

UNIVERSIDAD DE GRANADA
Programa de Doctorado en Ciencias de la Educación
Facultad de Ciencias de la Educación
Departamento de Didáctica de la Matemática



COMPRENSIÓN DEL MUESTREO POR ALUMNOS CHILENOS DE EDUCACIÓN SECUNDARIA

Tesis Doctoral
Karen Andrea Ruiz Reyes

Dirigida por:
Dr. José Miguel Contreras García

Granada, 2019

Editor: Universidad de Granada. Tesis Doctorales
Autor: Karen Andrea Ruiz Reyes
ISBN: 978-84-1306-427-7
URI: <http://hdl.handle.net/10481/59296>

UNIVERSIDAD DE GRANADA
Facultad de Ciencias de la Educación
Departamento de Didáctica de la Matemática

**COMPRENSIÓN DEL MUESTREO POR ALUMNOS CHILENOS DE
EDUCACIÓN SECUNDARIA**

Tesis Doctoral

Memoria realizada bajo la dirección del Doctor José Miguel Contreras García en el Departamento de Didáctica de la Matemática de la Universidad de Granada, que presenta **Karen Andrea Ruiz Reyes** para optar al grado de Doctor en el Programa de Doctorado de Ciencias de la Educación.

Fdo: Karen Andrea Ruiz Reyes

V° B° del Director

Dr. José Miguel Contreras García

RESUMEN

El propósito de esta investigación es evaluar la comprensión de las nociones básicas sobre muestreo por alumnos chilenos de educación secundaria. El creciente interés en los temas de inferencia como parte de la alfabetización estadística, ha motivado la elección de este tema como parte de esta tesis doctoral, es decir, este estudio, se dirige a completar los estudios anteriores desarrollados en otros contextos, para proporcionar información en el contexto chileno, donde no hay investigación previa sobre el razonamiento de los estudiantes en este tema.

Como marco teórico se emplea el enfoque ontosemiótico del conocimiento y la instrucción matemáticos, que aporta una serie de elementos que permiten caracterizar el significado institucional presente en las orientaciones curriculares como en los textos escolares chilenos, contribuyendo además a compararlo con el significado personal de los estudiantes que participan en esta investigación.

En una primera parte, se ha realizado un análisis semiótico de los objetos primarios relativos al muestreo presente tanto en las directrices curriculares chilenas, como también se ha elaborado una comparación con otras directrices a nivel internacional, además de establecerlos elementos del marco teórico que forman parte de esta investigación.

A continuación, se establece una visión general de las investigaciones previas que abarcan la comprensión del muestreo, útiles para establecer un banco inicial de preguntas para posteriormente elaborar el cuestionario de evaluación.

Como se mencionó anteriormente, se emplean además las herramientas del enfoque ontosemiótico para realizar un análisis de los objetos primario presentes en una serie de textos escolares chilenos de educación secundaria, para poder comparar si existe alguna relación o diferencia con los propuestos en las orientaciones curriculares.

Posteriormente, se describe el proceso de construcción y validación del cuestionario de evaluación de la comprensión del muestreo, considerando la información recogida del análisis de antecedentes como también del análisis de los textos escolares.

Por último, se presentan los resultados finales obtenidos al aplicar el cuestionario a una muestra de 1241 estudiantes de secundaria chilenos, de 8º año de primaria (13-14 años); 2º año de secundaria (15-16 años) y de 4º año de secundaria.

ABSTRACT

The aim of this research was to evaluate the understanding of the basic notions about sampling by Chilean secondary school students. The growing interest in inference issues as part of statistical literacy has motivated the choice of this topic as part of this doctoral thesis, that is, this study is aimed at completing previous studies developed in other contexts, to provide information about the Chilean context, where there is no previous research on students' reasoning of in this topic.

The theoretical framework used is the ontosemiotic approach to mathematical knowledge and instruction, which provides a series of elements that allow characterizing the institutional meaning present in curricular guidelines as in Chilean school textbooks, also helping to compare it with the personal meaning of students who participate in this research.

In the first part, a semiotic analysis of primary objects related to sampling has been carried out in both the Chilean curriculum guidelines, as well as a comparison with other international guidelines, in addition to establishing elements of the theoretical framework that are part of this research.

The following is an overview of the previous research covering sampling understanding, useful to establishing an initial bank of questions for later drafting the assessment questionnaire.

As mentioned earlier, the tools of the ontosemiotic approach are also used to perform an analysis of the primary objects present in a series of secondary school Chilean textbooks, to be able to compare whether there is any relationship or difference from those proposed in the curricular guidelines.

Subsequently, the process of construction and validation of the sampling understanding assessment questionnaire is described, taking into account the information collected from the background analysis as well as the analysis of the school textbooks.

Finally, the final results obtained by applying the questionnaire to a sample of 1241 Chilean secondary school students in grades 8th grade (13-14 years), 10th (15-16 years) and 12th (17-18 years) are presented.

ÍNDICE

	Pág.
RESUMEN	
ABSTRACT	
INTRODUCCIÓN	1
CAPITULO 1: EL PROBLEMA DE INVESTIGACIÓN	
1.1. Introducción	5
1.2. Importancia del muestreo	6
1.3. El muestreo en el currículo de educación primaria	7
1.3.1. El muestreo en el currículo chileno	8
1.3.1.1. La organización de la enseñanza en Chile	8
1.3.1.2. El muestreo en educación primaria	9
1.3.1.3. El muestreo en educación secundaria	10
1.3.2. Comparación con otros contextos	15
1.3.2.1. Orientaciones generales para el territorio español	15
1.3.2.2. Los estándares americanos, Common Core Initiative y el Proyecto GAISE	17
1.4. Marco Teórico	24
1.4.1. Actividad matemática. Objeto y significado	24
1.4.2. Tipos de Objetos Matemáticos	25
1.4.3. Significado institucional y personal	26
1.4.4. Evaluación de la Comprensión	27
1.4.5. Conflictos Semióticos	27
1.4.6. Idoneidad didáctica y sus tipos	28
1.5. Objetos matemáticos considerados en el trabajo	29
1.6. Objetivos del trabajo	32
1.7. Hipótesis iniciales	33
1.8. Organización de la investigación	34
1.9. Descripción general de la metodología	35
CAPÍTULO 2: ANTECEDENTES	
2.1. Introducción	39
2.2. Niveles de uso de los conceptos de inferencia	40
2.3. Comprensión de las propiedades de muestreo	41
2.3.1. Heurísticas y sesgos	42
2.3.2. Niveles de desarrollo en el razonamiento en el muestreo	46
2.3.3. Comprensión de la variabilidad en el muestreo	52
2.4. Comprensión de la distribución muestral	54
2.5. Comprensión del enfoque frecuencias de la probabilidad	56
2.5.1. Estimación de una probabilidad a partir de la frecuencia	56
2.5.2. Características de las secuencias aleatorias de resultados	60
2.6. Muestreo en el razonamiento inferencial informal de los estudiantes	61
2.7. Conclusiones del estudio de los antecedentes	62

CAPITULO 3: MUESTREO EN LOS LIBROS DE TEXTO DE EDUCACIÓN SECUNDARIA	
3.1	Introducción 65
3.2	Objetivos del estudio de los libros de texto 66
3.3	Metodología de análisis 67
3.3.1	Muestra de los textos analizados 68
3.3.2	Método y variables de análisis 69
3.4	Situaciones problema 71
3.5	Lenguaje matemático 80
3.5.1	Lenguaje verbal 82
3.5.2	Lenguaje numérico 83
3.5.3	Lenguaje simbólico 84
3.5.4	Lenguaje tabular 86
3.5.5	Lenguaje gráfico 89
3.6	Conceptos 98
3.7	Propiedades 109
3.8	Procedimientos 114
3.9	Argumentos 121
3.10	Conclusiones sobre el estudio de libros de texto 125
CAPÍTULO 4: CONSTRUCCIÓN DEL CUESTIONARIO DE EVALUACIÓN	
4.1.	Introducción 129
4.2.	Objetivos y clasificación del instrumento 130
4.3.	Fundamentos de la definición de la variable objeto de medición 132
4.4.	Elaboración de un banco inicial de ítems 133
4.4.1.	Proceso de selección y depuración 134
4.5.	Selección de ítems mediante juicio de expertos 134
4.5.1.	Ítems que componen el cuestionario piloto y análisis a priori 135
4.6.	Prueba piloto del cuestionario 157
4.6.1.	Descripción de la muestra de estudiantes y contexto 158
4.6.2.	Resultados del estudio piloto 158
4.7.	Aproximación al estudio de validez de contenido 171
4.8.	Conclusiones sobre el cuestionario 175
CAPÍTULO 5: ESTUDIO DE EVALUACIÓN	
5.1.	Introducción 179
5.2.	Objetivos 180
5.3.	Descripción de la muestra y contexto 181
5.4.	Metodología del análisis de los datos 183
5.5.	Resultados por ítem 184
5.6.	Diferencias entre niveles educativos 230
5.7.	Diferencias entre sexos 234
5.8.	Diferencias entre tipos de centro 234
5.9.	Conflictos semióticos detectados en el estudio de evaluación 237
5.10.	Conclusiones sobre los resultados del estudio de evaluación 240

CAPÍTULO 6. CONCLUSIONES	
6.1. Introducción	245
6.2. Conclusión respecto a los objetivos generales	245
6.3. Conclusión respecto a las hipótesis	248
6.4. Principales aportes	250
6.5. Limitaciones del estudio	250
6.6. Futuras líneas de investigación	251
REFERENCIAS	253

INTRODUCCIÓN

El interés progresivo por las nociones básicas de inferencia como parte de la alfabetización estadística, ha destacado en la actualidad en que estas nociones se incorporen paulatinamente en las orientaciones curriculares de educación secundaria en diversos países, por ejemplo, España, Estados Unidos y Chile (CCSI, 2010; MEC, 2015; MINEDUC, 2009; 2012; 2015a).

En esta tesis nos interesamos por el concepto de muestra, que introduce a la inferencia, básica en la investigación y la ciencia y establece un puente entre la estadística y probabilidad (Burrill y Biehler, 2011). Estos autores, señalan que es fundamental que los estudiantes logren su comprensión para no proyectar los errores relativos a la comprensión del concepto de muestreo en otros posteriores. Además, todo nuestro conocimiento y juicios sobre el mundo o las personas están basados en el muestreo, pues sólo podemos estudiar una parcela de la realidad en la que estamos interesados.

Sin embargo, diversos investigadores en educación estadística han propuesto un rol mucho más amplio y profundo para su tratamiento en la matemática escolar (Batanero, 2013; Makar y Ben-Zvi, 2011), destacando que los fundamentos del razonamiento estadístico. Algunos de estos autores indican que las principales ideas de la estadística inferencial, deberían establecerse desde los primeros años de educación secundaria (Batanero, 2013; Ben-Zvi, Bakker y Makar, 2015; Meletiou-Mavrotheris y Paparistodemou, 2015).

Tomar muestras representativas de datos y usar muestras para hacer inferencias sobre poblaciones desconocidas son el núcleo de las estadísticas. La comprensión de cómo varían las muestras (variabilidad del muestreo) es crucial para hacer estimaciones y decisiones razonadas basadas en datos (Ben-Zvi, Bakker y Makar, 2015).

Como señalan Watson y Moritz (2000), a pesar de la importancia del concepto de muestreo en el aprendizaje de la estadística, y su paulatina inclusión en los diferentes lineamientos curriculares, existe escasa investigación sobre el desarrollo de las cogniciones del muestreo por parte de los estudiantes en el campo de didáctica de la estadística, por tal motivo, esta investigación se dirige a completar los estudios anteriores, para proporcionar información en el contexto chileno, donde no hay información sobre el razonamiento de los estudiantes en este tema.

En el Capítulo 1, se ha elaborado una comparación de los lineamientos curriculares chilenos con otras directrices a nivel internacional (España y Estados Unidos), además, se caracterizan los elementos del marco teórico que se aplican en esta investigación y se presenta un análisis semiótico de los objetos primarios relativos al muestreo presente tanto en las directrices curriculares chilenas, como también, en las internacionales. Por otro lado, se plantea el problema de investigación el que concierne a la evaluación de la comprensión del muestreo por parte de los alumnos chilenos de Educación Secundaria junto con las características generales de la organización y la metodología a seguir.

En el Capítulo 2, se entrega una síntesis de los antecedentes que engloban la comprensión del muestreo, constituyendo así una base inicial de fundamentos clave para poder elaborar el banco de preguntas preliminar para posteriormente construir el cuestionario de evaluación.

Como se indicó previamente, se aplican en el Capítulo 3 las herramientas del enfoque ontosemiótico para elaborar un análisis de los objetos primario presentes en una serie de textos escolares chilenos de educación secundaria, con la finalidad de contrastar si existe alguna diferencia con los propuestos en los lineamientos curriculares. Se presentan las conclusiones parciales del análisis realizado.

A continuación, en el Capítulo 4 de esta memoria, se describe el proceso de construcción y validación del cuestionario de evaluación de la comprensión del muestreo. Para la elaboración del instrumento se ha tenido en cuenta tanto la información recopilada la literatura existente como también de los textos escolares analizados. Se establece la respuesta a priori de los ítems del cuestionario, como también, se presentan las características psicométricas del instrumento y las principales conclusiones de validación del cuestionario con una muestra de 148 estudiantes de 4º año de educación secundaria (17-18 años).

En el Capítulo 5, se presentan los resultados obtenidos al aplicar el cuestionario a una muestra de 1241 estudiantes de secundaria chilenos, de 8º año de primaria (13-14 años); 2º año de secundaria (15-16 años) y de 4º año de secundaria (17-18 años). Se presentan los resultados por ítem, como también una comparación de dichos resultados por nivel educativo. Además, se describen los principales conflictos semióticos detectados y las principales conclusiones parciales del estudio.

Finalmente, en el Capítulo 6 se detallan las conclusiones generales por objetivos e hipótesis. Además, se describen los principales aportes y limitaciones del estudio, como también, se expresan las líneas abiertas para realizar futuras investigaciones. Por otra parte, se incluye el listado de referencias utilizadas en la elaboración de esta tesis doctoral.

CAPITULO 1:

EL PROBLEMA DE INVESTIGACIÓN

- 1.1. Introducción
- 1.2. Importancia del muestreo
- 1.3. El muestreo en el currículo de educación secundaria
 - 1.3.1. El muestreo en el currículo chileno
 - 1.3.1.1. La organización de la enseñanza de la Matemática en Chile
 - 1.3.1.2. El muestreo en la educación primaria
 - 1.3.1.3. El muestreo en la educación secundaria
 - 1.3.2. Comparación con otros contextos
 - 1.3.2.1. Orientaciones generales para el territorio español
 - 1.3.2.2. Los Estándares Americanos, Common Core State Standard Initiative y el Proyecto GAISE
- 1.4. Marco teórico
 - 1.4.1. Actividad matemática. Objeto y significado
 - 1.4.2. Tipos de objetos matemáticos
 - 1.4.3. Significado institucional y personal
 - 1.4.4. Evaluación de la comprensión
 - 1.4.5. Conflictos Semióticos
 - 1.4.6. Idoneidad didáctica y sus tipos
- 1.5. Objetos matemáticos considerados en el trabajo
- 1.6. Objetivos del trabajo
- 1.7. Hipótesis iniciales
- 1.8. Organización de la investigación
- 1.9. Descripción general de la metodología

1.1. INTRODUCCIÓN

Este capítulo plantea el problema de investigación, el que corresponde a la evaluación de la comprensión del muestreo por parte de los estudiantes chilenos de Educación Secundaria. Se comienza el capítulo justificando la importancia del muestreo para la formación del estudiante y en la propia estadística. Seguidamente se presenta un análisis del modo en que el tema se considera en el currículo de Educación Secundaria en Chile y, como pauta de

comparación, se añade el estudio del mismo contenido en los currículos de España y los Estados Unidos.

Se continúa el capítulo con la explicación resumida del marco teórico que nos sirve de apoyo, que es el enfoque ontosemiótico de la cognición e instrucción matemáticas (Godino, 2002; 2012; 2017; Godino, Batanero y Font, 2007), destacando las principales herramientas de análisis que proporciona dicho marco teórico para el análisis de las producciones de los estudiantes, entre otros elementos que permiten caracterizar el análisis curricular previamente mencionado, para, posteriormente plantear los objetivos e hipótesis que pretende abarcar esta investigación. Se finaliza el capítulo con un resumen de las características generales de la organización y de la metodología utilizada en esta memoria.

1.2. IMPORTANCIA DEL MUESTREO

La teoría del muestreo analiza la forma de recoger las muestras para asegurar que las conclusiones que, a partir de ellas, se generalizan para la población de donde se han recogido, son fiables. Es por ello esencial en el estudio de la inferencia estadística, pues esta no puede aplicarse si las muestras no han sido elegidas en forma correcta (Harradine, Batanero y Rossman, 2011).

El muestreo fue incluido en el listado de ideas fundamentales de Heitele (1975) por ser la base de la inferencia, por su papel dentro del desarrollo de la estadística y por constituir un vínculo entre estadística y probabilidad, ya que se usa en las dos materias. Además, el avance del conocimiento científico sobre el mundo, según este autor, se basa en muestras, porque no podemos explorar la realidad completa del tema de interés; por ello las diferentes ramas de las ciencias, así como en gestión y economía se deben tomar muestras para realizar estudios en dichas áreas.

Los estudiantes se encuentran con muestras cuando leen la prensa o navegan por Internet o la televisión, pues con frecuencia se dan noticias de elecciones, estudios científicos o temas similares, cuyas conclusiones se han obtenido a partir de estudios muestrales. Es por ello importante que adquieran una capacidad mínima para comprender el muestreo y las condiciones adecuadas en que debe realizarse, pues dicho conocimiento forma parte de la *cultura estadística* (Batanero, 2004). Destacamos además la importancia que se le atribuye a la *alfabetización estadística* otros autores como Gal (2002), quien señala que “la *alfabetización estadística* se describe como la capacidad de interpretar, evaluar críticamente y comunicarse

sobre información y mensajes estadísticos” (p. 1) o Watson (2006) quien menciona que “el término *alfabetización estadística* se utiliza para enfatizar que el propósito del currículo escolar no debe ser generar estadísticas, sino preparar graduados escolares alfabetizados estadísticamente que estén preparados para participar en la toma de decisiones sociales” (p. xx).

En el tema del muestreo los estudiantes pueden aplicar sus conocimientos anteriores sobre probabilidad, tales, como, por ejemplo, las ideas de probabilidad condicional e independencia (Batanero y Borovcnick, 2016). Además, la comprensión adecuada del muestreo es fundamental para los estudiantes que vayan a continuar más adelante estudios universitarios, ya que en la mayoría de los mismos se requieren conocimientos de temas de inferencia estadística, tales como el intervalo de confianza o el contraste de hipótesis. Dichos temas demandan la comprensión de las ideas de muestra y distribución muestral, por lo que un fallo en estos conocimientos tendrá repercusión en los estudios posteriores. Todos estos puntos justifican el interés de investigar en la comprensión de ideas elementales del muestreo por parte de los estudiantes chilenos.

1.3. EL MUESTREO EN EL CURRÍCULO DE EDUCACIÓN SECUNDARIA

Watson (2004) señala, que, aunque el muestreo es el fundamento de la inferencia estadística, rara vez se logra profundizar en este tema en los documentos del currículo escolar. Esto se debe principalmente a que se presenta en forma descriptiva y menos avanzada que la mayoría de los otros contenidos presentes en el currículo de matemáticas, lo que hace pensar que es difícil para los estudiantes comprender este concepto plenamente (NCTM, 2000, p. 54). Se cree que esta es una hipótesis de la autora, todavía por analizar y que en países como Chile el muestreo puede tener mayor relevancia que la descrita. Es por ello que será importante al comenzar nuestro trabajo, realizar un estudio de las directrices curriculares.

A continuación, se analizan las ideas relacionadas con el muestreo en los lineamientos curriculares de educación secundaria, con objeto de comprobar la conjetura de Watson. Primero se estudia el currículo chileno, puesto que el estudio de evaluación, que es el principal componente de la tesis doctoral, se lleva a cabo en este país. Como la tesis se realiza dentro de un programa de doctorado español y dentro de un grupo de investigación del mismo país, se sigue con el análisis del currículo español. Además, se efectúa una comparación con los estándares americanos, que tienen un gran prestigio y han servido de guía en el desarrollo de

la educación estadística en los países latinoamericanos. La finalidad es establecer un significado de referencia, que faculte posteriormente definir en forma adecuada el contenido del cuestionario de evaluación que se utilizará con estudiantes chilenos. Algunos resultados de este análisis se han publicado en Begué, Ruiz-Reyes, Gea y Batanero (2017), Ruiz-Reyes, Begué, Batanero y Contreras (2017), Ruiz-Reyes, Ruz, Molina-Portillo y Contreras (2018).

1.3.1. EL MUESTREO EN EL CURRÍCULO CHILENO

Puesto que la tesis se presenta en España, se comienza contextualizando brevemente la organización enseñanza en Chile, para informar al lector sobre las edades de los estudiantes en los diferentes cursos y la correspondencia de estos cursos con los contemplados en otros currículos.

1.3.1.1. LA ORGANIZACIÓN DE LA ENSEÑANZA DE LA MATEMÁTICA EN CHILE

La Ley General de Educación (2009) fija que la enseñanza básica (primaria) es de carácter obligatorio y compone de ocho cursos, comenzando los niños dicha etapa educativa a la edad de seis años, a partir del año 2027, se pretende una reducción en la duración de esta etapa educativa, pasando de 8 a 6 cursos, como en el caso de España. La asignatura de matemática se organiza en cinco ejes temáticos: Números y Operaciones; Patrones y Álgebra, Geometría, Medición; Datos y Probabilidades (MINEDUC, 2012), los cuales se estructuran por objetivos de aprendizaje y criterios de evaluación, los cuales se detallarán más específicamente en el siguiente apartado de este capítulo

La enseñanza media (secundaria) en Chile se compone de cuatro cursos, de primero a cuarto medio y tiene carácter obligatorio. Primero y segundo medio corresponden, por las edades de los estudiantes de 3° y 4° de ESO y tercero y cuarto medio corresponden a 1° y 2° de Bachillerato en España. Los conocimientos se organizan en cuatro ejes temáticos: Números, Álgebra y funciones, Geometría y Probabilidad y Estadística.

Para no confundir al lector, dividiremos tanto la enseñanza primaria y secundaria en 6 cursos, como en el caso de España, es decir, la enseñanza primaria de los cursos de 1° a 6° básico y la enseñanza secundaria quedará establecida en esta memoria por los cursos 7° y 8° Básico y 1°, 2°, 3° y 4° medio.

1.3.1.2. EL MUESTREO EN LA EDUCACIÓN PRIMARIA

En Chile se introducen desde la enseñanza primaria algunas ideas sobre muestreo. Es en el eje *Datos y Probabilidades*, de la asignatura de matemática, donde se presentan algunas ideas informales sobre muestreo, su objetivo es “que todos los estudiantes registren, clasifiquen y lean información dispuesta en tablas y gráficos, y que se inicien en temas relacionados con el cálculo de las probabilidades de sucesos sencillos” (p. 219, MINEDUC, 2018). Se sugiere que “es necesario que los niños conozcan y apliquen algunas encuestas y cuestionarios sencillos, por medio de la formulación de preguntas relevantes para ellos, que estén basadas en sus experiencias e intereses” (p. 219, MINEDUC, 2018). Posteriormente se debieran recoger datos, y después registrarlos y hacer predicciones a partir de ellos (MINEDUC, 2012; 2018). Más concretamente, en los Programas de Estudio se proponen los objetivos de aprendizaje para el Muestreo los cuales se describen en la Tabla 1.1.

Tabla 1.1. Objetivos de aprendizaje del muestreo en el currículo chileno de educación primaria

Curso	Objetivos de Aprendizaje
Cuarto Básico (9 – 10 años)	- Realizar encuestas, analizar los datos y comparar con los resultados de muestras aleatorias, usando tablas y gráficos. (MINEDUC, 2018; p. 244)
Quinto Básico (10-11 años)	- Utilizar diagramas de tallo y hojas para representar datos provenientes de muestras aleatorias. (MINEDUC, 2018; p. 249)
Sexto Básico (11 – 12 años)	- Comparar distribuciones de dos grupos, provenientes de muestras aleatorias, usando diagramas de puntos y de tallo y hojas. (MINEDUC, 2018; p. 254)

Aparecen las primeras ideas de muestreo desde 4° básico, donde se sugiere la realización de encuestas y comparar datos de muestras aleatorias empleando conceptos básicos de estadística descriptiva (tablas y gráficos). Estas actividades de representación de datos de muestras continúan en los cursos de 5° de primaria y 6° de primaria, especialmente con la construcción del diagrama de tallo y hojas, donde se hace énfasis en el uso de muestras aleatorias para comparar distribuciones de dos grupos.

Tabla 1.2. Criterios de evaluación del muestreo en el currículo chileno en educación primaria

Curso	Criterios de Evaluación
Cuarto Básico	- Realizan encuestas de su interés; por ejemplo: actividades en su tiempo libre, preferencias de tipo de música, club de fútbol, etc. - Comparan los resultados de sus encuestas con otros cursos del colegio, con resultados publicados en diarios y revistas, etc. (MINEDUC, 2013; p. 129)
Quinto Básico	- Obtienen muestras aleatorias y las representan en diagramas de tallo y hojas. - Completan diagramas de tallo y hojas en que están representados datos correspondientes a muestras aleatorias. (MINEDUC, 2013; p. 143)
Sexto Básico	- Usan diagramas de puntos para responder preguntas. - Construyen diagramas de puntos para obtener y comparar distribuciones. - Construyen diagramas de tallo y hojas para obtener y comparar distribuciones. (MINEDUC, 2013; p. 132)

Como se observa, el contenido es mucho más detallado en los criterios de evaluación listados en la Tabla 1.2, presentados en los distintos Programas de Estudio por cada curso, que en los objetivos de aprendizaje propuestos en las Bases Curriculares. Se especifican de forma informal propiedades de las muestras, y se introducen las técnicas de cálculo y se enfatiza la comparación e interpretación. También se incluyen actividades informales de inferencia. Así, en 4° Básico se sugiere realizar encuestas y comparar los resultados de dichas encuestas realizadas con muestras aleatorias.

1.3.1.3. EL MUESTREO EN LA EDUCACIÓN SECUNDARIA

Es en secundaria, donde el estudio del muestreo cobra más importancia en Chile, puesto que en varios objetivos se pide relacionar los estadísticos de las muestras con los parámetros de la población. De acuerdo a las Bases Curriculares correspondientes a esta etapa educativa, el eje de *Probabilidad y Estadística* (MINEDUC, 2015), incluye el muestreo, tiene por objetivo que “los estudiantes realicen diferentes tipos de análisis, establezcan inferencias y obtengan información a partir de datos estadísticos recogidos por ellos mismos u obtenido de diferentes fuentes” (p.100). Se desea preparar estudiantes críticos, que logren emplear la información estadística en sus argumentos para avalar sus decisiones y opiniones. Por ejemplo, se desea que los estudiantes puedan detectar las interpretaciones erróneas en gráficos publicados en la prensa y las eventuales manipulaciones intencionadas que se realizan con la información estadística presente en la prensa o en Internet. En probabilidad, se aspira a que los estudiantes calculen de manera intuitiva la probabilidad en sucesos sencillos y que establezcan la probabilidad de ocurrencia de eventos en forma teórica y empírica. Por tanto, se asume el estudio de las nociones clásica y frecuencial de la probabilidad. Asimismo, se pretende que los estudiantes confeccionen modelos probabilísticos basados en situaciones aleatorias.

Además, pretende que los alumnos planteen “experimentos de muestreo aleatorio para inferir sobre características de poblaciones” (p.100), tomen muestras y registren datos recogidos por ellos mismos; utilicen medidas de tendencia central, de posición y de dispersión para resolver problemas.

En síntesis, las Bases Curriculares (2015) señalan que:

El enfoque de este eje radica en la interpretación y visualización de datos estadísticos, en las medidas que permitan comparar características de poblaciones y en la realización de simulación y el estudio de experimentos aleatorios sencillos, para construir desde ellos la teoría y modelos probabilísticos. Al final de este ciclo el estudiante debe comprender el rol de la probabilidad en la sociedad, utilizando herramientas de la

estadística y de la probabilidad. (p. 100).

En la Tabla 1.3 se detallan los objetivos de aprendizaje del muestreo referentes a la educación secundaria, que se recogen solo en cinco de los seis cursos que forman esta etapa educativa. En estos objetivos hemos considerado también las técnicas combinatorias, porque se definen usando la idea de muestra con y sin reemplazo y por la estrecha relación entre el muestreo y los conceptos combinatorios resaltados por Heitele (1975). Vemos que los objetivos incluyen, además de la idea de población y muestra, la de estimación y la relación entre los estadísticos (como la media muestral) y los parámetros (como la media poblacional), aunque usando un lenguaje menos formal, incluyendo la aplicación de la tecnología para la verificación de conjeturas sobre el muestreo. También se considera el cálculo de algunos estadísticos de las muestras, conocidos los parámetros de la población o al contrario y de manera informal se hace referencia a la ley de los grandes números (aplicada a la media).

Tabla 1.3. Objetivos de aprendizaje del muestreo en el currículo chileno en educación secundaria

Curso	Objetivos de Aprendizaje
Séptimo Básico (12 – 13 años)	<ul style="list-style-type: none"> - Estimar el porcentaje de algunas características de una población desconocida por medio del muestreo. (MINEDUC, 2013; p. 13) - Representar datos obtenidos en una muestra mediante tablas de frecuencias absolutas y relativas, utilizando gráficos apropiados, de manera manual y/o con software educativo. (MINEDUC, 2013; p. 13) - Mostrar que comprenden las medidas de tendencia central y el rango: <ul style="list-style-type: none">) Determinando las medidas de tendencia central para realizar inferencias sobre la población. (MINEDUC, 2013; p. 13)) Utilizándolos para comparar dos poblaciones. (MINEDUC, 2013; p. 13)
Octavo Básico (13 – 14 años)	<ul style="list-style-type: none"> - Mostrar que comprenden las medidas de posición, percentiles y cuartiles: <ul style="list-style-type: none">) Utilizándolas para comparar poblaciones. (MINEDUC, 2013; p. 16) - Explicar el principio combinatorio multiplicativo. (MINEDUC, 2013; p. 16)
Primero Medio (14 – 15 años)	<ul style="list-style-type: none"> - Obtener la cardinalidad de espacios muestrales y eventos, en experimentos aleatorios finitos, usando más de una estrategia. - Calcular la media aritmética de las medias de muestras de igual tamaño, extraídas desde una población. (MINEDUC, 2011; p. 70 – 71) - Formular y verificar conjeturas, en casos particulares, acerca de la relación que existe entre la media aritmética de una población de tamaño finito y la media aritmética de las medias de muestras de igual tamaño extraídas de dicha población, con y sin reemplazo. (MINEDUC, 2011; p. 70 – 71)
Segundo Medio (15 – 16 años)	<ul style="list-style-type: none"> - Emplear elementos del muestreo aleatorio simple para inferir sobre la media de una población. (MINEDUC, 2011; p. 80 - 81) - Calcular medias muestrales. (MINEDUC, 2011; p. 80 - 81) - Verificar que, a medida que el número de pruebas crece, la media muestral se aproxima a la media de la población. (MINEDUC, 2011; p. 80 - 81)
Cuarto Medio (17 – 18 años)	<ul style="list-style-type: none"> - Aplicar distribuciones normales para resolver problemas de la vida diaria. (MINEDUC, 2015; p. 116 - 117) - Estimar la media poblacional de una distribución normal sobre la base de niveles de confianza dados. (MINEDUC, 2015; p. 116 - 117) - Verificar mediante ejemplos concretos que la media \bar{X} de muestras aleatorias del tamaño n, extraídas de una población, se distribuye aproximadamente normal, si se aumenta el tamaño de la muestra. (MINEDUC, 2015; p. 116 - 117)

En 7° básico se sugieren actividades de inferencia sobre las características de una población, a partir de los resultados de muestras aleatorias. Igualmente se pide diferenciar muestras aleatorias y no aleatorias para sugerir la importancia de la aleatorización en el muestreo, y la representatividad de las muestras se trabaja de manera informal. También se incluye el cálculo de las medidas de tendencia central para deducir características de la población o como medidas de comparación entre dos poblaciones.

En 8° Básico, se hace hincapié en el cálculo de las medidas de posición para comparar dos poblaciones, también se emplea el diagrama de caja como herramienta útil para realizar dicha comparación.

En 1° medio se introducen las técnicas combinatorias que son definidas mediante la metáfora del muestreo y permitirán calcular el número de muestras posibles a partir de una población finita. También se pide obtener diferentes muestras y comparar su media con la de la población para apreciar la variabilidad en el muestreo.

En 2° medio, los estudiantes obtienen muestras aleatorias y comienzan de manera informal a trabajar la distribución muestral de la media. También deben apreciar el efecto del tamaño de la muestra sobre la variabilidad de la media muestra.

En 4° medio, se fomenta la elaboración de “conjeturas sobre el tipo de distribución al que tienden las medias muestrales, con lo que se introduce la noción de distribución muestral por medio de herramientas tecnológicas. Además, se incluye el uso de criterios acerca de la representatividad de la muestra en el análisis crítico de las inferencias realizadas a partir de estudios estadísticos o simulaciones” (Ruiz-Reyes, Ruz, Molina-Portillo y Contreras, 2018, p. 509)

Como vemos en la Tabla 1.4 los criterios de evaluación, al igual que en la enseñanza primaria, son más específicos en cuanto a los contenidos y procedimientos que los estudiantes deben alcanzar, proporcionando así una visión más completa y amplia de lo que se debe enseñar a los estudiantes en cada curso.

Tabla 1.4. Criterios de evaluación sobre muestreo en el currículo chileno

Curso	Criterios de Evaluación
Séptimo Básico	<ul style="list-style-type: none"> - Infieren sobre la composición de una población incógnita pequeña mediante un muestreo aleatorio, por medio de temas que interesen, como encuestas entre los estudiantes del 7° nivel. - Infieren porcentajes representativos de la muestra y luego comparan con la realidad. - Identifican las muestras aleatorias y no aleatorias a base de ejemplos dados. - Elaboran modelos para el muestreo aleatorio en la población del 7° nivel. - Analizan las muestras obtenidas para ver coincidencias o diferencias. - Confeccionan tablas de frecuencias absolutas y relativas de los datos obtenidos en las muestras. (MINEDUC, 2016a; p.164)
Octavo Básico	<ul style="list-style-type: none"> - Reconocen cuándo es adecuado utilizar alguna de las medidas de posición para analizar una muestra. - Comparan muestras de poblaciones, utilizando algunas de las medidas de tendencia. (MINEDUC, 2016b; p.169)
Primero Medio	<ul style="list-style-type: none"> - Determinan la cardinalidad de un espacio muestral utilizando el principio multiplicativo en diversos experimentos aleatorios. - Obtienen el número de muestras aleatorias posibles de un tamaño dado que se pueden extraer, sin reposición, desde una población de tamaño finito, aplicando el número combinatorio. - Seleccionan la técnica combinatoria apropiada para resolver problemas que involucren el cálculo de probabilidades, acorde a los requerimientos de cada problema. - Establecen estrategias para determinar el número de muestras de un tamaño dado, con o sin reemplazo, que se pueden extraer desde una población de tamaño finita. - Calculan el promedio de cada una de las muestras de igual tamaño extraídas desde una población. - Calculan el promedio de todos los promedios de muestras de igual tamaño extraídas desde una población. - Realizan diferentes comparaciones entre la media de una población con la media de cada uno de los promedios de muestras de igual tamaño extraídas desde una población. - Conjeturan acerca de la relación que existe entre la media de una población y el promedio de cada uno de los promedios de muestras de igual tamaño extraídas desde una población. - Verifican, utilizando herramientas tecnológicas, la conjetura formulada. (MINEDUC, 2011a; p. 70 - 72)
Segundo Medio	<ul style="list-style-type: none"> - Producen muestras aleatorias de una población, utilizando diferentes métodos. - Emplean medios computacionales para hacer inferencias de una población. - Calculan la media muestral de pruebas independientes de experimentos probabilísticos. - Realizan experimentos con medias muestrales y establecen resultados - Extraen muestras de una población y calculan sus medias. - Analizan las medias obtenidas de las muestras, cuando el número de datos de las muestras aumenta. (MINEDUC, 2011b; p. 80 - 81)
Cuarto Medio	<ul style="list-style-type: none"> - Determinan, para un dato de una muestra aleatoria de población normal, la ubicación de este en los intervalos $[\mu - s, \mu + s]$, $[\mu - 2s, \mu + 2s]$, $[\mu - 3s, \mu + 3s]$. - Sacan una muestra del tamaño n, determinan la media X y construyen un intervalo de confianza con un nivel de confianza dado utilizando la desviación estándar - Determinan para muestras aleatorias no binomiales ni normales: <ul style="list-style-type: none"> - la media poblacional μ, - la desviación estándar poblacional σ, - la media de la distribución de medias aritméticas $\bar{\mu}$, - la desviación estándar de la distribución de medias aritméticas $\bar{\sigma}$. (MINEDUC, 2015b, p. 116 - 117)

Durante el proceso de elaboración de esta tesis doctoral, se han implementado a partir del año 2016 nuevos cambios curriculares en educación secundaria, en el eje de Probabilidad y Estadística para los cursos de primero y segundo medio (MINEDUC, 2015a), por lo que se considera pertinente incluirlos también en esta memoria (ver Tabla 1.5).

Tabla 1.5 Objetivos de aprendizaje del muestreo en el currículo chileno en educación secundaria

Curso	Objetivos de Aprendizaje
Primero Medio	<ul style="list-style-type: none"> - Registrar distribuciones de dos características distintas, de una misma población, en una tabla de doble entrada y en una nube de puntos. (MINEDUC, 2015c; p. 120) - Comparar poblaciones mediante la confección de gráficos “xy” para dos atributos de muestras, de manera concreta y pictórica: <ul style="list-style-type: none">) Utilizando nubes de puntos en dos colores.) Separando la nube por medio de una recta trazada de manera intuitiva. (MINEDUC, 2015c; p. 120)
Segundo Medio	<ul style="list-style-type: none"> - Utilizar permutaciones y la combinatoria sencilla para calcular probabilidades de eventos y resolver problemas. (MINEDUC, 2015d; p. 125)

Como se ve en la Tabla 1.5, en el caso de primero medio, se introducen nuevos conceptos, relacionados con registrar información de distribuciones de una misma población en tablas de doble entrada y a su vez comparar datos a través de las nubes de puntos, contenidos que no eran considerado en los lineamientos anteriores. En el caso de tercero y cuarto medio, no se registran cambios, pues aún están vigentes las directrices curriculares detalladas previamente en Tabla 1.4.

Al comparar estos objetivos de aprendizaje con los anteriores al cambio curricular, podemos apreciar que, en el currículo del 2012, se daba prioridad en segundo año medio a la obtención de muestras y conceptos asociados, quedando estos ahora solamente ahora restringidos a los conceptos más básicos desarrollados en séptimo básico. Es decir, se eliminan los conceptos relacionados con muestreo aleatorio y sus propiedades, como también, su comparación con otros métodos de muestreo, para solo trabajar temas de combinatoria y de probabilidades.

Se detallan los nuevos criterios de evaluación en la Tabla 1.6, que describen mayormente los contenidos a abordar en primero y segundo medio con los cambios mencionados en la Tabla 1.5.

Tabla 1.6. Criterios de evaluación sobre muestreo en el currículo chileno en educación secundaria

Curso	Criterios de Evaluación
Primero Medio	<ul style="list-style-type: none"> - Registran datos de dos características provenientes de una o de dos poblaciones, en tablas de doble entrada, y representan los datos mediante nubes de puntos en dos colores. - Argumentan acerca de coherencias o diferencias entre nubes de puntos de diferentes poblaciones. - Trazan de manera intuitiva la recta que separa de mejor forma la nube de puntos en dos poblaciones. (MINEDUC, 2016c, p. 161)
Segundo Medio	<ul style="list-style-type: none"> - Realizan permutaciones de hasta cinco elementos, con material concreto o pictóricamente. - Combinan las permutaciones con el sorteo al azar, con o sin reposición. (MINEDUC, 2016d, p. 154)

En resumen, el contenido sobre muestreo en los documentos curriculares analizados presenta el tema de forma bastante completa, aunque con una aproximación informal, teniendo en cuenta las edades de los estudiantes y no se sostiene la conjetura enunciada por Watson (2004). Será también importante el análisis de los textos escolares, que se realiza en el Capítulo 3, para complementar y finalizar el estudio de la presentación de los conceptos relacionados con el muestreo a los estudiantes de esta etapa educativa.

1.3.2. COMPARACIÓN CON OTROS CONTEXTOS

El análisis que se ha presentado en las anteriores secciones de este capítulo nos informa sobre el modo que el muestreo se presenta en la Educación Secundaria en Chile, pero para completar este informe conviene comparar con la situación en otros países. Este apartado expone una síntesis de los documentos curriculares de España, que es el país donde se desarrolla la tesis doctoral. También se analizan las disposiciones presentadas por el NTCM y el Proyecto GAISE, que son directrices mundiales en educación estadística para completar nuestro informe sobre la importancia del muestreo en la Educación Secundaria.

1.3.2.1. ORIENTACIONES GENERALES PARA EL TERRITORIO ESPAÑOL

Se presentan en la Tabla 1.7 las directrices curriculares propuestas en España en el currículo actual de Educación Secundaria Obligatoria (MECD, 2015) para el Bloque de Estadística y Probabilidad.

Tabla 1.7. Contenidos relacionados con el muestreo en el currículo español

Curso	Contenido
1° y 2° ESO	Población e individuo. Muestra. Variables estadísticas. Variables cualitativas y cuantitativas. Frecuencias absolutas y relativas. (MEC, 2015, p.413).
3° curso Matemáticas orientadas a las enseñanzas académicas	Fases y tareas de un estudio estadístico. Población, muestra. Variables estadísticas: cualitativas, discretas y continuas. Métodos de selección de una muestra estadística. Representatividad de una muestra. (MEC, 2015, p.394).
3° curso Matemáticas orientadas a las enseñanzas aplicadas	Fases y tareas de un estudio estadístico. Población, muestra. Variables estadísticas: cualitativas, discretas y continuas. Métodos de selección de una muestra estadística. Representatividad de una muestra. (MEC, 2015, p.403).
4° curso Matemáticas orientadas a las enseñanzas académicas	Introducción a la combinatoria: combinaciones, variaciones y permutaciones. Identificación de las fases y tareas de un estudio estadístico (MEC, 2015, p. 398)

Como se aprecia en la Tabla 1.7, a partir de 3° curso hay dos orientaciones posibles de las matemáticas (orientadas a las enseñanzas académicas si se piensa seguir el Bachillerato y orientadas a las enseñanzas aplicadas, cuando el objetivo es continuar con formación

profesional).

En dicho currículo español, los temas que abordan el muestreo están presentes en los cuatro años de Educación Secundaria Obligatoria, pero no se considera la distribución muestral (tema de 2º de Bachillerato) debido a que, en el contexto de esta investigación, solo haremos referencias a los conceptos asociados al muestreo.

Tabla 1.8. Criterios y estándares de evaluación sobre muestreo en el currículo español

Curso	Criterios de evaluación	Estándares de aprendizaje evaluables
1º y 2º ESO	Formular preguntas adecuadas para conocer las características de interés de una población y recoger, organizar y presentar datos relevantes para responderlas, utilizando los métodos estadísticos apropiados y las herramientas adecuadas, organizando los datos en tablas y construyendo gráficas, calculando los parámetros relevantes y obteniendo conclusiones razonables a partir de los resultados obtenidos. (MEC, 2015, p.413).	Define población, muestra e individuo desde el punto de vista de la estadística, y los aplica a casos concretos. Reconoce y propone ejemplos de distintos tipos de variables estadísticas, tanto cualitativas como cuantitativas Organiza datos, obtenidos de una población, de variables cualitativas o cuantitativas en tablas, calcula sus frecuencias absolutas y relativas, y los representa gráficamente. (MEC, 2015, p.413).
3º curso Matemáticas orientadas a las enseñanzas académicas	Elaborar informaciones estadísticas para describir un conjunto de datos mediante tablas y gráficas adecuadas a la situación analizada, justificando si las conclusiones son representativas para la población estudiada. (MEC, 2015, p.394). Analizar e interpretar la información estadística que aparece en los medios de comunicación, valorando su representatividad y fiabilidad (MEC, 2015, p.394).	Distingue población y muestra justificando las diferencias en problemas contextualizados. Valora la representatividad de una muestra a través del procedimiento de selección, en casos sencillos. (MEC, 2015, p.394). Utiliza un vocabulario adecuado para describir, analizar e interpretar información estadística de los medios de comunicación. (MEC, 2015, p.394).
3º curso Matemáticas orientadas a las enseñanzas aplicadas	Elaborar informaciones estadísticas para describir un conjunto de datos mediante tablas y gráficas adecuadas a la situación analizada, justificando si las conclusiones son representativas para la población estudiada. (MEC, 2015, p.403). Analizar e interpretar la información estadística que aparece en los medios de comunicación, valorando su representatividad y fiabilidad. (MEC, 2015, p.403).	Distingue población y muestra justificando las diferencias en problemas contextualizados. Valora la representatividad de una muestra a través del procedimiento de selección, en casos sencillos. (MEC, 2015, p.403). Utiliza un vocabulario adecuado para describir, analizar e interpretar información estadística de los medios de comunicación. (MEC, 2015, p.403).
4º curso Matemáticas orientadas a las enseñanzas académicas	Elaborar e interpretar tablas y gráficos estadísticos, así como los parámetros estadísticos más usuales, en distribuciones unidimensionales y bidimensionales, utilizando los medios más adecuados (lápiz y papel, calculadora u ordenador), y valorando cualitativamente la representatividad de las muestras utilizadas. (MEC, 2015, p.398).	Aplica en problemas contextualizados los conceptos de variación, permutación y combinación. Interpreta un estudio estadístico a partir de situaciones concretas cercanas al alumno. Selecciona una muestra aleatoria y valora la representatividad de la misma en muestras muy pequeñas (MEC, 2015, p.398).

