



Centro UC
Medición - MIDE

ORIENTACIONES PARA EL DESARROLLO DE ENCUESTAS

CENTRO DE MEDICIÓN MIDE UC



Santiago, Julio de 2017

mideuc.cl



ÍNDICE

1. PRESENTACIÓN.....	3
Modo de trabajo	3
2. ORIENTACIONES DE LA COMISIÓN	4
Definiciones Comunes	5
3. MARCO CONCEPTUAL	8
4. MODELO DE ANÁLISIS	10
5. CONSTRUCCIÓN DE PREGUNTAS	15
Diseño de Preguntas	16
Tipos de preguntas	18
Desafíos de una buena pregunta	20
Recomendaciones generales.....	26
6. MUESTREO	28
Muestreos Probabilísticos	30
Muestreos No Probabilísticos.....	33
7. Nota Final.....	35
8. BIBLIOGRAFÍA.....	36
9. ANEXO 1.....	37
Encuesta sobre Hábito Nutritivos y Vida Saludable	37
10. Anexo 2.....	40
Nota metodológica sobre información que debería reportarse al informar resultados de encuestas.....	40

1. PRESENTACIÓN

Modo de trabajo

Este documento fue elaborado como resultado del trabajo de la comisión de encuestas, conformada por representantes de distintas áreas, que funcionó durante el año 2016 y principios del 2017.

La comisión fue requerida por la Directora Ejecutiva del Centro de Medición, y liderada por MIDE Sociedad. En ella se buscó generar orientaciones en base a criterios comunes, para organizar de mejor forma el trabajo de diseño, aplicación y reporte de encuestas.

Participaron de esta comisión Flavio Cortés y Sofía Retamal (Área Sociedad¹), José Manuel de la Maza (Área Organizaciones), María Paz Espinosa (Área Educación), Francisco Silva (ADP) y como invitados a sesiones, David Torres y Susana Valenzuela. A cargo de la coordinación del trabajo de la comisión y el documento final estuvo Flavio Cortés y, mientras que responsables por temas fueron: Marco conceptual, José Manuel de la Maza; Modelo de Análisis, María Paz Espinosa; Construcción de encuestas, Sofía Retamal y Mariella Concha; y Muestreo, Flavio Cortés.

Siendo muy amplia la gama de tópicos y aspectos en el ámbito de encuestas², se definieron aquellos que se consideraron como temas prioritarios para la labor de definición de criterios:

1. Marco conceptual de la encuesta
2. Modelo de análisis de encuestas
3. Construcción de preguntas de encuestas
4. Muestreo

El trabajo se organizó asignando al encargado de cada tema las siguientes tareas:

- Revisar las encuestas realizadas por MIDE UC en los aspectos que el tema consideraba.

¹ También participó Mariella Concha que en ese entonces (1er semestre de 2016) formaba parte de MIDE Sociedad.

² No se abordó el proceso de pilotaje del cuestionario de una encuesta debido a que otra comisión de trabajo de MIDE UC se encargó de este aspecto específico. En todo caso se contó con la presentación de Francisco Silva al respecto.

- Revisar y analizar los estándares de MIDE respecto al tema (en especial el sistema de gestión de calidad, SGC)³.
- Buscar estándares de referencia más específicos en la bibliografía o en recomendaciones de organismos internacionales.
- Generar una propuesta a presentar y debatir en sesión.

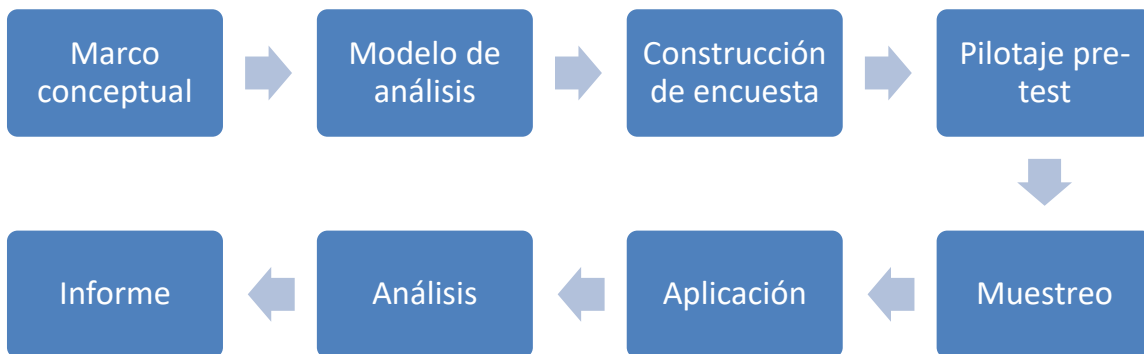
De este modo se quiso abordar aspectos que parecen más críticos, aunque también se tuvo en cuenta elementos pragmáticos, es decir, tiempo y experiencia acumulada. Cualquier otro proceso podrá ser abordado en una futura comisión. Para el desarrollo del presente documento se tuvo en consideración las clases que dicta el equipo de Investigación de Mide UC en el curso de “Introducción a la Medición en las Ciencias Sociales” (2016), incorporando terminología y referencias aportadas por los docentes, por lo que se recomienda leer este documento en conjunto con las clases dictadas (http://www.mideuc.cl/intranet2/?page_id=42).

2. ORIENTACIONES DE LA COMISIÓN

La comisión no buscó generar estándares o normas de trabajo, ya que para eso están los establecidos en el modelo de gestión de calidad. Lo que sí se propuso fue aportar con conocimiento técnico a quienes se inician en la elaboración de encuestas, o a quienes quieren sistematizar mejor su trabajo, dando orientaciones y recomendaciones específicas que pudieran ser de utilidad en distintos momentos del proceso de elaboración de encuestas.

El proceso de elaboración de encuestas puede ser entendido en base a las siguientes etapas:

³ Por cierto, queda abierta la posibilidad de revisar estas recomendaciones en el caso que MIDE UC decida revisar la forma en como explicita y norma un sistema que promueva la calidad de la gestión.



Cada etapa puede ser subdividida, y si bien el proceso es presentado aquí como lineal, habrá instancias en las cuales será necesario o recomendable volver a etapas anteriores. Un ejemplo común es cuando después del pilotaje se vuelve nuevamente sobre la etapa de construcción - creando nuevas preguntas-, o diseños de muestreo que en realidad pueden realizarse en una etapa previa o posterior, es decir independientemente de la construcción de preguntas.

Definiciones Comunes

Dado que suele usarse el término encuesta para distintos referentes, se entregan algunas definiciones que sirven de punto de partida semántico:

Encuesta: Es una técnica de estudio cuantitativa que se produce en un proceso que generalmente contempla un muestreo, levantamiento de datos y análisis estadísticos. También llamada encuesta social (denominación normalmente usada en ciencias sociales y particularmente por la sociología para investigación social mediante encuestas). Este método recoge información primaria de una población y, a diferencia de otras técnicas, lo que se conoce no son individuos, sino individuos-valores (de las respuestas) como indica Asún (en Canales, M. Ed. 2006). Otra definición que pone énfasis en la estandarización y la capacidad de generalizar los datos plantea que una encuesta es “un método de investigación por el cual la información es recabada al presentarle a las personas preguntas sobre un tópico específico y el procedimiento de recolección de los datos es estandarizado y se encuentra bien definido (siendo) la información obtenida acerca de un

subconjunto de la población de interés con la intención de generar estadísticas que sean generalizables al total de la población” (Groves, Fowler, Couper, et al., 2009)

Cuestionario: Es el instrumento que, en el caso de la encuesta, es fundamental como medio de levantamiento de datos. Está confeccionado para ser respondido por un participante, y se trata de una “conversación” en que las opiniones del entrevistado se registran en forma cuantitativa cuando la encuesta se realiza cara a cara. Por cierto los cuestionarios también pueden administrarse en plataformas online o telefónicamente. También los cuestionarios pueden incluir escalas de actitudes en su interior (Canales, M. Ed., 2006). El cuestionario y los protocolos que definen cómo este se aplica, permiten estandarizar el procedimiento de levantamiento de la información, buscando que este proceso no se sesgue por las particulares condiciones en que ocurre el levantamiento (situación de la entrevista, características del entrevistador, etc.).

A partir de estas definiciones podemos entender una encuesta como la técnica de indagación cuantitativa en que, utilizándose un cuestionario principalmente integrado por preguntas cerradas y semicerradas y un procedimiento sistemático y estandarizado de aplicación, se busca obtener información que interese al investigador sobre los comportamientos, actitudes, valores, creencias u opiniones de una población. Lo que hace a esta técnica tan usada es que, si la selección de los participantes a los que se les aplica el cuestionario recurre a un muestreo probabilístico⁴, los datos que se obtengan pueden ser generalizados desde el grupo más reducido que se encuestó, a la población general o sea son datos representativos de la población⁵.

