiento que sugerimos para construir un instrumento de medición es el siguiente, especialmente para quien se inicia en esta materia.

PASOS

- a) **LISTAR LAS VARIABLES** que se pretende medir u observar.
- b) **REVISAR SU DEFINICIÓN CONCEPTUAL Y COMPRENDER SU SIGNIFICADO.** Por ejemplo, comprender bien qué es la motivación intrínseca y qué dimensiones la integran.
- c) **REVISAR CÓMO HAN SIDO DEFINIDAS OPERACIONALMENTE LAS VARIABLES,** esto es, cómo se ha medido cada variable. Ello implica comparar los distintos instrumentos o maneras utilizadas para medir las variables (comparar su confiabilidad, validez, sujetos a los cuales se les aplicó, facilidad de administración, veces que las mediciones han resultado exitosas y posibilidad de uso en el contexto de la investigación).
- d) **ELEGIR EL INSTRUMENTO O LOS INSTRUMENTOS (YA DESARROLLADOS) QUE HAYAN SIDO FAVORECIDOS POR LA COMPARACIÓN Y ADAPTARLOS AL CONTEXTO DE LA INVESTIGACIÓN.**

En este caso sólo deben seleccionarse instrumentos cuya confiabilidad y validez se reporte. No se puede uno fiar de una manera de medir que carezca de evidencia clara y precisa de confiabilidad y validez. Cualquier investigación sería reportar la confiabilidad y validez de su instrumento de medición. Recuérdese que la primera varía de 0 a 1 y para la segunda se debe mencionar el método

³⁷ Véase el tema de con-elación en el siguiente capítulo.

utilizado de validación y su interpretación. De no ser así no podemos asegurar que el instrumento sea el adecuado. Si se selecciona un instrumento desarrollado en otro país, deben hacerse pruebas piloto más extensas (véase el paso G). También, no debe olvidarse que traducir no es validar un instrumento, por muy buena que sea la traducción.

O en caso de que no se elija un instrumento ya desarrollado, sino que se prefiera construir o desarrollar uno propio, debe pensarse en cada variable y sus dimensiones, y en indicadores precisos e ítems para cada dimensión. La figura 9.6 es un ejemplo de ello:

FIGURA 9.6 EJEMPLO DE DESARROLLO DE ÍTEMS

DEFINICIÓN	VARIABLE	OPERACIONAL	INDICADORES	DIMENSIONES	ITEM
S Coordinación, entre organizacio nes compradoras y proveedoras, desde el punto de vista de las primeras. Grado percibido mutuo de esfuerzo invertido para no provocar problemas a la otra parte al interferir en sus deberes y respon-sabilidades. Grado percibido mutuo de interés y buena voluntad de ambas partes. Coordinación de conflictos, ¿Qué tanto se esfuerza su empresa por no provocar problemas con sus proveedores?	1.	Se	esfuerza	al	mínimo posible.
2.	Se		esfuerza		poco.
3. Se esfuerza medianamente posible	¿Qué	tanto	se	esfuerzan	sus proveedores
por no provocar problemas con su empresa?	1.	Se esfuerzan	al	mínimo	posible.
esfuerzan poco.	3.	Se esfuerzan	medianamente.	4.	Se esfuerzan mucho.
esfuerzan al máximo posible.	Coordinación de no interferencia,	¿Cuánto se esfuerza su empresa	por no interferir en los deberes y responsabilidades de sus proveedores?	5.	Se esfuerza al máximo

posible.	4.	Se	esfuerza
mucho.	DEFINICIÓN	VARIABLE OPERACIONAL	INDICADORES DIMENSIONES I
TEMS	3. Se esfuerzan medianamente.	2. Se esfuerzan poco.	1. Se esfuerza
al mínimo posible.	¿Cuánto se esfuerzan sus proveedores por no interferir con los deberes y responsabilidades de su empresa?	5. Se esfuerzan al máximo posible.	4. Se esfuerzan
mucho.	3. Se esfuerzan medianamente.	2. Se esfuerzan poco.	1. Se
esfuerzan al mínimo posible.	Coordinación de objetivos, ¿Cuánto se esfuerza la empresa por trabajar junto con sus proveedores —de manera constante— para alcanzar objetivos comunes?	4. Se esfuerza mucho.	3. Se
esfuerza medianamente.	5. Se esfuerza al máximo posible.	2. Se esfuerza poco.	1. Se esfuerza al mínimo
posible.	¿Cuánto se esfuerzan los proveedores por trabajar junto con su empresa —de manera constante— para alcanzar objetivos comunes?	1. Se esfuerza al mínimo	posible.

DEFINICIÓN	VARIABLE OPERACIONAL	INDICADORES	DIMENSIONES	ITEMS
mucho.	Coordinación de objetivos.	5. Se esfuerzan al máximo posible.	4. Se esfuerzan	1.
Se esfuerzan al mínimo posible.	3. Se esfuerzan medianamente.	diariamente.	2. Se esfuerzan poco.	1.
bien el trato con sus proveedores?	establecidas	están de	las rutinas para	empresa
Medianamente establecidas.	5. Muy bien establecidas.	4. Bien establecidas.	3.	3.
establecidas.	2. Mal establecidas.	1. Muy mal	1. Muy mal	establecidas.
organizaciones.	Frecuencia de la interacción entre organizaciones.	Lapsos de interacciones entre	representantes.	Estableciendo'
promedio recibe	aproximado su	¿qué empresa	tan	seguido
visita de los representantes de sus proveedores verdaderamente importantes?				la
día	12. Una vez al día.	11. Tres veces por semana.	10. Dos veces por	13. Vanas veces al
semana.	9. Una vez a la semana.	8. Tres veces al mes.	7. Dos veces al	mes.
mes.	6. Una vez al mes.			

DEFINICIÓN	VARIABLE OPERACIONAL	INDICADORES	DIMENSIONES	ITEMS
5. Una vez cada dos meses.	4. Una vez cada cuatro meses.	3. Una vez		
cada seis meses.	2. Una vez al año.	1. Otra (especifique).	Estableciendo un	
promedio aproximado	¿qué tan seguido recibe su empresa la visita de los representantes de sus	proveedores poco importantes?	13. Varias veces al día	12. Una vez al día.
veces por semana.	10. Dos veces por semana.	9. Una vez a la semana.	8. Tres veces al	11. Tres
mes.	7. Dos veces al mes.	6. Una vez al mes.	5. Una vez cada dos meses.	4. Una
vez cada cuatro meses.	3. Una vez cada seis meses.	2. Una vez al año.	1. Otra	
(especifique).	Llamadas telefónicas, Estableciendo un promedio aproximado	¿qué tan seguido		
le llaman por teléfono a su empresa los representantes de sus proveedores muy importantes?				

DEFINICIÓN	VARIABLE OPERACIONAL	INDICADORES	DIMENSIONES	ITEMS
13. Varias veces al día	12. Una vez al día.	11. Tres veces por semana.	10. Dos veces por	

semana. 9. Una vez ala semana. 8. Tres veces al mes. 7. Dos veces al mes. 6. Una vez al mes. 5. Una vez cada dos meses. 4. Una vez cada cuatro meses. 3. Una vez cada seis meses. 2. Una vez al año. 1. Otra (especifique). Etcétera.

En este segundo caso, debemos asegurarnos de tener un suficiente número de ítems para medir todas las variables en todas sus dimensiones. Ya sea que se seleccione un instrumento previamente desarrollado y se adapte o bien, se construya uno, éste constituye la versión preliminar de nuestra medición. Versión que debe pulirse y ajustarse, como se verá más adelante.

e) **INDICAR EL NIVEL DE MEDICIÓN DE CADA ÍTEM Y, por ende, EL DE LAS VARIABLES.**

Existen **CUATRO NIVELES DE MEDICIÓN ampliamente conocidos:**

1. *Nivel de medición nominal.* En este nivel se tienen dos o más categorías del ítem o variable. Las categorías no tienen orden o jerarquía. Lo que se mide es colocado en una u otra categoría, lo que indica solamente diferencias respecto a una o más características.

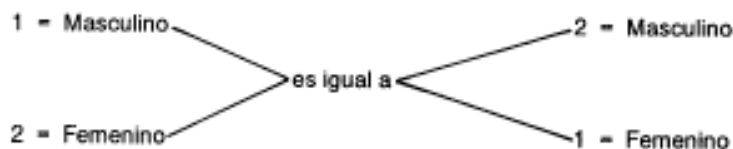
Por ejemplo, la variable sexo de la persona tiene sólo dos categorías: masculino y femenino (si la variable fuera “práctica sexual” podría haber tal vez más, pero sexo sólo tiene dos categorías). Ninguna de las categorías tiene mayor jerarquía que la otra, las categorías únicamente reflejan diferencias en la variable. No hay orden de mayor a menor.



Si les asignamos una etiqueta o símbolo a cada categoría, éste exclusivamente identifica a la categoría. Por ejemplo:

* = Masculino
z = Femenino

Si usamos numerales es lo mismo:



Los números utilizados en este nivel de medición tienen una función puramente de clasificación y *no* se pueden manipular aritméticamente. Por ejemplo, la afiliación religiosa es una variable nominal, si pretendiéramos operarla aritméticamente tendríamos situaciones tan ridículas como ésta:

1 = Católico
 2 = Judío 1+2=3
 3 = Protestante
 4 = Musulmán Un católico + un judío = protestante?
 5 = Otros

No tiene sentido.

Las variables nominales pueden incluir dos categorías (se les llama dicotómicas) o tres o más categorías (se les llama categóricas). Ejemplos de variables nominales dicotómicas sería el sexo y el tipo de escuela a la que se asiste (privada-pública); y de nominales categóricas tendríamos a la afiliación política (Partido A, Partido B,...), la carrera elegida, la raza, el departamento o provincia o estado de nacimiento y el canal de televisión preferido.

2. *Nivel de medición ordinal.* En este nivel se tienen varias categorías, pero además éstas mantienen un orden de mayor a menor. Las etiquetas o símbolos de las categorías sí indican jerarquía. Por ejemplo, el prestigio ocupacional en los Estados Unidos ha sido medido por diversas escalas que ordenan a las profesiones de acuerdo con su prestigio, por ejemplo:³⁸

<u>Valor en la escala</u>	<u>Profesión</u>
90	Ingeniero químico.
80	Científico de ciencias naturales <excluyendo la Química>.
60	Actor
50	Operador de estaciones eléctricas de potencia.
02	Manufactureros de tabaco.

90 es más que 80, 80 más que 60, 60 más que 50 y así sucesivamente (los números —símbolos de categorías— definen posiciones). Sin embargo, las categorías no están ubicadas a intervalos iguales (no hay un intervalo común). No podríamos decir con exactitud que entre un actor (60) y un operador de estaciones de poder (50) existe la misma distancia —en prestigio— que entre un científico de las ciencias naturales (80) y un ingeniero químico (90). Aparentemente en ambos casos la distancia es 10, pero no es una distancia real. Otra escala³⁹ clasificó el prestigio de dichas profesiones de la siguiente manera:

<u>Valor en la escala</u>	<u>Profesión</u>
98	Ingeniero químico.
95	Científico de ciencias naturales (excluyendo la Química).
84	Actor.
78	Operador de estaciones eléctricas de potencia.
13	Manufactureros de tabaco.

Aquí la distancia entre un actor (84) y un operador de estaciones (78) es de 6, y la distancia entre un ingeniero químico (98) y un científico de ciencias naturales (95) es de 3.

Otro ejemplo sería la posición jerárquica en la empresa:

Presidente	10
Vicepresidente	9
Director General	8
Gerente de Área	7
Subgerente o Superintendente	6
Jefe	5
Empleado A	4
Empleado B	3
Empleado C	2
Intendencia	1

Sabemos que el Presidente (10) es más que el Vicepresidente (9), éste más que el Director General (8), a su vez este último más que el Gerente (7) y así sucesivamente; pero no puede precisarse en cada caso cuánto más. Tampoco podemos utilizar las operaciones aritméticas básicas: No podríamos decir que 4 (Empleado A) + 5 (jefe) = 9 (Vicepresidente), ni que 10 (Presidente) \pm 5 (jefe) = 2 (Empleado C). Sería absurdo, no tiene sentido.

3. *Nivel de medición por intervalos.* Además de haber orden o jerarquía entre categorías, se establecen intervalos iguales en la medición. Las distancias entre categorías son las mismas a lo largo de toda la escala. Hay intervalo constante, una unidad de medida.



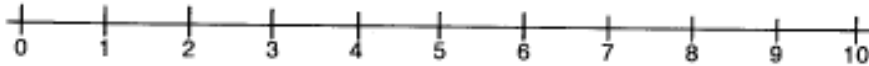
Por ejemplo: Una prueba de resolución de problemas matemáticos (30 problemas de igual dificultad). Si Ana Cecilia resolvió 10, Laura resolvió 20 y Brenda 30. La distancia entre Ana Cecilia y Laura es igual a la distancia entre Laura y Brenda.

Sin embargo, el cero (0) en la medición, es un cero arbitrario, no es real (se asigna arbitrariamente a una categoría el valor de cero y a partir de ésta se construye la escala). Un ejemplo clásico en ciencias naturales es la temperatura (en grados centígrados y Fahrenheit), el cero es arbitrario, no implica que realmente haya cero (ninguna) temperatura (incluso en ambas escalas el cero es diferente).

³⁹ Nam et al. (1975).

Cabe agregar que diversas mediciones en el estudio del comportamiento humano no son verdaderamente de intervalo (y. g., escalas de actitudes, pruebas de inteligencia y de otros tipos), pero se acercan a este nivel y se suele tratarlas como si fueran mediciones de intervalo. Esto se hace porque este nivel de medición permite utilizar las operaciones aritméticas básicas (suma, resta, multiplicación y división) y algunas estadísticas modernas, que de otro modo no se usarían. Aunque algunos investigadores no están de acuerdo en suponer tales mediciones como si fueran de intervalo (pero estos investigadores son minoría).

4. *Nivel de medición de razón.* En este nivel, además de tenerse todas las características del nivel de intervalos (intervalos iguales entre las categorías y aplicación de operaciones aritméticas básicas y sus derivaciones), el cero es real, es absoluto (no es arbitrario). Cero absoluto implica que hay un punto en la escala donde no existe la propiedad.



Es real

Ejemplos de estas mediciones serían la exposición a la televisión, el número de hijos, la productividad, las ventas de un producto y el ingreso.

Desde luego, hay variables que pueden medirse en más de un nivel, según el propósito de medición. Por ejemplo, la variable “antigüedad en la empresa”:

<u>Nivel de medición</u>	<u>Categorías</u>
— De razón	En días (0 a K días)
— Ordinal	Bastante antigüedad
	Regular antigüedad
	Poca antigüedad

Es muy importante indicar el nivel de medición de todas las variables e ítems de la investigación, porque dependiendo de dicho nivel se selecciona uno u otro tipo de análisis estadístico (por ejemplo, la prueba estadística para correlacionar dos variables de intervalo es muy distinta a la prueba para correlacionar dos variables ordinales).

Así, es necesario hacer una relación de variables, ítems y niveles de medición.

f) **INDICAR LA MANERA COMO SE HABRÁN DE CODIFICAR LOS DATOS** en cada ítem y variable. **CODIFICAR** los datos significa asignarles un valor numérico que los represente. Es **decir, a las categorías de cada ítem y variable se les asignan valores numéricos que tienen un significado**. Por ejemplo, si tuviéramos la variable “sexo” con sus respectivas categorías, “masculino” y “femenino”, a cada categoría le asignaríamos un valor. Éste podría ser:

<u>Categoría</u>	<u>Codificación (valor asignado)</u>
— Masculino	1
— Femenino	2

Así, Carla Magaña en la variable sexo sería un “2”. Luis Gerardo Vera y Rubén Reyes serían un “1”, Verónica Larios un “2” y así sucesivamente.

Otro ejemplo sería la variable “horas de exposición diaria a la televisión”, que podría codificarse de la siguiente manera:

<u>Categoría</u>	<u>Codificación (valor asignado)</u>
— No ve televisión	0
— Menos de una hora	1
— Una hora	2
— Más de una hora, pero menos de dos	3
— Dos horas	4
— Más de dos horas, pero menos de tres	5
— Tres horas	6

— Más de tres horas, pero menos de cuatro	7
— Cuatro horas	8
— Más de cuatro horas	9

Es necesario insistir que cada ítem y variable deberán tener una *codificación* (códigos numéricos) para sus categorías. Desde luego, hay veces que un ítem no puede ser codificado a priori (precodificado) porque es sumamente difícil conocer cuáles serán sus categorías. Por ejemplo, si en una investigación fuéramos a preguntar: ¿Qué opina del programa económico que recientemente aplicó el Gobierno? Es posible que las categorías encontradas podrían ser muchas más de las que nos imaginemos y resultaría difícil predecir con precisión cuántas y cuáles serán. En estos casos la codificación se lleva a cabo una vez que se aplica el ítem (*a posteriori*). A lo largo de este capítulo se profundizará en la forma de codificar y sus implicaciones. Por el momento, lo importante es que se comprenda el significado de codificar y que el instrumento de medición, antes de aplicarse, debe ir precodificado hasta donde sea posible (codificar los ítems cuyas categorías sean conocidas de antemano).

La codificación es necesaria para poder cuantitativamente analizar los datos (aplicar análisis estadístico). A veces se utilizan letras o símbolos en lugar de números (*, A, Z).

g) **UNA VEZ QUE SE INDICA EL NIVEL DE MEDICIÓN DE CADA VARIABLE E ÍTEM Y QUE SE DETERMINA SU CODIFICACIÓN, SE PROCEDE A APLICAR UNA “PRUEBA PILOTO” DEL INSTRUMENTO DE MEDICIÓN. Es decir, se aplica a personas con características semejantes a las de la muestra o población objetivo de la investigación.**

En esta prueba se analiza si las instrucciones se comprenden y si los ítems funcionan adecuadamente. Los resultados se usan para calcular la confiabilidad —y de ser posible la validez— del instrumento de medición. La *prueba piloto* se realiza con una pequeña muestra (inferior a la muestra definitiva). Los autores aconsejamos que cuando la muestra sea de 2000 más, se lleve a cabo la prueba piloto con entre 25 y 60 personas. Salvo que la investigación exija un número mayor.

h) **SOBRE LA BASE DE LA PRUEBA PILOTO, EL INSTRUMENTO DE MEDICIÓN PRELIMINAR SE MODIFICA, AJUSTA Y SE MEJORA, LOS INDICADORES DE CONFIABILIDAD Y VALIDEZ SON UNA BUENA AYUDA. Y ESTAREMOS EN CONDICIONES DE APLICARLO.**

Este procedimiento general para desarrollar una medición debe —desde luego— adaptarse a las características de los diferentes tipos de instrumentos de que disponemos en el estudio del comportamiento, los cuales veremos a continuación.

9.6. ¿DE QUÉ TIPOS DE INSTRUMENTOS DE MEDICIÓN O RECOLECCIÓN DE LOS DATOS DISPONEMOS EN LA INVESTIGACIÓN SOCIAL?

En la investigación del comportamiento disponemos de diversos tipos de instrumentos para medir las variables de interés y en algunos casos se pueden combinar dos o más métodos de recolección de los datos. A continuación describimos —brevemente— estos métodos o tipos de instrumentos de medición.

9.6.1. Escalas para medir las actitudes

Una *actitud* es una predisposición aprendida para responder consistentemente de una manera favorable o desfavorable respecto a un objeto o sus símbolos (Fishbein y Ajzen, 1975; Oskamp, 1977). Así, los seres

humanos tenemos actitudes hacia muy diversos objetos o símbolos, por ejemplo: actitudes hacia el aborto, la política económica, la familia, un profesor, diferentes grupos étnicos, la Ley, nuestro trabajo, el nacionalismo, hacia nosotros mismos, etcétera.

Las actitudes están relacionadas con el comportamiento que mantenemos en torno a los objetos a que hacen referencia. Si mi actitud hacia el aborto es desfavorable, probablemente no abortaría o no participaría en un aborto. Si mi actitud es favorable a un partido político, lo más probable es que vote por él en las próximas elecciones. Desde luego, las actitudes sólo son un indicador de la conducta, pero no la conducta en sí. Es por ello que las mediciones de actitudes deben interpretarse como “síntomas” y no como “hechos” (Padua, 1979). Por ejemplo, si detecto que la actitud de un grupo hacia la contaminación es desfavorable, esto no significa que las personas están adoptando acciones para evitar contaminar el ambiente, pero sí es un indicador de que pueden ir las adoptando paulatinamente. La actitud es como una “semilla”, que bajo ciertas condiciones puede “germinar en comportamiento”.

Las actitudes tienen diversas propiedades, entre las que destacan: dirección (positiva o negativa) e intensidad (alta o baja), estas propiedades forman parte de la medición.

Los métodos más conocidos para medir por escalas las variables que constituyen actitudes son: el método de escalamiento Likert, el diferencial semántico y la escala de Guttman. Hablemos de cada método.

Escalamiento tipo Likert⁴⁰

Este método fue desarrollado por Rensis Likert a principios de los treinta; sin embargo, se trata de un enfoque vigente y bastante popularizado. Consiste en un conjunto de ítems presentados en forma de afirmaciones o juicios ante los cuales se pide la reacción de los sujetos a los que se les administra. Es decir, se presenta cada afirmación y se pide al sujeto que externé su reacción eligiendo uno de los cinco puntos de la escala. A cada punto se le asigna un valor numérico. Así, el sujeto obtiene una puntuación respecto a la afirmación y al final se obtiene su puntuación total sumando las puntuaciones obtenidas en relación a todas las afirmaciones.

Las afirmaciones califican al objeto de actitud que se está midiendo y deben expresar sólo una relación lógica, además es muy recomendable que no excedan de —aproximadamente— 20 palabras.