Los temas de muestreo aparecen desde primer curso, pero en el 4º curso no hay ningún contenido relacionado con las matemáticas orientadas a las enseñanzas aplicadas, mientras en las matemáticas orientadas a las enseñanzas académicas se incluyen los conceptos combinatorios. El contenido es bastante más limitado que en el caso de Chile, cumpliéndose en este caso la conjetura de Watson (2004).

En bachillerato, se encuentra que algunas nociones de muestreo se exploran en el segundo curso, en la opción de matemáticas aplicadas a las ciencias sociales II. Pero no se consideran estos cursos en nuestro estudio. Cabe mencionar que en primer curso tampoco se mencionan temas relacionados con el muestreo. Otra diferencia es que no consideran actividades de inferencia informal en el currículo español ni se menciona la ley de los grandes números. Las actividades de comparación e interpretación también son mucho menos explícitas en las directrices curriculares.

Completamos el estudio con la presentación en la Tabla 1.8 de los criterios y estándares de evaluación sobre muestreo, que también detallan el contenido. Se ve que el énfasis se pone en la diferenciación entre población y muestra y en el concepto de representatividad, así como en los métodos de selección de muestras, además, también se abordan los temas combinatorios.

Se presenta una comparación más completa entre los lineamientos curriculares chilenos y los españoles en Begué, Ruiz-Reyes, Batanero y Gea (2017) y en Ruiz-Reyes, Begué, Batanero y Contreras (2018).

1.3.2.2. LOS ESTÁNDARES AMERICANOS, COMMON CORE STATE STANDARDS INITIATIVE Y EL PROYECTO GAISE

En este apartado se consideran los lineamientos sugeridos por los Estándares Americanos y las propuestas de Common Core State Standards Initiative y el Proyecto GAISE sobre los principales objetivos de enseñanza de la probabilidad y el muestreo.

Estándares Americanos

Los Principios y Estándares para la Educación Matemática (NCTM, 2000) proponen una serie de descriptores para la educación matemática en la escuela. Dentro del estándar de Análisis de Datos y Probabilidad, se observan los lineamientos que rigen la enseñanza de la Estadística. Seguidamente, describimos las principales ideas para la enseñanza del concepto de

muestreo.

En el estándar de análisis de datos y probabilidad, se destaca la presencia del objetivo: desarrollar y evaluar inferencias y predicciones basadas en datos, en el cual “los estudiantes deberían llegar a comprender los elementos básicos del análisis estadístico: seleccionar una muestra adecuada, recoger datos de esta muestra, describir la muestra y hacer inferencias razonables que relacionen la muestra y la población”. (NCTM, 2000, p.53)

Por otro lado, también se menciona que “al final de los niveles medios y en la enseñanza secundaria, los alumnos deberían utilizar las ideas de selección de muestras e inferencia estadística, y empezar a comprender que hay maneras de cuantificar el grado de certeza de los resultados estadísticos” (NCTM, 2000, p. 54), es decir, deberían formular preguntas y diseñar experimentos o encuestas, con el fin de comparar características dentro de una población o entre poblaciones

Además, los estudiantes deberían usar simulaciones para aprender sobre distribuciones muestrales y hacer inferencias informales. En particular, deberían saber que las técnicas estadísticas básicas se utilizan para controlar la calidad en el mundo laboral. Al terminar la escuela secundaria deberían estar capacitados para juzgar la validez de los argumentos basados en datos, como los que aparecen en la prensa (NCTM, 2000, p.54).

Los estándares americanos señalan que “al recoger y representar datos, los estudiantes deberían tener por guía el deseo de responder a preguntas basándose en los datos. Durante el proceso, deberían hacer observaciones, formular inferencias y conjeturas, y proponer nuevas preguntas”. (NCTM, 2000, p.256)

Se pretende que los estudiantes puedan emplear todas las herramientas que le ofrece la estadística descriptiva, sobre todo en la elaboración de gráficos, como es el caso de las nubes de puntos “las que son útiles para detectar y examinar relaciones entre dos características de una población”. (NCTM, 2000, p.256)

Mediante experiencias apropiadas, los alumnos empezarían a entender que muchos de los conjuntos de datos son muestras de poblaciones mayores. Pueden observar diversas muestras extraídas de la misma población (diferentes clases de su escuela, por ejemplo), o comparar estadísticas sobre sus propias muestras para parámetros conocidos de la población (por ejemplo, comparar el tamaño medio de las familias de su clase con el de su ciudad). Y pueden reflexionar sobre qué aspectos afectan a la representatividad de una muestra (si representa bien o no a la población de la que se extrajo), y empezar a notar cómo pueden variar

las muestras de una misma población. (NCTM, 2000, p. 185).

La Tabla 1.9 detalla las expectativas u objetivos propuestos para los niveles en donde se trabaja explícitamente el concepto de muestreo, donde ya en los grados 6 a 8 se sugieren actividades de inferencia informal y en los 9 a 12 actividades de exploración de la variabilidad muestral y relación con el tamaño de la muestra.

Tabla 1.9. El muestreo en los Estándares Americanos

Etapa	Expectativas
3 - 5	Z Describir la forma y las características importantes de un conjunto de datos, y comparar conjuntos que tengan relación, poniendo el énfasis en cómo se distribuyen los datos. Z Proponer y justificar conclusiones y predicciones basadas en datos, y diseñar estudios para investigarlas más a fondo. (NTCM, 2000, p.180)
6 - 8	Z Formular preguntas, diseñar estudios y recoger datos sobre una característica compartida por dos poblaciones, o sobre diferentes características de una misma población. Z Utilizar observaciones relativas a las diferencias entre dos o más muestras, para formular conjeturas sobre las poblaciones de las que se han extraído. Z Formular conjeturas sobre las posibles relaciones entre dos características de una muestra, a partir de nubes de puntos de los datos y líneas de ajuste aproximadas. (NTCM, 2000, p.252)
9 - 12	Z Utilizar simulaciones para explorar la variabilidad de muestras estadísticas de una población conocida, y para construir distribuciones muestrales. Z Comprender cómo las muestras estadísticas reflejan los valores de los parámetros de la población, y utilizar las distribuciones muestrales como base para inferencias informales. (NTCM, 2000, p.328)

Common Core State Standards Initiative

Los estándares que se proponen en CCSSI (2010) delimitan lo que los alumnos deben entender y realizar en el estudio de las matemáticas a nivel escolar. En el caso de la enseñanza de la Estadística y Probabilidad, se presentan en la Tabla 1.10 los objetivos de aprendizajes propuestos y los indicadores de evaluación para el muestreo. Se destaca, además, que los temas de estadística y probabilidad se inician en 6° Grado y su énfasis en temas relacionados con el muestreo es en séptimo grado como también en la Educación Secundaria.

Tabla 1.10. Objetivos de Aprendizaje propuestos en Common Core State Standards Initiative

Curso	Objetivos de Aprendizaje
7° Grado (12 – 13 años)	– Usar el muestreo aleatorio para sacar conclusiones acerca de una población. – Sacar conclusiones comparativas informales sobre dos poblaciones. – Investigar y desarrollar procesos aleatorios, usar, y evaluar los modelos de probabilidad. (CCSSI, 2010, p.47)
Educación secundaria (14 años en adelante)	– Comprender y evaluar procesos subyacentes a experimentos estadísticos. – Hacer inferencias y justificar conclusiones de encuestas por muestreo, experimentos y estudios observacionales. (CCSSI, 2010, p.80)

En el Grado 7, los alumnos complementan sus conocimientos sobre distribuciones de una variable para realizar la comparación entre dos variables y formular preguntas sobre las diferencias entre poblaciones. Empiezan el estudio informal del muestreo aleatorio para generar conjuntos de datos, además de comprender que las estadísticas se pueden usar para obtener información sobre una población examinando una muestra de ésta y aprender acerca de la importancia de muestras representativas para hacer generalizaciones e inferencias. Por otro lado, se espera que los estudiantes generen distintas muestras (o muestras simuladas) de igual tamaño para estimar la variación.

En secundaria se estudia la variabilidad de los datos. Como señalamos en Begué, Ruiz-Reyes, Gea y Batanero (2017):

La aleatorización tiene dos usos importantes: a) la recogida de datos a partir de una muestra aleatoria de una población hace que sea posible sacar conclusiones válidas sobre toda la población; b) la asignación al azar de diferentes tratamientos a individuos permite comparar la eficacia de dichos tratamientos. La forma en que se recogen los datos es importante para extraer conclusiones de los mismos, por lo que hay que tenerla en cuenta en la revisión crítica de la estadística en los medios de comunicación. (p. 7)

En este sentido, se espera que los estudiantes comprendan que los estadísticos son una herramienta útil en el proceso de realizar inferencias de los los parámetros de la población basándose en una muestra de ésta seleccionada aleatoriamente; decidan si un modelo específico es consistente con los resultados de un proceso de generación de datos dado, por ejemplo, usando simulación; reconozcan los propósitos y diferencias entre las encuestas por muestreo, los experimentos y los estudios observacionales y explicar cómo la aleatorización se relaciona con cada uno; utilizar los datos de una encuesta por muestreo para estimar la media o proporción de una población; desarrollar un margen de error mediante el uso de modelos de simulación para muestreo aleatorio; utilizar los datos de un experimento aleatorio para comparar dos tratamientos (condiciones experimentales); usar simulaciones para decidir si las diferencias entre parámetros son significativas y finalmente, evaluar informes basados en datos.

Proyecto GAISE

El Proyecto GAISE (Franklin et al. 2007) fue elaborado siguiendo los principios y estándares del NCTM y ha sido publicado por la American Statistical Association (ASA) bajo el nombre de “*Pautas para la Evaluación e Instrucción en Educación Estadística*”.

De acuerdo a Ruiz (2013):

Uno de sus principales objetivos es la *alfabetización estadística*, cuyo propósito es instruir,

desde los primeros años de escolaridad, alumnos que sepan razonar a partir de datos empíricos; comprender y explicar la variación de los datos; percibir, cuantificar y justificar el comportamiento de los datos. Consideran tres niveles de desarrollo (A, B, C), a través de cuatro procesos: formular preguntas, recolectar datos, analizar datos e interpretar resultados. (p.16)

El muestreo aparece en los niveles B y C. Este proyecto señala que en la formación de los estudiantes deben considerarse las siguientes ideas principales: un ciudadano informado por encuestas necesita entender que los resultados fueron determinados a partir de una muestra de la población estudiada, que la fiabilidad de los resultados depende de cómo se seleccionó la muestra y que los resultados están sujetos a errores de muestreo. El ciudadano estadísticamente alfabetizado debe entender el comportamiento de las muestras “aleatorias” y ser capaz de interpretar un “margen de error de muestreo”.

En la Tabla 1.11 se observan los contenidos relacionados con el muestreo de acuerdo a los niveles de desarrollo propuestos en estos lineamientos, los cuales abarcan tres grandes grupos o componentes del proceso: (1) recolectar datos; (2) analizar datos; (3) interpretar resultados.

Por otro lado, se plantea que las principales interpretaciones que deben considerarse del concepto de *Variabilidad*, que aparece de cuatro formas, de las cuales las tres primeras nos interesan en el estudio:

-) *La variabilidad natural*: inherente a la naturaleza, pues los individuos son diferentes. De ahí que se consideren las variables estadísticas o aleatorias para describir un rasgo en una muestra o población, por ejemplo, el peso de los individuos.
-) *La variabilidad de muestreo*: Puesto que no se estudia a toda la población, se toma una muestra de tamaño dado, pero si se utiliza una segunda muestra del mismo tamaño, aunque sea parecida no será exactamente igual.
-) *La variabilidad de medición*: Las medidas siempre tienen un error aleatorio.
-) *Variabilidad inducida*: Es cuando se comparan las medidas de sujetos donde se manipula un factor; por ejemplo, un medicamento y se compara su efecto en dos grupos. Este tipo de variabilidad no se tiene en cuenta en este estudio.

Tabla 1.11. Muestreo en los niveles de desarrollo del Proyecto GAISE

Componente del proceso	Nivel B	Nivel C
II. Recolectar datos	<ul style="list-style-type: none"> - Los estudiantes diseñan y llevan a cabo encuestas de muestra no aleatorias y comienzan a utilizar la selección aleatoria. - Los estudiantes diseñan y llevan a cabo experimentos comparativos y comienzan a usar tareas aleatorias. (Franklin et al. 2007, p.14) 	<ul style="list-style-type: none"> - Los estudiantes deben entender lo que constituye una buena práctica al realizar una encuesta por muestreo. - Los estudiantes deben poder diseñar e implementar un plan de recopilación de datos para estudios estadísticos, incluidos estudios observacionales, encuestas por muestreo y experimentos comparativos simples. (Franklin et al. 2007, p.14)
III. Analizar datos	<ul style="list-style-type: none"> - Los estudiantes amplían su comprensión de una distribución de datos. - Los estudiantes cuantifican la variabilidad dentro de un grupo. - Los estudiantes comparan dos o más distribuciones usando diferentes representaciones gráficas y resúmenes numéricos. - Los estudiantes reconocen el error de muestreo. (Franklin et al. 2007, p.14) 	<ul style="list-style-type: none"> - Los estudiantes deben entender cómo se utilizan las distribuciones de muestreo (desarrolladas a través de la simulación) para describir la variabilidad de muestra a muestra de las estadísticas de la muestra. (Franklin et al. 2007, p.14)
IV. Interpretar Resultados	<ul style="list-style-type: none"> - Los estudiantes reconocen que una muestra puede no ser representativa de una población más grande. - Los estudiantes reconocen la variabilidad del muestreo en los estadísticos de resumen, como la media muestral y la proporción muestral (Franklin et al. 2007, p.15) 	<ul style="list-style-type: none"> - Los estudiantes deben ser capaces de interpretar el margen de error asociado con una estimación de una característica de la población. (Franklin et al. 2007, p.15)

Igualmente se resalta la importancia del *muestreo aleatorio*. Una muestra representativa es aquella en la que las características relevantes de los miembros de la muestra son generalmente las mismas que las de la población. La selección inadecuada o parcial de la muestra tiende a favorecer sistemáticamente ciertos resultados y puede producir conclusiones erróneas. El muestreo aleatorio es una manera de eliminar el sesgo en la selección de la muestra, y tiende a producir muestras representativas.

Además, al seleccionar una muestra, los estudiantes deben desarrollar una comprensión básica del rol que juega la probabilidad en la selección aleatoria, y en la asignación aleatoria al realizar un experimento.

Los estudiantes también necesitan entender el concepto de independencia. Dos resultados son independientes si nuestra evaluación de la probabilidad de que se produzca un resultado no

se ve afectada por el conocimiento de que el otro resultado se ha producido. La selección aleatoria (con reemplazo) de una población asegura que las observaciones en una muestra son independientes.

En el Nivel B, los estudiantes deben experimentar las consecuencias de la selección no aleatoria y desarrollar una comprensión básica de los principios involucrados en los procedimientos de selección aleatoria.

La comprensión de los conceptos estadísticos del Nivel B le permite al estudiante comenzar a apreciar que el análisis de datos es un proceso de investigación que consiste en formular sus propias preguntas, recopilar datos apropiados a través de varias fuentes (censos, encuestas de muestreo aleatorias y no aleatorias y experimentos comparativos con asignación aleatoria), analizando datos a través de gráficos y medidas de resumen simples, e interpretando resultados con miras a la inferencia a una población basada en una muestra.

Los estudiantes de nivel C deben convertirse en expertos en el uso de herramientas estadísticas como parte natural del proceso de investigación. Esto implica que los estudiantes deben poder seleccionar las técnicas de resumen apropiadas para el tipo de datos disponibles, producir estos resúmenes y describir en contexto las características importantes de los datos.

También deberían poder sacar conclusiones de los datos y apoyar estas conclusiones utilizando evidencia estadística. Los estudiantes deben ver a las medidas estadísticas como herramientas poderosas que les permiten responder preguntas y tomar decisiones informadas. Además, deben entender las limitaciones de las conclusiones basadas en datos de encuestas y experimentos de muestra, y deben poder cuantificar la incertidumbre asociada con estas conclusiones utilizando el margen de error y las propiedades relacionadas de las distribuciones de muestreo.

1.4. MARCO TEÓRICO

A continuación, se exponen algunos elementos teóricos que tendremos en cuenta en el trabajo y que se han tomado del enfoque Ontosemiótico, desarrollado en la Universidad de Granada principalmente (Godino, 2002, 2012, 2017; Godino y Batanero, 1994; 1998; Godino, Batanero, y Font, 2007). Dicho modelo teórico consta de varios componentes o dimensiones: epistémica, cognitiva, instruccional y normativa. Esta tesis doctoral se centrará en las componentes epistémica y cognitiva, con la teoría de los objetos matemáticos, sus tipos y facetas, así como los significados personales e institucionales y la faceta normativa (teoría de la idoneidad didáctica).

1.4.1. ACTIVIDAD MATEMÁTICA. OBJETO Y SIGNIFICADO

En el marco teórico citado se parte de la idea primitiva de problema y de la actividad matemática desarrollada en la resolución de problemas (por ejemplo, la actividad matemática que tiene lugar cuando surge un problema de muestreo, tal como decidir el tamaño de una muestra para estimar la media de una población). En dicha actividad matemática aparecen una serie de prácticas operativas o discursivas (acciones que tienen lugar para dar solución al problema dado) y de las cuáles surgen los objetos matemáticos y se pueden ejecutar de forma conjunta o individual.

Godino, Batanero, y Font (2007) denominan “*institución* al grupo de personas que, además de resolver problemas comunes, comparte instrumentos, reglas y modos de funcionamiento” (p.129).

1.4.2. TIPOS DE OBJETOS MATEMÁTICOS

En las prácticas matemáticas intervienen objetos matemáticos (Godino, 2002) cuya definición o propiedades el alumno debe recordar y que en el marco teórico se clasifican de la forma siguiente:

) *Lenguajes* (términos, expresiones, notaciones, tablas, gráficos) en sus diversos registros (escrito, oral, gestual) (Godino, Batanero y Font, 2007, p.130). Para este estudio, se consideran en esta categoría los términos *población, muestra, parámetro, estadístico, representatividad, sesgo*. Respecto al lenguaje simbólico, tabular y gráfico se estudiarán ejemplos en el análisis de los textos escolares.

-) *Situaciones-problemas*: son las aplicaciones matemáticas o extra-matemáticas o de las que surge el objeto (Vera, 2015, p.48). Para este estudio se puede considerar: estimar algunas características de una población desconocida por medio del muestreo, determinar el número de muestras de un tamaño dado, que se pueden extraer desde una población de tamaño finito, con y sin reemplazo, entre otras.
-) *Conceptos- definición* son introducidos mediante definiciones o descripciones (Vera, 2015, p.48). Por ejemplo: parámetro muestral, parámetro poblacional.
-) *Proposiciones* son los teoremas y propiedades que ligan los conceptos (Vera, 2015, p.48). En este caso podemos mencionar el Teorema Central del Límite, o que la muestra es un subconjunto de la población.
-) *Procedimientos* algoritmos, operaciones, técnicas de cálculo (Vera, 2015, p.48). Para este estudio se pueden hallar en obtener todas las posibles muestras de una población finita, comparar datos de una muestra con una población, decidir si un método de muestreo es adecuado o diferenciar tipos de muestreo.
-) *Argumentos* usados para justificar o explicar las proposiciones y procedimientos, o las soluciones a los problemas (Vera, 2015, p.48). Para nuestro estudio, el estudiante debe entregar sus argumentos en los problemas abiertos.

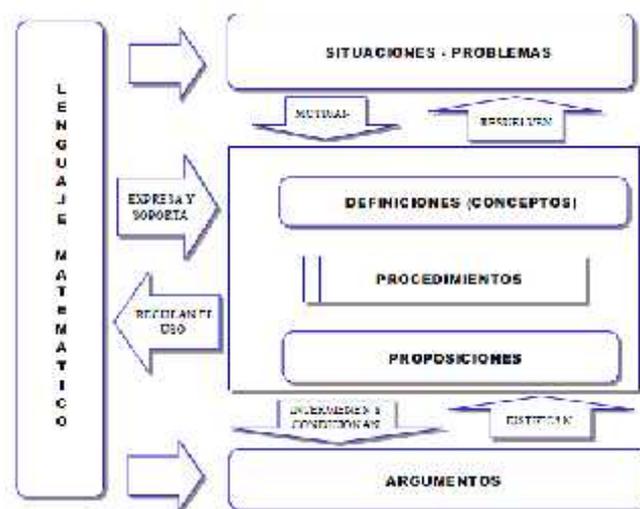


Figura 1.1. Configuración de objetos primarios (Godino *et al*, 2008, p.7)

Estos tipos de objetos primarios (ver Figura 1.1) extienden la clasificación tradicional que solo distingue entre conocimiento conceptual y procedimental en la enseñanza, a la vez que muestran la complejidad de la labor del profesor. Godino, Font y Wilhelmi (2008) señalan que estos objetos se relacionan entre sí mediante configuraciones epistémicas (en el caso de las

instituciones) o cognitivas (en el caso de las personas) (Vera, 2015, p.48).

1.4.3. SIGNIFICADO INSTITUCIONAL Y PERSONAL

Godino y Batanero (1994) señalan que “el significado (institucional o personal) de un objeto al sistema de prácticas de las cuáles surge el objeto, es decir, el significado de los objetos está íntimamente ligado a los problemas y a la actividad realizada para su resolución, no pudiéndose reducir este significado del objeto a su mera definición matemática” (p.5).

Por otro lado, Godino (2002) diferencia en varios tipos de significados institucionales y personales. Siguiendo lo señalado por Gómez (2014):

Dentro del significado institucional, se distinguen el *holístico* (sistema de prácticas en el sentido más amplio), el *referencial* (base de una investigación, que suele ser más restringido que el anterior), el *pretendido* (fijado para la planificación de un proceso de enseñanza), el *evaluado* (subsistema que utiliza el docente para valorar los aprendizajes), el *implementado* (sistema de prácticas implementado por el docente). Dentro del significado personal, identifican el significado *global* (todo el sistema potencial de prácticas subjetivas), el *declarado* (prácticas mostradas en un proceso de evaluación), el *logrado* (las mostradas que coinciden con las pretendidas institucionalmente). En un proceso de estudio, interesan también los significados iniciales y finales de los estudiantes que se observan al comparar sus sistemas de prácticas personales antes y después de la instrucción (Font, Godino, y Gallardo, 2013) (p.7-8).

Esta tipología de significados se puede representar gráficamente a través de la Figura 1.2, donde Godino, Font y Wilhelmi (2008) representan “las relaciones dialécticas entre enseñanza y aprendizaje, lo cual da como consecuencia el acoplamiento progresivo entre los significados personales e institucionales” (Godino, Batanero, Font, 2009, p 6).



Figura 1.2. Tipos de significados en el enfoque ontosemiótico (Godino, Font y Wilhelmi, 2008, p. 5)

1.4.4. EVALUACIÓN DE LA COMPRENSIÓN

Godino (1996) sugiere un modelo que considera dos ejes, para analizar la comprensión que tiene un grupo de sujetos de un cierto objeto matemático:

-) *Eje descriptivo*: señala las características o elementos que componen los objetos que se desean valorar, en este estudio, serán los contenidos de muestreo que abarca el cuestionario de evaluación.
-) *Eje procesual*: señala los niveles o etapas indispensables para valorar que un alumno logra la comprensión; en el cuestionario de evaluación del muestreo serán las respuestas correctas establecidas en el análisis a priori de cada situación problema.

Como señala Vera (2015) siguiendo lo propuesto por Godino (1996):

La comprensión personal de un determinado objeto por parte de un sujeto, no puede ser observada directamente, pero sí indirectamente desde las prácticas personales. Por tanto, en el enfoque ontosemiótico la evaluación de la comprensión concierne al estudio de la relación entre los significados personales e institucionales. Esta evaluación, por tanto, depende de la institución desde donde se lleva a cabo, ya que es la que determina si un sujeto “comprende” el significado de un objeto. (p.51)

En síntesis, de acuerdo a Godino, Batanero y Font (2007) “se considera que un sujeto comprende un determinado objeto matemático cuando lo usa de manera competente en diferentes prácticas” (p.123).

En esta investigación, guiándonos por las orientaciones presentadas en el marco teórico, se entenderá la comprensión personal de los estudiantes de secundaria chilenos que participan en el estudio de la evaluación de la comprensión del muestreo, por medio del análisis de las respuestas (entendidas como prácticas matemáticas), que ejecuta el sujeto para dar solución a los ítems planteados.

1.4.5. CONFLICTOS SEMIÓTICOS

De acuerdo a Godino, Batanero y Font (2007):

Un *conflicto semiótico* es cualquier disparidad o discordancia entre los significados atribuidos a una expresión por dos sujetos (personas o instituciones). Si la disparidad se produce entre *significados institucionales* hablamos de conflictos semióticos de tipo *epistémico*, mientras que, si la disparidad se produce entre *prácticas* que forman el significado personal de un mismo sujeto, los designamos como conflictos semióticos de tipo *cognitivo*. (p. 133)

En nuestro estudio de evaluación de la comprensión del muestreo por alumnos de secundaria chilenos, se establecerán los posibles conflictos semióticos a partir de las respuestas entregadas por los alumnos a las distintas preguntas del cuestionario de evaluación, como todos aquellos elementos que inducen la disparidad entre los significados personales y los significados institucionales, es decir, los significados que no coinciden con los considerados correctos desde el punto de vista institucional.

1.4.6. IDONEIDAD DIDÁCTICA Y SUS TIPOS

En la Figura 1.3, se presenta la caracterización de la noción de idoneidad didáctica, en donde intervienen las dimensiones que se describen a continuación:

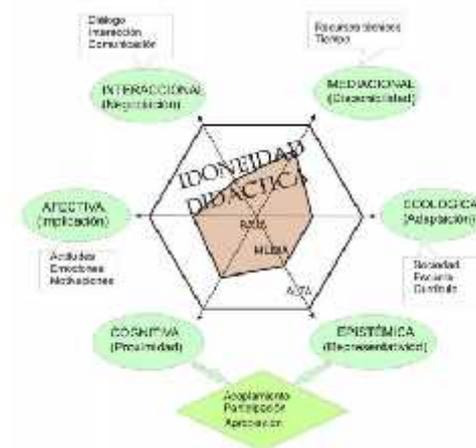


Figura 1.3. Idoneidad Didáctica (Godino, 2013, p. 116)

-) *Idoneidad epistémica*: Representatividad de los significados institucionales implementados (o pretendidos), respecto de un significado de referencia. Sería si los significados de los objetos presentes en el cuestionario son adecuados desde el punto de vista matemático.
-) *Idoneidad cognitiva*: Grado en que los significados pretendidos/implementados son asequibles a los alumnos, así como si los significados personales logrados por los alumnos son los significados pretendidos por el profesor.
-) *Idoneidad interaccional*: Grado en que la organización de la enseñanza permite identificar conflictos semióticos y resolverlos durante el proceso de instrucción. Esta componente no se analizará en nuestro estudio.
-) *Idoneidad mediacional*: Disponibilidad y adecuación de los recursos necesarios para el desarrollo del proceso de enseñanza-aprendizaje.

-) *Idoneidad emocional*: Interés y motivación del alumnado en el proceso de estudio.
-) *Idoneidad ecológica*: Grado en que el proceso de estudio se ajusta al proyecto educativo del centro, la escuela y la sociedad y a los condicionamientos del entorno en que se desarrolla. Esta componente no se analizará en nuestro estudio.

1.5. OBJETOS MATEMÁTICOS CONSIDERADOS EN EL TRABAJO

En las Tablas 1.12, 1.13, 1.14, 1.15 y 1.16 se presentan los tipos de objetos primarios (Godino, 2002) considerados en nuestro trabajo. Constituyen el significado institucional de referencia en la investigación y se ha determinado a partir del análisis curricular citado en las secciones anteriores. Resultados parciales de esta comparación se presentan en Ruiz-Reyes, Begué, Batanero y Contreras (2017), los cuales se han completado en el análisis resumido a continuación.

Tabla 1.12. Problemas considerados en los currículos

Tipo de objeto	Objetos	Currículo chileno	Currículo español	NCTM	CCSSI	GAISE	
Problemas	Estimar algunas características de una población desconocida por medio del muestreo.	x	x	x	x	x	
	Determinar el número de muestras de un tamaño dado, que se pueden extraer desde una población de tamaño finito, con y sin reemplazo	x	x	x	x	x	
	Relacionar la proporción de una población y la proporción de las muestras	x			x	x	
	Relacionar la media de una población y la media aritmética de las muestras	x			x	x	
	Identificar las muestras aleatorias y no aleatorias	x	x	x	x	x	
	Elaboran modelos para el muestreo aleatorio	x	x		x	x	
	Valorar la representatividad de una muestra	x	x	x	x	x	
	Comparar dos o más conjuntos de datos a través de las medidas de tendencia central	x		x	x	x	
	Comparar poblaciones y muestras usando medidas de posición.	x		x	x	x	
	Registrar distribuciones de dos características distintas de una misma población en una nube de puntos.	x		x	x	x	
	Registrar distribuciones de dos características distintas, de una misma población, en una tabla de doble entrada.	x			x	x	
	Comparar poblaciones mediante gráficos de dispersión para dos variables utilizando puntos con colores o separando la nube con una recta de manera intuitiva	x			x	x	x

En la Tabla 1.12 se muestran las principales situaciones problema presentes en los diferentes documentos curriculares, las cuales están más explícitas en el currículo chileno y en el proyecto GAISE como en el Common Core State Standar Initiative, en cambio, en el currículo español se exhiben descriptores más generales y no tan específicos.

Tabla 1.13. Lenguaje considerado en los currículos

Tipo de objeto	Objetos	Currículo chileno	Currículo español	NCTM	CCSSI	GAISE	
Lenguaje	Términos, expresiones, fórmulas y expresiones algebraicas	x	x	x	x	x	
	Gráficos	Diagrama de Barra	x	x	x	x	x
		Diagrama de Sectores	x	x	x	x	x
		Histogramas	x	x	x	x	x
		Diagrama de Caja	x	x	x	x	x
		Diagrama de árbol	x	x	x	x	x
	Tablas de frecuencia, listados de datos	x	x	x	x	x	

Como se puede apreciar en la Tabla 1.13, los distintos tipos de lenguaje están presente en todos los currículos, ya que desde los niveles de primaria se introducen paulatinamente las distintas herramientas de la estadística descriptiva (tablas, gráficos, estadísticos de resumen, etc.), las cuales son necesarias para posteriormente en secundaria apuntar al análisis de datos y a obtener inferencias a partir de la información proporcionada en los distintos tipos de lenguaje.

Tabla 1.14. Conceptos considerados en los currículos

Tipo de objeto	Objetos	Currículo chileno	Currículo español	NCTM	CCSSI	GAISE
Conceptos	Conceptos combinatorios: variaciones, permutaciones y combinaciones	x	x		x	
	Experimento aleatorio	x	x	x	x	x
	Espacio muestral (Finito o infinito, Sucesos equiprobables o no, Probabilidad conocida o no)	x	x	x	x	x
	Variable estadística, distribución	x	x	x	x	x
	Muestra (Con y sin reemplazamiento, Aleatoria, sistemática, estratificada, no probabilística, Ordenadas o no, Repitiendo elementos o no)	x	x	x	x	x
	Tamaño de la muestra	x	x	x	x	x
	Población (Elemento, Tamaño, finita o no)	x	x	x	x	x
	Probabilidad (Clásica, Frecuencial)	x	x	x	x	x
	Media Muestral	x	x	x	x	x
	Representatividad Muestral	x	x	x	x	x
	Sesgo en muestreo			x		x
	Variabilidad			x	x	x

La mayoría de los conceptos (ver Tabla 1.14) y propiedades (ver Tabla 1.15) abordadas por el proyecto GAISE y el CCSSI se encuentran desarrolladas tanto en el currículo chileno como en el español. Por otro lado, los conceptos de sesgo y variabilidad no se presentan de una

manera explícita en el currículo chileno, en cambio, en los otros lineamientos se presentan de una manera más detallada.

Tabla 1.15. Propiedades consideradas en los currículos

Tipo de objeto	Objetos	Currículo chileno	Currículo español	NCTM	CCSSI	GAISE
Propiedades	Principio Multiplicativo	x	x			x
	Población finita o infinita	x	x	x	x	x
	Representatividad de una muestra en función del método de extracción	x	x	x	x	x
	Variabilidad de la muestra en función del tamaño	x	x	x	x	x
	Aproximación de la frecuencia relativa a la probabilidad mediante la simulación o experimentación	x	x	x	x	x

En el caso de los conceptos combinatorios, estos se ven reflejados más explícitamente en el currículo chileno, español y CCSSI, en cambio, en los otros lineamientos internacionales no son mencionados. Los conceptos principales de muestra y algunas de sus propiedades son prácticamente abordados por todos los documentos curriculares, salvo por la propiedad del principio multiplicativo que no aparece explícita en los estándares americanos.

Algunos procedimientos (ver Tabla 1.6), tales como obtener todas las posibles muestras de una población finita y calcular el número de muestras diferentes de una población finita, son presentados de manera más explícita en el currículo chileno y español.

En el currículo español, el proyecto GAISE como en los estándares americanos (NCTM, 2000), se consideran las fases de un estudio estadístico como uno de los contenidos a enseñar, en el caso del currículo chileno, estas fases son más explícitas en el Texto escolar de 7° Básico, información que se presenta en los resultados del Capítulo 3.

Cabe destacar que todos los lineamientos curriculares consideran emplear herramientas tecnológicas para la enseñanza de los temas estadísticos, como también, para verificar resultados.

En el caso de los argumentos (ver Tabla 1.16), es preciso señalar que, en el caso del currículum español, los ejemplos y contraejemplos sólo están presentes explícitamente en los contenidos de los cursos de 1° y 2° de Bachillerato.

Tabla 1.16. Procedimientos y argumentos considerados en los currículos

Tipo de Objeto	Currículo chileno	Currículo español	NCTM	CCSSI	GAISE	
Procedimientos	Realizar encuestas	x	x	x	x	x
	Representar la distribución de frecuencias mediante tablas o gráficas	x	x	x	x	x
	Calcular las medidas de centralización (media)	x	x	x	x	x
	Construir Diagramas de árbol (espacio muestral)	x	x	x	x	
	Obtener todas las posibles muestras de una población finita	x	x		x	
	Calcular el número de muestras diferentes de una población finita	x	x		x	
	Estimar una proporción, media o frecuencia esperada en una muestra conocida la población	x	x	x	x	x
	Estimar una proporción, media o frecuencia esperada en una población, conocida la muestra	x	x	x	x	x
	Dada una muestra de resultados, decidir si el método de selección fue aleatorio	x	x	x	x	x
	Decidir si un método de muestreo es adecuado o diferenciar tipos de muestreo	x	x	x	x	x
	Dados los resultados de una muestra y el tamaño de la población decidir la composición más probable de la población	x			x	x
	Calcular el promedio de todos los promedios de muestras de igual tamaño extraídas desde una población y analizar la relación	x	x		x	
	Analizar las medias obtenidas de las muestras, cuando el número de datos de las muestras aumenta	x	x		x	
	Determinar las fases de un estudio estadístico	x	x	x	x	x
	Argumentos	Utilizar herramientas tecnológicas	x	x	x	x
Uso de Ejemplos y contraejemplos		x	x		x	x
Apoyo gráfico para comprobación de propiedades		x			x	x
Razonamiento inductivo a partir de datos		x			x	x
Verificar, utilizando herramientas tecnológicas, las conjeturas formuladas.		x	x	x	x	x

1.6. OBJETIVOS DEL TRABAJO

Esta investigación se dirige a completar los estudios anteriores, para proporcionar información en el contexto chileno, donde no hay información sobre el razonamiento de los estudiantes. Persigue tres objetivos en cada uno de los cuales pretendemos aportar información original:

Objetivo 1. Analizar la presencia del concepto de muestreo en las directrices curriculares de Educación Secundaria chilena (MINEDUC, 2013, 2015), comparando con las directrices españolas y otras internacionales y analizando su implementación en los textos escolares chilenos.

Con ello se determinará el significado de referencia de nuestro trabajo y la idoneidad didáctica de los textos. Este estudio proporcionará la base para determinar de modo objetivo el contenido semántico del cuestionario de evaluación.

Objetivo 2. Construir un instrumento válido y fiable para evaluar la comprensión del muestreo en los alumnos chilenos de Educación Secundaria, siguiendo unos criterios metodológicos rigurosos.

Una vez fijado el contenido del cuestionario, con un proceso riguroso se construirá y validará un cuestionario adecuado, siguiendo los pasos descritos en la metodología, asegurando calidad en el instrumento (McDonald, 2013). La necesidad del instrumento se justifica por no haber encontrado uno en los antecedentes que sirva para evaluar la comprensión del significado de referencia fijado para el muestreo.

Objetivo 3. Aplicar el cuestionario construido para evaluar la comprensión que alcanzan los estudiantes chilenos al comienzo y final de la educación secundaria sobre el muestreo y describir las principales dificultades observadas.

Mediante un estudio de evaluación con estudiantes chilenos de secundaria trataremos de determinar los elementos del significado del muestreo que son comprendidos en cada uno de los grupos analizados y se identificarán los posibles conflictos semióticos que muestren los estudiantes. Se trata de ver si los objetivos planteados en los lineamientos curriculares y los textos escolares son alcanzados por los alumnos chilenos.

1.7. HIPÓTESIS INICIALES

Esta tesis pretende profundizar sobre el concepto de muestra, que introduce a la inferencia, básica en la investigación y la ciencia y constituye un nexo entre la estadística y probabilidad. Además, todo nuestro conocimiento y juicios sobre el mundo o las personas están basados en el muestreo, pues sólo podemos estudiar una parcela de la realidad en la que estamos interesados. Otras razones, como inexistencia de investigaciones similares y aplicación de los resultados en la mejora del currículo se han expuesto anteriormente. A continuación, se mencionan las hipótesis que asumimos en nuestra investigación.

H1. Existen diferencias entre la presentación del muestreo incluidas en una selección de textos escolares y Bases Curriculares de educación secundaria chilena y el expuesto en otras directrices curriculares internacionales (específicamente, las americanas y españolas).

En los últimos años, las directrices curriculares internacionales como el proyecto GAISE, o los Common Core State Standar Initiative, han dado énfasis al concepto de muestreo como uno de los contenidos básicos de la educación estadística. Nuestra conjetura es que el contenido en el currículo chileno es más amplio que el sugerido en estas directrices.

H2. Existen dificultades en la comprensión del concepto de muestreo por parte de los alumnos chilenos, que van desapareciendo, pero no totalmente al avanzar la edad. En concreto, se esperan los siguientes resultados, que se han encontrado en investigaciones previas en otros contextos:

H2.1: En tareas de generación de muestras (formar muestras, conocida la población), la conjetura es que los estudiantes elaboran muestras que reproducen el valor medio en la población, pero no tienen en cuenta la variabilidad muestral.

H2.2: En tareas de estimación (de la media de una población) a partir de la muestra, una proporción importante de estudiantes aplicarán la equiprobabilidad (dando un valor al azar). Se observarán también dificultades al aplicar la proporcionalidad inversa.

1.8. ORGANIZACIÓN DE LA INVESTIGACIÓN

Esta investigación se organiza en dos partes, en la primera, se exponen los fundamentos teóricos, y en la segunda, se recogen los estudios empíricos realizados (ver figura 1.4).

Los fundamentos teóricos incluyen los siguientes lineamientos:

-) *Análisis curricular*, el cual se ha detallado en este mismo capítulo, que brinda un panorama general de los contenidos que se abordan en los diferentes cursos de los estudiantes participantes en el Estudio y ofrece una perspectiva apropiada para el posterior análisis de los libros de texto.
-) *Marco teórico*, en este trabajo se utilizaren algunos elementos del enfoque ontosemiótico desarrollados por Godino y colaboradores (Godino, 2002; Godino y Batanero, 1994; 1998; Godino, Batanero y Font, 2007). El análisis curricular como el marco teórico, se presentan en este capítulo.
-) *Antecedentes de la investigación* los que permiten vincular este trabajo con otros anteriores. Los antecedentes se exponen en el Capítulo 2 de este documento.

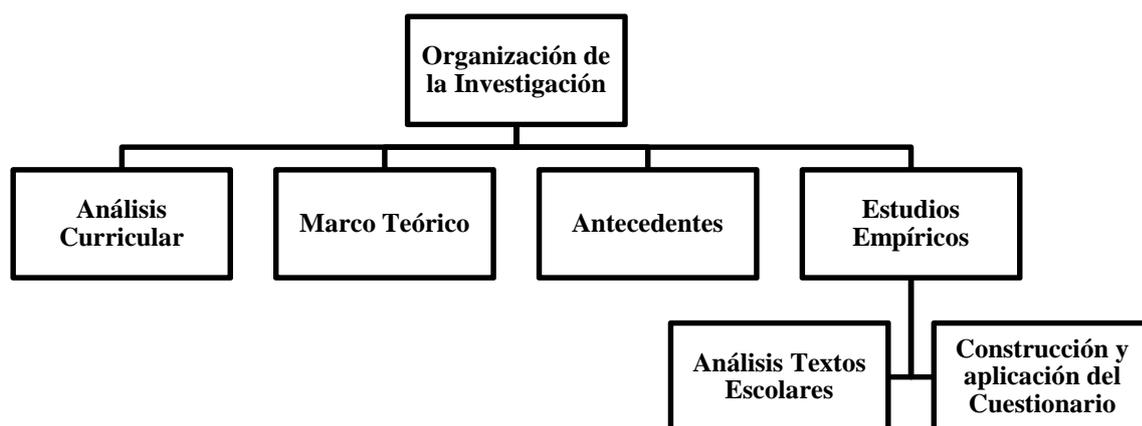


Figura 1.4. Organización de la investigación. Elaboración propia.

Los estudios empíricos que se realizarán son los siguientes:

-) *Análisis de textos escolares*, se analizará una muestra de textos escolares chilenos, en los que estudiamos detalladamente los distintos tipos de objetos matemáticos que forman parte del marco teórico, y la forma en que se presentan e interrelacionan.
-) *Construcción del cuestionario*, se elaborará un cuestionario para la recopilación de datos, ya que no se ha encontrado uno en los antecedentes.
-) *Aplicación del cuestionario*, se administrará a una muestra de alumnos de secundaria chilenos y se analizarán sus respuestas tanto cualitativa como cuantitativamente.

1.9. DESCRIPCIÓN GENERAL DE LA METODOLOGÍA

Se trata de una investigación mixta con componentes cualitativos y cuantitativos. Dentro del contexto chileno es exploratoria, pues no se tienen antecedentes en dicho contexto. Es aplicada pues se espera que sus resultados permitan la mejora de la planificación docente. Tiene un componente documental (estudio de documentos curriculares y textos) y una parte de investigación de campo, donde se recogerán datos con un cuestionario. La investigación se organiza en tres etapas:

1. Análisis del muestreo en las orientaciones curriculares y los textos escolares.

Puesto que la evaluación del estudiante se debe conectar con la enseñanza recibida, se analizará los documentos que establecen el currículo de Educación Secundaria en Chile,

identificando las orientaciones y metodología sobre el muestreo. Como comparación se analizarán también las orientaciones curriculares españolas y algunas propuestas internacionales como el Common Core State Standar Initiative y el Proyecto GAISE. Además, se realizará un análisis de contenido de una colección de textos escolares de matemática, para identificar los contenidos incluidos relacionados con el muestreo.

2. Construcción de un cuestionario de evaluación.

Partiremos de poblaciones discretas (por ejemplo, bolas en urnas, niños en una clase, etc.) e interesándonos por la proporción. Se consideran dos tipos de tareas: a) generación de muestras, conocida la población y b) estimación de la proporción en la población, a partir de datos de una muestra. Variaremos las variables en la tarea (sucesos equiprobables o no; forma en que se dan los datos, etc.).

Una vez fijado el contenido del cuestionario, se iniciará una recopilación de ítems posibles, tomados del análisis de los textos escolares y de investigaciones previas. Posteriormente, se seleccionarán los ítems del cuestionario mediante juicio expertos. Para ello se cuenta con la colaboración de investigadores en Educación Estadística a los que se mandará un cuestionario para que evalúen la calidad de los ítems y la correspondencia de cada uno con contenidos específicos del cuestionario. Una vez preparado el cuestionario se realizará una prueba piloto para determinar su legibilidad y el tiempo requerido en responderlo, así como sus características psicométricas (índices de dificultad y discriminación; fiabilidad; validez) (McDonald, 2013).