Por otra parte, una indagación puede utilizar cuestionarios y no tratarse propiamente de una encuesta. Una distinción que resulta aclaratoria al respecto es la que Cronbach (1960) realiza entre instrumentos de *desempeño máximo*⁶, y de *desempeño típico*. Son instrumentos de

⁴ En el muestreo probabilístico se habla de una selección aleatoria de las muestras en el que cada elemento de la población tiene una probabilidad conocida no nula de ser seleccionado, con lo cual cada elemento de la muestra represente a un sector de la población y su totalidad a toda la población (Pimienta Lastra, Rodrigo, 2000)

⁵ Por cierto, de la definición anterior no se desprende que las encuestas solo deben utilizar técnicas de muestreo representativo. Cuando no es posible contar con una muestra representativa, los datos que estas arrojen no podrán ser atribuidos al total de la población y sólo describirán al grupo respondiente..

⁶ Que miden habilidades y logros y estos últimos tienen que ver con desempeños de contenidos que son explícitamente enseñados en procesos educacionales o de capacitación (Cronbach, 1990)

desempeño máximo las pruebas de conocimientos que son explícitamente enseñados en procesos educacionales o de capacitación y los test de habilidades que miden el mejor desempeño en un área que no ha sido explícitamente enseñada. En estos instrumentos evaluativos hay respuestas correctas e incorrectas y en ellos se espera un esfuerzo de desempeño.

Los instrumento de desempeño típico miden conductas que son típicas para el participante (características de personalidad, intereses o actitudes) y allí no existen respuestas correctas o incorrectas y se espera respuestas que representen, lo más cercanamente posible, lo que las personas piensan, hacen o creen. En ambos instrumentos se pueden usar cuestionarios pero como se ve con muy diferentes características y propósitos.

Las encuestas, utilizando la categorización anterior, se acercan más a un instrumento de desempeño típico pues interrogan sobre preferencias, opiniones, conductas, intereses e información factual y algunas veces incorporan escalas de actitudes (para relacionarlas con opiniones, conductas preferencias o intereses). Incluso en algunos casos se le presentan al respondiente preguntas de conocimiento (por ejemplo, en las encuestas políticas grado de conocimiento de figuras políticas), pero no se espera que el respondiente haga un máximo esfuerzo para contestar correctamente, ni la pregunta se formula para prioritariamente evaluar el grado de conocimiento del participante encuestado.

De esta forma, escalas o inventarios con fines de evaluación individual o grupal donde se espera medir un aprendizaje, comportamiento o *desempeño* del participante que responde usan un cuestionario, pero el marco conceptual requiere un modelo que describa el aprendizaje, comportamiento o desempeño, por lo que el diseño del análisis y las decisiones de aplicación (muestra), varían considerablemente de las normalmente usadas para una encuesta social.

A continuación se entregarán recomendaciones para las distintas fases componentes de una encuesta: generación de marcos conceptuales, modelo de análisis, la construcción del cuestionario y la estrategia de muestreo.

3. MARCO CONCEPTUAL

El desarrollo de un marco conceptual sirve al propósito de definir el o los conceptos centrales a ser evaluados y para identificar las principales variables que deben ser consideradas para responder a las preguntas del estudio. Las variables de un cuestionario pueden ser las más simples y comunes, tales como edad, sexo, NSE, o más específicas al tema de investigación, como por ejemplo en un estudio de satisfacción, definir las dimensiones que componen la satisfacción en un cierto dominio (por ejemplo la atención en salud, la satisfacción estudiantil, etc.)

Estándares actuales: La elaboración de un marco conceptual está contemplada en el sistema de gestión de calidad de MIDE UC en términos generales. Por ello, en este proceso, se entregarán consejos útiles para quien se enfrenta a la siempre compleja tarea de desarrollar una conceptualización de constructo. Esta generalmente supone una negociación con el requirente basado en los objetivos del proyecto. En casos en que esto no sea posible, como ocurre en las licitaciones, se recomienda explicitar en la propuesta, las preguntas y productos a entregar.

Algunas de las recomendaciones son:

1. Determinar cuáles son las **inferencias claves** que se quieren formular con el instrumento a construir (las que se deben constituir a partir de las preguntas que la investigación busca responder).
2. Identificar qué **modelo teórico y modelo de análisis** nos permitirá responder dichas preguntas, lo que a su vez permitirá definir el diseño del cuestionario, qué análisis hacer, y la cantidad y tipo de evidencia empírica que se requiere para responder.
3. Explicitar el modelo teórico que identifica **atributos o variables relevantes** que deben ser consideradas en el proyecto e incorporadas en la encuesta.

Como se observa, el proceso de desarrollo de un marco conceptual debe estar basado en un proceso de definición con el requirente, una revisión de literatura científica especializada, y un análisis inicial que permita orientar y prospectar el análisis posterior.

La revisión de literatura suele ser un proceso complejo, en que resulta difícil distinguir lo relevante de lo que no lo es. Para apoyar la revisión, pueden ser útiles las orientaciones de Hernández-Sampieri, Fernández-Collado y Baptista, (2014).

- i. ¿Se acudió a un par de bancos de datos y se pidieron referencias por lo menos de cinco años atrás?
- ii. ¿Se indagó en directorios, motores de búsqueda y espacios en internet?
- iii. ¿Se revisaron por lo mínimo cuatro revistas científicas que tratan el tema de interés? ¿Se consultaron de cinco años atrás a la fecha?
- iv. ¿Se buscó en algún lugar donde había tesis y disertaciones sobre el tema de interés?
- v. ¿Se localizaron libros sobre el tema en al menos dos buenas bibliotecas (físicas o virtuales)?
- vi. ¿Se consultó con más de una persona que sepa algo del tema?
- vii. Si aparentemente no se descubrieron referencias, ¿se contactó a alguna asociación científica del área en la cual se encuentra enmarcado el problema de investigación?

Otra útil guía para revisar literatura es la que ofrece Rowley y Slack (2004) en el artículo “Conducting a Literature Review”.

Como fuentes de literatura se recomienda realizar búsquedas a través del sistema de bibliotecas UC. Periódicamente hay talleres en estas bibliotecas para aprender a buscar, además de información disponible sobre cómo hacer búsquedas en bases académicas. Entre las búsquedas, resulta una buen práctica el incluir documentos de organismos internacionales como BID, OMS, UNICEF, UNESCO, dependiendo de la temática. Las tesis suelen ser excelentes referentes para identificar bibliografía relevante de una temática o problema de estudio. Es muy útil consignar en esta etapa también, y a partir de la revisión, otras encuestas que sirvan como fuente de validación, así como fuente de insumos para construcción, y referencia para el análisis.

4. MODELO DE ANÁLISIS

El plan o modelo de análisis es la propuesta de cómo se responderán las preguntas a partir de las técnicas de análisis de datos.

Estándares actuales: La elaboración de un diseño de análisis está contemplada en el sistema de gestión de calidad en términos generales. En SGC, para la organización de la construcción se pide definir un plan y metas, y en el análisis psicométrico de ítems, se indican análisis que debieran hacerse normalmente (confiabilidad, puntaje estándar y estructura factorial).

Desde la perspectiva de la comisión, la elaboración del modelo de análisis se puede dar como un proceso más diferenciado de lo que actualmente aparece en el sistema de gestión de calidad. En este sentido, se propone un modelo de análisis que se desarrolle en continuidad con el marco conceptual o teórico.

Si el **marco conceptual** identifica las variables o atributos relevantes que deben ser considerados en el cuestionario, el **modelo de análisis** debe responder a la pregunta acerca de cómo es que cada uno de esas variables/atributos será incorporado en el análisis de los datos empíricos y, por lo tanto, indica cómo se utilizará en concreto el modelo teórico en los análisis.

1. Para una buena definición del atributo a medir se recomienda tener en cuenta que el constructo o atributo no es lo mismo que la/las preguntas. El atributo es latente (no lo podemos medir directamente) y entre atributo e indicadores hay un proceso de asignación. El atributo puede ser psicológico, sociológico, de la salud, etc... Ejemplo de estos serían “buen docente”, “clima organizacional”, “empatía”, “confianza política”, “calidad de vida”, “hábitos saludables”, “bienestar empresarial”, mientras que a las preguntas se le asigna la capacidad de medir la variable latente.
2. La medición puede involucrar la medición de un atributo o varios, o de un atributo multi o unidimensional. Para definir cómo se debiera medir es relevante la revisión de literatura, no pudiendo tenerse a priori por mejor una opción que otra. Por ejemplo,

hay mediciones unidimensionales de satisfacción estudiantil, y también multidimensionales, encontrándose puntos de controversia entre una opción y otra.

3. El atributo a medir debe ser definido según si será una variable cuantitativa, ordinal o categórica. Esto es una definición y no un dato “dado” (p.e. ¿“inteligencia” es un atributo cuantitativo o categórico?) Para ello se debe revisar literatura y justificar la legitimidad de medir el atributo de un modo u otro.

