

Diseños cuantitativos: análisis e interpretación de la información

Miguel Aigner A.

Centro de Estudios de Opinión
Facultad de Ciencias Sociales y Humanas
Universidad de Antioquia
Medellín, Colombia 2006



1 8 0 3



Miguel Aigner Aburto
Centro de Estudios de Opinión -CEO-
ISBN: 958-655-958-0

Primera edición
febrero de 2006

Diseño de cubierta
Luisa Fernanda Bernal Bernal

Diseño, diagramación, impresión y terminación
Imprenta Universidad de Antioquia

Rector Universidad de Antioquia
Dr. Alberto Uribe Correa

Decana Facultad de Ciencias Sociales y Humanas
Dra. Luz Stella Correa Botero

Director Centro de Estudios de Opinión
Dr. Jaime Ruiz Restrepo

Jefe de Departamento de Sociología
Dr. Jhony Vahos Arias

Impreso y hecho en Colombia / Printed and made in Colombia
Prohibida la reproducción total o parcial de esta obra sin la autorización escrita
del autor y del Centro de Estudios de Opinión CEO, Universidad de Antioquia.

Universidad de Antioquia
Facultad de Ciencias Sociales y Humanas
Departamento de Sociología
Centro de Estudios de Opinión
Fondo editorial -CEO-
Teléfono: (574) 210 57 75. Telefax: (574) 210 57 75
E-mail: ceo@quimbaya.udea.edu.co
Página web: <http://ceo.udea.edu.co>
Apartado aéreo 1226. Medellín, Colombia

Imprenta Universidad de Antioquia
Teléfono: (574) 210 53 30. Telefax: (574) 210 53 32
E-mail: imprenta@quimbaya.udea.edu.co

Contenido

	Pág.
Introducción	9
Los diseños cuantitativos	13
Objetivos	13
Diseño de la investigación	14
El planteamiento de un problema de investigación	15
La formulación de hipótesis de investigación	17
El proceso de comprobación de hipótesis de investigación	18
Operacionalización de hipótesis de investigación	19
Ejercicio de autoevaluación	23
2. Tipos de diseños de investigación social	25
Objetivos	25
Introducción	25
Diseños experimentales	26
Diseños no-experimentales	26
La calidad interna de la información: la validez y confiabilidad de los datos	27
Ejercicio de autoevaluación	31
3. El muestreo	32
Objetivo	32
Determinación de la muestra	32
Tipos de muestreo aleatorio o probabilístico	33
Muestreo no aleatorio - no probabilística	40

4. La recolección de la información	41
Objetivo	41
Introducción	41
La recolección de la información	41
Conceptos básicos de la recolección de información	42
Técnicas de recolección de la información	45
Medición escalar	45
La entrevista social	51
La encuesta social	51
Ejercicio de auto-evaluación	65
5. Análisis de datos cuantitativos	67
Introducción	67
Objetivos	67
La medición en los datos cuantitativos	67
Los niveles de medición y los estadísticos apropiados a cada nivel	67
Sistematización de la información	73
Procesamiento de la información	73
6. Análisis de la información	78
Introducción	78
Objetivos	78
Procedimientos para la sistematización de la información	79
Manejo del editor de datos SPSS®	80
Análisis descriptivo de los datos	84
Técnicas de representaciones gráficas	111
Interrelación entre variables	113
Inferencia estadística	145
Medidas de relaciones entre variables	166
Análisis multivariado	185

Regresión múltiple	187
Análisis factorial	192
Glosario	202
Bibliografía	208
Anexos: archivos de trabajo	211

Introducción

En el campo de la investigación social generalmente se habla de diseños cuantitativos y cualitativos. Aquí nos ocuparemos de los diseños cuantitativos, los cuales son una forma específica de aproximación sistemática al estudio de la realidad, los cuales se apoyan principalmente en categorías numéricas que utilizan como base una información factible de cuantificar. Desde esta perspectiva, el análisis se realiza a través de diferentes formas de interrelación de los datos cuantificados y en especial, tomando en cuenta las tendencias que muestra el comportamiento de ellos. En otros términos, este texto hace énfasis en el proceso investigativo que prioriza el uso de una información cuantificada, desde el planteamiento del problema hasta el análisis de la información.

El modelo pedagógico tradicional para la enseñanza de la metodología de la investigación no integra al proceso concreto, acciones y ejercicios que permitan ver en forma práctica el uso de herramientas como la estadística y la informática, las que vienen convirtiéndose día a día en auxiliares importantes de la investigación social, aspecto que pretendemos subsanar a lo largo de este texto que es una combinación de un libro y un programa informático que se complementan el uno al otro.

Debido a lo anterior, se busca en primer lugar brindar una visión técnica y práctica del proceso de investigación social, que se apoya preferentemente en los diseños cuantitativos tratando de acercarnos a la práctica real del investigador, para así responder a la tendencia vigente del entorno, al énfasis profesional y a la especialización, lo que lleva a mejorar la eficacia y eficiencia del desempeño profesional.

En segundo lugar, pretende presentar el amplio panorama de los diseños cuantitativos como una variedad de estrategias que se adaptan, en forma creativa, imaginativa y sagaz, a los distintos objetos de investigación, de tal manera que los llamados *problemas teóricos-metodológicos, técnicos y empíricos*, no obstaculicen su aplicación. También se insiste en la tendencia de combinar diversos diseños metodológicos y la necesaria complementariedad entre teorías, metodologías y técnicas, dada la complejidad del estudio de lo social.

Por tal motivo, se reitera la necesidad de familiarizar al investigador social con una estrategia investigativa que no sólo evite, por un lado, las dificultades metodológicas de los diseños cuantitativos, así como la unilateralidad del momento cuantitativo.

De este modo se desea acercar a los estudiantes a las herramientas estadísticas e informáticas aplicadas a la investigación social, para que puedan utilizarlas en el análisis de los datos más representativos, permitiendo la lectura e interpretación de los resultados obtenidos.

Este texto está conformado por seis capítulos cuyo contenido es el siguiente:

- El primero hace una introducción general a los diseños cuantitativos, brindando los elementos necesarios para el abordaje de los capítulos siguientes.
- El capítulo dos presenta los diferentes tipos de diseños de investigación utilizados en la recolección y medición de información cuantitativa.
- El capítulo tres presenta una introducción a los diseños muestrales.
- El capítulo cuatro expone algunos conceptos y técnicas básicas en la recolección de datos, haciendo énfasis en el uso de la técnica de la encuesta social.
- El capítulo cinco da cuenta del proceso de medición en los datos cuantitativos y la sistematización de la información.
- El capítulo sexto trata del análisis de datos cuantitativos, aquí se sintetiza lo que se debe conocer para poder seleccionar la prueba estadística apropiada para cada tipo y problema de investigación. También se enfatiza en el manejo de la información, para su aplicación en procesos de sistematización, utilizando los instrumentos que facilita el programa SPSS®¹ y a la vez se revisan algunas herramientas estadísticas.

Los contenidos de cada capítulo van acompañados de *ejemplos y de autoevaluaciones*, que ayudan al lector en la comprensión de temas tratados, así como en la identificación del grado de adquisición de su aprendizaje. Es igualmente importante tener en cuenta los *Procedimientos SPSS®* que detallan paso a paso como se utilizan las herramientas que ofrece este programa de manejo y gestión de datos cuantificados.

1 SPSS® es un programa de archivo y gestión de datos numéricos que integra procedimientos estadísticos con gráficos, asistiendo al usuario en la manipulación, organización y archivo mediante menús y submenús.

Aunque este texto es el resultado de muchos años de trabajo personal en el campo de la investigación social, tanto en la docencia como en la práctica, no habría sido posible llevarla a cabo si durante este tiempo no hubiera contado con inestimables ayudas. Quiero hacer constar, en primer lugar, mi reconocimiento a un gran profesional de la investigación social el profesor Jaime Ruiz Restrepo, con quién realice mis primeras aportaciones al diseño y análisis de datos de encuestas. Durante los últimos quince años he sido parte del Centro de Estudios de Opinión CEO que dirige él y hemos tenido la ocasión de desarrollar y aplicar buena parte de los conocimientos de sistematización y análisis de datos que se incluyen en este texto.

Son muchas las personas individuales a las que debo ideas, comentarios y críticas y lamento no poder nombrarlas a todas.

Finalmente, no quiero dejar de agradecer la confianza y aporte que el Fondo Editorial del Centro de Estudios de Opinión de la Universidad de Antioquia puso en mi persona, al permitirme desarrollar un ambicioso y pionero proyecto editorial que puede mostrar actualmente una colección de valioso material docente.

1. Los diseños cuantitativos

Objetivos

1. Actualizar los conocimientos y competencias en los campos de la metodología y técnicas de investigación social.
2. Analizar casos de investigación con ejemplos para el uso de las metodologías y técnicas propuestas.

El punto de partida de la investigación social es, desde luego, la existencia de un problema que requiere alguna forma de respuesta, de explicación o de solución. El propósito final de este proceso es la respuesta esperada, tanto desde el punto de vista de la formalidad teórica como desde la metodológica. Para obtener la respuesta a ese problema, el investigador recurre a métodos y técnicas adecuadas e históricamente probadas.

Sin embargo, la aplicación de métodos y técnicas específicas en la investigación social, no se hace de manera rígida o mecánica, de tal modo que la creación de conocimientos resulte, de la sumisión estricta a un conjunto de reglas de procesamiento preestablecidas e inmodificables. La investigación social es un proceso donde la creatividad, el compromiso del analista, la imaginación, las ideas, la argumentación, la perspicacia y la sagacidad son instrumentos válidos y confiables. Toda *formalización metodológica* tiene sólo un carácter instrumental, indicativo, protocolario y provisional, que facilita el desarrollo del proyecto de investigación.

Considerar la metodología, como una simple lista de pasos o etapas estandarizadas, sería ir en contra de la esencia misma de la metodología científica, porque ésta no es un conjunto de recomendaciones que se puedan aplicar siempre.

El investigador debe estar atento, de manera continua, a los posibles errores y desviaciones que pueden conducirlo al uso indiscriminado de algunas

técnicas de investigación. En este caso, una sólida formación metodológica, le permitirá una selección adecuada, de algunos de estos procedimientos o la limitación de su aplicación a ciertos temas y aspectos de la realidad social.

Así, la metodología que se utiliza para realizar un trabajo investigativo es algo que se necesita comprender y aprender. Requiere de cierto conocimiento de la realidad en estudio; cierto *capital de base* que le permita al analista plantearse inquietudes y posibles respuestas sobre el tema, porque no existe *un conjunto formalizado único de protocolos o normas* que se puedan aplicar mecánicamente a cualquier situación problemática.

El término *método-científico*, es entendido aquí como un concepto general que comprende diversas maneras de abordar un problema, de conocer y de acercarse a la realidad, guiadas por principios o modelos teórico-metodológicos que permiten considerar los resultados logrados como conocimiento científico. Sin embargo, existen elementos específicos comunes a toda investigación, especialmente en los llamados diseños cuantitativos, que ayudan a pensar y planear mejor, tales como *los aspectos científicos técnicos de un proyecto* – ¿Qué hacer y cómo hacerlo? – y los aspectos de administración y control del proyecto – estrategia operativa.

Dada la diferencia entre el accionar real del investigador y la exposición didáctica de la misma, es posible distinguir diversas etapas en el proceso de investigación.

Diseño de la investigación

En todo proceso investigativo se debe definir y diseñar un programa que sirva de guía para el desarrollo del estudio. El diseño de investigación implica no sólo plantear el problema, además se debe elaborar un plan para la fase operativa que tiene que ver con la especificación de la metodología y de las técnicas para recolectar y validar la información, necesaria para darle una respuesta al estudio.

El plan operativo de la investigación comprende un conjunto de actividades como:

- Identificación, análisis y formulación del problema a investigar.
- Determinación del tipo de diseño de investigación a utilizar.

- Especificación de la hipótesis u objetivos del estudio.
- Diseño de *la matriz de variables-dimensiones*: definición, clasificación y operacionalización de las variables del estudio.
- Selección de las fuentes de información y diseño de la muestra.
- Diseño del instrumento para la recolección de la información.
- Recolección y sistematización de los datos.
- Análisis e interpretación de los resultados.
- Elaboración del informe final.

El planteamiento de un problema de investigación

Un problema de investigación es un proceso, hecho o situación a la que se le quiere encontrar respuesta o solución. En algún momento de la investigación es posible que el analista social se vea enfrentado a dos situaciones. La primera tiene que ver con el discurso pedagógico, el cual insinúa que el proceso de investigación sigue un camino unidireccional – punto de partida y punto de llegada –; y la segunda, es aquella donde la realidad lo lleva de lo abstracto a lo concreto, o de adelante hacia atrás y viceversa; es decir, el mismo proceso va sugiriendo los pasos a seguir.

¿Cuál es el problema de investigación? En la realidad esas diferentes concepciones o suposiciones acerca de cuál es el problema, cuál es su tratamiento teórico- metodológico, cuál es la respuesta, son expresiones de diferentes puntos de vista teóricos, metodológicos e ideológicos que se evidencian en los llamados paradigmas de investigación social.

El problema emerge de un tema global mediante un proceso deductivo. Una primera aproximación lo lleva a identificar y aislar la situación problemática con el fin de especificar qué se va concretamente investigar. Obsérvese el siguiente ejemplo:

El director de bienestar estudiantil en una universidad pública, está interesado en conocer el grado de aceptación, de los programas culturales por parte de la comunidad estudiantil. Este es un problema de investigación identificado y delimitado dentro de una temática global y se tiene conciencia de la presencia de un vacío específico de conocimiento.

En otras palabras, se identifica un problema cuando se le somete a un *proceso de análisis*. Aquí lo que hace el investigador es plantearse una serie de interrogantes, como los siguientes:

En cuáles circunstancias aparece el problema; qué elementos básicos lo componen; [...] qué interrelaciones existen entre los elementos componentes: cuáles elementos dependen de otros; qué mediaciones existen en esas relaciones; qué aspectos del problema se conocen; cuáles están identificados; cuáles faltan por hacerlo; qué tipo de explicaciones a partir de modelos conceptuales nos permiten tener una mejor comprensión de ellos [...] es decir, se trata de desmenuzar el problema²

Contextualizado el problema de investigación, delimitados y examinados sus elementos o variables, y definida la forma como se interrelacionan, el paso siguiente es *su formulación*, es decir, se debe precisar qué es lo que se va estudiar: tanto los elementos, características o variables que en él interviene, como las relaciones existentes entre ellas y que están a la espera de una explicación cierta y comprobable.

El planteamiento del problema de investigación es una actividad fundamentalmente práctica, puesto que sólo se aprende haciéndolo, teniendo en cuenta que el proceso comienza a partir de la identificación de un problema de investigación, cuestión que requiere no solamente de la capacidad de *ver* los problemas y de *la actitud* de respuesta o solución, sino que además, hay que sumarle una cierta capacidad de imaginación, sentido común y manejo creativo de esa realidad que se quiere escudriñar.

En el planteamiento de un problema se concretan básicamente dos situaciones. La primera, aquella idea que comprende un esquema o modelo que se debe investigar (objeto general de estudio) y la forma de *interpretar* los hechos, (aunque en muchas ocasiones tales *esquemas o modelos* no se explicitan); y la segunda, *una metodología* – procedimientos, protocolos y técnicas – que van a permitir al investigador encontrar una respuesta.