3. Estudio de evaluación de la comprensión de los estudiantes.

Se tomará una muestra de estudiantes de los cursos 8° básico (equivalente a 2° de Educación Secundaria Obligatoria en España), 2° medio (4° de ESO en España) y 4° medio (2° de Bachillerato en España) para analizar la progresión de su comprensión con la edad. Se eligen estos niveles escolares porque de acuerdo a las modificaciones curriculares, en ellos se incluyen algunas ideas de muestreo. Se eligieron 5 o más grupos clase de cada nivel; en total unos 1241 alumnos, 247 de octavo básico, 529 de segundo medio y 438 estudiantes de cuarto medio, de 4 colegios y 2 liceos (institutos de secundaria) que representen las clases sociales y económicas adecuadamente de la ciudad de Osorno, Chile.

Se realizará un análisis cuantitativo y cualitativo de las respuestas. En el análisis cuantitativo se incluye índices definitivos de dificultad y discriminación para cada tarea; coeficiente de fiabilidad Alpha de Crombach; puntuación total; diferencia por curso; análisis factorial (para estudiar las principales dimensiones del cuestionario). Se justificará igualmente la validez de contenido, discriminante y de constructo del cuestionario.

En el análisis cualitativo se identificarán los principales conflictos semióticos presentes en las respuestas erróneas. Para ello se elegirá una respuesta típica en cada error encontrado con frecuencia apreciable (más del 10% de la muestra) y se realizará un análisis semiótico, como sugiere Godino (2002).

CAPÍTULO 2.

ANTECEDENTES

- 2.1. Introducción
- 2.2. Niveles de uso de los conceptos de inferencia
- 2.3. Comprensión de las propiedades de muestreo
 - 2.3.1. Heurísticas y sesgos
 - 2.3.2. Comprensión de las propiedades de las muestras
 - 2.3.3. Comprensión de la variabilidad en el muestreo
- 2.4. Comprensión de la distribución muestral
- 2.5. Comprensión de la probabilidad, desde el punto de vista frecuencial
 - 2.5.1. Estimación de una probabilidad a partir de la frecuencia
 - 2.5.2. Características de las secuencias aleatorias de resultados
- 2.6. Muestreo en el razonamiento inferencial informal de los estudiantes
- 2.7. Conclusiones del estudio de los antecedentes

2.1. INTRODUCCIÓN

A continuación se presenta un resumen de las principales investigaciones relacionadas con la comprensión del muestreo y sus propiedades, y de conceptos cuya comprensión es necesaria para entenderlo, como la distribución muestral. Para realizar esta síntesis, hemos partido de las contenidas en Begué (2016), Ben-Zvi, Bakker y Makar (2015), Harradine, Batanero y Rossman (2011) y Tauber (2011).

Primero, se estudiarán los distintos niveles del uso del concepto de inferencia estadística, indicando en el análisis la importancia y complejidad del concepto de distribución y las principales dificultades que pueden presentar los alumnos.

Segundo, se realizará una revisión de las investigaciones centradas en la comprensión de las propiedades del muestreo. Dicho análisis señala las principales concepciones erróneas que presentan los sujetos sobre las propiedades de la muestra y del

muestreo. Se mencionan además, los trabajos más relevantes realizados con alumnos, centrados en analizar la comprensión del concepto de muestra que resulta al llevar a cabo un muestreo con ayuda de la simulación, especialmente, con el apoyo de software especializado.

Finalmente, se abordan las investigaciones enfocadas en la comprensión del significado frecuencial de la probabilidad.

2.2. NIVELES DE USO DE LOS CONCEPTOS DE INFERENCIA

Actualmente, ha aumentado el interés por la enseñanza de la inferencia estadística, disciplina que nos proporciona diferentes herramientas para afrontar la toma de decisiones en diversos ámbitos.

Si analizamos los lineamientos curriculares en el ámbito escolar, vemos que, en los cursos inferiores, los conceptos estadísticos enseñados principalmente se demarcan fundamentalmente a los de estadística descriptiva. No obstante, en los últimos años, distintos investigadores en educación matemática sugieren un rol más amplio y más profundo para la estadística en la matemática escolar (Makar y Ben-Zvi, 2011), es por ello, que sugieren que los fundamentos del razonamiento estadístico, como también, las ideas fundamentales de la estadística inferencial, deberían incorporarse desde los primeros años de escolaridad (Batanero, 2013; Ben-Zvi, Bakker y Makar, 2015; Meletiou-Mavrotheris y Papanastasiou, 2015).

Siguiendo las ideas planteadas por Watson (2004), si bien el muestreo es un concepto fundamental de la inferencia estadística, es un concepto que inusualmente logra una alta presencia en los documentos curriculares a nivel escolar, debido a que el tema es más descriptivo y menos numérico que el resto del currículo de matemática o porque se considera difícil de comprender por los estudiantes, dado que el desarrollo del razonamiento relacionado al muestreo está ligado al desarrollo de habilidades de alfabetización y razonamiento social en lugar de fomentar el desarrollo de habilidades de cálculo numérico.

Harradine, Batanero y Rossman (2011) indican que la estadística inferencial consta de tres elementos fundamentales que interactúan entre sí: (1) el proceso de razonamiento, (2) los conceptos y (3) los cálculos asociados. Además, realzan la importancia de los conceptos: probabilidad, muestreo aleatorio, parámetro, distribución de valores de una

muestra estadística, intervalos de confianza, hipótesis nula y alternativa, p valor, nivel de significación y lógica de inferencia. Así mismo, mencionan que la tecnología juega un rol esencial en el proceso de cálculo ya que actualmente los softwares facilitan dicho proceso. Es por ello, que recalcan que los profesores de estadística deben enseñar los tres componentes y no solo los mecanismos de inferencia, porque las principales dificultades para entender la inferencia estadística se encuentran en los procesos de razonamiento y en los conceptos.

Burrill y Biehler (2011), indican que la teoría del muestreo, estudia cómo se seleccionan las muestras y sus propiedades, para obtener, a partir de ellas, conclusiones con cierto grado de certeza. Al considerar el muestreo como un concepto esencial y en la inferencia estadística, es importante desarrollar su comprensión, pues es un concepto crucial para seguir profundizando en temas como los intervalos de confianza y el contraste de hipótesis y no proyectar errores de comprensión previos en dichos conceptos.

De acuerdo a Batanero, Godino, Green, Holmes y Vallecillos (1994) el muestreo es uno de los conceptos centrales de la inferencia estadística, además, mencionan que la comprensión de dicho concepto implica distinguir la diferencia entre la representatividad y la variabilidad. Además, resumen los principales errores y sesgos asociados a la comprensión del muestreo, algunos de los cuales se describirán en las secciones siguientes de este capítulo.

2.3. COMPRENSIÓN DE LAS PROPIEDADES DEL MUESTREO

Dentro de las ideas estocásticas fundamentales, Heitele (1975) distingue al muestreo como uno de los conceptos más relevantes, ya que lo estima como parte esencial de la conexión entre la estadística y la probabilidad. Al mismo tiempo, plantea que los procedimientos vinculados al muestreo son parte de nuestra vida cotidiana, ya que nuestro conocimiento se establece por medio de la percepción de fragmentos de la realidad, que, en algunas circunstancias, puede ser muy difícil de llegar a observarla completamente.

Siguiendo estas ideas, Moses (1992) señala que en la inferencia estadística, una muestra otorga *alguna* información sobre la población, aumentando así nuestro conocimiento sobre esta última. Es decir, la inferencia estadística, puede entenderse como un conjunto de métodos para aprender a partir de la experiencia.

Según Batanero, Godino, Vallecillos, Green, y Holmes (1994), la comprensión del muestreo supone vincular dos ideas supuestamente contradictorias: (1) *representatividad muestral*, la cual señala que una muestra aleatoria de tamaño suficiente tendrá características semejantes a las correspondientes características de la población y (2) *variabilidad muestral*, la cual admite que las muestras pueden cambiar en su composición y características.

2.3.1. HEURÍSTICAS Y SESGOS

Tauber (2001) indica que dentro de la investigación educativa, el empleo de heurísticas, es el campo más estudiado en las investigaciones psicológicas. Éstas se definen como mecanismos que sirven para reducir la “incertidumbre que induce nuestra limitación para afrontar la complejidad de las situaciones inciertas” (Pérez Echeverría, 1990). En el razonamiento probabilístico, una heurística simplifica la información del problema y ayuda a resolverlo.

Dentro de las investigaciones desarrolladas en psicología sobre razonamiento estadístico, destaca el trabajo de Kahneman, Slovic y Tversky (1982), quienes evidencian los fallos persistentes, que se producen al realizar juicios en condiciones de incertidumbre probabilística. Sugirieron además que, no somos aptos para producir juicios probabilísticos adecuados y recurrimos a tácticas mentales inconscientes o “*heurísticas*”, que disminuyen la complejidad del problema eliminando una parte de los datos. Sin embargo, causan errores y son resistentes al cambio. Éstos son sistemáticos y se repiten, además en pocas oportunidades pueden corregirse, aún en personas que han recibido instrucción significativa en probabilidad.

Kahneman, Slovic y Tverky (1982) explican “tres heurísticas fundamentales en los juicios probabilísticos: “representatividad, disponibilidad y ajuste y anclaje” (p.1124), además de los sesgos asociados. A continuación, se describen las principales heurísticas y sesgos asociados a la comprensión del muestreo. La heurística de “ajuste y anclaje” no es considerada en nuestro estudio.

Heurística de representatividad

La heurística más presente en el campo de la inferencia, es la heurística de la *representatividad*. Los sujetos que la manifiestan creen que la “verosimilitud de un suceso considerando únicamente su representatividad respecto a la población a la cual pertenece”

(Tversky y Kahneman, 1974, p.1124), es decir, como mencionan Harradine, Batanero y Rossman (2011) y Watson (2004), los alumnos presentan dificultades con la idea de la variabilidad en las poblaciones, tienen demasiada confianza en muestras pequeñas, y no consideran la importancia del tamaño de la muestra en muestras aleatorias, lo que puede generar conflictos al momento de seguir profundizando en los conceptos de inferencia estadística.

Tversky y Kahneman (1974) mencionan que encontramos dos sesgos asociados a dicha heurística: *la insensibilidad al tamaño de la muestra y la falacia del jugador*. El primero de ellos, la *insensibilidad al tamaño de la muestra*, donde una persona acepta que “una muestra, aunque sea pequeña, e independientemente de cómo se haya escogido, siempre es representativa de la población”.

Kahneman, Slovic y Tversky (1982) señalan que este tipo de personas, también cree en la *ley de los pequeños números*, es decir, creen que, al seleccionar muestras no suficientemente grandes, la distribución muestral se distribuye de la misma manera que la distribución de la población, independientemente del tamaño de la muestra.

El segundo de ellos, la *falacia del jugador*, señala que las personas creen en “que el resultado de un experimento aleatorio afectará en la probabilidad de sucesos futuros” (Begué, Batanero, Ruiz y Gea, 2019, p.60). En este caso, se espera que la probabilidad de obtener el suceso contrario aumente, después de una racha de un mismo suceso, por ejemplo, los sujetos creen que después de una racha larga de caras, es más probable obtener una cruz.

Estudios sobre el efecto de la enseñanza en la representatividad

Serrano (1993) analizó cómo influye la *heurística de la representatividad* en el trabajo con simulación de experimentos aleatorios. Muestra los resultados obtenidos en entrevistas individuales, a diez alumnos de magisterio y otros diez estudiantes de bachillerato (14 años), para comparar si manifiestan en sus respuestas la heurística de la representatividad. El objetivo del estudio fue examinar si el enfoque frecuencial propuesto en el currículum, es verdaderamente fácil para los alumnos de secundaria. Para ello, introduce el enfoque frecuencial de la probabilidad, elaborando diversas series de actividades didácticas. Los resultados evidencian que no se reflejan grandes diferencias en cuanto a las concepciones probabilísticas incorrectas, entre los futuros profesores y los

alumnos de bachillerato. Indican además que los alumnos se orientan por los resultados específicos de las simulaciones realizadas y varían sus nociones iniciales de acuerdo a los resultados que van obteniendo en dichas simulaciones.

En un estudio posterior, Serrano (1996) contrasta el uso de diferentes heurísticas, empleando una muestra compuesta por 140 estudiantes de 14 años que no tuvieron instrucción previa de probabilidad y otros 140 de 17 años al finalizar la secundaria, los cuales habían estudiado probabilidad en cursos anteriores.

A través de cuestionarios y el análisis de argumentos de los estudiantes, utilizando las técnicas de análisis de contenido y de análisis multivariante, el investigador observa que no se emplea tan comúnmente la *heurística de la representatividad* como mencionan otros investigadores, debido a que las soluciones de los estudiantes esencialmente se deben al *enfoque en el resultado* (Konold, 1989). También se observa que al aumentar la edad de los participantes, éstos usan de mejor manera el concepto de representatividad (Batanero, Serrano y Garfield, 1996; Serrano, Batanero, Ortiz y Cañizares, 1998).

Igualmente delMas (1988) implementa una tarea de enseñanza de conceptos probabilísticos con alumnos de enseñanza secundaria, empleando la experimentación y la simulación. Aplica un pretest y un postest para evaluar el aprendizaje de 78 estudiantes, centrándose en la estabilidad de la heurística de la representatividad. Otras experiencias y sus análisis, son reseñadas en Garfield y delMas (1989, 1990). Dichos trabajos, concluyen que los estudiantes que en el pretest utilizaban en parte una estrategia frecuencial correcta, mostraban una leve mejora con la instrucción, por otro lado, hubo un mayor número de participantes que usaban en el postest esa misma estrategia, de donde se infiere que la instrucción tiene un carácter efectivo ya que mejora los resultados de los estudiantes. No obstante, varios alumnos que emplearon en el pretest una estrategia correcta alternaron en el postest a una estrategia incorrecta, además, algunas respuestas de los estudiantes parecen no mejorar con la instrucción ya que en ambos test aplicaban estrategias incorrectas.

Dentro de las investigaciones que han estudiado los errores en situaciones probabilísticas donde se emplea la idea de muestreo, encontramos a Pollatsek, Konold, Well y Lima (1991) quienes examinan las concepciones sobre el muestreo de alumnos y expertos en estadística. Los resultados evidencian que los estudiantes tienen

concepciones erróneas sobre el muestreo y en sus respuestas se evidencia la heurística de la representatividad.

Heurística de la disponibilidad

Tversky y Kahneman (1974) indican que en la *heurística de la disponibilidad* “hay situaciones en las que las personas evalúan la frecuencia de una clase o la probabilidad de un evento por la facilidad con la que se pueden traer a la mente casos o sucesos” (p.1127), es decir, el cálculo de probabilidades se ve influenciada por los datos previos que ese sujeto tiene sobre el suceso en estudio.

Sesgos asociados a la heurística de la disponibilidad

Esta heurística está asociada a tres sesgos, los cuales se describen más detalladamente en Begué, Batanero, Ruiz y Gea (2019) y se resumen a continuación:

(1) *el sesgo de equiprobabilidad* (Lecoutre, 1992) “que consiste en pensar que los resultados de cualquier fenómeno aleatorio, en particular, cualquier valor de un estadístico muestral, son igualmente probables” (p.55). Chernoff y Russel (2012) lo describen como “la *falacia de la composición*, que consiste en transferir a un todo, una propiedad que se cumple en una de sus partes” (p.55).

(2) *el enfoque del resultado* (Konold, 1989), “que explica el resultado de un experimento desde el enfoque subjetivo de la probabilidad. En otros términos, el sujeto no interpreta la situación desde un enfoque frecuencial, porque identifica en forma demasiado exacta frecuencia relativa con probabilidad” (p.55).

(3) *la ilusión de control* (Langer, 1982) consiste en no distinguir entre los juegos de azar y los de habilidad, es decir, las personas consideran que puede controlar el azar lo que conlleva a no comprender la aleatoriedad. (p.55).

2.3.2. NIVELES DE DESARROLLO EN EL RAZONAMIENTO SOBRE EL MUESTREO

A continuación, se describen los principales trabajos de investigación desarrollados con alumnos de educación primaria y secundaria a cerca de la comprensión del muestreo.

Algunos investigadores (por ejemplo, Shaughnessy, 1992; Shaughnessy, Garfield, y Greer, 1996, Watson, 1997; 2004; Jacobs, 1997; Watson y Kelly, 2005) han sugerido que los escolares tienen dificultades con la comprensión del concepto de muestreo y tratan de definir grados de desarrollo de este tipo de razonamiento.

El estudio de Watson (1997) establece tres niveles de jerarquización para la comprensión del concepto de muestra: (1) comprensión de los términos asociados a la toma de muestras; (2) aplicación y comprensión del muestreo en diversos contextos, y (3) habilidades críticas necesarias para rebatir afirmaciones sobre muestras elegidas sin un fundamento estadístico apropiado.

En la investigación desarrollada por Jacobs (1997), se realizan dos estudios para indagar acerca del entendimiento informal de los problemas de muestreo en el contexto de la interpretación y valoración de los resultados de la encuesta. Esta autora, analizó la comprensión de los alumnos de cuarto y quinto grado (9 - 11 años), evaluando un problema con encuestas que emplean diferentes tipos de métodos de selección de muestras.

El estudio 1, fue diseñado para identificar categorías de las concepciones de la toma de muestras de los niños a partir de entrevistas. Se entrevistaron individualmente a 17 niños. Se les pidió evaluar una variedad de métodos de muestreo y sacar conclusiones cuando se les presentaron encuestas múltiples. En el estudio 2, 110 niños completaron tareas escritas para confirmar las categorías de respuesta identificadas en el estudio 1 y para determinar la prevalencia de las categorías de respuesta en una muestra mayor.

Los estudiantes evaluaron los métodos de muestreo que se centran en la posibilidad de sesgo, la equidad, problemas prácticos, o en los resultados. Todos los niños utilizan varios tipos de lógicas de evaluación, y el foco de sus evaluaciones variaron un poco por el contexto y el tipo de método de muestreo (restringido, auto-seleccionado, o aleatorio). Utilizaron justificaciones efectivas (imparciales) más a menudo en el contexto escolar y las razones se centraron en los resultados más frecuentemente en contextos fuera de la escuela. Presentaron más dificultades para detectar el sesgo de los métodos de muestreo auto-seleccionados que con los métodos de muestreo restringidos debido a la autoselección fue inicialmente más justa (es decir, todos tuvieron la oportunidad de participar). Los estudiantes prefieren el muestreo aleatorio estratificado que el muestreo aleatorio simple, ya que querían asegurarse de que se incluyeran todos los tipos de individuos.

A la hora de sacar conclusiones de varios estudios, los niños: (1) consideran la calidad de la encuesta; (2) agregan todas las encuestas independientemente de la calidad; (3) utilizan sus propias opiniones e ignoran todos los datos de la encuesta; o (4) se niegan

a sacar conclusiones. Incluso cuando los niños fueron capaces de identificar los posibles sesgos, a menudo ignoran la calidad de la encuesta a la hora de sacar conclusiones a partir de múltiples encuestas.

Como se ha señalado anteriormente, a pesar de la importancia que se le atribuye al concepto de muestreo, Watson y Moritz (2000) consideran que existe escasa investigación sobre el desarrollo de las cogniciones del muestreo por parte de los estudiantes en el campo de didáctica de la estadística, a pesar de su paulatina incorporación en las diferentes directrices curriculares.

Watson y Moritz (2000) examinan cómo los estudiantes construyen el concepto de muestra, aplicando y analizando una entrevista con preguntas abiertas sobre el muestreo a 62 alumnos de 3º, 6º y 9º grado (edades entre 8-9, 11-12 y 14-15 años, respectivamente). Los resultados reflejan diferentes niveles de desarrollo creciente en cuanto a los niveles propuestos por Watson (1997). Los autores consideran que hay un mejor rendimiento a medida que va aumentando la edad de los participantes del estudio. Los estudiantes de 3º, manifiestan nociones muy elementales y particulares de las muestras, específicamente de aquellas ideas derivadas de experiencias de la vida diaria, con productos de muestra (por ejemplo, la muestra de algún producto en el supermercado). Además, confiaban en obtener conclusiones sobre la población basadas en los datos de muestras muy pequeñas, con escasa preocupación sobre el sesgo. Los alumnos de 6º, expresan una diversidad de creencias sobre el tamaño de la muestra y el método de muestreo. Por el contrario, la mayoría de los estudiantes de 9º consideraban la variación en la población, es decir, manifestaban la necesidad de contar con una muestra suficientemente grande y representativa. Sin embargo, a menudo no identificaban el sesgo asociado a la muestra.

Watson (2004) realiza un estudio del razonamiento de los estudiantes de secundaria sobre el muestreo, aplicando entrevistas longitudinales con 38 estudiantes a lo largo de 3 o 4 años; de grados 6 al 12 (11-18 años). Inicialmente los estudiantes estaban en 3º de primaria (8), 6º de primaria (18), 2º de educación secundaria (7) y 4º de secundaria (5). Les interroga para analizar su comprensión acerca del concepto de muestra, cómo se deben recoger muestras, y el poder representativo de una muestra basada en su tamaño. Utiliza el modelo establecido en el estudio de Watson (1997), ver Tabla 2.1.

Tabla 2.1. Niveles de comprensión del muestreo según Watson (1997)

Nivel 1. Comprende la terminología			
Categoría	1.	Muestras pequeñas sin selección	Dan ejemplos de muestras Pueden describir una muestra como una parte pequeña Sugieren muestras de menos de 15 elementos No dan método de selección o dan un método idiosincrático
Categoría	1.	Muestras pequeñas con selección primitiva	Dan ejemplos de muestras Pueden describir una muestra como una parte pequeña Sugieren muestras de menos de 15 elementos Sugieren selección por métodos aleatorios sin describirlo o con descripción sencilla
Nivel 2. Comprende la terminología en contexto			
Categoría	3.	Muestras pequeñas con pre-selección	Dan ejemplos de muestras Pueden describir una muestra como una parte pequeña Sugieren muestras de menos de 15 elementos Sugieren selección de personas por peso, por ejemplo, delgados, gordos o de peso normal
Categoría	4.	Muestras equívocas	Dan ejemplos de muestras Muestran indiferencia sobre el tamaño de la muestra, basándose en aspectos irrelevantes Combinan tamaños pequeños con método apropiado de selección o sensibilidad al sesgo o a grandes muestras con métodos de selección inapropiados
Categoría	5.	Muestras grandes con selección distribuida	Dan ejemplos de muestras Pueden describir una muestra como una parte pequeña Pueden referirse al promedio Sugieren muestras de al menos 20 elementos o un porcentaje de la población Sugieren métodos de selección aleatorios o geográficos
Nivel 3. Cuestionamiento crítico			
Categoría	6.	Muestras grandes, sensibilidad al sesgo	Dan ejemplos de muestras, a veces de encuestas Pueden describir una muestra como una parte pequeña Pueden referirse al promedio o al valor representativo Sugieren muestras de al menos 20 elementos o un porcentaje de la población Sugieren métodos de selección aleatorios o geográficos Se preocupan por evitar el sesgo en el muestreo Identifican muestras sesgadas en noticias o resultados de encuestas

Los alumnos prefieren elegir muestras seleccionadas con algún procedimiento sesgado, es decir, muestras entregadas voluntariamente, sobre a las seleccionadas mediante un método de muestreo aleatorio. Los niños van progresando en las diferentes categorías a lo largo del estudio, es decir, mejoran en general desde la primera a la última entrevista. Los alumnos de escuela primaria inicialmente están en las dos primeras, los de comienzos de la secundaria entre la 1 y la 5 y los del final de secundaria entre la tres y la seis. Al finalizar el estudio estos chicos tienen 3-4 años más y han pasado a las categorías 2-6 (niños de primaria), 1 a 6 (principios de secundaria) y 4 a 6 (fin de secundaria).

En esta investigación, se han utilizado, principalmente dos marcos teóricos:

- J) El primero se basa en la taxonomía SOLO de Biggs y Collis (1991), la que se fundamenta en la estructura de los resultados de aprendizaje observados. En esta taxonomía se analizan las respuestas según la mayor complejidad estructural: a) *Uniestructural* (U), si las respuestas emplean elementos individuales de las tareas y no reconocen las contradicciones si se produjeran; b) *Multiestructural* (M), cuando las respuestas utilizan más de un elemento de forma secuencial, y a menudo reconocen las contradicciones, pero son incapaces de resolverlas; y c) *Relacional* (R), si las respuestas de los elementos de las tareas para producir soluciones completas están libres de contradicciones integradas.
- J) El segundo marco, debido a Watson (1997) usa tres grados de la jerarquía de la cultura estadística aplicada a la toma de muestras: a) *Nivel 1*, relacionado con la comprensión de la terminología asociada con la toma de muestras; b) *Nivel 2*, consideró la aplicación y comprensión de la terminología de muestreo teniendo en cuenta el contexto, en particular los contextos sociales que aparecen en los medios de comunicación y c) *Nivel 3*, se asoció con las habilidades críticas necesarias para cuestionar las afirmaciones sobre muestras hechas sin fundamento estadístico adecuado. En la Tabla 2.1 se reproduce las características de cada nivel, que se dividen a su vez en seis categorías de razonamiento.

Bakker (2004) examina un secuencia de enseñanza con estudiantes de secundaria (13-14 años), basado en el uso de representaciones gráficas de las muestras obtenidas, donde va aumentando el tamaño de las muestras. Señala que el proceso de enseñanza contribuyó a que los alumnos percibieran que un mayor tamaño de la muestra implica una menor variabilidad en el muestreo, lo que posibilita mejorar las conclusiones acerca de las características de la población

Watson y Kelly (2005), obtienen resultados semejantes a lo descrito en Watson (2004), en una investigación aplicada a 639 alumnos australianos, de los grados 3º, 5º, 7º y 9º (8-14 años), los cuales fueron encuestados como parte de un estudio más amplio, incluidas las preguntas sobre muestreo. Cuando se les pregunta a los estudiantes por la definición del término “muestra”, la mayoría proporciona una única idea asociada con el concepto o solo dieron un ejemplo (40.2%), otros no fueron capaces de definir el término en absoluto (32.5%). Una mayor cantidad de estudiantes de 7º (12.6%) y 9º (11.6%) alcanzaron el nivel óptimo de respuesta a dicha pregunta, y aproximadamente un

porcentaje igual de estudiantes en 5° y 7° grado respondieron parcialmente correcto (18.2% y 21.2%, respectivamente). En 3°, la mayoría fueron respuestas incorrectas (63.6%), pero este porcentaje fue disminuyendo en los cursos posteriores, a 29.3% en 5°, 2.5% en 7° y 18.3% en 9°.

El trabajo de Begué (2016), tiene como objetivo analizar la comprensión de algunas ideas elementales de muestreo con estudiantes de Educación Secundaria Obligatoria de 2° y 4° curso. La muestra estuvo constituida por un total de 302 alumnos, de los cuales 157 son de 2°ESO (12-13 años) y 145 de 4°ESO (15-16 años).

Esta investigación evalúa la comprensión acerca de la variabilidad en el muestreo y del efecto del tamaño de la muestra sobre dicha variabilidad. Analiza además, la comprensión de la relación entre el valor de una proporción en la población y la frecuencia relativa esperada en muestras tomadas de dicha población. Para ello se aplica y estudia un cuestionario desde la reflexión sobre la respuesta esperada para cada uno de los ítems que lo constituye. También se examinan las respuestas para cada ítem y se realiza un estudio comparativo de las respuestas según el grupo participante, identificando los posibles sesgos o concepciones erróneas acerca de los conceptos evaluados.

El instrumento de evaluación está constituido por cuatro ítems, dos de ellos demandan la producción de muestras grandes y los otros dos de pequeñas. Además, en dos de los ítems la asignación de probabilidad se hace de forma frecuencial, mientras que los otros dos se apoyan en el significado clásico de la probabilidad a través de la regla de Laplace. Los fenómenos aleatorios que se describen los dos primeros ítems, a diferencia de los otros dos, no presentan el principio de equiprobabilidad.

Generalmente, los estudiantes tienen una buena intuición del valor esperado, es decir, la proporción de resultados para el caso favorable pedido se acerca a la proporción teórica. No obstante, alrededor de un 30% dan como valor esperado el 50% incluso en los experimentos donde la equiprobabilidad no es clara, como es lanzar una chincheta y observar si cae de punta o hacia abajo.

Otra de las conclusiones planteadas en esta investigación, es que los alumnos presentan una mejor intuición o comprensión del muestreo cuando el experimento aleatorio está asociado a un número de ensayos menor y los sucesos elementales son equiprobables. Por otro lado, en las dos tareas en que la muestra es grande los estudiantes, independiente del curso, proporcionan muestras con una variabilidad extrema, siendo la

frecuencia mayor en el grupo de 2ºESO. Se observa por tanto, que los estudiantes no presentan una apropiada comprensión de la variabilidad cuando el tamaño de la muestra es mayor, es decir, los estudiantes no presentan una intuición adecuada de la Ley de los Grandes Números.

Finalmente, la autora señala que los alumnos no comprenden el efecto del tamaño de la muestra sobre la variabilidad en el muestreo, resultado que aparece en las investigaciones previas de Serrano (1996), Shaughnessy et al. (2004), Gómez (2014) y Gómez et al. (2014).

2.3.3. COMPRENSIÓN DE LA VARIABILIDAD EN EL MUESTREO

Moore (2004) sostiene que realizar inferencias en estadística implica “ir más allá de los datos disponibles para elaborar conclusiones acerca de un “universo más amplio”; al hacer una inferencia siempre se debe tener en cuenta que la variación está siempre presente, por lo que las inferencias son inciertas” (García-Ríos, 2013, p. 344).

Shaughnessy, Watson, Moritz, y Reading (1999) presentan una mezcla conocida de los objetos de color (por ejemplo, 50% de rojo, 50% de otros colores) para obtener muestras de la misma a 700 estudiantes de enseñanza secundaria de tres países. La mayoría reconocen la variabilidad en el número de “rojos” que serán obtenidos en las repetitivas muestras. Sin embargo, los estudiantes difieren en la forma en que presentan variabilidad en sus predicciones, y en las razones de sus predicciones. Estos investigadores también encontraron que los estudiantes de secundaria superior que habían estudiado probabilidad tenían una mayor tendencia a no tener en cuenta la variación en el tipo de predicciones sobre el muestreo, que los estudiantes de la escuela secundaria inferior.

La investigación realizada por Shaughnessy, Canada y Ciancetta (2003) resume el pensamiento de 84 estudiantes de secundaria sobre la variabilidad en tres tareas estocásticas que involucran ensayos sucesivos, aplicando un cuestionario. Las tres tareas de ensayos sucesivos fueron: (1) la toma de muestras, en donde se debe extraer un puñado de 10 caramelos de un recipiente con 100 unidades, de los cuales 60 son de color rojo y 40 de color amarillo; (2) se realiza el experimento de lanzar un dado 60 veces completando una tabla de frecuencias de dichos lanzamientos y se pide argumentar el porqué de dichas respuestas; y (3) se muestra una ruleta, cuya mitad es blanca y la otra

mitad es gris, solicitando que se escriba una lista de 6 series correspondiente a 50 giros. En sus respuestas a las tareas de los dados, la mayoría se centra únicamente en la probabilidad teórica de un único resultado, $1/6$ para cualquier número, mientras que eran mucho más propensos a considerar una serie de posibles los resultados, en la toma de muestras o en la ruleta. Los resultados indican que hubo una tendencia muy fuerte en estos estudiantes a no reconocer la variabilidad en la predicción de la distribución de frecuencias de los resultados del problema de los dados.

Por otra parte, estos estudiantes no fueron consistentes en los tres entornos de trabajo en sus predicciones de la variabilidad en los resultados de los ensayos sucesivos. De los 84 estudiantes en el estudio, sólo 14 de ellos (16%) predijo listas razonables de los resultados para las tres tareas. Son propensos a predecir resultados constantes para los ensayos sucesivos en una situación de probabilidad conocida como la tarea de los dados, y descuidan la cuestión de la variabilidad en las frecuencias de los resultados individuales. Aunque algunos estudiantes reconocieron la variabilidad en las tareas de los dados, estos estudiantes estaban en minoría. Los autores creen que parte de la razón de tal razonamiento se debe a la forma en que la probabilidad se enseña en las escuelas, pues con demasiada frecuencia se apresura llevar a los estudiantes al cálculo de la probabilidad de eventos individuales o probabilidades de resultados particulares, sin tener en cuenta la variación en los resultados que puede ocurrir en ensayos sucesivos reales, es decir, rara vez se da a los estudiantes la oportunidad de desarrollar su intuición para un rango probable de resultados en las situaciones de los ensayos sucesivos.

Shaughnessy, Ciancetta y Canada (2004) estudian la comprensión de la variabilidad en el muestreo en 272 alumnos de 6° a 12° grado (10 - 19 años), empleando problemas de extracción al azar de caramelos de dos colores (60% rojos y 40% amarillos), con cambios en el tamaño de la población y de la muestra. El 25% de estos alumnos esperaban obtener el mismo resultado en dos muestras repetidas del mismo tamaño y algunos estudiantes propusieron en sus muestras resultados muy poco probables (como 100% de amarillos). Los alumnos tienden a creer que los resultados extremos, en muestras reiteradas, ocurrirán con mucha más frecuencia de lo que realmente lo hacen en la práctica. Los autores señalan que la influencia de la enseñanza de probabilidad puede interferir con lo que los estudiantes piensan acerca de la variabilidad, llevando a algunos de ellos a sobreestimar la variabilidad muestral, ya que piensan en la posibilidad de resultados muy variados, independientemente del tamaño de la muestra; mientras que

otros conceden escasa variabilidad al conjunto de posibles resultados en un muestreo, es decir, los estudiantes poseen un pensamiento aditivo, como hay más “rojos”, potenciando el sesgo de enfoque del resultado (Konold, 1989), además, dos tercios de los estudiantes no razonaron proporcionalmente.

Saldanha y Thompson (2002) investigaron acerca del desarrollo del pensamiento de los estudiantes al participar en un experimento de enseñanza diseñado para apoyar su concepción de muestreo como un esquema de ideas interrelacionadas que incluyen la selección aleatoria repetida, la variabilidad entre las estadísticas de la muestra y la distribución. En este estudio, participaron 27 estudiantes de 11° y 12° grado (16 - 18 años), en un experimento de enseñanza de 9 sesiones de clase, donde se abordan ideas de muestra, distribuciones de muestreo y márgenes de error. La instrucción hizo hincapié en dos temas generales y relacionados: 1) el proceso de selección aleatoria puede repetirse en condiciones similares, y 2) los juicios sobre los resultados del muestreo pueden basarse en patrones de frecuencia relativa que surgen en colecciones de resultados de muestras similares. Estos temas estaban destinados a apoyar el desarrollo de una interpretación distributiva del muestreo y la probabilidad.

Estos autores, descubrieron que la mayoría de los alumnos tenían bastantes dificultades para concebir el muestreo en términos de tres niveles distintos: población, muestra y recopilación de estadísticos de muestra. Observan, además, que los estudiantes de bajo rendimiento confundían el número de personas en una muestra con el número de muestras extraídas, asimismo, la mayor parte de los alumnos no logran diferenciar los tres niveles de uso del concepto de distribución, interpretando erróneamente los resultados de las simulaciones como un porcentaje de personas en vez de un porcentaje de la proporción de la muestra. Por otro lado, los estudiantes tuvieron dificultad para imaginar una distribución muestral debido a su tendencia a juzgar la representatividad de una muestra solo en relación con la proporción poblacional subyacente, los investigadores conjeturan que esto se debe principalmente, a que la noción de muestra de estos estudiantes era aditiva, es decir, tendían a ver una muestra simplemente como un subconjunto de una población y a ver muestras múltiples como subconjuntos múltiples.

2.4. COMPRENSIÓN DE LA DISTRIBUCIÓN MUESTRAL

Los procedimientos de la inferencia estadística son empleados para elaborar conclusiones sobre de una determinada población; utilizando en general la distribución muestral, es decir, la distribución de un estadístico como la media muestral en todas las posibles muestras de un tamaño dado de una población. Esta distribución se forma cuando se considera el valor de la media en todas las posibles muestras teóricas de la población.

Chance, Delmas y Garfield (2004) realizaron una investigación donde aplicaron cinco estudios a 114 alumnos universitarios, encontraron que mientras que los estudiantes pudieron observar los comportamientos y patrones de comunicación en el comportamiento (por ejemplo, mayor es el tamaño de la muestra más pequeña la variación) mostrado por los applets de muestreo aleatorio, los alumnos no entendían por qué se produjo el comportamiento. Los autores observaron que, después de la exposición a los applets, los estudiantes fueron capaces de sugerir las distribuciones posibles de muestras para un determinado tamaño de muestra. Estos estudiantes no tienen una clara distinción entre la distribución de una muestra de datos y la distribución de las medias de las muestras. La simple exposición a los applets no es suficiente para garantizar el aprendizaje. Los autores concluyen que: (a) los alumnos necesitan familiarizarse más con el proceso de toma de muestras, (b) las actividades relacionadas con los applets tienen que ser estructuradas y no estructuradas, y (c) los estudiantes necesitan discutir sus observaciones después de una actividad por lo que podrían llegar a centrarse en qué observaciones son más importantes, que observaciones importantes no hicieron y cómo se conectan las observaciones importantes.

Saldahna y Thompson (2002) estudiaron las concepciones de 27 estudiantes de grados 11° y 12° de secundaria, que participan en actividades que utilizan applets para simular un muestreo aleatorio repetido de una población. La actividad requiere que los estudiantes elijan al azar una muestra de una población, calculen una proporción de la muestra y repitan este proceso una y otra vez. Los autores señalan que un gran número de los estudiantes han mostrado grandes dificultades en la concepción de un muestreo repetido en términos de tres niveles distintos: población, muestra, recogida de estadísticas de la muestra. Estas dificultades llevaron a muchos estudiantes a malinterpretar el resultado de una simulación como un porcentaje de la gente en lugar de un porcentaje de las proporciones de la muestra.

En el estudio realizado por Vanhoof, Castro Sotos, Onghena, y Verschaffel (2007) investigan si el uso del software desarrollado por Chance, delMas y Garfield (2004) contribuye al aprendizaje de conceptos como la distribución muestral, con tres grupos de estudiantes universitarios (en total 221) , por medio de la generación de muestras de diferentes poblaciones con un tamaño dado (escogidos por el alumno) y representa en diferentes ventanas la última muestra obtenida y la distribución muestral, conforme se va generando. Destacan que el pos-test evidencia una mejora significativa de la comprensión de la distribución muestral en estos estudiantes, sin embargo, únicamente aproximadamente el 50% de los sujetos manifiestan una apropiada comprensión.

En el estudio realizado por Ko (2016) analiza la comprensión de la distribución muestral, con alumnos pertenecientes a dos grupos, el primero de 5° grado (10 a 11 años), con 65 estudiantes, de los cuales 34 presentan altas capacidades hacia las matemáticas y el segundo, conformado por estudiantes de 8° grado (13 a 14 años), con una muestra de 65 alumnos, 36 de los cuales tienen altas capacidades en matemática. En el 5° grado los alumnos estudiaron la construcción de tablas y gráficos (de barras, diagramas tallo hoja), como también a calcular estadísticos de tendencia central (media). En el de 8° grado los estudiantes aprendieron gráficos de sectores, tablas de frecuencias (relativas y acumuladas), histogramas y polígonos de frecuencias, y probabilidad. Los estudiantes responden un cuestionario compuesto por dos problemas, el primero involucra la variabilidad de la distribución muestral y en el segundo deben estipular cuál distribución de la media del muestreo (4 casos) corresponde a una muestra dada. Para analizar la información proporcionada por los estudiantes, se siguen dos pasos: (1) se establecen categorías de análisis a partir de los términos y expresiones expresadas en el lenguaje empleado en los argumentos que entregan los estudiantes y (2) se clasifican la respuestas empleando los niveles de análisis de la Taxonomía SOLO (Biggs y Collis, 1982) descritos anteriormente en esta memoria (ver página 48).

2.5. COMPRENSIÓN DEL ENFOQUE FRECUENCIAL DE LA PROBABILIDAD

Gómez, Batanero y Contreras (2014) señalan que:

“la investigación sobre la enseñanza-aprendizaje del significado frecuencial de la probabilidad aborda principalmente dos puntos: (a) la posibilidad de estimar una probabilidad teórica a partir de datos de frecuencias y (b) la comprensión de las características de las secuencias de resultados aleatorios y de la convergencia” (p. 211).

Ambas posibilidades están muy relacionadas con la idea de muestreo, por lo que las vamos a comentar a continuación.

2.5.1. ESTIMACIÓN DE UNA PROBABILIDAD A PARTIR DE LA FRECUENCIA

Green (1983) evalúa la comprensión de la estimación frecuencial de la probabilidad, con estudiantes ingleses (11 - 16 años), en cuyo cuestionario incorpora un ítem sobre el experimento de lanzar 100 chinchetas al aire. Realizando un pre-test de la pregunta, encuesta a 66 alumnos para ver qué resultado esperaban obtener, sin sugerir una experimentación previa; el 61% expresaron predilección por resultados equiprobables. Posteriormente, modificó el ítem, pidiendo a los estudiantes seleccionar los resultados posibles entre cinco alternativas, dando la información de las frecuencias obtenidas en un experimento previo (68 chinchetas con la punta hacia abajo). El 17% de la muestra de 2930 alumnos, proporcionan una correcta estimación de tipo frecuencial, la mayoría (64%) prefieren dar resultados equiprobables y los restantes consideran otros resultados. Posteriormente, Cañizares (1997) obteniendo resultados similares (15% correctas y 64% con sesgo de equiprobabilidad Lecoutre (1992), al administrar el mismo ítem a 253 niños españoles (11-14 años).

Valdéz (2016) realiza un estudio sobre el razonamiento probabilístico informal de un grupo de 30 alumnos de bachillerato (17 – 18 años), quienes ya habían recibido un curso formal sobre probabilidad, al responder a tres versiones de un cuestionario. El autor buscaba analizar si los estudiantes diferencian y ponen en relación los significados clásico y frecuencial de la probabilidad. Se recolectaron los datos a partir de tres cuestionarios y cuatro entrevistas. El diseño de los cuestionarios tiene como base los instrumentos empleados por Cañizares (1997), Metz (1998) y Stohl y Tarr (2002).

Las situaciones que componen cada cuestionario se plantean en un contexto familiar de urnas, en el que los cálculos son sencillos. Las entrevistas, tienen como propósito esclarecer y profundizar en los razonamientos que los estudiantes manifiestan de manera escrita, además de observar si dichos razonamientos se modifican o se mantienen al llevar a cabo la simulación computacional de algunas situaciones que conforman los cuestionarios. Se procuró entrevistar a quienes tuvieron un mejor desempeño en los cuestionarios. Las entrevistas fueron hechas por el investigador a cargo del estudio y se

llevaron a cabo fuera del horario de clase, con una duración de 90 minutos cada una, fueron video-grabadas y transcritas para su análisis.

El Cuestionario I está conformado por tres situaciones; cada una, a su vez, compuesta por tres preguntas. En la primera cuestión, a partir de una urna con 4 bolas se les pregunta la composición de una muestra de 1000 extracciones con reemplazamiento, el resultado en la extracción 1001 y cuál fue el primer resultado. En la segunda, a partir de dos urnas con bolas y la frecuencia de blancas y negras obtenidas 1000 extracciones de cada una se les pregunta cuál urna elegiría si quiere sacar una bola negra en la extracción 1001 y asignar una probabilidad a este suceso en cada urna. En la tercera, se da una secuencia aleatoria de 10 bolas blancas y negras en cada urna y se les pregunta cuál urna elegiría si quiere sacar una bola negra en la extracción 11 y asignar una probabilidad a este suceso en cada urna. Se pide argumentar todas las respuestas.

Los tres cuestionarios son semejantes pero se buscaba observar el impacto que tiene en las soluciones entregadas por los alumnos la modificación de ciertas variables de las tareas, con el objetivo de tener una imagen más amplia de sus razonamientos. De esta manera, en el Cuestionario I la composición de la urna es desconocida, pero se da el número total de bolas; en el Cuestionario II el contenido de las urnas es conocido. Por su parte, en el Cuestionario III sólo se consideran situaciones con la distribución del contenido de cada una de las urnas B y C es desconocido, así como su total.

En resumen, de acuerdo con las respuestas de los alumnos, se pueden distinguir los siguientes resultados de los análisis de los cuestionarios:

-) Aceptan la variabilidad en la muestra cuando la distribución de bolas en la urna es equiprobable, y la ignoran al hacer predicciones sobre el posible resultado de una extracción.
-) Reconocen el modelo equiprobable en la distribución de las bolas, pero consideran a la variabilidad en la muestra como excesiva si se aparta un poco de lo predicho por el modelo; es decir, que la mayor proporción de bolas negras en la muestra marca una tendencia sobre la composición de la urna, aunque las frecuencias observadas sean próximas a las esperadas.
-) Para estimar la probabilidad de un suceso asignan la frecuencia relativa, teniendo a las muestras como su único referente, aunque conozcan la composición de la urna (ausencia de un conocimiento informal de la ley de los grandes números).

- J) Eligen la urna que da mayor probabilidad de un resultado con base en los datos observados, pero excluyen la incertidumbre que subyace en el resultado de los ensayos. Además no logran dar un valor de la probabilidad en forma cuantitativa.
- J) No logran coordinar los modelos con las muestras respectivas, debido a que no aceptan como natural la variabilidad que se observa (Ley de los pequeños números).

Este autor, crea a partir de la respuesta de los estudiantes, una jerarquía de clasificación para dichas respuestas, con cuatro niveles del desarrollo del razonamiento probabilístico informal. En la Tabla 2. se muestra la clasificación por nivel de las características generales que se observaron en el estudio de Valdéz (2016), respecto a las ideas de Variabilidad, Aleatoriedad e Independencia.