Un aspecto fundamental para determinar las opciones de análisis, es la de identificar si se trata de un estudio a) descriptivo o inferencial, b) si se analiza una o más variables, y c) el tipo de variables de que se trata. En este último punto, y a modo de corregir errores comunes en la elaboración de un modelo de análisis, se recomienda distinguir cómo se está tratando la variable en el nivel del *atributo*, de los *datos recabados* y *del análisis*. Así por ejemplo podemos concebir que la confianza política es un atributo que se expresa cuantitativamente de 1 a 100, pero en una encuesta la medición se operacionaliza en una escala ordinal del tipo “mucho-algo-poco-nada” y finalmente en el análisis se puede plantear una situación de “confianza versus desconfianza”. Entonces no necesariamente habrá continuidad entre el carácter categórico, ordinal o cuantitativo entre la conceptualización del atributo, los datos recabados y el análisis que se hace con los datos recabados. A modo de recordatorio, los tipos de variables puede verse en la siguiente tabla.

Ilustración 1: niveles de las variables, tomado del curso de Introducción a la medición

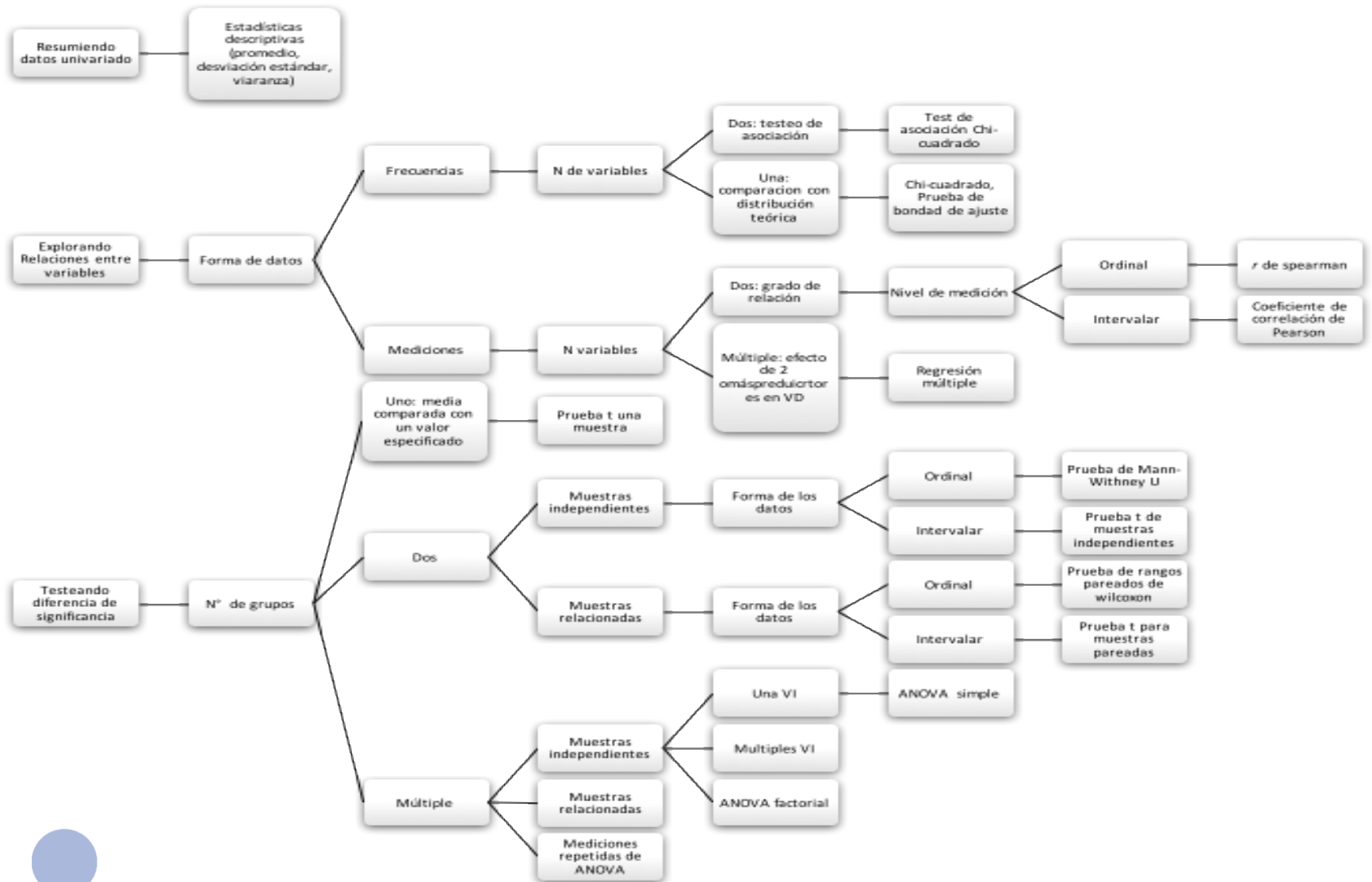
Tipo A	≠	Tipo B	≠	Tipo C	:	categorial
Bajo	<	Medio	<	Alto	:	ordinal
2		3		4	:	cuantitativo

Como planteamos, una variable del atributo puede ser de distinto tipo en los datos o ser modificado al momento del análisis. Un ejemplo de esto es el desempeño docente: esta es una variable **ordinal** (donde se concibe que se puede ser mejor o peor al desempeñarse); la variable

desempeño se mide **cuantitativamente** en una escala de 1 a 4; y a partir del análisis se representa tales puntajes en **categorías**: destacado, competente, básico e insatisfactorio.

Luego de definir el atributo, cabe especificar los análisis posibles de realizar (plan de análisis), los que a su vez determinarán los reportes. Para seleccionar el análisis adecuado, se recomienda tener en cuenta los modelos de análisis, descriptivos, pruebas estadísticas o técnicas multivariantes, como se presenta en la ilustración 2, donde VD es Variable dependiente y VI variable independiente, y NM nivel de medición. Se recomienda, como un recurso útil para elaborar un modelo de análisis, usar modelos gráficos que representen las relaciones posibles entre variables.

Ilustración 2 : Árbol de técnicas de análisis.





5. CONSTRUCCIÓN DE PREGUNTAS

Este proceso incluye el diseño y evaluación de las preguntas (cerradas, semicerradas o abiertas) a incluir en un cuestionario con un determinado modo de aplicación (presencial, telefónico u on-line).

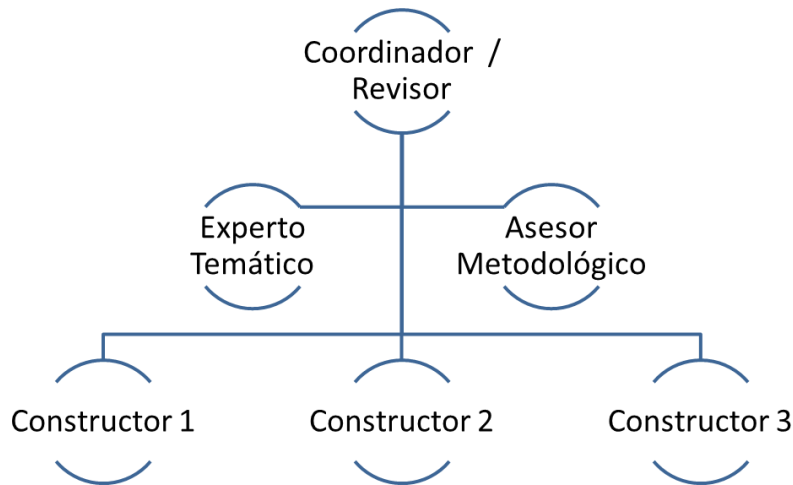
Estándares actuales: Aparecen indicaciones muy generales en los procesos del sistema de gestión de calidad (PS.03: Desarrollo de Instrumentos o Procedimientos), en particular, se encuentra definido que se planifique el proceso de construcción incluyendo el número de ítems por dimensión.

Además, se describen procesos de construcción según la organización del trabajo. En este punto se refieren a construcción individual y en equipo. En este último caso, se plantea el análisis ente pares y la figura de un coordinador/responsable y visto bueno del director de área. Adicionalmente también existe la figura del experto que valida el instrumento en su versión final.

Desde la perspectiva del equipo de trabajo, y dada la experiencia de MIDE Sociedad al respecto, se propone un proceso de construcción con las siguientes funciones: coordinador-revisor, constructores, asesor metodológico y experto temático. El rol de asesor metodológico puede encontrarse dentro de la misma institución, siendo probable que en algunos casos el experto temático sea un profesional que se deba buscar fuera de la institución.

En este proceso de construcción y pilotaje, el coordinador de construcción se puede asesorar directamente en un experto metodológico.

Esquema de Construcción de Preguntas



Se asigna la tarea de construcción de preguntas a profesionales mientras el experto temático propone orientaciones y revisa las preguntas construidas básicamente desde el contenido de las mismas, mientras el asesor metodológico analiza si las preguntas están bien formuladas desde un punto de vista de coherencia lógica y metodológica. El experto temático puede en casos excepcionales contribuir en la construcción de preguntas y el experto metodológico puede ser una figura senior en la unidad a cargo de la construcción como un Jefe de Proyecto o el Director del Área.