Al investigador se le pueden presentar diversas alternativas teórico-metodológicas dentro de las cuales es posible encontrar la ansiada respuesta. Surgirán entonces, varias preguntas a responder. Ejemplo: en el problema planteado acerca del grado de aceptación de los programas culturales por parte de la comunidad estudiantil, cabe preguntarse:

2 Facundo D. Ángel H. El proyecto de investigación. Serie aprender a investigar. Módulo 5. Bogotá: Procesos Editoriales ICFES. 1987. Pág. 22.

- ¿A qué tipo de programas culturales se están refiriendo?
- ¿Qué tipo de información se requiere para responder a lo planteado?
- ¿Qué se entiende por grado de aceptación?
- ¿Es un estudio comparativo? Por ejemplo, respecto al grado de aceptación en años anteriores.
- ¿Es un estudio evaluativo?
- ¿Es sólo un estudio descriptivo?

Cada una de estas preguntas requiere de diferentes diseños metodológicos y distintos tipos de información. Igualmente, antes de recolectar la información se deberá diseñar de acuerdo con la técnica seleccionada, una guía de entrevista, una guía de observación o un cuestionario. También habrá que planificar el proceso de recolección y análisis de la información y por último, analizar e interpretar los resultados a la luz de un marco teórico.

La formulación de las hipótesis de investigación

Toda investigación se orienta por hipótesis que hacen el papel de guías de trabajo, las cuales pueden ser explícitas o implícitas.

Las principales funciones de una hipótesis son:

- Conducir la investigación: La hipótesis [...] guía toda su marcha, desde la caracterización de la población a la cual se refiere y necesita para su contratación, la información a recoger, el manejo del trabajo de campo y las técnicas adecuadas para el análisis de los datos.³
- Especificar cuáles hechos se observan y qué relación se establecerá entre ellos.
- Ofrecer una suposición o explicación tentativa al problema de investigación.

3 Briones, Guillermo. Métodos y técnicas avanzadas de investigación aplicadas a la educación y a las ciencias sociales. Módulo 4. Análisis e interpretación de datos cuantitativos. Bogotá: Corporación Editorial Universitaria de Colombia. 1989. Pág. 54

Se supone que se ha identificado un problema de investigación como el que a continuación se presenta: Es conveniente desmontar el sentido de pertenencia y arraigo sobre aspectos físicos y más bien, resaltar la necesidad de sondear elementos que son posibles de incorporar para restablecer o crear formas de convivencia ciudadana con el fin de lograr una mejor ciudad para vivir. Bajo esta perspectiva, los procesos demográficos, los cambios en la organización familiar, en el sistema educativo [...] en la cotidianidad, configuran el universo de factores que en la ciudad de Medellín asumen perfiles diferentes, no sólo por determinaciones socioeconómicas y socioespaciales, sino ante todo por la perceptible incoherencia de los patrones y valores de comportamiento y cultura que encarna cada generación según el modelo y la temporalidad inscrita.⁴

Es posible apreciar cómo en el mismo planteamiento del problema se enuncian explícitamente algunas respuestas, se hacen algunas afirmaciones o suposiciones tentativas, es decir, una hipótesis necesaria para comprobar a través de un proceso investigativo.

Antes de formular una hipótesis es necesario reunir suficientes antecedentes y conclusiones teóricas, comparar un número tan grande como sea posible de información asociada, disponer para ello de la documentación pertinente; estableciendo así la hipótesis que requiere seguir el camino de lo conocido a lo supuesto.

El proceso de comprobación de una hipótesis de investigación

Para que una hipótesis sea comprobable debe reunir las siguientes características:

- Ser conceptualmente operacionalizable, es decir, debe tener un referente empírico observable.
- Estar relacionada con un marco de referencia teórico-conceptual, que facilite el análisis e interpretación de los datos, que van a permitir comprobar o refutar la hipótesis⁵.

4 Ruiz, Jaime; Aignerren, Miguel; Ochoa, Jaime; De Los Ríos, Héctor; Parra, Carlos. Imágenes de identidad, integración y conflicto entre la comunidad estudiantil de la Universidad de Antioquia. Medellín: Universidad de Antioquia. 1991. Págs. 24-25

5 En los diseños cuantitativos, generalmente, cuando se habla de comprobar hipótesis significa someterla a prueba por medio de observaciones.

– Permitir el uso de técnicas que faciliten las observaciones.

En un estudio acerca de un programa de seguridad comunitaria⁶, la hipótesis es que la violencia responde a una doble ruptura: la primera es aquella entre la clase política, en cuyo seno buscan replegarse los dirigentes comunitarios reivindicativos, y la vida comunitaria que cotidianamente rechaza a los políticos tradicionales; la segunda es la ruptura entre los barrios populares, donde están confinados los sectores de menores ingresos urbanos, y el resto de la sociedad, cuya cohesión parece concertarse cada vez más sobre la base de esta discriminación geográfica y social. Además, hay una noción implícita de un factor causal y de una consecuencia: los conceptos variable dependiente y variable independiente reflejan esta situación.

Por *variable* entendemos aquellos enunciados formados por significados teóricos que permiten transmitir ideas referidas a características de sujetos o hechos que de alguna forma cambian. Por variable independiente entenderemos esa propiedad, esa situación o hecho que se supone es el causante o explica el efecto señalado. Por variable dependiente entenderemos los efectos o la consecuencia de un hecho o proceso.

En el ejemplo dado, los factores causales o determinantes –variable independiente– serían la ruptura entre la clase política y la vida comunitaria y la segregación geográfica y socioeconómica. La consecuencia –variable dependiente– sería un antimovimiento social, cuya expresión más visible es la violencia juvenil.

El analista social debe demostrar si la relación enunciada en la hipótesis es válida o inválida. En la investigación cuantitativa, las hipótesis no siempre son enunciadas explícitamente. Aún así, es necesario que la pregunta central del problema de investigación este presente, aunque sea sólo en la mente del analista, con el fin de no perder la orientación de lo que se desea.

Operacionalización de las hipótesis de investigación

Un problema común en un enunciado hipotético es la existencia de conceptos que pueden ser entendidos de diversas formas por los interlocutores,

6 Tironi, Eugenio. Pobladores e integración social. Propositiones 14. Santiago de Chile: Ediciones Sur. Pág. 64. Agosto 1997

lo cual va a determinar graves sesgos en la información recolectada. Existen otros conceptos que por su abstracción son de difícil operacionalización, como por ejemplo: *autoridad, tranquilidad, felicidad, ansiedad, eficiencia, miedo*, etc.

Con el fin de probar una hipótesis, se hace necesario no sólo tener un conocimiento del problema a investigar, sino también el marco de referencia teórico-conceptual que lo define. En el ejercicio sobre la relación entre rupturas políticas, segregación socio espacial y violencia barrial (medida por índices de conductas delincuenciales), implícitamente se encuentra una suposición en el sentido de que la reconstrucción del tejido social y político permitiría erradicar a los antisociales y sus manifestaciones delincuenciales. Esta afirmación parece ser explícita, pero en la práctica no es un planteamiento operacional: habría que definir operacionalmente los conceptos: *ruptura política, segregación socio espacial, erradicación, conductas delincuenciales*, de tal forma que permitan ser traducidos en referentes empíricos observables en el proceso de diseño del instrumento para la recolección de la información.

Definición, tipos y medida de las variables

Una variable es una característica o propiedad de un sujeto, hecho social o situación, cuyos valores, que son objeto de estudio, pueden cambiar de una unidad de análisis a otra. Las variables pueden asumir diferentes *valores*, lo que se puede observar en el siguiente ejemplo:

Variables	Valores
Nivel o grado de escolaridad	Primaria incompleta
	Primaria completa
	Secundaria incompleta
	Secundaria completa
	Educación superior incompleta
	Educación superior completa

Según *la relación entre causa y efecto*, las variables pueden dividirse, por el papel que desempeñen en la relación causal, en independientes, dependientes e intervinientes:

- *Variable independiente*. Es una variable que se utiliza para explicar el comportamiento de otra que pasa a ser dependiente. Se denomina también variable

explicativa o predictora. Por ejemplo, el nivel de ingresos puede ser explicado por el nivel de escolaridad, que en este caso es la variable independiente.

- *Variable dependiente*. Es aquella cuyo comportamiento es explicado por otra u otras variables independientes. Se la denominaba también *variable a explicar*. Por ejemplo, el comportamiento lector (variable dependiente) puede ser explicado por la edad del individuo.
- *Variable interviniente*. Supone que la relación entre dos variables está mediada por otra u otras que determinan o influyen, si es el caso, los efectos o las características de las primeras. Cabe preguntarse cómo decidir ¿cuál es la variable interviniente?, teniendo en cuenta que es evidente que en cada hipótesis se encuentran potencialmente muchas variables que pueden intervenir. La respuesta está en la pregunta acerca de cuáles factores, (de acuerdo al criterio teórico-metodológico del analista), están más relacionados, tanto con la variable independiente como con la dependiente; por ejemplo, el comportamiento lector puede ser no solamente explicado por la edad del individuo, sino además por el nivel de escolaridad.

En función de los valores cuantitativos que se pueden adoptar, las variables se clasifican en:

- *Variable continua*. Es aquella que puede tomar cualquier valor numérico; por ejemplo, el monto de los salarios de una población en estudio o el rendimiento académico expresado en promedios.
- *Variable discreta*. Es aquella que sólo puede tomar un número finito de valores distintos, por ejemplo, el número de hijos.
- *Variable dicotómica*. Aquella que sólo puede tomar dos valores; por ejemplo, el sexo (hombre - mujer), una afirmación o negación (sí - no), la evaluación de un test (aprobado - reprobado) o una apreciación o conocimiento (verdadero - falso).

Identificación de las unidades de análisis

En los diseños cuantitativos, la unidad de análisis es un dato — observable y externo — que se convierte en la unidad de información básica. Está asociado al objeto de estudio, por ejemplo: individuos, grupos, regiones, hechos y procesos. La identificación de las unidades de análisis es un paso importante

en el planteamiento del problema de investigación. Al respecto se recomienda ser cuidadoso en cuanto a su naturaleza; no siempre en la investigación social las unidades están definidas en términos de individuos o conjuntos de individuos. Las unidades pueden definirse en términos de los productos de los individuos, sus comportamientos o pueden ser de entes geográficos o físicos. Además también son unidades de análisis: los artefactos o productos del hombre tales como periódicos, automóviles y viviendas, entre otros, así como países, regiones, ciudades, comunas, barrios, etc. Cada una de esas unidades supone la existencia de una población o universo de tales objetos y cada unidad puede ser descrita, clasificada o caracterizada en términos de las dimensiones significativas para el estudio.

Observar el siguiente ejemplo

Problema de investigación: la deserción académica estudiantil en los primeros semestres en la Universidad de Antioquia.

Antecedentes: Uno de los mayores problemas que afecta la educación superior en Colombia, es la deserción académica de los estudiantes universitarios en sus primeros semestres de estudios. Entre las posibles causas de la llamada *mortalidad académica* están: la falta de orientación profesional, el bajo nivel de la educación secundaria, los bajos recursos económicos de las familias o el bajo rendimiento académico en los primeros semestres universitarios. Muchos de estos estudiantes que desertan de la Universidad se vinculan al mercado laboral en empleos mal remunerados, o se inscriben en institutos tecnológicos en donde hacen una carrera tecnológica, a la cual debieron haber llegado directamente, sin la experiencia frustrante de la universidad.

Determinar la unidad de análisis, las variables en juego y los valores que asumen:

1. La unidad de análisis (sujeto-objeto) del problema de investigación son los estudiantes universitarios en sus primeros semestres académicos de estudio.
2. Las variables son: la mortalidad académica estudiantil es un problema que está determinado, entre otros, por los siguientes factores que pueden ser explicados por las variables independientes: la falta de orientación profesional, la calidad de la enseñanza secundaria, los recursos económicos de las familias y el bajo rendimiento académico en los primeros semestres universitarios.
3. Valores de la variable recursos económicos familiares: las familias de los estudiantes que ingresan a la universidad, podrían dividirse en familias de recursos económicos: bajos, medio-bajos; medios; medio-altos y altos.

Ejercicio de auto-evaluación

En un estudio sobre *los programas de seguridad barrial*, se hace referencia a las políticas y las circunstancias socioeconómicas, más determinantes en la contención de la violencia barrial. Es posible suponer que hay otros factores socioeconómicos asociados al proceso que se está analizando y que pueden tener un impacto independiente de la existencia o no de programas de seguridad comunitaria, tales como el desempleo en grupos de jóvenes. Esta clase de factores intervinientes deben ser tenidos en cuenta cuando se planea el diseño de la investigación, aun cuando se decida que ellos no son lo suficientemente importantes para ser incluidos en el análisis. Igualmente, factores tales como los culturales e ideológicos, que pueden ser un reflejo de opiniones, actitudes y conductas formadas por los medios de comunicación, deben ser tenidos en cuenta en los resultados esperados.

Desarrollo: Se recomienda a partir del tema objeto de estudio sobre *los programas de seguridad barrial*, u otro problema de investigación que le parezca más afín, desarrollar y enriquecer una matriz de variables, a partir de la experiencia que Uds. tengan sobre el problema de investigación. En síntesis, esta recomendación rescata el aporte de los participantes en la investigación y en ese sentido son las sugerencias metodológicas; la idea no es limitarse a responderlas de manera concisa de acuerdo a lo señalado en los ejemplos del texto, es decir, no se trata de responder unas líneas sobre ¿cuál es la variable independiente, dependiente o cuál es la interviniente? sino que se desarrolle el texto, atendiendo a la multicausalidad de lo social, una matriz de variables⁷ pertinentes al estudio reseñado sobre los programas de seguridad barrial o formulando otro problema que parezca más cercano a sus desempeños profesionales, siempre tendiendo en perspectiva el diseño de un instrumento de recolección de información *válido y confiable*.

Por ejemplo, de la siguiente hipótesis del estudio reseñado, que hace referencia a cómo las formas no tradicionales de solidaridad barrial, han

7 Diseño de una matriz de variables – conformada por conceptos, variables e indicadores que hipotéticamente *miden* el objeto de estudio.

aminorado ciertas conductas delincuenciales, señale cuáles son las variables independientes, las dependientes y las intervinientes.

Variable (s) independiente (s)

Variable (s) dependiente (s)

Mencione algunas variables intervinientes o de control que permitirían *poner a prueba* la consistencia de la relación original enunciada en la primera hipótesis.

Variable interviniente (s)

Variable interviniente (s)

Variable interviniente (s)

2. Tipos de diseños de investigación social

Objetivo

Familiarizar al estudiante con algunos de los tipos de diseños de investigación social.

Introducción

En términos metodológicos, se plantean diseños experimentales y no experimentales. Hay elementos que diferencian los diseños de tipo experimental o casi experimental de los denominados *no experimentales*. En los primeros genera una situación artificial y controlable a diferencia de los no experimentales.

La elección de un tipo de diseño u otro depende del problema a resolver, de los objetivos y del contexto que rodea el trabajo de investigación. Éstos no sólo pueden combinarse sino también complementarse.

Generalmente existen dos clases de factores que inciden sobre la validez de los resultados obtenidos, en ambos tipos de diseños de investigación. En primer lugar los factores históricos, los cuales hacen referencia a sucesos o procesos acaecidos durante el tiempo del experimento, y a la vez pueden haber tenido cierto grado de influencia en ese último; el segundo, es el efecto del factor instrumental, el cual debe controlarse en los estudios de series de tiempo, en el sentido en que la dimensión conceptual, al ser operacionalizada en un proceso de medición, realmente mida en todas las circunstancias, lo que se quiere medir.

Es necesario dejar claro que estos criterios de clasificación de los diferentes tipos de diseños, sólo tiene una justificación pedagógica. La práctica real de la

investigación muestra generalmente inteligentes combinaciones de diseños, que responden a los objetivos e intereses del investigador y no a la sujeción a protocolos metodológicos absolutos.