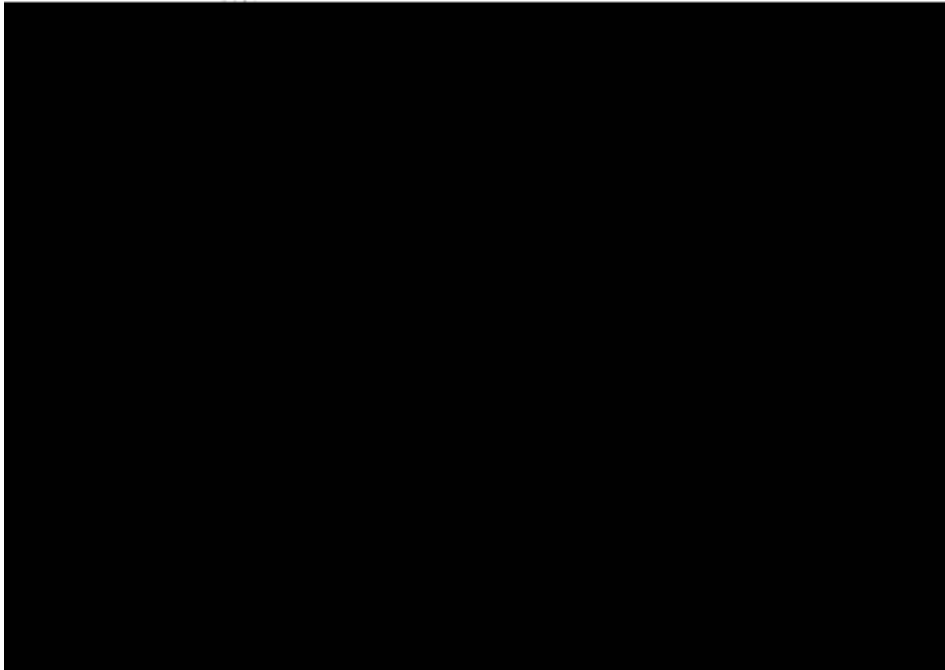
EJEMPLO

<u>Objeto de actitud medido</u>	<u>Afirmación</u>
El voto ciudadano responsable”	“Votar es una obligación de todo

⁴⁰ Para profundizar en esta técnica se recomienda consultar a Likert (1976a o 1976b), Seiler y Hough (1976) y Padua (1979).

En este caso la afirmación incluye 8 palabras y expresa una sola relación lógica (X—Y). Las alternativas de respuesta o puntos de la escala son cinco e indican cuánto se está de acuerdo con la afirmación correspondiente. Las alternativas más comunes se presentan en la figura 9.7. Debe recordarse que a cada una de ellas se le asigna un valor numérico y sólo puede marcarse una opción. Se considera un dato inválido a quien marque dos o más opciones.

FIGURA 9.7



Asimismo, pueden hacerse distintas combinaciones como “totalmente verdadero” o “completamente no”. Y las alternativas de respuesta pueden colocarse horizontalmente —como en la figura 9.7— o verticalmente.

EJEMPLO

- Muy de acuerdo
- De acuerdo
- Ni de acuerdo, ni en desacuerdo
- En desacuerdo
- Muy en desacuerdo

O bien utilizando recuadros en lugar de paréntesis:

Definitivamente sí

Probablemente sí

Indeciso

Probablemente no

Definitivamente no

Es indispensable comentar que el número de categorías de respuesta debe ser el mismo para todas las afirmaciones.

DIRECCIÓN DE LAS AFIRMACIONES

Las afirmaciones pueden tener dirección: *favorable o positiva* y *desfavorable o negativa*. Y esta dirección es muy importante para saber cómo se codifican las alternativas de respuesta.

Si la afirmación es positiva significa que califica favorablemente al objeto de actitud, y entre los sujetos estén más de acuerdo con la afirmación, su actitud es más favorable.

EJEMPLO

“El Ministerio de Hacienda ayuda al contribuyente a resolver sus problemas en el pago de impuestos”.

Si estamos ‘muy de acuerdo’ implica una actitud más favorable hacia el Ministerio de Hacienda que si estamos “de acuerdo”. En cambio, si estamos “muy en desacuerdo” implica una actitud muy desfavorable. Por lo tanto, *cuando las afirmaciones son positivas se califican comúnmente de la siguiente manera:*

- (5) Muy de acuerdo
- (4) De acuerdo
- (3) Ni de acuerdo, ni en desacuerdo
- (2) En desacuerdo
- (1) Muy en desacuerdo

Es decir, estar más de acuerdo implica una puntuación mayor.

Si la afirmación es negativa significa que califica desfavorablemente al objeto de actitud, y entre los sujetos estén más de acuerdo con la afirmación, su actitud es menos favorable, esto es, más desfavorable.

EJEMPLO

“El Ministerio de Hacienda se caracteriza por obstaculizar al contribuyente en el pago de impuestos”.

Si estamos “muy de acuerdo” implica una actitud más desfavorable que si estamos de “de acuerdo” y así sucesivamente. En contraste, si estamos “muy en desacuerdo” implica una actitud favorable hacia el Ministerio de Hacienda. Rechazamos la frase porque califica negativamente al objeto de actitud. Un ejemplo cotidiano de afirmación negativa sería: “Luis es un mal amigo”, entre más de acuerdo estemos con la afirmación, nuestra actitud hacia Luis es menos favorable. Es decir, estar más de acuerdo implica una puntuación menor. *Cuando las afirmaciones son negativas se califican al contrario de las positivas.*

EJEMPLO

- (1) Totalmente de acuerdo
- (2) De acuerdo
- (3) Ni de acuerdo, ni en desacuerdo
- (4) En desacuerdo
- (5) Totalmente en desacuerdo

En la figura 9.8. se presenta un ejemplo de una escala Likert para medir la actitud hacia un organismo tributario.⁴¹

⁴¹ El ejemplo fue utilizado en un país latinoamericano y su confiabilidad total fue de 0.89; aquí se presenta una versión reducida de la escala original. El nombre del organismo tributario que aquí se utiliza es ficticio.

FIGURA 9.8

EJEMPLO DE UNA ESCALA LIKERT

LAS AFIRMACIONES QUE VOY A LEERLE SON OPINIONES CON LAS QUE ALGUNAS PERSONAS ES

TAN DE ACUERDO Y OTRAS EN DESACUERDO. VOY A PEDIRLE QUE ME DIGA POR FAVOR QUE TAN DE ACUERDO ESTÁ USTED CON CADA UNA DE ESTAS OPINIONES.

1. *“EL PERSONAL DE LA DIRECCIÓN GENERAL DE IMPUESTOS NACIONALES ES GROSERO AL ATENDER AL PÚBLICO”.*

- 1) Muy de acuerdo 3) Ni de acuerdo, ni en desacuerdo
2) De acuerdo 4) En desacuerdo 5) Muy en desacuerdo

2. *“LA DIRECCIÓN GENERAL DE IMPUESTOS NACIONALES SE CARACTERIZA POR LA DESHONESTIDAD DE SUS FUNCIONARIOS”.*

- 1) Muy de acuerdo 3) Ni de acuerdo, ni en desacuerdo
2) De acuerdo 4) En desacuerdo 5) Muy en desacuerdo

3. *“LOS SERVICIOS QUE PRESTA LA DIRECCIÓN GENERAL DE IMPUESTOS NACIONALES SON EN GENERAL MUY BUENOS”.*

- 5) Muy de acuerdo 3) Ni de acuerdo, ni en desacuerdo
4) De acuerdo 2) En desacuerdo 1) Muy en desacuerdo

4. *“LA DIRECCIÓN GENERAL DE IMPUESTOS NACIONALES INFORMA CLARAMENTE SOBRE CÓMO, DÓNDE Y CUÁNDO PAGAR LOS IMPUESTOS”.*

- 5) Muy de acuerdo 3) Ni de acuerdo, ni en desacuerdo
4) De acuerdo 2) En desacuerdo 1) Muy en desacuerdo

5. *“LA DIRECCIÓN GENERAL DE IMPUESTOS NACIONALES ES MUY LENTA EN LA DEVOLUCIÓN DE IMPUESTOS PAGADOS EN EXCESO”.*

- 1) Muy de acuerdo 3) Ni de acuerdo, ni en desacuerdo
2) De acuerdo 4) En desacuerdo 5) Muy en desacuerdo

6. *“LA DIRECCIÓN GENERAL DE IMPUESTOS NACIONALES INFORMA OPORTUNAMENTE SOBRE CÓMO, DÓNDE Y CUÁNDO PAGAR LOS IMPUESTOS”.*

- 5) Muy de acuerdo 3) Ni de acuerdo, ni en desacuerdo
4) De acuerdo 2) En desacuerdo 1) Muy en desacuerdo

7. *“LA DIRECCIÓN GENERAL DE IMPUESTOS NACIONALES TIENE NORMAS Y PROCEDIMIENTOS BIEN DEFINIDOS PARA EL PAGO DE IMPUESTOS”.*

- 5) Muy de acuerdo 3) Ni de acuerdo, ni en desacuerdo
4) De acuerdo 2) En desacuerdo 1) Muy en desacuerdo

8. *“LA DIRECCIÓN GENERAL DE IMPUESTOS NACIONALES TIENE MALAS RELACIONES CON LA GENTE PORQUE COBRA IMPUESTOS MUY ALTOS”.*

- 1) Muy de acuerdo 3) Ni de acuerdo, ni en desacuerdo
2) De acuerdo 4) En desacuerdo 5) Muy en desacuerdo

Como puede observarse en la figura 9.8, las afirmaciones 1, 2, 5 y 8 son negativas (desfavorables) y las afirmaciones 3, 4, 6 y 7 son positivas (favorables).

FORMA DE OBTENER LAS PUNTUACIONES

Las puntuaciones de las escalas Likert se obtienen sumando los valores obtenidos respecto a cada frase. Por ello se le denomina escala aditiva. La figura 9.9 constituiría un ejemplo de cómo calificar una escala de Likert:

FIGURA 9.9

EJEMPLO DE CÓMO CALIFICAR UNA ESCALA LIKERT

1. El personal de la dirección general de impuestos nacionales es grosero al atender al público.

<input checked="" type="checkbox"/> 1) Muy de acuerdo	<input type="checkbox"/> 3) Ni de acuerdo, ni en desacuerdo	
<input type="checkbox"/> 2) De acuerdo	<input type="checkbox"/> 4) En desacuerdo	<input type="checkbox"/> 5) Muy en desacuerdo

2. La dirección general de impuestos nacionales se caracteriza por la deshonestidad de sus funcionarios.

<input checked="" type="checkbox"/> 1) Muy de acuerdo	<input type="checkbox"/> 3) Ni de acuerdo, ni en desacuerdo	
<input type="checkbox"/> 2) De acuerdo	<input type="checkbox"/> 4) En desacuerdo	<input type="checkbox"/> 5) Muy en desacuerdo

3. Los servicios que presta la dirección general de impuestos nacionales son en general muy buenos.

<input type="checkbox"/> 5) Muy de acuerdo	<input type="checkbox"/> 3) Ni de acuerdo, ni en desacuerdo	
<input type="checkbox"/> 4) De acuerdo	<input type="checkbox"/> 2) En desacuerdo	<input checked="" type="checkbox"/> 1) Muy en desacuerdo

4. La dirección general de impuestos nacionales informa claramente sobre cómo, dónde y cuándo pagar los impuestos.

<input type="checkbox"/> 5) Muy de acuerdo	<input checked="" type="checkbox"/> 3) Ni de acuerdo, ni en desacuerdo	
<input type="checkbox"/> 4) De acuerdo	<input type="checkbox"/> 2) En desacuerdo	<input type="checkbox"/> 1) Muy en desacuerdo

5. La dirección general de impuestos nacionales es muy lenta en la devolución de impuestos pagados en exceso.

<input checked="" type="checkbox"/> 1) Muy de acuerdo	<input type="checkbox"/> 3) Ni de acuerdo, ni en desacuerdo	
<input type="checkbox"/> 2) De acuerdo	<input type="checkbox"/> 4) En desacuerdo	<input type="checkbox"/> 5) Muy en desacuerdo

6. La dirección general de impuestos nacionales informa oportunamente sobre cómo, dónde y cuándo pagar los impuestos.

<input type="checkbox"/> 5) Muy de acuerdo	<input type="checkbox"/> 3) Ni de acuerdo, ni en desacuerdo	
<input type="checkbox"/> 4) De acuerdo	<input type="checkbox"/> 2) En desacuerdo	<input checked="" type="checkbox"/> 1) Muy en desacuerdo

7. La dirección general de impuestos nacionales tiene normas y procedimientos bien definidos para el pago de impuestos.

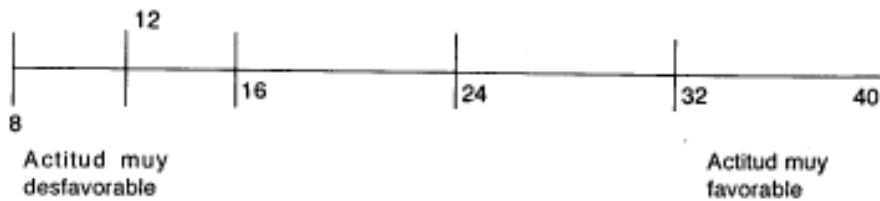
<input type="checkbox"/> 5) Muy de acuerdo	<input type="checkbox"/> 3) Ni de acuerdo, ni en desacuerdo	
<input type="checkbox"/> 4) De acuerdo	<input checked="" type="checkbox"/> 2) En desacuerdo	<input type="checkbox"/> 1) Muy en desacuerdo

8. La dirección general de impuestos nacionales tiene malas relaciones con la gente porque cobra impuestos muy altos.

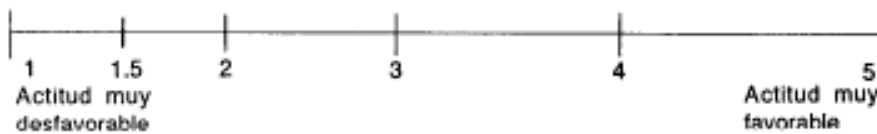
<input checked="" type="checkbox"/> 1) Muy de acuerdo	<input type="checkbox"/> 3) Ni de acuerdo ni en desacuerdo	
<input type="checkbox"/> 2) De acuerdo	<input type="checkbox"/> 4) En desacuerdo	<input type="checkbox"/> 5) Muy en desacuerdo

VALOR = 1 + 2 + 1 + 3 + 1 + 1 + 2 + 1 = 12

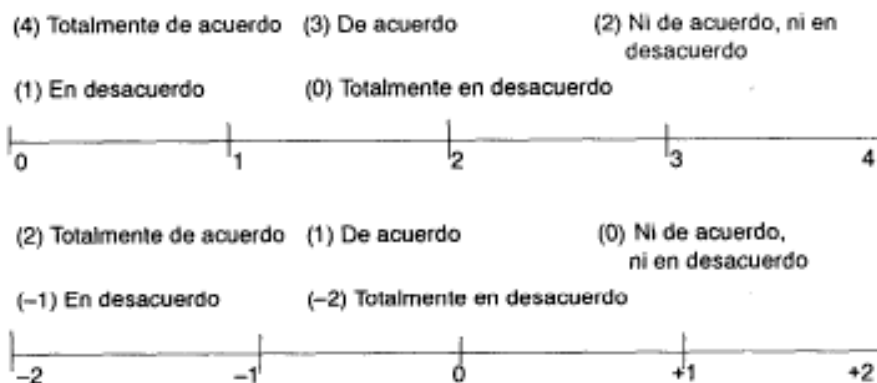
Una puntuación se considera alta o baja según el número de ítems o afirmaciones. Por ejemplo, en la escala para evaluar la actitud hacia el organismo tributario la puntuación mínima posible es de 8 (1+1+1+1+1+1+1+1) y la máxima es de 40 (5+5+5+5+5+5+5+5), porque hay ocho afirmaciones. La persona del ejemplo obtuvo “12”, su actitud hacia el organismo tributario es más bien sumamente desfavorable, veámoslo gráficamente:



Si alguien hubiera tenido una puntuación de 37 (5+5+4+5+5+4+4+5) su actitud puede calificarse como sumamente favorable. En las escalas Likert a veces se califica el promedio obtenido en la escala mediante la sencilla fórmula $\frac{PT}{NT}$ (donde PT es la puntuación total en la escala y NT es el número de afirmaciones), y entonces una puntuación se analiza en el continuo 1-5 de la siguiente manera, con el ejemplo de quien obtuvo 12 en la escala ($\frac{12}{8} = 1.5$):



La escala Likert es, en estricto sentido, una medición ordinal sin embargo, es común que se le trabaje como si fuera de intervalo. Asimismo, a veces se utiliza un rango de 0 a 4 o de -2 a +2 en lugar de 1 a 5. Pero esto no importa porque se cambia el marco de referencia de la interpretación. Veámoslo gráficamente.



Simplemente se ajusta el marco de referencia, pero el rango se mantiene y las categorías continúan siendo cinco.

OTRAS CONSIDERACIONES SOBRE LA ESCALA LIKERT

A veces se acorta o incrementa el número de categorías, sobre todo cuando los respondientes potenciales pueden tener una capacidad muy limitada de discriminación o por el contrario muy amplia.

EJEMPLOS (CON AFIRMACIONES POSITIVAS)

(1) De acuerdo	(0) En desacuerdo	
(3) De acuerdo	(2) Ni de acuerdo ni en desacuerdo	(1) En desacuerdo
(7) Totalmente de acuerdo	(6) De acuerdo	(5) Indeciso, pero más bien de acuerdo
(4) Indeciso, ni de acuerdo ni en desacuerdo	(3) Indeciso, pero más bien en desacuerdo	
(2) En desacuerdo	(1) Totalmente en desacuerdo	

Si los respondientes tienen poca capacidad de discriminar pueden incluirse dos o tres categorías. Por el contrario, si son personas con un nivel educativo elevado y capacidad de discriminación, pueden incluirse siete categorías. Pero debe recalcarse que el número de categorías de respuesta debe ser el mismo para todos los ítems, si son tres, son tres categorías para todos los ítems o afirmaciones. Si son cinco, son cinco categorías para todos los ítems.

Un aspecto muy importante de la escala Likert es que asume que los ítems o afirmaciones miden la actitud hacia un único concepto subyacente, si se van a medir actitudes hacia varios objetos, deberá incluirse una escala por objeto aunque se presenten conjuntamente, pero se califican por separado. En cada escala se considera que todos los ítems tienen igual peso.

COMO SE CONSTRUYE UNA ESCALA LIKERT

En términos generales, una escala Likert se construye *generando un elevado número de afirmaciones que califiquen al objeto de actitud y se administran a un grupo piloto para obtener las puntuaciones del grupo en cada afirmación. Estas puntuaciones se correlacionan con las puntuaciones del grupo a toda la escala (la suma de las puntuaciones de todas las afirmaciones), y las afirmaciones cuyas puntuaciones se correlacionen significativamente con las puntuaciones de toda la escala, se seleccionan para integrar el instrumento de medición.* Asimismo, debe calcularse la *confiabilidad* y *validez* de la escala.

PREGUNTAS EN LUGAR DE AFIRMACIONES

En la actualidad, la escala original se ha extendido a preguntas y observaciones. Por ejemplo:

¿CÓMO CONSIDERA USTED AL CONDUCTOR QUE APARECE EN LOS PROGRAMAS?

5	Muy buen conductor	4	Buen conductor	3	Regular
2	Mal conductor	1	Muy mal conductor		

Esta pregunta se hizo como parte de la evaluación de un video empresarial.

Otro ejemplo sería una pregunta que se hizo en una investigación para analizar la relación de compra-venta

en empresa de la Ciudad de México (Paniagua, 1986). De ella se presenta un fragmento en la figura 9.10.

FIGURA 9.10
EJEMPLO DE LA ESCALA LIKERT APLICADA A UNA PREGUNTA PARA ELEGIR SUS
PROVEEDORES, ¿QUÉ TAN IMPORTANTE ES...?

Indispensable	Sumamente importante	Medianamente importante	Poco importante	No se toma en cuenta
(5)	(4)	(3)	(2)	(1)

- Precio
- Forma de pago (contado-crédito)
- Tiempo de entrega
- Lugar de entrega

Indispensable importante	Sumamente importante	Medianamente importante	Poco importante	No se to
(5)	(4)	(3)	(2)	(1)

- Garantía del producto
- Servicio de reparación
- Prestigio del producto (marca)
- Prestigio de la empresa proveedora
- Comunicación que se tiene con la(s) persona(s) que representan al proveedor
- Apego del proveedor a los requerimientos legales del producto
- Cumplimiento el proveedor con las especificaciones
- Información que sobre el producto proporcione el proveedor
- Tiempo de trabajar con el proveedor
- Entrega del producto en las condiciones acordadas
- Calidad del producto
- Personalidad de los vendedores

Las respuestas se califican del mismo modo que ya hemos comentado.

MANERAS DE APLICAR LA ESCALA LIKERT

Existen dos formas básicas de aplicar una escala Likert. La primera es de manera *auto administrada*: se le entrega la escala al respondiente y éste marca respecto a cada afirmación, la categoría que mejor describe su

reacción o respuesta. Es decir, marcan su respuesta. La segunda forma es *la entrevista*; un entrevistador lee las afirmaciones y alternativas de respuesta al sujeto y anota lo que éste conteste. Cuando se aplica vía entrevista, es muy necesario que se le entregue al respondiente una tarjeta donde se muestran las alternativas de respuesta o categorías. El siguiente es un ejemplo que se aplica a la pregunta de la figura 9.10:

Indispen- sable	Sumamente importante	Medianamen- te importante	Poco importante	No se toma en cuenta
--------------------	-------------------------	------------------------------	--------------------	-------------------------

Al construir una escala Likert debemos asegurar que las afirmaciones y alternativas de respuesta serán comprendidas por los sujetos a los que se les aplicará y que éstos tendrán la capacidad de discriminación requerida. Ello se evalúa cuidadosamente en la prueba piloto.

*Diferencial semántico*⁴²

El diferencial semántico que desarrollado originalmente por *Osgood, Suci y Tannenbaum (1957)* para explorar las dimensiones del significado. Pero hoy en día *consiste en una serie de adjetivos extremos que califican al objeto de actitud ante los cuales se solicita la reacción del sujeto*. Es decir, éste tiene que calificar al objeto de actitud en un conjunto de adjetivos bipolares, entre cada par de adjetivos se presentan varias opciones y el sujeto selecciona aquella que refleje su actitud en mayor medida.

EJEMPLOS DE ESCALAS BIPOLARES

Objeto de actitud: Candidato "A"

justo: ____ : ____ : ____ : ____ : ____ : ____ : injusto

Debe observarse que *los adjetivos son "extremos"* y que *entre ellos hay siete opciones de respuesta*. Cada sujeto califica al candidato "A" en términos de esta escala de adjetivos bipolares.

Osgood, Suci y Tannenbaum (1957) nos indican que, si el respondiente considera que el objeto de actitud se relaciona *muy estrechamente* con uno u otro extremo de la escala, la respuesta se marca así:

justo: X : ____ : ____ : ____ : ____ : ____ : ____ : injusto

o de la siguiente manera:

justo: _____ : ____ : ____ : ____ : ____ : ____ : X : injusto

Si el respondiente considera que el objeto de actitud se relaciona *estrechamente* con uno u otro extremo de la escala, la respuesta se marca así (dependiendo del extremo en cuestión):

justo: _____ : X : ____ : ____ : ____ : ____ : ____ : injusto

justo: _____ : _____ : ____ : ____ : X : ____ : ____ : injusto

Si el respondiente considera que el objeto de actitud se relaciona *mediante* con alguno de los extremos, la respuesta se marca así (dependiendo del extremo en cuestión):

justo: _____ : _____ : X : ____ : ____ : ____ : ____ : injusto

justo: _____ : _____ : _____ : ____ : X : ____ : ____ : injusto

Y si el respondiente considera que el objeto de actitud ocupa una posición neutral en la escala (ni

justo ni injusto en este caso), la respuesta se marca así:

justo: _____ : _____ : _____ : X : _____ : _____ : _____ : injusto

Es decir, en el ejemplo, cuanto más justo considere al candidato “A” más me acerco al extremo “justo”, y viceversa, entre más injusto lo considero más me acerco al extremo opuesto.