Diseño de Preguntas

La construcción de preguntas se concibe como una tarea que es consecuencia de todos los pasos anteriores, es decir, de las decisiones y definiciones tanto conceptuales como de análisis, entre otros. Esto es lo que suele plasmarse en la “tabla de especificaciones”. Más concretamente, una tabla puede desglosarse desde el concepto (atributo, variable), las dimensiones, indicadores y categorías de respuesta. La tabla de especificaciones suele ser más extensa y las preguntas por construir cubren una parte del concepto. Un ejemplo de lo anterior puede advertirse en la tabla 1.

Tabla 1. *Especificaciones violencia política*

CONCEPTO	DIMENSIONES	INDICADORES/VARIABLES	CATEGORÍAS
Violencia política en manifestaciones públicas	Actitud ante la violencia política	Posicionamiento ante la afirmación: "Cuando una persona es detenida por desórdenes en una manifestación, es porque lo provocó"	Muy de acuerdo; De acuerdo; Término medio; En desacuerdo; Muy en desacuerdo
		Posicionamiento ante la afirmación: "Todas las marchas políticas terminan en desmanes"	Muy de acuerdo; De acuerdo; Término medio; En desacuerdo; Muy en desacuerdo
		Posicionamiento ante la afirmación: "los que se involucran en política tienden a crear conflictos innecesarios"	Muy de acuerdo; De acuerdo; Término medio; En desacuerdo; Muy en desacuerdo
	Percepción de la violencia política	Valoración de la existencia de violencia política en la sociedad	0-10 (0=No hay violencia política; 10=Hay mucha violencia política)
		Percepción de la evolución de la violencia de género	Antes había más violencia política; Igual; Ahora hay más violencia política
	Agentes de la violencia	Actores responsables de la violencia	Sociedad; externos (infiltrados); policía; gobiernos, manifestantes
	Perfil político	Tendencia política	Izquierda; centro; derecha
		Sexo	Hombre; Mujer
		Edad	18-100
		Generación	Nacido antes del golpe 1973; nacido después del golpe 1973
Nivel de estudios		Sin estudios; Estudios no oficiales; Estudios primarios; Estudios secundarios; Estudios universitarios	

Es importante considerar antes de construir preguntas, que la formulación de ellas puede variar en su forma dependiendo si ellas son formuladas para ser presentadas en un cuestionario autoaplicado o para formar parte de una batería de preguntas a ser aplicadas por un tercero, sea presencialmente o telefónicamente. Al respecto, se recomienda tenerse en mente las siguientes consideraciones:

1. Los cuestionarios autoaplicados deben ser simples y entregar TODA la información que se requiera para responder, sea en la pregunta como en otros apoyos semánticos, como títulos de un conjunto de preguntas (p.e. "Evaluación de servicios médicos"), bajadas de

texto (p.e. “para responder esta pregunta considere las experiencias negativas”) o explicaciones ad-hoc de cómo debe entenderse la pregunta.

2. Los cuestionarios de aplicación telefónica deben ser concisos por la expectativa del encuestado de uso de su tiempo menor al de una encuesta presencial. Idealmente las preguntas debieran ser leídas y esperar la respuesta espontánea, sin que sea necesario leer las alternativas de respuesta (pero ello supone que el entrevistador está capacitado para hacer la selección de la alternativa adecuada).
3. Los cuestionarios aplicados mediante encuesta presencial permiten mayor tiempo y más preguntas, el uso de tarjetas y otros medios de apoyo para responder preguntas (pero son mucho más onerosos en su aplicación).

Tipos de preguntas

Resulta relevante que la decisión de los tipos de preguntas tome en cuenta las múltiples opciones de diseño. Para una mejor distinción, proponemos diferenciar las preguntas por su uso, su contenido, su nivel de medición y sus alternativas de respuesta, como se presenta en las siguientes tablas.

Tabla 2. Clasificación de preguntas según su uso

Filtro, control y rompehielo	Consistencia y amortiguador	Sustantivas y batería
<ul style="list-style-type: none"> • Filtro: son las que se incluyen previamente a fin de eliminar a los participantes que no les afecte ésta. • De Control: su finalidad es asegurarse el interés y buena fe del encuestado y la veracidad y fiabilidad de sus respuesta. • Rompehielos: para iniciar el cuestionario e interesar al encuestado con él, o para pasar de un tema a otro 	<ul style="list-style-type: none"> • De Consistencia: para comprobar la consistencia de las respuestas del encuestado. Suelen ser preguntas similares pero redactadas de forma diferente y presentadas de forma espaciada para comprobar la congruencia de las respuestas • Amortiguador: preguntas indirectas que se usan para atenuar respuestas en temas difíciles antes de formular preguntas más directas 	<ul style="list-style-type: none"> • Sustantivas: son las preguntas para cuestiones básicas y las referente a las cuestiones investigadas • Batería de Preguntas: conjunto de preguntas sobre la misma cuestión, que se complementan unas a otras, enfocando diversos aspectos del tema a indagar.

Fuente: elaboración propia

Tabla 3. Clasificación de preguntas según su contenido

Relativas a comportamientos	Relativas a Información factual	Relativas a actitudes, opiniones, expectativas , intereses, creencias
<ul style="list-style-type: none"> Se pregunta respecto a una conducta o comportamiento que el participante ha realizado o está realizando (p.e. ¿votó usted en las últimas elecciones?) Hay que precisar que se obtiene el reporte de la conducta no la conducta en sí 	<ul style="list-style-type: none"> Se obtiene información sobre el respondiente o su familia respecto a datos como sexo, ingreso, bienes y servicios que posee, vivienda, etc. 	<ul style="list-style-type: none"> Se pretende caracterizar al participante en cuanto miembro de una categoría respecto a actitudes, opiniones, expectativas , intereses, o creencias

Fuente: elaboración propia

Tabla 4. Clasificación de preguntas según su respuesta

Preguntas abiertas	Preguntas cerradas de respuesta múltiple	Preguntas cerradas de respuesta única
<ul style="list-style-type: none"> Ventaja de no restringir la opinión de los participantes a una serie de alternativas prefijadas, pero con costos mayores de procesamiento posterior (codificación). En un cuestionario hay una o muy pocas de estas preguntas. 	<ul style="list-style-type: none"> Se presenta una serie de alternativas de respuesta, donde el encuestado puede elegir más de una En general sirve para estudiar temas no excluyentes. 	<ul style="list-style-type: none"> Se permite escoger sólo una alternativa Alternativas deben: ser excluyentes (que no se solapen); ser exhaustivas (que cubran todas las posibles opciones).

Fuente: elaboración propia

Tabla 5. Clasificación según nivel de medición

Respuesta nominal dicotómica y no dicotómica	Respuesta ordinal	Respuesta intervalar
<ul style="list-style-type: none"> Serie acotada de respuestas que no se encuentran ordenadas entre sí. Ejemplo: ¿podría decirme la religión a la que usted se siente más cercano?, alternativas de respuesta: a) católica, b) evangélica, c) mormona, d) etc. 	<ul style="list-style-type: none"> Las alternativas de respuesta constituyen un continuo ordenado desde la máxima aceptación, frecuencia o cercanía a un tema, hasta el máximo rechazo, lejanía o mínima frecuencia de ocurrencia del fenómeno. Es usual que la pregunta no esté redactada en forma interrogativa, sino como 	<ul style="list-style-type: none"> Las alternativas de respuesta están constituidas por un continuo numérico frente al cual el participante debe posicionarse. La ventaja de este tipo de preguntas es que permite una amplia variabilidad de respuestas con sólo una pregunta, siendo más sensible a

Respuesta nominal dicotómica y no dicotómica	Respuesta ordinal	Respuesta intervalar
<ul style="list-style-type: none"> Existe el riesgo de no ofrecer una alternativa que interprete las opciones del respondiente 	<p>una afirmación.</p> <ul style="list-style-type: none"> Se debe decidir cuántas alternativas incluir y si se redacta una alternativa intermedia. 	<p>las actitudes, deseos o conductas del participante, al contrario que las preguntas dicotómicas.</p>

Fuente: elaboración propia

Desafíos de una buena pregunta

La Comisión reconoce que el proceso de construcción de preguntas en la literatura es poco visible y parece ser la parte más “artesanal” del diseño de un cuestionario. Sin embargo este proceso requiere mucha atención, ya que es fuente de error no estadístico.

Como estándar de buena pregunta, se propone tomar en consideración lo siguiente respecto de los participantes a encuestar (Fowler y Fowler, 1995):

- Definir objetivos y especificar el tipo de respuesta para obtener lo que se espera de la pregunta.
- Asegurarse de que todos los participantes dispongan de un conocimiento común del significado de la pregunta.
- Asegurarse que los participantes puedan saber cómo responder en base a su conocimiento disponible. Algunas barreras para ello son:
 - Nunca haber tenido la información necesaria para responder la pregunta.
 - Tener algún tipo de información, pero estar imposibilitado para recordar la información con exactitud o el detalle requerido por la pregunta.
 - Dificultades en la exactitud de ubicar eventos en el tiempo.
- Asegurarse que los participantes puedan contestar en los términos de las opciones de respuesta.
- Hacer preguntas para que los participantes estén dispuestos a contestar con precisión.