Diseños experimentales

Estos diseños se utilizan generalmente para estudiar conductas sociales, y busca controlar el impacto de factores externos, en una relación hipotética de dos grupos seleccionados al azar, uno *experimental* y otro *de control*, midiendo esa proyección en espacios temporales diferentes. En este tipo de diseño existe un control más riguroso, se analizan relaciones experimentales o controladas y se pueden *manipular* las variables por separado o conjuntamente con otras para conocer sus efectos. Los principales diseños experimentales son:

Diseño con un grupo de control no equivalente. En el cual los grupos experimentales y de control no han sido formados al azar; como por ejemplo, los funcionarios de la oficina de planeación y catastro municipal. El control de variables extrañas al experimento se hace mediante la elección de un grupo, que tenga el mayor grado posible de características similares con el grupo experimental.

Diseño de antes y después con grupo de control. En este diseño se busca seleccionar intencionalmente a los grupos experimentales y de control, con el criterio de reunir, en términos generales, las características del grupo experimental.

Diseño de antes y después. Este tipo de diseño señala la necesidad de comparar conocimientos, actitudes o comportamientos sociales, teniendo como referente una información previa, puesto que es evidente que un diseño descriptivo sería incapaz de evaluar, por ejemplo, el grado de aceptación de una nueva reglamentación estudiantil sin tener un referente.

Post-test con sólo comparación. En esta clase de diseño los sujetos del estudio son comparados con un grupo de control después de haber participado en un programa. Obviamente no hay pre-test. Los estudios de opinión son un típico ejemplo del diseño de post-test.

Diseños no experimentales

En la práctica investigativa de lo social, es difícil lograr un control de todos los factores internos y externos que pueden afectar una información. En

aquellas situaciones en las que no es factible el control y manipulación de las variables, se encuentran otros diseños metodológicos alternativos, como los diseños no experimentales. Entre estos podemos mencionar:

Diseños descriptivos. Conocidos también como diseños transeccionales. Tienen como objetivo describir características de una población sujeto de estudio y analizar su incidencia o interrelación en un momento dado. Son una fotografía de un individuo o colectivo.

Diseños correlacionales. Tienen como objetivo describir las relaciones de unos factores o variables con otros, considerados como determinantes; es decir, predicen la relación entre dos variables, ninguna de las cuales es controlada o manipulada por el investigador. Se caracterizan porque permiten identificar, medir e interrelacionar múltiples variables simultáneamente, no importando establecer relaciones de causa-efecto entre ellas.

Diseños longitudinales. En algunas circunstancias, el interés del analista es observar cambios a través del tiempo de ciertas variables o en las relaciones entre éstas. Los diseños longitudinales recolectan información en períodos específicos para hacer inferencias respecto “al cambio, sus determinaciones y consecuencias”¹. Existen tres tipos de diseños longitudinales: de tendencia, evolución, de grupo y panel. Se diferencian por las características de las unidades de análisis – grupos específicos o población en general. Un buen ejemplo de estos diseños son los sondeos electorales, tipo panel, que siguen las preferencias electorales a través del tiempo.

Los diseños de panel, también llamados *series de tiempo*, en los cuales no se requiere de un grupo de control, consisten en una serie de mediciones periódicas que se hacen a las unidades de análisis del estudio, antes y después que se aplica el tratamiento. En este tipo de diseños se puede intentar un cálculo en la influencia de los factores externos mediante un sistema de medidas a lo largo del estudio.

La calidad interna de la información: la validez y confiabilidad de los datos

Por *validez interna* se considera, el grado en que las mediciones empíricas reflejan la dimensión conceptual que se quiere medir y por *confiabilidad* se

1 Hernández S., Roberto; Fernández C. Carlos; Baptista L. Pilar. Metodología de la investigación. México: McGraw Hill. 1992. Pág. 196.

entiende la consistencia de los puntajes o la ausencia de errores en mediciones sucesivas. La validez y confiabilidad de un diseño cuantitativo se fundamenta en la posibilidad de controlar o tener en cuenta los siguientes factores:

Maduración: Un diseño investigativo cuya unidad de análisis son escolares que mental y biológicamente van desarrollándose, deberá tomar en cuenta que independiente de cualquier tipo de programa didáctico, los escolares van incrementando sus habilidades de todo tipo, si esto no se considera se determinará un sesgo en la medición.

Selección: Así el grupo de estudio sea homogéneo en edad y nivel de escolaridad, no lo es en términos individuales, en el sentido de los procesos diferenciados de socialización familiar y profesional. Es dado pensar que hijos de profesionales asociados a las áreas del conocimiento en las ciencias exactas tendrán habilidades mayores o más desarrolladas. Esta particular diferenciación sesgará los puntajes, independientemente del programa didáctico.

Mortalidad: Es de creer que en un grupo como el del estudio señalado anteriormente, habrá mortalidad y deserción académica, lo cual sesgará los resultados en la prueba del post-test.

Repetición del cuestionario: Es posible que los sujetos del estudio logren cierta habilidad en responder los cuestionarios diseñados para la recolección de la información, dada la familiarización con él al serle aplicado los pre-test, lo cual incidirá en los resultados del post-test. Es posible que los puntajes de medición tienda a subir como producto de ciertas características del cuestionario del post-test que facilitan las respuestas, o porque la habituación de los participantes del experimento con los asistentes de la investigación, cree una situación que influya en las respuestas acertadas.

Historicidad: También es posible que durante el período de tiempo que ha transcurrido en el estudio, influencias externas a la escuela incentiven el desarrollo de “habilidades matemáticas” en los escolares; por ejemplo, que en un canal cultural de televisión presenten programas dedicados a desarrollar esas habilidades que se están midiendo con el experimento.

En el siguiente ejemplo se evidencia un problema con la calidad interna de la información.

La Secretaría de Educación Municipal, está interesada en mejorar el rendimiento académico de los estudiantes de los establecimientos educativos

de enseñanza secundaria mediante el uso de programas pedagógicos innovadores. Se ha propuesto evaluar, por medio de un estudio, un programa académico asistido por técnicas informáticas y se ha llegado a la conclusión de que el desarrollo de la habilidad matemática, de los estudiantes asistidos mediante programas tutoriales informáticos, es mayor que la de aquellos que no han tenido acceso a ellos.

Se asume, para los efectos de este ejemplo, que el programa tutorial es didáctico, que se ha definido operacionalmente el concepto *habilidad matemática*, y que se cuenta con un cuestionario que logra *medir* esa habilidad.

Desde el punto de vista metodológico del diseño de la investigación, se decide poner a prueba la habilidad matemática del grupo de estudio, antes y después de haberlo sometido al experimento pedagógico, con el fin de intentar medir su desarrollo en habilidades matemáticas. Al final del programa, ambos grupos serán sometidos a una prueba post-test. Es de suponer que el grupo experimental puede ganar en puntajes, comparado con los obtenidos en *el pre-test*, y que la ganancia-aumento del grupo de control tendrá un incremento menor.

Lamentablemente, esta sencilla estrategia investigativa, lleva a que los resultados que se encuentren no tengan mayor validez. Y esto puede explicarse por lo que los metodólogos denominan la no confiabilidad en la calidad interna de la información².

—*Criterios metodológicos de validación de la información*

Para superar posibles problemas con la validez y confiabilidad, se recomienda no sólo tener en cuenta los factores que se enunciaron, sino, también asegurar algunos criterios metodológicos de control, como los que se dan en ejemplo a continuación:

- Definir un listado de posibles participantes en el experimento, teniendo como base características sociodemográficas homogéneas, tales como: la edad, el lugar de residencia, el estrato social, etc.
- Seleccionar un grupo de estudio mediante una técnica de muestreo. Se divide el grupo: la mitad integrara el grupo experimental y la otra, el grupo de control. A cada uno de los participantes de cada grupo se le aplicará

2 La calidad interna de la información depende de la confiabilidad y de la validez.

un pre-test para medir en ese momento sus habilidades matemáticas. Es posible que los resultados sean relativamente similares en cada grupo.

Teniendo presente estos factores internos y controlándolos, se lograrán resultados más validos. Pero el problema no se reduce sólo a estos factores, sino que también tienen peso los llamados *factores de calidad externa*, La calidad externa de la información depende de la representatividad, la cual tiene que ver con el grado en que los resultados, a partir de la muestra, pueden ser atribuidos a la población en general. Así los factores de calidad externa se refieren a la dificultad de generalizar los resultados, producto de una situación experimental controlada, a otro segmento de la población, sin tomar algunos criterios de control.

Ejercicio de auto-evaluación

1. Supóngase que usted es un analista social que ha sido contratado para evaluar un programa de reinserción de jóvenes delincuentes. Uno de los objetivos del programa es dar una oportunidad a jóvenes no reincidentes. Para ser admitidos, deben ser sometidos a un escrutinio de sus antecedentes y predisposición a participar en el programa. Una vez admitidos, los cargos penales son cesados mediante un convenio entre las partes, que compromete al participante a restituir, mediante actividades comunitarias, los servicios recibidos. Partidarios del programa arguyen que es menos costoso, para el erario público la implementación de este tipo de metodologías de rehabilitación, que seguir los trámites judiciales tradicionales; al igual, que es exitoso en términos de la no reincidencia. Ante la posibilidad que el programa se termine, se le solicita al analista que reúna toda la información que pueda, apoyada en datos concretos representativos, confiables y válidos que permitan mostrar que el programa sí está produciendo los resultados esperados.

De acuerdo al anterior problema de investigación, formule un plan operativo de la investigación³ provisorio que proporcione una guía del cómo desarrollar metodológica y técnicamente el estudio.

En el diseño de la matriz de variables y dimensiones – variables – que tentativamente abarcan el problema de investigación y las hipótesis u objetivos planteados, asegúrese de ser lo más exhaustivo posible.

2. ¿Cuáles son los factores más comunes que afectan *la calidad interna de la información* en estudios que utilizan diseños metodológicos⁴ de tipo post-test con comparación y el diseño de antes y después con grupo de control? Explícite, todos aquellos factores determinantes que usted considere relevantes, que pueden incidir en la calidad de la información recolectada.

3 Un plan operativo o un diseño operacional de una investigación es un conjunto de etapas del proceso de planificación y conducción de la investigación.

4 Estos diseños experimentales buscan controlar el impacto de factores externos en una relación hipotética de dos grupos seleccionados al azar, uno experimental y otro de control, midiendo ese impacto en espacios temporales diferentes

3. El muestreo

Objetivo

Describir los diferentes tipos de muestra y la forma más general de calcular el tamaño de una muestra para una población dada.

Determinación de la muestra

El objetivo final de una investigación social no se reduce a la descripción de procesos de intervención, diagnósticos o evaluaciones sobre poblaciones particulares, en muchas ocasiones, lo que se pretende es generalizar los resultados de esos procesos.

La mayoría del conocimiento existente se debe a experiencias basadas en inferencias a partir de la observación y el análisis de un limitado número de unidades de análisis. A ese número limitado de unidades de análisis se denominan muestra¹. Cuando se analizan todos los integrantes de una población² se tiene un *censo*. Cuando se analizan algunos casos, pero que cumplen la condición de ser representativos de toda la población, se tiene una *muestra poblacional*. Ambos procedimientos tienen ventajas y desventajas. En el censo se tiene la ventaja de incluir a todos los miembros de la población, por lo que no hay necesidad de generalizar los resultados; la desventaja es que levantar un censo de población puede ser casi imposible por cuestiones de tiempo y recursos. La muestra tiene como ventaja que reduce la cantidad de recursos y tiempo requeridos para recolectar la información necesaria; la desventaja principal está asociada con *el grado de representatividad de la información* recogida, la cual se requiere para extender la inferencia al resto de la población.

Por ejemplo, tanto el costo de cada entrevista personal como el límite de tiempo disponible, no permiten realizar un censo o una cobertura total de la

1 Los estadísticos usan el concepto muestra para describir una porción seleccionada de la población.

2 Generalmente se usa la palabra población no sólo para referirse a personas sino a todos los elementos de un universo que se han escogido para el estudio.

población de votantes; en estos casos, se recomienda usar muestras cuando no es posible medir o contar cada una de las unidades de análisis de la población. De la calidad y representatividad de la muestra dependerá la validez y confiabilidad de la información recolectada.

En los diseños cuantitativos, el proceso de escogencia de una muestra es importante. De allí que haya que definir quiénes van a ser estudiados, esto debe hacerse a partir de los objetivos del estudio.

Matemáticamente es posible describir las características de una población a través de estadísticos³ muestrales.

El *muestreo* es la técnica con que se determina el tamaño y los elementos que integrarán la población muestral, a fin de que cumpla la condición de ser representativa de toda la población. Hay varios tipos de muestreo y cada uno tiene diferentes fórmulas aritméticas para calcular el tamaño de la muestra, así como distinto método de elegir los integrantes de ésta.

La tabla 3.1 Ejemplifica unos tamaños muestrales teóricos de acuerdo al grado de confianza⁴ requerido:

Los métodos de selección de las unidades de análisis que componen una población muestral y pueden clasificarse en dos tipos: muestreo aleatorio o probabilístico y muestreo no aleatorio.

Tipos de muestreo aleatorio o probabilístico

Muestreo aleatorio es aquel método de selección de una población muestral que se basa en la teoría de la probabilidad; en el cual todos los elementos de la población tienen una probabilidad de ser escogidos en la muestra. Este puede ser:

- *Muestreo aleatorio probabilístico simple*. Es aquel en que cada individuo dentro de la población tiene la misma probabilidad de ser seleccionado como integrante de la muestra. Así, en una población de tamaño N , cada individuo de la población total tiene una probabilidad de ser elegido $1/N$. Los elementos de la muestra se eligen en forma aleatoria. Por ejemplo,

3 Un estadístico es una característica de una muestra representativa.

4 El grado de confianza es la diferencia entre la unidad y el nivel de significación con el que se contrasta una hipótesis. Se expresa en tanto por ciento. Por ejemplo, si el nivel de significación es del 0,05, el nivel de confianza será del $1 - 0,05 = 95\%$

Tabla 3.1 Tamaño muestral teórico

Tamaño de la población	± 1%	± 2%	± 3%	± 4%	5%
> 1000	1/2*	1/2	1/2	75	278
2.000	1/2	1/2	696	462	322
3.000	1/2	1.344	787	500	341
4.000	1/2	1.500	842	522	350
5.000	1/2	1.622	879	536	357
10.000	4.899	1.963	964	566	370
20.000	6.489	2.144	1.013	583	377
50.000	8.057	2.291	1.045	593	381
100.000	8.763	3.245	1.056	597	383
500.000 y más	9.423	2.390	1.065	600	384

*En estos casos más de la mitad de la población o universo se requeriría como población muestral

Fuente: H. P., Hill; J. L., Roth y H. Arkin, *Sampling in Auditing*, New York, The Ronald Press, 1982.

se requiere seleccionar una muestra representativa de un grupo de 3.000 estudiantes matriculados en la Facultad de Ciencias Sociales. Existen tres maneras de hacerlo; por *selección intencional*, por ejemplo: con el listado de estudiantes, elegir uno de cada 10. Por *sorteo*: se seleccionan escogiendo al azar un número predeterminado de estudiantes (10%). *Aleatoriamente*, se asigna un número a cada uno de ellos de 1 a 3000 y se seleccionan cien números de una tabla de números aleatorios.