Algunos ejemplos de adjetivos se muestran en la figura 9.11.

FIGURA 9.11
EJEMPLOS DE ADJETIVOS BIPOLARES

fuerte-débil	poderoso-impotente
grande-pequeño	vivo-muerto
bonito-feo	joven-viejo
alto-bajo	rápido-lento
claro-oscuro	gigante-enano
caliente-frío	perfecto-imperfecto
costoso-barato	agradable-desagradable
activo-pasivo	bendito-maldito
seguro-peligroso	arriba-abajo
bueno-malo	útil-inútil
dulce-ácido	favorable-desfavorable
profundo-superficial	agresivo-tímido

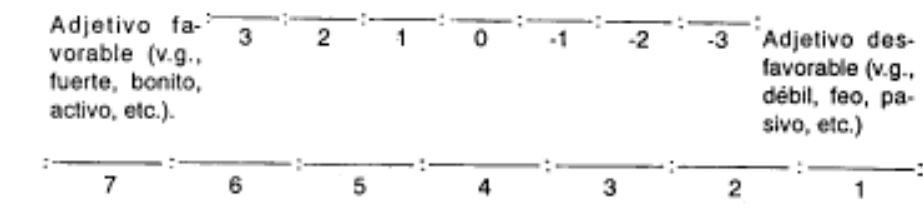
La figura 9.11 presenta sólo algunos ejemplos, desde luego hay muchos más que han sido utilizados o que pudieran pensarse. La elección de adjetivos depende del objeto de actitud a calificar, los adjetivos deben poder aplicarse a éste.

CODIFICACIÓN DE LAS ESCALAS

Los puntos o categorías de la escala pueden codificarse de diversos modos, éstos se uresentan en la fwura 9.12.

FIGURA 9.12

MANERAS COMUNES DE CODIFICAR EL DIFERENCIAL SEMÁNTICO.



En los casos en que los respondientes tengan mayor capacidad de discriminación, se pueden reducir las categorías a cinco opciones, por ejemplo:

sabroso : _____ : _____ : _____ : _____ : _____ : desabrido

FIGURA 9.14

EJEMPLO DE CÓMO CALIFICAR UN DIFERENCIAL SEMÁNTICO

sabroso	:	—	:	X	:	—	:	—	:	—	:	—	:	—	:	—	:	insipido
rico	:	X	:	—	:	—	:	—	:	—	:	—	:	—	:	—	:	pobre
suave	:	—	:	X	:	—	:	—	:	—	:	—	:	—	:	—	:	áspero
balanceado	:	—	:	X	:	—	:	—	:	—	:	—	:	—	:	—	:	desbalanceado

valor = 6 + 7 + 6 + 6 = 25

Su interpretación depende del número de ítems o pares de adjetivos. Asimismo, en ocasiones se califica el promedio obtenido en la escala total $\left[\frac{\text{Puntuación total}}{\text{número de ítems}} \right]$. Y se pueden utilizar distintas escalas o diferenciales semánticos para medir actitudes hacia varios objetos. Por ejemplo, podemos medir con cuatro pares de adjetivos la actitud hacia el candidato “A”, con otros tres pares de adjetivos la actitud respecto a su plataforma ideológica y con otros seis pares de adjetivos la actitud hacia su partido político. Tenemos tres escalas, cada una con distintos pares de adjetivos para medir la actitud en relación a tres diferentes objetos.

El diferencial semántico es estrictamente una escala de medición ordinal, pero es común que se le trabaje como si fuera de intervalo.

*Escalograma de Guttman*⁴³

Este método para medir actitudes fue desarrollado por *Louis Guttman*. Se basa en el principio de que algunos ítems indican en mayor medida la fuerza o intensidad de la actitud. *La escala está constituida por afirmaciones*, las cuales poseen las mismas características que en el caso de Likert. Pero *el escalograma garantiza que la escala mide una dimensión única*. Es decir, cada afirmación mide la misma dimensión de la misma variable, a esta propiedad se le conoce como “*unidimensionalidad*”.

Algunos autores consideran que el escalograma más que ser un método de medición de actitudes, es una técnica para determinar si un conjunto de afirmaciones reúnen los requisitos de un tipo particular de escala (v.g., Edwards —1957—).

Para construir el escalograma *es necesario desarrollar un conjunto de afirmaciones pertinentes al objeto de actitud. Estas deben variar en intensidad*. Por ejemplo, si pretendiéramos medir la actitud hacia la calidad en el trabajo dentro del nivel gerencial, la afirmación: “La calidad debe vivirse en todas las actividades del trabajo y en el hogar” es más intensa que la afirmación: “La calidad debe vivirse sólo en las actividades más importantes del trabajo”. Dichas afirmaciones *se aplican a una muestra a manera de prueba piloto. Y una vez administradas se procede a su análisis*. Cabe agregar que *las categorías de respuesta* para las afirmaciones, *pueden variar entre dos* (“de acuerdo-en desacuerdo”, sí-no , etcétera) *o más categorías* (v.g., las mismas categorías que en el caso de Likert).

TÉCNICA DE CORNELL

La manera más conocida de analizar los ítems o afirmaciones y desarrollar el escalograma es la *técnica Cornell* (Guttman, 1976). En ella se procede a:

1. *Obtener el puntaje total* de cada sujeto en la escala.
2. *Ordenar a los sujetos* de acuerdo con su puntaje total (del puntaje mayor al menor, de manera vertical descendente).
3. *Ordenar a las afirmaciones* de acuerdo con su intensidad (de mayor a menor y de izquierda a

O
O
O
O
O
O
O DA
11k ED
Á~L DA
11k ED
1~L

X
X
X
X
X
X
X
X DA
11k
X
X
X
X
X
X
X
X ED
(0) DA
(1) ED
(0)

X
X *
*
*

X
 X 4
 4
 4
 3
 3
 3
 2
 2
 2
 2
 1
 1
 0
 0 DA = De acuerdo o 1, ED = En desacuerdo o 0

Como puede observarse en la tabla 9.1, los sujetos están ordenados por su puntuación en la escala total. Asimismo, las frases deben estar ordenadas por su intensidad (en el ejemplo, A tiene mayor intensidad que B, B mayor que C y C mayor que D) y también sus categorías se encuentran jerarquizadas de acuerdo con su valor de izquierda a derecha. Hay que recordar que si la afirmación es negativa, la codificación se invierte (“La calidad es poco importante para el desarrollo de una empresa”, “de acuerdo” se codificaría con cero y “en desacuerdo” con uno). En el ejemplo de la tabla 9.1 tenemos cuatro afirmaciones positivas.

Los sujetos que estén “de acuerdo” con la afirmación “A”, que es la más intensa, es muy probable que también estén de acuerdo con las afirmaciones “B”, “C” y “D”, ya que su intensidad es menor. Los individuos que respondan “de acuerdo” a la afirmación “B”, tenderán a estar “de acuerdo” con “C” y “D” (afirmaciones menos intensas), pero no necesariamente con “A”. Quienes estén “de acuerdo” con “C”, lo más probable es que se encuentren “de acuerdo” con “D”, pero no necesariamente con “A” y “B”.

Debe observarse que el sujeto número 1 estuvo “de acuerdo” respecto a las cuatro afirmaciones. Los sujetos 2 y el 3 respondieron de igual forma. Las puntuaciones de todos ellos equivalen a 4 (1+1+1+1). Los sujetos 4, 5 y 6 obtuvieron una puntuación de 3, pues estuvieron “de acuerdo” con tres afirmaciones y así sucesivamente (los últimos dos sujetos estuvieron “en desacuerdo” respecto a todas las afirmaciones). Idealmente, los sujetos que obtienen una puntuación total de 4 en esta particular escala han respondido “de acuerdo” a las cuatro afirmaciones. Los individuos que alcanzan una puntuación total de 3 han respondido estar “en desacuerdo” con la primera afirmación pero están “de acuerdo” con las demás afirmaciones. Quienes reciben una puntuación de 2 manifiestan estar “en desacuerdo” con los dos primeros ítems pero “de acuerdo” con los dos últimos. Los sujetos con puntuación de 1, han respondido “en desacuerdo” a las tres primeras afirmaciones y “de acuerdo” a la última. Finalmente aquellos que hayan estado “en desacuerdo” respecto a las cuatro afirmaciones, tienen una puntuación total de 0.

Los sujetos se escalan de manera perfecta, sin que nadie rompa el patrón de intensidad de las afirmaciones: si están “de acuerdo” con una afirmación más intensa, de igual manera lo estarán con las menos intensas. Por

X
X X
X
X
X

X
X X

X
X

X
X
X 4
3
3
2
1
0

Los *errores* se detectan analizando las respuestas que rompen el patrón y para ello se establecen los “*puntos de corte*” en la tabla donde se cruzan las afirmaciones y sus categorías con las puntuaciones totales. En el ejemplo de la tabla 9.1 los “puntos de corte” serían los que se muestran en la tabla 9.3.

TABLA 9.3
EJEMPLO DE ESTABLECIMIENTO DE LOS PUNTOS
DE CORTE EN LA TÉCNICA DE CORNELL

SUJETO	AFIRMACIONES	PUNTUACIONES TOTALES	A	B	C	D	DA
(1)	ED						
(0)	DA						
(1)	ED						
(0)	DA						
(1)	ED						
(0)	DA						
(1)	ED						
(0)		1					
2							
3							
4							
5							
6							
7							
8							
9							
10							

X
X
X
X X
X
X
X
X
X
X
X
X
X
X
X
X
X

X
X 4
4
4
3
3
3
2
2
2
2
2
1
1
0

0 DA = De acuerdo o 1, ED = En desacuerdo o 0

No se aprecia ninguna inconsistencia. En cambio, en la tabla 9.4 se aprecian cuatro inconsistencias o errores, las respuestas están desubicadas respecto a los *puntos de corte*, rompen el patrón de intensidad (los errores están encerrados en un círculo).

TABLA 9.4
EJEMPLOS DE ERRORES RESPECTO A LOS PUNTOS DE CORTE

SUJETO	AFIRMACIONES	PUNTUACIONES TOTALES	A	B	C	D	DA
(1)	ED						
(0)	DA						

(1) ED

(0) DA

(1) ED

(0) DA

(1) ED

(0) 1

2

3

4

5

6

7

8

9

10

11

12

13

14 X

X

X

X

X

X

X

X

X

X

X

X

X

X X

X

X

X

X

X

X

.....
X
X
X
X
X
X
X
X X
X
X
X
X
X
X
X
X
X
X

X
X
X
X X
X
X
X
X
X
X
X
X
X
X

X
X
X

X 4
 4
 4
 4
 3
 3
 3
 2
 2
 2
 1
 1
 0
 0 - - - - - = puntos de corte (líneas punteadas)

Como se han mencionado anteriormente, cuando el número de errores es excesivo, la escala no presenta *reproductividad* y no puede aceptarse. *La reproductividad se determina mediante un coeficiente*. La fórmula de este coeficiente es:

$$\text{Coeficiente de reproductividad} = \frac{\text{Número de errores o inconsistencias}}{\text{Número total de respuestas}}$$

donde el número total de respuestas = número de ítems o afirmaciones X número de sujetos. Por lo tanto, la fórmula directa sería:

$$\text{Coeficiente de reproductividad} = \frac{\text{Número de errores}}{(\text{número de ítems})(\text{número de sujetos})}$$

En el ejemplo de la tabla 9.4 tendríamos que el coeficiente de reproductividad es:

$$Cr = 1 - \frac{4}{(4)(14)}$$

$$Cr = 1 - 0.07$$

$$Cr = .93$$

El *coeficiente de reproductividad oscila entre 0 y 1*, y cuando equivale a .90 o más nos indica que el número de errores es tolerable y la escala es unidimensional y se acepta. Cuando dicho coeficiente es menor a .90 no se acepta la escala. Guttman originalmente recomendó administrar un máximo de 10 a 12 ítems o afirmaciones a un mínimo de 100 personas (Black y Champion, 1976).

Una vez determinado el número de errores aceptable mediante el *coeficiente de reproductividad*, se procede a aplicar la escala definitiva (si dicho coeficiente fue de .90 o más, esto es, el error permitido no excedió al 10%) o a hacer ajustes en la escala (reconstruir ítems, eliminar ítems que estén generando errores, etcétera). Los cinco pasos mencionados son una especie de prueba piloto para demostrar que la escala es unidimensional y funciona.

CODIFICACIÓN DE RESPUESTAS

Cuando se aplica la versión definitiva de la escala los resultados se *codifican* de la misma manera que en la escala Likert, dependiendo del número de categorías de respuesta que se incluyan. Y al igual que la escala Likert y el diferencial semántico, *todos los ítems deben tener el mismo número de categorías de respuesta. Este es un requisito de todas las escalas de actitud.* Asimismo, se considera una respuesta inválida a quien marque dos o más opciones para una misma afirmación. El *escalograma de Guttman es una escala estrictamente ordinal pero que se suele trabajar como si fuera de intervalo.* Puede aplicarse mediante *entrevista* —con uso de tarjetas conteniendo las opciones o categorías de respuesta— o puede ser *auto administrada*.

9.6.2. Cuestionarios

Tal vez el instrumento más utilizado para recolectar los datos es el cuestionario. *Un cuestionario consiste en un conjunto de preguntas respecto a una o más variables a medir.*

¿Qué tipos de preguntas puede haber?

El contenido de las preguntas de un cuestionario puede ser tan variado como los aspectos que se midan a través de éste. Y básicamente, podemos hablar de *dos tipos de preguntas: “cerradas” y “abiertas”.*

Las *preguntas “cerradas”* contienen categorías o alternativas de respuesta que han sido delimitadas. Es decir, se presentan a los sujetos las posibilidades de respuesta y ellos deben circunscribirse a éstas. Las preguntas “cerradas” pueden ser dicotómicas (dos alternativas de respuesta) o incluir varias alternativas de respuesta. Ejemplos de preguntas cerradas dicotómicas serían: ¿Estudia usted actualmente?

- Si
- No

¿Durante la semana pasada vio la telenovela “Los Amantes”?

- Si
- No

Ejemplos de preguntas “cerradas” con varias alternativas de respuesta serían: ¿Cuánta televisión ves los domingos?

- No veo televisión
- Menos de una hora
- 1 o 2 horas
- 3 horas
- 4 horas
- 5 horas o más

¿Cuál es el puesto que ocupa en su empresa?

- Director General ; Presidente o Director
- Gerente/ Subdirector
- Subgerente / Superintendente
- Coordinador
- Jefe de área
- Supervisor
- Empleado
- Obrero
- Otro

Si usted tuviera una elección, ¿preferiría que su salario fuera de acuerdo con su productividad en el trabajo?

- () Definitivamente sí
() Probablemente sí
() No estoy seguro
() Probablemente no
() Definitivamente no

Como puede observarse, en las preguntas “cerradas” las categorías de respuesta son definidas a priori por el investigador y se le presentan al respondiente, quien debe elegir la opción que describa más adecuadamente su respuesta. Las escalas de actitudes en forma de pregunta caerían dentro de la categoría de preguntas “cerradas”.

Ahora bien, hay preguntas “*cerradas*”, donde el respondiente puede seleccionar más de una opción o categoría de respuesta.

EJEMPLO

Esta familia tiene:

- ¿Radio?
¿Televisión?
¿Videocasetera?
¿Teléfono?
¿Automóvil o camioneta?
Ninguno de los anteriores

Algunos respondientes pudieran marcar una, dos, tres, cuatro o cinco opciones de respuesta. Las categorías no son mutuamente excluyentes. Otro ejemplo sería la siguiente pregunta:

De los siguientes servicios que presta la biblioteca, ¿cuál o cuáles utilizaste el semestre anterior? (*Puede señalar más de una opción.*)

DE LA SALA DE LECTURA:

- No entré
A consultar algún libro
A consultar algún periódico
A estudiar
A pasar trabajos a máquina
A buscar a alguna persona
Otros, especifica

DE LA HEMEROTECA:

- No entré
A consultar algún periódico
A usar las videocassetteras
A estudiar
A hacer trabajos
A sacar fotocopias
A leer algún libro
Otros, especifica

DEL MOSTRADOR DE PRÉSTAMOS:

- No fui
- A solicitar un libro
- A solicitar una tesis
- A solicitar un periódico
- A solicitar grabositas
- A solicitar maquinas de escribir
- A solicitar equipo audiovisual
- A solicitar asesoría para la localización de material
- Otros, especifica

En otras ocasiones, el respondiente tiene que jerarquizar opciones. Por ejemplo: ¿cuál de los siguientes conductores de televisión considera usted el mejor?, ¿cuál en segundo lugar?, ¿cuál en tercer lugar?⁴⁴

⁴⁴ Conductores ficticios.

- LEM
- BCC
- MME

O bien debe asignar un puntaje a una o diversas cuestiones.

EJEMPLO

¿CUÁNTO LE INTERESA DESARROLLAR? (Indique de 1 a 10 en cada caso según sus intereses)

- _____ Administración de sueldos y compensaciones.
- _____ Salud, seguridad e higiene.
- _____ Administración y negociación de contratos.
- _____ Relaciones con sindicatos.
- _____ Habilidades de comunicación ejecutiva.
- _____ Programas y procesos sobre calidad/productividad.
- _____ Calidad de vida en el trabajo.
- _____ Teoría de la organización.
- _____ Administración financiera.
- _____ Desarrollo organizacional innovación.
- _____ Técnicas de investigación organizacional.
- _____ Estructura organizacional (tamaño, complejidad, formalización).
- _____ Sistemasde información y control.
- _____ Auditoria administrativa.
- _____ Planeación estratégica.
- _____ Sistemas de computación.
- _____ Mercadotecnia y comercialización.
- _____ Otros (especificar):

En cambio, *las preguntas “abiertas”* no delimitan de antemano las alternativas de respuesta. Por lo cual el número de categorías de respuesta es muy elevado. En teoría es infinito.

EJEMPLO

¿Por qué asiste a psicoterapia?

¿Qué opina del programa de televisión “Los Cazadores”?

¿De qué manera la directiva de la empresa ha logrado la cooperación del sindicato para el proyecto de calidad?

¿ Usamos preguntas cerradas o abiertas?

Cada cuestionario obedece a diferentes necesidades y problemas de investigación, lo que origina que en cada caso el tipo de preguntas a utilizar sea diferente. Algunas veces se incluyen solamente preguntas “cerradas”, otras veces únicamente preguntas “abiertas” y en ciertos casos ambos tipos de preguntas. *Cada clase de pregunta tiene sus ventajas y desventajas.* Las cuales se mencionan a continuación.

Las preguntas “cerradas” son fáciles de codificar y preparar para su análisis. Asimismo, estas preguntas requieren de un menor esfuerzo por parte de los respondientes. Éstos no tienen que escribir o verbalizar pensamientos, sino simplemente seleccionar la alternativa que describa mejor su respuesta. Responder a un cuestionario con preguntas cerradas toma menos tiempo que contestar a uno con preguntas abiertas. Si el cuestionario es enviado por correo, se tiene una mayor respuesta cuando es fácil de contestar y requiere menos tiempo completarlo. La *principal desventaja* de las preguntas “cerradas” reside en que *limitan las respuestas de la muestra* y —en ocasiones— ninguna de las categorías describe con exactitud lo que las personas tienen en mente, no siempre se captura lo que pasa por las cabezas de los sujetos.

Para poder formular preguntas “cerradas” es necesario anticipar las posibles alternativas de respuesta. De no ser así es muy difícil plantearlas. Asimismo, el investigador tiene que asegurarse que los sujetos a los cuales se les administrarán, conocen y comprenden las categorías de respuesta. Por ejemplo, si preguntamos qué canal de televisión es el preferido, determinar las opciones de respuesta y que los respondientes las comprendan es muy sencillo. Pero si preguntamos sobre las razones y motivos que provocan esa preferencia, determinar dichas opciones es algo bastante más complejo.

Las preguntas “abiertas” son particularmente útiles cuando no tenemos información sobre las posibles respuestas de las personas o cuando esta información es insuficiente. También sirven en situaciones donde se desea profundizar una opinión o los motivos de un comportamiento. *Su mayor desventaja* es que son más difíciles de codificar, clasificar y preparar para su análisis. Además, pueden presentarse sesgos derivados de distintas fuentes: por ejemplo, quienes tienen dificultades para expresarse oralmente y por escrito pueden no responder con precisión lo que realmente desean o generar confusión en sus respuestas. El nivel educativo, la capacidad de manejo del lenguaje y otros factores pueden afectar la calidad de las respuestas (Black y Champin, 1976). Asimismo, responder a preguntas “abiertas” requiere de un mayor esfuerzo y tiempo.

La elección del tipo de preguntas que contenga el cuestionario depende del grado en que se puedan anticipar las posibles respuestas, los tiempos de que se disponga para codificar y si se quiere una respuesta más precisa o profundizar en alguna cuestión. Una recomendación para construir un cuestionario es que se analice variable por variable qué tipo de pregunta o preguntas pueden ser más confiables y válidas para medir a esa variable, de acuerdo con la situación del estudio (planteamiento del problema, características de la

muestra, análisis que se piensan efectuar, etcétera).

¿Una o varias preguntas para medir una variable?

En ocasiones sólo basta una pregunta para recolectar la información necesaria sobre la variable a medir. Por ejemplo, para medir el nivel de escolaridad de una muestra, basta con preguntar: ¿Hasta qué año escolar cursó? o ¿cuál es su grado máximo de estudios? En otras ocasiones es necesario elaborar varias preguntas para verificar la consistencia de las respuestas. Por ejemplo, el nivel económico puede medirse preguntando: Aproximadamente ¿cuál es su nivel mensual de ingresos? y preguntando: Aproximadamente, ¿cuántos focos eléctricos tiene en su casa?⁴⁵ Además de preguntar sobre propiedades, inversiones, puesto que ocupa la fuente principal de ingresos de la familia (generalmente, el padre), etcétera.

⁴⁵ En varios estudios se ha demostrado que el nivel de ingresos está relacionado con el número de focos de una casa/habitación. El número de focos está vinculado con el número de cuartos de la casa, extensión de ésta, presencia de focos en el jardín de la casa, candiles y otros factores.

Al respecto, *es recomendable hacer solamente las preguntas necesarias* para obtener la información deseada o medir la variable. Si una pregunta es suficiente no es necesario incluir más, no tiene sentido. Si se justifica hacer varias preguntas, entonces es conveniente plantearlas en el cuestionario. Esto último ocurre con frecuencia en el caso de variables con varias dimensiones o componentes a medir, en donde se incluyen varias preguntas para medir las distintas dimensiones. Se tienen varios indicadores.