Por la importancia de la temática, y el vasto tipo de preguntas, la comisión sugiere se tenga en vista como paradigma la “psicología de la respuesta”. En este marco se consideran las respuestas

a preguntas de una encuesta como una actividad compleja, identificando el tipo de pregunta y los procesos cognitivos asociados al acto de responder. Tourangeau, Rips y Rasinski (2000) proponen analizar el proceso cognitivo de elaborar una respuesta a una pregunta en distintas fases:

- Codificación y almacenamiento de la información
- Comprensión de la pregunta
- Recuperación de la información
- Juicio y estimación
- Reporte de respuesta

En el análisis secuencial graficado más abajo (ilustración 3), es posible observar que se pueden producir procesos no lineales. También se pueden producir “baches” en cualquier punto, o vueltas atrás que afectarán la respuesta (cuando las opciones de respuesta no representan la opinión del respondiente, este puede volver atrás en busca de nueva información en su memoria). También, y a pesar de haber hecho toda la secuencia, se podría preferir no dar la respuesta elegida, sino elegir otra por deseabilidad social.

Ilustración 3. Proceso de respuesta



Veremos cada fase con algunas dificultades y sugerencias:

En la fase de **comprensión** el encuestado:

- Atiende a la pregunta y a las instrucciones que la acompañan; le asigna un significado inicial a la pregunta e identifica la información buscada por el encuestador

Dificultades

El encuestado puede distraerse durante la encuesta y no poner atención durante la formulación de la pregunta, perdiendo parte de su formulación. La persona puede no tomar nota de la información relevante o incluso, advirtiéndolo que hay instrucciones, no darse el tiempo para leerlas (si la encuesta es auto-aplicada). Por otra parte, la pregunta puede incluir términos poco familiares para el respondiente, o términos que pueden ser entendidos de diferente forma por distintas personas. También la pregunta puede estar formulada con una sintaxis compleja o contener detalladas explicaciones que sean difíciles de entender. Incluso términos familiares pueden acarrear dificultades (p. e., “hermanos”, ¿incluye hermanastros?).

Sugerencias

- Administrar los estímulos -visuales, la pregunta, las opciones de respuesta, su diagramación- de manera que el encuestado se concentre en la pregunta.
- Hacer preguntas simples, evitando sintaxis complicadas (oraciones subordinadas, etc...) Preguntas que cubran múltiples posibilidades se pueden descomponer en preguntas simples que cubran una sola posibilidad.
- Evitar conceptos vagos y entregar ejemplos cuando estos conceptos deben ser mencionados.
- Reemplazar adjetivos vagos en rangos que especifican posibilidades exactas de probabilidad o frecuencia.

En la fase de **recuperación** de la información relevante el encuestado:

- Rememora información relevante de la memoria de largo plazo adoptando una estrategia de recuperación de los datos, generando claves específicas para gatillar

recuerdos y trata de completar las memorias parciales que tiene de los hechos realizando inferencias

Dificultades

El proceso de recuperación involucra generalmente acceder al recuerdo de información relevante de la memoria de largo plazo. Normalmente al recordar hay inferencias (por ejemplo no se recuerda cuánto se consume cierto alimento diario sino que se estima). Esto aplica también a preguntas de actitud (p. e., la pregunta ¿considera que los diputados usan bien su tiempo?, supondría recordar eventos que informen sobre aquello). El “material” del recuerdo tiene características que afectan este proceso:

- a) cuán distintivo es el evento a recordar (p. e., es diferente inquirir a una madre sobre las veces que se ha vacunado al hijo en el último año que la cantidad de veces que ha acudido a una sucursal bancaria en similar período dado la distintividad que tiene el primer tipo de preguntas);
- b) el grado de ajuste entre los términos preguntados y la codificación original en la memoria;
- c) si el recuerdo es muy lejano o más bien cercano en el tiempo (p.e. conductas realizadas la semana pasada en contraposición a aquellas que se realizaron tres años atrás);
- d) si el recurso de memoria es de primera o segunda mano (p.e., las encuestas de hogar suelen preguntar al jefe o jefa de hogar por información de la pareja e hijos).

Muchas de las preguntas de recuerdo tratan de que se recuerde una conducta o eventos junto con una regularidad u ocurrencia, p.e. “¿cuántas veces ha sentido dolor en la última semana?”

Sugerencias

- Estimular el recuerdo haciendo preguntas similares a la pregunta objetivo.
- Ayudar a recuperar eventos preguntados para un tiempo situado usando calendarios, puntos de referencia personal, eventos de vida, etc...

- Adaptar la duración del tiempo de referencia a las propiedades del evento (p. e. preguntar por consumo de bebidas diario, no mensual)

En la etapa de **juicio o evaluación** el encuestado usa información relevante para:

- Realizar juicios considerando la completitud o exactitud de la recolección de información, juicios sobre las inferencias realizadas, juicios sobre las inferencias usadas para completar brechas de memoria; integración de todos los elementos recuperados en un solo juicio general.

Dificultades en preguntas sobre datos y duraciones

- Las personas tienen dificultad recordando fechas exactas y pueden reportar eventos que ocurrieron antes del tiempo de referencia especificado en la pregunta; también incorrectamente, pueden conjeturar una fecha con un rango incierto de tiempo, y redondear información temporal a valores típicos.

Dificultades en preguntas de actitud

- Los respondientes pueden no tener ninguna opinión sobre lo que se les pregunta.
- Las respuestas pueden verse afectadas por el fraseo de las preguntas o su ordenamiento.
- Las personas pueden moderar o retener un juicio si perciben que la información con que cuentan no es suficiente.

Sugerencias

- Evitar preguntar por “lo habitual”, ya que puede haber habitualidades simultáneas (p.e si trotó tanto algunos fines de semana como días de semana, y distintas opciones de habitualidad son trotar el fin de semana, y trotar los días de semana, podría elegir ambas).
- En preguntas de actitud es recomendable usar escalas que tengan componentes positivos y negativos (p.e. me fastidia leer textos de estudio complejos; las clases expositivas me aburren; me entusiasman los talleres de estudio donde se puede opinar y reflexionar grupamente).

En la etapa de **reportar y seleccionar respuesta** el participante:

- Ubica su respuesta en la escala apropiada u opción de respuesta y “edita” la respuesta por consistencia, aceptabilidad u otros criterios.
-
- En ciertas categorías las personas difieren en su aproximación a dar una respuesta. Por ejemplo, pasar de “habitual” a “casi siempre” si preguntan por cuantas veces salimos a comer puede ser respondido con paradigmas distintos; y en preguntas de actitud, pasar de “totalmente de acuerdo” a “de acuerdo” también es variable según la manera de afrontar la respuesta. (p.e. cuántas veces comer fuera de casa pasa de ser “habitual” a “casi siempre”). Todo esto se relaciona con la estrategia usada por el respondiente: simplemente contestar lo más rápido posible la opción que le parezca más aceptable o tratar de encontrar la opción que los represente.
 - El nivel de disposición del encuestado para efectivamente responder la pregunta, puede ser bajo, lo que puede inducir a respuestas del tipo “no sé” o responder lo que se cree que es más aceptable para el encuestador o para la convención social (deseabilidad social).
-
- Sugerencias
 - Cuando sea posible, especificar las categorías de respuesta de manera que sea claro a qué se refieren, y no expresen una preferencia social (p.e que una cierta frecuencia sea positiva y otra negativa).
 - Generar un setting donde sea posible contestar con libertad y en casos de preguntas o temas que plantean problemas relacionados con temas morales o de intimidad personal, o aluden a actitudes u opiniones contra-culturales, asegurar confidencialidad y, si es necesario, disponer de formas de emitir opinión o preferencia sin el conocimiento del encuestador.

Recomendaciones generales

Adicionalmente se discutieron algunos tópicos que se reportan a continuación:

¿Cuántas preguntas por dimensión son adecuadas?

Hay que considerar que para una medida unidimensional un buen número de preguntas es 4 ítems, de modo que con esa cantidad ya se puede calcular un coeficiente de consistencia. .

¿Es necesario incluir la alternativa “otro”?

La alternativa “otro” puede ser usada inadecuadamente para cubrir varias alternativas que no se explicitan y finalmente en los análisis suele perderse como información. Es preferible –si es posible- identificar previamente las principales alternativas de respuesta mediante un estudio piloto.

¿Las escalas Likert son la única manera de preguntar por actitudes?

La comisión estima que es relevante considerar que las escalas Likert fueron creadas especialmente para medir actitudes⁷, para lo cual se detalla información que servirá para la mejor toma de decisión de construcción.

En escalas de actitud se pide al participante que señale, dentro de una serie graduada de ítems, su grado de aceptación o preferencia. No existe una respuesta válida y a diferencia de la pregunta de opinión donde se inquiriere sobre una postura más estática que representa una posición mental consciente y manifiesta sobre algo o alguien, la actitud implica una disposición a la acción, un estado de disposición psicológica que incita al individuo a reaccionar de una manera característica frente a determinados estímulos.

Las escalas de actitud son varias, pero comparten el tener dirección (positiva o negativa) e intensidad (alta o baja), siendo los métodos más usados el escalamiento tipo Likert y el diferencial semántico (escala en la que una persona evalúa un tópico valorándolos sobre la base de un conjunto de adjetivos bipolares –fuerte v/s débil, p.e.)