- *Muestreo aleatorio probabilístico sistemático*. En este método los elementos se escogen aplicando un criterio de selección preestablecido y uniforme. Por ejemplo: en la selección de 300 estudiantes de un grupo de 3.000 se puede utilizar el siguiente procedimiento. Con el listado de 3.000 (total de la población de tamaño N), se seleccionan un número de n^5 elementos (300) a partir de un intervalo de selección Q ($Q = N/n$). Entonces, $Q = 3000/300 = 10$. El valor resultante indica que de cada diez personas será elegida una hasta completar el listado de tres mil.

5 Valor que se determina intencionalmente.

- *Muestreo estratificado*. Para este tipo de muestreo, la población se divide en subpoblaciones, grupos o estratos que se presentan en la población con base en algún criterio o atributo, como el lugar de residencia, la edad, el género, etc. Luego se selecciona una muestra para cada subpoblación o estrato. Los elementos se escogen al azar dentro de cada subpoblación o estrato.

El siguiente diseño muestral estratificado se ha extraído del estudio *Imágenes de identidad, integración y conflicto entre la comunidad estudiantil de la Universidad de Antioquia*⁶

Dada la suposición que se hizo en el estudio de la distribución de las características en la población y ante la decisión lógica de estudiar éstas en una muestra de estudiantes, se optó por el empleo de la metodología del muestreo estratificado. Para el efecto, los estudiantes fueron agrupados en tres estratos o niveles de agrupamiento:

- Estrato 1 o nivel inicial: estudiantes en semestre 1º, 2º y 3º
- Estrato 2 o nivel intermedio: estudiantes en semestres 4º, 5º y 6º
- Estrato 3 o nivel terminal: estudiantes en semestres 7º y más

El número de estudiantes en cada nivel, así como la ponderación de cada uno, se presenta en la tabla 3.2.

Tabla 3.2 Número de estudiantes según nivel de agrupamiento y ponderación

Nivel de agrupamiento	Número de estudiantes	Ponderación nivel. wh
1	7.179	0,49
2	4.026	0,27
3	3.559	0,24
Total	14.764	1,00

Para la investigación, los tres estratos o agrupamientos presentan diferentes niveles de información, entonces se descartó el procedimiento más utilizado en este tipo de muestreo, a saber, la división proporcional de la muestra total en los niveles de acuerdo con la ponderación de ellos. Por estas consideraciones se estableció el siguiente orden o clasificación: nivel 2, nivel 3 y nivel 1. Los

6 Ruiz, Jaime, Aigner, Miguel, Ochoa Jaime, De Los Ríos Héctor. Parra Carlos. *Imágenes de identidad, integración y conflicto entre la comunidad estudiantil de la Universidad de Antioquia*. Editorial Universidad de Antioquia: Medellín. 1991: p. 5.

pesos muestrales asignados a éstos fueron: 0,5; 0,3 y 0,2. La tabla 3.3 presenta la distribución de las ponderaciones por nivel:

Tabla 3.3 Ponderación poblacional y muestral de los niveles

Nivel	Ponderación wh	Ponderación wh1
1	0,49	0,2
2	0,27	0,5
3	0,24	0,3

Luego del análisis de diversas alternativas, el grupo investigador decidió fijar como grado de confiabilidad el 90% y un error máximo de 3,5% en la estimación de la proporción, lo cual a su vez representa un tamaño de muestra de 875 estudiantes, distribuidos por nivel de la siguiente manera:

170 para el Nivel 1

450 para el Nivel 2

255 para el Nivel 3

Tabla 3.4 Tamaño de la muestra por nivel según ponderación

Nivel	Afijación proporcional		Afijación establecida		Afijación real	
	Ponder	Muestra	Ponder	Muestra	Ponder	Muestra
1	0,49	398	0,2	162	0,25	202
2	0,27	219	0,5	406	0,43	349
3	0,24	195	0,3	244	0,32	261

Como se anotó en la selección de estudiantes, la muestra total de cada nivel se repartió entre los diversos programas académicos en forma proporcional al peso de ellos en el nivel académico correspondiente.

- *Muestreo estratificado proporcional.* Este tipo de muestreo es un caso particular de muestreo estratificado, en el cual se escoge *en cada grupo, el número*

de elementos que es proporcional al tamaño del estrato. Es el más utilizado, dado que las particularidades de una población — sexo, edad, nivel de escolaridad, nivel de ingresos— pueden influir en la selección de los elementos que integran la muestra.

En el ejemplo presentado, una muestra de 300 estudiantes representa el 10% del grupo de 3.000. Se quiere establecer una estratificación proporcional en función de su rendimiento académico. Así, el estrato A será para los estudiantes con un promedio entre 3,0 y 3,5; el estrato B para estudiantes con promedios que varíen entre 3,6 y 4,0; y el estrato C para los estudiantes con promedios mayores de 4,0.

Se sabe que el tamaño de la muestra — 300 — representa el 10% de la población. En este estudio, se quiere determinar el peso de cada estrato en que se dividió la población estudiantil, según su rendimiento académico. A la población de los tres estratos se les aplica idéntico factor proporcional (10%).

Desagregada la población estudiantil de la Facultad de Ciencias Sociales, de acuerdo a un criterio proporcional, la distribución se presenta en la tabla 3.5

Tabla 3.5 Distribución de estudiantes de la Facultad de Ciencias Sociales según nivel de estratificación socioeconómica.

Estrato	Población	Muestra de un 10%
A	770	77
B	1.500	150
C	730	73
Total	3.000	300

— *Muestreo por conglomerados.* En esta clase de muestreo, se subdivide la población en grupos denominados *conglomerados*. La diferencia de este muestreo con respecto al estratificado, es que se deben censar todos los integrantes de los conglomerados escogidos y no sólo algunos elegidos al azar. Los conglomerados se pueden seleccionar en forma aleatoria o proporcionalmente al tamaño.

Por ejemplo, se cuenta con un listado de los 14.476 estudiantes matriculados en los programas académicos que ofreció la institución universitaria en el primer semestre.

Supóngase que el listado está ordenado alfabéticamente y se seleccionan los estudiantes que integrarán la muestra por medio de una tabla de números aleatorios. Después se busca cada alumno en su facultad y programa académico correspondiente, se le asesora en el diligenciamiento del cuestionario diseñado para la recolección de la información, etc. Dado que este procedimiento es difícil y poco práctico, resulta entonces más sencillo escoger algunos programas académicos y aplicar cuestionarios a todos los estudiantes. En este caso cada programa académico es *un conglomerado*.

El tamaño de la muestra de este ejemplo se obtuvo de esta manera:

El tamaño de muestra para proporciones está dado por:

$$n_o = \frac{PQ}{V}$$

Donde:

n_o = tamaño de la muestra

P = probabilidad para que se realice el evento

Q = probabilidad para que no se realice el evento

V = varianza del estimador

La varianza del estimador es:

$$V = \frac{d^2}{T^2}$$

Donde:

d = error permitido en los datos (está en función de la precisión deseada)

t^2 = valor de la abscisa en el eje X de una distribución normal, tal que deje en la parte central un área igual a la confianza deseada. Así, reemplazando 2 en 1, tenemos:

$$n_o = \frac{t^2 P Q}{d^2}$$

b) Se establece que se requiere un 95% de confianza y que el error no sea mayor del 4%.

Se busca en la tabla⁷ del área bajo la curva normal y se identifica que “t” = 1,96, o sea, aproximadamente 2. Se estima que existe un 50% de probabilidad de que se efectúe el evento y, por tanto, un 50% de probabilidad de que no ocurra ($Q = 100 - P$).

Se tiene al sustituir estos valores en la ecuación 3:

$$n_o = \frac{(4)(50)(50)}{16} = 625$$

c) Para poblaciones finitas, se ajusta al tamaño de la muestra, pues en ocasiones resulta mayor que el de la población. El ajuste para poblaciones finitas está dado por la ecuación:

$$n = \frac{n_o}{1 + \frac{n_o}{N}}$$

Donde:

n = tamaño ajustado de la muestra

n_o = tamaño no ajustado de la muestra

N = tamaño de la población

Entonces sustituyendo los valores en la ecuación 4, obtenemos:

$$n = \frac{625}{1 + \frac{625}{14.756}} = 599,21$$

⁷ La tabla del Área bajo la Curva Normal indica la proporción del área de la curva que corresponde a puntajes Z, con áreas que van desde 0.00 hasta 5.00

Así, el tamaño de la muestra (n) = 599 estudiantes representativos de la población total estudiantil

Observar que el cociente n_o / N es crítico para tamaños pequeños de población; pero casi no tiene efecto en poblaciones grandes, como en este caso, donde la población es igual a 14.476 estudiantes.

Muestreo no aleatorio - no probabilística

Muestreo no aleatorio es un método de selección de una muestra poblacional sin utilización de aleatoriedad. El criterio de escogencia esta sustentado en el conocimiento de la población, de los objetivos del estudio y del tipo de diseño metodológico requerido. En este tipo de muestreo se utiliza el conocimiento y la opinión personal del investigador para identificar aquellos elementos de la población que deberán estar incluidos en la muestra. Este muestreo puede ser:

— *Muestreo dirigido - no probabilístico*. Tales muestras pretenden seleccionar unidades de análisis que según el criterio del experto, son representativas de una población en estudio; por ejemplo: mujeres en edad fértil entre 15 y 19 años.

— *Muestreo por cuotas*. Se utiliza en estudios en donde el criterio de desagregación de la población son ciertas características de sus variables demográficas, por ejemplo: jóvenes profesionales menores de treinta años. El analista selecciona las personas con esas características hasta completar la cuota definida intencionalmente.

4. La recolección de la información

Objetivo

Este capítulo ofrece una información general sobre las técnicas más usuales de recolección de información. Además, se introduce al lector en el instrumento básico de recolección de información de la técnica de la encuesta social: el cuestionario.

La recolección de la información

En el proceso de investigación, la etapa de la recolección de información ocupa un lugar básico. Previo a esta etapa, debemos resolver operativamente varios aspectos que detallamos a continuación:

- Un estudiado planteamiento de un problema de investigación y un enunciado claro y conciso de la(s) hipótesis o del (los) objetivo(s).
- ¿Qué tipo de información requerimos para responder a las preguntas de investigación?
- ¿Qué tipo de instrumentos de recolección necesitamos y, si es el caso, cuándo construir nuestros propios instrumentos de recolección de información y cuándo no?
- ¿Cuál va ser el plan o matriz de análisis?

Así, el proceso de recolección de información implica las siguientes etapas:

- Diseñar o adaptar un instrumento válido y confiable de recolección de información.

- Definir una estrategia de recolección de la información.
- Determinar la población informante – diseño de la muestra.
- Realizar el trabajo de campo: aplicación del instrumento de recolección de información.
- Validar y criticar la información recolectada.
- Codificar y sistematizar los datos.
- Analizar e interpretar la información.

Con respecto al tipo de información requerida, se debe utilizar en algunas ocasiones información primaria, es decir, aquella que el investigador recoge directamente por medio de un contacto inmediato con su objeto de análisis. En otras *información secundaria*, aquella que el investigador recoge a partir de estudios ya realizados por otros investigadores. La distinción entre estos dos tipos de información señala la *técnica* para recolectar el tipo de datos requeridos:

Tabla 4.1 Técnicas de recolección de información

Primaria	Secundaria
Encuesta social	Análisis de contenido documental
Entrevista social	Análisis de medios audiovisuales
Observación científica	Análisis de datos agregados
Experimentación	Análisis etnográfico

Conceptos básicos de la recolección de información

El proceso de selección de la información necesaria y pertinente no es estrictamente un proceso intuitivo; al menos no lo es en los diseños cuantitativos. Las dimensiones conceptuales utilizadas en la formulación del problema y las hipótesis están señalando qué tipo de información se requiere. En términos metodológicos, se puede afirmar que, existen unos lineamientos básicos para recolectar la información – datos – que se sintetizan en una estructura tripartita común a cualquier tipo de datos: las unidades de análisis, las variables y los valores de las variables, y es la siguiente:

En el estudio ya citado sobre las apreciaciones de los estudiantes sobre su mundo universitario, es posible observar, desde el planteamiento del problema, cuál es el sujeto-objeto del estudio — en este caso son los estudiantes de la institución — las hipótesis que explícita o implícitamente se plantean; qué tipo de información será pertinente, cuáles son las dimensiones principales del estudio.

Analicemos apartes de la introducción que nos indica estos tres elementos que componen la estructura tripartita del dato:

Una gran parte de los estudios referidos a los estudiantes universitarios ha puesto su atención, bien, en la descripción de las características sociales, económicas y culturales de los estudiantes y su entorno familiar, o bien, en las manifestaciones del movimiento estudiantil. Estas perspectivas de análisis comúnmente han renunciado al estudio de esos aspectos que para la investigación sociológica son importantes, a saber las prácticas estudiantiles (académicas, existenciales, institucionales, universitarias e institucionales sociales) frente a la vida universitaria, sus potencialidades de integración y de conflicto¹.

La *unidad de análisis* en esta investigación es el estudiante de pregrado matriculado en programas académicos presenciales que ofrece la Universidad de Antioquia. Primero, se identifica un sujeto-objeto del estudio, a continuación, puede preguntarse ¿qué es lo que se quiere saber de esta unidad de análisis? El principio básico del diseño del estudio de opinión que presentamos, es la descripción de estos elementos, relaciones, situaciones, prácticas sociales, académicas y existenciales que se manifiestan en el contexto de la Universidad de Antioquia.

Se ha definido que la unidad de análisis es el estudiante matriculado en programas de pregrado en la Universidad de Antioquia. Segundo, se especifica una serie de *variables* o características asociadas a la unidad de análisis (estudiante universitario) que interesa conocer:

- Los aspectos sociodemográficos y académicos generales de la población estudiantil
- Las imágenes que proyecta la Universidad
- Factores de integración y conflicto
- La vida cotidiana estudiantil en la Universidad

1 Ruiz, Jaime, et. al. Op Cit. Pág. 45.

– Los niveles de participación del estudiante

Finalmente, se debe especificar las categorías o *valores de las variables*.

La tabla 4.2 sintetiza la estructura de los datos en el estudio referido, en algunas de las variables consideradas:

Tabla 4.2 Matriz de variables

Unidad categorías de análisis	Variables	Valores de las variables
Estudiante	Composición de la población por sexo	Masculino
		Femenino
	Estado civil de los estudiantes	Soltero
		Casado
		Separado
		Viudo
		Unión libre
		Horas de estudio diario
	1 - 2	
	3 - 4	
4 o más		
Estratificación socioeconómica	1 - Bajo - Bajo	
	2 - Bajo	
	3 - Medio Bajo	
	4 - Medio	
	5 - Medio Alto	
	6 - Alto	

Veamos ahora en la tabla 4.3 el formato que adquiere la estructura tripartita del dato, en una situación real:

- Se sabe que la unidad de análisis son los estudiantes.
- En este caso específico, *la variable* hace referencia a la percepción que se tiene sobre los problemas que vive la Universidad de Antioquia.

- Los valores que asume la variable serán las diferentes categorías de respuestas percibidas con su respectivo peso porcentual dentro de las opiniones.

Tabla 4.3 El principal problema en el interior de la ciudad universitaria y fuera de ella

Problema	Ciudad Universitaria	Fuera de ciudad universitaria
Inseguridad interna	24,2%	33,0%
Deficiente dotación de bibliotecas y centros de documentación	4,2	2,4
Deficientes ayudas audiovisuales	2,0	1,2
Bajo nivel docente del profesorado	15,3	10,0
Exceso de profesores de cátedra	9,6	4,1
Débil vocación del estudiante	9,0	12,3
Deficientes servicios asistenciales	4,2	6,2
Poca oferta de cupos	10,0	9,3
Deficiente comunicación entre las directivas y estudiantes	17,0	16,4
Otros	4,2	5,0

Fuente: Ruiz, Jaime, Aignerren, Miguel, Ochoa Jaime, De Los Ríos Héctor, Parra Carlos. *Imágenes de identidad, integración y conflicto entre la comunidad estudiantil de la Universidad de Antioquia*. Editorial Universidad de Antioquia. Medellín: 1991: p. 96.