En el Nivel 1, los alumnos no reconocen la aleatoriedad, la independencia, ni la variabilidad; piensan que es posible hacer predicciones determinísticas, buscan en los datos claves que les permitan predecir los resultados siguientes y piensan que las tendencias en muestras pequeñas son significativas, es decir, no interpretan apropiadamente la variabilidad. En este nivel, los alumnos no han comenzado a desarrollar un razonamiento probabilístico informal, no obstante conocen las definiciones de probabilidad clásica y frecuencial.

En el Nivel 2, los alumnos comienzan a mostrar rasgos incipientes que involucran las grandes ideas de probabilidad, pero sin mucha consistencia. Hacen predicciones deterministas, pero las matizan con expresiones que indican incertidumbre como “puede ser que...”, etc., expresando que ya distinguen que los resultados no son totalmente predecibles. No observan la independencia, pues buscan indicios en los datos para predecir lo que ocurrirá en subsiguientes ensayos, ignorando el modelo; por lo mismo, utilizan el enfoque frecuencial para asignar probabilidades. Finalmente, piensan que las diferencias pequeñas en muestras grandes son significativas; esto se debe a que, generalmente, razonan con frecuencias absolutas y no con frecuencias relativas.

En el Nivel 3, los alumnos comienzan a mostrar un razonamiento probabilístico informal de manera más sistemática; reconocen que no se pueden hacer predicciones deterministas debido a la naturaleza aleatoria del experimento; la asignación de probabilidades es con base en el modelo, ignorando los resultados anteriores tanto en el caso de muestras pequeñas como grandes; y estiman intuitivamente, sin criterios

numéricos, que siempre hay pequeñas diferencias en las frecuencias de los eventos en las muestras grandes.

En el Nivel 4, los alumnos manifiestan un razonamiento probabilístico informal consolidado, en él ubicaríamos a los estudiantes que además de mostrar los rasgos del nivel 3, les asocian los términos aleatoriedad, independencia y variabilidad de manera adecuada. Reconocen que no pueden predecir un resultado porque el experimento es aleatorio, ignoran alguna tendencia en la muestra porque son ensayos independientes y reconocen la idea de que la variabilidad de las frecuencias es grande en muestras pequeñas y que dicha variabilidad disminuye cuando la muestra crece.

Tabla 2.3. Niveles de respuesta con relación a las grandes ideas de probabilidad, según Valdéz (2016)

	Nivel 1	Nivel 2	Nivel 3	Nivel 4
Variabilidad	Esperan resultados sin variabilidad y/o en muestras pequeñas consideran las diferencias entre las frecuencias y las esperadas como significativas	Creer que pequeñas diferencias son significativas en muestras grandes	Estiman que diferencias relativamente pequeñas no son significativas cuando la muestra es grande	Reconocen que la variabilidad es grande cuando la muestra es pequeña, pero es poca cuando la muestra es grande (Observan el papel del tamaño de la muestra)
Aleatoriedad	Hacen una predicción determinista	Hacen una predicción determinista matizándola con lenguaje probabilístico	Reconocen que no se puede predecir el resultado con exactitud	Reconocen que no se puede predecir con certeza el resultado de un ensayo, pero sí la estabilidad de las frecuencias alrededor de un valor (probabilidad) en el largo plazo
Independencia	Asignan probabilidades con base en las frecuencias relativas, pero introducen elementos no pertinentes	Asignan Probabilidades con base en las frecuencias relativas, ignorando el modelo	Asignan probabilidades con base en el modelo e ignoran los resultados previos (muestra)	Reconocen la independencia, por tanto, ignoran lo ocurrido y asignan probabilidades con base en el modelo

2.5.2. CARACTERÍSTICAS DE LAS SECUENCIAS ALEATORIAS DE RESULTADO

En el enfoque frecuencial, se espera que los estudiantes realicen experiencias y registren datos de secuencias de resultados aleatorios. Por ello es necesario comprender las características de dichas secuencias, en las que por un lado se debe mantener aproximadamente una proporción de cada resultado cercana a la probabilidad teórica y por otro tener la suficiente variabilidad. En este punto se parece al muestreo, pero a

diferencia de las investigaciones centradas en el muestreo, en las que incluimos en este apartado también se estudian aspectos como la independencia de resultados o las rachas de un mismo resultado.

Las investigaciones sobre este punto utilizan dos tipos de tarea: en una de ellas (*tareas de generación*) se pide a los estudiantes escribir una secuencia que parezca aleatoria para otra persona y analizando esta secuencia se estudian las ideas de la persona sobre la aleatoriedad.

Este tipo de tareas fue indagado por Green (1991), quien pidió a un grupo de 305 alumnos (7 a 11 años), crear un listado de 50 resultados que permitieran simular los lanzamientos sucesivos de una moneda. Los alumnos manifiestan una clara falta de comprensión de la independencia, ya que anotan un número igual de caras y cruces, además de producir rachas demasiado cortas. Los estudiantes producen secuencias con dispersión más pequeña de lo esperado para una secuencia realmente aleatoria, es decir, sus respuestas se inclinan a subestimar la variabilidad en las secuencias generadas.

Un segundo tipo de tarea consiste en dar varias secuencias de resultados para preguntar al sujeto cual de ellas es aleatoria (*tareas de reconocimiento*). Una investigación de este tipo la efectúa Serrano (1996), quien preguntó a 277 sujetos (13 - 17 años), cómo reconocer algunas secuencias aleatorias entre distintas opciones. Les pidió razonar su elección. Analizando las razones dadas, detectó tres puntos importantes en la comprensión de la aleatoriedad de las secuencias:

(1) *variabilidad local*: se esperan frecuentes alternancias entre distintos sucesos, (2) *regularidad global*: se espera convergencia de la frecuencia relativa a la probabilidad teórica y (3) *grado de discrepancia entre la distribución esperada y la observada*: se espera un ajuste casi exacto (Gómez, Batanero y Contreras, 2014, p.48).

Lo cual no será correcto en general, sobre todo si la longitud de la secuencia es pequeña.

Dentro de los aspectos poco comprendidos identificó los siguientes:

(a) la estimación frecuencial de una probabilidad no sirve para predecir la ocurrencia del resultado en un único experimento; (b) uso de razones frecuenciales para explicar la ocurrencia de un suceso poco probable; y (c) estimación de la frecuencia relativa de un suceso en una serie futura de experimentos cuando se conoce su probabilidad (Gómez, Batanero y Contreras, 2014, p.48).

2.6. EL MUESTREO EN EL RAZANAMIENTO INFORMAL DE LOS ESTUDIANTES

Entre los diversos modelos utilizados dentro de esta línea, citaremos el de Zieffler, Garfield, delMas y Reading (2008) y Makar y Rubin (2009), estos autores plantean cuatro categorías para la caracterización de del razonamiento inferencial informal:

(1) Obtener una conclusión que vaya más allá de los datos, (2) Usar los datos como evidencia, (3) Usar el lenguaje probabilístico para expresar la incertidumbre acerca de la conclusión y (4) Utilizar e integrar el conocimiento estadístico disponible (Begué, Batanero, Ruiz y Gea, 2019, p.63).

La investigación realizada por García-Ríos (2013), analiza los argumentos proporcionados por 16 estudiantes de bachillerato (15-17 años) a un problema que involucra contrastes de hipótesis, que tiene por objetivo indagar sobre los niveles de razonamiento de inferencia estadística informal. En concreto, las cuestiones planteadas demandan al estudiante determinar si un tratamiento es efectivo, para lo cual se les proporciona los datos de una muestra. El análisis de las respuestas se realiza en torno a las cuatro componentes descritas por Zieffler, Garfield, delMas y Reading (2008), para las que se identifican una serie de niveles.

Uno de los principales hallazgos que menciona García-Ríos (2013) es “que las respuestas de los estudiantes se basan más en sus creencias y en sus conocimientos personales acerca del contexto y no en los datos del problema” (p.343). Además, una gran cantidad de estudiantes emplean en sus argumentos los datos de la muestra, pero no son capaces de elaborar una conclusión adecuada. También, el autor detectó dificultades en el uso del lenguaje probabilístico y en la identificación de un modelo de referencia.

Meletiou-Mavrotheris y Papparistodemou, (2015), analizan el razonamiento inferencial informal de 69 alumnos de 4° a 6° grado (9 a 11 años) sobre el conocimiento previo de conceptos tales como: la muestra, su tamaño, el método de selección de muestras y el sesgo. Un 85% de los estudiantes habían oído la palabra muestra en contextos no vinculados a la escuela (tienda, farmacia, etc.), sin embargo, casi la mitad de ellos (46%), no proporcionaron explicaciones acerca de dicho concepto en la tarea inicial o dieron alguna respuesta incompleta. Casi el 40%, proporcionaron definiciones que reflejaban la idea de parte, sin el conjunto asociado que caracteriza la relación de muestra - población, o cómo una prueba de algo; sólo el 12% dieron respuestas más elaboradas y reconocieron la relación parte-todo entre la muestra y la población, es decir,

reconocen la validez de elegir una muestra en lugar de consultar a todos los sujetos de la población.

2.7. CONCLUSIONES DEL ESTUDIO DE LOS ANTECEDENTES

Se aprecia en el análisis de las investigaciones previas relacionadas con el concepto de muestreo que los estudiantes manifiestan una comprensión insuficiente de este concepto, debido a “que algunos estudiantes conciben las muestras como subconjuntos excluyentes de la misma población y no como subconjuntos que pueden contener elementos comunes” (Begué, Batanero, Ruiz y Gea, p.70). Esto se debe principalmente, porque no distinguen la relación parte-todo que es esencial para identificar la muestra de la población, como se evidencia en estudios como los de Watson (2004) y Meletiou-Mavrotheris y Papparistodemou (2015), entre otros.

Por otro lado, algunas investigaciones (por ejemplo, García-Ríos, 2013) señalan que el rol del contexto también es un factor que influye en la comprensión del concepto de muestra, debido a que algunos estudiantes no lo consideran importante y basan los argumentos en creencias personales y subjetivas de lo que entienden por muestra.

Vanhoof et al. (2007) y Serrano (1993) consideran que no es posible determinar si el uso de simulaciones como apoyo a trabajo con distribuciones muestrales influya en las mejoras de los aprendizajes de los sujetos que participaron en su estudio, debido a que estos sujetos constaban con instrucción previa en estadística o basado en sus conocimientos a partir de su experiencia.

Otro aspecto de interés para los docentes de matemática es la caracterización de las heurísticas y sesgos mencionados, que independientemente del nivel educativo, son observados en los estudios previos descritos en este capítulo.

Por otro lado, algunas de las situaciones propuestas en estos antecedentes se tendrán en cuenta para elaborar el banco inicial de ítems para conformar el cuestionario de evaluación dirigido a estudiantes chilenos de educación secundaria, previa adaptación y/o traducción considerando los niveles educativos a los que se dirigirá el estudio.

Esta síntesis de investigaciones previas presenta una visión general de cómo se ha evaluado la comprensión del muestreo en diferentes niveles educativos, con diversos métodos de investigación, además, permite caracterizar las principales dificultades que

presentan los estudiantes al enfrentarse a los temas de estadística inferencial, en concreto los que involucran al muestreo, como también, sugieren distintas actividades que permiten diseñar y planificar secuencias didácticas que permitan contribuir a la mejora de los aprendizajes de los estudiantes.

CAPITULO 3:

MUESTREO EN LOS LIBROS DE TEXTO DE EDUCACIÓN SECUNDARIA

- 3.1 Introducción
- 3.2 Objetivos del estudio de los libros de texto
- 3.3 Metodología de análisis
 - 3.3.1 Muestra de textos analizados
 - 3.3.2 Método y variables de análisis
- 3.4 Situaciones problema
- 3.5 Lenguaje matemático
 - 3.5.1 Lenguaje verbal
 - 3.5.2 Lenguaje numérico
 - 3.5.3 Lenguaje simbólico
 - 3.5.4 Lenguaje tabular
 - 3.5.5 Lenguaje gráfico
 - 3.5.6 Lenguaje icónico
- 3.6 Conceptos
- 3.7 Propiedades
- 3.8 Procedimientos
- 3.9 Argumentos
- 3.10 Conclusiones sobre el estudio de libros de texto

3.1.INTRODUCCIÓN

En esta sección se lleva a cabo el análisis de una colección de textos escolares de educación secundaria chilenos, centrandó la atención en la identificación de los diferentes objetos matemáticos ligados al muestreo y la forma en que cada libro los contempla. Este estudio servirá para señalar semejanzas y diferencias en la forma que se abordan el muestreo en los textos y las directrices curriculares presentadas en el Capítulo 1.

Los libros de texto constituyen un material didáctico de especial relevancia para los estudiantes, siendo un soporte que vehicula el conocimiento académico que se pretende transmitir en la escuela (Escolano, 2009). Constituye, según Herbel (2007), un paso intermedio y necesario entre el currículo planteado en los lineamientos curriculares y el puesto en práctica en la sala de clases. Ferreira y Mayorga (2010) resaltan su importancia,

al concebir el texto escolar como un recurso que ayuda al alumno en su proceso de aprendizaje y es una pauta para el proceso de enseñanza docente. Cordero y Flores (2007), por su parte, indican que la acción en el aula el discurso matemático escolar es establecido frecuentemente por el texto escolar, que es un importante factor para determinar la enseñanza.

Todas estas consideraciones, nos han llevado a fundamentar la construcción de nuestro cuestionario con un análisis previo de los textos escolares empleados por los estudiantes a los que va dirigido. La finalidad principal del estudio es de precisar mejor el significado institucional pretendido del muestreo en nuestro trabajo, teniendo en cuenta la forma en que este tema se presenta en los libros y completando así el estudio curricular citado. Este significado institucional será la base para elegir las tareas que formarán parte de nuestro estudio de evaluación.

En lo que sigue, se describen los objetivos perseguidos, la muestra de textos utilizados, la metodología del análisis y los resultados en cada tipo de objetos matemáticos descritos en el marco teórico. Se finaliza el capítulo con algunas conclusiones sobre los principales hallazgos de este análisis.

3.2. OBJETIVOS DEL ESTUDIO DE LOS LIBROS DE TEXTO

Como se ha mencionado anteriormente, el principal objetivo de este estudio es diagnosticar qué dificultades tienen los alumnos chilenos de educación secundaria en la comprensión del muestreo. Los resultados del estudio de evaluación se podrán interpretar mejor si lo relacionamos con la enseñanza recibida. Por ello, queremos que el cuestionario tenga en cuenta las tareas que han completado durante la enseñanza, el lenguaje utilizado, los conceptos y propiedades que han estudiado, los procedimientos y tipos de argumentos presentados.

El objetivo principal es, por tanto, completar la definición del significado institucional del muestreo presentado a los estudiantes, que posteriormente se comparará con los significados personales, presentados por medio de las prácticas matemáticas utilizadas por los alumnos en el cuestionario. Este objetivo se puede descomponer en los siguientes objetivos parciales:

1. Identificar las situaciones problemas relativas al muestreo que se enseñan en los textos escolares utilizados por los estudiantes de la muestra.

2. Determinar los diferentes tipos de lenguajes matemáticos presentados en los textos escolares usados por los estudiantes de la muestra.
3. Establecer los distintos conceptos relacionados con el muestreo, expuestos en los textos escolares analizados.
4. Identificar las principales propiedades que sobre los anteriores conceptos aparecen en la muestra de textos escolares analizados.
5. Determinar los procedimientos que se asocian al muestreo en la serie de textos estudiados.
6. Establecer los tipos de argumentos expuestos en la serie de libros de texto estudiada y que sirven para introducir los conceptos o propiedades o bien para describir y ejemplificar los procedimientos.

3.3. METODOLOGÍA DE ANÁLISIS

El análisis de los textos escolares, forma parte de la investigación cualitativa pues, se centra en el estudio sistemático de documentos y se interesa por variables cualitativas (Kirk y Miller, 1986). Entre dichas variables podemos introducir las consideradas en nuestro análisis, que se refieren a la forma en que los diferentes objetos matemáticos que iremos analizando se presentan o no en los textos.

Según Bisquerra (2004) el procedimiento seguido en el análisis de los textos escolares es de tipo inductivo, ya que se comienza con la exploración de casos particulares, es decir, de unos pocos libros de texto, y el objetivo es producir generalizaciones (entendidas en el sentido en que se pueden concebir en una investigación cualitativa) desde observaciones sistemáticas del entorno. Además, es de tipo descriptiva, pues no se cambia ninguna variable, pues se restringe a examinar y especificar los fenómenos observados, también es una investigación aplicada, porque está enfocada a lograr criterios para mejorar el progreso curricular.

Siguiendo a López (2002), realizamos un análisis interno de los textos, para destacar sus características fundamentales; se trata de un análisis intensivo, el cual examina con minuciosidad ciertos documentos, en vez de analizar con menor especificación una muestra más amplia. Para establecer si están o no presentes los diferentes objetos

matemáticos se usan el análisis de contenido y el análisis semiótico, los que se describen más adelante.

3.3.1. MUESTRA DE TEXTOS ANALIZADOS

Para seleccionar los libros de texto se utilizó un muestreo intencional, basado en una selección controlada y de acuerdo a determinadas características que queremos tenga nuestra muestra (Hernández, Fernández y Baptista, 2014). En nuestro caso son libros de texto que estuvieron vigentes en el periodo académico 2016-2018, que es cuando se ha realizado el análisis y han sido editados de acuerdo a las directrices curriculares vigentes en este periodo en Chile. Aunque la edición puede cambiar en el curso que se tomen los datos, se han tomado los que usaron los estudiantes de la muestra el curso anterior a la muestra de datos, con lo que si acaso, variará la edición usada en el último nivel, pero no en los anteriores.

Para nuestro trabajo se optó por elegir una muestra reducida de libros de texto, por tanto, es una investigación de corte cualitativo y exploratorio, pues permite desarrollar un análisis más profundo y teniendo en cuenta también, que el análisis de los textos escolares no es el objetivo principal de la tesis, sino que se usa para fundamentar la construcción del cuestionario, como también aportar mayor información para establecer los significados institucionales que complementan el análisis curricular desarrollado en el Capítulo 1. Los libros de texto a analizar fueron los disponibles en el portal web de los textos escolares chilenos http://catalogotextos.mineduc.cl/catalogo_textos-web/mvc/login/login?tipo=docente, en donde se encuentran disponibles para los docentes de primaria y secundaria los textos escolares vigentes para las distintas asignaturas del sistema escolar chileno.

Se analizó la serie completa, es decir, un libro de cada nivel educativo, correspondiente a los cursos de 7° Básico a 4° año medio. En el transcurso del desarrollo de esta memoria, considerando los ajustes curriculares incorporados paulatinamente desde el año 2016, se han incluido en nuestro análisis final, con las modificaciones efectuadas en los niveles de 1° y 2° medio.

En la Tabla 3.1 listamos los textos utilizados, a los que se ha dado un código que será empleado durante este Capítulo para referenciar al texto respectivo. Como es un análisis cualitativo, la muestra es intencional; se ha elegido de forma conveniente para los

finde de la investigación. Se ha examinado una muestra que abarque los casos más típicos y únicos, para captar distintos aspectos que caractericen el contexto planteado en los textos escolares (Martínez, 2006).

Tabla 3.1. Muestra de libros de texto analizada

Curso	Libro	Código
7° Básico	Merino, R., Muñoz, V., Pérez, B. y Rupin, P. (2015). Matemática 7° Básico. Santiago de Chile: S. M.	L1
8° Básico	Catalán. D., Pérez, B., Prieto, C. y Rupin, P. (2015). Matemática 8° Básico. Santiago de Chile: S. M.	L2
1° Medio	Del Valle, J., Muñoz, G. y Santís, M. A. (2013). Matemática 1° Medio. Santiago de Chile: S. M.	L3
2° Medio	Muñoz, G. y Rupin, P. y Jiménez, L. (2013). Matemáticas 2° Medio Santiago de Chile: S.M.	L4
3° Medio	Saiz, O. y Blumenthal, V. (2018). Matemáticas 3° Medio. Santiago de Chile: Cal y Canto	L5
4° Medio	Muñoz, G., Gutiérrez, V. y Muñoz, S. (2013). Matemáticas 4° Medio. Santiago de Chile: Santillana.	L6
1° Medio	Galasso, B., Maldonado, L. y Marambio, V. (2016). Matemáticas 1° Medio. Santiago de Chile: Santillana.	L7
2° Medio	Chacón, A., García, G., Rupin, P., Setz, J y Villena, M. (2017). Matemáticas 2° Medio. Santiago de Chile: S.M.	L8

3.3.2. MÉTODO Y VARIABLES DE ANÁLISIS

El análisis se basa principalmente en un análisis de contenido, que parte de la hipótesis de que cualquier texto puede dividirse en unidades de análisis y estas unidades permiten catalogarse en un número pequeño de categorías. Esta división se hace en función de variables que fija el investigador, y su contenido permite elaborar conclusiones (Krippendorff, 1997).

El análisis de contenido es una técnica de estudio documental que suple eventualmente las apreciaciones personales del análisis de documentos escritos utilizando métodos estandarizados, cuyo propósito es transformar los contenidos analizados en los documentos en datos (León y Montero, 2002). Es una metodología cualitativa específica para estudiar estos documentos escritos en forma sistemática y exhaustiva con la finalidad de producir generalizaciones (Zapico, 2007).

El método se ha utilizado para el análisis de textos escolares en diversas investigaciones sobre educación estadística; por ejemplo, sobre probabilidad en educación secundaria (Ortiz, 2002); medidas de posición central (Mayén, 2009) y en probabilidad en educación primaria (Gómez, 2014) y en gráficos estadísticos (Díaz-Levicoy, 2018). Se siguen los pasos sugeridos por los anteriores autores, que se indican a continuación:

1. Establecer un sistema de variables y categorías de estudio para desarrollar la codificación de los datos. Para ello se elaboró de una lista de objetos matemáticos (situaciones problema, elementos lingüísticos, conceptos, propiedades, procedimientos y argumentos) a partir del análisis curricular realizado en el Capítulo 1. El estudio de los textos trata de comprobar cuáles de estos objetos se presentan en los libros o si se añaden otros diferentes. Además, se pretende mostrar la forma particular en que se presentan.
2. Seleccionar las unidades del texto que interesen para el estudio. Para ello se identificaron las páginas o los capítulos de textos escolares donde se incluyen contenidos de muestreo, que son nuestros datos primarios.
3. Dividir el texto en secciones independientes (párrafos, ejemplos, ejercicios) que se usan como unidades de análisis. Con ello los datos se transforman en unidades de información que pueden ser analizadas por separado.
4. Establecer si están presentes los diferentes objetos matemáticos hallados en el significado institucional pretendido identificado en el Capítulo 1, cotejando el contenido de estas páginas con la lista elaborada en el paso 1. Al aparecer algún objeto matemático nuevo se incorpora en la lista. Codificación de los datos siguiendo un proceso inductivo y cíclico donde se revisa los datos obtenidos para darles sentido.
5. Seleccionar y analizar ejemplos para caracterizar los objetos matemáticos hallados en los textos escolares.
6. Elaborar resúmenes (tablas) de los distintos objetos matemáticos identificados en cada texto escolar, que proporcionen conclusiones sobre el significado institucional pretendido de las series analizadas.

3.4.SITUACIONES PROBLEMA

Se comienza con la delimitación del campo de problemas relativo al concepto de muestreo, identificando las situaciones problemas que involucran dicho concepto. Las situaciones problemas tienen gran importancia en el enfoque ontosemiótico (Godino, 2002; Godino y Batanero, 1994; Godino, Batanero y Font, 2007) pues este enfoque sugiere que los objetos matemáticos surgen al resolver problemas, como entidades culturales de las prácticas, personales y/o institucionales. Los problemas matemáticos y sus soluciones pueden ser particulares de una persona (prácticas personales), pero como indica Godino (2018) también pueden ser “compartidos en el seno de instituciones o colectivos específicos implicados en el estudio de ciertas clases problemas” (prácticas institucionales) (p.30). Una de ellas es la institución docente y el análisis de los textos escolares permite identificar los problemas considerados relevantes en dicha institución.

La clasificación teórica de los principales campos de problemas que otorgan sentido al muestreo en los textos escolares chilenos, se fundamenta en el estudio previo de los documentos curriculares, donde se identificaron los siguientes campos de problemas:

) *Estimar algunas características de una población desconocida por medio del muestreo.*

Con frecuencia se desea estudiar un fenómeno aleatorio que viene caracterizado por una distribución de probabilidad, que depende de uno o varios parámetros, por ejemplo, la media. Pero no es posible recolectar los datos de toda la población, sino que hemos de contentarnos con una muestra aleatoria de la misma en donde se calcula un estadístico que sirve para estimar el parámetro. (Olivo, 2008, p.28)

En los textos escolares analizados, se presentan algunas actividades que solicitan estimar características de la población, por ejemplo, la media o la proporción y las encontramos en los cursos 7° y 2° medio. En la Figura 3.1, aparece una situación problema que pretende que los estudiantes estimen el porcentaje de peces de cada especie de acuerdo a los resultados planteados en una tabla de frecuencias que reúne la extracción de diez muestras. Se espera que los estudiantes calculen el porcentaje global en el conjunto de muestras y usen este valor como estimador del porcentaje en la población.

Aunque en el curso 1° medio no aparecen explícitamente problemas de este tipo, si se consideran implícitamente pues se introduce la probabilidad en sentido frecuencial, que es estimada a partir de la frecuencia relativa en muestras de resultados. Por tanto, en esta introducción se estima la probabilidad (que es un parámetro) a través de la frecuencia relativa (que es un estadístico en la muestra).

3. Una muestra de 50 peces se repitieron 10 veces y en cada oportunidad reposición los peces al lago. A continuación se presentan los resultados.

	1º	2º	3º	4º	5º	6º	7º	8º	9º	10º
Especie A	35	28	22	31	30	32	34	27	29	33
Especie B	15	22	28	19	20	18	16	23	21	17

¿Podrían estimar el porcentaje de peces de cada especie en el lago? Expliquen su procedimiento.

A repetir un experimento muchas veces, los porcentajes se estabilizan, es decir se observa que tienden a acercarse a un valor determinado.

Figura 3.1. Estimar una característica de la población (L1, 2015, p. 289)

) *Determinar el número de muestras de un tamaño dado, que se pueden extraer desde una población de tamaño finito, con o sin reemplazo.* Esta actividad consiste, bien en listar todas las posibles muestras con unas características dadas, o en deducir su número sin listarlas. Se puede incluir bien como repaso del tema de muestreo o como parte del estudio de la combinatoria, ya que las operaciones combinatorias se definen en estos libros de textos a partir de muestras con y sin reemplazamiento (tomadas de una colección o población de objetos).

Para la pregunta 2 (ver Figura 3.2), que es un ejemplo de este campo de problemas, puede ser necesario aclarar si las muestras son con o sin reemplazamiento y repasar los conceptos combinatorios (combinaciones o variaciones). También se debe sugerir a los estudiantes en que, para tamaños pequeños de la población, primero calculen la cantidad de muestras y luego las determinen, para poder verificar que efectivamente han calculado el número correcto.

2. Calcula la cantidad de muestras de tamaño m que se pueden extraer de una población de tamaño p , con los datos dados.

a. $p = 6; m = 2$	e. $p = 18; m = 15$
b. $p = 6; m = 5$	f. $p = 20; m = 7$
c. $p = 12; m = 3$	g. $p = 24; m = 11$
d. $p = 15; m = 10$	h. $p = 25; m = 21$

Figura 3.2. Cálculo del número de muestras de una población finita (L4, 2013, p. 275)

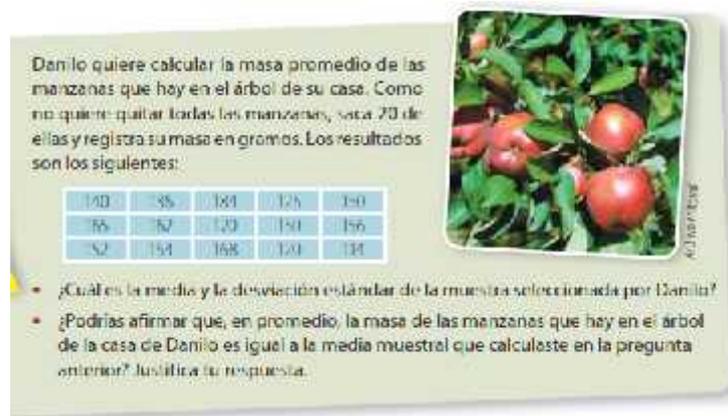
) *Relacionar la media de una población y la media aritmética de las muestras.* Este tipo de problemas puede tomar dos formas posibles: o bien se da la media de la población y se pide deducir las medias de algunas muestras o bien el caso contrario. Así, en la Figura 3.3 se pide comparar los promedios de varias muestras, estimar a partir de ellos la media de la población y finalmente sacar una conclusión. Se espera que los estudiantes aprecien que las medias de las muestras se parecen a la de la

población, pero varían entre sí, es decir apliquen las propiedades de representatividad y variabilidad muestral en este caso.

6. Extrae 5 muestras aleatorias de las alturas de tus compañeros de curso y calcula el promedio de cada una de ellas. Compara los resultados. ¿Son parecidos entre sí? Calculen el promedio de estatura del curso, y la media aritmética de los promedios de las muestras anteriores. ¿Qué puedes concluir?

Figura 3.3. Relacionar la media aritmética de una población y media aritmética de las medias de muestras (L4, 2013, p. 279)

) *Relacionar la proporción de una población y la proporción de las muestras.* Como en el caso, anterior, este campo de problemas puede tomar dos formas posibles: o bien se da la proporción de la población y se pide deducir las proporciones de algunas muestras o bien el caso contrario. Asimismo, en la Figura 3.4 se pide, en primer lugar, calcular la media y la desviación estándar de la muestra que se proporciona en los datos del enunciado, posteriormente, se solicita a los estudiantes comparar el promedio de la población con la media muestral. Se espera que los estudiantes aprecien que las proporciones de las muestras se parecen a la de la población, pero varían entre sí.



Dani lo quiere calcular la masa promedio de las manzanas que hay en el árbol de su casa. Como no quiere quitar todas las manzanas, saca 20 de ellas y registra su masa en gramos. Los resultados son los siguientes:

120	85	184	125	150
75	82	120	150	156
82	154	168	120	134

- ¿Cuál es la media y la desviación estándar de la muestra seleccionada por Dani lo?
- ¿Podrías afirmar que, en promedio, la masa de las manzanas que hay en el árbol de la casa de Dani lo es igual a la media muestral que calculaste en la pregunta anterior? Justifica tu respuesta.

Figura 3.4. Relacionar la proporción de la población y proporción de una muestra. (L6, 2013, p. 314)

) *Identificar las muestras aleatorias y no aleatorias.* En este campo de problemas se describen diferentes métodos de muestreo y se pide al estudiante decidir si el método de selección es o no aleatorio. La finalidad es que reconozcan en qué condiciones una muestra puede ser utilizada para realizar una inferencia sobre una población. Un ejemplo se presenta en la Figura 3.5, donde ninguno de los métodos es un muestreo aleatorio simple.

- 1** Identifica cuáles de los siguientes casos corresponde a un muestreo aleatorio simple:
- Eligir a tres deportistas de cada curso para generar un estudio acerca de la frecuencia con que hacen deporte los alumnos del liceo.
 - Escoger por votación de todos los alumnos del curso al mejor compañero.
 - Escoger a 20 de los 400 trabajadores de una empresa para responder una encuesta sobre el clima laboral, de modo que el apellido de cada trabajador comience con una letra diferente.

Figura 3.5. Reconocer muestras no aleatorias (L4, 2013, p. 290)

) *Elaborar modelos para el muestreo aleatorio.* Este problema es en cierto sentido inverso al anterior, pues en lugar de pedir reconocer si una muestra es aleatoria, se trata de pensar o construir un método para genera una muestra aleatoria. Generalmente se pide al estudiante elaborar con Excel, un simulador aleatorio de un cierto experimento de muestreo que se describe en la actividad. Por ejemplo, en la Figura 3.6 se explica cómo simular con Excel el lanzamiento de tres monedas y a continuación se pide generar 100 lanzamientos aleatorios de las cuatro monedas.

3. Mediante una planilla de cálculo realiza una simulación del lanzamiento de tres monedas para calcular la probabilidad experimental o empírica de obtener exactamente 3 caras.

Paso 1: Escribe en las celdas A1, B1 y C1, moneda 1, moneda 2 y moneda 3 respectivamente.

Paso 2: Genera la simulación del lanzamiento de una moneda escribiendo en la celda A2 la función =ALEATORIO.ENTRE(0;1), esta función generará de manera aleatoria, el número 0 o el número 1. Considera como 0 = sello y 1 = cara. Luego, copia esta función en las celdas B2 y C2.

Paso 3: Escribe en la celda D2 la función =SUMA(A2:C2), la cual permitirá sumar la cantidad de ocasiones en que la suma de las caras de las monedas es 3.

Paso 4: Escribe en la celda E3 la función =CONTAR.SI(D:D;3), la cual permitirá contar los lanzamientos en que aparecen en las tres caras o los tres valores de 1, cuya suma sea 3.

Paso 5: Genera 100 lanzamientos de las monedas, copiando las celdas A2, B2, C2 y D2 hasta las celdas A102, B102, C102 y D102. De esta manera obtendrás la simulación para 100 lanzamientos.

Figura 3.6. Modelos de muestreo aleatorio (L3, 2013, p. 300)

Este tipo de actividad, permite de forma fácil y rápida efectuar experimentos para enseñar a los alumnos como la frecuencia relativa, se aproxima a la probabilidad teórica, aplicando la regla de Laplace, proponiendo así nuevo método para calcular la probabilidad de un evento la utilización por medio de la frecuencia relativa.

) *Valorar la representatividad de una muestra.* Se enuncia una situación de muestreo y se pide valorar la representatividad de la muestra a partir de la descripción del método de muestreo. Un ejemplo se incluye en la Figura 3.7 que pide comparar un muestreo aleatorio simple con un muestreo estratificado (que podría también ser aleatorio). En este tipo de situaciones problemas, por lo general, se trabaja una interpretación

cualitativa de la representatividad, sin llegar a formalizar mediante alguna fórmula matemática.

A partir del enunciado de la situación problema los estudiantes deben analizar dos afirmaciones planteadas para la selección de la muestra para una encuesta y posteriormente comentar una serie de preguntas que apuntan a evaluar la representatividad de dicha muestra.

Figura 3.7. Representatividad de la muestra (L1, 2015, p. 288)

Encontramos en la Figura 3.8, otra actividad en la que se les solicita a los estudiantes evaluar la representatividad de una muestra a partir de los gráficos circulares que representan una cierta característica en la población y la muestra respectivamente. No se indica el método de muestreo, ni el tamaño de la muestra; teniendo en cuenta la variabilidad de la proporción muestral, es posible que el muestreo sea aleatorio y los resultados solo reflejen la variabilidad de la distribución muestral de la proporción.



Figura 3.8. Analizar la representatividad de la muestra (L1, 2015, p. 290)

Con este tipo de situaciones que abordan el concepto de representatividad de la muestra, por lo general, los estudiantes suelen pensar que entre mayor sea el número de

la muestra, mayor será su representatividad. Se debe hacer hincapié entonces, en que la representatividad está directamente ligada con que los resultados de la muestra se asemejen a los de la población sin olvidar la variabilidad razonable debida al tamaño de muestra.

) *Comparar dos o más conjuntos de datos a través de las medidas de tendencia central.*
 En la situación planteada en la Figura 3.10, se les pide a los estudiantes comparar la distribución de las edades de dos grupos que participan en un taller de baile, para ellos, los alumnos disponen de la información de la situación problema como de los datos resumidos en una tabla, y deben emplear las medidas de tendencia central para responder a la pregunta planteada y así obtener conclusiones de las dos muestras presentadas.

El departamento de deportes de un municipio inició las inscripciones para el taller de zumba. Cada inscrito tiene derecho a asistir a una clase semanal, siendo estas los lunes (grupo 1) o los miércoles (grupo 2). Para conocer el impacto de este taller, la profesora registró las edades (años) de los asistentes durante la primera semana.

¿Cómo es la distribución de las edades en cada uno de los grupos?

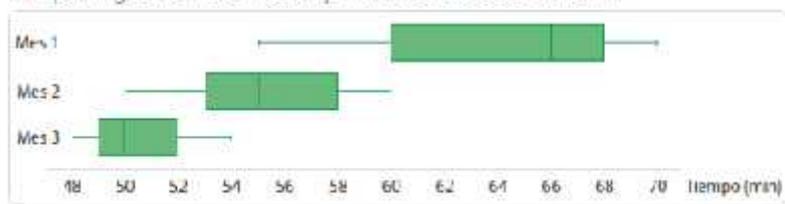
Día	Asistentes	Edad promedio	Mediana	Edad mínima	Edad máxima	Rango
Lunes	11	25	16	14	45	31
Miércoles	11	25	27	21	28	7

Figura 3.9. Comparar muestras usando medidas de tendencia central. (L1, 2015, p. 324)

) *Comparar poblaciones y muestras usando medidas de posición*

En la Figura 3.10, vemos una situación problema en la que describe tres muestras de datos obtenidas del entrenamiento mensual realizado por un ciclista, las cuales se expresan mediante de diagramas de caja, en donde se pueden observar los valores máximos y mínimos, como también los percentiles 25, 50 y 75. Se les consulta a los estudiantes establecer algunas conclusiones a partir del rendimiento mostrado por el ciclista. Este es un problema de aplicación que involucra las medidas de posición y los conceptos de muestras estudiadas anteriormente, en donde los estudiantes deben usar sus conocimientos para traspasar la información del gráfico a números y comparar dichas cantidades para realizar inferencias de la información entregada en la situación problema.

Un ciclista entrena habitualmente como una forma de recrearse y para mantener un buen estado de salud. Él va de su casa a su trabajo en bicicleta y ha ido registrando lo que demora para constatar el progreso de su rendimiento. A partir de estos registros, construyó los siguientes diagramas de caja para representar los tiempos registrados durante los primeros 3 meses de su rutina.



Situación Comparando diagramas de caja
 ¿Qué conclusiones pueden extraerse de la evolución del rendimiento del ciclista a partir de los tres diagramas de caja?

Figura 3.10. Comparar muestras usando medidas de posición. (L2, 2015, p. 326)

) Registrar distribuciones de dos características distintas de una misma población en una nube de puntos. Una nube de puntos corresponde a la gráfica de un conjunto de pares ordenados en el plano cartesiano, donde las coordenadas de cada punto corresponden a una variable cuantitativa en estudio (L7, 2016, p. 228). En la figura 3.11, se les solicita a los estudiantes representar en una nube de puntos, los datos entregados en la tabla, para ello deben distinguir qué características serán las que se convierten en el par ordenado correspondiente que deben registrar en el plano cartesiano y así generar la gráfica solicitada.

La siguiente tabla corresponde a datos obtenidos mediante una encuesta que se realizó a 10 personas sobre su masa corporal y edad.

Persona	1	2	3	4	5	6	7	8	9	10
Edad	10	13	15	14	11	17	19	15	17	11
Masa corporal	28	34	43	40	31	52	63	42	53	31

Representa los datos de la tabla en una nube de puntos.

Figura 3.11. Distribución de dos características distintas de una misma población en una nube de puntos.

(L7, 2016, p. 228)

) Registrar distribuciones de dos características distintas, de una misma población, en una tabla de doble entrada. Cuando se requiere analizar dos características se utilizan las tablas de doble entrada, ya que permiten organizar los datos de forma ordenada y conveniente (L7, 2016, p. 234).

En un estudio se quiere determinar si es más probable que un niño tenga asma si tiene padres fumadores que aquel cuyos padres no son fumadores. La tabla de frecuencias resume los resultados.

Tipo de familia	Frecuencia
Padres fumadores e hijo con asma	280
Padres fumadores e hijo sin asma	45
Padres no fumadores e hijo con asma	32
Padres no fumadores e hijo sin asma	143

	Padres fumadores	Padres no fumadores
Hijo con asma		
Hijo sin asma		

- Completa la tabla de contingencia a partir de la tabla de frecuencias.
- Según los datos, ¿cuál sería la conclusión del estudio? ¿En qué tabla te fijaste?

- Escribe una diferencia entre las tablas. Considera, por ejemplo, la cantidad de variables y sus categorías.

Figura 3.12. Registrar distribuciones de dos características distintas, de una misma población, en una tabla de doble entrada. (L7, 2016, p. 234)

En la Figura 3.12, encontramos una situación problema en que piden transformar la información entregada en una tabla como listado de datos, a una tabla de doble entrada, para apreciar mejor los datos proporcionados. Además, se les solicita a los estudiantes responder un listado de preguntas y obtener conclusiones a partir de los datos presentados en las tablas.

) *Comparar poblaciones mediante gráficos de dispersión para dos variables utilizando puntos con colores o separando la nube con una recta de manera intuitiva. Una nube de puntos permite realizar comparaciones entre dos poblaciones cuando se relacionan dos variables cuantitativas. Para esto, basta con representar los datos de ambas poblaciones en el mismo gráfico, con distintos colores para distinguirlas, y en la misma escala (L7, 2016, p. 240).*

En la Figura 3.13, se presenta una situación problema con los datos de dos poblaciones presentados en una tabla de listado de datos, en una primera actividad se les pide a los estudiantes realizar una nube de puntos para caracterizar el comportamiento de

los datos, tanto para hombres como mujeres. Posteriormente, se les solicita responder a una serie de preguntas de cada uno de los conjuntos de datos (hombres y mujeres), para que luego propongan una forma de representar los datos para poder realizar una comparación de las poblaciones.

De una población se extrae una muestra de 12 hombres y 12 mujeres a los cuales se les preguntó su edad y se les midió el IMC (índice de masa corporal). Los datos se registraron en las siguientes tablas:

Mujer	Edad	IMC	Hombre	Edad	IMC
1	35	29	1	22	19
2	45	31	2	39	25
3	18	27	3	25	22
4	25	28	4	40	26
5	29	30	5	28	20
6	36	29	6	32	27
7	57	34	7	51	24
8	20	30	8	33	22
9	45	27	9	44	23
10	11	20	10	19	16
11	54	31	11	58	26
12	41	25	12	51	24

• Representa, para los hombres y para las mujeres, las variables edad - IMC en una nube de puntos:

Figura 3.13. Comparar poblaciones mediante gráficos de dispersión para dos variables. (L7, 2016, p. 238)

Se presenta en la Tabla 3.2 la síntesis de los principales campos de problemas encontrados en los textos escolares chilenos examinados. Como podemos apreciar en dicha tabla, las situaciones problemas abordadas en cada libro van aumentando el grado de profundidad y dificultad en que son tratados los diferentes temas relacionados al muestreo de acuerdo al nivel de enseñanza. En L1, están relacionadas con características de la población y la representatividad de la muestra, en cambio en L3, están relacionadas con el número de muestra y la modelización o simulación de situaciones y en L4 se profundizan las ideas de muestreo aleatorio simple. Finalmente, en L5 no encontramos situaciones problema relacionadas al muestreo.

Tabla 3.2. Situaciones problema propuestas en los libros

Situaciones problemas	L1	L2	L3	L4	L5	L6	L7	L8
Estimar algunas características de una población desconocida por medio del muestreo.	x							
Determinar el número de muestras de un tamaño dado, que se pueden extraer desde una población de tamaño finito, con o sin reemplazo.			x	x		x		x
Relacionar la media de una población y la media aritmética de las muestras.				x		x		
Relacionar la proporción de una población y la proporción de las muestras.						x		
Identificar las muestras aleatorias y no aleatorias				x				
Elaboran modelos para el muestreo aleatorio			x					
Valorar la representatividad de una muestra	x							
Comparar dos o más conjuntos de datos a través de las medidas de tendencia central	x						x	
Comparar poblaciones y muestras usando medidas de posición		x						
Registrar distribuciones de dos características distintas de una misma población en una nube de puntos.							x	
Registrar distribuciones de dos características distintas, de una misma población, en una tabla de doble entrada.							x	
Comparar poblaciones mediante gráficos de dispersión para dos variables utilizando puntos con colores o separando la nube con una recta de manera intuitiva							x	

3.5.LENGUAJE MATEMÁTICO

Incluimos en el lenguaje matemático los términos, símbolos y todas las representaciones de los conceptos y propiedades asociados al tema o que se usan para reseñar datos y problemas. En el trabajo matemático son necesarios para representar los objetos abstractos o para operar con ellos para dar solución a los problemas, en la entrega de las soluciones a otra persona o para generalizar sus resultados.

En los Decretos de Enseñanzas Mínimas (MEC, 2007) se sugiere que la competencia matemática es:

La habilidad para utilizar y relacionar los números, sus operaciones básicas, los símbolos y las formas de expresión y razonamiento matemático, tanto para producir e interpretar distintos tipos de información, como para ampliar el conocimiento sobre aspectos cuantitativos y espaciales de la realidad, y para resolver problemas relacionados con la vida cotidiana y con el mundo laboral (p.20).