Las escalas de actitud pueden referir a importancia, acuerdo, verdad, frecuencia, prioridad, probabilidad, y aceptabilidad (una útil síntesis es la de Vagias, 2006).

⁷ Se alude a lo que actualmente se denomina medición explícita de las actitudes, que es el tipo de mediciones que mejor se adapta al formato de una encuesta. En contraste, en los últimos 20 años se han popularizado en la psicología social mediciones implícitas de las actitudes (especialmente el Implicit Association Test –IAT- que requiere un instrumental computacional para administrarlas).

Los tramos que se distinguen en una escala pueden ser etiquetados completamente (p.e. totalmente inaceptable, inaceptable, ligeramente inaceptable, neutral, ligeramente aceptable, aceptable, totalmente aceptable) a etiquetados extremos bipolares (p.e. totalmente inaceptable - totalmente aceptable) incluye definir cuántos niveles, y las categorías de los mismos. También se puede etiquetar con un solo concepto (etiquetado unipolar), por ejemplo, extremadamente feliz – algo feliz – nada feliz.

Los etiquetados también incluyen formatos más simples para niños o personas que no conocen el idioma o que sean analfabetos, como rostros esquematizados, como lo muestra el ejemplo de más abajo⁸.

Ilustración 4



Ante esta variedad la comisión propone lo siguiente:

a) **¿Cuántas alternativas incluir en una escala?**

Se sugiere considerar no incluir demasiadas alternativas en escala sin justificación. Se da como propuesta un mínimo de tres alternativas y un máximo absoluto de siete, aunque se insiste en consignar que esta decisión debe ser según lo que se requiere, habiendo escalas de 10 u 11 puntos. También está la opción que no definen alternativas en una escala sino que presenta un continuo “termómetro de respuesta”, en que sólo se identifican los extremos y el encuestado puede aplicar una marca en cualquier punto del continuo.

b) **¿Se debe redactar una alternativa intermedia?**

c) El debate es poco claro. No existe una recomendación precisa, aunque se sabe que colocar una alternativa intermedia tiene fundamentalmente dos defectos: 1) puede permitir a las personas no tomar posición o “fugarse” de tomar opciones difíciles o poco deseables

⁸ Escala usada en Nadia del Carmen Hernández (2015). Musicoterapia para el control de la ansiedad en los tratamientos odontológicos en niños con síndrome de Down. Tesis para optar al título de Odontopediatría. Universidad Científica del Sur. Lima, Perú.

(especialmente relevante al medir conceptos complicados como racismo, intolerancia, conductas sexuales, etc.) y, 2) no siempre encontramos en el español una palabra que exprese claramente el sentido intermedio que queremos. Por ejemplo “indeciso”, “no sé”, “en medio”, “ni de acuerdo ni en desacuerdo”, no necesariamente implican estar en la mitad del continuo. Se recomienda decidir considerando que si se omite la alternativa “neutral” se obliga al encuestado a definirse en una posición. También debe considerarse si es pertinente omitir un punto medio en relación al constructo que se busca medir. En todo caso, típicamente en la escala Likert de 5 punto se etiquetan todos los puntos para evitar que se tienda a preferir las alternativas etiquetadas de las no etiquetadas.

d) **Usar alternativa “no sé” con precaución.**

Puede ir al final de la escala, y en algunos casos sólo debe formularse si el encuestado no asume previamente ninguna posición.

e) Se propone **considerar el nivel educacional y edad** del participante para decidir usar una escala Likert basada en una afirmación, dado que requiere un proceso cognitivo más complejo. Una alternativa más simple si se requiere puede ser una escala graduada de otro tipo (frecuencia, agrado, etc.)

f) Escribir **preguntas con dirección tanto positiva como negativa** para evitar sesgo de aquiescencia, es decir, la tendencia a contestar preguntas de la misma forma, y efecto halo, tendencia a contestar del mismo modo en base a un atributo general. En este caso es importante considerar que preguntas negativas no son las que contienen un NO en una afirmación, sino que invierten la valoración, p. e. “en general me aburro en los eventos sociales”, cuando se quiere saber cuán intensa son las relaciones vinculares de una población.

g) Usar enunciados con afirmación para **un solo aspecto (simplicidad)**.

h) **Considerar que palabras o frases con absolutos** (siempre, nunca, nada) para los extremos de las escalas, puede implicar que se escojan menos por la tendencia a evitar extremos.

6. MUESTREO

Es el proceso para determinar la cantidad de unidades necesarias de conocer para generalizar en una población dada.

Estándares actuales: El procedimiento PS.04 “Aplicación de Instrumentos o Procedimientos” en el Instructivo “Criterios de Aplicación de Instrumentos” de MIDE UC no se refiere a la técnica de muestro, y el documento externo “Criterios de aplicación de instrumentos” se habla de muestra pero no de muestreo. Por su parte, en el documento de “Estándares Internacionales para la Medición y Evaluación” (de Mide UC, sept. 2009) se señala que “la composición de cualquier muestra de examinados a partir de la cual se obtenga evidencia de validación, debe ser descrita en detalle, incluyendo los principales aspectos sociodemográficos y características de desarrollo”. También se indica “los procedimientos utilizados para hacer el muestreo de los participantes seleccionados para hacer los análisis de confiabilidad y los estadísticos descriptivos que caracterizan a esta muestra también deben ser reportados, y cuando se desarrollan pruebas piloto debe documentarse los procedimientos para seleccionar la muestra y las principales características de ésta”. Finalmente, se señala que “cuando corresponda, la muestra debe ser lo más representativa posible de la población a la cual se dirige el test”. Este tipo de fraseo sugiere que se está aludiendo a muestreo no probabilístico y a un instrumento de desempeño máximo.

La comisión consideró que resulta pertinente dar una vista general que caracterice qué es un muestreo y los distintos tipos de muestreo existentes.

Una muestra es un conjunto de elementos (individuos, grupos, categorías) que pertenecen a un grupo mayor y que un investigador selecciona para estudiar, con la expectativa que estudiando el grupo más pequeño (muestra) se logrará información sobre el grupo más grande (población). Para poder generalizar afirmaciones (hacer inferencias) desde la muestra a la población, la muestra debe ser probabilística lo que significa que se debe realizar una selección aleatoria de las muestras en el que cada elemento de la población tiene una probabilidad conocida no nula (distinta de 0 ó 1) de ser seleccionado, con lo cual cada elemento de la muestra represente a un sector de la población y su totalidad a toda la población.

Una muestra seleccionada con un muestreo probabilístico es aquella que es similar a la población de donde es extraída en todos los aspectos, salvo en el tamaño. Evidentemente, al extraer una

muestra de una población y estimar un valor para cualquier aspecto (p.e., la media de la altura de los participantes), va a haber alguna variación entre, en este caso, la media de la altura de la muestra -estadígrafo- y la media de la altura de la población -parámetro-, dado que al generar una muestra se selecciona sólo una de todas las muestras posibles y es en base a esa muestra, y no a todas las muestras posibles, que se realiza la estimación. La ventaja de usar una muestra probabilística es que se puede estimar la magnitud esperada del error muestral y éste refiere a la “dispersión del error de muestreo que ocurre cuando se muestrea repetidamente una población” (Ritchey, 2008, p. 211)

Por el contrario, en el caso de muestras no probabilísticas no existen reglas para estimar la magnitud de dicho error (y por eso, las encuestadoras en esos casos usan expresiones como que el error **habría sido** del 3% si la muestra hubiera sido aleatoria).

El paso inicial del proceso de muestreo es constituir un marco muestral, consiste en un listado de todos y cada uno de los elementos que conforman la población. A continuación, se debe seleccionar el tipo de muestreo, entre muestreo aleatorio simple; muestreo sistemático; muestreo aleatorio estratificado (con variantes de muestreo proporcional estratificado); muestreo de conglomerados. Veremos brevemente cada uno.

Muestreos Probabilísticos

Muestreo Aleatorio Simple

Como su nombre lo indica, es el más simple de los muestreos y consiste en que cada elemento de la muestra tiene igual probabilidad de ser seleccionado en ésta (método de selección de igual probabilidad). Una forma de extraer los elementos de la población para constituir la muestra consiste en utilizar una tabla de números aleatorios (que es una lista de número dispuestos de manera aleatoria), o un programa informático que genere números aleatorios (<https://www.random.org/>) y el analista o investigador tiene entonces que asignar previamente

una numeración a los elementos de la muestra para poder asociar cada elemento al número que la tabla o programa reporta. También se puede usar Excel con la función “Aleatorio”⁹.