EJEMPLO

La empresa Comunicometría, S. C., realizó una investigación para la Fundación Mexicana para la Calidad Total, A.C. (1988), con el propósito de conocer las prácticas, técnicas, estructuras, procesos y temáticas existentes en materia de Calidad Total en México. El estudio fue de carácter exploratorio y constituyó el primer esfuerzo por obtener una radiografía del estado de los procesos de calidad en dicho país.

En esta investigación se elaboró un cuestionario que medía el grado en que las organizaciones mexicanas aplicaban diversas prácticas tendientes a elevar la calidad, la productividad y la calidad de vida en el trabajo. Una de las variables importantes era el “grado en que se distribuía la información sobre el proceso de calidad en la organización”. Esta variable se midió a través de las siguientes preguntas:

A. Por lo que respecta a los programas de información sobre calidad, ¿cuáles de las siguientes actividades se efectúan en esta empresa?

- (1) Planeación del manejo de datos sobre calidad.
- (2) Formas de control.
- (3) Elaboración de reportes con datos sobre calidad.
- (4) Evaluación sistemática de los datos sobre calidad.
- (5) Distribución generalizada de información sobre calidad.
- (6) Sistemas de autocontrol de calidad.
- (7) Distribución selectiva de datos sobre calidad.

B. Sólo a quienes distribuyen selectivamente datos sobre calidad: ¿A qué niveles de la empresa?

C. Sólo a quienes distribuyen selectivamente datos sobre calidad: ¿A qué funciones?

1

0

D. ¿Qué otras actividades se realizan en esta empresa para los programas de información sobre calidad? _____

En este ejemplo, las preguntas “B” y “C” se elaboraron para ahondar sobre los receptores o usuarios de los datos en aspectos del control de calidad distribuidos selectivamente. Se justifica el hacer estas dos preguntas, ayuda a tener mayor información sobre la variable. Cuando se tienen varias preguntas para una misma variable se dice que se tiene una “batería de preguntas”.

¿Las preguntas van precodificadas o no?

Siempre que se pretendan efectuar análisis estadísticos es necesario codificar las respuestas de los sujetos a las preguntas del cuestionario, y debemos recordar que esto significa asignarles símbolos o valores numéricos a dichas respuestas. Ahora bien, cuando se tienen preguntas “cerradas”, es posible codificar “a priori” o precodificar las alternativas de respuesta e incluir esta precodificación en el cuestionario (tal y como lo hacíamos con las escalas de actitudes).

EJEMPLOS DE PREGUNTAS PRECODIFICADAS

¿Tiene usted inversiones en la Bolsa de Valores?

Sí

No

Cuando se enfrenta usted a un problema en su trabajo, para resolverlo recurre generalmente a:

- (1) Su superior inmediato
- (2) Su propia experiencia
- (3) Sus compañeros
- (4) Los manuales de políticas y procedimientos
- (5) Otra fuente _____
(especificar)

En ambas preguntas, las respuestas van acompañadas de su valor numérico correspondiente, han sido precodificadas. Obviamente *en las preguntas “abiertas” no puede darse la precodificación*, la codificación se realiza posteriormente, una vez que se tienen las respuestas. Las preguntas y alternativas de respuesta precodificadas tienen la ventaja que su codificación y preparación para el análisis son más sencillas y requieren de menos tiempo.

¿Qué características debe tener una pregunta?

Independientemente de que las preguntas sean abiertas o cerradas y sus respuestas estén precodificadas o no, hay una serie de características que deben cubrirse al plantearlas:

A. *Las preguntas deben ser claras y comprensibles para los respondientes. Deben evitarse términos confusos o ambiguos y como menciona Rojas (1981, p.138) no es nada recomendable sacrificar la claridad por concisión. Es indispensable incluir las palabras que sean necesarias para que se comprenda la pregunta. Desde luego, sin ser repetitivos o barrocos. Por ejemplo, la pregunta: ¿ve usted televisión? es confusa, no delimita cada cuánto. Sería mucho mejor especificar: ¿acostumbra usted ver televisión diariamente? o ¿cuántos días durante la última semana vio televisión? y después preguntar los horarios, canales y contenidos de los programas.*

B. *Las preguntas no deben incomodar al respondiente.* Preguntas como: ¿acostumbra consumir algún tipo de bebida alcohólica?, tienden a provocar rechazo. Es mejor preguntar: ¿algunos de sus amigos acostumbran consumir algún tipo de bebida alcohólica? y después utilizar preguntas sutiles que indirectamente nos indiquen si la persona acostumbra consumir bebidas alcohólicas (v.g., ¿cuál es su tipo de bebida favorita?, etcétera). Y hay temáticas en donde a pesar de que se utilicen preguntas sutiles, el respondiente se sentirá molesto. En estos casos, pueden utilizarse escalas de actitud en lugar de preguntas o aún otras formas de medición. Tal es el caso de temas como el homosexualismo, la prostitución, la pornografía, los anticonceptivos y la drogadicción.

C. *Las preguntas deben preferentemente referirse a un sólo aspecto o relación lógica.* Por ejemplo, la pregunta: ¿acostumbra usted ver televisión y escuchar radio diariamente?, expresa dos aspectos, puede confundir. Es mucho mejor dividirla en dos preguntas, una relacionada con la televisión y otra relacionada con la radio.

D. *Las preguntas no deben inducir las respuestas* (Rojas, 1981, p. 138). Preguntas tendenciosas o que dan pie a elegir un tipo de respuesta deben evitarse. Por ejemplo: ¿considera usted a Ricardo Hernández el mejor candidato para dirigir nuestro sindicato?, es una pregunta tendenciosa, induce la respuesta. Lo mismo que la pregunta: ¿los trabajadores mexicanos son muy productivos? Se insinúa la respuesta en la pregunta. Resultaría mucho más conveniente preguntar:

¿Qué tan productivos considera usted —en general— a los trabajadores mexicanos?

Sumamente productivos	Más bien productivos	Más bien improductivos	Sumamente improductivos
-----------------------	----------------------	------------------------	-------------------------

E. *Las preguntas no pueden apoyarse en instituciones, ideas respaldadas social-mente ni en evidencia comprobada.* Es también una manera de inducir respuestas. Por ejemplo, la pregunta: La Organización Mundial de la Salud ha realizado diversos estudios y concluyó que el tabaquismo provoca diversos daños al organismo, ¿usted considera que fumar es nocivo para su salud? Esquemas del tipo: “La mayoría de las personas opinan que...”, “La Iglesia considera...”, “Los padres de familia piensan que...”, etcétera, no deben anteceder a una pregunta, sesgan las respuestas.

E. *En las preguntas con varias alternativas o categorías de respuesta y donde el respondiente sólo tiene que elegir una, puede ocurrir que el orden en que se presenten dichas alternativas afecte las respuestas de los sujetos* (v.g., tiendan a favorecer a la primera o a la última alternativa de respuesta). Entonces resulta conveniente rotar el orden de lectura de las alternativas de manera proporcional. Por ejemplo, si preguntamos: ¿cuál de los siguientes tres candidatos presidenciales considera usted que logrará disminuir verdaderamente la inflación? Y el 33.33% de las veces que se haga la pregunta se menciona primero al candidato “A”, el 33.33% se menciona primero al candidato “B” y el restante 33.33% al candidato “C”.

G. *El lenguaje utilizado en las preguntas debe ser adaptado a las características del respondiente* (tomar en cuenta su nivel educativo, socioeconómico, palabras que maneja, etcétera). Este aspecto es igual al que se comentó sobre las escalas de actitudes.

¿Cómo deben ser las primeras preguntas de un cuestionario?

En algunos casos es conveniente iniciar con preguntas neutrales o fáciles de contestar, para que el respondiente vaya adentrándose en la situación. *No se recomienda comenzar con preguntas difíciles de responder o preguntas muy directas.* Imaginemos un cuestionario diseñado para obtener opiniones en torno al aborto que empiece con una pregunta poco sutil tal como: ¿Está usted de acuerdo en que se legalice el aborto en este país? Sin lugar a dudas será un fracaso.

A veces los cuestionarios pueden comenzar con preguntas demográficas sobre el estado civil, sexo, edad, ocupación, nivel de ingresos, nivel educativo, religión, ideología, puesto en una organización o algún tipo de

afiliación a un grupo, partido e institución. Pero en otras ocasiones es mejor hacer este tipo de preguntas al final del cuestionario, particularmente cuando los sujetos puedan sentir que se comprometen al responder al cuestionario. Cuando construimos un cuestionario es indispensable que pensemos en cuáles son las preguntas ideales para iniciar. Éstas deberán lograr que el respondiente se concentre en el cuestionario.

¿De qué está formado un Cuestionario?

Además de las *preguntas y categorías de respuestas*, un cuestionario está formado por *instrucciones* que nos indican cómo contestar, por ejemplo:

Hablando de la mayoría de sus proveedores en qué medida conoce usted (MOSTRAR TARJETA UNO Y MARCAR LA RESPUESTA EN CADA CASO):

Comp.

(5) Bast.

(4) Reg.

(3) Poco

(2) Nada

- (1) •¿Las políticas de su proveedor?
- ¿Sus finanzas (estado financiero)?
- ¿Los objetivos de su área de ventas?
- ¿Sus programas de capacitación para vendedores?
- ¿Número de empleados de su área de ventas?
- ¿Problemas laborales?
- ¿Los métodos de producción que tienen?
- ¿Otros clientes de ellos?
- ¿Su índice de rotación personal?

¿Tiene este ejido o comunidad, ganado, aves o colmenas que sean de propiedad colectiva? (CIRCULE LA RESPUESTA)

Si 1
(continúe)

No 2
(pase a 30)

¿Se ha obtenido la cooperación de todo el personal o la mayoría de éste para el proyecto de calidad?

1 Si
(pase a la pregunta 26)

2 No
(pase a la pregunta 27)

Las instrucciones son tan importantes como las preguntas y es necesario que sean claras para los usuarios a quienes van dirigidas. Y una instrucción muy importante es agradecer al respondiente por haberse tomado el tiempo de contestar el cuestionario. También, es frecuente incluir una carátula de presentación o una carta donde se expliquen los propósitos del cuestionario y se garantice la confidencialidad de la información, esto ayuda a ganar la confianza del respondiente. A continuación en la figura 9.16 se presentan algunos textos ilustrativos de cartas introductorias a un cuestionario.

FIGURA 9.16
EJEMPLOS DE CARTAS

Buenos días (tardes):

Estamos trabajando en un estudio que servirá para elaborar una tesis profesional acerca de la Biblioteca de la Universidad Anáhuac.

Quisiéramos pedir tu ayuda para que contestes a unas preguntas que no llevarán mucho tiempo. Tus respuestas serán confidenciales y anónimas.

Las personas que fueron seleccionadas para el estudio no se eligieron por su nombre sino al azar.

Las opiniones de todos los encuestados serán sumadas y reportadas en la tesis profesional, pero nunca se reportarán datos individuales.

Te pedimos que contestes este cuestionario con la mayor sinceridad posible. No hay respuestas correctas ni incorrectas.

Lee las instrucciones cuidadosamente, ya que existen preguntas en las que sólo pueden responder a una opción; otras son de varias opciones y también se incluyen preguntas abiertas.

Muchas gracias por tu colaboración.

BUENOS DÍAS (TARDES)

COMUNICOMETRÍA ESTÁ HACIENDO UNA ENCUESTA CON EL PROPÓSITO DE CONOCER UNA SERIE DE OPINIONES QUE SE TIENEN ACERCA DE ESTA EMPRESA, Y PARA ELLO LE PEDIRÍA FUERA TAN AMABLE DE CONTESTAR UNAS PREGUNTAS. NO LE TOMARÁ MÁS DE 20 MINUTOS. LA INFORMACIÓN QUE NOS PROPORCIONE SERÁ MANEJADA CON LA MÁS ESTRICTA CONFIDENCIALIDAD. DESDE LUEGO, NO HAY PREGUNTAS DELICADAS.

La manera en que pueden distribuirse las preguntas, categorías de respuestas e instrucciones es variada. Algunos prefieren colocar las preguntas a la izquierda y las respuestas a la derecha, con lo que se tendría un formato como el siguiente:

¿Cómo se llama a este producto?	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
¿Qué le gusta más de este producto?	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
¿Cuál es su opinión sobre el precio?	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
¿Recomendaría este producto a un amigo?	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
¿Hay algo más que le gustaría comentar?	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

Otros dividen el cuestionario por secciones de preguntas y utilizan un formato horizontal. Por ejemplo:

PRESENTACIÓN
PREGUNTAS SOBRE MOTIVACIÓN INTRÍNSECA:
¿ Es un _____ ? <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
¿ Es un _____ ? <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
PREGUNTAS SOBRE SATISFACCIÓN LABORAL
ETC.

O combinan diversas posibilidades, distribuyendo preguntas que miden la misma variable a través de todo el cuestionario. Cada quien puede utilizar el formato que desee o juzgue conveniente, lo importante es que sea totalmente comprensible para el usuario; las instrucciones, preguntas y respuestas se diferencien, no resulte visualmente tedioso y se pueda leer sin dificultades.

¿De qué tamaño debe ser un cuestionario?

No existe una regla al respecto, aunque como menciona Padua (1979), si es muy corto se pierde información y si resulta largo puede resultar tedioso de responder. En este último caso, las personas pueden ‘negarse a responder o —al menos— no completar el cuestionario. El tamaño depende del número de variables y dimensiones a medir, el interés de los respondientes y la manera como es administrado (de este punto se hablará en el siguiente inciso). Cuestionarios que duran más de 35 minutos pueden resultar tediosos a menos que los respondientes estén muy motivados para contestar (v.g., cuestionarios de personalidad, cuestionarios para obtener un trabajo). *Una recomendación* que puede ayudarnos para evitar un cuestionario más largo de lo requerido es: “No hace preguntas innecesarias o injustificadas”.

¿Cómo se codifican las preguntas abiertas?

Las preguntas abiertas se codifican una vez que conocemos todas las respuestas de los sujetos a las cuales se les aplicaron o al menos las principales tendencias de respuestas en una muestra de los cuestionarios aplicados.

El procedimiento consiste en encontrar y darle nombre a los patrones generales de respuesta (respuestas similares o comunes), *listar estos patrones y después asignar un valor numérico o símbolo a cada patrón*. Así, un patrón constituirá una categoría de respuesta. Para cerrar las preguntas abiertas, se sugiere el siguiente procedimiento, basado parcialmente en Rojas (1981, pp.1 50-151):

1. *Seleccionar determinado número de cuestionarios* mediante un método adecuado de muestreo, asegurando la representatividad de los sujetos investigados.
2. *Observar la frecuencia con que aparece cada respuesta* a la pregunta.
3. *Elegir las respuestas* que se presentan con mayor frecuencia (patrones generales de respuesta).
4. *Clasificar las respuestas elegidas* en temas, aspectos o rubros, de acuerdo con un criterio lógico, cuidando que sean mutuamente excluyentes.
5. *Darle un nombre o título a cada tema*, aspecto o rubro (patrón general de respuesta).
6. *Asignarle el código a cada patrón* general de respuesta.

Por ejemplo, en la investigación de Comunicometría (1988) se hizo una pregunta abierta:

¿De qué manera la alta gerencia busca obtener la cooperación del personal para el desarrollo del proyecto de

calidad?

Las respuestas fueron múltiples pero pudieron encontrarse los siguientes patrones generales de respuesta:

CÓDIGOS	CATEGORÍAS	(PATRONES	O				
RESPUESTAS		CON	MAYOR				
FRECUENCIA	DE	MENCIÓN) NÚM.	DE	FRECUENCIAS	DE	MENCIÓN	1
	Involucrando al personal y comunicándose con él.	28	2	Motivación e integración.	20		

3 Capacitación en general. 12 4 Incentivos ; recompensas. 11 5 Difundiendo el valor “calidad” o la filosofía de la empresa. 7 6 Grupos o sesiones de trabajo. 5 7 Posicionamiento del área de calidad o equivalente. 3 8 Sensibilización. 2 9 Desarrollo de la calidad de vida en el trabajo. 2 10 Incluir aspectos de calidad en el manual de inducción. 2 11 Enfatizar el cuidado de la maquinaria. 2 12 Trabajando bajo un buen clima laboral. 2 13 Capacitación “en cascada”. 2 14 Otras. 24

Como varias categorías o patrones tenían solamente dos frecuencias, éstos a su vez pudieron reducirse a:

CATEGORÍAS Involucrando al personal y comunicándose con él. Motivación e integración / mejoramiento del clima laboral. Capacitación. Incentivos / recompensas.
Difundiendo el valor “calidad” o la filosofía de la empresa. Grupos o sesiones de trabajo. Otras.

Al “cerrar” preguntas abiertas y ser codificadas, debe tenerse en cuenta que un mismo patrón de respuesta puede expresarse con diferentes palabras. Por ejemplo, ante la pregunta: ¿Qué sugerencias podría hacer para mejorar al programa “Estelar”? Las respuestas: “Mejorar las canciones y la música”, “cambiar las canciones”, “incluir nuevas y mejores canciones”, etc., pueden agruparse en la categoría o patrón de respuesta “modificar la musicalización del programa”.

¿En qué contextos puede administrarse o aplicarse un Cuestionario?

Los cuestionarios pueden ser aplicados de diversas maneras:

A) *Autoadministrado*. En este caso el cuestionario se les proporciona directamente a los respondientes, quienes lo contestan. No hay intermediarios y las respuestas las marcan ellos. Por ejemplo, si los respondientes fueran una muestra de los estudiantes de la Licenciatura en Comunicación de Bogotá, se acudiría a ellos y se les entregarían los cuestionarios. Los estudiantes se autoadministrarían el cuestionario. Obviamente que esta manera de aplicar el cuestionario es impropia para analfabetas, personas que tienen dificultades de lectura o niños que todavía no leen adecuadamente.

B) *Por entrevista personal*. En esta situación, un entrevistador aplica el cuestionario a los respondientes (entrevistados). El entrevistador va haciéndole las preguntas al respondiente y va anotando las respuestas. Las instrucciones son para el entrevistador. Normalmente se tienen varios entrevistadores, quienes deberán estar capacitados en el arte de entrevistar y conocer a fondo el cuestionario, y no deben sesgar o influir las respuestas.

C) *Por entrevista telefónica*. Esta situación es similar a la anterior, solamente que la entrevista no es “cara a cara” sino a través del teléfono. El entrevistador le hace las preguntas al respondiente por este medio de comunicación.

D) *Autoadministrado y enviado por correo postal; electrónico o servicio de mensajería*. En este caso también los respondientes contestan directamente el cuestionario, ellos marcan o anotan las respuestas, no hay intermediario. Solamente que no se entregan los cuestionarios directamente a los respondientes (“en propia mano”) sino que se les envían por correo u otro medio, no hay retroalimentación inmediata, si los sujetos tienen alguna duda no se les puede aclarar en el momento.

Consejos para la administración del cuestionario, dependiendo del contexto

Cuando se tiene población analfabeta, con niveles educativos bajos o niños que apenas comienzan a leer o no dominan la lectura, el método más conveniente de administración de un cuestionario es por entrevista. Aunque hoy en día ya existen algunos cuestionarios muy gráficos que usan escalas sencillas. Como por ejemplo:



En desacuerdo



Neutral



De acuerdo

Con *trabajadores* de niveles de lectura básica se recomienda utilizar entrevistas o cuestionarios autoadministrados sencillos que se apliquen en grupos con la asesoría de entrevistadores o supervisores capacitados.

En algunos casos, con *ejecutivos* que difícilmente pueden dedicarle a un solo asunto más de 20 minutos, se pueden utilizar cuestionarios autoadministrados o entrevistas telefónicas. Con estudiantes suelen funcionar los cuestionarios autoadministrados.

Asimismo, algunas asociaciones hacen encuestas por correo y ciertas empresas envían cuestionarios a sus ejecutivos y supervisores mediante el servicio interno de mensajería o por correo electrónico. Cuando el cuestionario contiene unas cuantas preguntas (su administración no toma más de 4 minutos —o máximo 5—), la entrevista telefónica es una buena alternativa.

Ahora bien, sea cual fuere la forma de administración, *siempre debe haber uno o varios supervisores que verifiquen que se están aplicando correctamente los cuestionarios*. Cuando un cuestionario o escala es aplicado(a) de forma masiva suele denominarse “*encuesta*”.

Cuando se utiliza la *entrevista telefónica* se debe tomar en cuenta el *horario*. Ya que si hablamos sólo a una hora (digamos en la mañana), nos encontraremos con unos cuantos subgrupos de la población (v.g., amas de casa).

Cuando *lo enviamos por correo o es autoadministrado* directamente, las instrucciones deben pecar de precisas, claras y completas. Y debemos dar instrucciones que motiven al respondiente para que continúe contestando el cuestionario (v.g., ya nada más unas cuantas preguntas, finalmente).

Asimismo, cabe señalar que cuando se trata de *entrevista personal*, el lugar donde se realice es importante (oficina, casa-habitación, en la calle, etc.). Por ejemplo, Jaffe, Pasternak y Grifel (1983) realizaron un estudio para comparar —entre otros aspectos— las respuestas obtenidas en dos puntos diferentes: en el hogar y en puntos de venta. El estudio se interesaba en la conducta del comprador y los resultados concluyeron que se pueden obtener datos exactos en ambos puntos, pero la entrevista en los puntos de compra-venta es menos costosa.

Las *entrevistas personales* requieren de una *atmósfera apropiada*. El entrevistador debe ser amable y tiene

que generar confianza en el entrevistado. Cuando se trata de entrevistados del sexo masculino, mujeres simpáticas y agradables suelen resultar excelentes entrevistadores. Quien responde a una entrevista debe concentrarse en las preguntas y estar relajado. Y después de una entrevista debe prepararse un informe que indique si el sujeto se mostraba sincero, la manera como respondió, el tiempo que duró la entrevista, el lugar donde se realizó, las características del entrevistado, los contratiempos que se presentaron y la manera como se desarrolló la entrevista, así como otros aspectos que se consideren relevantes.

La elección del contexto para administrar el cuestionario deberá ser muy cuidadosa y dependerá del presupuesto de que se disponga, el tiempo de entrega de los resultados, los objetivos de la investigación y el tipo de respondientes (edad, nivel educativo, etc.).

Estas maneras de aplicar un cuestionario pueden hacerse extensivas a las escalas de actitudes, sólo que es mucho más difícil en el caso de que se administren por teléfono (pocas frases y alternativas claras de respuesta).

Cuando los cuestionarios son muy complejos de contestar o aplicar, suele utilizarse un manual que explica a fondo las instrucciones y cómo debe responderse o ser administrado.

¿Cuál es el proceso para construir un cuestionario?

Siguiendo los pasos para construir un instrumento de medición tendríamos la figura 9.17.