Estos Decretos le atribuyen una gran importancia al uso tanto del lenguaje matemático como también del lenguaje natural para entregar información y soluciones al resolver de problemas. Destacamos el interés en el uso del lenguaje matemático en los distintos lineamientos curriculares analizados en el Capítulo 1 (Tabla 1.13). También en el nuevo currículo español (MECD, 2015) se indica que el uso del lenguaje formal

matemático, permite la argumentar y explicar dichos fenómenos y comunicar con mayor precisión los conocimientos. Además, se han encontrado los siguientes Estándares de aprendizaje relacionados con el lenguaje:

J *Matemáticas. 1º y 2º ESO*: Expone y defiende el proceso seguido además de las conclusiones obtenidas, utilizando distintos lenguajes: algebraico, gráfico, geométrico y estadístico-probabilístico. (MECD, 2015, p.409)

J *Matemáticas orientadas a las enseñanzas académicas. 3º ESO*: Utiliza un vocabulario adecuado para describir, analizar e interpretar información estadística de los medios de comunicación. (MECD, 2015, p.394)

J *Matemáticas aplicadas a las Ciencias Sociales*: Usa el lenguaje, la notación y los símbolos matemáticos adecuados al contexto y a la situación. Utiliza argumentos, justificaciones, explicaciones y razonamientos explícitos y coherentes. (MECD, 2015, p.386)

De acuerdo con las Bases Curriculares (MINEDUC, 2013) uno de los propósitos del currículo de Matemática radica en que “los estudiantes sean capaces de transitar entre los distintos niveles de representación (concreto, pictórico y simbólico), traduciendo situaciones de la vida cotidiana a lenguaje formal o utilizando símbolos matemáticos para resolver problemas o explicar situaciones concretas” (MINEDUC, 2013, p. 87), permitiendo así que las expresiones matemáticas adquieran un sentido cercano para los alumnos.

La utilización adecuada del lenguaje facilita el aprendizaje significativo y contribuye al desarrollo de un adecuado razonamiento en el estudiante. Godino, Batanero y Font (2007) en el enfoque ontosemiótico, indican que el lenguaje matemático “es fundamental puesto que se supone que los objetos matemáticos emergen de las prácticas de una persona o institución al resolver problemas, y en estas prácticas se requiere el lenguaje” (Gómez, 2014, p.53). Así, los alumnos han de alcanzar un amplio manejo del lenguaje para comprender las diversas situaciones problemas planteadas, solucionar tareas, expresar y argumentar sus procedimientos para llegar a dicha solución.

Como señalan Cordero y Flores (2007), el discurso matemático escolar viene frecuentemente determinado por el lenguaje del texto escolar, “que regula las acciones de enseñanza y aprendizaje sobre cada tema” (Gea, 2014, p.73). A continuación, se muestran diversas caracterizaciones de los tipos de lenguaje encontrados en los textos escolares analizados, para ello se emplean variables previamente usadas en investigaciones previas como las desarrolladas por Ortiz, Batanero y Serrano (2001); Gea (2014) y Gómez (2014),

entre las cuales se destacan: “términos y expresiones verbales; notación simbólica y expresiones algebraicas; y representaciones tabulares y gráficas” (Gea, 2014, p106).

Los textos de secundaria contienen representaciones de tipo diverso: verbal, numérico, simbólico, tabular y gráfico. A continuación, analizamos diferentes tipos de lenguaje que aparecen en los textos.

3.5.1. LENGUAJE VERBAL

En primer lugar, siguiendo la metodología de análisis propuesta por Gea (2014), se analiza “el lenguaje verbal, es decir, las palabras y expresiones empleadas en los libros para denotar los conceptos, propiedades, y plantear o describir problemas, procedimientos y justificaciones o argumentos” (p.106).

Rothery (1980) señala que existen en la enseñanza de las matemáticas, tres clases de expresiones y que las dificultades en el aprendizaje propenden a emerger, principalmente, con las primeras dos categorías:

- a. “Términos matemáticos específicos que, normalmente, no forman parte del lenguaje cotidiano” (Gómez, 2014, p.53); un ejemplo sería el de distribución muestral y otro el de esperanza matemática.
- b. “Palabras usadas en matemáticas y el lenguaje ordinario, aunque no con el mismo significado” (Gómez, 2014, p.53); por ejemplo, imposible o seguro, que se usan en el lenguaje cotidiano para describir sucesos cuya probabilidad es respectivamente muy pequeña o muy grande, mientras en matemáticas las probabilidades han de ser 0 y 1.
- c. “Palabras con significados iguales o muy próximos en ambos contextos” (Gómez, 2014, p.53), como, por ejemplo, muestra o población.

En la Tabla 3.3, se muestran las palabras encontradas en los textos analizados, donde vemos que entre las categorías indicadas por Rothery, la más frecuente es la de términos específicos. Como cada uno de estos términos específicos corresponde a un concepto o a una propiedad, el listado evidencia la complejidad del tema, porque es una gran cantidad de conceptos y propiedades que el estudiante debe aprender y aplicar, los cuales se puntualizan en las secciones siguientes.

Tabla 3.3. Palabras encontradas en el análisis de los textos

Tipo	Ejemplos
Términos específicos	Aleatoria, cardinalidad, censo, combinación, conjunto, diagrama de árbol, distribución, encuesta, espacio muestral, estudio estadístico, experimento aleatorio, frecuencia (absoluta, relativa, acumulada), función de probabilidad, ley de los grandes números, media, medias muestrales, muestreo aleatorio simple, permutación, porcentaje, principio multiplicativo, probabilidad experimental, regla de Laplace, representatividad, subconjunto, variable aleatoria, variable estadística, variación.
Ordinarios con diferente significado	Dato, gráfico, inferencias, muestra, población, procedimiento, promedio, rango, representativa, tabla.
Ordinarios con el mismo significado	Analizar, asocia, calcular, clasifica, completa, conceptos, construye, datos, decisiones, definición, describe, determina, estimar, estrategias, evalúa, identificar, interpretar, investiga, justifica, observar, predicciones, recolecta información, recopilar, representar, resuelve, utiliza.

3.5.2. LENGUAJE NUMÉRICO

También encontramos diversas expresiones en lenguaje numérico, específicamente números enteros, decimales y fracciones.

Gómez (2014) señala que “los números enteros se introducen desde la enseñanza primaria con función nominal, ordinal o numérica” (p.55). En el primer caso, en el tema del muestro el entero representa un valor de la variable aleatoria de interés o un suceso en el experimento; por ejemplo, en el lanzamiento de un dado, al preguntar “*Al elegir el 6 o el 1, ¿es más difícil ganar que al elegir cualquier otro número? Justifiquen su respuesta*” (L1, 2015, p. 338). En el segundo caso, el número entero señala un orden “*¿Se puede considerar al 7° básico como un subconjunto de los estudiantes del colegio? ¿Por qué?*” (L1, 2015, p. 284), o “*Lanzamiento 1°*” (L1, 2015, p. 238,). En el tercer caso, “el entero representa un conteo, ya sea de casos favorables o posibles en el significado clásico; o de frecuencias absolutas en el significado frecuencial” de la probabilidad (Gómez, 2014, p56).

En los niveles seleccionados, las fracciones se utilizan, principalmente “para representar el valor de la probabilidad o la frecuencia relativa” (Gómez, 2014, p.56). Por ejemplo, “*Vicente y Nicole averiguan que la probabilidad teórica de obtener un 3 al lanzar un dado es igual a $\frac{1}{6}$ [...]*” (L3, 2013, p. 299).

En algunos casos, en primero medio, se hace la equivalencia de estas fracciones con números decimales en la estimación frecuencial de la probabilidad, por ejemplo: “*La*

frecuencia relativa de lanzar un dado y obtener 3 es, aproximadamente, 16,8%, es decir, la probabilidad experimental es de 0,168.” (L3, 2013, p. 299).

Podemos apreciar que los distintos conjuntos numéricos, se utilizan en todos los cursos de la serie analizada (Tabla 3.4), así mismo, “se presenta el uso de fracciones para representar probabilidades o frecuencias relativas” (Gómez, 2014, p.56).

En la Tabla 3.4 distinguimos que los textos analizados presentan los diferentes tipos de lenguaje numérico, que no tienen diferencias en los diferentes niveles.

Tabla 3.4. Tipos de números incluidos en los libros de texto

Tipos de números	L1	L2	L3	L4	L5	L6	L7	L8
Números enteros	X	X	X	X	X	X	X	X
Números decimales	X	X	X	X	X	X	X	X
Fracción representa probabilidad	X	X	X	X	X	X	X	X
Fracción representa frecuencia relativa	X	X	X	X	X	X	X	X

3.5.3. LENGUAJE SIMBÓLICO

Gea (2014) indica que las notaciones simbólicas:

Permiten una comunicación comprimida entre individuos, pudiendo trabajar a un alto nivel de complejidad. La autora cita a Skemp (1980), quien diferencia las siguientes funciones de los símbolos matemáticos: comunicación, registro del conocimiento, formación de nuevos conceptos, confección de clasificaciones múltiples correctas, explicación, facilitar la actividad reflexiva, mostrar estructuras matemáticas, automatizar las manipulaciones rutinarias, recuperar información de la memoria y generar una actividad creativa. (p.80)

En el análisis de los textos escolares, se ha hallado gran diversidad de símbolos, al igual que Ortiz (2002).

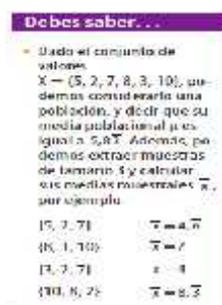


Figura 3.14. Cálculo de Media Poblacional y Media Muestral (L4, 2013, p. 284)

Como se puede apreciar en la Figura 3.14, vemos que en un mismo párrafo se utilizan distintos símbolos. Así, como el símbolo \bar{x} lo ocupan para distinguir la media

poblacional de la media muestral, aparece también la letra X para describir una variable y el símbolo de conjunto para englobar los valores de la variable.

Tabla 3.5. Lista de símbolos encontrados en los textos analizados

	L1	L2	L3
Números y operaciones	12, 20%, 1° $\frac{8}{84} = 0,095$	46, 2, 30%; 53,4	31, 41, 51, 51, 51, 51 = 900 8, $\bar{3}$, 3!, \$40990, 16,8%, 17, $\frac{19}{27}$
Letras	A, B, n.	A, B, C	n, a, b, c, E,
Índices	M_i	Q1, Q2, Q3	A_i, B_i, f_r
Símbolos	=, %,		f, X, \bar{x} , \varnothing , $\{ \}$, $\{ \}$, =, %,
Expresiones	$n = 20, f, F, f_r, X \frac{f}{n},$ $f\% = f_r \cdot 100$ $\bar{x}, M_o, \frac{n+1}{2},$	P% Q1, Q2, Q3 mínimo, máximo	$AB = BA, f_r, X \frac{f}{n}$
Conjuntos	A, B, $\{C, S\}$	$E = \{1,2,3,4,5,6, \}$	$\varnothing, X \sum \text{cara, sello*}$ $A, X \sum, 3, 5, 7, 9, 11, 13, 15, 17^*$
Fórmulas combinatorias			$P_2^3, X \frac{3!}{2!}, X \frac{3 \cdot 2 \cdot 1}{2 \cdot 1} = \frac{6}{2}, X 3$ $n! \cdot X n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot 3 \cdot 2 \cdot 1$ $C_r^n, X \frac{n!}{(n-r)! \cdot r!}, \frac{P_k^n}{C_k^n}$ $C_p^n, X \frac{n \Gamma p Z 1}{p}$ $V_r^n, X \frac{n!}{(n-r)!}, P_5, P_{(4)}, V_5^6,$ $C_3^4, P_{(a,b,c)}^n, X \frac{n!}{a!b!c!}$
Intervalos		[2,4[•2,5•
Pares ordenados	(1,2)		
Fórmula de la probabilidad	$f(A) = \frac{n^\circ \text{ de } A}{n^\circ \text{ de } S} = \frac{f}{i_1}$	$f(1) = \frac{n^\circ \text{ de } A}{n^\circ \text{ de } S} = \frac{f}{i_1}$	$P(A) = \frac{n^\circ \text{ de casos favorables}}{n^\circ \text{ casos totales}}$ $P(A) = \frac{n^\circ \text{ de veces que ocurre el evento}}{n^\circ \text{ de veces que se realiza el experimento}}$

En los textos chilenos, podemos encontrar diferentes cuadros resúmenes que presentan definiciones o conceptos que son expresados en lenguaje simbólico, que reproducimos en las Tablas 3.5 y 3.6.

Otros símbolos frecuentemente usados son la igualdad, el porcentaje, la coma decimal y el símbolo de periodo, de fracción, símbolos combinatorios, símbolos conjuntistas para señalar el espacio muestral, el uso de paréntesis para señalar los intervalos en tablas de frecuencias para datos agrupados, y algunos parámetros estadísticos como la media y la media poblacional.

Tabla 3.6. Lista de símbolos encontrados en los textos analizados

	L4	L7	L8
Números y operaciones	\$100, \$500	6,1; 20%	$5 \cdot 4 = 20$ $5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$
Letras	p, m,		ABCD
Índices			C_3^9
Símbolos	$\bar{x}, \mu, X, \{ \}, =$		
Expresiones	$p = 6; m = 2$		
Conjuntos	$X \times \sum 7, 8, 2, 4, 5, 12, 2, 0, 10, 12^*$	$C, D, C \cup D, C \cap D$ $\mathcal{C} = \{(2, \mathcal{C}), (2, \mathcal{D}), (2, \mathcal{S}), (2, \mathcal{S})\}$	
Fórmulas combinatorias	$n!, P_n \times n!, V_n^m \times \frac{m!}{f_m Z n A}$ $C_n^m \times \frac{m!}{f_m Z n A} \times \frac{m}{n}$		$F_n = n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot 2 \cdot 1$ $P_n \times n! \times V_n^m \times \frac{m!}{f_m Z n A}$ $C_n^m \times \frac{m!}{f_m Z n A} \times \frac{m}{n}$
Intervalos			
Pares ordenados	(1,1)	(4,8;6,0)	
Fórmula de la probabilidad	$f_r(A) = \frac{\text{n}^\circ \text{ de casos favorables al suceso } A}{\text{n}^\circ \text{ de casos posibles}}$	$P(C \cup D) = \frac{\#(C \cup D)}{\#\Omega}$	
Funciones	$f(x) : \mathbb{R} \rightarrow \bullet 0, 1'$		$P(X = x)$

Se puede apreciar en las Tablas 3.5 y 3.6, que dependiendo del curso el uso de símbolos va adquiriendo mayor complejidad. Se observa que el símbolo \bar{x} se ocupa para distinguir la media muestral de la media poblacional \sim . Aparece también la letra X para describir una variable y el símbolo de conjunto para englobar los valores de la variable. Para denotar al espacio muestral se utiliza el símbolo \mathcal{C} . En el caso de la variable aleatoria se utiliza el símbolo $x(\mathcal{C})$.

Otros símbolos frecuentemente usados son la igualdad, la coma decimal y el símbolo de periodo, de fracción, símbolos combinatorios, el símbolo de factorial, la notación de intervalos para las tablas de frecuencia de datos continuos.

Cabe mencionar que los símbolos presentes de los libros de tercer medio (L5) y cuarto medio (L6), presentan principalmente símbolos relacionados con la probabilidad condicional, distribución binomial y con la distribución normal, temas que no son considerados en este estudio.

3.5.4. LENGUAJE TABULAR

La tabla es característica de la enseñanza de la estadística y permite estructurar información, presentando, la correspondencia existente entre los diferentes tipos de frecuencia y los valores de la variable (Ortiz, 2002). Para leerla o construirla el niño

necesita entender la forma en que se debe realizar su construcción, distinguiendo los distintos significados que tienen los números incorporados en el cuerpo de la tabla. Hemos encontrado varios tipos de tabla. Como señala Cazorla (2002), las tablas estadísticas pueden emplearse para presentar información además de ser una herramienta útil para el análisis de datos. Hemos encontrado varios tipos de tabla.



Figura 3.15. Estimación de la probabilidad mediante la frecuencia (L3, 2013, p. 298)

En primer lugar, tenemos *tablas de frecuencia*, que pueden ser absolutas o relativas, encontradas también en Ortiz (2002) y Gómez (2014). Además de la columna de datos se añade otra con la frecuencia que corresponde a cada valor, representado en la tabla. Dichas tablas se usan para acercar a los estudiantes a la estimación de la probabilidad, mediante el cálculo de las frecuencias relativas. Se presenta un ejemplo en la Figura 3.15 que exhibe los datos asociados a los experimentos de lanzar un dado y una moneda con sus correspondientes frecuencias absolutas respectivamente.

En otras secciones del libro de texto, se plantean situaciones problema en donde se muestra una *tabla de frecuencias para datos agrupados* y se les solicita a los estudiantes calcular la frecuencia relativa de los datos presentados en ella (Figura 3.16).

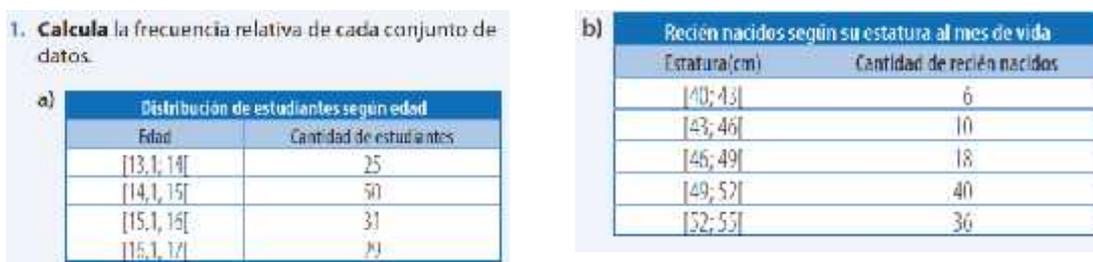


Figura 3.16. Cálculo de la Frecuencia Relativa (L3, 2013, p. 300)

Un segundo tipo de tabla corresponde a *simple listado de datos*, también identificado por Gómez (2014), “en el que cada fila (o columna) de la tabla representa los datos de cada una de las variables observadas en cada individuo de la muestra” (Gea,

2014, p.111). Además, se usan *tablas de doble entrada*, que no aparecen en los trabajos citados, en las que los estudiantes deben completar datos de experimentos sin reposición para el estudio de muestras. En la Figura 3.17 se presentan estos dos tipos de tabla. Estas tablas tienen mayor dificultad de interpretación pues las celdas se refieren a las modalidades de una variable bidimensional.

La tabla de contingencia o de doble entrada es aquella que sirve para contar la cantidad de individuos u objetos con dos tipos de características o variables cualitativas. Está conformada por filas y columnas. Las filas están formadas por las categorías de una variable, y las columnas, por las de la otra variable. En cada una de las casillas formadas se ubica la cantidad de datos que tienen ambas características simultáneamente (L7, 2016, p. 234).

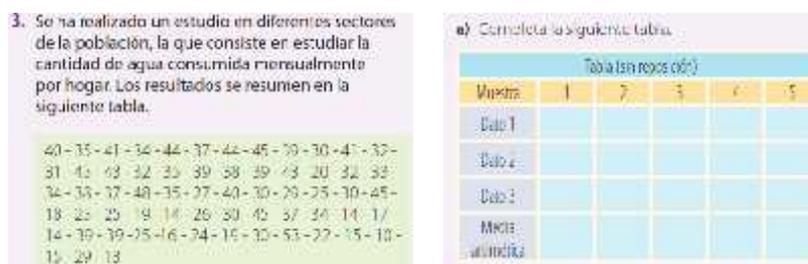


Figura 3.17. Selección de Muestras (L4, 2013, p. 286)

Por otro lado, en situaciones problemas del cálculo de probabilidad encontramos también *tablas de doble entrada* (ver Figura 3.18), las que brindan información estadística de dos sucesos relacionados entre sí, en este caso el sexo de los alumnos y la especialidad técnica a la cual pertenecen.

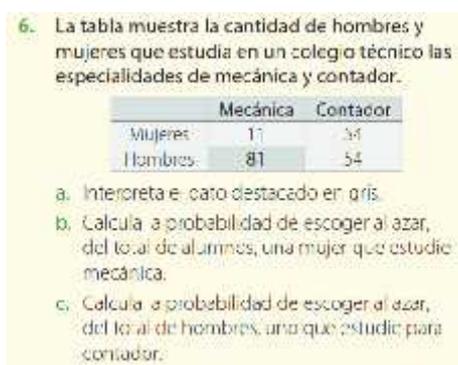


Figura 3.18. Tabla de doble entrada (L1, 2015, p. 348)

Asimismo, encontramos *Tablas que organizan los cálculos*, que se suelen usar para indicar los pasos necesarios para estimar la probabilidad mediante el cálculo de la frecuencia relativa de cada experimento (Figura 3.19). En otros casos se dan vacías y se

utilizan para que el estudiante anote los resultados de un experimento y a partir de ellos organice los cálculos.

Paso 2 Calcular la frecuencia relativa en ambos experimentos.

Resultados del dado	Frecuencia absoluta	Frecuencia relativa
1	90	18%
2	81	16,2%
3	87	16,8%
4	96	19,2%
5	80	16%
6	66	13,8%
Total	500	100%

Resultados de la moneda	Frecuencia absoluta	Frecuencia relativa
Cara	263	52,6%
Sello	237	47,4%
Total	500	100%

Figura 3.19. Estimación de la Probabilidad Frecuencial (L3, 2013, p. 298)

A modo de síntesis, podemos clasificar los distintos tipos de tablas que aparecen en la serie de textos analizados, como se muestra en la Tabla 3.7.

Tabla 3.7. Tipos de tablas encontradas en los textos

Tipos de Tablas	L1	L2	L3	L4	L5	L6	L7	L8
Tabla de frecuencia sin agrupar	x	x	x		x		x	x
Tabla de frecuencias agrupadas	x	x	x	x	x		x	
Tablas que organizan los cálculos		x	x	x	x		x	x
Listado de datos	x	x	x	x	x	x	x	x
Tablas de doble entrada	x		x	x	x		x	x

Como se aprecia en la Tabla 3.7, los libros de primero medio (L3 y L7) son los más completos en presentar la información en los distintos tipos de tabla, igualmente en el caso del libro de tercero medio (L5). En el caso del de séptimo básico (L1), solo considera tablas de frecuencia para datos agrupados o no y listados de datos, y las tablas de doble entrada se utilizan para el cálculo de probabilidades. En octavo básico (L2) el único tipo de tabla que no encontramos es de doble entrada. En el de segundo medio (L4), se consideran todos los ejemplos excepto las tablas de datos sin agrupar, en cambio, en el otro texto de segundo medio (L8), no se encuentran tablas de frecuencias agrupadas. Finalmente, en el caso del libro de cuarto medio (L6), solo se observan tablas de listados de datos, ya que principalmente se abordan temas de distribuciones y cálculos de probabilidades.

3.5.5. LENGUAJE GRÁFICO

También analizado por Ortiz (2002) y Gómez (2014), el lenguaje gráfico es muy frecuente en los libros de texto, bien para proporcionar diferentes tipos de información o como instrumento de análisis de datos. Se considera a los gráficos estadísticos como un

recurso apropiado para transmitir ideas, indicar las relaciones entre diversas variables, de utilizar con funciones instrumentales. Ello es debido a la cantidad de información que pueden representar en forma eficiente (Cazorla, 2002).

Podemos destacar que los principales tipos de gráficos que aparecen en los textos escolares chilenos para apoyar el estudio de los conceptos relacionados con muestreo son: gráfico de barras, gráfico circular (o de sectores), diagrama de árbol, entre otros. Muchos de estos gráficos ya han sido utilizados por los estudiantes en la educación primaria, por lo que, en principio, su lectura debiera ser sencilla para el estudiante.

Diagrama de árbol

Fischbein (1975) señala que los diagramas de árbol son un recurso para resolución de problemas que de acuerdo a Ortiz (2002)

permite sugerir e inculcar la generalización iterativa (extensión de una cierta propiedad en modo iterativo), que es la base de la inducción matemática. También posibilita la generalización constructiva, es decir la adaptación de un problema a otros derivados de él (p. 238).

Este tipo de gráficos, facilitan la enumeración para determinar el espacio muestral o señalar las posibles muestras y también posibilita la resolución de problemas de probabilidad compuesta.

Al igual que en Ortiz (2002), en el caso de los conceptos combinatorios (permutación, variación, combinación), se propone realizar diagramas de árbol para visualizar todos los resultados posibles de un experimento aleatorio o todas las muestras posibles de una población.



Figura 3.20. Diagrama de árbol (L3, 2013, p. 280)

En la Figura 3.20 se muestra un ejemplo, que sirve para formar todas las posibles variaciones (recorridos que realizan los camiones de Santiago a Concepción) y para explicar el principio multiplicativo, que permite determinar la cantidad total de caminos diferentes que se pueden ejecutar.

Gráfico de barras

Arteaga (2011) define este tipo de gráfico como:

Es una representación gráfica que puede ser usada para representar la distribución de frecuencia de variables cualitativas, cuantitativas discretas o incluso variables continuas, si los intervalos de valores se han transformado en categorías. (p. 8)

Este tipo de gráfico representa la frecuencia en una barra de altura proporcional a la misma.

Encontramos la utilización del gráfico de barras en situaciones problemas de aplicación del concepto combinación (Figura 3.21), en donde el estudiante debe inferir el error que se presenta en dicha solución. Para ello debe leer los datos del diagrama y revisar la solución ofrecida a la derecha de la gráfica.

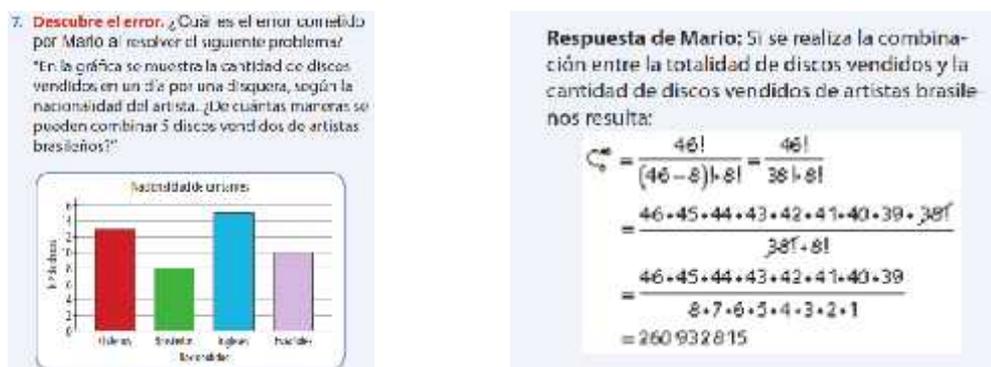


Figura 3.21. Gráfico de barras y combinaciones (L3, 2013, p. 291)

Histograma

Un histograma (ver Figura 3.22) “es una representación gráfica en forma de barras continuas, en las que sus alturas son proporcionales a la frecuencia absoluta de los intervalos representados” (L3, 2013, p. 250).

En la Figura 3.22, vemos el Histograma y Polígono de frecuencia de una situación problema en donde se trabajan las distintas representaciones de datos agrupados en intervalos.



Figura 3.22. Histograma y polígono de frecuencia (L3, 2013, p. 249)

Polígono de Frecuencias

Un polígono de frecuencias (ver Figura 3.22). corresponde a la línea poligonal, que se obtiene al unir los puntos referidos a las marcas de clase y la frecuencia absoluta de cada intervalo. (L3, 2013, p. 250)

El polígono de frecuencia acumulada (ver Figura 3.23) se representa uniendo los puntos referidos al límite superior y frecuencia acumulada de cada intervalo. También se conoce con el nombre de Ojiva. (L3, 2013, p.250)

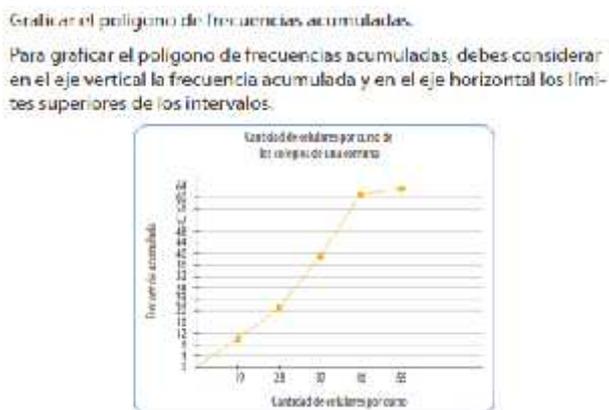


Figura 3.23. Polígono de frecuencias acumulada (L3, 2013, p.250)

Gráficos circulares o de sectores

Se usa principalmente para representar variables cualitativas, y la frecuencia de las diferentes categorías se representa mediante la amplitud de un sector circular proporcional a dicha frecuencia. En los textos, para el estudio de la probabilidad clásica, se plantean problemas que aplica de la Regla de Laplace mediante la utilización de gráficos

circulares. Así, en la Figura 3.22 se usa posteriormente para identificar en cuál de las ruletas se puede aplicar la Regla de Laplace.



Figura 3.24. Gráfico Circular (L3, 2013, p. 296)

En la Figura 3.24, lo estudiantes deben determinar que es más fácil poder calcular la Regla de Laplace con las ruletas (a) y (b), ya que es posible enumerar los casos favorables y los casos posibles, en cambio en la ruleta (c) no podemos determinar a simple vista los casos favorables ni los casos posibles.

Gráfico de Líneas

De acuerdo a Arteaga (2011), un gráfico de líneas (ver Figura 3.25):

Se usa para representar, bien frecuencias de una variable cualitativa o bien valores numéricos de una serie de datos, y no para representar variables cuantitativas. Un gráfico de líneas usa puntos conectados por líneas para mostrar cómo cambia el valor de algo (a lo largo del tiempo o del valor de la variable). (p.12)

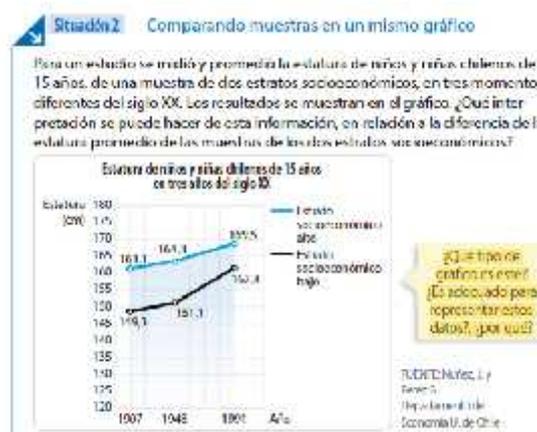


Figura 3.25. Gráfico de líneas (L2, 2015, p.297)

Gráfico de Caja

Un diagrama de cajas o Boxplot (ver Figura 3.26):

Es un gráfico que muestra la distribución de los datos, dividiendo estos en cuatro partes iguales mediante los cuartiles. Para construir un Boxplot se dibuja una caja que va desde Q_1 hasta Q_3 . Dentro de ella se traza una línea vertical en la mediana. Luego, se trazan líneas desde la caja a los valores mínimo y máximo. (L3, 2013, p. 269)

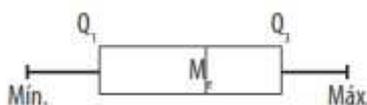


Figura 3.26. Diagrama de caja (L3, 2013, p. 269)

Gráfico de puntos

El gráfico de puntos (ver Figura 3.27), de acuerdo a Arteaga (2011) posibilita:

representar conjuntos pequeños de datos y puede ser fácilmente construido (Moore, 1999). Se usa para variables cuantitativas, tanto discretas, como continuas. [...]El gráfico de puntos permite detectar tendencias en conjuntos pequeños de datos, así como los valores extremos o intervalos de frecuencia nula. Este tipo de gráfico puede ser utilizado también para representar variables cualitativas, representando en el eje de abscisas los valores de cada una de las modalidades de la variable en estudio y poniendo sobre cada una de dichas modalidades tantos puntos como frecuencia absoluta corresponda a cada variable. (p.12)

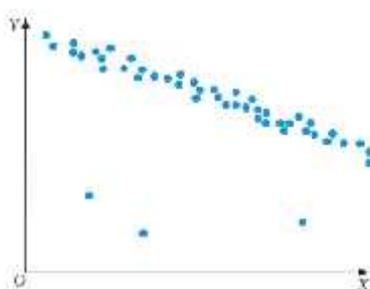


Figura 3.27. Nube de puntos. (L7, 2016, p. 231)

En la Figura 3.28, podemos apreciar la comparación de dos diagramas de puntos que representan los puntajes obtenidos por dos cursos de 7° básico, en las competencias de la semana del colegio, en este gráfico, los alumnos deben determinar los puntajes totales obtenidos por cada curso y el puntaje más alto del 7°A. La principal dificultad para contestar dichas preguntas, tiene relación con ponderar cada puntaje por su frecuencia absoluta.

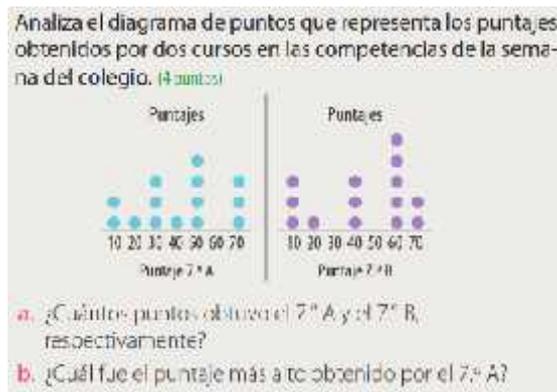


Figura 3.28. Gráfico de puntos (L1, 2015, p. 282)

Diagrama de Venn

Ortiz (2002) señala que de acuerdo a Fischbein (1975), los diagramas de Venn son un tipo particular de diagramas. Estos “son representaciones de fenómenos y relaciones entre los mismos” (p.237) Este autor señala que los diagramas “se construyen artificialmente para representar o modelar un sistema original existente en la realidad” (Ortiz, 2002, p.237).

En los libros de texto de matemáticas, aparecen frecuentemente los diagramas y se usan habitualmente “como elemento heurístico en la resolución de problemas (Love y Pimm, 1996)” (Ortiz, 2002, p.237). El diagrama como objetivo pretende afianzar la forma de pensar y orientar el interés del lector. Siguiendo a Ortiz (2002), este autor señala que “debido a su naturaleza icónica, se olvida continuamente que es preciso saber “leer” el diagrama” (p.237). De las características intuitivas del diagrama, destacamos:

-)] Ofrece una representación global, sinóptica de una estructura o un proceso y ello contribuye a la globalidad de su comprensión. (Ortiz, 2002, p.237)
-)] Es un punto ideal entre la interpretación conceptual y la expresión práctica de una cierta realidad. (Ortiz, 2002, p.237)

En la Figura 3.29, podemos apreciar un diagrama de Venn utilizado para apoyar la definición de muestra, por lo que se ejemplifica el concepto de subconjunto, ya que la muestra es un subconjunto de la población.

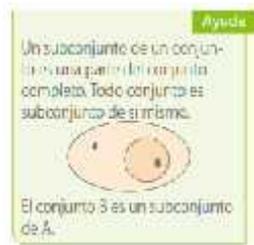


Figura 3.29. Diagrama de Venn (L1, 2015, p. 284)

Fotografías

En el trabajo de Ortiz (2002) se analizan también las ilustraciones o fotografías que se incluyen para atraer la atención del alumno hacia el tema. En nuestro caso, en los distintos textos escolares, podemos encontrar diferentes fotografías para contextualizar alguna situación problema o como aparece en la Figura 3.30 para representar los elementos de una población o muestra.



Figura 3.30. Población y muestra (L1, 2015, p. 286)

Representaciones icónicas

Finalmente, encontramos el lenguaje icónico para representar esquemáticamente objetos, como, por ejemplo, hemos visto anteriormente las ruletas (ver Figura 3.22). Se trata de hacer concreta de alguna manera una situación para que el estudiante la perciba con menor abstracción. También encontramos iconos para representar bolas en urnas en el trabajo con las permutaciones, como en la Figura 3.31.



Figura 3.31. Extracción de bolitas en una urna - Permutación (L3, 2013, p. 284)

Como podemos apreciar en la Figura 3.31, el lenguaje icónico sigue estando presente en los libros de educación media, pues como los lineamientos de las Bases Curriculares hacen hincapié en ofrecer a los alumnos la relación entre lo concreto, pictórico y simbólico, este tipo de lenguaje se utiliza para ejemplificar la extracción de bolas de una urna.

Esquemas

En algunos de los distintos textos escolares analizados, se presenta a modo de síntesis de la unidad de estadística y probabilidad, un esquema o mapa conceptual, en donde los estudiantes deben completar dicho mapa, con un listado de conceptos (ver Figura 3.32).

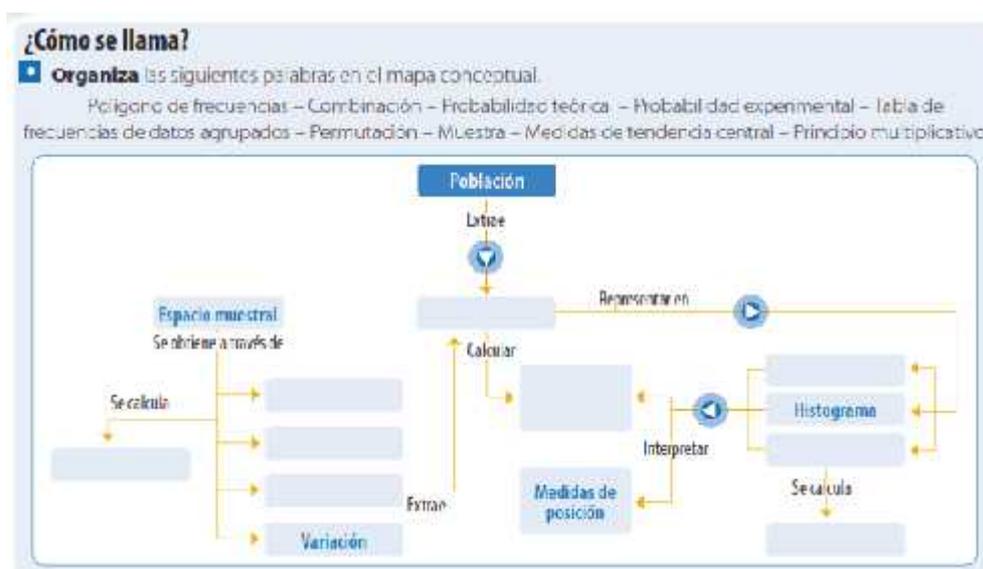


Figura 3.32. Esquema o mapa conceptual (L3, 2013, p. 309)

En la Tabla 3.8, podemos apreciar la síntesis de las distintas representaciones gráficas presentes en los textos analizados. Las fotografías y las representaciones icónicas son los únicos que aparecen en todos los textos analizados. Como podemos apreciar, L3 contiene una gama más amplia de representaciones, pasando de los distintos tipos de gráficos a las distintas representaciones como esquemas y fotografías. Consideramos que L4 presenta menor cantidad de representaciones gráficas debido a que los temas presentados son más abstractos e implican una mayor explicación de propiedades matemáticas como también señalan los cálculos a efectuar en las distintas situaciones planteadas.

En L1 y L2 hallamos una mayor cantidad de representaciones, sobre todo las referidas a los distintos tipos de gráficos estudiados en estos niveles educativos. Por otro lado, L3 incluye representaciones gráficas más complejas como el diagrama de árbol y grafos, diagrama de caja, histograma y polígono de frecuencias como también, aparece el uso de las representaciones contextualizadas como los esquemas, fotografías o íconos.

En L5 encontramos un mayor uso de diagrama de Venn pues se abordan temas de probabilidad condicional. L6 presenta una menor cantidad de representaciones gráficas ya que se plantean situaciones problemas relativas a la distribución normal.

En L7 se aprecia una mayor cantidad de representaciones gráficas como el diagrama de árbol y los gráficos de puntos, además de emplear el Diagrama e Venn para el cálculo de probabilidades y comprobar propiedades, en cambio. L8 incluye además del diagrama del árbol, gráficos de barra, histogramas y polígonos de frecuencia.

Tabla 3.8. Tipos de representaciones gráficas encontradas en los textos

Tipo de Gráfico	L1	L2	L3	L4	L5	L6	L7	L8
Diagrama en árbol	x	x	x	x			x	x
Gráfico de barras	x	x	x	x	x			x
Gráfico circular	x	x	x		x			
Gráfico de puntos	x						x	
Gráficos de líneas	x	x						
Gráfico de caja		x	x					
Histograma			x			x		x
Polígono de Frecuencias	x		x					x
Diagrama de Venn	x			x	x		x	
Esquemas	x	x	x	x	x	x	x	
Fotografías	x	x	x	x	x	x	x	x
Representaciones icónicas	x	x	x	x	x	x	x	x

3.6.CONCEPTOS

En la actividad matemática que realiza el sujeto, precisa recordar distintos fundamentos básicos o conceptos con los que opera y utiliza al dar solución a un problema, por medio de sus definiciones. Los conceptos y propiedades son elementos básicos que se demandan en el aprendizaje de las matemáticas. Aunque el conocimiento conceptual y el conocimiento procedimental, como señala Gómez (2014) los conceptos “están vinculados, el conocimiento conceptual es más generalizable, pues no está unido a un ejemplo concreto, incluye la comprensión implícita o explícita de los principios de un dominio dado y sus interrelaciones (Rittle-Johnson, Siegler y Alibali, 2001)”. (p.67)

El aprendizaje conceptual requiere un largo periodo de tiempo, por lo que Skemp (1980) sugiere que, en lugar de expresar solamente la definición de un concepto, es

necesario incluir adecuadamente los conceptos por medio de una amplia gama de ejemplos y de actividades apropiadamente secuenciadas. A continuación, describimos la manera en que se exponen los conceptos relacionados con el muestreo en los textos escolares analizados.

Conceptos combinatorios: variaciones, permutaciones y combinaciones

En nuestro estudio hemos encontrado conceptos combinatorios, ligados a los tipos de muestreo con y sin reemplazamiento. Como se ha indicado, las operaciones combinatorias de variaciones y permutaciones se suelen definir utilizando el modelo de selección, que es equivalente al muestreo y por tanto consideramos estos conceptos parte del significado del muestreo. Este conjunto de conceptos forma parte de las ideas estocásticas fundamentales propuestas por Heitele (1975), las cuales tienen que guiar la enseñanza de los sujetos durante todo el proceso de escolarización, incluida la formación universitaria.

Usualmente los conceptos combinatorios se introducen usando ideas de muestreo. Vemos que primero se plantea un problema; un ejemplo se presenta para el caso del concepto de combinación, donde su explicación surge a partir de la idea de seleccionar una muestra de dos personas de una población de cuatro, situación problema mostrada en la Figura 3.33 a la que acompaña una fotografía que alude al contexto (realizar una entrevista).

Israel está realizando un estudio en el que debe entrevistar grupos de dos personas escogidas al azar de un total de 4, Andrés, Berta, Carla y Daniela. Esta situación se traduce a elegir muestras de 2 personas de una población de 4 personas. ¿De cuántas maneras podría Israel elegir estas muestras?



Figura 3.33. Combinación (L3, 2013, p. 288)

Una vez planteadas las preguntas, se utiliza el diagrama de árbol para representar los datos de acuerdo a lo solicitado en el enunciado anterior. Las diferentes bifurcaciones del árbol (ver Figura 3.34) representan cada elección que se debe hacer y las ramas que salen de cada una las posibilidades de dicha elección. Se aclara en una nota que el orden es irrelevante.

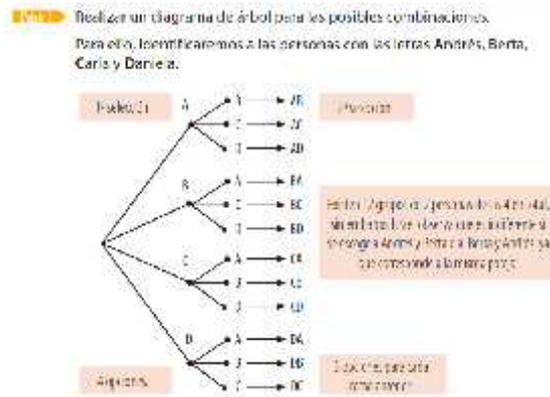


Figura 3.34. Diagrama de árbol para las posibles combinaciones (L3, 2013, p. 288)

Obtenidos todos los casos posibles, se descartan los casos repetidos, como se muestra en el paso 2, Figura 3.35 en que con el signo de igualdad y usando notación de letras (inicial del nombre de cada alumno) se muestran los casos repetidos.

Paso 2 Descartar las parejas que se repiten.

Combinaciones descartadas		Combinaciones que quedan	
AB = BA	AC = CA	AB	BC
CD = DC	BC = CB	AC	BD
AD = DA		AD	CD

Israel puede elegir entre 6 **combinaciones** de personas para su entrevista, es decir, puede elegir 6 muestras distintas de una población de 4 individuos.

Figura 3.35. Combinaciones (L3, 2013, p. 288)

Finalmente, se presenta un resumen con la definición matemática del concepto de combinación, donde podemos observar que se insiste en que el orden no importa (Figura 3.36). En forma similar se presentan el resto de conceptos combinatorios.

En resumen
 El número de combinaciones que se pueden efectuar con una cantidad n de elementos desde un conjunto de n elementos son los distintos grupos de r elementos que se pueden hacer con los n elementos, en este caso no interesa el orden con el cual se extraen o escogen los elementos.

Figura 3.36. Definición de combinación (L3, 2013, p. 288)

Conceptos probabilísticos

En los textos escolares analizados se recuerdan algunos conceptos probabilísticos cuando se introduce el muestreo, específicamente: experimentos aleatorios, suceso, espacio muestral y probabilidad.

) *Experimento Aleatorio.* Un experimento aleatorio es aquí que puede dar lugar a diferentes resultados y no podemos predecir cuál de ellos será el que ocurra. En la Figura 3.37, se plantea una actividad grupal (tres estudiantes) para realizar el experimento aleatorio de “lanzar un dado”, donde los estudiantes deben anotar sus resultados en la tabla respectiva, escribiendo el número que obtenido y quién ganó. Posteriormente, deben responder una serie de preguntas en relación a dicho experimento. No se da una definición explícita de experimento aleatorio, aunque si se define lo que es un experimento aleatorio equiprobable.