Muestreo Sistemático

Técnica usada para muestrear aleatoriamente listas de elementos, sea que la lista esté ordenada o no. Se determina el tamaño del intervalo de la muestra (k), luego se selecciona un punto de partida aleatorio entre 1 y k , y a continuación se selecciona cada elemento para incluirlo en la muestra. Para que esta técnica presente el error estándar asociado a un muestreo simple se debe precaver que la lista esté aleatoriamente ordenada. Para que quede más claro, podemos suponer un ejemplo de 100 trabajadores en un centro de estudios de los cuales se desea seleccionar una muestra de 10, entonces se divide la población ($N=100$) por la muestra deseada ($n=10$) y el resultado es el intervalo de la muestra ($k=10$). Para obtener el punto de partida se puede utilizar la tabla o programa de números aleatorios y el primer número que ésta arroje entre 1 y 100 es el primer elemento seleccionado, y el segundo seleccionado es ese número más k y así sucesivamente hasta tener los 10 elementos de la muestra.

En este tipo de muestreo si la lista de elementos que constituye el marco muestral se ordena de acuerdo a los niveles de una variable categórica (p.e. sexo masculino y femenino; religión católica, protestante, mormona, otros; nivel socio-económico bajo, medio y alto) o a los valores de una variable cuantitativa, el muestreo será más eficiente que el muestreo sistemático simple descrito en el párrafo anterior. Esto ocurre porque al generar estratos internamente homogéneos se reduce el error estándar. A este subtipo de muestreo se le denomina muestreo sistemático estratificado.

Muestreo Aleatorio Estratificado

Técnica por la cual la población es dividida en grupos mutuamente excluyentes (estratos) y luego se seleccionan, mediante muestreo aleatorio simple o muestreo sistemático, muestras de cada

⁹ =ALEATORIO.ENTRE(indicar número menor;indicar número mayor)

uno de los estratos (p.e.: en una muestra de establecimientos educacionales establecimientos particulares pagados, particulares subvencionados y públicos). La variable que ordena la población dividiéndola en estratos se denomina variable de estratificación (en el ejemplo, la dependencia administrativa de los establecimientos) y puede ser una variable categórica (religión o sexo), o cuantitativa (ingreso o edad). Como es evidente este tipo de muestreo requiere que los estratos se constituyan a partir de las variables de interés (dependientes) en la encuesta y en ese caso este muestreo es más eficiente que el aleatorio simple.

Un caso particular de este tipo de muestreo es el muestreo estratificado proporcional (también denominado muestreo estratificado con afijación proporcional) en que el analista diseña un marco muestral estratificado, luego determina los tamaños muestrales por estrato de modo que estos tamaños son proporcionales a los tamaños de los estratos poblacionales, y luego selecciona una muestra aleatoria del tamaño apropiado en cada estrato. Obviamente cuando se usa esta técnica se deben conocer las proporciones poblacionales de la o las variables de estratificación y éstas deben estar relacionadas con el objeto de interés del estudio, caso en el cual es conveniente utilizarlo porque es más eficiente que el muestreo aleatorio simple.

Sin embargo, hay casos en un investigador necesita seleccionar una muestra estratificada no proporcional, es decir, una muestra que sus proporciones respecto de la variable de estratificación son diferentes de las proporciones de la población. Las razones para realizar este tipo de muestreo es comparar grupos más que realizar generalizaciones estadísticas desde la muestra a la población, o cuando en la población existen grupos muy pequeños que probablemente no tendrán elementos seleccionados en la muestra usando otras técnicas de muestreo (p.e. población de origen étnico aymara en una muestra general de la población). La prevención que hay que tener al usar este muestreo es que no se pueden hacer generalizaciones directamente desde la muestra a la población a menos que se usen factores de ponderación de los datos, entregándole pesos mayores a los estratos más grandes y pesos menores a los estratos más pequeños, para reflejar de ese modo las reales proporciones de los estratos en la población.

Muestreo de conglomerados

Este es un tipo de muestreo en donde se selecciona aleatoriamente conglomerados a partir de un marco muestral de conglomerados (ciudades, escuelas, comunas, vecindarios, etc...) más que de elementos individuales (alcaldes, estudiantes, vecinos). Un conglomerado entonces es una unidad colectiva que incluye múltiples elementos. Una vez seleccionada aleatoriamente una muestra de conglomerados desde la población, se requiere desarrollar marcos muestrales para aquellos conglomerados (no siendo necesario desarrollar dicho marco para toda la población). El caso más simple de esta técnica es monoetápica y supone seleccionar aleatoriamente un conjunto de conglomerados de la población y luego incluir todos los individuos de los conglomerados seleccionados en la muestra. En una muestra de conglomerados bietápica, primero se seleccionan aleatoriamente un conjunto de conglomerados, y luego, en la segunda etapa, una muestra de elementos es extraída aleatoriamente de cada uno de los conglomerados seleccionados en la primera etapa. Dado que no todos los conglomerados pueden ser del mismo tamaño, se pueden administrar procedimientos donde la probabilidad de seleccionar el conglomerado sea proporcional a su tamaño.

A diferencia del muestreo estratificado, donde es deseable la homogeneidad de los elementos del estrato, en el muestreo por conglomerados es deseable la heterogeneidad de los elementos componentes del conglomerado. Este tipo de muestreo requiere que los conglomerados sean de gran tamaño, siendo sumamente útil cuando se requiere muestrear población geográficamente dispersas y se requiere aplicación de encuestas presenciales (p. e. un muestreo por conglomerados podría dejar fuera las regiones extremas y centrarse en las de mayor concentración poblacional en una encuesta nacional).

Muestreos No Probabilísticos

En algunos casos es posible utilizar muestreos no probabilísticos en tanto el objetivo perseguido **no** sea lograr un cuadro representativo de la población. En estos casos la muestra es seleccionada incluyendo personas que estén disponibles o que pueden ser reclutadas para los fines del estudio,

y ello es legítimo en tanto no se hagan generalizaciones estadísticas imputables a la población a partir de los resultados obtenidos en la muestra.

Existen diferentes **tipos de muestreos no aleatorios** y acá distinguiremos el muestreo por cuotas, el muestreo intencionado, el muestreo por referencias, y el muestreo voluntario.

El primero (**muestreo por cuotas**) involucra la identificación de grupos en la población para luego definir el número de personas que debería ser incluida en cada grupo (cuota) para formar una muestra “ideal”, siendo la selección de las unidades de la muestra realizada por conveniencia hasta llenar la cuota de cada grupo.

En el caso del **muestreo intencionado**, el analista o investigador conoce las características de la población objetivo y luego busca participantes específicos que tengan dichas características para incluirlos en la muestra.

El **muestreo por referencias o referenciado** es usado principalmente por aquellos analistas que tienen como propósito investigar poblaciones de difícil detección o selección de participantes y para contrarrestar aquello se recurre a que la muestra se constituya a través de referencias que puedan dar los participantes. Se distinguen así dos técnicas: el **muestreo por red** (Granovetter, 1976) y el **muestreo por “bola de nieve”**. En el primero el investigador comienza obteniendo una muestra probabilística de una población que es probable que tenga algún grado de relación con la población objetivo y consiguientemente cada unidad de la muestra reporta otro participante que podría formar parte de la muestra. En el caso del muestreo por bola de nieve se le pide a cada persona que reporte a otras que podrían encuestarse y así se engrosa de forma no lineal el número de participantes.

Finalmente, en el **muestreo voluntario** el investigador realiza una convocatoria a través de un medio que puede ser un cartel, aviso en medios de comunicación o mensaje en redes sociales, para participar en la aplicación de un instrumento, en este caso, una encuesta.

Por cierto, al describir los distintos tipos de muestreo la Comisión no pretende que el analista, jefe de proyecto o director de área dominen la técnica de muestreo y realicen en todos los casos la selección de los participantes de la población. Sólo ello es posible con muestras acotadas en el caso del muestreo aleatorio simple o del muestreo sistemático. En los casos que se requiere un proceso más sofisticado de distinguir si la técnica de muestreo conveniente es por conglomerados o estratificado o en los casos que se requiere al muestrear múltiples pasos (bietápico o trietápico) es necesario el asesoramiento por un profesional calificado.

7. Nota Final

Al finalizar este documento, en el Anexo 1 entregamos una ficha de encuesta, que ejemplifica los diferentes aspectos aquí mencionados y que puede ilustrar cómo enfrentar los diferentes requerimientos al hacer una encuesta y en el Anexo 2 se entregan una serie de recomendaciones que en la literatura se presentan sobre qué aspectos deberían ser informados al reportar los resultados de una encuesta para que el lector pueda tener un juicio informado sobre los alcances y el valor científico de la encuesta. Como es evidente, la mayor parte las encuestas en nuestro país no reporta ni una mínima parte de los aspectos allí mencionados lo que implica que las personas e investigadores no pueden tener un juicio informado sobre los datos presentados.

8. BIBLIOGRAFÍA

Canales, M. (Ed.) (2006) Metodologías de investigación social. Lom ediciones. Santiago.

Cronbach, L. J. (1960). Essentials of psychological testing. New York: HarperCollins Publishers

Fowler J, y Fowler, Jr. (1995). Improving Survey Questions. Design and Evaluation. Center for Survey Research, Boston: University of Massachusetts

Granovetter, M. (1976). Network Sampling: Some First Steps. The American Journal of Sociology, 81, 1287-1303

Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., Tourangeau, R. (2009). Survey Methodology. Hoboken: John Wiley & Sons, Inc.