Técnicas de recolección de información

A continuación se presentan algunas de las técnicas de recolección de información más utilizadas en los diseños cuantitativos, porque facilitan el manejo de la información cuantificable.

Medición escalar

Existe un tipo de información que por sus características — subjetivas, del orden de la opinión, de las imágenes y los valores — no permite ser medida en su cabalidad con instrumentos tradicionales. En estos casos, *la medición*

escalar coloca al encuestado en la posibilidad de expresar su opinión dentro de una gradación que mide la intensidad de aquella. Esta técnica permite una mayor aprehensión de lo cualitativo del hecho social que se estudia.

*“Una escala es una técnica que permite medir la cantidad de una propiedad, llamada actitud hacia algo, poseída por un conjunto de personas.”*²

Se dice que una medición es integral cuando, abarca adecuadamente todos los aspectos del concepto en estudio. A menudo, esto es difícil de realizar con una sola medida, tal como con una pregunta en un cuestionario.

La escala³ más aconsejable es aquella que presenta *proposiciones* que deben ser evaluadas en un continuo de categorías de no más de 5 valores.

En muchas circunstancias, un investigador podrá obtener varias mediciones y combinarlas en un índice o en una escala. Veamos el diseño de una escala que permita evaluar el desempeño académico de los estudiantes del Programa de Sociología de la Universidad de Antioquia, medido por su mayor o menor dedicación a un conjunto de actividades que determinan su rendimiento académico. A los estudiantes se les hizo la pregunta: ¿Cuánto tiempo le dedica usted a las siguientes actividades relacionadas con su desempeño académico? Para efectos del estudio se diseñó la tabla 4.4.

Una matriz de respuestas, en la cual cada alternativa de respuesta tiene un valor o peso dentro de la escala, permite ordenar a los encuestados en un índice o jerarquía. En el ejemplo anterior, se puede hacer una escala que va desde seis puntos — desde el estudiante que declara que *nunca* realiza las actividades enumeradas — hasta dieciocho puntos — lo cual sería el caso del estudiante que afirma realizar *siempre* todas las actividades.

Etapas del diseño de la escala

“Un tipo de escala muy utilizada en la investigación social, es aquella destinada a jerarquizar objetos, (personas, grupos, procesos, roles, ocupa-

2 Escalante. A. Carlos, La medición de las actitudes. Colección Ciencias Sociales. Teorías y Métodos. Bogotá: Tercer Mundo. 1983. Pág. 11

3 “Se puede definir una escala como un conjunto de ítems o declaraciones construidos por el investigador según ciertas definiciones conceptuales previas y utilizadas para cuantificar características o propiedades del comportamiento” Ídem. Pág. 11

Tabla 4.4 Tiempo de dedicación al trabajo académico

Objeto de actitud medido	Afirmaciones de los estudiantes		
	Siempre - 3	Ocasional - 2	Nunca - 1
Desempeño académico - Valores			
Estudiar notas de clases	{ }	{ }	{ }
Revisar la bibliografía	{ }	{ }	{ }
Asistencia a clases	{ }	{ }	{ }
Consultar textos	{ }	{ }	{ }
Trabajo en equipo	{ }	{ }	{ }
Consultar al profesor	{ }	{ }	{ }
Totales	{ }	{ }	{ }

ciones) y valores en una cierta propiedad o variable. Tal variable puede ser el prestigio, [...] la eficiencia, etc"⁴

El procedimiento es el siguiente:

De acuerdo a los objetivos de la investigación, se eligen unos sujetos objeto de estudio y unas variables que se expresan en forma de proposiciones, índices alfanuméricos y de las cuales se quiere una cierta jerarquización.

Como instrumento de recolección de información es un cuestionario que se nos presenta con un conjunto de preguntas que tienen una estructura de ítems o proposiciones utilizados para cuantificar características o variables de lo que se quiere medir. De ahí la importancia de la *selección de las afirmaciones o proposiciones de la escala*. En esta primera etapa del diseño de la escala hay que brindar especial cuidado en la validez y relación de *las afirmaciones* que se van a utilizar para medir la dimensión conceptual del estudio, ya que si un aspecto importante es omitido, la escala puede ser considerada inválida. Las afirmaciones califican al objeto de la actitud que se está midiendo⁵.

La mejor estrategia para probar si existen *omisiones* en las afirmaciones o proposiciones sobre la dimensión del estudio, es incluir la categoría *otras* y *especifique cuáles* en la encuesta piloto. Aquellos valores más frecuentemente mencionados deberán ser incluidos en el cuestionario final.

4 Briones Guillermo. Métodos y técnicas de investigación para las ciencias sociales. México: Trillas. 1982. Pág. 138

5 Hernández, R. et al. Op Cit. Pág.264.

Interrelacionar las afirmaciones. Si todas las afirmaciones tienen como objetivo medir el mismo concepto, ellas deben estar relacionadas una a una y en conjunto. La escala debe ser unidimensional, esto es, medir una sola dimensión del hecho, proceso, actitud o conducta que se está investigando.

En el caso del modelo de escala utilizada para medir la intensidad de tiempo, que el estudiante le dedica a un conjunto de actividades que se relacionan con su desempeño académico, se está tratando de diferenciar entre estudiantes *que siempre realizan una actividad*, con aquellos que *nunca la hacen*; quienes logran un puntaje alto (tres puntos) en dos o más respuestas, deberían obtener semejantes puntajes al contestar las siguientes – con algunas excepciones. Si las tendencias muestran en la selección una dispersión en las respuestas, es evidente que se están midiendo dos o más dimensiones del desempeño académico, lo cual es un error.

Valorización de la escala. Otro paso seguir es asignar un puntaje numérico a cada respuesta obtenida, esto implica decidir acerca del valor de cada respuesta y luego sumaras. A continuación, las respuestas deberán ser agrupadas en algunas pocas categorías, tales como *alta*, *media* o *baja*. Al examinar el total de los puntajes, usualmente los extremos – *altos* o *bajos*, *siempre* o *nunca* – son escasos.

Análisis de los resultados obtenidos mediante una escala. La lectura de los datos es básicamente un ejercicio aritmético, de sumatoria de los puntajes obtenidos por cada uno de los que respondieron, permitiendo así algunas inferencias sobre la actitud o conjunto de actitudes que se pretendió medir.

Algunas clases de escalas

- *Escala Likert.* Esta es una técnica de medición multidimensional⁶ que trabaja generalmente con datos a nivel ordinal⁷ y que permite medir una actitud y el comportamiento consecuente con ésta. Se basa en la premisa según la cual la intensidad de las apreciaciones, opiniones o imágenes de un individuo puede ser escalada dentro de ciertas categorías. Así, en lugar de preguntar abiertamente a las personas acerca de un tema, ellas son enfrentadas a unos enunciados o afirmaciones cuyos contenidos, por

6 Una característica propia de la escala Likert es que no es unidimensional; los enunciados o afirmaciones o ítems no necesariamente deben estar relacionados temáticamente unos a otros.

7 Al igual que la escala Guttman, se trabaja con datos ordinales; existen escalas de medición como la *escala Thurstone* que pueden producir mediciones a nivel de intervalos.

sus características, los colocan ante la posibilidad de tener que manifestar sus opiniones, dentro de una escala de categorías posibles de respuestas y la variación de tales categorías; pueden abarcar, por ejemplo, un rango desde *Muy de acuerdo*, *De acuerdo*, *Indiferente*, *Desacuerdo*, hasta *Muy en desacuerdo*. Un aspecto distintivo de la escala Likert es que asume que *las afirmaciones* miden la actitud hacia un único objeto – concepto subyacente. El procedimiento para diseñar y codificar una escala tipo Likert es similar a las que se han presentado anteriormente: debe construirse al menos, con diez afirmaciones o enunciados que califiquen el objeto de actitud.

Un ejemplo de los resultados de una escala Likert aplicada a una pregunta para seleccionar proveedores “¿Qué tan importante para su empresa es...?” se presenta en la tabla 4.5

Tabla 4.5 Nivel de importancia de un producto

Características del producto	Muy importante (5)	Importante (4)	Medianamente - importante (3)	Poco importante (2)	- No es importante (1)
Calidad producto	X				
Precio	X				
Forma pago		X			
Tiempo entrega			X		
Garantía		X			
Marca del producto			X		
Prestigio marca		X			

Las respuestas se califican de acuerdo con los valores preestablecidos y las puntuaciones de las escalas Likert se obtienen sumando los valores obtenidos respecto a cada *frase o afirmación*. Por ejemplo, el valor obtenido es: $5 + 5 + 4 + 3 + 4 + 4 + 4 = 29$

Una puntuación se considera alta o baja según el número de afirmaciones. Por ejemplo, en la escala para evaluar la actitud para seleccionar un proveedor para la empresa, la puntuación mínima posible es 7 y la máxima es 35. La persona que respondió la pregunta obtuvo veintinueve puntos, luego se puede concluir que su actitud hacia los proveedores es más bien favorable – de acuerdo a unos criterios preestablecidos.

Véase otro ejemplo. En el estudio de opinión *Imágenes de identidad, integración y conflicto entre la comunidad estudiantil de la Universidad de Antioquia*⁸, se le solicitó a los estudiantes su valorización acerca de lo que el estudio denomina *factores integradores*, en términos de las categorías *adecuadas*, *inadecuadas* e *indiferentes*. La tabla 4.6 muestra los resultados obtenidos en este estudio mediante un formato escalar:

Tabla 4.6 Valoración de las prácticas existenciales dentro y fuera de la ciudad universitaria

Actividades	Adecuadas		Inadecuadas		Indiferentes	
	C.U. %	F.U. %	C.U. %	F.U. %	C.U. %	F.U. %
Teatro	91,0	89,1	2,3	5,4	6,8	1,0
Conciertos	88,3	88,8	3,6	2,0	8,1	1,0
Deportivas	53,1	57,65	14,7	11,7	2,2	30,7
Recreativas	54,2	54,1	13,7	15,6	31,1	30,3
Cine	9,1	16,6	61,9	46,3	29,0	37,0

C.U.= En el interior de la ciudad universitaria. -F.U.= Fuera de la ciudad universitaria.

Fuente: Ruiz, Jaime, et. al., op cit. 1991: Pág. 24

- *Escala de Guttman*. Esta es una técnica que presenta múltiples escogencias a lo largo de una dimensión. El propósito de esta escala, es medir la *unidimensionalidad actitudinal* – mide sólo una dimensión – de acuerdo con el supuesto de que la actitud íntegra está contenida en una sola dimensión. Es decir, los enunciados o afirmaciones deben estar relacionados unos con otros. Los encuestados son enfrentados a responder a un conjunto de opciones o alternativas, que se pueden dar en la dimensión actitudinal. Estas opciones se presentan en una especial disposición, de tal manera que las alternativas o preguntas midan, la intensidad de la apreciación o la opinión.

8 Fuente: Ruiz, Jaime, Aignerren, Miguel, Ochoa Jaime, De Los Ríos Héctor, Parra Carlos. *Imágenes de identidad, integración y conflicto entre la comunidad estudiantil de la Universidad de Antioquia*. Editorial Universidad de Antioquia. Medellín: 1991: p. 24.

Supongamos que estamos midiendo actitudes frente a las prácticas abortivas. Para estos efectos se diseñó un continuo de ítems relacionados con el tema:

¿Está usted de acuerdo o no con que el aborto sea permitido en las siguientes circunstancias?

- Cuando la vida de la mujer está en peligro.
- En caso de incesto o violación.
- Cuando el feto muestra clínicamente riesgos graves de salud.
- Cuando hay consentimiento paterno.
- Cuando la mujer no puede criar responsablemente el futuro niño.
- Cuando la madre lo considere pertinente.

La entrevista social

En ocasiones, la información que se requiere, ya sea por sus características, complejidad – creencias, valores – o por razones de orden psicológico o ideológico, e incluso histórico, sólo es posible reconstruirla mediante un proceso conversacional. Si bien este tipo de técnica no permite, en términos generales, recolectar una información factible de cuantificar, se convierte en una útil fuente de referencia, sustentación y complementación en los diseños cuantitativos.

El instrumento principal de esta técnica de recolección primaria es *la guía de entrevista o guión*, diseñado con base en los objetivos, las hipótesis o las temáticas preestablecidas, lo cual permite obtener una información válida gracias a la posibilidad de replantear, aclarar la pregunta o utilizar ayudas audiovisuales, facilitando registro más confiable y completo de las respuestas.

La encuesta social

Es una técnica destinada a obtener información primaria de un grupo representativo de individuos para proyectar los resultados sobre la totalidad de la población en estudio. Su objetivo es medir de la forma más acertada

los conocimientos, las actitudes, las creencias y los comportamientos de las personas. La posibilidad de lograr lo anterior depende básicamente de la precisión, pertinencia y concisión de las preguntas.

Los elementos constitutivos de la encuesta social son:

- Planteamiento de un problema de investigación.
- Determinación de las hipótesis de trabajo u objetivos.
- Definición y operacionalización de los conceptos y variables.
- Diseño del cuestionario.
- Diseño muestral.
- Trabajo de campo.
 - Pre-test o prueba piloto.
 - Recolección, crítica y validación de la información.
- Sistematización.
- Análisis e interpretación de la información.
- Informe final.

Las temáticas propias de la encuesta social son:

- Características de la muestra poblacional: composición social, composición del hogar, acceso a medios de consumo colectivo, características demográficas.
- Opinión pública: actitudes, comportamientos e ideologías.
- Conocimientos: tendencias electorales, mercadeo social y económico, etc.

En síntesis, la encuesta social:

1. Permite mayor cobertura en la población del estudio y mayor representatividad de ésta. Los resultados se pueden generalizar gracias a las técnicas de muestreo y la inferencia estadística.
2. Hace posible traducir la información al lenguaje de variables/preguntas; así se puede codificar la información mediante un sistema alfanumérico, facilitando la cuantificación.

3. Es una técnica complementaria a la observación social⁹ y la entrevista social. Sus diferencias pueden verse en la tabla 4.7.

Tabla 4.7 Diferencias entre las técnicas de la encuesta social, la observación y la entrevista.

	<i>Observación /Entrevista</i>	<i>Encuesta Social</i>
Nº. de unidades de análisis	Limitado	Ilimitado
Nº. de dimensiones variables del estudio	Limitado	Ilimitado
Representatividad	Menor	Mayor
Precisión cuantificable	Menor	Mayor
Cualidad del hecho	Mayor	Menor

El Cuestionario

El *cuestionario* es el instrumento básico de recolección de información de la técnica de la encuesta social. Este instrumento consiste en un formulario con un conjunto de preguntas estandarizadas aplicadas a determinados sujetos o unidades de análisis.

– *El proceso de diseño de un cuestionario.* La base de las preguntas de un cuestionario son los conceptos y variables mediante los cuales se ha definido un problema de investigación. Pero para pasar de los conceptos y variables a las preguntas es necesario contar con una guía de lo que se debe observar.