Un aspecto muy importante que es necesario mencionar, reside en que cuando se construye un cuestionario —al igual que otros instrumentos de medición— se debe ser consistente en todos los aspectos. Por ejemplo, si se decide que las instrucciones vayan en mayúsculas o algún tipo de letra especial, todas las instrucciones deberán ser así. Si se prefiere que los códigos de las categorías de respuesta van en recuadro, todas deberán ajustarse a esto. Si no se es consistente, algunos respondientes o entrevistadores pueden desconcertarse.

9.6.3. Análisis del contenido

¿QUÉ ES Y PARA QUÉ SIRVE EL ANÁLISIS DE CONTENIDO?

De acuerdo con la definición clásica de Berelson (1952), *el análisis de contenido es una técnica para estudiar y analizar la comunicación de una manera objetiva, sistemática y cuantitativa*. Krippendorff (1982) extiende la definición del análisis de contenido a una *técnica de investigación para hacer inferencias válidas y confiables de datos con respecto a su contexto*. Algunos autores consideran al análisis de contenido como un diseño. Pero más allá de como lo definamos, es una técnica muy útil para analizar los procesos de comunicación en muy diversos contextos. El *análisis de contenido* puede ser aplicado virtualmente a cualquier forma de comunicación (programas televisivos o radiofónicos, artículos en prensa, libros, poemas, conversaciones, pinturas, discursos, cartas, melodías, reglamentos, etcétera). Por ejemplo, puede servir para analizar la personalidad de alguien, evaluando sus escritos; conocer las actitudes de un grupo de personas mediante el análisis de sus discursos; indagar sobre las preocupaciones de un pintor o un músico; compenetrarse con los valores de una cultura; o averiguar las intenciones de un publicista o propagandista.

USOS DEL ANÁLISIS DE CONTENIDO

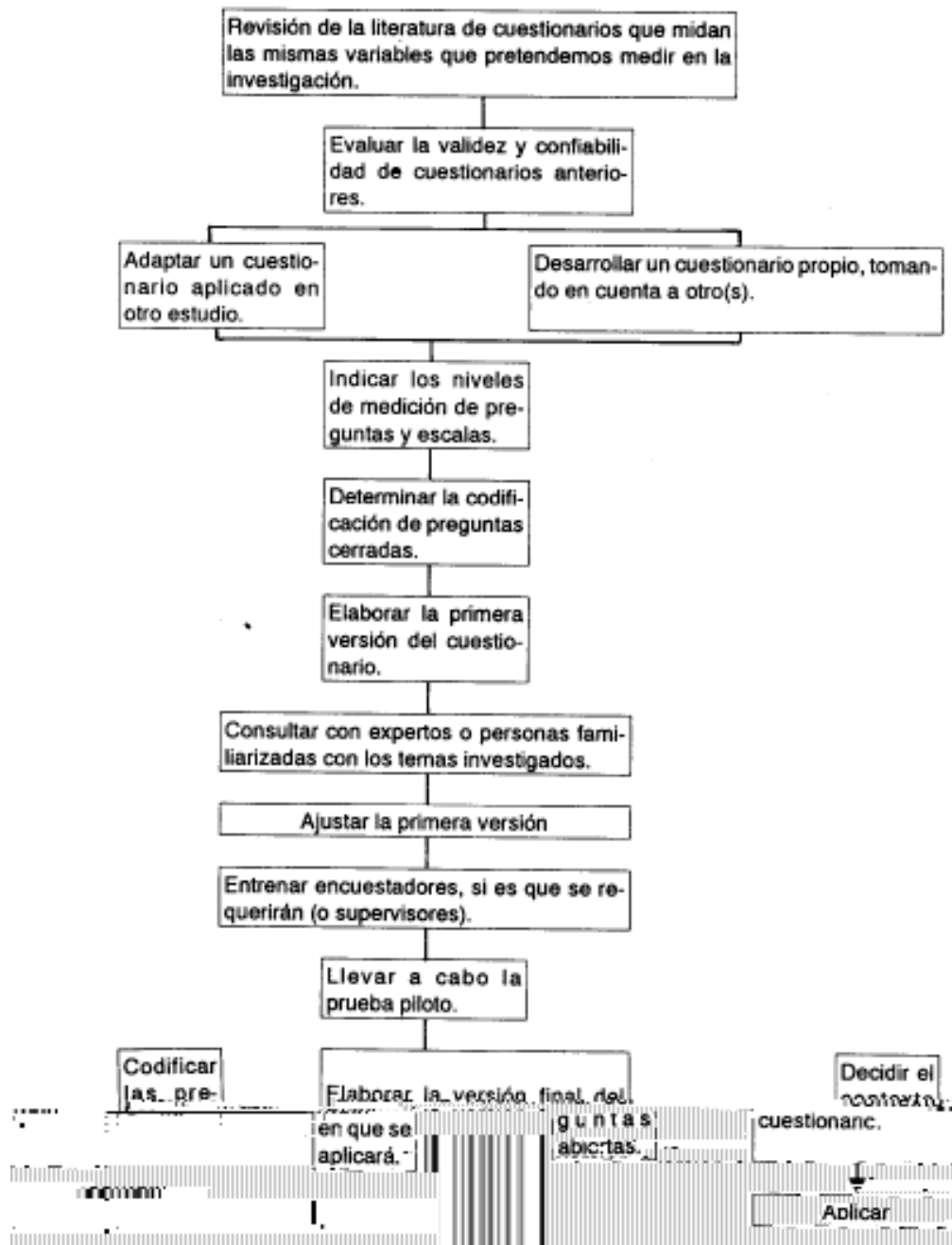
Berelson (1952) señala varios usos del análisis de contenido, entre los que destacan:

- 1) *Describir tendencias* en el contenido de la comunicación.
- 2) *Develar inferencias* en el contenido de la comunicación (entre personas, grupos, instituciones, países).
- 3) *Comparar mensajes*, niveles y medios de comunicación.
- 4) *Auditar el contenido* de la comunicación y compararlo contra estándares u objetivos.
- 5) *Construir y aplicar estándares de comunicación* (políticas, normas, etc.).

- 6) *Exponer técnicas publicitarias y de propaganda.*
- 7) *Medir la claridad de mensajes.*
- 8) *Descubrir estilos de comunicación.*
- 9) *Identificar intenciones, apelaciones y características de comunicadores.*
- 10) *Descifrar mensajes ocultos y otras aplicaciones a la inteligencia militar y a la seguridad política.*
- 11) *Revelar “centros” de interés y atención para una persona, un grupo y una comunidad.*
- 12) *Determinar el estado psicológico de personas o grupos.*
- 13) *Obtener indicios del desarrollo verbal (v.g., en la escuela, como resultado de la capacitación, el aprendizaje de conceptos).*
- 14) *Anticipar respuestas a comunicaciones.*
- 15) *Reflejar actitudes, valores y creencias de personas, grupos o comunidades.*
- 16) *Cerrar preguntas abiertas.*

FIGURA 9.17

PROCESO PARA CONSTRUIR UN CUESTIONARIO



El *análisis de contenido* puede utilizarse para ver si varias telenovelas difieren entre sí en cuanto a su carga de contenido sexual, para conocer las diferencias ideológicas entre varios periódicos (en términos generales o en referencia a un tema en particular), para comparar estrategias propagandísticas de partidos políticos contendientes, para contrastar a través de sus escritos a diferentes grupos que asisten a psicoterapia, para comparar el vocabulario aprendido por niños que se exponen a mayor contenido televisivo en relación con niños que ven menos televisión, para analizar la evolución de las estrategias publicitarias a través de algún medio respecto a un producto (v.g., perfumes femeninos de costo elevado); para conocer y comparar la posición de diversos presidentes latinoamericanos respecto al problema de la deuda externa; para comparar estilos de escritores que se señalan como parte de una misma corriente literaria; etcétera.

¿CÓMO SE REALIZA EL ANÁLISIS DE CONTENIDO?

El *análisis de contenido* se efectúa por medio de la *codificación*, que es *el proceso a través del cual las características relevantes del contenido de un mensaje son transformadas a unidades* que permitan su descripción y análisis preciso. Lo importante del mensaje se convierte en algo susceptible de describir y analizar. Para poder codificar es necesario definir el *universo* a analizar, las *unidades de análisis* y las *categorías de análisis*.

UNIVERSO

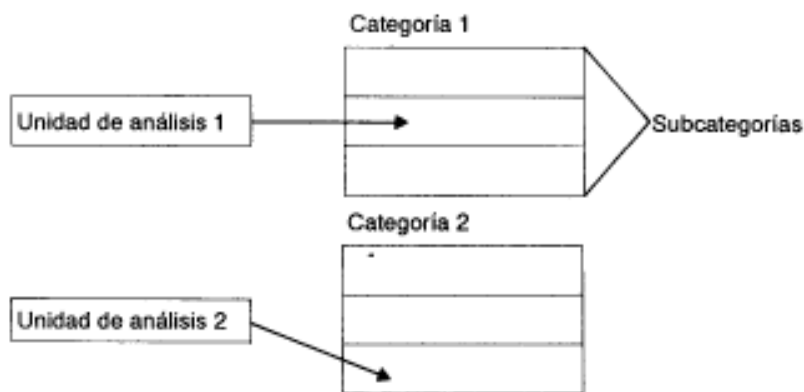
El universo podría ser la obra completa de Franz Kafka, las emisiones de un noticiario televisivo durante un mes, los editoriales publicados en un día por cinco periódicos de una determinada ciudad, todos los capítulos de tres telenovelas, los discos de Janis Joplin, Jimi Hendrix y Bob Dylan, los escritos de un grupo de estudiantes durante un ciclo escolar, los discursos pronunciados por varios contendientes políticos durante el último mes previo a la elección, escritos de un grupo de pacientes en psicoterapia, las conversaciones grabadas de 10 parejas que participan en un experimento sobre interacción matrimonial. El universo, como en cualquier investigación, debe delimitarse con precisión.

UNIDADES DE ANÁLISIS

Las unidades de análisis constituyen segmentos del contenido de los mensajes que son caracterizados para ubicarlos dentro de las categorías. Berelson (1952) menciona cinco unidades importantes de análisis:

- 1) *La palabra*. Es la unidad de análisis más simple, aunque como señala Kerlinger (1975), puede haber unidades más pequeñas como letras, fonemas o símbolos. Así, se puede medir cuántas veces aparece una palabra en un mensaje (v.g., veces que en un programa televisivo de fin de año se menciona al Presidente).
- 2) *El tema*. Éste se define a menudo como una oración, un enunciado respecto a algo. Los temas pueden ser más o menos generales. Kerlinger (1975, p. 552) utiliza un excelente ejemplo para ello: “Las cartas de adolescentes o estudiantes de colegios superiores pueden ser estudiadas en sus expresiones de autoreferencia. Este sería el tema más extenso. Los temas que constituyen éste podrían definirse como cualesquiera oraciones que usen “yo”, “mí” y otros términos que indiquen referencia al yo del escritor. Así, se analizaría qué tanta autoreferencia esté presente en dichas cartas.
Si los temas son complejos, el análisis del contenido es más difícil, especialmente si se complica al incluirse más de una oración simple.
- 3) *El ítem*. Tal vez es la unidad de análisis más utilizada y puede definirse como la unidad total empleada por los productores del material simbólico (Berelson 1952). Ejemplos de ítems pueden ser un libro, una editorial, un programa de radio o televisión, un discurso, una ley, un comercial, una carta amorosa, una conversación telefónica, una canción o la respuesta a una pregunta abierta. En este caso lo que se analiza es el material simbólico total.
- 4) *El personaje*. Un individuo, un personaje televisivo, un líder histórico, etcétera. Aquí lo que se analiza es el personaje.
- 5) *Medidas de espacio-tiempo*. Son unidades físicas como el centímetro-columna (por ejemplo, en la prensa), la línea (en escritos), el minuto (en una conversación o en radio), el periodo de 20 minutos (en una interacción), el cuadro (en televisión), cada vez que se haga una pausa (en un discurso).

Estas unidades se enclavan, colocan o caracterizan en categorías, esto podría representarse de la siguiente manera:

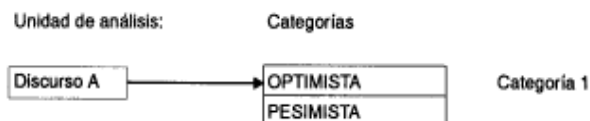


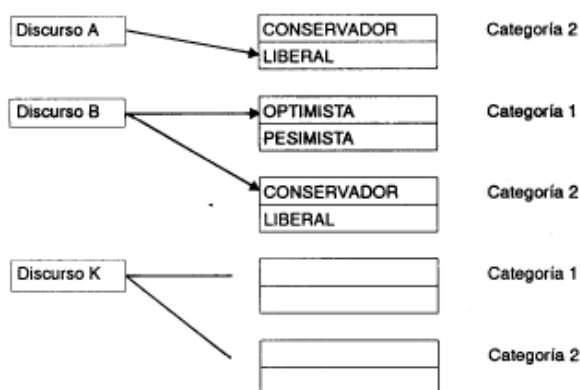
Respecto a la pregunta: ¿qué unidad debe seleccionarse? Esto depende de los objetivos y preguntas de investigación. Sin embargo, Berelson (1952) sugiere lo siguiente:

- a) En un solo estudio se pueden utilizar *más de una unidad de análisis*.
- b) Los cálculos de palabras y las unidades amplias, como el ítem y las medidas de espacio-tiempo, son más adecuadas en los análisis que dan énfasis a asuntos definidos.
- c) Las *unidades* amplias y las más definidas son válidas para la aceptación o rechazo en una categoría.
- d) Las unidades amplias generalmente requieren de menos tiempo para su codificación que las unidades pequeñas, referidas a las mismas categorías y materiales.
- e) Debido a que los temas u oraciones agregan otra dimensión al asunto, la mayoría de las veces son más difíciles de analizar que las palabras y las *unidades* amplias.
- f) El tema es adecuado para análisis de significados y las relaciones entre éstos.

CATEGORÍAS

Las *categorías* son los niveles donde serán caracterizadas las unidades de análisis. Tal y como menciona Holsti (1968), son las “*casillas o cajones*” en las cuales son clasificadas las unidades de análisis. Por ejemplo, un discurso podría clasificarse como optimista o pesimista, como liberal o conservador. Un personaje de una caricatura puede clasificarse como bueno, neutral o malo. En ambos casos, la unidad de análisis es categorizada. Veámoslo esquemáticamente:





Es decir, cada unidad de análisis es categorizada o encasillada en uno o más sistemas de categorías. Por ejemplo, en un estudio citado por Krippendorff (1982) se analizaron 2 430 actos de violencia televisada, en cada acto el personaje principal (unidad de análisis) era categorizado como:

- Bueno, neutral o malo (sistema 1).
- Asociado con hacer cumplir la ley, no era relacionado con la ley o era presentado como un delincuente o criminal (sistema 2).

En este caso tenemos que la *unidad de análisis* es el comportamiento del personaje durante el acto televisivo, y las *categorías* eran dos: bondad del personaje y carácter involucrado. A su vez, las subcategorías de la bondad del personaje eran tres: bueno, neutral y malo. Y las subcategorías del carácter involucrado también eran tres: asociado con cumplir la ley, no relacionado con la ley y un criminal. Esto podría representarse así:



La selección de categorías también depende del planteamiento del problema.

TIPOS DE CATEGORÍAS

Krippendorff (1982) señala *cinco tipos de categorías*:

- 1) *De asunto o tópico*. Las cuales se refieren a cuál es el asunto, tópico o tema tratado en el contenido (¿de qué trata el mensaje o la comunicación?).

EJEMPLO

- Analizar el último informe del Secretario o Ministro de Hacienda o Finanzas.

Categoría: Tema financiero

Subcategorías: Deuda

Impuestos
Planeación hacendaria
Inflación
Etc.

2) *De dirección.* Estas categorías se refieren a cómo es tratado el asunto (¿positiva o negativamente?, ¿favorable o desfavorable?, ¿nacionalista o no nacionalista?, etc.). Por ejemplo:

Comparar la manera como dos noticiarios televisivos hablan de la posibilidad de una moratoria unilateral en el pago de la deuda externa de Latinoamérica.

Categoría: Tono en el tratamiento de la deuda externa.

Subcategorías: A favor de la moratoria unilateral.

En contra

Neutral

3) *De valores.* Se refieren a categorías que indican qué valores, intereses, metas, deseos o creencias son revelados. Por ejemplo:

- Al estudiar la compatibilidad ideológica de matrimonios, se podría analizar la ideología de cada cónyuge pidiéndoles un escrito sobre temas que puedan reflejar valores (sexo, actitud hacia la pareja, significado del matrimonio).

Categoría: Ideología del esposo.

Subcategorías: Muy tradicional

Más bien tradicional

Neutral

Más bien liberal

Muy liberal

4) *De receptores.* Estas categorías se relacionan con el destinatario de la comunicación (¿a quién van dirigidos los mensajes?).

EJEMPLO

- Analizar a quiénes se dirige más un líder sindical en sus declaraciones a los medios de comunicación durante un periodo determinado.

Categoría: Receptores a quienes dirige el mensaje.

Subcategorías: Opinión pública en general

Presidente

Gabinete económico

Gabinete agropecuario

Gobierno en general

Sector empresarial

Obreros afiliados a su sindicato

Obreros no afiliados a su sindicato

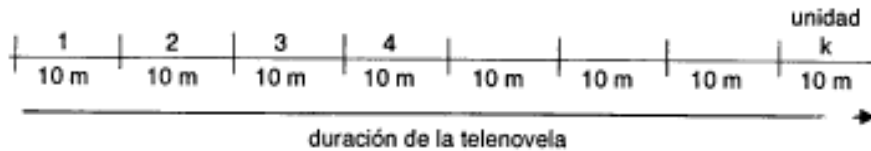
Obreros en general (afiliados y no afiliados)

Etc.

5) *Físicas.* Son categorías para ubicar la posición y duración o extensión de una unidad de análisis. De posición pueden ser por ejemplo la sección y página (en prensa), el horario (en televisión y radio). De duración, los minutos (en una interacción, un comercial televisivo, un programa de radio, un discurso), los centímetros / columna (en prensa), los cuadros en una película, etc. No se debe confundir las medidas de espacio-tiempo con las categorías físicas. Las primeras son unidades de análisis, las segundas constituyen categorías.

EJEMPLO

Cada periodo de 10 minutos de una telenovela lo voy a considerar la unidad de análisis.



La unidad 1 la coloco en categorías.

La unidad 2 la coloco en categorías.

La unidad k la coloco en categorías.

Los minutos (cada 10) los considero una unidad.

Pero puedo tener categorías de tiempo. Por ejemplo, para analizar comerciales en radio:

Categoría: Duración del comercial.

Subcategorías: 10 segundos o menos

11-20 segundos

21-30 segundos

Más de 30 segundos.

En ambos casos se usa el tiempo, pero en el primero como unidad y en el segundo como categoría. Es muy diferente.

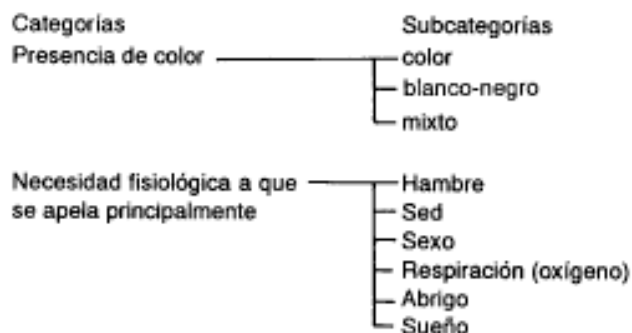
REQUISITOS DE LAS CATEGORÍAS

En un análisis de contenido se suelen tener varias categorías, pero éstas deben cumplir con los siguientes *requisitos*:

1) Las categorías y subcategorías deben ser *exhaustivas*. Es decir, abarcar todas las posibles subcategorías de lo que se va a codificar. Por ejemplo, la categoría “Ideología del esposo” no podría prescindir de la subcategoría “neutral”.

2) Las subcategorías deben ser *mutuamente excluyentes*, de tal manera que una unidad de análisis puede caer en una y sólo una de las subcategorías de cada categoría. Por ejemplo, un personaje no puede ser “bueno” y “malo” a la vez.

Aunque con respecto a las *categorías* no siempre son mutuamente excluyentes. Por ejemplo, al analizar comerciales televisivos podríamos tener —entre otras categorías— las siguientes:



Una unidad de análisis (un comercial) puede caer en una subcategoría de “Presencia de color” y en otra

subcategoría de la categoría “Necesidad fisiológica a que se apela” (color y sed). Pero no puede caer en dos subcategorías de la misma categoría “hambre” y “sed”, a menos que generáramos la subcategoría “hambre y sed”:

Hambre y sed

Desde luego, en ciertos casos especiales, puede interesar al analista de contenido un sistema de categorías donde éstas no sean mutuamente excluyentes. Pero no es lo normal.

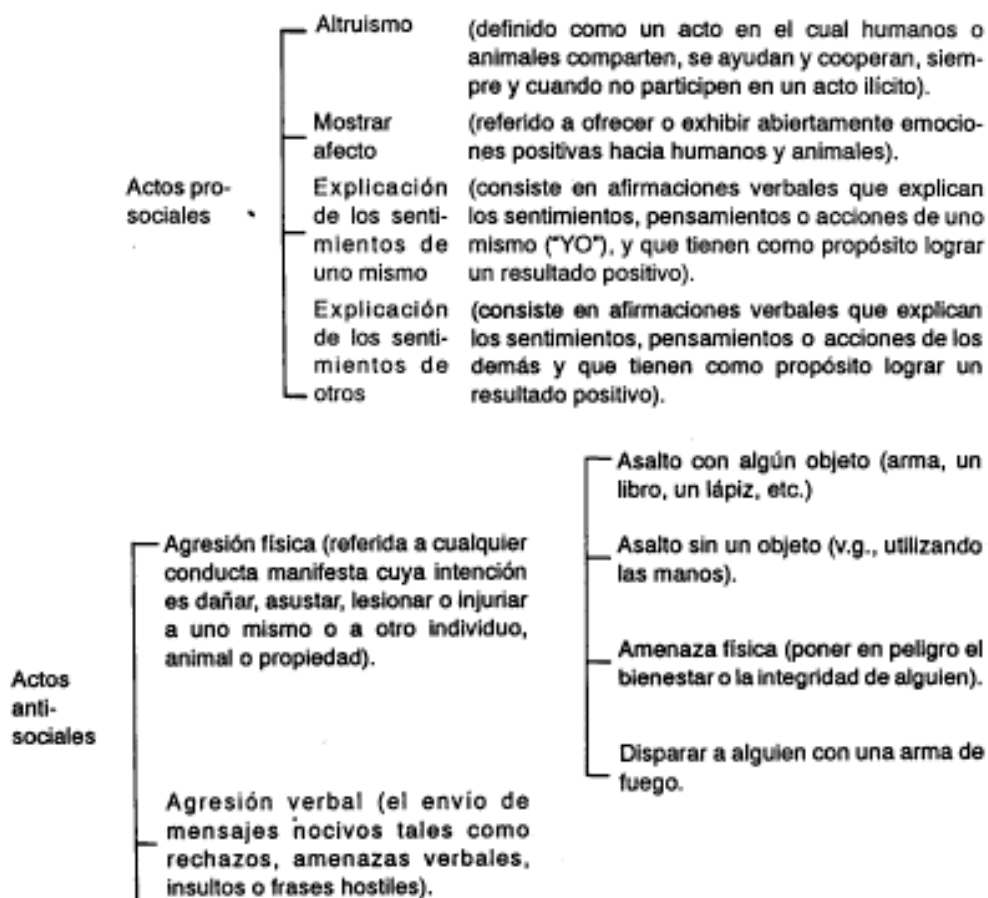
3) Las categorías y subcategorías *deben derivarse del marco teórico y una profunda evaluación de la situación.*

EJEMPLO DE UN ANÁLISIS DE CONTENIDO

Para ejemplificar el análisis de contenido y específicamente la generación de categorías acudiremos a un estudio de Greenberg, Edison, Korzenny, Fernández-Collado y Atkin (1980). El estudio consistió en un análisis de contenido de las series televisadas por las tres grandes cadenas de los Estados Unidos: ABC, CBS y NBC. Se analizaron diversos programas durante tres periodos de 1975 a 1978, para evaluar el grado en que la televisión norteamericana mostraba actos prosociales y antisociales como medida de la violencia televisada. Las categorías y subcategorías eran las siguientes:

FIGURA 9.18

EJEMPLO DE CATEGORÍAS



Robo (la acción de tomar intencional y deliberadamente una propiedad que pertenece a otra persona o institución, sin derecho o permiso).
 Engaño (engaño intencional de alguien con el propósito de provocar el detrimento de un individuo, grupo o institución (v.g., fraude, mentiras, asumir la identidad de otros).