Hernández – Sampieri, R, Fernandez-Collado, C., Baptista, P. (2014) Metodología de la Investigación. México D.F.: McGraw Hill

Pimienta Lastra, Rodrigo (2000) Encuestas probabilísticas vs. no probabilísticas. Política y Cultura, núm. 13, 2000, pp. 263-276. Universidad Autónoma Metropolitana Unidad Xochimilco, D.F. México. Recuperado el 04.07.2017 desde <http://www.redalyc.org/articulo.oa?id=26701313>

Ritchey, Ferris J. (2008). Estadística para las ciencias sociales (ed. 2), Madrid: McGraw Hill.

Rowley, J. y Slack, F. (2004) Conducting a Literature Review. The American Journal of Maternal/Child Nursing. 27 (6), 31 – 39.

Tourangeau, R., Rips, L.J., Rasinski, K. (2000). The Psychology of Survey Response. New York: Cambridge University Press.

Vagias, M. (2006). Likert-type scale response anchors. Clemson International Institute for Tourism & Research Development, Department of Parks, Recreation and Tourism Management. Clemson University. Recuperado el 03.06.2017 desde <https://www.uc.edu/content/dam/uc/sas/docs/Assessment/likert-type%20response%20anchors.pdf>

9. ANEXO 1

Encuesta sobre Hábito Nutritivos y Vida Saludable

Una empresa comercializadora de alimentos saludables quiere conocer cuan saludables se comporta la población en general. Se define el requerimiento como “Conocer hábitos de vida saludable y nutrición de la población por sexo y género”.

Marco conceptual

Para conformar este se revisan medidas de vida saludable de la Organización Mundial de la Salud, y otras organizaciones, y se analizó con la contraparte según los propósitos.

El modelo conceptual propone un concepto de hábitos “conductas recurrentes que inciden en bienestar físico, mental y social” en base a un concepto de promoción.

Se define promoción desde la OMS, consignando sus prerrequisitos o condiciones fundamentales.

Se revisa un estudio en Colombia que agrupó hábitos saludables en **conductas de riesgos** para 6 dimensiones.

Se encuentra coincidencia entre hábitos saludables y comportamientos detectados por la FAO en una tendencia social actual (aumento de la prosperidad).

Se recogen datos de la Encuesta Nacional de Salud de Chile, el Sondeo de Vida Sana del INJUV levantando conclusiones relevantes como la diferencia entre hombres y mujeres, edad y hábitos alimentarios.

Se aborda la dimensión de **hábitos alimentarios**, consignando la definición de la OMS y describiendo lo que es el IMC, además de describir gráficos de la Escuela de Salud Pública de Harvard respecto a lo que debe considerarse alimentación saludable (pirámide de alimentación saludable y plato de alimentación saludable) describiendo sus recomendaciones.

Luego se aborda la dimensión de **actividad física y ejercicio** consignando lo que indica la Organización Mundial de la Salud (OMS). Además, la contraparte solicita considerar la diferencia entre lo que las personas **perciben** de su propio comportamiento, con una medición objetivo de comportamientos saludable.

En resumen, las variables relevantes son las que permiten distinguir en distintos grupos de la población su autopercepción global y los hábitos de vida saludables en las dimensiones de

hábitos alimentarios, actividad física conductas de riesgo. A partir del marco conceptual se define la tabla de especificaciones.

Modelo de análisis

En el proyecto mencionado anteriormente para la empresa comercializadora de alimentos saludables se preparó una tabla de especificaciones que consigna datos relevantes para construir las preguntas así como para analizarlos. Los hábitos saludables quedan dimensionados, y la **autopercepción** como una variable de dos medidas: antes y después (al inicio y al final del cuestionario, habiendo respondido preguntas sobre hábitos). Se formulan las siguientes preguntas en torno a la **autopercepción al inicio** del cuestionario:

- ¿Existen diferencias entre la autopercepción previa y los índices sobre alimentación y actividad física respectivamente?.
- ¿Cuán asociadas están las diferentes mediciones sobre alimentación, actividad física y vida saludable?.
- ¿Existen diferencias en la autopercepción previa y los índices de alimentación y actividad física al diferenciar por sexo, grupo etario y/o grupo socioeconómico (GSE)?
- ¿Existen diferencias en la autopercepción previa y los índices de alimentación y actividad física entre hombres y mujeres, entre, grupos etarios y/o GSEs? (por ejemplo, las mujeres obtienen un índice más bajo que su autopercepción, en cambio los hombres no).

La variable de autopercepción se definió mediante distintos niveles en las etapas del proyecto.

- a) Concepto de atributo en nivel categórico (poco saludable/saludable/muy saludable),
- b) Preguntas y datos resultante en nivel cuantitativo (cuán saludable son sus hábitos en escala de 1 a 7), y
- c) En el Análisis se desarrolla una transformación para poder comparar puntajes en base a una *transformación lineal*, donde el mínimo de la escala original (1) se transforma en el mínimo de la escala comparable (0) y el máximo 7 se convierte en 3 respectivamente. Luego por cada punto de la escala original, la escala comparable aumenta medio punto. De esta forma tenemos (1=0, 2=0.5, 3=1, 4=1.5, 5=2, 6=2.5, 7=3). En este caso se mantiene como variable cuantitativa pero en el reporte previsto se propone presentarla como variable categorial colapsando en dos niveles: a) los

que consideran llevar hábitos de vida saludable versus b) los que no creen llevar hábitos de vida saludable.

Ficha de encuesta caso ejemplo empresa nutrición

Universo: Personas de ambos sexos, de grupos socioeconómicos ABC1, C2, C3 y D (se excluye el grupo socioeconómico E o de extrema pobreza), de entre 18 y 64 años de edad, residentes en hogares particulares del Gran Santiago y principales ciudades de Chile (Antofagasta, Viña del Mar, Valparaíso, Concepción, Talcahuano y Temuco).

Población representada: 5,9 millones de personas según las proyecciones de población para el año 2013 del INE-CELADE (ajustadas de acuerdo a los porcentajes de GSE usados de manera estándar por el mercado nacional), lo que corresponde al 60% de la población urbana adulta del país.

Método de Recolección: Encuesta presencial en hogares particulares.

Muestra: 809 personas*. Selección probabilística en sus tres etapas (manzana-hogar-persona) con muestreo aleatorio y estratificado no proporcional por grupo socioeconómico y ciudad. Se aplica tabla aleatoria tipo Kish para la selección de la persona a encuestar en caso que en el hogar visitado haya más de una que cumpla la condición de género y edad en el hogar.

Nivel de precisión: El error muestral se estima en $\pm 3,4\%$ considerando varianza máxima y un 95% de confianza.

Fecha de Terreno: La encuesta se aplicó entre el 23 de Julio y el 13 de Agosto de 2013.

Instrumento de Medición: Cuestionario compuesto por 52 preguntas breves, en su mayoría cerradas.

10. Anexo 2

Nota metodológica sobre información que debería reportarse al informar resultados de encuestas

Las encuestas deberían ir acompañadas de una nota metodológica donde se explicita la siguiente información:

1. Se debe proporcionar acceso a los ítems del cuestionario utilizados ya sea en el cuerpo del informe, en un anexo o en un link.
2. Se debe informar de las propiedades psicométricas de la herramienta, reportando la fiabilidad y/o validez de un instrumento de encuesta recién desarrollado.
3. Se debe describir el proceso de desarrollo y/o describir los métodos utilizados para la pre-prueba de la herramienta (pilotaje).
4. En caso que los instrumentos de la encuesta requirieron puntuación, proporcionar una descripción de los procedimientos de puntuación.
5. Los estudios deben describir la representatividad de la muestra (presentar un cálculo del tamaño de la muestra u otra justificación del tamaño de la muestra) y especificar el margen de error de la encuesta para un determinado nivel de confianza.
6. Se debe describir el modo de administración de la encuesta (es decir, en persona, por correo, Internet, teléfono o una combinación de éstos).
7. Se debe proporcionar información sobre el número de intentos de contacto realizados y cuántos de éstos fueron exitosos (tasa de rechazo).
8. Se debe describir quién fue identificado como la organización o grupo que financia y cual fue presentado a los participantes de investigación para su participación en la encuesta.

9. Se debe informar acerca de los métodos utilizados para analizar los datos (es decir, una descripción de las variables que se analizaron, cómo se manipularon y los métodos estadísticos utilizados).
10. Se debe proporcionar los métodos utilizados para el análisis del error de no respuesta, el cálculo de las tasas de respuesta o el manejo de los datos de los elementos que faltan.
11. Se debe proporcionar una definición o límite de corte para completar parcialmente los cuestionarios (cuándo un cuestionario se considerará en el análisis y cuán se le desechará).
12. Proporcionar una tasa de respuesta definida y definir la tasa de respuesta reportada.
13. Analizar y reportar si los no respondientes difirieron de los encuestados.
14. Resumir los resultados con respecto a los objetivos.
15. Explicitar los procedimientos de consentimiento para los participantes de la investigación.
16. Si la encuesta forma parte de un proceso de investigación científica, informar si el estudio recibió la revisión del correspondiente consejo de ética.