El esquema de la guía es el siguiente:

Concepto → Variable → Indicadores → Pregunta

El proceso de convertir una variable en indicadores se le denomina *operacionalización*. Los indicadores son referentes que permiten hacer medibles y observables conceptos o variables. En esta dirección, es la traducción de

9 Técnica de recolección de información primaria que permite obtener datos que supone la comprobación de la situación de interés, registrándose hechos, acciones o comportamientos mediante una guía de observación. Generalmente se la clasifica en estructurada y no estructurada.

conceptos o ideas a términos cuyos referentes son observables, empíricos y medibles. Las definiciones operacionales son proposiciones que especifican un procedimiento para determinar el valor numérico de una variable dada en casos concretos. Es traducir un concepto-variable en indicadores.

En el ejemplo¹⁰ que a continuación se presenta, se puede observar el proceso que va desde la definición y operacionalización de los conceptos, pasando por el diseño de los indicadores hasta la formulación de las preguntas – es decir, la construcción del cuestionario.

Problema de investigación: *los estudiantes de la Universidad de Antioquia y sus prácticas existenciales*. Gran parte de los estudios referidos a los estudiantes universitarios han dirigido su atención a la descripción de las características sociales, económicas y demográficas de los estudiantes y de sus grupos familiares, o en las manifestaciones del movimiento estudiantil. Estas perspectivas de análisis comúnmente han renunciado a un aspecto que para la investigación social es importante como es el caso de las llamadas prácticas estudiantiles – académicas, existenciales, institucionales – frente a la vida universitaria.

El objetivo de esta investigación es identificar las manifestaciones más corrientes de estas prácticas existenciales. Este análisis involucra una doble visión: las prácticas mismas y la caracterización social de los actores.

Para el diseño del cuestionario se procede de la siguiente manera: a título de ejemplificación se trabaja el concepto de *prácticas existenciales*. Del mismo planteamiento del problema se extrae su definición operacional: son aquellas prácticas que tienen que ver con la existencia individual del estudiante, con su vida cotidiana, con las dimensiones no universitarias de su vida. Esta definición sugiere varias dimensiones a analizar, tales como el proceso de reproducción social del estudiante, las características sociodemográficas, la vida cultural, el uso del tiempo libre –recreación– y la composición social, etc. Para simplificar, vamos sólo a *describir* las características sociodemográficas del estudiante de la Universidad de Antioquia. La siguiente tarea es idear y construir indicadores observables a partir de los conceptos dados anteriormente en el planteamiento del problema. Luego se formulan las preguntas (véase tabla 4.8)

10 Ruiz R., Jaime, et al. Op Cit.

Tabla 4.8 Matriz de variables, indicadores y preguntas

Variables	Indicadores	Preguntas
Distribución por sexo	Masculino	Consignar sexo
	Femenino	
Distribución por edad	15 – 19	¿Qué edad tiene usted?
	20 - 24	
	25 - 29	
	30 y más	
	soltero	
Estado civil	casado	¿Cuál es su estado civil?
	separado	
	viudo	
	unión libre	

- *Preguntas del cuestionario.* Con ellas se busca obtener información acerca de lo que una persona conoce, cree o espera, siente o desea, intenta hacer, hace o ha hecho. Correlativas a las preguntas están las respuestas y sobre éstas, tradicionalmente, la discusión se centra en su *calidad externa e interna*. La *calidad externa* se refiere a la posibilidad de generalizar los resultados encontrados. En cambio, la *calidad interna* de la información recolectada se refiere a dos situaciones: *la validez y la confiabilidad*. La primera, procura que las preguntas reflejen realmente lo que se quiere medir, y la segunda indica que un cuestionario o una pregunta es confiable cuando se aplica repetidamente con la misma metodología, en idéntica población, produciendo iguales resultados.
- *Tipos de preguntas.* Los cuestionarios pueden presentar dos tipos básicos de preguntas: según el contenido y según la forma.

Según *su contenido*, se pueden distinguir dos tipos:

1. *Preguntas relacionadas con hechos* susceptibles de confirmarse mediante la comparación o valiéndose de información secundaria considerada válida, como también, partiendo de situaciones vividas por el interlocutor.

Por ejemplo: “¿Ha consumido usted, por lo menos una vez en su vida, alguna droga adictiva como marihuana o cocaína?”

2. *Preguntas de opinión.* Éstas tienen que ver con las opiniones, actitudes o preferencias del encuestado.

Por ejemplo: ¿comparando con el año anterior, durante lo corrido del año, usted percibe que la seguridad en los espacios públicos de la ciudad ha aumentado o ha disminuido?

Según su forma, podemos hablar de *preguntas abiertas o cerradas*, o de una mezcla de ambas. Las *preguntas cerradas* dan al encuestado la oportunidad de escoger una respuesta, entre una serie de opciones o categorías preestablecidas; las *preguntas abiertas* no proponen respuestas posibles y dejan espacio para que cada cual responda.

Un ejemplo de cuestionario con preguntas cerradas, con alternativas de respuesta preestablecidas y codificadas en el mismo instrumento sería:

¿Con quién vive?	
1. Con sus padres y hermanos	1. ___
2. Con otros parientes	2. ___
3. Con hermanos mayores	3. ___
4. En una residencia	4. ___
5. Con su esposa(o)	5. ___
6. Con amigos	6. ___
7. Solo	7. ___

En términos generales, la ventaja de utilizar preguntas cerradas en un cuestionario radica, en la facilidad de codificación y comparabilidad. Otra ventaja de este tipo de preguntas en temas como: participación política, frecuencia de la actividad sexual, ingresos, edad, etc., es que los encuestados prefieren responder escogiendo una de las alternativas ofrecidas antes que tener que expresar verbalmente respuestas incómodas. Además, las opciones preestablecidas que acompañan a las preguntas cerradas pueden ayudar a clarificar la pregunta misma al encuestado, así se evitan problemas de mala interpretación y de respuestas inconsistentes.

Desde la perspectiva impresionista de los diseños cualitativos, se insiste que con las preguntas cerradas se fuerza a los encuestados a limitar sus apreciaciones dentro de categorías que supuestamente no cubren la cualidad del hecho o proceso sobre el cual se está inquirendo y, en consecuencia, la

respuesta tiene menos significado y menos capacidad de aprehensión. Esta imagen distorsionada tiene asidero tanto en el abusivo uso de preguntas cerradas con escasas alternativas – dos o tres categorías – como en el uso de la categoría *otros*, sin permitir que el encuestado señale y consigne la posible respuesta alternativa.

En cuanto a los cuestionarios con preguntas abiertas, la ventaja es su utilidad para explorar procesos o problemas desconocidos parcial o totalmente por los investigadores. Este tipo de formato da la posibilidad al encuestado, de describir la razón de sus ideas. Algunas veces, en el proceso de construcción de un cuestionario es difícil especificar todas las posibles respuestas a una determinada pregunta. En esta situación, una pregunta abierta es un excelente instrumento para lograr cubrir todas las posibilidades de respuesta para posteriormente codificarlas. Estas preguntas son de gran valor cuando se quiere medir el conocimiento de los encuestados sobre un determinado hecho o personas involucradas en el.

Lo extraño es que la principal desventaja de las preguntas abiertas esta relacionada con su ventaja principal, según la cual, el encuestado puede extenderse en sus respuestas, pero esto puede dificultar enormemente el proceso de validación, crítica y codificación de la información obtenida. Ambas situaciones son relevantes si finalmente se decide que se procesará esa información utilizando un programa informático especializado, y esto es difícil pues las preguntas abiertas deben cerrarse tipificando¹¹ las distintas respuestas, para después incluirlas en hojas de codificación antes de su digitación en un programa informático.

Cuando no es posible que la respuesta a la pregunta sea respondida, es frecuente formular una combinación de preguntas, dejando la posibilidad de una respuesta abierta como se presenta en el siguiente ejemplo:

Ocupación de la persona que contribuye principalmente al sostenimiento del hogar:

- | | |
|----------------------|----------------|
| 1. Empleado | 4. Comerciante |
| 2. Obrero | 5. Profesional |
| 3. Por cuenta propia | 6. Otro |

Especificar: _____

11 El proceso de tipificación consiste en encontrar cuáles son las respuestas más comunes, asignarles un código o clave alfanumérico y después, con esas claves, codificar las respuestas de todos los cuestionarios.

– *La composición y orden de las preguntas de un cuestionario.* La redacción, el ordenamiento y disposición de las preguntas son pasos básicos en el diseño de un cuestionario. No debe olvidarse que ellas son estímulos verbales, visuales o escritos que deben despertar interés sobre el tema y, por tanto, deben ir creando una atmósfera de confianza y motivación para contestar las siguientes. Por eso es tan importante el ordenamiento de las preguntas.

Elizabeth Noelle, refiriéndose a lo que ella llama “el cuestionario con buenos modales”, dice lo siguiente:

*Cuestionarios que pueden tener buenos o malos modales, pueden ser cortos o atentos, toscos o egoístas” [...] por ejemplo, la elección de las preguntas de contacto al comienzo, o la colocación de una pregunta interesante o alegre al fin de la entrevista [...] Pero, sin darnos cuenta, las preguntas de buenos modales en la entrevista se convierten, en condiciones cuyo cumplimiento determina inmediatamente la calidad de los resultados”.*¹²

El orden en la presentación de las preguntas es otro criterio metodológico que se debe tener en cuenta, puesto que puede influir en la confiabilidad y validez de las respuestas. Es así como, al ordenar las preguntas dentro del formato del cuestionario, se debe tener en cuenta la influencia de las preguntas previas en las respuestas siguientes — *contaminación de las preguntas*.

Según se ubique una pregunta al principio o al final del cuestionario, ésta puede provocar respuestas diferentes. A medida que se va desarrollando el interrogatorio, el encuestado reflexiona, se familiariza, se forma una opinión sobre el tema, lo cual incide sobre el contenido de su respuesta. Un asunto diferente es que se le formulen preguntas inesperadas en relación con su posición dentro del cuestionario, dicha posición debe ser cuidadosamente escogida.

Ghiglione¹³, refiriéndose a lo que él denomina *el efecto de halo*, afirma que éste se manifiesta “cuando se solicita primero un juicio global: ciertas personas se sentirían incoherentes si, al haber expresado de entrada un juicio favorable, manifestaran después críticas o reservas sobre aspectos particulares”. De igual forma, previene acerca de los problemas que conlleva “formular múltiples preguntas sobre el mismo tema o sobre temas próximos; tal es el caso cuando se pretende construir una escala de actitud o se desea explorar los diferentes aspectos de un tema y analizar las relaciones entre las respuestas”.

12 Noelle, N., Elizabeth. Encuestas en la sociedad de masas. Madrid: Alianza. 1970. Pág. 64

13 Ghiglione, Rodolfo, Las encuestas sociológicas, teoría y práctica. México: Trillas. 1989

Es necesario, entonces combinar enunciados favorables y desfavorables".¹⁴ A veces es recomendable la alternación de temas en las preguntas para lograr la confianza y motivación del encuestado, de tal forma que se logren respuestas de mayor profundidad.

Para prevenir *la contaminación* de las preguntas se pueden seguir estos criterios:

- Dispersar las preguntas en el conjunto del cuestionario, alejando aquellas que arriesgan contagiarse.
- Se puede disponer las preguntas en forma de "embudo", es decir, colocando primero aquellas preguntas más generales y después progresivamente las particulares y especiales, evitándose así las sugerencias de unas a otras.

Limitaciones de la técnica de la encuesta social

En el caso particular de los *sondeos de opinión*, éstos son una foto instantánea, una radiografía de la opinión pública. La calidad de la fotografía depende no sólo de la calidad de la máquina fotográfica, sino también de la capacidad creativa y de manejo que de aquella haga el fotógrafo —no hay máquinas fotográficas automáticas que produzcan encuestas en serie. El investigador define el campo u objeto, la abertura del lente, la distancia, revela y enmarca la instantánea. Teóricamente sólo registra y mide la opinión.

Como se está hablando de procesos en donde la intervención del hombre desempeña un papel importante, pueden darse fotos movidas, oscuras, fuera de foco o simplemente mutiladas. Ante esta situación, cabe preguntarse si esto es producto de un mal manejo técnico o si es el efecto de los sujetos-objetos de la fotografía, que por sus propias características, sean imposibles de *captarlos* o que haya cierta intencionalidad de manipulación con preguntas que tergiversen o inciten las respuestas. Esto último tiene que ver con el uso de los datos obtenidos: puede ser la materia prima para hacer un fotomontaje para crear opinión —consensos consumistas o políticos, etc.— o ser una técnica que permita oscultar la opinión, suministrar elementos de juicio a los hacedores de políticas. En síntesis, un estudio de opinión *capta*

14 Idem. pág. 8

los acuerdos o desacuerdos de los pensamientos de los demás con relación a hechos sociales, coyunturas, imágenes, etc.

Se afirma entonces que los sondeos mediante encuestas son como una fotografía instantánea y ello implica que en la próxima fotografía todo será diferente. Es por esto que se recomienda, especialmente en los estudios de opinión sobre tendencias electorales, que haya continuidad y periodicidad en el tiempo de los sondeos.

También se presentan algunas limitaciones técnicas con las preguntas, por lo que se recomienda tener en cuenta las características de las que se van a utilizar para medir las variables del estudio y sus dimensiones.

Cuando se comienza a formular un conjunto de preguntas a partir de las dimensiones del estudio, se deben tener en cuenta algunas recomendaciones con el fin evitar sesgos en las respuestas, tales como evitar *preguntas dobles* – dos preguntas en una – como se ve a continuación: ¿Esta usted de acuerdo con la declaración del gobierno, que la política de apertura económica es irreversible y que se debe congelar el gasto público para facilitar el proceso?

Con la pregunta anterior se pueden dar situaciones como las siguientes: una persona puede estar de acuerdo con la política de apertura, pero en desacuerdo con la reducción del gasto público y viceversa, estar en desacuerdo con la política de apertura y de acuerdo con una mayor intervención del Estado, reflejado en una mayor gasto fiscal.

Para el analista, el problema es similar: ¿cómo interpretar una respuesta que señale un acuerdo? ¿Éste es aplicable a ambos componentes de la pregunta o sólo a uno de ellos?

Del enunciado de la pregunta citada surgen otros problemas, sobre los cuales se deben tomar algunas precauciones, o al menos tener conciencia de sus efectos. A pesar del lógico cuidado y esfuerzo del investigador en el proceso de construcción del cuestionario, al utilizar conceptos cuyas definiciones operacionalizables, permitirían una mayor y mejor aprehensión de la cualidad que se está estudiando, se deben evitar conceptos o palabras con múltiples significados o interpretaciones que puedan distorsionar los objetivos que se quieren lograr con tal o tales preguntas.

Un típico ejemplo de una pregunta conceptualmente ambigua es: ¿Qué opina sobre la legitimidad del Estado y la gobernabilidad democrática?

Tal formato de pregunta no sólo es conceptualmente ambiguo, sino que también es ambigua en lo que realmente se quiere inquirir y además, es *doble*.

Como se aprecia, una pregunta mal planteada puede de manera implícita abarcar más de un tema, también ser ambigua o vaga conceptualmente. La sofisticación conceptual, como el uso de anglicismos o conceptos técnicos, al igual que el uso de modismos regionales o un vocabulario muy académico, parecen ser problemas recurrentes en el diseño de un cuestionarios.

Esta sugerencia metodológica toca con otro problema, como es el que cada cuestionario debe ser construido de acuerdo no sólo con los objetivos del estudio, sino también con las características de quien responde – características demográficas, socioeconómicas, etc. Se sugiere usar un vocabulario apropiado acorde con la población estudiada.