Experimento con un dado

Realiza el siguiente experimento:

- 1. Cada uno lanza una vez el dado.
- 2. Anota el resultado obtenido en una hoja de papel con el número que se obtiene a cada lanzamiento.
- 3. Suma los resultados de los lanzamientos.
- 4. Anota el resultado obtenido en la tabla y responde a las preguntas que se indican.

La cantidad:	El número que sale:	¿Quién ganó?
10		
20		
30		
40		
50		
60		
70		
80		
90		

1. ¿Qué números pueden elegir al lanzar un dado?
2. Al elegir el 6 o el 1, ¿es más difícil ganar que al elegir cualquier otro número? Justifiquen sus respuestas.
3. ¿Existe algún número con el cual, al elegirlo, haya más posibilidades de ganar? Justifiquen sus respuestas.
4. ¿Se pueden sacar a quien ganó el juego de acuerdo a la elección de sus números? Justifiquen sus respuestas.

Figura 3.37. Experimento aleatorio (L1, 2015, p. 338)

) *Espacio muestral.* “Se define como el conjunto de todos los posibles resultados en un experimento aleatorio” (L4, 2013, p. 378) y puede ser *finito o infinito*. Suele denotarse por la letra E o Ω . En la Figura 3.38, se solicita determinar el espacio muestral de cuatro experimentos.

4 Determina el espacio muestral (todos los casos posibles) de los siguientes experimentos:

- a. Elección al azar de un color de la bandera chilena.
- b. Se lanza al aire una moneda de \$100 y una de \$500.
- c. Se escoge un día al azar del mes de febrero, y se anota su número.
- d. Se lanza un dado dos veces, y se anota la diferencia entre los puntajes obtenidos.

Figura 3.38. Espacio Muestral (L4, 2013, p. 275)

En la Figura 3.39, se les solicita a los estudiantes relacionar cada experimento aleatorio con su espacio muestral, uniendo ambos mediante una línea. Podemos apreciar que esta actividad es de menor complejidad que la planteada en la Figura 3.33, ya que los elementos del espacio muestral están anotados en extensión, en cambio, la actividad anterior, presenta una mayor complejidad ya que se deben establecer todos los elementos del espacio muestral de cada experimento.

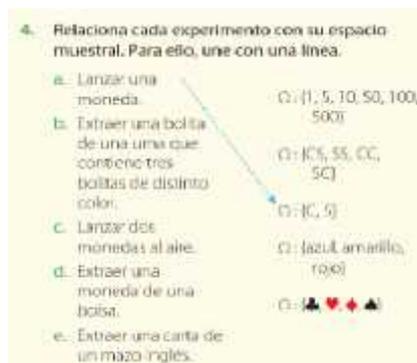


Figura 3.39. Espacio muestral (L1, 2015, p.340)

) *Sucesos y sucesos equiprobables.* Un suceso o evento A es cualquier subconjunto del espacio muestral. Se definen que los resultados de un experimento aleatorio son equiprobables, si todos poseen la misma probabilidad de ocurrir (por ejemplo, en L3, p. 247). Seguidamente se proponen ejercicios con sucesos equiprobables. Así, en la Figura 3.40 se considera el espacio muestral del experimento aleatorio “lanzar dos monedas”, para obtener la probabilidad en distintos sucesos.

1. Resuelve los siguientes problemas.

a) Se considera el experimento aleatorio “lanzar dos monedas” cuyo espacio muestral es:

$\{CC, SC, SS, CS\}$

- ¿Cuál es la probabilidad de obtener dos caras?
- ¿Cuál es la probabilidad de obtener dos sellos?
- ¿Cuál es la probabilidad de que se obtenga una cara y un sello?
- ¿Cuál es la probabilidad de que al menos se obtenga una cara?

Figura 3.40. Sucesos equiprobables (L3, 2013, p. 296)

) *Probabilidad clásica.* Hemos encontrado definida la probabilidad, tanto desde el enfoque clásico, como desde el frecuencial. “Cuando un experimento aleatorio tiene resultados equiprobables, se calcula la probabilidad de un evento mediante la regla de Laplace. Esto se conoce como probabilidad teórica o clásica. Para calcular la probabilidad teórica de un evento A se utiliza la expresión”:

$$P(A) = \frac{\text{n}^\circ \text{ casos favorables}}{\text{n}^\circ \text{ casos totales}} \quad (\text{L3, 2013, p.293})$$

En la Figura 3.41, se plantea una situación problema en un contexto cotidiano para los estudiantes, la realización de un experimento aleatorio con el juego de cartas “Uno” y se les plantea una interrogante para analizar las afirmaciones de dicha situación.

Joshua y Nadia están realizando un experimento aleatorio con el juego de cartas UNO. Se trata de sacar una carta sin mirar. Joshua afirma que la probabilidad de obtener una carta con un número es de $\frac{19}{27}$, mientras que Nadia indica que la probabilidad de extraer una carta que posea un símbolo es de $\frac{8}{27}$. ¿Es correcto lo que afirma cada uno? ¿Por qué?



Figura 3.41. Probabilidad Clásica (L3, 2013, p.294)

Luego, se recuerda una definición del concepto de equiprobabilidad (Figura 3.42).

Si un experimento cualquiera puede dar lugar a un número finito de resultados posibles y no existe razón que privilegie un resultado por sobre otros, es decir, los resultados son **equiprobables** (todos poseen la misma probabilidad de ocurrir), se puede calcular la probabilidad de un evento aleatorio según la regla de Laplace, realizando los siguientes pasos:

Figura 3.42. Resultados Equiprobables (L3, 2013, p.294)

Se introduce la Regla de Laplace a través de una serie de pasos que se describen en la Figura 3.43, en donde se definen los elementos del experimento aleatorio, se calcula el número de casos favorables para cada evento, se calcula el número de casos totales y finalmente, se calcula la probabilidad de cada evento mediante la Regla de Laplace.

1. Definir los eventos del experimento aleatorio.
 Experimento aleatorio: Sacar una carta al azar del juego de cartas UNO.
 Evento A: Sacar una carta del naipe UNO y obtener un número.
 Evento B: Sacar una carta del naipe UNO y obtener un símbolo.

2. Calcular el número de casos favorables para cada evento.
 Para el evento A existen 4 colores con los números del 0 al 9, y con los números del 1 al 9, por lo tanto, existen:

7 casos
 $10 \times 4 + 9 \times 4 = 76$
 76 casos corresponden a números.

Para el evento B existen 3 símbolos con 2 cartas de cada color y 2 símbolos con 4 cartas cada uno, es decir, en total existen:

10 casos
 $3 \times 2 \times 4 + 2 \times 4 = 10$
 10 cartas en total que corresponden a símbolos.

3. Calcular el número de casos totales.
 El naipe UNO tiene 108 cartas en total.

4. Calcular el cociente entre el número de casos favorables y el número de casos totales para obtener la probabilidad de cada evento.

$P(A) = \frac{\text{n}^\circ \text{ de casos favorables}}{\text{n}^\circ \text{ de casos totales}} = \frac{\text{cantidad de cartas con números}}{\text{total de cartas del naipe}} = \frac{76}{108} = \frac{19}{27}$

$P(B) = \frac{\text{n}^\circ \text{ de casos favorables}}{\text{n}^\circ \text{ de casos totales}} = \frac{\text{cantidad de cartas con símbolos}}{\text{total de cartas del naipe}} = \frac{10}{108} = \frac{5}{54}$

Aunque estos datos (de azar), ya que como los eventos del experimento aleatorio tienen equiprobables (todos con misma probabilidad para sacar cualquier carta al azar) las probabilidades mencionadas son correctas.

Figura 3.43. Probabilidad clásica (L3, 2013, p.294 - 295)

) **Probabilidad frecuencial.** “La probabilidad frecuencial (o empírica) de un evento A, se calcula mediante el cociente entre la cantidad de veces que ocurre el evento y la cantidad de veces que se realiza el experimento, es decir, se estima a partir de la frecuencia relativa, aunque no se suele enfatizar la idea de que de este modo sólo se obtiene una estimación”:

$$P(A) \approx \frac{\text{n}^\circ \text{ veces que ocurre el evento A}}{\text{n}^\circ \text{ de veces que se realiza el experimento}} \quad (\text{L3, 2013, p.298})$$

En la Figura 3.15, se describe una situación problema para el cálculo de probabilidades al realizar dos juegos: el lanzamiento de un dado y el lanzamiento de una moneda, registrando las repeticiones de dichos experimentos 500 veces en una tabla de frecuencias para cada juego. Para descubrir quien tiene mayor probabilidad de ganar, se realiza la repetición de los experimentos 500 veces y se registra los resultados en tablas de frecuencias, para después calcular la probabilidad frecuencial del experimento “lanzar un dado y que aparezca 3” y “que aparezca cara al lanzar una moneda” (ver Figura 3.44).

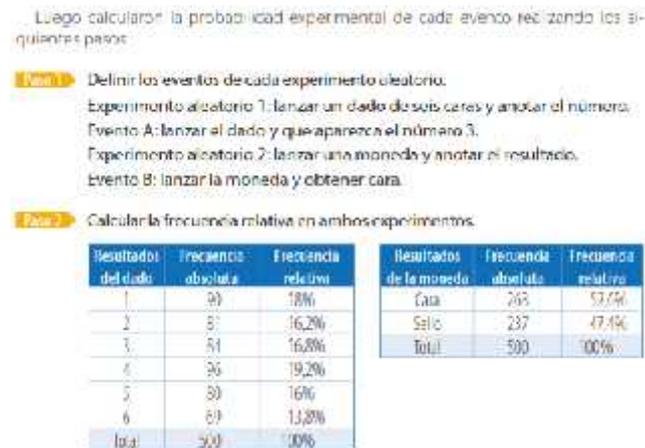


Figura 3.44. Estimación de la probabilidad por la frecuencia relativa (L3, 2013, p.298)

Finalmente, se comparan las frecuencias relativas de ambos experimentos y se concluye que el experimento “lanzar una moneda y obtener cara”, tiene una mayor probabilidad de ocurrencia (ver Figura 3.45).

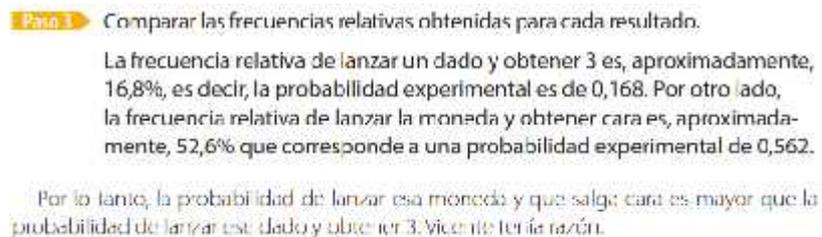


Figura 3.45. Comparación de frecuencias relativas (L3, 2013, p.298)

Conceptos de muestreo

J) *Población*. “Se llama población al conjunto total que se encuentra en estudio, y puede estar formado por personas, objetos, casos, números, etc.” (L4, 2013, p.276). En los libros el término no se define formalmente, sino se deja implícito su significado en los diferentes problemas y actividades. Así, en la Figura 3.46, se solicita determinar la población objeto de estudio y una posible muestra de ésta. Se espera que los estudiantes, con la ayuda del profesor diferencien ambos conceptos.

- 1** Determina en cada caso la población y una posible muestra de ella.
- Una fábrica de yogur quiere investigar sobre la calidad de sus productos.
 - Diego necesita saber el precio de un kilogramo de carne, para una comida familiar.
 - Ximena estudia respecto del tamaño de las hormigas que habitan en un insectario.
 - Dario desea saber si una ciudad cuenta con suficientes lluvias, para realizar una plantación.

Figura 3.46. Población y muestra (L4, 2013, p. 275)

) **Muestra.** Una muestra se define como: “Una muestra es un subconjunto de una población”. A través de las muestras se pueden inferir datos de la población (L1, 2015, p. 247). Para conseguir este objetivo se proponen ejercicios donde los estudiantes deban diferenciar la muestra y la población. Por ejemplo, en la Figura 3.47, se presentan dos gráficos de sectores para mostrar los resultados de una encuesta, correspondientes a un colegio y a un curso de 7° básico. Se plantean una serie de preguntas para que los estudiantes distingan la variable estudiada, infieran conclusiones a partir de la información presentada en los gráficos, identifiquen el concepto de muestra, diferencien entre la población y muestra, y empleen de manera intuitiva el concepto de representatividad muestral.

¿Qué es una población y una muestra?

Objetivo: El tuyo y el mío

¿Pueden explicar cómo se selecciona una muestra?

Una muestra es un subconjunto de la población. Se puede seleccionar una muestra en un colegio y se comparan los resultados con los de toda la población.

- Identifiquen qué se está investigando en la encuesta.
- ¿Qué pueden inferir acerca de los estudiantes de este colegio?
- ¿Se puede considerar al 7.º básico como un subconjunto de los estudiantes del colegio? ¿Por qué?
- ¿Por qué varían los resultados al considerar solo el 7.º básico?
- ¿Podrían decir que el 7.º básico es un buen representante del estudio realizado en todo el colegio? Justifiquen su respuesta.

Figura 3.47. Ejercicios para reforzar el concepto de muestra (L1, 2015, p. 284).

) **Tipos de muestra: aleatoria, sistemática, estratificada, no probabilística.** Los textos introducen los diferentes tipos de muestreo una vez que se ha trabajado la diferencia entre muestra y población. “Se llama muestreo aleatorio simple a la elección de una muestra de una población, de modo que cada elemento de la población tiene la misma probabilidad de ser escogido”. (L4, 2013, p.277)

Un colegio instalará una gradería alrededor de una cancha para ubicar a los alumnos para ver un partido muy importante de un campeonato interescolar. La empresa encargada desea conocer la masa promedio de los estudiantes para hacer estudios respecto a la resistencia del material. El colegio no tiene tiempo de registrar la masa de sus 600 estudiantes, por lo que decide tomar algunas muestras. Para ello:



Paso 1 Asigna un número a cada estudiante, de 1 a 600.

Alumnos = {1, 2, 3, ..., 598, 599, 600}

Paso 2 Escoge al azar, 10 números del 1 al 600 (utilizando una tómbola con bolitas numeradas, por ejemplo), y escoge a los estudiantes que corresponden a dichos números. Por ejemplo:

Muestra = {233, 573, 592, 427, 234, 591, 395, 84, 137, 161}

Paso 3 Registra las masas de estos estudiantes, y calcula el promedio de ellas.

Masas = {41,1 - 45,9 - 52,7 - 58,9 - 58,0 - 59,5 - 40,7 - 54,7 - 47,8 - 54,1}

Promedio: 50,9

Se ha escogido una muestra aleatoria, en la que cada estudiante tenía igual probabilidad de salir, sin considerar a qué curso pertenecían, edad, etc. Por esto, habíamos de muestreo aleatorio simple. El promedio obtenido se llama media muestral (que se anota \bar{x}), y se espera que, para un tamaño de muestra no demasiado pequeño, su valor sea cercano a la media poblacional (anotada como μ), es decir, el promedio de toda la población.

Figura 3.48. Muestreo Aleatorio Simple (L4, 2013, p. 276)

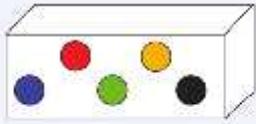
Como podemos apreciar en la Figura 3.48, para introducirlo, se describe una situación problema contextualizada que involucra el concepto de muestreo aleatorio simple. Se aprovecha la misma situación para introducir los conceptos de muestra aleatoria y media muestral. Se describe la población de estudiantes involucrados y se explica a través de tres pasos cómo se puede extraer una muestra aleatoria de estudiantes, para posteriormente registrar los datos requeridos para calcular la media muestral. Finalmente se añade una nota explicando las características del muestreo aleatorio simple.

No se introducen explícitamente otros tipos de muestreo, aunque en el libro L4 se incluye un ejercicio en que se pide averiguar las características de otros tipos de muestreo:

Conexiones: Averigua en qué consisten los siguientes métodos de muestreo. Aleatorio sistemático. Aleatorio estratificado. Aleatorio por conglomerados. Da, para cada uno, dos ejemplos en que sería mejor utilizar estos tipos de muestreo en lugar del muestreo aleatorio simple. (L4, 2013, p. 279)

) *Muestras ordenadas.* En el caso de considerar o no el orden, al extraer una muestra, hacemos uso de algunos conceptos combinatorios como las permutaciones y variaciones en las que importa el orden en que son extraídos los elementos, en cambio en las combinaciones no importa el orden. Este tema los textos los contemplan en la parte dedicada al estudio de la combinatoria. En la Figura 3.49, se presenta un problema de permutación en el que se deben realizar diferentes extracciones de bolas de colores de una urna, para responder una serie de cuestiones, dependiendo de si se quiere seguir un determinado orden en la selección de elementos.

b) Amalia participa en un juego que consiste en adivinar el orden de extracción de 5 bolitas desde una urna que contiene 5 bolitas de colores.



- ¿Cuántas extracciones distintas existen?
- Si Amalia dice que la primera bolita es de color azul, ¿cuántas extracciones distintas existen para que ella gane?
- Si Amalia dice que la primera bolita es azul y la segunda verde, ¿cuántas extracciones distintas existen para que ella gane?
- Si Amalia dice que la última bolita es de color amarillo, ¿cuántas extracciones distintas existen para que ella gane?

Figura 3.49. Muestras ordenadas y no ordenadas (L3, 2013, p. 287)

J) *Muestras con y sin reemplazamiento.* Podemos encontrar muestras con o sin reemplazamiento de elementos. Se tiene un muestreo con reemplazo si se obtiene un elemento de la población, si observa la característica que se trata de estudiar, se apunta ésta como un elemento de la muestra, y el elemento observado se devuelve a la población antes de sacar el siguiente elemento para la muestra. En este caso de muestreo con reemplazo es factible escoger cada miembro de la población más de una vez. Por el contrario, si una vez que un elemento de la población se elige para entrar en la muestra, se deja permanentemente fuera de la población, no será posible escoger cada miembro de la población más de una vez, y se tendrá un muestreo sin reemplazo (Olivera y Zúñiga, 1987). Por ejemplo, en las extracciones sin reposición se extrae la bolita y no se vuelve a ingresar a la urna. En la Figura 3.50, se plantea una situación problema sobre variación, en donde se discute sobre los elementos repetidos o no de los ordenamientos que se pueden hacer con los números del 1 al 6.

3. Resuelve los siguientes problemas.

a) ¿Cuántos números de tres cifras se pueden formar con las cifras 1, 2, 3, 4, 5 y 6 sin que se repita ninguna? ¿Y si se repite cada uno en dos ocasiones? ¿Cuántos de esos números terminarán en 26?

Figura 3.50. Muestra con repetición de elementos (L3, 2013, p. 287)

En la Figura 3.51 se plantea un experimento de extracción de bolas de colores de una urna, en donde a través de una representación icónica de la urna y las bolas de colores, los estudiantes deben responder una serie de cuestiones que plantean distintas extracciones, con diferentes condiciones y con la extracción de elementos con o sin reemplazamiento.

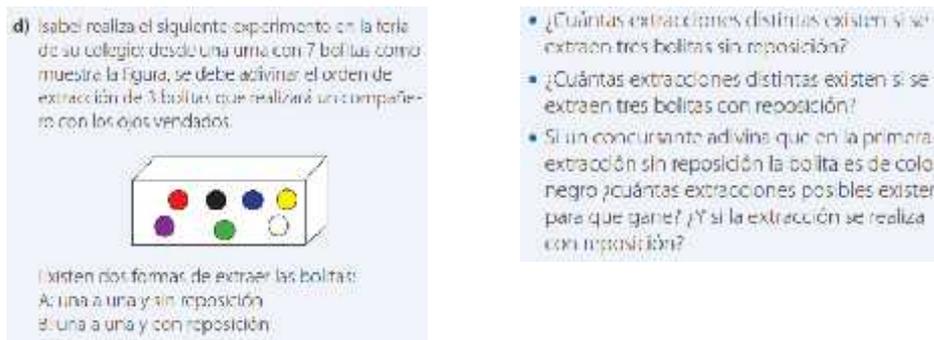


Figura 3.51. Muestras con y sin remplazamiento (L3, 2013, p. 287)

J) *Tamaño de la muestra.* Corresponde a la cantidad de elementos que compone la muestra y a igual que otros conceptos no se define explícitamente, pero se introduce mediante su uso en diferentes ejercicios. Por ejemplo, la Figura 3.52 propone determinar todas las muestras posibles de un tamaño dado de acuerdo a los elementos de diferentes conjuntos, además, se pide calcular la media muestral de cada muestra extraída.

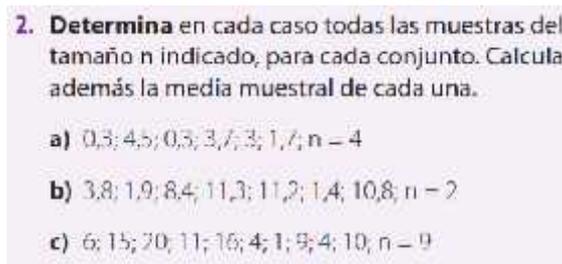


Figura 3.52. Tamaño de la muestra (L4, 2013, p. 278)

J) *Media Muestral.* Corresponde al promedio de los valores obtenidos de una muestra. En la Figura 3.53 se plantea una situación en donde los estudiantes deben seleccionar distintas muestras y luego calcular la media muestral, para finalmente generalizar que “se espera, para un tamaño de muestra no demasiado pequeño, el valor de la media muestral (\bar{x}) sea cercano a la media poblacional (anotada como μ), es decir, el promedio de toda la población” (L4,2013, p.276).

Por otro lado, en las figuras 3.63 y 3.64 también apreciamos una situación problema en donde se les solicita a los estudiantes calcular la media muestral de distintas muestras de una población finita.

En un colegio se asigna a cada uno de los alumnos que los evaluadores tienen por curso 13 evaluaciones en el siguiente cuadro:

$$40 - 35 - 41 - 36 - 44 - 37 - 46 - 45 - 38 - 30 - 41 - 32 - 31 - 45 - 41 - 32 - 35 - 39 - 38 - 39 - 41 - 30 - 32 - 33 - 34 - 38 - 37 - 48 - 35 - 27 - 40 - 32 - 39 - 35 - 30 - 45 - 38 - 25 - 35 - 38 - 4 - 36 - 31 - 43 - 37 - 34 - 34 - 37 - 34 - 37 - 39 - 25 - 6 - 24 - 39 - 30 - 23 - 22 - 15 - 10 - 15 - 20 - 12$$

A partir de esto se calcula el promedio o media poblacional de los datos obtenidos (esto le corresponde a 51,98%) de calificaciones por curso.

1. Completen la tabla escogiendo cinco muestras de tamaño 3 y calculen la media aritmética de cada una de las muestras.
2. Realicen en sus cuadernos una tabla como la anterior, pero esta vez escojan 5 muestras de tamaño 10, 15 y 20.
3. Calculen la media muestral para cada una de las muestras de tamaño 10, luego para las de tamaño 15 y finalmente para las de tamaño 20.
4. Comparan las medias muestrales obtenidas en los puntos 1 y 3. ¿Qué sucede con los valores de las medias muestrales a medida que aumenta el tamaño de la muestra?
5. ¿Qué relación observan entre los valores de las medias muestrales a medida que el tamaño de la muestra aumenta y la media poblacional de la población de calificaciones?

Tabla (ejemplar)					
Muestra	1	2	3	4	5
Def. 1					
Def. 2					
Def. 3					
Media muestral					

Figura 3.53. Media muestral (L4, 2013, p. 284)

Como podemos ver en la Tabla 3.9, en L1 distinguimos conceptos relacionados con el muestreo, así como también experimentos aleatorios y cálculo de probabilidades. En L2 hallamos una mayor cantidad de conceptos ligados al cálculo de probabilidades. En L3, encontramos conceptos combinatorios, cálculo de probabilidades y los relacionados a medias muestrales y poblacionales. En L4 se desarrollan mayormente los conceptos relacionados al muestreo aleatorio y los tipos de muestras. En L5 se aborda el cálculo de probabilidades. En L6 solo se presentan temas relacionados a la distribución normal. Finalmente, en L7 distinguimos también un mayor empleo de conceptos asociados al cálculo de probabilidades y en L8 además de trabajar conceptos combinatorios también aborda algunos de probabilidad.

Tabla 3.9. Conceptos analizados en los libros de texto

Conceptos	L1	L2	L3	L4	L5	L6	L7	L8
Conceptos combinatorios			x	x				x
Experimento aleatorio	x	x	x	x			x	
Espacio muestral	x	x	x	x			x	
Sucesos, sucesos equiprobables	x	x	x	x			x	
Probabilidad clásica	x	x	x	x			x	x
Probabilidad frecuencial	x		x		x		x	x
Población	x		x				x	
Muestra	x		x	x			x	
Tipos de muestras				x				
Muestras ordenadas			x	x				x
Muestras con reemplazamiento			x	x				x
Tamaño de muestra	x			x				
Media Muestral				x				

3.7. PROPIEDADES

De acuerdo a Godino, Font, Contreras y Wilhelmi (2006) las propiedades son relaciones entre objetos que los enriquecen y tienen naturaleza institucional, pues vienen

fijadas en la matemática. En los textos analizados, las propiedades, se presentan de manera informal y mayormente utilizando significados parciales de las mismas, o bien se expresan mediante fórmulas. Usualmente solo se incluyen los enunciados y muy pocas demostraciones. Distinguimos y describimos las siguientes.

) *Principio multiplicativo.* Este principio es esencial en combinatoria, porque permite determinar el número posible de muestras de una población sin tener que listarlas. Se expresa en la forma siguiente: “Si la realización de un proceso se divide en k etapas y cada etapa se puede realizar de n_1, n_2, \dots, n_k formas, entonces todo el proceso se puede realizar de $n_1 \cdot n_2 \cdot \dots \cdot n_k$ ” (L3, 2013, p.281). En la Figura 3.54, se presenta una situación problema, en donde se debe aplicar la noción de principio multiplicativo, para determinar las diferentes formas en que los 4 equipos de tenis queden ubicados en la tabla de posiciones

a) En un torneo de tenis se tienen cuatro equipos, A, B, C y D que se disputan el tercer y cuarto lugar. ¿De cuántas formas posibles estos equipos pueden quedar ubicados en la tabla de posiciones?

Figura 3.54. Principio multiplicativo (L3, 2013, p. 286)

) *Equiprobabilidad de sucesos elementales.* Esta propiedad permite calcular probabilidades mediante la regla de Laplace, como se muestra en la Figura 3.42, se describe en forma sencilla a los estudiantes “Si un experimento cualquiera puede dar lugar a un número finito de resultados posibles y no existe razón que privilegie un resultado por sobre otros, es decir, los resultados son equiprobables (todos poseen la misma probabilidad de ocurrir) ...” (L3, 2013, p. 294)

La mayoría de las situaciones de azar mostradas en los textos emplean el concepto de equiprobabilidad, lo cual, aunque es comprensible por ser situaciones más sencillas puede dar lugar a que los estudiantes adquieran el sesgo de equiprobabilidad según Godino, Batanero y Cañizares (1988). Sin embargo, en los libros analizados se insiste en mostrar ejemplos donde no se aplica, para evitar este sesgo. En la figura 3.55, se le solicita al estudiante analizar si los experimentos son o no equiprobables. En la Figura. 3.40 vemos también un suceso equiprobable como es el lanzamiento de dos monedas.



Figura 3.55. Experimentos equiprobables (L1, 2015, p. 344)

Además, podemos apreciar que en la Figura 3.56, se propone una actividad de activación de conocimientos previos, ya que al inicio de la unidad se proporciona una definición de “sucesos equiprobables” y se les solicita a los estudiantes identificar para cada experimento los eventos equiprobables.

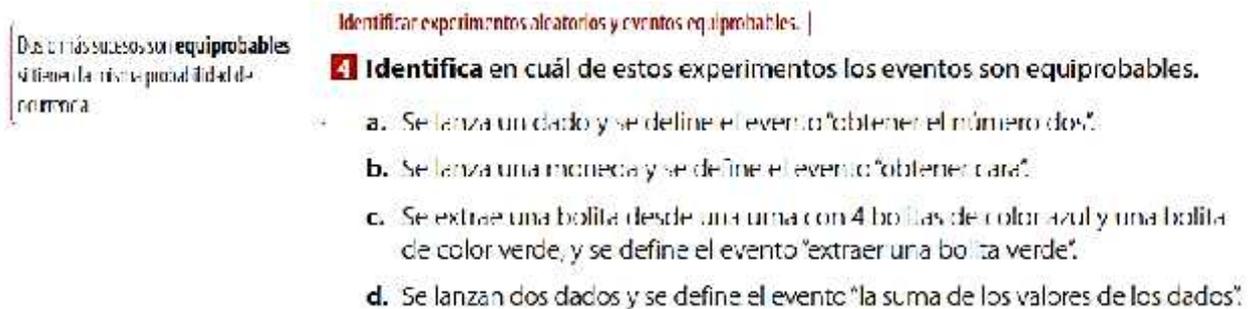


Figura 3.56. Eventos equiprobables (L3, 2013, p. 247)

) *Regla de Laplace.* Una vez introducido el principio de equiprobabilidad, en L1 y L3 se presenta el cálculo de probabilidades mediante la regla de Laplace, en la que la probabilidad se presenta mediante el cociente entre casos favorables y posibles, como se muestra en la Figura 3.43. En las Figuras 3.40 y 3.41, se describen situaciones que ejemplifican el uso de dicha regla.

) *Ley empírica de los grandes números.* Como señala Ortiz (2002), en la vida diaria es frecuente poder observar ejemplos de convergencia empírica de la frecuencia relativa hacia la probabilidad teórica de un suceso. En el nivel escolar que tratamos la ley de los grandes números se incluye sólo informalmente; en los textos analizados también

encontramos enunciada la Ley de los grandes números, como vemos en la Figura 3.57.

Ley de los grandes números

La ley de los grandes números establece que la frecuencia relativa de un suceso tiende a estabilizarse en torno a un número a medida que la cantidad de veces que se realiza un experimento aleatorio crece indefinidamente.

$$f_n(A) = \frac{\text{número de casos favorables al suceso } A}{\text{número de casos posibles}} \rightarrow P(A)$$

Donde $P(A)$ denota la probabilidad de ocurrencia del suceso A .

Figura 3.57. Ley de los grandes números (L4, 2013, p. 285)

Además, se proponen actividades a desarrollar con la ayuda de la hoja de cálculo Excel, para aplicar la simulación y mediante ella observar dicha propiedad (ver figura 3.58).

4. Utiliza Excel/ para responder.

En el experimento A: lanzar una moneda, hay dos posibles resultados: cara (C) o sello (S).

- ¿Qué funciones te permitirán generar 2000 resultados aleatorios?
- ¿A qué valor tienden las frecuencias relativas de cada resultado?

Figura 3.58. Situación problema ley de los grandes números (L4, 2013, p.286)

-) *Población finita o infinita.* En todos los textos analizados, apreciamos que las situaciones problemas estudiadas consideran mayoritariamente poblaciones finitas. Encontramos varios ejemplos en la Figuras 3.2, 3.52, 3.53, 3.63 y 3.64.
-) *Representatividad de una muestra en función del método de extracción.* L1 propone estudiar la representatividad de la muestra de manera cualitativa, señalando que, al escoger una muestra, es necesario que esta sea la adecuada o lo más representativa posible, dependiendo del tipo de información que se quiera recoger. Ello va a depender del método de muestreo. “Por ejemplo, si se quiere estimar el tiempo que demora un estudiante en llegar desde su casa al colegio, no se obtendrán los mismos resultados si se encuestan a tres estudiantes que si se encuestan a 90” (L1, 2015, p.288) o si el método es o no aleatorio. En los problemas planteados en la Figura 3.59, los estudiantes deben determinar si la muestra es representativa de acuerdo a los datos presentados en cada enunciado.

<p>7. Debatan en grupos de cuatro estudiantes (dos a favor y dos en contra) las siguientes situaciones:</p> <p>a. Se desea investigar las dificultades que presentan los alumnos en el aprendizaje de la matemática. Para ello se toma una muestra formada por 60% de alumnos provenientes de colegios particulares, 30% de alumnos de colegios subvencionados y 10% de alumnos de colegios municipales.</p> <p>¿Es representativa la muestra obtenida para la investigación que se desea realizar sobre las dificultades en el aprendizaje de la matemática?</p>	<p>h. Supongamos que para estudiar el presupuesto familiar (ingresos y gastos del hogar) en nuestro país, se toma una muestra formada por 30% de hogares con ingresos bajos, 50% con ingresos medios y 20% con ingresos altos.</p> <p>Si desean responder a cuánto ascienden los gastos de los hogares de acuerdo a su tipo de ingreso (bajo, medio, alto), ¿es representativa la muestra obtenida para responder la pregunta?</p>
---	--

Figura 3.59. Representatividad de la muestra (L1, 2015, p. 291)

) *Variabilidad de la muestra en función del tamaño.* Una de las principales características que caracteriza el pensamiento del estadístico tiene que ver con la valoración de la variabilidad y de las causas de variabilidad, por ejemplo, el tamaño muestral (delMas, 2004). En el transcurso de las tareas sugeridas en la Figura 3.53, se puede apreciar como progresivamente se espera que los estudiantes identifiquen la disminución de la variabilidad inducida por el muestreo al aumentar el tamaño muestral.

) *Aproximación de la frecuencia relativa a la probabilidad mediante la simulación o experimentación.* En la Figura 3.60, vemos problemas que involucran la utilización de la planilla de cálculo para crear modelos de experimentos aleatorios. Se pretende enseñar a los alumnos como la frecuencia relativa, se aproxima a la probabilidad teórica al realizar repeticiones del experimento y como una forma de calcular la probabilidad de un evento.

4. Aplica los pasos anteriores, realiza las siguientes actividades y luego responde.

a) Haz una simulación para 100 lanzamientos de tres monedas. ¿Cuál es la probabilidad experimental para el evento "obtener tres caras"?

b) Haz una simulación para 500 lanzamientos de tres monedas. ¿Cuál es la probabilidad experimental para el evento "obtener tres caras"?

c) Haz una simulación para 1000 lanzamientos de tres monedas. ¿Cuál es la probabilidad experimental para el evento "obtener tres caras"?

d) A medida que aumentas los lanzamientos, ¿a qué valor piensas que se acercará la probabilidad experimental? Corrobora tus conclusiones realizando una simulación para 2000, 3000, 4000 y 5000 simulaciones.

Figura 3.60. Probabilidad a través de la simulación (L3, 2013, p. 300)

En la Tabla 3.10 encontramos las principales propiedades presentes en los textos analizados, L1 propone una mayor cantidad de situaciones problemas que tratan sobre la representatividad de las muestras y otras relacionadas al cálculo de probabilidades, L2 aborda principalmente propiedades de probabilidad, en L3 las propiedades tienen relación con el principio multiplicativo y las de aproximación de la frecuencia relativa a la probabilidad mediante la simulación, en L4, se profundiza en propiedades de la variabilidad de las muestras. Como hemos mencionado anteriormente, L5 y L6 abordan temas que no son considerados en este estudio, por lo que no mencionamos las propiedades que ahí aparecen. En L7 se abordan una serie de propiedades sobre unión e intersección de eventos, como también, las reglas aditivas y multiplicativas de la probabilidad, que no son relevantes para nuestro estudio. Finalmente, en L8 encontramos nuevamente el principio multiplicativo y algunas propiedades relativas al cálculo de probabilidades.

Tabla 3.10. Propiedades encontradas en los libros analizados

Propiedades	L1	L2	L3	L4	L5	L6	L7	L8
Principio multiplicativo		x	x					x
Equiprobabilidad de sucesos elementales	x	x	x					
Regla de Laplace	x	x	x					x
Ley empírica de los grandes números	x		x	x				x
Población finita o infinita			x	x				x
Representatividad de una muestra en función del método de extracción	x							
Variabilidad de la muestra en función del tamaño				x				
Aproximación de la frecuencia relativa a la probabilidad mediante la simulación o experimentación	x		x	x				

3.8.PROCEDIMIENTOS

Según Gea (2014) al tratar de resolver las situaciones problemas relacionadas con un cierto objeto matemático se necesitan operaciones y técnicas que deben ser aprendidas por el estudiante. En este Capítulo, se han estudiado también los procedimientos, es decir, las operaciones, algoritmos, técnicas de cálculo (Godino, Batanero y Font, 2007, p130). En los textos analizados encontramos los siguientes:

) *Realizar encuestas.* Uno de los principales métodos empleados para la recopilación de datos en la unidad de estadística de los textos analizados, es la realización de encuestas. En la Figura 3.61, se les plantea a los estudiantes elaborar y aplicar una encuesta en el curso al cual pertenecen. Para lo cual deben identificar diferentes

conceptos como la variable de estudio, la viabilidad de realizar la encuesta, el método para seleccionar muestras, resumir y organizar la información obtenida en tablas de frecuencias y obtener conclusiones a partir de los datos obtenidos.

Taller 2 El estudio estadístico

En grupos de 4 personas realicen la siguiente actividad:

Elijan un tema que les guste investigar sobre los estudiantes de su curso y elaboren una encuesta dirigida a obtener información sobre ese tema (ejemplo: tipo de música que escuchan).

1. Indiquen la variable en estudio.
2. ¿Qué aspectos deben considerar al momento de realizar la investigación?
3. Describan el método o criterio que utilizaron para seleccionar a los alumnos que serán encuestados.

4. Resuman y organicen la información obtenida en una tabla como siguiente:

Tipo de música	Cantidad de estudiantes
Rochota	
Electrónica	
Pop	
Otro/a	

5. ¿Qué pueden inferir de acuerdo a su encuesta y las realizadas por compañeros sobre los estudiantes de su curso?

Figura 3.61. Realizar Encuestas. (L1, 2015, p. 285)

) *Representar la distribución de frecuencias de los datos de una muestra mediante tablas o gráficas.* Este procedimiento se presenta en la muestra de textos analizada como una herramienta para el cálculo de probabilidad frecuencial. En la Figura 3.16 podemos ver dos tablas de frecuencias para datos agrupados en intervalos, con sus respectivas frecuencias absolutas. Se les pide a los estudiantes calcular las frecuencias relativas de cada conjunto de datos.

) *Construir un diagrama de árbol.* Como describimos anteriormente en la Sección 3.5.5, el diagrama de árbol se utiliza para obtener todos los resultados posibles de un experimento aleatorio, o todas las muestras posibles. En la Figura 3.62, se les solicita a los estudiantes completar el diagrama de árbol para las diferentes combinaciones que se pueden formar (con y sin sentido), de la palabra RIO. Además, se plantean una serie de cuestiones relativas o a la selección de determinados elementos del conjunto de datos, como también, la aplicación del principio multiplicativo.

4. **Completa** el diagrama de árbol de acuerdo con la situación dada, y luego responde las preguntas.

Las letras de la palabra RIO, se pueden combinar para escribir distintas palabras con o sin sentido y sin que se repitan según el siguiente diagrama de árbol:

a) ¿Cuántas palabras sin importar su sentido, se pueden formar con la palabra RIO?

b) ¿Cuántas palabras con sentido se pueden formar con la palabra RIO?

c) Si se agrega la letra S a la palabra RIO, ¿cuántas palabras con o sin sentido y sin que se repitan se podrían formar? Comprueba con el principio multiplicativo.

Figura 3.62. Diagrama de árbol. (L3, 2013, p. 282)

) *Obtener todas las posibles muestras de una población finita.* En la Figura 3.2, como en la figura 3.63, se les solicita a los estudiantes determinar todas las posibles muestras de una población finita para un tamaño dado de la muestra.

2. Determina en cada caso todas las muestras del tamaño n indicado, para cada conjunto. Calcula además la media muestral de cada una.

a) 0,3; 4,5; 0,3; 3,7; 3; 1,7; $n = 4$

b) 3,8; 1,9; 8,4; 11,3; 11,2; 1,4; 10,8; $n = 2$

c) 6; 15; 20; 11; 16; 4; 1; 9; 4; 10; $n = 9$

Figura 3.63. Obtener todas las muestras de tamaño n (L4, 2013, p. 278)

) *Calcular el número de muestras diferentes de una población finita.* En la Figura 3.64, se entregan los datos de una población finita de 10 elementos, y se les solicita a los estudiantes determinar 5 muestras de distintos tamaños, Además de calcular la media muestral.

3. Determina 5 muestras de distinto tamaño del conjunto $X = \{7, 8, 2, 4, 5, 12, 2, 0, 10, 12\}$, y calcula en cada caso la media muestral.

Figura 3.64. Calcular muestras diferentes de una población finita (L4, 2013, p. 275)

) *Estimar una proporción, media o frecuencia esperada en una muestra conocida la población.* En el ejemplo de la Figura 3.65, se conoce el resultado de evaluar la variable “promedio en la asignatura de matemática” en todos los estudiantes del curso, por lo tanto, la población de interés es conocida y se desea estimar el valor esperado de dicha variable por medio de una muestra de 5 estudiantes (a) y luego compararla con su verdadero valor (b).

4. En un curso de 20 alumnos los promedios en la asignatura de matemática son:

1,8; 5,3; 6,6; 5,8; 3,5; 4,3; 5,8; 3,7; 5,6; 6,3;
5,2; 4,5; 5,4; 6,1; 5,4; 4,1; 6,3; 6,6; 4,6; 5,7

La profesora desea inferir acerca de la media con una muestra de 5 alumnos:

a. Utiliza muestreo aleatorio simple para escoger una muestra de 5 alumnos. Explica tu procedimiento.

b. Calcula el promedio de todo el curso, y compara este con la media muestral anterior. ¿Es una buena aproximación?

Figura 3.65. Estimar la media conocida la población. (L4, 2013, p. 290)

) *Estimar una proporción, media o frecuencia esperada en una población, conocida la muestra.* En la Figura 3.66, se presenta el ejemplo de una tarea donde se debe seleccionar en qué condiciones la estimación de la media es más cercana a su verdadero valor poblacional.

- b. Se extraen 5 muestras de tamaño 3 desde una población de tamaño 10. ¿En cuál de las siguientes situaciones el promedio de las muestras se encontraría más cercana a la media de la población? Justifica tu respuesta.
- Las muestras son de la estatura de los integrantes de un equipo de básquetbol de 10 estudiantes.
 - Las muestras son de la estatura de 10 personas que caminan por la calle.

Figura 3.66. Estimar la media conocida la muestra. (L3, 2013, p. 303)

) *Decidir si un método de muestreo es adecuado o diferenciar tipos de muestreo.* En la Figura 3.67, se plantea una actividad de investigación sobre algunos tipos de muestreo, para que el estudiante se familiarice con los términos y busque ejemplos para poder diferenciar dichos tipos de muestreo, ya que principalmente, L4 describe y profundiza en el muestreo aleatorio simple.

- 10. Conexiones:** Averigua en qué consisten los siguientes métodos de muestreo.
- Aleatorio sistemático.
 - Aleatorio estratificado.
 - Aleatorio por conglomerados.
- Da, para cada uno, dos ejemplos en que sería mejor utilizar estos tipos de muestreo en lugar del muestreo aleatorio simple.

Figura 3.67. Diferenciar los tipos de muestreo. (L4, 2013, p. 279)

) *Dados los resultados de una muestra y el tamaño de la población decidir la composición más probable de la población.* Como se puede ver en la Figura 3.1, dados los resultados y el tamaño de la población de una muestra, se puede inferir cuál es la composición más probable de la población.

) *Calcular el promedio de todos los promedios de muestras de igual tamaño extraídas desde una población y analizar la relación.* Debido a la expresión matemática que define al promedio, divide a la suma de las observaciones en la cantidad de datos

considerados, si las muestras analizadas tienen el mismo tamaño, calcular el promedio de cada una de ellas y promediarlos entre sí, es equivalente a calcular el valor medio de todas las observaciones juntas como si fueran una sola muestra.

Sean k muestras diferentes de tamaño n , con promedios $\bar{x}_1, \dots, \bar{x}_k$ respectivamente, el valor medio \bar{x} de todas ellas equivale a promediar los kn observaciones por separado. El procedimiento anterior se puede ver ejemplificado en la pregunta (b) de la Figura 3.68, donde se debe analizar la relación entre calcular la media de los promedios de cada una de las semanas (muestras) y compararla con la media de los ahorros de esas cuatro semanas juntas.

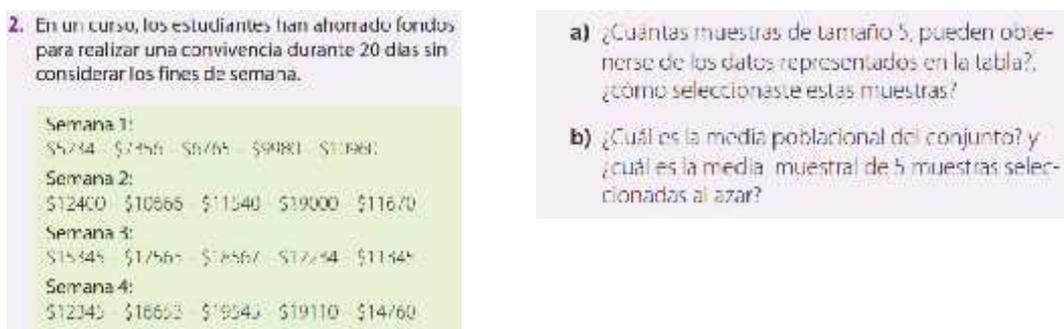


Figura 3.68. Calculan el promedio de todos los promedios de muestras de igual tamaño extraídas desde una población. (L4, 2013, p. 286)

Además, llama la atención el uso de la expresión “media poblacional” en la pregunta (b), ya que en el enunciado del problema no es explícito que el ahorro se realizará únicamente durante 20 días, lo que puede generar confusión con la noción de población cuando años más tarde los estudiantes deban enfrentarse a los procedimientos inferenciales.