El estudio consideró como unidad de análisis a la conducta, cada vez que una conducta se presentaba era codificada.

Cuando se crean las categorías, éstas deben ser definidas con precisión y es necesario explicitar qué se va a comprender en cada caso y qué habrá de excluirse.

El análisis de contenido consiste en asignar cada unidad a una o más categorías. De hecho, el producto de la codificación son frecuencias de cada categoría. Se cuenta cuántas veces se repite cada categoría o subcategoría (cuántas unidades de análisis entraron en cada una de las categorías). Por ejemplo, Greenberg et al. (1980, pp.113) encontraron los resultados que se muestran en la Tabla 9.5.

TABLA 9.5
 EJEMPLO DE LOS RESULTADOS DE LA CODIFICACIÓN DE
 ACUERDO AL ESTUDIO DE GREENBERG et al. (1980)

Incidencia de actos antisociales en los tres periodos

	Año 1 *		Año 2*		Año 3*	
	f	%	f	%	f	%
1. Agresión física						
A. Asalto con un objeto	466	(15.7)	248	(10.8)	370	(13.6)
B. Asalto sin un objeto	111	(3.7)	159	(6.9)	177	(6.5)
C. Amenaza física	180	(6.1)	233	(10.1)	135	(5.0)
D. Disparar	106	(3.6)	75	(3.2)	74	(2.7)
E. Otras	128	(4.3)	171	(7.4)	130	(4.8)
II. Agresión verbal	1 629	(55.0)	1 099	(47.6)	1 464	(54.0)
III. Robo	61	(2.1)	72	(3.1)	44	<1.6)
IV. Engaño	283	(9.5)	251	(10.9)	319	(11.8)
Total	2964		2308	(100.0)	2713	(100.0)
Horas analizadas		(68.5)		(58)		(63)

* Año 1 incluyó de octubre de 1975 hasta que se grabó un episodio de cada una de las series existentes (1976), año 2 igual pero en 1976-1977 y Año 3 igual pero de 1977-1978.

f = número de casos o unidades.

% = porcentajes.

¿CUÁLES SON LOS PASOS PARA LLEVAR A CABO EL ANÁLISIS DE CONTENIDO?

Ya hemos mencionado tres:

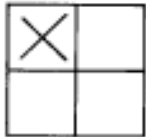
1. *Definir con precisión el universo y extraer una muestra representativa.*
2. *Establecer y definir las unidades de análisis.*
3. *Establecer y definir las categorías y subcategorías que representen a las variables de la investigación.*

Los demás pasos serían:

4. *Seleccionar a los codificadores.* Los codificadores son las personas que habrán de asignar las unidades de análisis a las categorías. Deben ser personas con un nivel educativo profesional (estudiantes a nivel de licenciatura como mínimo).
5. *Elaborar las hojas de codificación.* Estas hojas contienen las categorías y los codificadores anotan en ellas cada vez que una unidad entra en una categoría o subcategoría.

En la figura 9.19 se muestran los elementos de una hoja de codificación.

Las *hojas de codificación* pueden incluir elementos más precisos para describir el material. Por ejemplo, para analizar artículos de periódicos: Nombre del periódico, fecha de publicación, tipo de artículo (reportaje, entrevista, editorial, artículo de fondo, carta, inserción pagada, artículo general), signatario del artículo, sección donde se publicó, página, volumen de publicación (año, número o equivalente), ubicación,



tamaño (en centímetros/columna), nombre del codificador, día de codificación, etc.

Una hoja puede servir para una o varias unidades, según que nos interese o no tener datos específicos de cada unidad (normalmente se prefiere tener la información por unidad).

Asimismo, las *categorías y subcategorías* deben estar *codificadas* con sus respectivos valores numéricos. En la hoja de codificación de la figura 9.19, 1 significa acto prosocial y 2 acto antisocial (y 1.1.0 es “altruismo”, 1.2.0 “mostrar afecto”, etcétera). Tres cifras en cada caso porque como se verá en el apartado sobre codificación, una variable debe tener el mismo número de columnas para todas sus categorías.

6. *Proporcionar entrenamiento de codificadores.* Este entrenamiento incluye que los codificadores se familiaricen y compenetren con las variables, comprendan las categorías y subcategorías y entiendan las definiciones de ambas. Además, debe capacitarse a los codificadores en la manera de codificar y debe discutirse ampliamente con ellos las diferentes condiciones en que puede manifestarse o estar presente cada categoría y subcategoría. Asimismo, los codificadores deben comprender completamente en qué consiste la unidad de análisis.

FIGURA 9.19

ELEMENTOS DE UNA HOJA DE CODIFICACIÓN

Indicador de quién codificó número, iniciales o letras

Fecha en que se codificó el material

Frecuencias (veces que se repite cada categoría)

Descripción del material que se analizará (discurso, nota periodística, sesión terapéutica, etc.)

CODIFICADOR: 1		MATERIAL A ANALIZAR:		CAPÍTULO DE LA SERIE		
FECHA: 29-XI-88		DURACIÓN: 30 MINUTOS		"ALF" (CBS)		
CATEGORÍAS			FRECUENCIAS	TOTALES		
ACTOS PROSOCIALES	ALTRUISMO (1.1.0.)			2		
	MOSTRAR AFECTO (1.2.0.)			5		
	EXPLICACIÓN DE LOS SENTIMIENTOS DE UNO (1.3.0.)			3		
	EXPLICACIÓN DE LOS SENTIMIENTOS DE OTROS (1.4.0.)	.		0		
ACTOS ANTISOCIALES	AGRESIÓN FÍSICA (2.1.)	Asalto con objeto (2.1.1.)			0	
		Asalto sin objeto (2.1.2.)			0	
		Amenaza física (2.1.3.)			0	
		Disparar (2.1.4.)			0	
		Otros (2.1.5.)			0	
	AGRESIÓN VERBAL (2.2.0.)			3		
	ROBO (2.3.0.)			0		
	ENGAÑO (2.4.0.)			1		
	TOTAL					14

Suma de frecuencias en cada categoría o subcategoría

Al reverso pueden solicitarse comentarios del codificador.

Total de frecuencias o unidades

7. *Calcular la confiabilidad de los codificadores.* Una vez que se lleva a cabo el entrenamiento, los codificadores realizan una codificación provisional de una parte representativa del material (el mismo material para todos los codificadores), para ver si existe consenso entre ellos. Si no hay consenso no puede efectuarse un análisis de contenido confiable.

Para lo anterior se calcula la confiabilidad de cada codificador (individual) y la confiabilidad entre codificadores.

El cálculo de la ~~confiabilidad~~ individual de los codificadores depende de si tenemos uno o varios de éstos.

A. Si se dispone de un solo codificador (porque el material es reducido), se observan las diferencias de la codificación del mismo mensaje hecha por el codificador en dos tiempos diferentes. Si las diferencias son muy pequeñas, el codificador es individualmente confiable. Este tipo de confiabilidad es llamado “confiabilidad intra-codificador”. La cual mide la estabilidad de la prueba y re-prueba de un codificador a través del tiempo.

Otro método consistiría en que el codificador trabaje una parte representativa del material y después aplicar a su codificación (resultados) la siguiente fórmula:

$$\text{Confiabilidad individual} = \frac{\text{Número de unidades de análisis catalogadas correctamente por el codificador}}{\text{Número total de unidades de análisis}}$$

Suponiendo que un mensaje conste de 20 unidades y se logren correctamente 20, la confiabilidad será de 1

$\left[\frac{20}{20} \right]$ que es el máximo de confiabilidad. Si el codificador sólo pudo codificar adecuadamente 15 de los 20,

la confiabilidad sería de $0.75 \left[\frac{15}{20} \right]$.

B. Si se dispone de varios codificadores, la confiabilidad individual puede determinarse así: Se les pide a todos los codificadores que codifiquen el mismo material, se toman los resultados de todos menos los de uno y se compara la codificación de éste contra la del resto. Así se procede con cada codificador.

También puede aplicarse a todos los codificadores la fórmula mencionada para calcular la confiabilidad individual y quien se distancie del resto se considera un caso poco confiable.

EJEMPLO

Codificador A	Codificador B	Codificador C	Codificador D
0.89	0.93	0.92	0.67

El codificador “D” tiene baja confiabilidad.

El cálculo de la confiabilidad intercodificadores se realiza por pares de codificadores (parejas). Se pide a cada pareja formada que codifique el material, se comparan los resultados obtenidos por las parejas, se cuenta el número de acuerdos entre las parejas, se determina el número de unidades de análisis y se aplica la siguiente fórmula:

$$\text{Confiabilidad} = \frac{\text{Número total de acuerdos entre dos parejas}}{\text{Número total de unidades de análisis codificadas}}$$

Después se suman los resultados de esta fórmula y se divide entre el número de comparaciones, que depende del número de parejas.

Veámoslo con un ejemplo:

Las parejas codifican el material:

Codificador A

Pareja 1 ~~codificación de material~~ = Resultado 1
 Codificador B

Codificador C
 Pareja 2 codificación de material = Resultado 2
 Codificador E

Codificador D
 Pareja 3 codificación de material = Resultado 3
 Codificador F

Se determina el número de acuerdos entre las parejas (un acuerdo consiste en que dos parejas codifican en la misma categoría a una misma unidad de análisis):

Pareja	Número de unidades de análisis codificadas	Número de acuerdos entre parejas
1	18	Entre parejas 1 y 2 = 17
2	8	Entre parejas 1 y 3 = 16
3	17	Entre parejas 2 y 3 = 16

Se aplica la fórmula de confiabilidad entre parejas

$$C_{1y2} = \frac{17}{18} = 0.94$$

$$C_{1y3} = \frac{16}{18} = 0.88$$

$$C_{2y3} = \frac{16}{18} = 0.88$$

Debe observarse que no hubo consenso total entre cuántas unidades de análisis podían distinguirse en el material (la pareja 3 distinguió 17 unidades y las parejas 1 y 2 distinguieron 18). En este caso se toma en cuenta para la fórmula de confiabilidad entre parejas, el máximo número de unidades de análisis codificadas por alguna de las parejas. Si fuera:

Pareja	Número de unidades de análisis codificadas	Número de acuerdos entre parejas
A	25	se toma el máximo
B	22	

La fórmula sería $C_{AB} = \frac{21}{25} = 0.84$

Se obtiene la *confiabilidad total* (que es la suma de las confiabilidades entre parejas sobre el número de comparaciones). En nuestro ejemplo:

$$C_T = \frac{C_{1y2} + C_{1y3} + C_{2y3}}{3}$$

$$C_T = \frac{0.94 + 0.88 + 0.88}{3}$$

No es conveniente tolerar una *confiabilidad* menor que 0.85 (ni total ni entre dos parejas) y de ser posible debe superar el 0.89. Al igual que con otros instrumentos de medición, la confiabilidad oscila entre 0 (nula confiabilidad) y 1 (confiabilidad total).

En el análisis de contenido una *baja confiabilidad puede deberse a que las categorías y/o unidades de análisis no han sido definidas con claridad y precisión, a un deficiente entrenamiento o a inhabilidad de los codificadores*. Cuando se obtiene una baja confiabilidad debe detectarse y corregirse el problema.

Asimismo, es conveniente calcular la confiabilidad a la mitad de la codificación (con el material codificado) y al finalizar ésta.

8. *Efectuar la codificación*. Lo que implica contar las frecuencias de repetición de las categorías (número de unidades que entran en cada categoría).
9. *Vaciar los datos de las hojas de codificación y obtener totales para cada categoría*.
10. *Realizar los análisis estadísticos apropiados*.

9.6.4. Observación

¿QUÉ ES Y PARA QUÉ SIRVE LA OBSERVACIÓN?

La observación consiste en el registro sistemático, válido y confiable de comportamientos o conducta manifiesta. Puede utilizarse como instrumento de medición en muy diversas circunstancias. Haynes (1978) menciona que es un método más utilizado por quienes están orientados conductualmente. Puede servir para determinar la aceptación de un grupo respecto a su profesor, analizar conflictos familiares, eventos masivos (v.g., la violencia en los estadios de fútbol), la aceptación de un producto en un supermercado, el comportamiento de deficientes mentales, etcétera.

Como método para recolectar datos es muy similar al análisis de contenido. De hecho, éste es una forma de observación del contenido de comunicaciones. Es por ello que en este apartado algunos conceptos sólo serán mencionados, pues han sido tratados en el apartado sobre análisis del contenido.

PASOS PARA CONSTRUIR UN SISTEMA DE OBSERVACIÓN

Los pasos para construir un sistema de observación son:

1. *Definir con precisión el universo de aspectos, eventos o conductas a observar*. Por ejemplo, si nuestro interés es observar los recursos con que cuentan las escuelas de un distrito escolar debemos definir lo que concebimos como recurso escolar". Un universo podría ser el comportamiento verbal y no verbal de un grupo de alumnos durante un semestre. Otro universo sería las conductas de un grupo de trabajadores durante sus sesiones en círculos de calidad o equipos para la calidad, en un periodo de un año. O bien las conductas agresivas de un grupo de esquizofrénicos en sesiones terapéuticas.
2. *Extraer una muestra representativa de los aspectos, eventos o conductas a observar*. Un repertorio suficiente de conductas para observar.
3. *Establecer y definir las unidades de observación*. Por ejemplo, cada vez que se presenta una conducta agresiva, cada minuto se analizará si el alumno está o no atento a la clase, durante dos horas al día (7:00 a 9:00 horas), el número de personas que leyeron el tablero de avisos de la compañía, etcétera. El concepto de unidad de análisis es el mismo que en el análisis de contenido solamente que en la observación se trata de

conductas, eventos o aspectos.

4. *Establecer y definir las categorías y subcategorías de observación.* Estas categorías son similares a las definidas para el análisis de contenido. Y la observación también consiste en asignar unidades a categorías y subcategorías de observación.

EJEMPLO DEL CUARTO PASO

En el caso del estudio citado —al hablar de la manipulación de variables independientes en experimentos en la página 114— para probar la hipótesis: “*A mayor grado de información sobre la deficiencia mental que el sujeto normal maneje, mostrará menor evitación en la interacción con el deficiente mental*” (Naves y Poplawsky, 1984): Las unidades de análisis eran cada 10 segundos. La interacción entre la persona normal y el actor que hacía el papel de “deficiente mental” duraba tres minutos. La variable dependiente fue “evitación de la interacción” y las categorías fueron cuatro (Naves y Poplawsky, 1984, Pp. 107-109):

1. **DISTANCIA FÍSICA:** se refiere a si el sujeto experimental aumenta o disminuye su distancia hacia el interlocutor a partir de la distancia que inicialmente debía ocupar; esta distancia inicial estuvo delimitada por los asientos que el actor y el sujeto debían ocupar y, según la teoría, es la distancia en la que dos extraños en una situación de comunicación, pueden interactuar cómodamente. Las dimensiones que esta variable adquiere son el acercarse (afiliación) con valor de “1” o el alejarse (evitación) del actor (deficiente mental) con valor de “0”, mediante inclinaciones corporales o bien modificando por completo su distancia.

2. **MOVIMIENTOS CORPORALES QUE DENOTAN TENSION:** esta variable se orienta a captar los movimientos que el sujeto está realizando como índice de tensión (evitación) con valor de 0” o de relajación (afiliación) que experimenta, con valor de “1”. En esta variable específicamente se analizan movimientos de pies y piernas a un ritmo acelerado, ademanes con brazos y manos (como el estarse rascando, picando, etc.) y la postura en general del sujeto.

3. **CONDUCTA VISUAL DEL SUJETO:** que según lo estipulado en esta investigación adquiere dos dimensiones:

a) dirigida hacia el interlocutor (afiliación), con valor de “1”.

b) dirigida hacia cualquier otra parte (evitación), con valor de “0”.

4. **CONDUCTA VERBAL:** este indicador está compuesto por el contenido verbal del discurso del sujeto hacia el deficiente y se orienta primordialmente al formato del discurso; incluye dos modalidades:

frases u oraciones completas (afiliación), con valor de “1”.

frases dicótomas y silencios (evitación), con valor de “0”.

La modalidad de frases dicótomas incluye respuestas monosilábicas (sí, no), murmulos, sonidos guturales, etc., y los silencios que se interpretan como respuestas dicótomas (respuesta de evitación) son los silencios no naturales en el discurso, aquellos en los que expresamente el sujeto se queda un periodo en silencio. La conducta verbal se mide a través del diálogo que sostenga el sujeto para con el deficiente mental; es decir, en respuesta al guión que el actor interpreta para con cada sujeto (que es idéntico para todos) y en las intervenciones que el propio sujeto realice. En un principio se pensó que además de medir la conducta verbal en cuanto a su formato, convendría medirla en cuanto a su contenido también; es decir, si las frases emitidas por él en respuesta a lo expresado por el deficiente revelaban un contenido positivo o negativo. Esta modalidad no fue incluida por la dificultad que presenta el obtener una medición objetiva.

Finalmente, cabe establecer que en los cuatro indicadores (variables) adoptados para medir la evitación en la interacción se establecieron dimensiones cuya medición fuese dicotómica; es decir, que las opciones de respuesta para cada variable únicamente pudiesen ser codificadas bajo la escala de 0 — 1; esto es, evitación o no evitación. La razón por la cual dichas variables no adquieren más opciones de respuesta obedece a la dificultad por detectar conductas de evitación tanto verbales como no verbales.

LAS SUBCATEGORÍAS PUEDEN SER ESCALAS DE ACTITUDES

FIGURA 9.20

EJEMPLO DE UNA HOJA DE CODIFICACIÓN PARA OBSERVAR CONDUCTA

Nombre _____ Grupo _____
 Fecha _____

CATEGORÍA	SUBCATEGORÍA	0'	10'	20'	30'	40'	50'	1'	10'	20'	30'	40'	50'	2'	10'	20'	30'	40'	50'	3'	
DISTANCIA FÍSICA	ALEJAMIENTO (0)																				
	ACERCAMIENTO (2)																				
	ESTÁTICO (1)																				
MOVIMIENTOS CORPORALES																					
	TENSIÓN (0)																				
	RELAJACIÓN (2)																				
	NINGUNO (1)																				
CONDUCTA VISUAL																					
	AL SUJETO (1)																				
	A OTRA PARTE (0)																				
CONDUCTA VERBAL																					
	F. COMPLETAS (1)																				
	F. DICOTOMAS (0)																				
	SILENCIOS (0)																				

Nota: Se agregaron las subcategorías "estático" en la categoría "distancia física" y "ninguno" en la categoría "movimientos corporales".

8. *Proporcionar entrenamiento de codificadores* (en las variables, categorías, subcategorías, unidades de análisis y el procedimiento de codificar, así como sobre las diferentes maneras como puede manifestarse una categoría o subcategoría de conducta).

9. *Calcular la confiabilidad de los observadores* (intra-observador e interobservadores). Los procedimientos y fórmulas pueden ser las mismas que las vistas en el apartado sobre el análisis de contenido, lo único que cambia es la palabra codificador(es)", "codificación", "codificada(s)"; por "observador(es)", "observación", "observada(s)".

Por ejemplo:

$$\text{Confiabilidad} = \frac{\text{Número de unidades de análisis catalogadas correctamente por el observador individual}}{\text{Número total de unidades de análisis}}$$

$$\text{Confiabilidad} = \frac{\text{Número total de acuerdos entre dos parejas}}{\text{entre parejas} \quad \text{Número total de unidades de análisis observadas}}$$

Haynes (1978, p. 160) proporciona otra fórmula para calcular la confiabilidad entre observadores o el grado de acuerdo interobservadores (Ao).

$$A_o = \frac{I_a}{I_a + I_d}$$

Donde “Ia” es el número total de acuerdos entre observadores e “Id” es el número total de desacuerdos entre observadores. Un “Acuerdo” es definido como la codificación de una unidad de análisis en una misma categoría por distintos observadores. Se interpreta como cualquier *coeficiente de confiabilidad* (0 a 1).

10. *Llevar a cabo la codificación por observación.*
11. *Vaciar los datos de las hojas de codificación y obtener totales para cada categoría.*
12. *Realizar los análisis apropiados.*

TIPOS DE OBSERVACIÓN

La observación puede ser *participante* o *no participante*. En la primera, el observador interactúa con los sujetos observados y en la segunda no ocurre esta interacción. Por ejemplo, un estudio sobre las conductas de aprendizaje de niños autistas, en donde una instructora interactúa con los niños y al mismo tiempo codifica.

VENTAJAS DE LA OBSERVACIÓN

Tanto la observación como el análisis de contenido tienen varias ventajas:

- 1) *Son técnicas de medición no obstrusivas.* En el sentido que el instrumento de medición no “estimula” el comportamiento de los sujetos (las escalas de actitud y los cuestionarios pretenden “estimular” una respuesta a cada ítem). Los métodos no obstrusivos simplemente registran algo que fue estimulado por otros factores ajenos al instrumento de medición.
- 2) *Aceptan material no estructurado.*
- 3) *Pueden trabajar con grandes volúmenes de datos (material).*

9.6.5. Pruebas e inventarios estandarizados

¿QUÉ SON LAS PRUEBAS ESTANDARIZADAS?