Otra situación de sesgo puede darse cuando la información pasa del investigador al divulgador. Con el auge de los estudios de tendencias electorales por medio de los sondeos de opinión, aparece también la polémica: ¿que tan confiables son? ¿Los resultados que arrojan crean o manipulan opinión? ¿Hay manipulación de los medios de comunicación por la forma que difunde la información? Al respecto, Christopher Hitchens señala: “Las encuestas fueron el fruto de una lucha no para descubrir la opinión, sino para manejarla. Fue un arma no para descubrir la opinión sino para manejarla. Fue un arma en las primeras guerras para frustrar los planes del obrerismo organizado y en la batalla contra el populismo; y más tarde fue el recurso favorito en el arsenal de la psicología de masas.”¹⁵

Hitchens también hace referencia a aquellas preguntas que pueden inducir *respuestas esperadas* o *de preferencia obligada*, que se presentan cuando el interrogador plantea una proposición tajante o que dan pie a elegir un sólo tipo de respuesta socialmente aceptable.

De igual manera, las preguntas demasiado escuetas o concisas en la formulación del interrogante y amplias o recargadas en los contenidos implícitos o explícitos (sobre la temática), según Hitchens, “pueden, así provocar una reacción nerviosa y suscitar respuestas impensadas”.¹⁶

15 Hitchens, Christopher; Cándido, Monzón Comp. Opinión Pública, comunicación y política. Las encuestas electorales. Madrid: Tecnos. 1996. Pág. 351.

16 Ídem. Pág. 352

Las preguntas acerca de temas personales o íntimos, tales como: preferencias políticas, la orientación sexual, el uso de anticonceptivos, la adicción a sustancias adictivas; deben ser diseñadas de tal forma que no necesariamente exijan una respuesta que involucre de modo explícito al encuestado y que no provoquen su rechazo. Se sugiere primero hacer las preguntas en términos que al encuestado se le interrogue acerca de lo que piensan *los otros* alrededor del tema (para que no se sienta directamente involucrado). Metodólogos como Hyman Korman afirman que en estos casos la respuesta reflejará realmente la opinión del encuestado más que la de *terceros*¹⁷. De igual forma, se sugiere utilizar las técnicas escalares de medición de actitudes, en lugar de preguntas directas para el tratamiento de este tipo de temáticas, así como para medir preferencias o niveles de satisfacción.

Formular preguntas de manera que implique aceptar ciertos presupuestos implícitos, puede convertirse en un instrumento de manipulación de la opinión, en el sentido de hacer admitir aspectos íntimos que con dificultad se reconocerían abiertamente, por ejemplo: Estamos estudiando los cambios recientes en las prácticas anticonceptivas de las universitarias. ¿Cuáles de estos anticonceptivos usa usted? Anticonceptivos orales, dispositivos intrauterinos, esterilización u otros?

Este tipo de preguntas ubica al interrogado en una situación comprometedora, al darse por sentado que la persona tiene una práctica anticonceptiva. Así mismo, vincular instituciones o personalidades a temáticas como una forma de respaldar socialmente la pregunta o como evidencia comprobada, puede afectar las respuestas. Lo anterior se constituye en una precaución metodológica para tener en cuenta al construir las preguntas si se quieren evitar sesgos en las respuestas.

Otro problema con las respuestas puede suceder cuando se confronta al encuestado con una serie de enunciados o afirmaciones sobre diferentes temas — técnicas escalares de medición de actitudes —, dado el diseño del cuestionario se le ofrecen idénticas categorías de respuesta como: *Muy de acuerdo, De acuerdo, Indiferente, Desacuerdo, Muy en desacuerdo*.

El encuestado puede estar de acuerdo con los primeros enunciados y luego por efecto del cansancio o reticencia al tema de la encuesta, declarar reiteradamente estar *De acuerdo* con los planteamientos o afirmaciones siguientes.

17 Korman Hyman. Técnicas para investigar la opinión pública. Department of Sociology. New York: State University of New York at Stony Brook. 1984. Pág. 9.

Este fenómeno es lo que se llama *desviación de convencimiento*: se tiende a responder *Sí* antes que *No*, *de acuerdo* antes que *En desacuerdo*.

Para evitar este tipo de respuestas *rutinarias y repetidas*, se sugiere que los enunciados o ítems relacionados con el fenómeno que se va a medir, sean redactados en forma alternada, de tal manera que el encuestado pueda estar de acuerdo con el primer enunciado y no necesariamente con el segundo, y así sucesivamente; de esta forma se le obliga a poner atención y a responder con cuidado. También el continuo de respuesta que representa una escala ordinal, de un mayor valor a un menor valor —de *Totalmente de acuerdo* a *Totalmente en desacuerdo*— deberá ser presentado alternadamente.

Hay temas como la inseguridad ciudadana, el terrorismo y el secuestro, que determinan un tipo de respuesta emotiva, sesgada e inducida posiblemente por los medios de comunicación. La sutileza en su formulación y la ubicación dentro del formato del cuestionario son recomendaciones a tener en cuenta. El hecho de ser presentadas al comienzo del cuestionario puede influir en las respuestas posteriores, tal vez a causa del esfuerzo del encuestado de aparecer consistente con sus primeras afirmaciones o con respuestas socialmente aceptables.

Semejante al anterior, es el fenómeno que se presenta en las encuestas de tendencias electorales o de imagen. Al colocar al comienzo del cuestionario preguntas directas sobre la imagen de una persona o de preferencias políticas, se motiva a que la tendencia a responder, *No sabe* o *No opina*, sea mayor que si estas preguntas estuviesen ubicadas en un orden posterior.

Lo anterior puede ser explicado por la tendencia del encuestado a responder de manera tal, que no se comprometa o de acuerdo a lo esperado socialmente. Esta situación aparentemente es superada, si este tipo de preguntas se hace en un momento en el cual, quien responde se siente más en confianza y motivado con los temas de la encuesta. Ahora que se ha entrado en el tema de las encuestas de tendencias electorales, en este campo se debe ser muy cuidadoso. Una de las principales características de este tipo de sondeo de opinión es que son dinámicas, porque permanentemente están retratando una imagen de algo tan voluble como la opinión pública.

Las cifras de las tendencias electorales que barajan los medios de comunicación, son justamente eso, tendencias y no inclinaciones electorales definitivas. De todas maneras, cada sondeo refleja más una imagen que una intención de voto; es decir, cuando a meses de la elección el encuestado responde el cuestionario no está pensando en su intención electoral definitiva.

Los metodólogos Webb y Hatry, en su texto *Application of Citizens Survey to Local Governments*, dan una serie de recomendaciones acerca del uso de los estudios de opinión¹⁸:

- Cuidado con los estudios de opinión sobre temas complejos que la gran mayoría de los ciudadanos no conoce o no tienen mayor información.
- Cuidado con las respuestas que reflejan opiniones o apreciaciones prematuras o tempranas, producto de un momento coyuntural de crisis societal o personal; son opiniones que se deben tener en cuenta pero dentro de ese contexto, al igual que aquellas que reflejan situaciones a más largo plazo.
- Cuidado con los estudios de opinión sobre temáticas que pueden interferir o inclinar las preferencias en una contienda electoral próxima.
- Cuidado con el enunciado conceptual de una pregunta. Lo que se diga o se deje de decir puede prestarse para malos entendidos.
- Cuidado con resultados no representativos, producto de inadecuados procedimientos de selección, diseño o contenido de la encuesta. Ejemplo de esto puede ser el rechazo generalizado a responder una determinada pregunta o cuestionario. Si esto sucede puede deberse a la inadecuada ubicación de la pregunta o a una población muestral sesgada; lo anterior daría motivos a considerar inválidos los resultados obtenidos.

Dedicar atención a estas recomendaciones básicas, permite aumentar la probabilidad, para que el encuestado interprete adecuadamente las preguntas y así las respuestas sean confiables y validas. Los estudios de opinión ayudan a averiguar qué piensa la gente; pero su mayor importancia, a largo plazo, esta en que cumplen una función importante en el proceso de comunicación del ciudadano con las instituciones.

18 Webb, K.; Hatry, H. *Handbook of Research Design and Social Measurement*. New York: David McKay Company. Inc. 1984. Págs. 33-39

Ejercicio de auto-evaluación

Diseñe un cuestionario acerca de una problemática conocida por usted que combine varios tipos de preguntas.

Proceda a formular las preguntas referidas a las dimensiones – variables que usted haya seleccionado – teniendo en cuenta las sugerencias metodológicas acerca de cómo formular una pregunta, de tal modo que la información recogida en el interrogatorio sea confiable y válida.

Nuevamente se ofrece la alternativa de desarrollar, a partir de un problema de investigación conocido o afín, los procesos de diseño de un cuestionario orientado a obtener una información cuantitativa. Se recomienda elaborar el cuestionario teniendo en cuenta:

- Los objetivos del estudio que determinan la matriz de variables, que más tarde constituirán el conjunto de preguntas del cuestionario.
- Incluir necesariamente, si es el caso, un conjunto básico y exhaustivo de preguntas o ítems de identificación del encuestado, tales como: sexo, edad, estado civil, oficio, posición ocupacional, estratificación socioeconómica de la vivienda de residencia habitual, escolaridad, nivel de ingresos, etc.

Otro conjunto de preguntas se elaboraría, siguiendo el modelo CAP que recomienda formular las preguntas dentro del cuestionario de acuerdo al siguiente ordenamiento, primero ítems referidos a conocimientos, actitudes y finalmente a prácticas o comportamientos sociales del entrevistado sobre el tema de estudio.

Se sugiere en este diseño incluir diversos tipos de preguntas, según sean los objetivos del estudio como: intentar medir actitudes, utilizando escalas de Likert o de Guttman, o preguntas formuladas bajo el criterio de la metodología del diseño pentadimensional¹⁹. También deberá incluir algunas

19 El *Diseño Pentadimensional* (5 ítems): se utiliza básicamente en el estudio de opiniones y actitudes unidimensionales. La utilización de una secuencia de preguntas, cerradas y abiertas en el diseño metodológico exige que la única variable-dimensión a analizar se desagregue en cinco componentes: *Familiaridad con el tema, Expresión de las actitudes individuales, Reacciones a propuestas específicas. Razones por las que sostienen tales opiniones y Grado de intensidad de las opiniones.*

preguntas de control y otras a partir de algunas variables consideradas intervinientes.

A modo de ejemplo, en el caso específico de un estudio de auditoría de un servicio asistencial, se hace necesario no sólo identificar al usuario sino también *medir* el conocimiento que tiene el usuario sobre el servicio ofertado. Posteriormente, habría que incluir un conjunto de preguntas que permitan explorar su apreciación sobre la calidad del servicio, (actitud) medida por sus niveles de satisfacción de las expectativas que tiene del servicio.

En el caso anterior, el cuestionario deberá incluir un conjunto de preguntas o ítems que arrojen índices de satisfacción del usuario, entendidos como la calificación que éste hace sobre la calidad del servicio que recibe. En síntesis, el cuestionario deberá contener un conjunto de preguntas que faciliten la definición de factores generadores de satisfacción e insatisfacción y de las necesidades y expectativas de los usuarios

Como se podrá apreciar, un cuestionario no es sólo una compilación de preguntas o ítems, sino que su estructuración debe responder a los objetivos del estudio y a las características de la unidad de análisis. Pretender analizar un hecho o proceso social a partir de la información suministrada por ocho o diez preguntas no es algo que se pueda recomendar ni aceptar metodológicamente hablando.

5. Análisis de datos cuantitativos

Objetivo

Familiarizarse con la medición en la investigación social, donde se utilizan preferentemente los diseños cuantitativos.

La medición en los datos cuantitativos

Este capítulo muestra, en términos generales, algunas consideraciones importantes sobre los niveles y criterios de medición utilizados en los diseños cuantitativos.

El análisis de la información se efectúa a partir de una base de datos sistematizados por un programa especializado de manejo y gestión de datos. Es por ello, que el énfasis se centra más en la interpretación de los datos que en los procedimientos realizados para obtenerlos. El proceso de análisis no sólo se reduce a la lectura de unos datos, sino también a encontrar aspectos que caractericen, relacionen y diferencien lo observado, aspectos que configuren procesos con el fin de facilitar el estudio de una situación.

La medición en la investigación social, es el proceso mediante el cual una situación, hecho o comportamiento social es observado sistemáticamente y representado por puntajes alfanuméricos.

Comúnmente se considera que la medición de las variables deben ser *objetiva* para que los resultados sean siempre los mismos, independientemente de quien los haga, y *cuantitativa*, en el sentido de usar puntajes numéricos.

Los niveles de medición y los estadísticos apropiados a cada nivel

Cuando se utiliza un diseño cuantitativo en un proceso investigativo, se pueden encontrar que unos datos se requerirán para su tratamiento y

posterior análisis, diversos niveles de medición. Es importante entender sus características y diferencias, puesto que cada uno exige procedimientos estadísticos¹. Contrario a las ciencias naturales, en donde las convenciones de medidas son más precisas como: el metro, el gramo, el litro, etc., en las Ciencias Sociales, las convenciones conceptuales denotan propiedades que no son generalmente tan precisas ni estructuradas en su resultado.

La medición se puede hacer en cuatro niveles, de acuerdo con la forma como los números — cantidades que se asignan a los objetos o individuos — se relacionan con la propiedad o característica que está siendo medida. Éstos están íntimamente asociados con la presencia acumulada de las siguientes cuatro características: clasificación, ordenamiento, intervalos iguales y cero absoluto o continuidad.

- *El nivel nominal*. Es un criterio de clasificación, que no implica orden, ni jerarquía. Las unidades de análisis — individuos, regiones, propiedad, situación o hechos — son simplemente clasificados en dos o más categorías, ejemplo: “estado civil” es una medida de nivel nominal que permite clasificar a las personas en diversas categorías: casados, solteros, viudos, separados, unión libre, sin que esta implique un orden o jerarquía.
- *El nivel ordinal*. Este nivel indica no sólo un criterio clasificador, sino también un sentido de orden. En la medición ordinal se puede afirmar que la categoría X es más o menos (+ -) que la categoría Y.

Ejemplo: al preguntarle a los estudiantes sus apreciaciones sobre los servicios que presta el Departamento de Bienestar Estudiantil de una universidad. Uno de los indicadores de la calidad de la prestación de los servicios podría ser el grado de conformidad con el servicio de citas telefónicas. En este caso podría presentarse un continuo de respuestas preestablecidas que representan una escala de satisfacción:

Muy satisfecho - Satisfecho - No tiene criterio o no sabe - Insatisfecho
- Muy insatisfecho

El conjunto de respuestas es un buen ejemplo de una medición ordinal, puesto que se clasifica y ordena a las personas que respondieron: existe un orden entre los estudiantes que expresaron una actitud *Muy satisfecho* con

1 Generalmente, los niveles de medición nominal y ordinal utilizan técnicas no-paramétricas y los niveles intervalos o de razón utilizan técnicas paramétricas.

los programas asistenciales del Departamento de Bienestar Estudiantil, con los que expresaron una actitud *Muy insatisfecho*.

Si se le asigna el número 1 (uno) a los que expresaron estar *Muy satisfechos*, el número 2 (dos) a los que están *Satisfechos* y así hasta la última categoría, donde el número 5 (cinco) se asigna a los que expresan el mayor grado de insatisfacción, nos encontraremos tanto con un criterio de ordenamiento como también de clasificación.

Otro ejemplo de medidas ordinales son las *escalas de actitudes*², en donde al encuestado se le presenta una lista de ítems o proposiciones relacionadas con la situación que se estudia conjuntamente, con un continuo de respuestas que representa las posibles valorizaciones que le asigna a la dimensión. La escala puede variar en un rango *de mayor a menor, de acuerdo en desacuerdo, de satisfacción a insatisfacción*.

Por ejemplo, en la tabla 5.1 se muestran mediante una estrategia escalar, las actitudes de los medellinenses sobre las causas de violencia obtenidas, en un sondeo de opinión.