-) *Analizar las medias obtenidas de las muestras, cuando el número de datos de las muestras aumenta.* Como se señala en la situación problema de la Figura 3.53, las cuestiones planteadas llevan a comparar los resultados de las medias muestrales para muestras de distintos tamaños, además deben compararse entre sí y con la media poblacional, para así extraer conclusiones sobre la variabilidad en el tamaño muestral.
-) *Determinar las fases de un estudio estadístico.* Cuando se aborda el tema de encuestas, L1 enfoca este método como un estudio estadístico. Como describimos anteriormente, en

la Figura 3.61, se les solicita a los alumnos confeccionar y aplicar una encuesta a su clase, identificando conceptos como la variable de estudio, la viabilidad de realizar la encuesta, el método para seleccionar muestras, resumir y organizar la información obtenida en tablas de frecuencias y obtener conclusiones a partir de los datos obtenidos.

J) *Utilizar herramientas tecnológicas.* En el programa de estudio de 2° año medio, para el uso de las Tecnologías de la Información y la Comunicación (TICs), en la asignatura de matemática, propone que es necesario que los estudiantes “procesen y organicen datos, utilizando plantillas de cálculo, y manipular la información sistematizada en ellas para identificar tendencias, regularidades y patrones relativos a los fenómenos estudiados en el sector” (MINEDUC, 2015, p. 14).

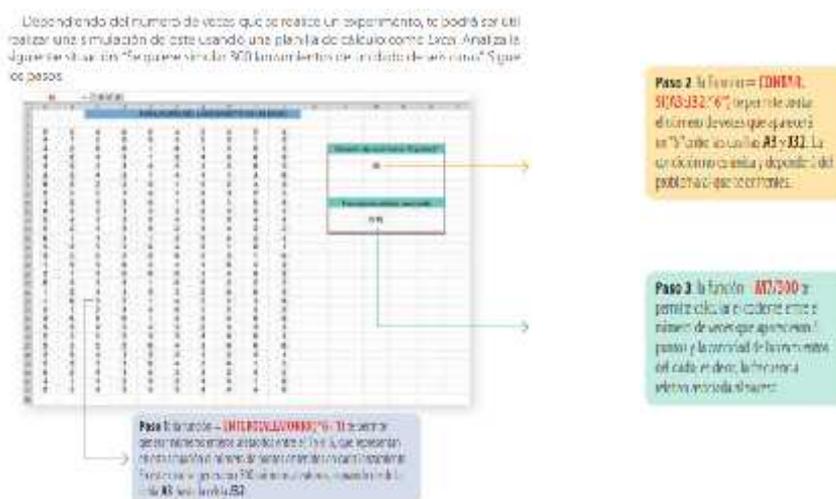


Figura 3.69. Uso de herramientas tecnológicas (L4, 2013, p. 285)

L4 sugiere realizar una simulación de una situación contextualizada usando una hoja de cálculo como Excel. En la Figura 3.69, se les solicita a los estudiantes analizar la siguiente situación: “Se quiere simular 300 lanzamientos de un dado de seis caras y se ejemplifica los pasos a seguir en la planilla de cálculo Excel. Finalmente, se les plantean las siguientes preguntas: “(a) Calcula la frecuencia relativa al suceso A: obtener un punto; (b) ¿A qué número se aproxima la casilla M13 si generas 1000 números aleatorios? Justifica.” (L4, 2013, p. 285)

El uso de las herramientas tecnológicas lo encontramos también en los otros dos textos analizados. En el caso de L1, podemos ver en la Figura 3.70 algunos pasos para realizar la actividad propuesta de la simulación de experimentos aleatorios, utilizando la hoja de cálculo Excel: “A continuación, realicen en parejas los siguientes pasos para simular la suma en el lanzamiento de dos dados. Luego, desarrollen las preguntas de la actividad” (L1, 2015, p. 354). Asimismo, en L3 encontramos actividades propuestas como se ve en las Figuras 3.6 y 3.60 sobre el cálculo de probabilidades a través de la simulación.

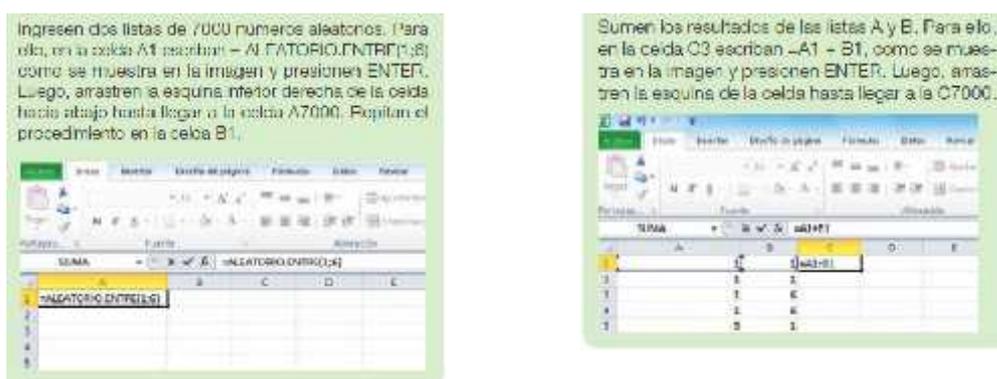


Figura 3.70. Pasos para realizar un experimento aleatorio (L1, 2015, p. 354)

En la Tabla 3.11, vemos los procedimientos que están presentes en los textos. L1 propone una mayor cantidad de situaciones problemas que tienen relación con la realización de encuestas, representar los datos mediante tablas o gráficos, construcción del diagrama de árbol, cálculo de probabilidades. En L2 hallamos procedimientos relacionados con representar la distribución de frecuencias en distintos tipos de gráficos, como también, la construcción de diagrama de árbol. En L3 encontramos procedimientos relacionados con el cálculo de la media, construcción de diagramas de árbol, comparar datos de una muestra con la población y obtener todas las posibles muestras de una población finita.

En L4 los procedimientos tienen relación con la obtención de muestras y los distintos tipos de muestreo. L5 no presenta procedimientos relacionados al muestreo. L6 presenta procedimientos relacionados con estimar una proporción, media o frecuencia esperada en una población, conocida la muestra. En L7 encontramos procedimientos ligados al cálculo de probabilidades y la construcción del diagrama de árbol. Finalmente, en L8 se encuentran procedimientos ligados a la representación de frecuencias en gráficos como la construcción de diagramas de árbol.

La mayoría de los textos analizados incluyen apartados que sugieren el uso de herramientas tecnológicas, como la hoja de cálculo Excel, para realizar simulaciones de las repeticiones de algunos experimentos aleatorios o generar muestras.

El procedimiento *Dada una muestra de resultados, decidir si el método de selección fue aleatorio*, señalado en el análisis curricular no se ha encontrado en un ejemplo dentro de las actividades de los textos escolares analizados.

Tabla 3.11. Procedimientos encontrados en los textos analizados

Procedimientos	L1	L2	L3	L4	L5	L6	L7	L8
Realizar encuestas	x							
Representar la distribución de frecuencias mediante tablas o gráficas	x	x	x					x
Construir un diagrama de árbol (espacio muestral)	x	x	x	x			x	x
Obtener todas las posibles muestras de una población finita				x				
Calcular el número de muestras diferentes de una población finita				x				
Estimar una proporción, media o frecuencia esperada en una muestra conocida la población				x				
Estimar una proporción, media o frecuencia esperada en una población, conocida la muestra			x			x		
Decidir si un método de muestreo es adecuado o diferenciar tipos de muestreo				x				
Dados los resultados de una muestra y el tamaño de la población decidir la composición más probable de la población	x			x				
Calcular el promedio de todos los promedios de muestras de igual tamaño extraídas desde una población y analizar la relación				x				
Analizar las medias obtenidas de las muestras, cuando el número de datos de las muestras aumenta.				x				
Determinar las fases de un estudio estadístico	x							
Utilizar herramientas tecnológicas	x	x	x	x		x		

3.9. ARGUMENTOS

Los argumentos son empleados para explicar y justificar las soluciones de los problemas, pueden involucrar un conjunto de procedimientos, algoritmos, propiedades, definiciones y objetos matemáticos que se articulan entre sí por medio de los razonamientos (Godino, Batanero y Font, 2007). Los estudiantes “deberían abandonar la escuela secundaria capacitados para juzgar la validez de los argumentos basados en datos, como los que aparecen en la prensa” (NTCM, 2000, p.54)

De acuerdo a la investigación realizada por Gómez (2014), existen cuatro tipos de argumentos para analizar textos escolares, los que se exponen a continuación.

Uso de ejemplos y contraejemplos

La forma más frecuente de argumentación presente en los libros de textos es la basada en ejemplos y contraejemplos. Esta argumentación no tiene una validez general, pero refuerza la comprensión del niño. Además, ayuda a desarrollar el pensamiento inductivo (examen de ejemplos) para posteriormente generalizar. La combinación de razonamiento inductivo y deductivo en la clase de matemáticas es recomendada por el NCTM (2000, p. 266).

Así, en la Figura 3.71, se usan ejemplos introducen los conceptos de muestreo aleatorio simple, media muestral y media poblacional. Finalmente, se presenta una síntesis de los conceptos involucrados en el problema anterior.

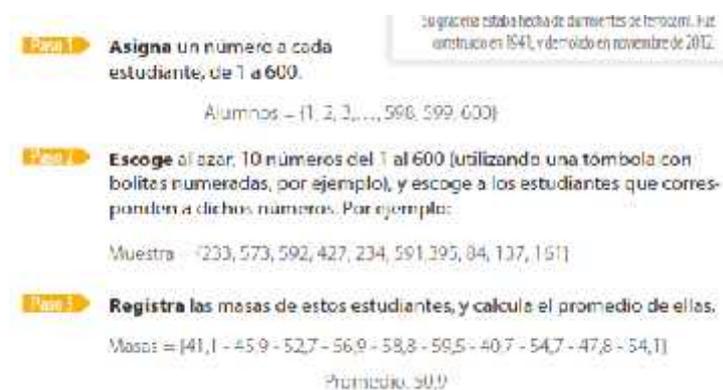


Figura 3.71. Uso de ejemplos para introducir el muestreo aleatorio simple (L4, 2013, p. 276)

En otros casos se parte un caso concreto y a continuación se generaliza una propiedad o definición, pero sin demostrarlo deductivamente. En los libros de texto se presentan algunos contenidos a partir de un ejemplo de una situación cotidiana, conocida por los alumnos, para relacionar el concepto con la definición respectiva. Un ejemplo se presenta en la Figura 3.41, donde se generaliza el cálculo de una probabilidad para un ejemplo particular a la regla de Laplace. Es frecuente en los libros coloquialmente hablar de “probabilidad teórica” para referirse a la probabilidad calculada por la regla de Laplace, porque los niños no llegan a ver nunca la probabilidad como límite de las frecuencias relativas (que también es una probabilidad teórica). Tan solo trabajan con las estimaciones de esta última y los textos la denominan “probabilidad experimental”.

Apoyo gráfico para comprobación de propiedades

En este tipo de argumento, se recurre a la visualización en lugar de usar lenguaje simbólico para discutir la verdad o falsedad de una afirmación o de una propiedad; es

muy frecuente en los textos. Particularmente, se destaca la utilización de los diagramas de árbol para introducir los temas combinatorios (permutaciones, variaciones, combinaciones), para facilitar la comprensión de dichos conceptos, a través de argumentos visuales. Esto ocurre en la Figura 3.72.



Figura 3.72. Variación (L3, 2013, p. 285)

Además de presentar el apoyo gráfico para explicar el concepto de variación, el texto escolar de 1º Medio, sugiere la fórmula del cálculo de la variación a través del principio multiplicativo (ver Figura 3.73). Esquemáticamente se visualizan los términos de la operación combinatoria variación, explicando su significado. Aunque se usa un ejemplo particular, la idea es generalizar a otros casos.

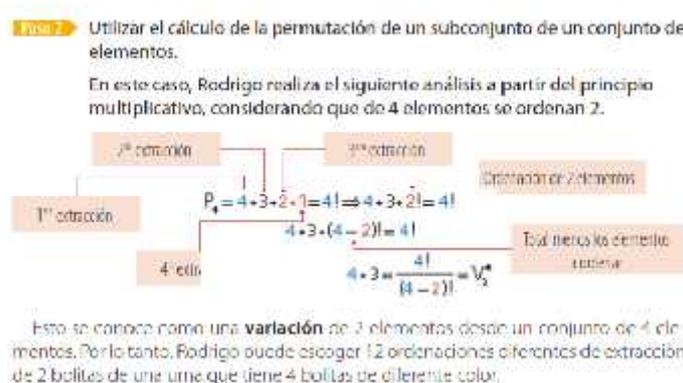


Figura 3.73. Variación, continuación ejemplo (L3, 2013, p. 285)

También podemos apreciar que en la Figura 3.20, se utiliza el diagrama de árbol como una herramienta para el cálculo de probabilidades, detallando los elementos que se deben considerar para lograr todas las combinaciones posibles o el espacio muestral de la situación planteada.

Razonamiento inductivo a partir de datos

Está presente en las situaciones problema derivadas de la interpretación de tablas de distribución de frecuencias, en la Figura 3.74 se puede apreciar la comparación entre

la probabilidad estimada de la frecuencia y su valor obtenido por la regla de Laplace, a través de la información proporcionada en la tabla de frecuencias, como también, los estudiantes deben ser capaces expresar el cálculo de probabilidades al lenguaje gráfico.

Taller

Reunense en grupos de 3 integrantes y realicen la siguiente actividad.

Vicente y Nicole averiguan que la probabilidad teórica de obtener un 3 al lanzar dado es igual a $\frac{1}{6}$ y se preguntaron: ¿cuál es la relación entre la probabilidad experimental y la probabilidad teórica de dicho experimento? Para responder a la pregunta realizaron el experimento 100, 500, 700 y 1000 veces registrando los resultados que aparecen en la siguiente tabla.

Número del dado	1	2	3	4	5	6
Número de lanzamientos	100	200	300	400	500	600
Número de veces que salió el número 3	16	35	47	61	75	82
Frecuencia relativa	0.16	0.175	0.157	0.153	0.15	0.137

1. Calculen la probabilidad experimental de obtener el número 3 en cada lanzamiento y completen la tabla.

Probabilidad experimental o frecuencia relativa	Número de lanzamientos			
	100	200	300	1000

2. Comparación un gráfico con los resultados obtenidos en la pregunta anterior.
3. ¿A qué valor tiende la probabilidad si se aumenta la cantidad de repeticiones experimentales?
4. Si comparan el valor encontrado en la pregunta anterior con la probabilidad teórica, ¿qué sucede? ¿Qué pueden concluir?

Figura 3.74. Comparación entre Probabilidad Frecuencial y Probabilidad Teórica (L3, 2013, p. 299)

Verifican, utilizando herramientas tecnológicas, las conjeturas formulada

En L6, se plantea una actividad en donde los estudiantes utilizarán GeoGebra para crear distribuciones de medias muestrales en parejas. Primero se les solicita abrir el programa, seleccionar la opción “Hoja de Cálculo y Gráficos”, y realizar la serie de indicaciones que se aprecian en la Figura 3.75.

- En la celda A1, escriban **AleatorioEntero(1, 100)** y luego pulsen **Enter**. Aparecerá un número entre 1 y 100. Hágalo azul.
- Seleccionen la celda A1 y, luego, pongan la flecha del mouse sobre el recuadro inferior derecho de color azul. Cuando la flecha se transforme en una cruz, hagan clic y, sin soltar el botón, arrastren el ratón hasta la celda A50. De esta manera aparecerán números aleatorios entre 1 y 100 en todas las celdas entre A1 y A50.
- Seleccionen las celdas A1 a A50 y, luego, pongan la flecha del mouse sobre el recuadro azul que aparece en la parte inferior derecha de la selección. Cuando la flecha se transforme en una cruz, hagan clic y, sin soltar el botón, arrastren el ratón hasta la celda Z50. Este conjunto de números correspondió a la población en estudio.
- En la celda A52, escriban **Media(A1:A50)** y luego pulsen **Enter**. El resultado obtenido corresponde a la media de la muestra conformada por los números de la columna A.
- Seleccionen la celda A52 y pongan la flecha del mouse sobre el recuadro azul que aparece en la parte inferior derecha de la selección. Cuando la flecha se transforme en una cruz, hagan clic y, sin soltar el botón, arrastren el ratón hasta la celda Z52. Así, obtendrán las medias de todas las muestras formadas por los números.
- Seleccionen todas las medias muestrales. Luego, hagan clic en el botón ubicado en la parte superior y, en la ventana que aparece, pulsen **Análisis**. Observarán un histograma de frecuencias con los datos seleccionados, tal como se muestra en la figura de la izquierda. Pueden variar la cantidad de clases del histograma moviendo la barra que se encuentra entre el gráfico.



Figura 3.75. Uso Geogebra (L6, 2013, p. 316)

Posteriormente, deben contestar a una serie de preguntas (ver Figura 3.76), para ver cómo se comportan los datos obtenidos.

2. A partir de lo que han realizado hasta el momento, y de lo que observan en el histograma, respondan las siguientes preguntas.
- ¿Cuál es el tamaño de la población total? ¿cuál es el tamaño de las muestras seleccionadas?
 - ¿Cuántas muestras utilizaron para representar la distribución de medias muestrales? ¿qué porcentaje del total de muestras de ese tamaño, posibles de extraer, utilizaron? Expliquen cómo lo calcularon.
 - ¿Cómo es la distribución de medias muestrales? ¿qué valores tienen más probabilidades de ocurrir? ¿cuáles tienen menos? ¿qué forma tiene la distribución?
 - A partir de lo anterior, ¿qué tipo de distribución creen que tienen las medias muestrales? Justifiquen su respuesta.

Figura 3.76. Conjeturas a resolver. (L6, 2013, p. 316)

En la Tabla 3.12, encontramos los cuatro tipos de argumentos presentes en los textos analizados. En la mayoría de los textos se plantean situaciones problema contextualizadas que permiten el uso de ejemplos y contraejemplos para explicar los distintos conceptos abordados. En L1 y L3 están mayormente presentes los argumentos de apoyo gráfico para comprobación de propiedades, utilizados, por ejemplo, en los distintos conceptos combinatorios o cálculo de probabilidades. L3 y L6 son los textos que presentan una mayor variedad de argumentos, mientras que en L8 predominan los argumentos de uso de ejemplos y contra ejemplos como el razonamiento inductivo a partir de datos. En los primeros seis textos analizados predominan actividades que incluyen la verificación de conjeturas utilizando herramientas tecnológicas, en cambio, L7 y L8, no incluyen este tipo de actividades.

Tabla 3.12. Argumentos encontrados en los textos

Argumentos	L1	L2	L3	L4	L5	L6	L7	L8
Uso de ejemplos y contraejemplos	x	x	x	x	x	x	x	x
Apoyo gráfico para comprobación de propiedades	x	x	x	x		x	x	
Razonamiento inductivo a partir de datos			x		x	x	x	x
Verifican, utilizando herramientas tecnológicas, las conjeturas formulada	x	x	x	x	x	x		

3.10. CONCLUSIONES SOBRE EL ESTUDIO DE LIBROS DE TEXTO

En la extensión de este capítulo se ha realizado un análisis minucioso considerando la presencia de los elementos primarios del marco teórico EOS (objetos primarios), por lo que hemos podido caracterizar cómo se presentan las nociones relacionadas al concepto de muestreo en la serie de libros de textos analizados.

Este análisis complementa el análisis curricular expuesto en el Capítulo 1, permitiéndonos así determinar y caracterizar el significado institucional de referencia de nuestro estudio. Se han analizados las actividades propuestas de ocho textos escolares

chilenos que abordan los diferentes elementos que intervienen progresivamente en el desarrollo del concepto de muestra.

Cabe mencionar que L5, el texto escolar de tercer año medio, aunque se ha incluido en el análisis, no presenta contenidos relacionados al muestreo, ya que los contenidos que se plantean están vinculados a la probabilidad condicional, función de probabilidad y distribución binomial.

En cuanto al lenguaje presente en los textos analizados, se puede mencionar que se observan una gran variedad de tipos de lenguaje, por ejemplo, el lenguaje verbal aporta una orientación a los términos específicos relacionados al muestreo, como también se ha podido identificar aquellos términos comunes que tienen el mismo o distinto significado (ver Tabla 3.3). Además, todos los textos escolares analizados emplean un lenguaje numérico y simbólico bastante completo dependiendo de la complejidad del tema abordado, para ciertos temas se han identificado igualmente la utilización de diferentes tipos de tablas (Tabla 3.7), asimismo, se han encontrado una mayor cantidad de representaciones gráficas (ver Tabla 3.8) que los propuestos en el currículo (ver Tabla 1.3, Capítulo 1).

Los conceptos de muestreo se abordan principalmente en L1, partiendo de los más simples como población y muestra, hasta llegar a abordar algunas propiedades como la representatividad de manera intuitiva. En L3 se consideran algunos conceptos combinatorios relacionados con la extracción de muestras de una población finita y muestras con o sin reposición. En L4 encontramos conceptos como muestra, su tamaño y los tipos de muestra. Con el cambio curricular, podemos señalar que en L7 incluye la comparación de muestras utilizando las nubes de puntos y L8 elimina los conceptos relativos al muestreo pasando a abordar concepto como variable aleatoria, función de probabilidad y distribución, entre otros relacionados al cálculo de probabilidades.

Se destaca además que conceptos como el sesgo en el muestreo y la variabilidad no son abordados por los textos escolares analizados, ya que en las orientaciones curriculares chilenas tampoco están presentes, en cambio, en las directrices internacionales sí son considerados relevantes para la enseñanza del muestreo.

Se han hallado más propiedades relacionadas con conceptos de probabilidad que igual están relacionadas al muestreo, como la equiprobabilidad, la Regla de Laplace y la ley de los grandes números, presentes en los textos escolares que las describen en el

análisis curricular. Además, encontramos principalmente una mayor cantidad de propiedades ligadas al cálculo de probabilidades que en cuanto al muestreo propiamente tal.

En el caso de los procedimientos, se distingue que todos los textos escolares analizados presentan una variedad de procedimientos ligados al muestreo como también al cálculo de probabilidades.

En cuanto a los tipos de argumentos empleados en las actividades presentes en los textos escolares analizados, se puede señalar que el texto escolar de segundo medio (L3, 2013) es el más completo de todos. Señalamos, además, que los libros que incluyen las modificaciones curriculares, no emplean las tecnologías para verificar conjeturas (L7 y L8), a diferencia de las ediciones anteriores en los mismos cursos. El argumento más recurrente en todos los libros analizados es el uso de ejemplos y contraejemplos.

CAPÍTULO 4:

CONSTRUCCIÓN DEL CUESTIONARIO DE EVALUACIÓN

- 4.1. Introducción
- 4.2. Objetivos y clasificación del instrumento
- 4.3. Especificación del contenido de la variable objeto de medición
 - 4.3.1. Fundamentos de la definición de la variable objeto de medición
 - 4.3.2. Tabla de especificaciones del cuestionario
- 4.4. Elaboración de un banco inicial de ítems
 - 4.4.1. Proceso de selección y depuración
- 4.5. Selección de ítems mediante juicio de expertos
 - 4.5.1. Ítems que componen el cuestionario piloto y análisis a priori
- 4.6. Prueba piloto del cuestionario
 - 4.6.1. Descripción de la muestra de estudiantes y contexto
 - 4.6.2. Resultados del estudio piloto
- 4.7. Aproximación al estudio de validez de contenido
- 4.8. Conclusiones sobre el cuestionario

4.1. INTRODUCCIÓN

En este capítulo se detalla la construcción del cuestionario de evaluación de la comprensión del muestreo orientado a estudiantes chilenos de secundaria que ha sido empleado en esta investigación y cuya elaboración ha sido parte de la misma. Se ha seguido un proceso sistemático y riguroso para la construcción de dicho instrumento, como lo indica la metodología propuesta por Batanero y Díaz (2005) y los criterios propuestos para el desarrollo de instrumentos de evaluación en psicología y educación (APA, AERA y NCME, 1999).

Se ha diseñado un cuestionario para medir la comprensión del muestreo en estudiantes de secundaria chilenos, ya que en el análisis de los antecedentes de esta investigación

(Capítulo 2), no se ha encontrado un estudio previo que incluya todas las variables que se consideran relevantes, incluyendo las encontradas en el análisis curricular y de los libros de textos realizado en los capítulos anteriores.

Se describe la construcción de un banco de ítems, que contiene los bloques de problemas que engloban a las variables que comprende el cuestionario, considerando, además, las diferencias de los estudiantes a quienes va dirigido. Se ha realizado esta recopilación a partir del análisis de los antecedentes, los ítems que forman parte de dichos estudios, como también, del análisis de los libros de textos descritos en el capítulo anterior. Una vez seleccionados los ítems, éstos fueron sometidos a juicio de expertos para su análisis y valoración. Dichas sugerencias fueron incorporadas para la mejora de las preguntas incluidas finalmente en el cuestionario.

Posteriormente, el instrumento piloto, se aplicó a una muestra de 148 alumnos chilenos de cuarto año de educación secundaria, para verificar la legibilidad de los ítems y el tiempo necesario para contestarlo. El análisis de las respuestas proporcionadas por los estudiantes ha permitido comprobar las expectativas sobre posibles errores, como también algunas de nuestras hipótesis iniciales.

4.2. OBJETIVOS Y CLASIFICACIÓN DEL INSTRUMENTO

Como se ha mencionado anteriormente, cada vez es más importante para la sociedad el conocimiento de aspectos de la inferencia estadística, como es el caso del muestreo, como componente de los conocimientos necesarios para formar ciudadanos estadísticamente cultos.

En el Capítulo 1, se ha analizado las diferentes directrices curriculares de varios países que han incluido nociones de muestreo en el ámbito de la educación secundaria. Como se trata de un tema que ha tenido una paulatina incorporación en este nivel educativo, los docentes necesitan tener herramientas que apoyen su labor, adicionalmente de lo que ofrecen los libros de textos chilenos, los cuales se analizaron en el Capítulo 3, y que presentan el tema adecuadamente.

Cabe señalar, que en el análisis realizado a las investigaciones descritas en el Capítulo 2, no se ha podido encontrar un instrumento que abarque la comprensión del muestreo en educación secundaria teniendo en cuenta las variables que se han definido tanto por el análisis curricular como con el análisis de los textos escolares, por lo que se ha diseñado un nuevo instrumento, del cual se describirán a continuación sus objetivos.

Objetivos de la construcción del instrumento

Como se ha descrito en el Capítulo 1, el objetivo formulado para la construcción del instrumento, es el siguiente:

Objetivo 2. Construir un instrumento válido y fiable para evaluar la comprensión del muestreo en los alumnos chilenos de Educación Secundaria, siguiendo unos criterios metodológicos rigurosos.

La necesidad de elaborar el instrumento se justifica por no haber encontrado uno en los antecedentes que sirva para evaluar la comprensión del significado de referencia fijado para el muestreo. Es decir, este instrumento es un valioso producto y aporte de esta investigación, que será útil para los docentes en el aula, como para otros investigadores que pueden emplearlo en nuevos estudios.

Clasificación del instrumento

Se ha elaborado un instrumento de medición que, mediante las respuestas de los alumnos a cada ítem planteado, proporcionará un acercamiento al conocimiento que ellos alcanzan, permitiendo así acceder a datos que no pueden ser verificados empleando otro tipo de observaciones (Barbero, 2002). Como se pretende evaluar la comprensión del muestreo que poseen los estudiantes chilenos de 8° básico, 2° medio y 4° medio, se ha caracterizado la manera en que afrontan el trabajo con los conceptos relacionados al muestreo por medio de la inferencia de las respuestas que entregan.

De acuerdo a Hernández, Fernández y Batista (2014), este tipo de instrumento corresponde a un *cuestionario*, que consiste en un conjunto de preguntas respecto de una o

más variables a medir (p. 217). El tipo de preguntas que se ha utilizado, son *preguntas abiertas*, que no delimitan las alternativas de respuesta. Son útiles cuando no hay suficiente información sobre las posibles respuestas de las personas (p. 220) .

4.3. FUNDAMENTOS DE LA DEFINICIÓN DE LA VARIABLE OBJETO DE MEDICIÓN

Como se ha señalado anteriormente, una vez delimitados los objetivos del cuestionario, se analizó la presencia del muestreo en las directrices curriculares chilenas, así como su presencia en los textos escolares analizados, tal como se puntualiza en los Capítulos 1 y 3. Además, en el Capítulo 2, se ha confeccionado una síntesis de las principales investigaciones previas relacionadas al muestreo que han permitido identificar algunos ítems de cuestionarios que podrían ser adaptados y usados en este estudio.

Siguiendo la información proporcionada en los análisis antes mencionados, se ha podido identificar las variables pertinentes para la elaboración y validación del cuestionario. Se han conformado ocho grupos o bloques de preguntas, dónde cada uno de ellos, tiene como objetivo medir alguno de los aspectos relacionados al muestreo, los cuales se describirán a continuación.

- 1) *Comprender la definición del concepto de muestra*: se espera que los estudiantes expliquen con sus palabras qué entienden por “muestra” y señalen algún contexto en que han oído dicha palabra. Además, deben interpretar el concepto de muestra en un contexto donde se conoce la población de sujetos analizados.
- 2) *Detectar sesgo en el método de muestreo*: se les solicita a los estudiantes indicar si el método de selección de muestra aplicando una encuesta es o no apropiado, además de entregar la justificación de su elección. Por otro lado, deben indicar cuál de los métodos de selección de muestras es más apropiado.
- 3) *Sugerir un método de muestreo*: se espera que los alumnos puedan especificar un método para obtener una muestra de un tamaño determinado de una población desconocida.
- 4) *Decidir si una muestra es representativa*: se proponen ejemplos de muestras y los estudiantes deben indicar si consideran o no la muestra representativa y justificar su respuesta.

5) *Identificar la muestra y la población*: dados algunos ejemplos de datos, se les solicita a los alumnos identificar una muestra y la población de cada uno de los datos proporcionados.

6) *Dada una proporción en la muestra determinar la proporción en la población*: se les pide a los estudiantes determinar la proporción de la población dadas la composición de algunas muestras. Además, dada una característica de la población se les pide encontrar la proporción de la muestra, obtener el número total de elementos que conforma la población, también determinar una muestra de los elementos que cumplen cierta característica.

7) *Dada la proporción en una población determinar la proporción en varias muestras*: dada la composición de la proporción de una población finita, se les pide determinar la cantidad de elementos de cada tipo en la población y proponer posibles resultados de la proporción de uno de los tipos de elementos que conforman la población. Por otro lado, dada la proporción de la población, deben señalar están de acuerdo o no con la proporción de la muestra justificando su elección.

8) *Obtener todas las muestras posibles de una población finita*: dada una población con un número finito de elementos, se les pide enumerar la composición de una muestra con reemplazo y otra sin reemplazo.

4.4. ELABORACIÓN DE UN BANCO INICIAL DE ÍTEMS

Para conformar el banco de ítems, se utilizó la revisión de las diversas investigaciones sobre el muestreo que forman parte del Capítulo 2; recociendo y analizando las tareas utilizadas, además de la caracterización de los estudiantes que forman parte de cada investigación y las respuestas obtenidas. En algunos casos, se hizo la traducción y adaptación de algunos ítems.

Se escogieron ítems de respuesta abierta, pues además de proporcionar la solución a las interrogantes planteadas también permiten contemplar los errores probables que pueden cometer los estudiantes.

Se completó esta recopilación, con algunos ítems tomados de los libros de texto analizados en el Capítulo 3, que abordan el muestreo en distintos niveles.

4.4.1. Proceso de selección y depuración

Una vez determinados los objetivos que conformarían los distintos tipos de ítems del cuestionario (ver apartado 4.3 de este Capítulo), se realizó una clasificación de los ítems encontrados en las investigaciones previas como también, los presentes en los textos escolares analizados, determinando los contenidos evaluados y su complejidad de respuesta. El banco de ítems se completó de manera iterativa, ampliando el conjunto de ítems, seleccionando los que cumplieran el objetivo de evaluación para completar dicho conjunto y cuyas situaciones problema pudiesen emplearse ya sea directamente o con modificaciones pequeñas de redacción o formato.

Dicho banco de ítems, se ha modificado gradualmente, mejorando la redacción y claridad de los enunciados de los ítems, eligiendo contextos comprensibles y variados, empleando datos sencillos y que la información presentada tanto en el enunciado como también en tablas, se exhibiera claramente.

4.5. SELECCIÓN DE ÍTEMS MEDIANTE JUICIO DE EXPERTOS

Acabada la construcción del banco de ítems, se realizó el procedimiento de juicio de expertos, en donde se validó la selección de los 11 ítems que conforman el cuestionario final.

La validez de expertos es el grado en que un instrumento realmente mide la variable de interés, de acuerdo con expertos en el tema. (Hernández, Fernández y Batista, 2014, p. 204)

En la Tabla 4.1, se resume la composición final del cuestionario, señalando cada objetivo (variable u objeto de medición), el número de ítem y la referencia a la fuente de donde se adaptó cada pregunta, incluyendo las de elaboración propia.

En el apartado 4.5.1 de este Capítulo, se presentan cada uno de los once ítems que conforman el cuestionario definitivo. Como se puede apreciar en la Tabla 4.1, se han conformado ocho grupos de preguntas, donde cada uno de ellos, tiene como objetivo medir alguno de los aspectos relacionados al muestreo, y contienen más de un ítem, siendo en total evaluadas 30 cuestiones, distribuidas en los ocho grupos de contenidos propuestos, las cuales se describen más detalladamente en el siguiente apartado.

Tabla 4.1. Investigaciones de referencia.

Grupo	Objetivo	Ítem	Fuente
1	Comprender la definición del concepto de muestra	1	Watson (2004);
2	Detectar sesgo en el método de muestreo	2	Meletiou-Mavrotheris y Paparistodemou (2015)
3	Sugerir un método de muestreo	3	Jacobs (1997); Watson (2004); Meletiou-Mavrotheris y Paparistodemou (2015)
4	Decidir si una muestra es representativa	4	Elaboración propia
5	Identificar la muestra y la población	5	Del Valle, Muñoz y Santís (2013), p. 247
6	Dada una proporción en la muestra determinar la proporción en la población	6	Muñoz, Rupin y Jimenez (2013), p. 275,
7	Dada la proporción en una población determinar la proporción en varias muestras	7	Merino, Muñoz, Pérez, y Rupin (2015), p.289
		8	Gómez (2014)
		9	Shaughnessy, Canada y Ciancetta (2003).
		10	Shaughnessy, Ciancetta y Canada (2004)
			García-Ríos (2013).
			García-Ríos y Sánchez (2015)
8	Obtener todas las muestras posibles de una población finita	11	Elaboración propia

4.5.1. ÍTEMS QUE COMPONEN EL CUESTIONARIO PILOTO Y ANÁLISIS A PRIORI

A continuación, se presentan los ítems que componen la prueba piloto y las respuestas que se esperan de los estudiantes. En cada uno de los ítems se describe la solución correcta esperada, las prácticas matemáticas que conlleva, un análisis semiótico de los objetos matemáticos implicados en su solución y los errores frecuentes.

Ítem 1. ¿Has oído la palabra *muestra* antes? Explica qué significa para ti.

En el Ítem 1, se espera que el estudiante sea capaz de explicar con sus palabras que es lo que entiende por “*muestra*” y mencionar algún contexto en el cual ha oído dicho concepto. En este caso, pueden contestar positivamente y señalar contextos de tipo social, profesional, etc., responder negativamente o no dar respuesta al ítem.

Este ítem, ha sido adaptado de Watson (2004), de una de las preguntas que conforma el protocolo de entrevista, como parte complementaria a un cuestionario escrito. Se ha considerado relevante la selección de este ítem porque puede proporcionar alguna idea sobre las concepciones que poseen los estudiantes sobre el concepto de muestra.

Responder la pregunta del Ítem 1, conlleva las siguientes prácticas matemáticas:

1. Leer el enunciado del problema.
2. Identificar el concepto pedido (muestra).
3. Recordar el significado del concepto muestra.
4. Reconocer o no el significado del concepto muestra.
5. Dar algún ejemplo contextualizado del concepto muestra.

En la Tabla 4.2 se ha confeccionado un análisis semiótico de las prácticas descritas y la variedad de objetos matemáticos implicados en la solución de esta pregunta.

Tabla 4.2. Objetos matemáticos ligados a la solución correcta del Ítem 1

Tipos de objetos	Objetos matemáticos
Situaciones problema	Comprender la definición del concepto de muestra
Lenguaje	Palabras del enunciado
Conceptos	Muestra
Propiedades	No se solicita propiedades en este ítem
Procedimientos	Leer e interpretar el enunciado Elaborar la respuesta a la pregunta planteada
Argumentos	Uso de ejemplos

Watson y Moritz (2000a, 200 b), proponen un análisis de las respuestas a una pregunta similar, en cuatro niveles, empleando la Taxonomía SOLO (Biggs y Collis, 1991). En el primero de ellos, los estudiantes no manifiestan un dominio de la tarea para este nivel, por ejemplo, dan ideas vagas o respuestas no relacionadas a lo que se les pregunta. En el segundo, al describir una muestra, por ejemplo, los estudiantes dan una característica única, que a veces involucra un ejemplo o una característica que define a una muestra como “un poco” o “algo para probar”.

En cambio, en el tercer nivel, los alumnos que demuestran una comprensión parcial del concepto de muestra, mencionan varios aspectos asociados con la tarea establecida, generalmente en secuencia. Respuestas a “¿Qué es una muestra?” incluyen una referencia al “todo”, es decir, que la muestra es una parte o una prueba, por ejemplo, mencionan que una muestra es “un poco de algo, no toda la cosa sino una pequeña parte de ella”. En el cuarto nivel, que corresponde a la respuesta correcta, los estudiantes se refieren a todos los elementos necesarios para resolver la tarea, describiendo que una muestra es una pequeña parte que representa a un todo. Por ejemplo, mencionan “una muestra es una pequeña porción de algo más grande”.

Ítem 2. En una investigación sobre los hábitos de ejercicio de los “estudiantes de secundaria chilenos”, algunos investigadores entrevistaron a una *muestra* de ellos.

1. ¿Qué significa la palabra *muestra* en esta frase?
2. ¿Por qué piensas que los investigadores seleccionan una *muestra* de estudiantes en vez de preguntar a todos?
3. ¿Estás de acuerdo con que los investigadores seleccionen una muestra de 10 estudiantes? ¿Cuántos elegirías tú?

En este ítem, se les solicita a los estudiantes aplicar el concepto de muestra y su representatividad a una situación contextualizada, con la finalidad de evaluar si comprenden o no el concepto de muestra. Como respuesta a la pregunta 1, los alumnos pueden indicar que una muestra es “un subconjunto de la población” o que corresponde a “un grupo de estudiantes de secundaria chilenos” o en su defecto, no entregar respuesta.

En el caso de la respuesta a la pregunta 2, dentro de las razones pueden indicar: “porque es difícil preguntar a todos”; “porque es un proceso costoso”; “porque lleva mucho tiempo” o no responden. Finalmente, como respuestas a la tercera pregunta, en el primer apartado pueden indicar: si, no o no contestar y en el segundo apartado dan un número en particular (por ejemplo: cantidad mayor a la muestra dada; o señalan el 50% de los estudiantes; o pueden considerar a todos los estudiantes).

El Ítem 2 ha sido adaptado de la investigación de Meletiou-Mavrotheris y Papanastasiou (2015), basado en un trabajo previo de Watson y Moritz (2000a, 2000b). Se ha considerado que este ítem permite observar las concepciones de los estudiantes respecto al concepto de muestra y su representatividad en una situación problema contextualizada.

Responder a la pregunta 1 de este ítem implica las siguientes prácticas matemáticas:

1. Leer el enunciado del problema.
2. Identificar el concepto pedido (muestra).
3. Recordar el significado del concepto muestra.
4. Relacionar el concepto muestra con el contexto de la situación planteada.
5. Describir que entienden por muestra en el contexto señalado.

En la Tabla 4.3 se ha confeccionado un análisis semiótico de las prácticas descritas y la variedad de objetos matemáticos implicados en la solución de esta pregunta.

Tabla 4.3. Objetos matemáticos ligados a la solución correcta de la pregunta 1 del Ítem 2

Tipos de objetos	Objetos matemáticos
Situaciones problema	Comprender la definición del concepto de muestra
Lenguaje	Palabras del enunciado
Conceptos	Encuesta Muestra
Propiedades	No se solicita propiedades en esta pregunta
Procedimientos	Leer e interpretar el enunciado Elaborar la respuesta a la pregunta planteada
Argumentos	No se solicitan argumentos en esta pregunta

Responder a la pregunta 2 de este ítem implica las siguientes prácticas matemáticas:

1. Leer el enunciado del problema.
2. Identificar el concepto (muestra).
3. Recordar el significado del concepto muestra.
4. Relacionar el concepto muestra con el contexto de la situación planteada.
5. Indicar como argumento las razones de factibilidad para realizar una encuesta seleccionando una muestra de la población objeto del estudio.

En la Tabla 4.4 se ha elaborado un análisis semiótico de las prácticas descritas y la variedad de objetos matemáticos implicados en la solución de esta pregunta.

Tabla 4.4 Objetos matemáticos ligados a la solución correcta de la pregunta 2 del Ítem 2

Tipos de objetos	Objetos matemáticos
Situaciones problema	Comprender la definición del concepto de muestra
Lenguaje	Palabras del enunciado
Conceptos	Encuesta Muestra
Propiedades	No se solicita propiedades en esta pregunta
Procedimientos	Leer e interpretar el enunciado Elaborar la respuesta a la pregunta planteada
Argumentos	Indicar las razones de factibilidad para realizar una encuesta

Responder a la pregunta 3 de este ítem implica las siguientes prácticas matemáticas:

1. Leer el enunciado del problema.
2. Identificar el concepto (muestra).
3. Recordar el significado del concepto muestra.
4. Relacionar el concepto muestra con el contexto de la situación planteada.
5. Indicar si están o no de acuerdo con un tamaño de muestra dado.
6. Proponer una muestra representativa de la población dada.

Tabla 4.5. Objetos matemáticos ligados a la solución correcta de la pregunta 3 del Ítem 2

Tipos de objetos	Objetos matemáticos
Situaciones problema	Comprender la definición del concepto de muestra
Lenguaje	Palabras del enunciado
Conceptos	Encuesta Muestra Representatividad de la muestra
Propiedades	No se solicita propiedades en esta pregunta
Procedimientos	Leer e interpretar el enunciado Proponer una muestra representativa de la población Elaborar la respuesta a la pregunta planteada
Argumentos	Indicar que la muestra propuesta no es representativa porque es muy pequeña

En la Tabla 4.5 se presenta un análisis semiótico de las prácticas descritas y la variedad de objetos matemáticos implicados en la solución de esta pregunta.

Meletiou-Mavrotheris y Papanastasiou (2015) consideran las categorías anteriormente descritas en el Ítem 1, descritas por Watson y Moritz (2000a) para clasificar las respuestas de la primera pregunta. Como respuesta correcta, en el nivel 1, consideran afirmaciones como la siguiente: *“te dan una pequeña parte del producto que anuncian para que pruebes. Si te gusta, entonces compras todo el producto”*; en el nivel 2, consideran afirmaciones como: *“No les preguntaron a todos los niños de la escuela primaria en Chipre, solo seleccionaron unos pocos para obtener una muestra”*; en el nivel 3, mencionan ideas como: *“Una muestra es un elemento que se utiliza solo para fines de prueba”*, finalmente, en el nivel 4, el que entrega respuestas incorrectas, o no hay respuesta o indican “No sé”.

En la pregunta 2, donde los estudiantes tuvieron que explicar por qué los investigadores usaron una muestra, las respuestas reflejan problemas prácticos y conocimiento del entorno social en el que se realiza el muestreo: *“Sería demasiado lento y tedioso visitar todas las escuelas primarias de Chipre y entrevistar a todos los niños”* (Meletiou-Mavrotheris y Papanastasiou, 2015).

En la tercera pregunta, se utilizó la jerarquía propuesta por Watson y Moritz (2000a), que también es utilizada en un trabajo posterior de Watson (2004) para caracterizar la comprensión de los estudiantes sobre conceptos de muestreo (ver Tabla 2.1, Capítulo 2). Algunas de las respuestas se clasifican, por ejemplo, en la categoría 1 aceptando un tamaño de muestra inferior a 15, en la categoría 4, sugirieron tamaños de muestra grandes, dieron justificaciones estadísticamente ingenuas y un grupo menor estaban en la categoría 6, los que

demonstraron una comprensión bastante bien desarrollada del muestreo, haciendo hincapié en la necesidad de una muestra suficientemente grande, y la adopción de métodos aleatorios o estratificados.