En la actualidad existe una amplia diversidad de pruebas e inventarios desarrollados por diversos investigadores para medir un gran número de variables. Estas pruebas tienen su propio procedimiento de aplicación, codificación e interpretación, y se encuentran disponibles en diversas fuentes secundarias y terciarias, así como en centros de investigación ~y difusión del conocimiento. Hay *pruebas para medir habilidades y aptitudes* (v.g., habilidad verbal, razonamiento, memoria, inteligencia, percepción, habilidad numérica), *la personalidad los intereses, los valores, el desempeño, la motivación, el aprendizaje, el clima laboral en una organización, etcétera*. También se puede disponer de *pruebas clínicas* para detectar conducta anormal, pruebas para seleccionar personal, pruebas para conocer las percepciones y/o opiniones de las personas respecto a diversos tópicos, pruebas para medir la autoestima y —en fin— otras muchas variables del comportamiento.⁴⁶

⁴⁶ Para conocer la diversidad de estas pruebas y sus aplicaciones se recomienda Anastasi (1982), Thorndike y Hagen (1980), Cronbach (1984) y Nunnally (1970). Son obras clásicas sobre medición y el manejo de pruebas estandarizadas. El problema en el uso de estas pruebas es que la mayoría ha sido desarrollada en contextos muy diferentes al latinoamericano, y en ocasiones su utilización puede ser inadecuada, inválida y poco confiable. Cuando se

utilice como instrumento de medición una prueba estandarizada es conveniente que *se seleccione una prueba desarrollada o adaptada por algún investigador para el mismo contexto de nuestro estudio y que sea válida y confiable* (debemos tener información a este respecto). En el caso de que elijamos una prueba diseñada en otro contexto, es necesario adaptarla y aplicar pruebas piloto para calcular su validez y confiabilidad, así como ajustarla a las condiciones de nuestra investigación. El instrumento o prueba debe demostrar que es válido y confiable para el contexto en el cual se va a aplicar.

Un tipo de pruebas estandarizadas bastante difundido lo constituyen las “*pruebas proyectivas*”, las cuales presentan estímulos a los sujetos para que respondan a ellos; después se pueden analizar las respuestas tanto cuantitativamente como cualitativamente y se interpretan. Estas pruebas miden proyecciones de los sujetos, como por ejemplo, la personalidad.

Dos pruebas proyectivas muy conocidas son el *Test de Rorschach* (que presenta a los sujetos manchas de tinta en tarjetas o láminas blancas numeradas y éstos relatan sus asociaciones e interpretaciones en relación a las manchas) y el *Test de Apercepción Temática* (que con un esquema similar al de Rorschach presenta a los sujetos cuadros que evocan narraciones o cuentos y las personas deben elaborar una interpretación).

Bastantes pruebas estandarizadas (v.g., las proyectivas) requieren de un entrenamiento considerable y un conocimiento profundo de las variables por parte del investigador que habrá de aplicarlas e interpretarlas. No pueden aplicarse con superficialidad e indiscriminadamente. La manera de aplicar, codificar, calificar e interpretar las pruebas estandarizadas es tan variada como los tipos existentes.

9.6.6. Sesiones en profundidad

¿QUÉ SON LAS SESIONES EN PROFUNDIDAD?

Un método de recolección de datos cuya popularidad ha crecido son las sesiones en profundidad. Se reúne a un grupo de personas y se trabaja con éste en relación a las variables de la investigación. Pueden realizarse una o varias reuniones. El procedimiento usual es el siguiente.

PASOS PARA REALIZAR LAS SESIONES DE GRUPO

1. *Se define el tipo de personas* que habrán de participar en la sesión o sesiones.
2. *Se detectan personas del tipo* elegido.
3. *Se invita* a estas personas a la sesión o sesiones.
4. *Se organizan la sesión o sesiones.* Cada sesión debe efectuarse en un lugar confortable, silencioso y aislado. Los sujetos deben sentirse cómodos y relajados. Asimismo, es indispensable planear cuidadosamente lo que se va a tratar en la sesión o sesiones (desarrollar una agenda) y asegurar los detalles (aún las cuestiones más sencillas como el servir café y refrescos).
5. *Se lleva a cabo cada sesión.* El conductor debe ser una persona entrenada en el manejo o conducción de grupos y debe crear “rapport” en el grupo (clima de confianza). Asimismo, debe ser un individuo que no sea percibido como “distante” por los participantes de la sesión y tiene que propiciar la participación de todos. La paciencia es una característica que deberá tener. Durante la sesión se pueden pedir opiniones, hacer preguntas, administrar cuestionarios, discutir casos, intercambiar puntos de vista, valorar diversos aspectos. Es conveniente que cada sesión se grabe en cinta o videocinta y después realizar análisis de contenido y observación. El conductor debe tener muy en claro la información o datos que habrán de recolectarse y evitar desviaciones del objetivo planteado.
6. *Elaborar el reporte de sesión.* El cual incluye principalmente datos sobre los participantes (edad, sexo, nivel educativo y todo aquello que sea relevante para el estudio), fecha y duración de la sesión, información completa del desarrollo de la sesión, actitud y comportamiento de los participantes hacia el conductor y la sesión en sí, resultados de la sesión y observaciones del conductor, así como una bitácora de la sesión.

7. Llevar a cabo la codificación y análisis correspondientes.

EJEMPLOS

Algunos ejemplos de la aplicación de este método podrían ser las sesiones en donde se evalúe a un nuevo producto, digamos un dulce. En estas sesiones se podría pedir a los participantes opiniones sobre el sabor, color, presentación, precio, cualidades, etcétera, del producto; discutir a fondo las propiedades, cualidades y carencias del producto: administrarles una escala de actitudes o un cuestionario, y hacer preguntas abiertas a cada participante. O bien sesiones para analizar la popularidad de varios candidatos políticos, evaluar el servicio y la atención recibida en un supermercado, indagar la percepción de un grupo de estudiantes sobre la calidad de la enseñanza recibida o conocer la opinión de los sectores de una comunidad sobre una reforma electoral.

También se pueden organizar sesiones con diferentes tipos de la población y mixtas. Por ejemplo, los autores participaron en un estudio para evaluar un programa televisivo que acababa de “salir al aire” con un nombre-formato nuevos. Además de realizarse una encuesta telefónica se organizaron varias sesiones (algunas con amas de casa, otras con estudiantes, también con trabajadores, empleados de oficina, ejecutivos, profesores, publicistas y otros grupos tipo; así como reuniones donde participaban —por ejemplo— un ama de casa de más de 60 años, un ama de casa de 50 años, un ama de casa más joven, un empleado de una oficina pública, una secretaria, un profesor, un dependiente de supermercado, un ejecutivo y dos estudiantes.

En las sesiones se profundizó en el formato del programa (música, manejo de cámaras, duración, manejo de comerciales, sonido, manera de presentar invitados, etc.), los conductores, el contenido, etcétera.

Normalmente en las sesiones participan de ocho a quince personas. No debe excederse de un número manejable de sujetos. El formato y naturaleza de la sesión o sesiones depende del objetivo y las características de los participantes.

9.6.7. Otras formas de recolección de los datos

¿QUÉ OTRAS MANERAS EXISTEN PARA RECOLECTAR LOS DATOS?

En ocasiones puede acudir a *archivos* que contengan los datos. Por ejemplo, podemos acudir a la alcaldía de algunas ciudades para solicitar datos relacionados con la violencia (si nuestra hipótesis fuera: “La violencia manifiesta en la ciudad de México es mayor que en la ciudad de Caracas”): número de asaltos, violaciones, robos a casa-habitación, asesinatos, etc. (datos por habitante, distrito y generales). También podríamos acudir a los hospitales y las diferentes procuradurías, etc. Otro ejemplo, sería consultar los archivos de una universidad y tomar los datos de inteligencia, personalidad u otras variables que nos interesen. Desde luego, a veces esta información no es accesible. En México hay un organismo que proporciona datos (incluso grabados en disco para computadora) sobre estadísticas nacionales, el Instituto Nacional de Estadística, Geografía e Informática (INEGI).

Asimismo, pueden utilizarse datos recolectados por otros investigadores, a lo que se conoce como “*análisis secundario*”. En este caso es necesario tener la certeza de que los datos son válidos y confiables, así como conocer la manera como fueron codificados. El intercambio de éstos es una práctica común entre investigadores. Además, existen métodos propios de las diferentes ciencias sociales como el *análisis de redes* para evaluar cómo se manifiesta la comunicación en un sistema social (quién se comunica con quién, quiénes distorsionan la información, como fluye la comunicación, quiénes son los líderes comunicativos, etc.), *sistemas de medición fisiológica*, *escalas multidimensionales* que miden a los sujetos en varias dimensiones (v.g., el sistema Galileo de J. Woelfel y E.L. Fink —1980—), como el medir la distancia psicológica entre los conceptos “patria,” “madre”, “presidente” y “nación”, etc., tomando en cuenta dimensiones cognitivas y emocionales. Y en fin otros métodos que escapan al nivel introductorio de este libro.

Para el manejo de archivos se recomienda consultar a Webb, Campbell y Schwartz (1966), para el análisis de redes a Rogers y Kincaid (1981) y para escalas multidimensionales a Norton (1980), Woelfel y Danes (1980) y, desde luego, las obras clásicas de Torgerson (1958) y Rummey, Shephard y Nerove (1972).

9.6.8. Combinación de dos o más instrumentos de recolección de los datos

¿PUEDE UTILIZARSE MÁS DE UN TIPO DE INSTRUMENTO DE MEDICIÓN?

En algunos casos, el investigador utiliza varias formas de medición para tener diferentes enfoques sobre las variables. Por ejemplo, el clima laboral en una organización puede medirse a través de una encuesta utilizando un cuestionario, pero además pueden realizarse varias sesiones en profundidad para solicitar opiniones sobre el clima laboral y los problemas existentes, observarse el comportamiento de los trabajadores y analizar el contenido de sus mensajes dirigidos a la organización (cartas de sugerencias, letreros pintados en los baños, quejas en sus reuniones, etc.).

9.7. ¿COMO SE CODIFICAN LAS RESPUESTAS A UN INSTRUMENTO DE MEDICIÓN?

Ya se ha venido mencionando que las categorías de un ítem o pregunta y las categorías y subcategorías de contenido u observación deben *codificarse a través de símbolos o números*. Y deben codificarse porque de lo contrario no puede efectuarse ningún análisis o solamente se puede contar el número de respuestas en cada categoría (v.g., 25 contestaron “sí” y 24 respondieron “no”). Pero el investigador se interesa en realizar análisis más allá de un conteo de casos por categoría y la mayoría de los análisis se llevan a cabo por computadora. Para ello es necesario transformar las respuestas en símbolos o valores numéricos. Los datos deben resumirse, codificarse y prepararse para el análisis.

También se comentó que las categorías pueden ir o no precodificadas (llevar la codificación en el instrumento de medición antes de que éste sea aplicado) y que las preguntas abiertas no pueden estar precodificadas. Pero en cualquier caso, una vez que se tienen las respuestas, éstas deberán codificarse.

La codificación de las respuestas implica cuatro pasos:

- 1) Codificar las categorías de ítems, preguntas y categorías de contenido u observación no precodificadas.
- 2) Elaborar el libro de códigos.
- 3) Efectuar físicamente la codificación.
- 4) Grabar y guardar los datos en un archivo permanente.

Veamos cada paso con algunos ejemplos.

Codificar

Si todas las categorías fueron precodificadas y no se tienen preguntas abiertas primer paso no es necesario. Éste ya se efectuó.

Si las categorías no fueron precodificadas y se tienen preguntas abiertas, asignarse los códigos o la codificación a todas las categorías de los ítems, preguntas o de contenido u observación. Por ejemplo:

Pregunta no precodificada

¿Practica usted algún deporte por lo menos una vez a la semana?

Sí

No

Se codifica

1

=

Sí

0 = No

Frase no precodificada

“Creo que estoy recibiendo un salario justo por mi trabajo”

() Totalmente
de acuerdo

() De acuerdo

() Ni de acuerdo,
ni en desacuerdo

() En desacuerdo
desacuerdo

() Totalmente en

desacuerdo

Se codifica

5 = Totalmente de acuerdo

4 = De acuerdo

3 = Ni de acuerdo, ni en desacuerdo

2 = En desacuerdo

1 = Totalmente en desacuerdo

Tratándose de preguntas abiertas ya se expuso cómo se codifican.

Libro de códigos

Una vez que están codificadas todas las categorías del instrumento de medición, se procede a elaborar el “libro de códigos”.

El *libro de códigos* es un documento que describe la localización de las variables y los códigos asignados a los atributos que las componen (categorías y/o subcategorías) (Babbie, 1979). Este libro cumple con dos funciones: i) es la guía para el proceso de codificación y ii) es la guía para localizar variables e interpretar los datos durante el análisis (Babbie, 1979). El libro de códigos puede conducirnos a los significados de los valores de las categorías de las variables.

Los elementos de un libro de códigos son: *variable, pregunta / ítem / tema, categorías-subcategorías, columna(s)*.

Supongamos que tenemos una escala Likert con tres ítems (frases):

“La Dirección General de Impuestos Nacionales informa oportunamente sobre cómo, dónde y cuándo pagar los impuestos”

(5) Muy de acuerdo

(4) De acuerdo

(3) Ni de acuerdo, ni en desacuerdo

(2) En desacuerdo

(1) Muy en desacuerdo

“Los servicios que presta la Dirección General de Impuestos Nacionales son en general muy buenos”

(5) Muy de acuerdo

- (4) De acuerdo
- (3) Ni de acuerdo, ni en desacuerdo
- (2) En desacuerdo
- (1) Muy en desacuerdo

“La Dirección General de Impuestos Nacionales se caracteriza por la deshonestidad de sus funcionarios”

- (1) Muy de acuerdo
- (2) De acuerdo
- (3) Ni de acuerdo, ni en desacuerdo
- (4) En desacuerdo
- (5) Muy en desacuerdo

El libro de códigos sería el que se muestra en la figura 9.21.

FIGURA 9.21
EJEMPLO DE UN LIBRO DE CÓDIGOS CON UNA ESCALA
DE ACTITUD TIPO LIKERT (TRES ÍTEMS)

VARIABLE	ÍTEM	CATEGORÍAS	CÓDIGOS	COLUMNA	— Actitud hacia la Dirección General
de Impuestos Nacionales	Frase 1	(informa)	— Muy de acuerdo		
		— De acuerdo			
		—Ni de acuerdo, ni en desacuerdo			
		—En desacuerdo			
		—Muy en desacuerdo			
		5			
		4			
		3			
		2			
	1 1	Frase 2			
	(servicios)	— Muy de acuerdo			
		— De acuerdo			
		—Ni de acuerdo, ni en desacuerdo			
		—En desacuerdo			
		—Muy en desacuerdo			
		5			
		4			
		3			
		2			
	1 2	Frase 3 (deshonestidad)	— Muy de acuerdo		
		— De acuerdo			
		—Ni de acuerdo, ni en desacuerdo			
		—En desacuerdo			
		—Muy en desacuerdo			
		1			
		2			
		3			
		4			
	5	3			

En el caso del estudio por observación de Naves y Poplawsky (1984) (figura 9.20), el libro de códigos sería el que se muestra en la figura 9.22.

FIGURA 9.22

EJEMPLO DE UN LIBRO DE CÓDIGOS EN EL CASO DEL ESTUDIO DE NAVES Y POPLAWSKY (1984)⁴⁷

LIBRO DE CÓDIGOS				
VARIABLE	CATEGORÍAS	SUBCATEGORÍAS	CÓDIGOS	COLUMNAS
— Número de sujeto	— Sujetos	—	01 a 30 (hubo 30 sujetos)	1 Y 2
— Tratamiento experimental	— Grupo cultural		1	3
	— Grupo socio-psicológico		2	
— Conducta de evitación	— Distancia física	— Alejamiento	0	4
		— Acercamiento	2	
		— Estático	1	
	— Movimientos corporales	— Tensión	0	5
		— Relajación	2	
		— Ninguno	1	
	— Conducta visual	— Al sujeto	1	6
		— A otra parte	0	
	— Conducta verbal	— Frases completas	1	7
		— Frases dicótomas o silencios	0	
— Codificador	— LRE	—	1	8
	— LEMM	—	2	
	— ÇFT	—	3	

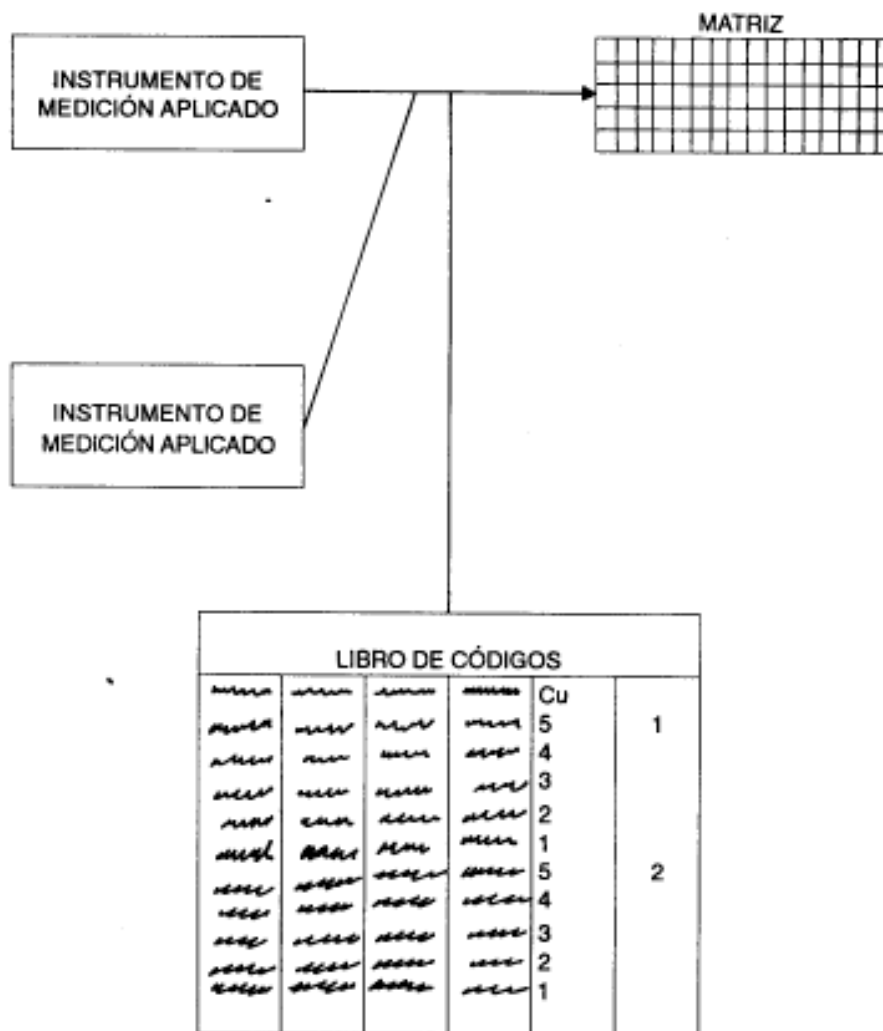
Es decir, el libro de códigos es un manual para el investigador y los codificadores. Los cuestionarios contestados, las escalas aplicadas, las hojas de codificación, las pruebas respondidas o cualquier otro instrumento de medición administrado son transferidos a una matriz, la cual es el conjunto de datos simbólicos o numéricos producto de la aplicación del instrumento. Esta matriz es lo que habrá de analizarse. El apartado “columna” dentro del libro de códigos tiene sentido en la matriz; veamos por qué. La matriz tiene renglones y columnas; los renglones representan casos o sujetos en la investigación, las columnas son los lugares donde se registran los valores en las categorías o subcategorías. Esto podría esquematizarse así:

MATRIZ DE DATOS

	Columna 1	Columna 2	Columna 3	
Caso 1				Valores de categorías / sub-categorías
Caso 2				
Caso 3				
Caso 4				
Caso k				

⁴⁷ Desde luego, Naves y Popiawsky (1984) para las categorías de “conducta de evitación” obtenían esta codificación cada 10 segundos (cada unidad de análisis), y sumaban el número de 1 (unos) y lo transformaban en porcentajes. Aquí suponemos que toda la interacción con el deficiente mental puede categorizarse y subcategorizarse.

Los resultados del instrumento de medición se transfieren a la matriz por medio del libro de códigos. El proceso puede representarse así:



Sin el *libro de códigos* no puede llevarse a cabo la transferencia. Vamos a suponer que hubiéramos aplicado la escala de actitud con tres ítems de la figura 9.21 a cuatro personas, obteniendo los siguientes resultados:

Persona 1

A continuación.
1. "La Dirección de Impuestos Nacionales informa claramente sobre cómo, donde y cuándo pagar los impuestos"

El sujeto obtuvo:

(5) Muy de acuerdo (3) Ni de acuerdo, (2) En desacuerdo ni en desacuerdo		4 (de acuerdo)
<input checked="" type="checkbox"/> De acuerdo (1) Muy en desacuerdo		
2. "Los servicios que presta la Dirección de Impuestos Nacionales son en general muy buenos"		
<input checked="" type="checkbox"/> Muy de acuerdo (3) Ni de acuerdo, (2) En desacuerdo ni en desacuerdo		
(4) De acuerdo (1) Muy en desacuerdo		5 (muy de acuerdo)
3. "La Dirección de Impuestos Nacionales se caracteriza por la deshonestidad de sus funcionarios"		
(1) Muy de acuerdo <input checked="" type="checkbox"/> Ni de acuerdo, (4) En desacuerdo ni en desacuerdo		
(2) De acuerdo (5) Muy en desacuerdo		3 (ni de acuerdo, ni en desacuerdo)

Persona 2

Obtuvo respectivamente: 3 (ni de acuerdo, ni en desacuerdo)
4 (de acuerdo)
3 (ni de acuerdo, ni en desacuerdo)

Persona 3

Obtuvo respectivamente: 4
4
4

Persona 4

Obtuvo respectivamente: 5
4
3

De acuerdo con el libro de códigos (figura 9.21), tendríamos la siguiente matriz (figura 9.23):

FIGURA 9.23

EJEMPLO DE MATRIZ DE DATOS PARA EL LIBRO DE CÓDIGOS DE LA FIGURA 9.21

	columna 1 (frase 1) (informa)	columna 2 (frase 2) (servicios)	columna 3 (frase 3) (deshonestidad)	
Persona 1	4	5	3	casos (en el ejemplo, sujetos)
Persona 2	3	4	3	
Persona 3	4	4	4	
Persona 4	5	4	3	

Valores de los sujetos en los ítems (en el ejemplo, frases) (categorías en las que cayeron transformadas a sus valores numéricos, es decir, codificadas)

En el ejemplo de Naves y Poplawsky (figura 9.22), a matriz sería la de la figura 9.24.