Tabla 5.1 Causas de aumento de la violencia

	Muy de acuerdo %	Acuerdo %	No opina	Desacuerdo %	Muy en desacuerdo
El aumento de la violencia es un hecho normal	4	37,2	2,9	46,5	8,8
La violencia son acciones aisladas exageradas por la prensa	5,5	48,6	5,8	35,5	4,6
La inseguridad y violencia son acciones de la delincuencia organizada	19,0	57,6	4,8	16,3	2,4

Fuente: M. Aigner. 1998

2 “Una actitud es una organización relativamente duradera de creencias en torno a un objeto o situación las cuales predisponen a reaccionar preferentemente de una manera determinada”. Rokeach, Milton. *The nature of human values*. New York: Free Press. 1973

La tabla anterior nos muestra el peso porcentual de las opiniones³ de un grupo de encuestados, frente a unas afirmaciones que se escuchan a diario, acerca de las posibles causas del aumento de la violencia en la ciudad de Medellín. Ahora bien, lo que interesa es individualizar estas actitudes ciudadanas y medir la intensidad. Si se considera que se le presentan nuevamente los tres ítems anteriores a un grupo de personas, y se les pide su grado de acuerdo o desacuerdo con las afirmaciones anteriores.

A cada encuestado se le presentó tres proposiciones o declaraciones referidas al tema, cada una con un continuo de posibles respuestas escalares que representaban una escala ordinal:

1. Muy de acuerdo; 2. De acuerdo; 3. Indiferente; 4. En desacuerdo; 5. Muy en desacuerdo.

Al sumar el valor de cada una de las respuestas, una persona podría recibir de tres a quince puntos: las personas que manifestasen estar *Muy en desacuerdo* con los planteamientos sobre las posibles causas del aumento de la violencia, podrían obtener un puntaje máximo de quince puntos; en cambio, aquellas que afirmasen, en todos los ítems, estar *Muy en desacuerdo* obtendrían un puntaje de tres puntos.

De esta forma podemos *ordenar* en una escala a cada persona encuestada, desde los que estaban *Muy de acuerdo* hasta los que están *Muy en desacuerdo*. Así se logra construir una escala ordinal.

- *Los niveles de medición de intervalo y de razón o proporción.* Estos permiten hacer mediciones exactas de distancia entre dos puntos cuyos intervalos son iguales. Las medidas de intervalo y de razón son generalmente producto de convenciones operacionales como: peso y estatura de las personas; horas de capacitación u horas de trabajo semanal; nivel de ingresos; distribución de la población por grupos de edades; años de escolaridad formal, etc. Los niveles de medición de razón o proporción, además de reunir todas las características anteriores, vale decir (clasificar, ordenar, distribuir en rangos o intervalos iguales entre categorías), tienen como punto de partida el cero absoluto, el cual indica que en la escala hay un punto de partida sin la propiedad o característica que se está midiendo. Esta situación es común en datos porcentuales o relativos, o en datos proporcionales:

3 Distribución de frecuencia

Otros ejemplos de estos cuatro niveles de medición son: Si un investigador tuviera que clasificar la afiliación partidista de los medellinenses utilizaría una escala nominal.

Un nivel de medición ordinal permite clasificar y ordenar las características socioeconómicas: lugar de residencia, nivel de ingresos familiares, estratificación socio económica y nivel de escolaridad de una población.

Si se quiere medir la edad en años cumplidos, de los estudiantes de la Facultad de Educación y Ciencias Sociales de una universidad, la escala más adecuada será la de razón. Y si lo que se desea es medir la población de mujeres en edad escolar agrupada en niveles etarios o de edades iguales (quinquenales), la escala más adecuada será la de intervalo.

Criterios de medición

La calidad interna y externa de la información depende en gran parte de criterios de validez, confiabilidad, representatividad, comprensibilidad y totalidad conceptual (operacionalidad conceptual); Estos criterios deben ser tenidos en cuenta por el analista social, desde el momento en que identifica un problema, hasta el proceso de análisis e interpretación de la información obtenida.

La Validez. Es el grado en el que un instrumento de medición realmente mide lo que debe medir. Un importante criterio para evaluar la calidad de un cuestionario radica en que lo que éste mida sea válido, ejemplo, si se desea medir la opinión sobre el pacto social de la Consejería Presidencial, se debe estar seguro que lo que se está sondeando es realmente la opinión ciudadana acerca de los planes y programas de dicho Proyecto, y no una opinión abstracta o global formada por la prensa o por sectores políticos o intelectuales.

¿Pero cómo se puede saber si la medición es realmente válida? En la práctica, no hay en las ciencias sociales un método único y confiable que permita concluir fehacientemente, que tal dimensión conceptual ha sido realmente medida. Existen unos criterios generales que se acercan más a lo intuitivo que a lo científico; por ejemplo, la técnica de validez de contenido, que exige que se abarquen todos los elementos o características posibles, que se suponen es parte de la dimensión conceptual que se pretende medir. Este criterio me-

todo lógico de validez se puede apreciar, cuando en un sondeo se pretende medir la opinión de los docentes de una universidad respecto al escalafón docente. Un sondeo de opinión como el planteado, no se puede reducir a medir sólo las apreciaciones acerca del recorrido y ascenso a través de los años, o sólo como un problema de méritos; se tendrían necesariamente que abarcar otras dimensiones, tales como: producción intelectual e investigativa, dedicación y experiencia profesional; capacitaciones y especializaciones logrados, etc.

– *Confiabilidad*. Una medida confiable es aquella que, aplicada sucesivamente a través del tiempo, producirá idénticos resultados, (se asume que el objeto medible no cambia a lo largo del tiempo). El criterio de confiabilidad es fuente continua de discusión y de sesgo en la investigación social; básicamente se manifiesta por el uso de preguntas vagas, complejas o no pertinentes con la temática que se está estudiando, o tal vez porque los encuestados no logran entender la pregunta. Aún las preguntas más sencillas pueden conllevar cierto grado de indeterminación y de desconfianza en las respuestas obtenidas.

¿Cómo se decide si una medida es confiable? Varias pruebas han sido desarrolladas para este fin. Una de las más utilizadas es el método de *prueba re-prueba*. Una medición confiable dará los mismos resultados cuando se aplica sucesivamente en diferentes tiempos, suponiendo que no se den cambios en el objeto-sujeto que se está midiendo.

¿Cómo se puede distinguir los criterios de validez de los de confiabilidad? El siguiente ejemplo puede ayudar a clarificar uno de otro.

En un estudio sobre la obra de equipamiento urbano del Instituto Metropolitano de Valorización, que comprende la ampliación de la frontera urbana (ampliación, apertura y amoblamiento de vías) en la zona suroriental por el sistema de valorización, se utilizó un diseño tipo panel, para evaluar las características socio económicas de los residentes.

Entre quienes fueron indagados con el mismo cuestionario cuatro meses más tarde, un porcentaje apreciable mostraron inconsistencias en sus respuestas sobre la edad declarada; el ingreso percibido; número de personas a cargo y bienes revelados. Si bien la mayoría se mantuvo firme en sus respuestas, esto muestra que la confiabilidad de la información obtenida en los llamados estudios de opinión es algo que se debe controlar o tener presente.

– *La Comprensibilidad*. Un analista generalmente está escribiendo informes para un lector que no es especialista en el tema; por lo tanto, es muy

importante escoger conceptos e indicadores que sean comprendidos por el ciudadano común. Medidas sobre conceptos o indicadores demasiado complejos, no importa que tan confiables y validos sean, no serán entendidos o tomados en cuenta.

– *Totalidad conceptual.* Es otra característica de una buena medición. Volvamos al ejemplo del estudio sobre seguridad comunitaria discutido; obviamente, la escala de cinco ítems o enunciados nos dio mejor información acerca de la actitud de los residentes con respecto al programa, que un cuestionario de una o dos preguntas. De igual forma, el uso de múltiples enunciados facilita la posibilidad metodológica de abarcar el máximo posible de apreciaciones, elementos o características de una dimensión.

Por ejemplo, se está interesado en estudiar la apreciación de los estudiantes sobre los espacios verdes de la ciudad universitaria. Una pregunta tan general como: ¿Qué opina sobre el estado de los espacios verdes en la ciudad universitaria? Sólo produciría confusión, porque algunos encuestados podrían pensar que se les pregunta sobre el uso de esos espacios, otros acerca del estado físico o de su mantenimiento. Lo anterior es un buen ejemplo de la necesidad de diseñar las preguntas de tal forma, que realmente investiguen acerca de lo que realmente se quiere indagar.

Este caso muestra que lo recomendable sería diseñar preguntas acerca de dimensiones específicas del entorno, de estos espacios verdes en los cuales el estudio está interesado. Tales dimensiones podrían ser: el estado o mantenimiento de los espacios verdes; la seguridad en ellos; el ornato; los accesos, la infraestructura disponible —asientos, mesas etc. Esto traerá ventajas también en los resultados, ya que aportará información específica. En síntesis, las respuestas que se obtengan dan cuenta en forma más integral o completa, de las dimensiones que se quieren analizar, si éstas abarcan la totalidad de las dimensiones del problema de investigación.

Sistematización de la información

Procesamiento de la información

El tratamiento de los datos cuantitativos es la fase del proceso de investigación posterior a la recolección de la información. Por lo general supone

un tratamiento informático de los datos que implica la digitación, grabación de los mismos en archivos de trabajo para su posterior manejo. Desde el momento que se comienza a estudiar el planteamiento de un problema de investigación hay que planear, dentro del diseño operacional, el proceso de sistematización de los datos. Este procesamiento requiere no sólo de un proceso de planeación, sino también de un plan adecuado de manipulación de datos, lo que determinará unos resultados que realmente sirvan para los objetivos del estudio.

El procesamiento de los datos es una necesidad muy usual en diversas áreas: gubernamental, de investigación, académica, etc., esto aunque los campos de aplicación sean diferentes, muchos de estos procesos son comunes. Dada esta estandarización de la demanda de necesidades, se han desarrollado programas especializados de manejo y gestión de datos. Existe una gran variedad de estos programas que difieren entre sí como: en aspectos de capacidad — número de datos y variables —, variedad de cálculos que pueden realizar y facilidad de su uso. Uno de estos programas es el Statistical Package for the Social Sciences (SPSS®), diseñado en forma modular; cada módulo con menús desplegados y cuadros de diálogo que permiten acceder a las distintas funciones que ofrece el programa. Este facilita dos de las tareas básicas del proceso de investigación social aplicada, es decir, la digitación y grabación de los datos obtenidos, y el análisis estadístico de los mismos.

Anteriormente se vio que los cuestionarios utilizan diversos tipos de preguntas. Interesa aquí analizar aquellos cuestionarios con preguntas cerradas (respuestas preestablecidas o precodificadas) y preguntas abiertas, en las cuales no se han previsto categorías de respuesta. Aquí interesa familiarizarnos con el tratamiento de las respuestas dadas a preguntas cerradas e incluso abiertas. Estas últimas exigen ser clasificadas en unas categorías de respuestas limitadas. En síntesis, la clave de un proceso de codificación consiste en no perder de vista qué es lo que se quiere conocer y una buena capacidad de síntesis y claridad conceptual.

Ante un conjunto de datos para analizar, surge la necesidad de elegir el tratamiento de sistematización, es decir, la elección de una metodología apropiada. El procesamiento de la información obtenida en el trabajo de campo, requiere de un plan adecuado de manipulación de los datos, para tal efecto, se deben realizar las siguientes actividades:

- *Validación de la información.* Al conjunto de actividades realizadas en la recolección de información recibe la denominación trabajo de campo.

Incluye la supervisión y el control de los errores de falta de respuesta, y el registro efectivo. Esto último, es lo que se denomina – *Validación de la información*. Es decir, una vez recolectada, se debe cotejar, corregir y validar mediante un procedimiento de crítica de la información registrada.

- *El proceso de codificación de la información*. Una vez validados los datos recogidos, se deben clasificar las respuestas individuales obtenidas. Esta etapa se denomina *codificación* y es de gran importancia, ya que la traducción de todas las respuestas a códigos alfanuméricos exige no sólo de un conocimiento del tema, sino también de los objetivos del estudio. Es un proceso básico ya que de él depende en gran medida la confiabilidad de la información recolectada. Los errores que se cometan en la codificación van a repercutir en el proceso de manejo y gestión de la información recolectada, ya que los valores de las variables van a quedar sesgados. Para evitar esto, el equipo de investigación debe diseñar un libro de codificación que será la base para codificar las respuestas obtenidas.

El proceso de codificación consiste, en la asignación de un código numérico para las diferentes categorías de respuestas para cada pregunta del cuestionario. El instrumento para transformar las respuestas en datos estadísticos, es el libro de codificación, cuadernillo que contiene la definición de las categorías de las variables en estudio y sus correspondientes códigos.

Algunas recomendaciones para tener en cuenta en el proceso de codificación son:

- Al diseñar los códigos se debe tener especial atención con los objetivos del estudio y la variedad de respuestas que pueden encontrarse en los cuestionarios.
- Recuerde siempre que por cada pregunta sólo se debe aceptar una sola respuesta, puesto que la mayoría de los programas estadísticos sólo reconocen esta forma, sin embargo, es posible que se acepten multirespuestas.
- Crear tantos códigos como posibles categorías de respuesta existan. No olvidar que un número superior a diez categorías dificultan enormemente la construcción de tablas y figuras. En el caso de clasificaciones de oficios, en donde es posible encontrar innumerables categorías, se sugiere crear tantos registros de preguntas como sea necesario, pero posteriormente recodificar o agrupar en no más de diez categorías, mediante un proceso de tipificación.

- Se recuerda que en los datos nominales, el valor numérico del código asignado no significa una jerarquía ni orden.
- *Diseño del libro de codificación.* Se recomienda diseñarlo de tal forma que aquellos que realicen el trabajo de codificación, tengan una guía precisa de cómo asignar un código a cada uno de los valores de las categorías de respuesta de cada pregunta (variable). El libro permite traducir cada categoría de respuesta a un número preestablecido (código numérico), el cual es vertido en la hoja de codificación. Un ejemplo de un libro de codificación se puede ver en la tabla 5.1.

Tabla 5.1 Ejemplo de un libro de codificación

Variables	Categorías	Código
Sexo	Masculino	1
	Femenino	2
	Soltero	1
Estado civil	Casado	2
	Viudo	3
	Separado	4
	Unión libre	5
Edad en grupos quinquenales	15 - 19	1
	20 - 24	2
	25 - 29	3
	30 - 34	4
	35 - 39	5
	40 - 44	6
	45 y más	7
Nº de años de vinculación	1 - 9	1
	10 - 19	2
	20 - 29	3
	30 - 39	4
	40 - 49	5
	50 y más	6

- *Transferencia de datos a una hoja de codificación.* Una hoja de codificación es una matriz formada por filas y columnas; una fila (horizontal) representa cada uno de los registros¹ del cuestionario y las columnas (verticales) representan el valor de cada variable.

Por ejemplo: la variable 1, *Sexo* es un campo que ocupa una columna en la hoja de codificación. En cada fila se anotará el valor que tome la variable *sexo* en cada uno de los sujetos del estudio. Los datos de la variable 2, *Estado civil* (campo que ocupa un espacio – un dígito –) se anotan en las filas de la columna 2 y así sucesivamente.

- *Generación de archivos de trabajo.* Superadas las etapas de crítica, validación y codificación de la información, se debe realizar la digitación contenida en las hojas de codificación en la base de datos. El proceso comprende las siguientes etapas: digitación o grabación, depuración y edición de los datos.

1 Cada registro representa un entrevistado