FIGURA 9.24

EJEMPLO HIPOTÉTICO DE MATRIZ DE DATOS PARA EL LIBRO DE CÓDIGOS DE LA FIGURA 9.22

	Columna 1 Número de sujeto	Columna 2 Número de sujeto	Columna 3 Tratamiento experimental	Columna 4 Distancia física	Columna 5 Movimientos corporales	Columna 6 Conducta visual	Columna 7 Conducta verbal	Columna 8 Codificador
S ₁	0	1	1	0	0	0	0	1
S ₂	0	2	1	0	0	0	0	1
S ₃	0	3	2	2	2	1	1	2
S ₄	0	4	1	1	0	0	0	2
S ₅	0	5	2	2	2	1	1	2
S ₆	0	6	1	0	1	1	0	3
S ₇	0	7	2	2	2	1	1	3
S ₈	0	8	2	2	2	1	1	3
S ₉	0	9	2	1	2	0	1	1
S ₁₀	1	0	1	2	1	0	0	2

El libro de códigos indica a los codificadores qué variable, ítem/categoría/subcategoría va en cada columna y qué valores debe anotar en cada columna, así como el significado de cada valor numérico.

Con el libro de códigos sabemos que el sujeto 1 es el "01", que pertenece al grupo cultural ("1" en la tercer columna), que tuvo una conducta de alejamiento en su distancia física ("0" en la cuarta columna), que sus movimientos corporales fueron de tensión ("0" en la quinta columna), que su conducta visual fue a otra parte, no vio al sujeto ("0" en la sexta columna), que dijo frases dicótomas y/o silencios ("0" en la séptima columna) y que fue codificado por LRE ("1" en la octava columna). Y así con cada sujeto.

Obsérvese que, sin el libro de códigos, no se puede codificar y una matriz de datos carece de significado. La siguiente matriz no nos dice nada sin el libro de códigos. Por ejemplo:

5	0	0	2	1	1	0	1	0	1	5
5	0	1	2	2	3	2	3	0	0	4
2	1	0	0	0	0	0	0	0	0	0
1	1	1	1	1	4	0	2	0	0	3
3	0	0	1	2	5	0	2	0	0	3
3	1	0	1	2	5	0	1	0	1	3
4	0	1	0	2	4	0	1	0	1	4
5	1	1	2	2	1	0	0	0	1	4
2	0	1	2	2	2	0	2	1	0	5

¿Qué significa cada columna, cada dígito? Está en clave y sólo podemos tener acceso a ella mediante el libro de códigos.

En el libro de códigos y en la matriz de datos, una variable, ítem, categoría o subcategoría puede abarcar una, dos o más columnas, dependiendo de lo que esté indicando. Veamos el siguiente libro de códigos y la matriz correspondiente (figura 9.25).

FIGURA 9.25

EJEMPLO DE LA VARIABILIDAD QUE PUEDEN TENER LAS COLUMNAS

LIBRO DE CÓDIGOS			
VARIABLE	CATEGORÍAS	CÓDIGOS	COLUMNA
— Número de sujeto	000 a 128	000 a 128	1, 2 y 3
— Sexo del respondiente	masculino	1	4
	femenino	2	
— Edad del respondiente	00 a 77	00 a 77	5 y 6
	No respondió	99	
— Ingresos del respondiente	— Salario mínimo o menos	1	7
	— De dos a tres veces el salario mínimo	2	
	— De cuatro a cinco veces el salario mínimo	3	
	— De seis a ocho veces el salario mínimo	4	
	— Nueve a once veces el salario mínimo	5	
	— Doce a quince veces el salario mínimo	6	
	— 16 o más veces el salario mínimo	7	
	— No respondió	9	
	— Tamaño del lugar en que vive (m ²)	0000 a 9998	
No respondió		9999	

MATRIZ DE DATOS

VARIABLES

	Número de sujeto			Sexo	Edad		Ingresos	Tamaño del lugar en que vive (residencia)			
Columnas Sujetos	1	2	3	4	5	6	7	8	9	10	11

LA MATRIZ SE LLENA DE ACUERDO A LOS RESULTADOS

Valores perdidos

Cuando las personas no responden a un ítem o contestan incorrectamente o no puede registrarse la información (v.g., no se pudo observar la conducta), se crea una o varias categorías de valores perdidos y se les asignan sus respectivos códigos.

EJEMPLO

Sí = 1	Sí = 1	
No = 2	No = 2	
No contestó = 3		o
Contestó incorrectamente = 4		Valor perdido por diversas razones = 9

Hasta el momento se han presentado, por razones didácticas, ejemplos resumidos de libros de códigos. Desde luego, un libro de códigos normalmente tiene más variables o categorías y consecuentemente más columnas (al igual que la matriz de datos). Hay matrices que pueden tener 500 o más columnas. Asimismo, debe recordarse que los renglones son casos (sujetos, escuelas, series de televisión, etc.), y a veces es necesario extender las columnas a otro renglón u otros renglones (en computadoras limitadas a manejar 80 columnas). Esto podría representarse así:

S_1	Col. 1	2	3	4	etc...	Col. 80
	Col. 81	82	83	84	etc...	Col. 160
	Col. 161	162	163	164	etc...	Col. 240
	Col. 241				etc...	Col. 320
S_2	Col. 1	2	3	4		Col. 80
	Col. 81	82	83	84		Col. 160
	Col. 161	162	163	164		Col. 240
	Col. 241					Col. 320
S_3						
S_k						

Codificación física

El tercer paso del proceso de codificación es la codificación física de los datos, es decir, el llenado de la matriz de datos. Esta codificación la efectúan los codificadores, a quienes se les proporciona el libro de códigos. Así, cada codificador va vaciando las respuestas en la matriz de datos, de acuerdo con el libro de códigos. El vaciado de la matriz de datos puede hacerse en "hojas de tabulación", las cuales tienen columnas y renglones. En la figura 9.26 se muestra un ejemplo de una de estas hojas.

Si no alcanzan con una hoja de tabulación utilizan las hojas necesarias para vaciar los datos de todos los casos. Por ejemplo, la primera hoja puede alcanzarnos para 24 casos, pero si tenemos 200 casos, habremos de utilizar 9 hojas. Cada hoja estará llena de dígitos.

GRABADO Y GENERACIÓN DE ARCHIVOS

Las hojas de tabulación pueden copiarse a un disco o una cinta magnética para computadora, o bien, pueden teclarse a un archivo dentro de una cuenta en una computadora. En cualquier caso, se crea un archivo o "file", el cual debe ser nombrado y contiene los datos codificados en valores numéricos en forma de matriz. El proceso se muestra en la figura 9.27.

También existen en la actualidad sistemas para la lectura óptica, los cuales pueden leer y almacenar los datos directamente de los cuestionarios (u otros instrumentos de medición) o de las hojas de tabulación. Desde luego, son sistemas costosos que requieren de lápices o tinta especiales. Asimismo, algunos investigadores con bastante experiencia pasan los datos directamente del instrumento aplicado al disco, cinta o computadora, pero se requiere de mucha práctica y personal capacitado.

Así, los datos han sido capturados en un archivo permanente y están listos para ser analizados mediante un programa de computadora. El proceso va desde la respuesta de los sujetos hasta un archivo que contiene una matriz (que es una matriz de valores numéricos que significan respuestas). La conducta y los valores de las variables han sido codificados.

FIGURA 9.26
EJEMPLO DE UNA HOJA DE TABULACIÓN

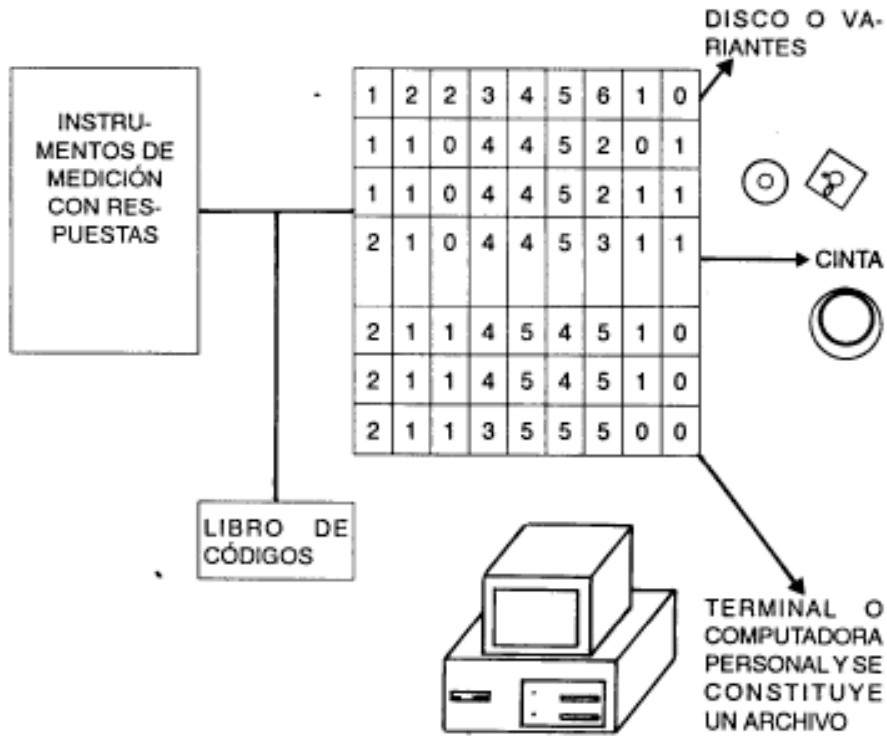
FECHA _____
INVESTIGACIÓN _____
CODIFICACIÓN _____
HOJA No. _____

Columns (variables) ↓ ↓ ↓ ↓ ↓

Rows (cases) ↑ ↑ ↑

FIGURA 9.27

PROCESO PARA CREAR UN ARCHIVO DE DATOS
(ARCHIVAR LA MATRIZ DE DATOS)



RESUMEN

1. Recolectar los datos implica seleccionar un instrumento de medición disponible o desarrollar uno propio, aplicar el instrumento de medición y preparar las mediciones obtenidas para que puedan analizarse correctamente.
2. Medir es el proceso de vincular conceptos abstractos con indicadores empíricos, mediante clasificación y/o cuantificación.
3. En toda investigación medimos las variables contenidas en las hipótesis.
Un instrumento de medición debe cubrir dos requisitos: confiabilidad y validez.
La confiabilidad se refiere al grado en que la aplicación repetida de un instrumento de edición al mismo sujeto u objeto, produce iguales resultados.
6. La validez se refiere al grado en que un instrumento de medición mide realmente la(s) variable(s) que pretende medir.
7. Se pueden aportar tres tipos de evidencia para la validez: evidencia relacionada con el contenido, evidencia relacionada con el criterio y evidencia relacionada con el constructo.
8. Los factores que principalmente pueden afectar la validez son: improvisación, utilizar instrumentos desarrollados en el extranjero y que no han sido validados a nuestro contexto, poca o nula empatía, factores de aplicación.
9. No hay medición perfecta, pero el error de medición debe reducirse a límites tolerables.
10. La confiabilidad se determina calculando un coeficiente de confiabilidad.
11. Los coeficientes de confiabilidad varían entre 0 y 1 (0 = nula confiabilidad, 1 —total confiabilidad).
12. Los procedimientos más comunes para calcular la confiabilidad son la medida de estabilidad, el método de formas alternas, el método de mitades partidas, el coeficiente alfa de Cronbach y el coeficiente

KR-20.

13. La validez de contenido se obtiene contrastando el universo de ítems contra los ítems presentes en el instrumento de medición.
 14. La validez de criterio se obtiene comparando los resultados de aplicar el instrumento de medición contra los resultados de un criterio externo.
 15. La validez de constructo se puede determinar mediante el análisis de factores.
 16. Los pasos genéricos para construir un instrumento de medición son:
 - Listar las variables a medir.
 - Revisar sus definiciones conceptuales y operacionales.
 - Elegir uno ya desarrollado o construir uno propio.
 - Indicar niveles de medición de las variables (nominal, ordinal, por intervalos y de razón).
 - Indicar cómo se habrán de codificar los datos.
 - Aplicar prueba piloto.
 - Construir versión definitiva.
 17. En la investigación social disponemos de diversos instrumentos de medición:
 - a) Principales escalas de actitudes: Likert, Diferencial Semántico y Guttman.
 - b) Cuestionarios (autoadministrado, por entrevista personal, por entrevista telefónica y por correo).
 - c) Análisis de contenido.
 - d) Observación.
- Pruebas estandarizadas (procedimiento estándar).
- Sesiones en profundidad.
- g) Archivos y otras formas de medición.
18. Las respuestas se codifican.
 19. La codificación implica:
 - a) Codificar los ítems o equivalentes no precodificados.
 - b) Elaborar el libro de códigos.
 - c) Efectuar físicamente la codificación.
 - d) Grabar y guardar los datos en un archivo permanente.

CONCEPTOS

BÁSICOS

Recolección de datos

Medición

Instrumento de medición

Confiabilidad

Validez

Coficiente de confiabilidad

Niveles de medición

Medida de estabilidad

Método de formas alternas

Método de mitades partidas

Coficiente alfa de Cronbach

Coficiente KR-20 de Kuder-Richardson

Evidencia relacionada con el contenido

Evidencia relacionada con el criterio

Evidencia relacionada con el constructo

Escalas de actitudes

Escala Likert

Diferencial semántico

Escalograma de Guttman

Cuestionarios

Análisis de contenido

Observación

Pruebas estandarizadas
Pruebas proyectivas
Sesiones en profundidad
Codificación
Codificador
Hojas de codificación
Matriz de datos
Hojas de tabulación
Archivo de datos

EJERCICIOS

1. Busque una investigación en algún artículo científico de una revista en ciencias sociales (ver apéndice uno) donde se incluya información sobre la confiabilidad y la validez del instrumento de medición. ¿El instrumento es confiable?, ¿qué tan confiable?, ¿qué técnica se utilizó para determinar la confiabilidad?, ¿es válido?, ¿cómo se determinó la validez?
2. Responda y discuta con ejemplos la diferencia entre confiabilidad y validez.
3. Defina ocho variables e indique su nivel de medición.
4. Suponga que alguien está tratando de evaluar la actitud hacia el Presidente de la República, construya un cuestionario tipo Likert con 10 ítems para medir dicha actitud e indique cómo se calificaría la escala total (5 ítems positivos y 5 negativos). Finalmente indique la dimensión que cada ítem pretende medir de dicha actitud (credibilidad, presencia física, etc.).
5. Construya un cuestionario para medir lo que usted considere conveniente (con preguntas demográficas y —por lo menos— 10 preguntas más), aplíquelo a 20 conocidos suyos, elabore el libro de códigos y la matriz de datos y vacíela en una hoja de tabulación elaborada por usted. Finalmente lea de la hoja de tabulación el significado de los dígitos de todas las columnas correspondientes a los 5 primeros casos.
6. Planee una sesión en profundidad (indique objetivos, procedimiento, sujetos tipo, agenda, etc.) y organícela con amigos suyos. Al final, autoevalúe su experiencia.
7. Diseñe una investigación (planteamiento del problema, hipótesis, diseño) donde utilice por lo menos dos tipos de instrumentos de medición para recolectar los datos.
8. ¿Cómo se podrían aplicar el análisis de contenido y las sesiones en profundidad para la evaluación de un programa educativo a nivel superior?

BIBLIOGRAFÍA SUGERIDA

- ANASTASI, A. (1982). *Psychological testing*. Nueva York, NY: MacMillan Publishing Co., Inc. Quinta Edición.
- BABBIE, E. R. (1979). *The practice of social research*. Belmont, CA.: Wadsworth Publishing Co., Inc.
- CARMINES, E. G. y ZELLER, R. A. (1979). *Reliability and validity assessment*. Beverly Hills, CA: Sage Publications Inc. Series: Quantitative Applications in the Social Sciences, Vol. 17.
- CRONBACH, L. J. (1984). *Essentials of psychological testing*. Nueva York, NY: Gardner Press, Inc.
- KRIPPENDORFF, K. (1982). *Content analysis*. Beverly Hills, CA: Sage Publishing Co., Inc.
- MILLER, D. C. (1977). *Handbook of research design and social measurement*. Nueva York, NY: Longman, Inc. Tercera edición.
- NUNNALLY, J.C. (1970). *Introduction to psychological measurement*. Nueva York, NY:

McGraw-Hill, Inc.

THORNDIKE, R. L. y HACEN, E. (1980). *Tests y técnicas de medición en psicología y educación*. México, D.F.: Editorial Trillas.

WEBB, E. J.; CAMPBELL, D. T., y SCHWARTZ, R.D. (1966). *Unobtrusive measures: nonreactive research in the social sciences*. USA: Rand McNally College Publishing Co.

WIERSMA, W. (1986). *Research methods in education*. Newton, Mass: Allyn and Bacon, Inc. Cuarta edición, capítulo 11 ("Measurement and data collection").

EJEMPLO

La televisión y el niño

FIGURA 9.28

CUESTIONARIOS QUE SE ADMINISTRARON A LOS NIÑOS DEL DISTRITO FEDERAL PARA EL EJEMPLO DE LA TELEVISIÓN Y EL NIÑO (SE INCLUYE PARCIALMENTE EN VERSIÓN PRECODIFICADA)

1. ¿Cada cuándo lees un periódico?

- _____ casi nunca (0)
- _____ una vez al mes (1)
- _____ una vez cada semana (2)
- _____ 2 o 3 veces a la semana (3)
- _____ 4 o 5 veces a la semana (4)
- _____ diario (5)

2. ¿Cada cuándo vas al cine?

- _____ una vez por semana (4)
- _____ una cada dos semanas (3)
- _____ una vez al mes (2)
- _____ una vez cada 2 o 3 meses (1)
- _____ casi nunca (0)

3. ¿Cuánto tiempo oyes el radio cada día?

- _____ no oigo el radio (0)
- _____ menos de 1 hora al día (1)
- _____ de 1 a 2 horas al día (2)
- _____ de 2 a 3 horas al día (3)
- _____ más de 3 horas al día (4)

4. Durante la semana pasada, ¿cuántos cuentos de monitos has leído? (número)

5. Durante la semana pasada, ¿cuántas revistas has leído, sin tomar en cuenta los cuentos de monitos?

_____ (número)

6. Sin contar los libros que hayas leído para la escuela, ¿cuántos libros has leído durante el mes pasado?

_____ (número)

7. ¿Ves tele todos los días?

Sí _____ No _____

(1) (0)

8. En los días que vas a la escuela, ¿cuánto tiempo ves televisión?

- _____ no veo televisión (0)
- _____ menos de 1 hora (1)
- _____ 1 o 2 horas (2)
- _____ 3 o 4 horas (3)
- _____ 5 horas o más (4)

9. ¿Como cuánto tiempo ves televisión los sábados?

- _____ no veo televisión (0)
- _____ menos de 1 hora (1)
- _____ 1 o 2 horas (2)
- _____ 3 o 4 horas (3)
- _____ 5 horas o más (4)

10. ¿Cuánta televisión ves los domingos?

- _____ no veo televisión (0)
- _____ menos de 1 hora (1)
- _____ 1 o 2 horas (2)
- _____ 3 o 4 horas (3)
- _____ 5 horas o más (4)

11. ¿Qué programas de televisión viste ayer?

12. ¿Cuándo ves más televisión?

- _____ entre semana (1)
- _____ los fines de semana (2)
- _____ siempre (3)

13. ¿Cuándo te gusta ver más televisión?

- _____ en la tarde (1)
- _____ en la noche (2)

14. ¿Qué prefieres hacer cuando no estás en la escuela?

- _____ estar con tus papás (1)
- _____ jugar (2)
- _____ leer (3)
- _____ salir a la calle (4)
- _____ ver la televisión (5)

15. ¿Qué haces cuando vas a ver televisión?

- _____ prendo la televisión para ver lo que hay (1)
- _____ veo el teleguio o el periódico, para ver qué programas hay en la

televisión (2)

_____ ya me sé de memoria lo que hay en la televisión (3)

_____ prendo la televisión y veo lo que sea (4)

16. Mientras estás viendo la televisión, ¿qué es lo que generalmente haces?

_____ como, juego, dibujo o hago cualquier cosa (1)

_____ hago la tarea (2)

_____ veo varios programas a la vez, cambiando de canal (3)

17. De la siguiente lista de programas marca con una palomita los que tú ves, y además qué tanto te gustan.

Me gusta mucho

(2)

Me gusta un poco

(1)

No lo veo

(0)

Porky

Mi marciano favorito

Mundo de juguete

Variedades Vergel

El hombre nuclear

Noticiero Domecq

Operación convivencia

Fútbol

Películas

Los Picapiedra

Hechizada

Ven conmigo

La criada bien criada

Kojack

En punto

Plaza Sésamo

El Oso Ruperto

Locos Adams

Una muchacha llamada Milagros

Los polivoces

Viaje al fondo del mar

24 horas

Universo 5

Clásicos infantiles

Nany y el profesor

Barata de primavera

El chavo del 8

El llanero solitario

Deporteve

Platícame un libro

18. ¿Generalmente con quién ves la televisión?

_____ solo (1)

_____ con mi papá (2)

_____ con mi mamá (3)

_____ con mis hermanos (4)

_____ con mis primos (5)

_____ con mis amigos (6)

_____ con la sirvienta (7)

19. La mayoría de las veces, ¿quién escoge los programas?

- _____ yo (1)
- _____ papá (2)
- _____ mamá (3)
- _____ hermanos (4)
- _____ primos (5)
- _____ amigos (6)
- _____ sirvienta (7)

20. Marca con una palomita, si estás de acuerdo o no, con las siguientes frases. Mi papá o mi mamá me regañan cuando vea mucha televisión.

Sí _____ No _____
(1) (0)

A veces me castigan sin ver televisión.

Si _____ No _____
(1) (0)

MI papá o mi mamá me prohíben ver algunos programas de televisión.

Sí _____ No _____
(1) (0)

Yo me voy a dormir a la hora que quiero.

Sí _____ No _____
(1) (0)

21. Veo la televisión porque me río mucho. siempre (2)

- _____ a veces (1)
- _____ casi nunca (0)

22. Veo la televisión porque es muy divertida. siempre (2)

- _____ a veces (1)
- _____ casi nunca (0)

23. Veo la televisión porque se me pasa el tiempo volando. siempre (2)

- _____ a veces (1)
- _____ casi nunca (0)

24. Veo la televisión porque me entretiene mucho. siempre (2)

- _____ a veces (1)
- _____ casi nunca (0)

25. Veo la televisión porque nunca me aburro. siempre (2)

- _____ a veces (1)
- _____ casi nunca (0)

26. Cuando ves la televisión por las razones de arriba, ¿qué tanto te gusta?

- _____ muchísimo (4)

_____ mucho (3)
_____ regular (2)
_____ un poco (1)
_____ casi nada (0)