Ítem 3. Los estudiantes de un colegio realizaron una encuesta para determinar el porcentaje de niños que reciclan en sus casas:

- Andrea preguntó a 80 estudiantes que son miembros del club de medioambiente.
- Elena envió un cuestionario a todos los niños del colegio y tomó los primeros 80 que contestaron.
- Pedro quiso el mismo número de niños y niñas. Así que preguntó a 5 niños y 5 niñas de cada curso para conseguir los 80 estudiantes que constituyen la muestra.
- Luis tenía los nombres de los 800 estudiantes en la escuela; puso cada nombre en un papel en un sombrero y sacó 80.
- Sara no conocía demasiados niños por lo que decidió encuestar a 80 niñas. Pero quería asegurarse de tener niñas de varias edades, por lo que tomó 10 niñas de cada nivel.

a. Completa la última columna donde preguntamos si el método de cada niño es apropiado.

Niño	A quien pregunta	¿Te parece un método apropiado? ¿Por qué?
Andrea	80 estudiantes del club de medioambiente.	
Elena	Envió un cuestionario a todos los niños del colegio y tomó los 80 que contestaron primero.	
Pedro	5 niños y 5 niñas de cada curso	
Luis	Puso el nombre de cada niño de la escuela (800) en un papel en un sombrero y sacó 80	
Sara	10 niñas de cada nivel.	

b. De acuerdo a los datos entregados en el enunciado, ¿Cuál crees que es la mejor manera de escoger a los niños para estimar el porcentaje de ellos que recicla en casa?

En el Ítem 3, se les pide a los estudiantes señalar si cada uno de los métodos de selección de muestras planteados son o no apropiados e indicar su justificación a la respuesta, además, en el segundo apartado deben indicar cuál de los métodos propuestos es el mejor para la selección de una muestra. En el caso del método de Andrea, en el método de Elena y finalmente en el método de Sara, el estudiante debe indicar que cada uno de esos métodos no son apropiados porque su selección de la muestra es sesgada, en cambio, en el método de Pedro y en el método de Luis deben indicar que el método si es apropiado, ya que el primero corresponde a un muestreo estratificado y el segundo a un muestreo aleatorio simple.

En el apartado b, deben señalar cuál de los métodos propuestos es el más apropiado, en este caso pueden señalar uno de los siguientes métodos: el método de Pedro o el método de Luis.

Para responder a cada uno de los métodos propuestos en la “pregunta a” de este ítem se deben realizar las siguientes prácticas matemáticas:

1. Leer el enunciado del problema y la pregunta planteada.
2. Recordar el significado del concepto muestra y de los métodos de muestreo.
3. Identificar el método de selección de la muestra en cada uno de los métodos propuestos.
4. Relacionar el método de muestreo con el contexto de la situación planteada.
5. Identificar el sesgo en el método de muestreo.
6. Señalar si el método de muestreo propuesto es o no apropiado.
7. Justificar por qué consideran o no apropiado cada uno de los métodos propuestos.

En la Tabla 4.6 se ha elaborado un análisis semiótico de las prácticas descritas y la variedad de objetos matemáticos implicados en la solución de esta pregunta.

Tabla 4.6. Objetos matemáticos ligados a la solución correcta de la “pregunta a” del Ítem 3

Tipos de objetos	Objetos matemáticos
Situaciones problema	Detectar sesgo en el método de muestreo
Lenguaje	Palabras del enunciado
Conceptos	Encuesta Muestra Métodos de muestreo Sesgo Representatividad Variabilidad
Propiedades	No se solicita propiedades en esta pregunta
Procedimientos	Leer e interpretar el enunciado Identificar el tipo de muestreo Identificar el sesgo en el método de muestreo Elaborar la respuesta a la pregunta planteada
Argumentos	Justificar por qué es apropiado o no el método de muestreo

Para responder a la pregunta b de este ítem, se necesitan las siguientes prácticas matemáticas:

1. Leer el enunciado del problema.
2. Recordar el significado del concepto muestra y de los métodos de muestreo.
3. Identificar el método de selección de la muestra en cada uno de los métodos propuestos.
4. Relacionar el método de muestreo con el contexto de la situación planteada.
5. Identificar el sesgo en el método de muestreo.
6. Señalar si el método de muestreo propuesto es o no apropiado.
7. Justificar por qué consideran o no apropiado cada uno de los métodos propuestos.

En la Tabla 4.7 se ha realizado un análisis semiótico de las prácticas descritas y la variedad de objetos matemáticos implicados en la solución de esta pregunta.

Tabla 4.7. Objetos matemáticos ligados a la solución correcta de la “pregunta b” del Ítem 3

Tipos de objetos	Objetos matemáticos
Situaciones problema	Detectar sesgo en el método de muestreo
Lenguaje	Palabras del enunciado
Conceptos	Encuesta Muestra Población Métodos de muestreo Sesgo Representatividad Variabilidad
Propiedades	No se solicita propiedades en esta pregunta
Procedimientos	Leer e interpretar el enunciado Identificar el tipo de muestreo Identificar el sesgo en el método de muestreo Elaborar la respuesta a la pregunta planteada
Argumentos	Justificar la elección del método de muestreo más apropiado

Este ítem ha sido adaptado del trabajo de Meletiou-Mavrotheris y Papanastasiou (2015) y del de Watson y Kelly (2005), enunciado que ha sido adaptado previamente de la investigación de Jacobs (1997), cuyo objetivo es conocer la comprensión de los estudiantes sobre el sesgo del muestreo en un contexto cercano a la realidad, como es la toma de encuestas.

Para el análisis de las respuestas del apartado a del Ítem 3, se proponen las categorías de análisis que se ha empleado en Ruiz-Reyes, Ruz, Molina-Portillo y Contreras (enviado), adaptadas del trabajo de Watson y Kelly (2005). En la categoría “*respuesta estadística apropiada*”, se describe la respuesta correcta a cada método, en “*ideas no centrales o imparcialidad*” se consideran las respuestas que evalúan cierta parte de la información, más específicamente en relación con la imparcialidad y el tamaño de la muestra, en el caso del “*análisis inapropiado*”, fueron críticas inapropiadas al método propuesto, centrándose en la imprecisión percibida del método aleatorio, la injusticia o la imparcialidad, falta de oportunidad y el tamaño pequeño de la muestra y las respuestas de “*lógica inapropiada*” señalan todos los razonamientos erróneos que realizan los estudiantes.

Ítem 4. Supongamos que quieres averiguar el porcentaje de niños que viene a tu escuela en diferentes medios de transporte: caminando, auto, micro, bicicleta, u otro. ¿Cómo tomarías una muestra de 50 niños de modo que los resultados sean representativos de toda la escuela?

Este Ítem corresponde a una elaboración propia, en donde se les pide a los estudiantes sugerir un método de muestreo, conocido el tamaño de la muestra, pero con el tamaño de la población desconocida. Con esta pregunta se pretende que los estudiantes sean capaces de seleccionar una muestra representativa de la población, aunque se desconozca el tamaño de esta última, se presentan los datos en un contexto cercano a la realidad de los estudiantes y con datos simples para facilitar la entrega de una respuesta.

Dentro de las posibles respuestas se encuentran las siguientes:

1. No contesta.
2. Incorrecto. Menciona una selección que no corresponde a un método de muestreo.
3. Menciona un método de muestreo, pero no justifica.
4. Menciona obtener una muestra de un grupo sesgado.
5. Menciona algún método de muestreo y justifica.

Para responder a este ítem se deben realizar las siguientes prácticas matemáticas:

1. Leer el enunciado del problema y la pregunta planteada.
2. Recordar el significado del concepto muestra y de los métodos de muestreo.
3. Identificar un método de selección de la muestra.
4. Relacionar el método de muestreo con el contexto de la situación planteada.
5. Señalar el método de muestreo propuesto.

En la Tabla 4.8 se ha elaborado un análisis semiótico de las prácticas descritas y la variedad de objetos matemáticos implicados en la solución de este ítem.

Tabla 4.8. Objetos matemáticos ligados a la solución correcta del Ítem 4

Tipos de objetos	Objetos matemáticos
Situaciones problema	Sugerir un método de muestreo
Lenguaje	Palabras del enunciado
Conceptos	Encuesta Muestra Métodos de muestreo Representatividad Variabilidad
Propiedades	No se solicita propiedades en esta pregunta
Procedimientos	Leer e interpretar el enunciado Aplicar un tipo de muestreo Elaborar la respuesta a la pregunta planteada
Argumentos	Justificar la elección del método de muestreo

Ítem 5. Analiza si las siguientes muestras son representativas de la población de estudio y completa la tabla con tus respuestas.

Muestra	¿Es representativa?	Explica por qué
1. Se seleccionan al azar 100 estudiantes, de un colegio de 300 alumnos, para calcular el promedio de notas de la población.		
2. Se les pregunta sobre la calidad del servicio del metro a las primeras 100 personas que toman el metro en la mañana.		
3. Se escogen los niños más altos y los más bajos de un curso para determinar el promedio de la estatura de los alumnos.		

El Ítem 5 ha sido adaptado y seleccionado de las situaciones problemas presentes en el texto escolar de primero medio (Del Valle, Muñoz y Santís, 2013, p. 247). El objetivo de este ítem es que los estudiantes puedan determinar si la muestra señalada es representativa o no y se le ha agregado la segunda parte de la pregunta, en donde deben justificar el porqué de su elección.

Tabla 4.9. Respuestas correctas Ítem 5

Muestra	¿Es representativa?	Explica por qué
1	Si	Porque se eligieron al azar.
2	No	Porque se pregunta solo a un sector de usuarios.
3	No	Porque no se considera la variabilidad.

En la Tabla 4.9 se presentan las respuestas correctas del Ítem 5, en donde solo se presenta una situación que plantea una muestra representativa y dos en que las muestras no son representativas de la población.

Para responder a cada una de las preguntas planteadas en este ítem se deben realizar las siguientes prácticas matemáticas:

1. Leer el enunciado del problema y la pregunta planteada.
2. Recordar el significado del concepto muestra y la representatividad de una muestra.
3. Identificar si la muestra es o no representativa.
4. Señalar que la muestra es o no representativa.
5. Justificar por qué consideran o no representativa cada una de las muestras propuestas.

En la Tabla 4.10 se ha confeccionado un análisis semiótico de las prácticas descritas y la variedad de objetos matemáticos implicados en la solución de esta pregunta.

Tabla 4.10. Objetos matemáticos ligados a la solución correcta del Ítem 5

Tipos de objetos	Objetos matemáticos
Situaciones problema	Decidir si una muestra es representativa
Lenguaje	Palabras del enunciado
Conceptos	Muestra Métodos de muestreo Representatividad Variabilidad
Propiedades	No se solicita propiedades en esta pregunta
Procedimientos	Leer e interpretar el enunciado Identificar si la muestra es representativa
Argumentos	Elaborar la respuesta a la pregunta planteada Justificar por qué es representativa la muestra

Ítem 6. Determina en cada caso la población y una posible muestra de ella. Completa la tabla con tus respuestas.		
Datos	Población	Muestra
1. Una fábrica de yogur quiere investigar sobre la calidad de sus productos.		
2. Diego necesita saber el precio de un kilogramo de carne, para una comida familiar.		
3. Ximena estudia respecto del tamaño de las hormigas que habitan en un insectario.		
4. Daniel desea saber si una ciudad cuenta con suficiente lluvia, para realizar una plantación.		

El Ítem 6 se ha tomado del texto escolar de segundo medio (Muñoz, Rupin y Jiménez, 2013, p. 275), en donde se entregan datos de distintas situaciones contextualizadas, y los estudiantes deben ser capaces de identificar la población y la muestra.

En la Tabla 4.11 se presentan las respuestas correctas del Ítem 6, señalando la población y muestra respectiva para cada uno de los datos proporcionados.

Tabla 4.11. Respuestas correctas Ítem 6

Datos	Población	Muestra
1. Una fábrica de yogur quiere investigar sobre la calidad de sus productos.	Todos los yogures producidos.	Un número determinado de cada sabor.
2. Diego necesita saber el precio de un kilogramo de carne, para una comida familiar.	Los precios de los tipos de carne de la carnicería.	Los precios de una variedad de cada tipo de carne (vacuno, pescado, ave)
3. Ximena estudia respecto del tamaño de las hormigas que habitan en un insectario.	Todas las hormigas en el insectario.	Cierta cantidad de hormigas del insectario.
4. Daniel desea saber si una ciudad cuenta con suficiente lluvia, para realizar una plantación.	Todas las ciudades del país.	Una ciudad por región.

Para responder a este ítem se deben realizar las siguientes prácticas matemáticas:

1. Leer el enunciado del problema y la pregunta planteada.
2. Recordar el significado de los conceptos población y muestra.
3. Identificar y señalar la población.
4. Identificar y señalar la muestra.

En la Tabla 4.12 se presenta un análisis semiótico de las prácticas descritas y la variedad de objetos matemáticos implicados en la solución de este ítem.

Tabla 4.12. Objetos matemáticos ligados a la solución correcta del Ítem 6

Tipos de objetos	Objetos matemáticos
Situaciones problema	Identificar la muestra y la población
Lenguaje	Palabras del enunciado
Conceptos	Muestra Población Representatividad
Propiedades	No se solicita propiedades en esta pregunta
Procedimientos	Leer e interpretar el enunciado Identificar la población Identificar la muestra Elaborar la respuesta a la pregunta planteada
Argumentos	Uso de ejemplos

Ítem 7. Se tomaron al azar cuatro muestras de estudiantes de un colegio. A continuación, se presentan los resultados.

Alumnos	35	29	24	32
Alumnas	15	21	26	18

¿Podrías estimar el porcentaje de alumnos en el colegio? Explica tu procedimiento.

Este ítem se ha adaptado de la situación problema planteada en el texto de séptimo básico (Merino, Muñoz, Pérez, y Rupin, 2015, p. 289). Este ítem apunta a evaluar que dada una proporción en la muestra los estudiantes puedan determinar la proporción en la población. En esta pregunta, los alumnos pueden realizar los cálculos de porcentaje con cada una de las muestras y de ahí sacar el promedio de dichos porcentajes, o utilizando otro método que sería calcular el porcentaje de la cantidad total de alumnos (varones). No se espera que expliquen el procedimiento empleado, debido a que están acostumbrados a realizar principalmente cálculos.

Tabla 4.13. Cálculos de la solución correcta al ítem 7

$$\frac{35}{50} = 0,7 = 70\%$$

$$\frac{28}{50} = 0,58 = 58\%$$

$$\frac{24}{50} = 0,48 = 48\%$$

$$\frac{32}{50} = 0,64 = 64\%$$

$$\frac{70 + 58 + 48 + 64}{4} = \frac{240}{4} = 60\%$$

Otro método: Calcular el porcentaje con los totales.

$$\frac{35 + 29 + 24 + 32}{200} = \frac{120}{200} = 0.6 = 60\%$$

Para responder a este ítem se deben realizar las siguientes prácticas matemáticas:

1. Leer el enunciado del problema y la pregunta planteada.
2. Identificar en qué proporción se presenta cada muestra.
3. Calcular el porcentaje de alumnos varones por grupo.
4. Calcular el promedio de los porcentajes en que se presentan los varones por grupo.
5. Señalar el porcentaje de alumnos varones que se espera en la población.

En la Tabla 4.14 se elabora un análisis semiótico de las prácticas descritas y la variedad de objetos matemáticos implicados en la solución de este ítem.

Tabla 4.14. Objetos matemáticos ligados a la solución correcta del Ítem 7

Tipos de objetos	Objetos matemáticos
Situaciones problema	Dada una proporción en la muestra determinar la proporción en la población
Lenguaje	Palabras del enunciado Tabla como listado de datos
Conceptos	Muestra Población Representatividad Proporción Porcentaje
Propiedades	No se solicita propiedades en esta pregunta
Procedimientos	Leer e interpretar el enunciado Identificar las muestras Calcular la proporción de cada muestra Calcular el promedio de los porcentajes de cada muestra Elaborar la respuesta a la pregunta planteada
Argumentos	Señalar el procedimiento llevado a cabo

Ítem 8. En un estanque hay peces, pero su dueño no sabe cuántos. Toma 200 peces y les pone una marca; los devuelve al estanque, donde se mezclan con el resto. Al día siguiente, el dueño toma 250 peces del estanque, y encuentra que 25 de ellos están marcados y el resto no.

1. ¿Qué proporción aproximada de peces del estanque están marcados?
2. ¿Cuál es el número aproximado de peces en el estanque?
3. Si el dueño saca ahora 100 peces más del estanque, ¿cuántos aproximadamente estarán marcados?

Este ítem ha sido adaptado de Gómez (2014, p.202), quien adaptó este problema de un experimento de enseñanza de Fishbein y Gazit (1984). Aunque en la investigación de Gómez (2014) el problema está enmarcado en el uso del significado frecuencial de la probabilidad, se ha considerado que puede aportar información relevante para nuestro estudio, pues se trabaja con la muestra de una población finita. Este ítem también pertenece al bloque 6, cuyo objetivo apunta a evaluar lo siguiente: dada una proporción en la muestra, determinar la proporción en la población.

La respuesta correcta a la pregunta 1 es: $\frac{25}{250} = 0,1 = 10\%$

La respuesta correcta a la pregunta 2 es: $\frac{200 \cdot 100}{10} = 2000$ peces.

La respuesta correcta a la pregunta 3 es: $\frac{100 \cdot 10}{100} = 10$ peces.

Para responder a la pregunta 1 de este ítem se deben realizar las siguientes prácticas matemáticas:

1. Leer el enunciado del problema y la pregunta planteada.
2. Identificar la muestra de peces seleccionada.
3. Identificar la proporción de peces marcados.
4. Calcular por medio de la regla de tres el porcentaje de peces marcados para determinar la proporción aproximada.
5. Señalar el porcentaje obtenido.

En la Tabla 4.15 se realiza un análisis semiótico de las prácticas descritas y la variedad de objetos matemáticos implicados en la solución de la pregunta 1 del Ítem 8.

Tabla 4.15. Objetos matemáticos ligados a la solución correcta de la pregunta 1 del Ítem 8

Tipos de objetos	Objetos matemáticos
Situaciones problema	Dada una proporción en la muestra determinar la proporción en la población
Lenguaje	Palabras del enunciado
Conceptos	Muestra Población Representatividad Proporción Porcentaje
Propiedades	No se solicita propiedades en esta pregunta
Procedimientos	Leer e interpretar el enunciado Identificar la muestra seleccionada Identificar la proporción de la muestra peces marcados Calcular la proporción de peces marcados Elaborar la respuesta a la pregunta planteada
Argumentos	No se precisa de argumentos

Para responder a la pregunta 2 de este ítem se deben realizar las siguientes prácticas matemáticas:

1. Leer el enunciado del problema y la pregunta planteada.
2. Identificar la muestra de peces marcados.
3. Identificar el porcentaje de peces marcados obtenido en la pregunta anterior.
4. Calcular, por medio de la regla de tres, el total de peces en el estanque, a partir de la muestra de peces marcados y su proporción.
5. Señalar el número total de peces obtenido.

En la Tabla 4.16 se elabora un análisis semiótico de las prácticas descritas y la variedad de objetos matemáticos implicados en la solución de la pregunta 2 del Ítem 8.

Tabla 4.16. Objetos matemáticos ligados a la solución correcta de la pregunta 2 del Ítem 8

Tipos de objetos	Objetos matemáticos
Situaciones problema	Dada una proporción en la muestra determinar la proporción en la población
Lenguaje	Palabras del enunciado
Conceptos	Muestra Población Representatividad Proporción Porcentaje
Propiedades	No se solicita propiedades en esta pregunta
Procedimientos	Leer e interpretar el enunciado Identificar la muestra de peces marcados Identificar el porcentaje de peces marcados Calcular el total de peces en el estanque Elaborar la respuesta a la pregunta planteada
Argumentos	No se precisa de argumentos

Para responder a la pregunta 3 de este ítem se deben realizar las siguientes prácticas matemáticas:

1. Leer el enunciado del problema y la pregunta planteada.
2. Identificar la muestra de peces seleccionados.
3. Identificar el porcentaje de peces marcados obtenido en la pregunta 1.
4. Calcular, por medio de la regla de tres, el número de peces marcados en la nueva muestra, a partir de la muestra de peces seleccionada y la proporción de peces marcados.
5. Señalar el número de peces marcados en esta nueva muestra.

En la Tabla 4.17 se realiza un análisis semiótico de las prácticas descritas y la variedad de objetos matemáticos implicados en la solución de la pregunta 3 del Ítem 8.

Tabla 4.17. Objetos matemáticos ligados a la solución correcta de la pregunta 3 del Ítem 8

Tipos de objetos	Objetos matemáticos
Situaciones problema	Dada una proporción en la muestra determinar la proporción en la población
Lenguaje	Palabras del enunciado
Conceptos	Muestra Población Representatividad Proporción Porcentaje
Propiedades	No se solicita propiedades en esta pregunta
Procedimientos	Leer e interpretar el enunciado Identificar la muestra de peces seleccionados Identificar el porcentaje de peces marcados Calcular el total de peces marcados en la nueva muestra Elaborar la respuesta a la pregunta planteada
Argumentos	No se precisa de argumentos

Ítem 9. Piensa que tienes un recipiente con 100 dulces; 60 son de frutilla, y 40 de limón. Los dulces están todos mezclados en el recipiente. Sacas un puñado de 10 dulces y cuentas el número de los dulces de frutilla.

1. ¿Cuántos serán de frutilla?
2. Supongamos que seis de tus compañeros repiten experimento, cada uno de ellos saca 10 dulces. (Después de cada extracción, los dulces se vuelven a colocar y se mezcla nuevamente).

Escribe en cada línea un número probable de dulces de frutilla que obtiene tu compañero.

Niño	Número de dulces de frutilla	Niño	Número de dulces de frutilla
Ana		Pedro	
Camila		María	
José		Juan	

Este ítem ha sido adaptado de la investigación de Shaughnessy, Canada y Ciancetta (2003), sobre la comprensión de los estudiantes de secundaria sobre la variabilidad de la muestra, en la que se les presenta a los alumnos tres tareas estocásticas que involucran

ensayos de repetición, la primera de ellas, pretende medir la predicción de resultados de muestras repetidas de una mezcla de caramelos, ítem que se considera pertinente para el cuestionario.

Aunque este ítem es tomado de una investigación en donde se estudia la variabilidad en la muestra de distintas tareas que implican la repetición de ensayos, se ha considerado adaptarla a nuestro objetivo, que es: dada la proporción en una población, determinar la proporción en varias muestras.

La respuesta a la pregunta 1 es: $\frac{60 \cdot 10}{100} = 6$ dulces de frutilla.

Las posibles respuestas a la pregunta 2:

1. No contesta.
2. Contesta todo igual a 6.
3. Contesta con variabilidad normal.
4. Contesta con variabilidad excesiva.

Para responder a la pregunta 1 de este ítem se deben realizar las siguientes prácticas matemáticas:

1. Leer el enunciado del problema y la pregunta planteada.
2. Identificar la proporción de los dos tipos de dulces.
3. Calcular, por medio de la regla de tres, el número de dulces de frutilla, a partir de la muestra seleccionada y la proporción de dulces de frutilla de la población inicial.
4. Señalar el número de dulces de frutilla en la muestra.

En la Tabla 4.18 se elabora un análisis semiótico de las prácticas descritas y la variedad de objetos matemáticos implicados en la solución de la pregunta 1 del Ítem 9.

Tabla 4.18. Objetos matemáticos ligados a la solución correcta de la pregunta 1 del Ítem 9

Tipos de objetos	Objetos matemáticos
Situaciones problema	Dada la proporción en una población, determinar la proporción en varias muestras
Lenguaje	Palabras del enunciado
Conceptos	Muestra Población Representatividad Variabilidad Proporción
Propiedades	No se solicita propiedades en esta pregunta
Procedimientos	Leer e interpretar el enunciado

Argumentos	Identificar la proporción de dulces de cada tipo Calcular el número de dulces de frutilla de la muestra Elaborar la respuesta a la pregunta planteada No se precisa de argumentos
------------	--

En el caso de la pregunta 2, los estudiantes deberían ser capaces de entregar muestras con una variabilidad normal, una variabilidad excesiva o todas las muestras iguales a la respuesta señalada en la primera pregunta de este ítem.

Para responder a la pregunta 2 de este ítem se deben realizar las siguientes prácticas matemáticas:

1. Leer el enunciado del problema y la pregunta planteada.
2. Identificar la proporción de dulces de frutilla en la muestra seleccionada.
3. Estimar el número probable de dulces de frutilla para cada nueva muestra.

En la Tabla 4.19 se elabora un análisis semiótico de las prácticas descritas y la variedad de objetos matemáticos implicados en la solución de la pregunta 2 del Ítem 9.

Tabla 4.19. Objetos matemáticos ligados a la solución correcta de la pregunta 2 del Ítem 9

Tipos de objetos	Objetos matemáticos
Situaciones problema	Dada la proporción en una población, determinar la proporción en varias muestras
Lenguaje	Palabras del enunciado Tabla que organiza información
Conceptos	Muestra Población Proporción Variabilidad
Propiedades	No se solicita propiedades en esta pregunta
Procedimientos	Leer e interpretar el enunciado Identificar la muestra de peces seleccionados Identificar la proporción de dulces de frutilla Estimar el número probable de dulces de frutilla en cada nueva muestra Elaborar la respuesta a la pregunta planteada
Argumentos	No se precisa de argumentos

Ítem 10. Una industria farmacéutica fabrica un producto llamado “*Elección del género*”, el cual, según la publicidad permite a las parejas incrementar sus posibilidades de tener una niña. Si 100 parejas usaron el medicamento y 52 tuvieron una niña, ¿Piensas que el medicamento es efectivo? ¿Por qué?

Este ítem ha sido adaptado de García-Ríos (2013, p.347), quien utiliza este problema para “explorar el razonamiento empleado por estudiantes de bachillerato para realizar inferencias informales” (García-Ríos, 2013, p.343) en la resolución de un problema de prueba

de hipótesis de proporciones. También pertenece al bloque 7, cuyo objetivo es: dada la proporción en una población, determinar la proporción en varias muestras.

Para responder a este ítem se deben realizar las siguientes prácticas matemáticas:

1. Leer el enunciado del problema y la pregunta planteada.
2. Calcular la proporción de parejas en que fue efectivo el medicamento.
3. Señalar que la muestra no es suficientemente representativa para afirmar que el medicamento fue efectivo.

En la Tabla 4.20 se presenta un análisis semiótico de las prácticas descritas y la variedad de objetos matemáticos implicados en la solución del Ítem 10.

Tabla 4.20. Objetos matemáticos ligados a la solución correcta del Ítem 10

Tipos de objetos	Objetos matemáticos
Situaciones problema	Dada la proporción en una población, determinar la proporción en varias muestras
Lenguaje	Palabras del enunciado
Conceptos	Muestra Población Proporción Variabilidad
Propiedades	No se solicita propiedades en esta pregunta
Procedimientos	Leer e interpretar el enunciado Calcular la proporción de parejas en que fue efectivo el medicamento Elaborar la respuesta a la pregunta planteada
Argumentos	Deductivo

García-Ríos (2013, p.348) propone tres categorías de análisis a las respuestas de este problema: (1) va más allá de los datos; (2) no va más allá de los datos; (3) No hay respuesta. Además, se señala que las respuestas de los estudiantes que *van más allá de los datos* se refieren a la efectividad del producto en la población y no de la muestra en particular; estas respuestas el autor las ha clasificado en cuatro categorías: 1) no funciona; 2) juzgan baja efectividad; 3) asignan porcentaje de efectividad y 4) juzgan que puede ser efectivo.

Por otro lado, las respuestas que *no van más allá de los datos*, se dividieron en dos categorías: 1) cuando concluyen que el producto pudo influir de alguna manera en las personas de la muestra; y 2) cuando concluyen que sí funciona el producto.

Ítem 11. Tenemos 4 libros numerados 1, 2, 3 y 4. Escribe todas las muestras posibles de dos libros *con* y *sin* reemplazamiento.

Este ítem es de elaboración propia, en el que se les solicita a los estudiantes señalar las diferentes muestras con y sin reemplazo que pueden determinar de una población finita de cuatro elementos. En la Tabla 4.21 se pueden ver las respuestas correctas del Ítem 11, con el listado de muestras de cada tipo.

Tabla 4.21. Respuestas correctas del Ítem 11

Muestras con reemplazo	Muestras sin reemplazo
12 - 21 - 31 - 41	12 - 23
13 - 23 - 32 - 42	13 - 24
14 - 24 - 34 - 43	14 - 34

Para responder a este ítem se deben realizar las siguientes prácticas matemáticas:

1. Leer el enunciado del problema y la pregunta planteada.
2. Elaborar el listado de muestras con reemplazo.
3. Elaborar el listado de muestras sin reemplazo.

En la Tabla 4.22 se presenta un análisis semiótico de las prácticas descritas y la variedad de objetos matemáticos implicados en la solución del Ítem 11.

Tabla 4.22. Objetos matemáticos ligados a la solución correcta del Ítem 11

Tipos de objetos	Objetos matemáticos
Situaciones problema	Obtener todas las muestras posibles de una población finita
Lenguaje	Palabras del enunciado
Conceptos	Muestra Población Conceptos combinatorios
Propiedades	Población finita Muestras con o sin reemplazo
Procedimientos	Leer e interpretar el enunciado Elaborar el listado de muestras con o sin reemplazo
Argumentos	No se precisa de argumentos

Síntesis: significado evaluado por el cuestionario

Resumiendo, el análisis a priori del cuestionario, se muestra en la Tablas 4.23 y 4.24 la recopilación de los diferentes objetos matemáticos primarios que entran en juego en las distintas prácticas matemáticas necesarias para resolver cada uno de los ítems. La Tabla 4.23 presenta el resumen de lo desarrollado en este apartado y proporciona la descripción del significado institucional evaluado en el instrumento.

Si se compara el significado institucional pretendido del muestreo en los textos analizados (ver Tablas 3.2; 3.3; 3.4; 3.7, Capítulo 3), con los bloques evaluados en el cuestionario, se puede apreciar que el cuestionario incluye algunas de dichas situaciones problema y los diferentes tipos de lenguaje considerados en la configuración epistémica que define el significado pretendido.

Cabe mencionar, que las situaciones problemas propuestas en el cuestionario son una selección acorde a la dificultad y objetivo de medición de las situaciones problemas encontradas en el análisis curricular (Capítulo 1), las investigaciones previas que respaldan esta investigación (Capítulo 2) y finalmente, algunas de las situaciones planteadas en los libros de texto (Capítulo 3).

En el instrumento, se han propuesto una serie de situaciones problemas relacionadas con los bloques descritos anteriormente (ver apartado 4.3 de este Capítulo y Tabla 4.1) que apuntan a evaluar progresivamente la comprensión del concepto de muestreo. Se señala que algunas de las situaciones propuestas en el análisis curricular, como también, las analizadas en los textos escolares, son un poco más complejas, y algunas apuntan a evaluar conjuntamente el muestreo con otros temas de estadística, por lo que no se han empleado en los bloques establecidos para la confección de este cuestionario.

En la Tabla 4.23 se sintetizan las situaciones problema, los tipos de lenguaje y los conceptos que tienen relación con cada uno de los ítems del cuestionario. Se destaca mayoritariamente el uso del concepto de muestra como el de representatividad.

Tabla 4.23. Síntesis de situaciones problema, lenguajes y conceptos evaluados en el cuestionario

Tipo de objeto	Objetos matemáticos	Ítem											
		1	2	3	4	5	6	7	8	9	10	11	
Situaciones problema	Comprender la definición del concepto de muestra	x	x										
	Detectar sesgo en el método de muestreo			x									
	Sugerir un método de muestreo				x								
	Decidir si una muestra es representativa					x							
	Identificar la muestra y la población						x						
	Dada una proporción en la muestra determinar la proporción en la población							x	x				
	Dada la proporción en una población determinar la proporción en varias muestras									x	x		
	Obtener todas las muestras posibles de una población finita												x
	Lenguajes	Palabras del enunciado	x	x	x	x	x	x	x	x	x	x	x
Tipos de números				x	x	x		x	x	x	x	x	x
Tablas como listado de datos								x					
Tabla que organiza información										x			
Conceptos	Población			x			x	x	x	x			x
	Muestra	x	x	x	x	x	x	x	x	x	x	x	x
	Encuesta		x	x	x								
	Métodos de muestreo			x	x	x				x			
	Porcentaje							x	x	x	x		
	Proporción							x	x	x	x		
	Representatividad de la muestra		x	x	x	x	x	x	x	x	x		
	Sesgo			x									
	Variabilidad			x	x	x				x	x		
	Conceptos combinatorios												x

En la Tabla 4.24 se aprecia que existe una gran cantidad de procedimientos empleados para responder a las diferentes preguntas del cuestionario. También se sintetizan algunos de los argumentos más empleados.

Para el caso de las propiedades, cabe mencionar que en varias preguntas no se precisa de propiedades para poder hallar la solución, es en el Ítem 11 donde se encuentran propiedades como “población finita” y “muestras con o sin reemplazo”.

Tabla 4.24. Síntesis de procedimientos y argumentos evaluados en el cuestionario

Tipo de objeto	Objetos matemáticos	Ítem												
		1	2	3	4	5	6	7	8	9	10	11		
Procedimientos	Leer e interpretar el enunciado	x	x	x	x	x	x	x	x	x	x	x	x	
	Elaborar la respuesta a la pregunta planteada	x	x	x	x	x	x	x	x	x	x	x	x	
	Proponer una muestra representativa de la población		x											
	Identificar el tipo de muestreo			x										
	Identificar el sesgo en el método de muestreo			x										
	Aplicar un tipo de muestreo				x									
	Identificar si la muestra es representativa					x								
	Identificar la población						x							
	Identificar la muestra						x		x	x				
	Identificar la proporción de la muestra							x	x	x	x			
	Calcular la proporción de la muestra								x		x			
	Elaborar el listado de muestras con o sin reemplazo									x		x		
	Argumentos	Uso de ejemplos	x					x						
		Indicar las razones de factibilidad para realizar una encuesta		x										
Justificar la elección del método de muestreo más apropiado			x	x	x									
Justificar por qué es representativa la muestra						x	x					x		
Señalar el procedimiento llevado a cabo								x						
Deductivo												x		

4.6. PRUEBA PILOTO DEL CUESTIONARIO

Una vez realizada la selección y respectiva modificación de los ítems que conforman el cuestionario, de acuerdo al juicio de expertos, se efectuó una prueba piloto del mismo, con la finalidad de corroborar el tiempo requerido para su implementación, así como la legibilidad por parte de estudiantes chilenos que finalizan Educación Secundaria. Esta información será de utilidad, ya que se pueden detectar las posibles dificultades en la comprensión de los enunciados, como también, tener una visión inicial de los posibles errores que se exhiben al responder a cada una de las tareas propuestas en los ítems.

Un análisis general del instrumento como la validación del mismo, se ha presentado en Ruiz-Reyes, Ruz, Contreras y Molina-Portillo (enviado).

4.6.1. Descripción de la muestra de estudiantes y contexto

Para validar el cuestionario, se ha considerado una muestra constituida por 148 estudiantes de cuarto año de secundaria (17-18 años), pertenecientes a tres centros de educación secundaria de la ciudad de Osorno, Chile, quienes cursaban el primer semestre del año 2018, y asisten a diferentes tipos de centros (ver Tabla 4.25): municipal (público), particular subvencionado (concertado), particular pagado (privado).

Tabla 4.25. Dependencia del centro educativo

Tipo Centro	Frecuencia	Porcentaje
Municipal	50	33,8
Particular Subvencionado	50	33,8
Particular Pagado	48	32,4
Total	148	100

Este grupo, está compuesto por 89 (60%) mujeres y 59 (40%) hombres. No se puede afirmar que los estudiantes que forman parte de este estudio hayan recibido la instrucción previa del concepto de muestra y sus propiedades, aunque este contenido esté presente en el currículo de educación primaria, ya que este punto no es analizado en la investigación realizada.

Los datos fueron recogidos por escrito como una actividad de la clase de matemáticas, completándose en un tiempo de 1 hora 30 minutos, con el investigador presente, quien explicaba el objetivo de la evaluación, como también, contestaba las dudas que pudieran surgir sobre la manera de completar el cuestionario.

4.6.2. Resultados del estudio piloto

En este apartado se presenta el análisis de las respuestas que entregan los estudiantes a cada uno de los ítems del cuestionario, además, de las ideas que expresan se mostrará el porcentaje de las respuestas correctas e incorrectas de cada pregunta (Ver Tabla 4.26).

Tabla 4.26. Respuestas correctas e incorrectas por ítem del cuestionario piloto

Ítem	Respuesta Incorrecta Frecuencia (%)	Respuesta Correcta Frecuencia (%)	Ítem	Respuesta Incorrecta Frecuencia (%)	Respuesta Correcta Frecuencia (%)
1	41 (27,7)	107 (72,3)	6.2	99 (66,9)	49 (33,1)
2.1	16 (10,8)	132 (89,2)	6.3	61 (41,2)	87 (58,8)
2.2	41 (27,7)	107 (72,3)	6.4	119 (70,4)	29 (19,6)
2.3	74 (50)	74 (50)	6.1.1	104 (70,3)	44 (29,7)
3.1	51 (34,5)	97 (65,5)	6.2.1	117 (79,1)	31 (20,9)
3.2	89 (60,1)	59(39,9)	6.3.1	96 (64,9)	52 (35,1)
3.3	24 (16,2)	124 (83,8)	6.4.1	127 (85,8)	21 (14,2)
3.4	57 (38,5)	91 (61,5)	7	81 (54,7)	67 (45,3)
3.5	46 (31,1)	102 (68,9)	8.1	72 (48,6)	76 (51,4)
3.6	52 (35,1)	96(64,9)	8.2	90 (60,8)	58 (39,2)
4	98 (66,2)	50 (33,8)	8.3	91 (61,5)	57 (38,5)
5.1	92 (62,2)	56 (37,8)	9.1	54 (36,5)	94 (63,5)
5.2	75 (50,7)	73 (49,3)	9.2	81 (54,7)	67 (45,3)
5.3	78 (52,7)	70 (47,3)	10	120 (81,1)	28 (18,9)
6.1	94 (63,5)	54 (36,5)	11	103 (69,6)	45 (30,4)

Cabe señalar que resultados parciales del análisis de los Ítems 1, 2 y 3 se han presentado en Ruiz-Reyes, Ruz, Molina-Portillo y Contreras (2019) y Ruiz-Reyes, Ruz, Molina-Portillo y Contreras (enviado). A continuación, se muestra una descripción más detallada del análisis y caracterización de las respuestas por ítem.

Ítem 1. ¿Has oído la palabra muestra antes? Explica qué significa para ti.

En la Tabla 4.27, se ve que en el Ítem 1, un porcentaje muy bajo de estudiantes no responde o no ha escuchado anteriormente el concepto de muestra (4.1%), mientras que un alto porcentaje contesta de forma afirmativa (95.9%). Por otro lado, como se aprecia en la Tabla 4.26, el (41) 27.7% de los estudiantes da una respuesta incorrecta, no respondiendo o indicando ejemplos en contextos poco claros, pero un porcentaje bastante alto de ellos, (107) el 72.3%, dan una respuesta correcta, identificando que han escuchado el concepto de muestra en algún contexto (salud, supermercado, tiendas, farmacias, etc.), aunque este no sea necesariamente en la clase de matemática.

Tabla 2.27. Respuestas de los estudiantes al Ítem 1

Respuestas	Frecuencia	Porcentaje
No responde	4	2,7
No. Nunca.	2	1,4
Si. Es una parte de un conjunto más grande	52	35,1
Si. Da ejemplos en diferentes contextos	47	31,8
Si. Da una explicación sin fundamento.	35	23,6
Si. Cuando dan a conocer algo	8	5,4
Total	148	100

Resultados similares se pueden verificar en los trabajos de Watson y Moritz (2000a) y Meletiou-Mavrotheris y Paparistodemou (2015), en este último un 85% de los alumnos han oído la palabra *muestra* en contextos externos a la escuela. Se destaca que, solo cinco estudiantes definen *muestra*, utilizando una definición en un contexto matemático, como *un subconjunto de la población*.

Ítem 2. En una investigación sobre los hábitos de ejercicio de los “estudiantes de secundaria chilenos”, algunos investigadores entrevistaron a una *muestra* de ellos.

1. ¿Qué significa la palabra muestra en esta frase?
2. ¿Por qué piensas que los investigadores seleccionan una *muestra* de estudiantes en vez de preguntar a todos?
3. ¿Estás de acuerdo con que los investigadores seleccionen una muestra de 10 estudiantes? ¿Cuántos elegirías tú?

Ante la primera pregunta del Ítem 2, se puede apreciar que los alumnos tienen alguna idea bastante clara de entender el concepto de *muestra* en un contexto cercano a ellos, ya que solo el 10.8% (16) no responde o responde incorrectamente, mientras que un 89.2% (132) responde correctamente (ver Tabla 4.26).

Se puede mencionar, además, que dentro del 10.8% que responden erróneamente (ver Tabla 4.28), solo dos estudiantes no responden, pero si 14 de ellos responden incorrectamente esta pregunta. Por otro lado, dentro del 89.2% de respuestas correctas, se aprecia que los alumnos mencionan ideas como *una pequeña parte* (56,8%), *un porcentaje* (8,1%), *una cantidad* (23,6%), y solamente un estudiante entrega una definición más precisa, entendiendo la palabra *muestra* como un *subconjunto de la población*. Estos resultados son mejores que los descritos por Meletiou-Mavrotheris y Paparistodemou (2015).

Tabla 4.28. Respuestas de los estudiantes a la pregunta 1, Ítem 2

Respuestas	Frecuencia	Porcentaje
No responde	2	1,3
Incorrecto	14	9,5
Una parte, una pequeña parte	84	56,8
Un porcentaje	12	8,1
Una cantidad	35	23,6
Un subconjunto de la población	1	0,7
Total	148	100

En la pregunta 2, el 72.3% (107) tienen una idea más clara de los factores que son importantes al momento de seleccionar una muestra de una población, por ejemplo, mencionan ideas acerca del excesivo tiempo que puede llevar el estudio, el costo de implementación de la entrevista o sobre el acceso a entrevistar a la población total. Así, un

27.7% (41) no contesta o contesta de manera incorrecta (ver Tabla 4.26).

En la Tabla 4.29 se presenta de forma más detallada como se caracterizan las respuestas de los estudiantes, solo 3 alumnos (2%) no responde a dicha pregunta, 38 de ellos (25.7%) responde incorrectamente. Dentro de los factores que los estudiantes consideran al momento de seleccionar una muestra, mencionan los siguientes: “porque es difícil preguntar a todos” (23%); “porque es costoso” (0,7%); “porque lleva mucho tiempo” (16,9%); “es más rápido o fácil” (14.2%), “porque obtienen los mismos resultados” (17.6%).

Tabla 4.29. Respuestas de los estudiantes a la pregunta 2, Ítem 2

Respuestas	Frecuencia	Porcentaje
No responde	3	2,0
Incorrecto	38	25,7
Porque es difícil preguntar a todos	34	23,0
Porque es costoso	1	0,7
Porque lleva mucho tiempo	25	16,9
Es más rápido o fácil	21	14,2
Porque obtienen los mismos resultados	26	17,6
Total	148	100

Si se observa el análisis de la pregunta 2, se aprecia que los resultados son mejores a los reportados por Meletiou-Mavrotheris y Papparistodemou (2015), donde la mayoría (65%) de los estudiantes dio respuestas que reflejaban problemas prácticos y conocimiento del entorno social en el que se realiza el muestreo: “Sería demasiado largo y tedioso visitar todas las escuelas primarias de Chipre y entrevistar a todos los niños”.

Tabla 4.30. Respuestas de los estudiantes a la pregunta 3a, ítem 2

Respuestas	Frecuencia	Porcentaje
No responde	2	1,4
Incorrecto	4	2,7
No estoy de acuerdo	74	50,0
Si estoy de acuerdo	68	45,9
Total	148	100,0

Considerando las respuestas a la pregunta 3, se ve en la Tabla 4.30 que un 4.1% no responde o da una información incorrecta, un 50% no está de acuerdo con que se elijan solamente 10 estudiantes como muestra, y un 45.9% si está de acuerdo con la afirmación planteada en la cuestión.

Se ve que tanto para respuestas incorrectas, como las correctas obtenemos el 50%. A pesar que en las preguntas anteriores, los alumnos muestran una comprensión del concepto

de muestra y las limitaciones de realizar un estudio preguntándole a todos los individuos de una determinada población, al momento de cuantificar los sujetos de la muestra sus concepciones no son tan acertadas, además, cabría señalar que también estas respuestas pueden ser influenciadas en que se trabaja con una población desconocida, la cual es difícil de cuantificar, lo que induce a un error en la interpretación de la cuestión planteada. Se puede mencionar que los estudiantes son susceptibles a la *insensibilidad al tamaño de la muestra* (Tversky y Kahneman, 1974), es decir, asumen que la muestra seleccionada, independientemente de su tamaño, siempre representa a la población a la que pertenece, lo que induce a graves errores de interpretación en los análisis estadísticos. Además, como indican Kahneman, Slovic y Tversky (1982), este tipo de personas, también cree en la *ley de los pequeños números*, es decir, creen que, al seleccionar muestras no suficientemente grandes, la distribución muestral se distribuye de la misma manera que la distribución de la población, independientemente del tamaño de la muestra.

Tabla 4.31. Respuestas de los estudiantes a la pregunta 3b, Ítem 2

Respuestas	Frecuencia	Porcentaje
No responde. Incorrecto	20	13,5%
10 estudiantes	33	22,3%
Más alumnos	68	45,8%
50%	6	4,1%
Todos	8	5,4%
Depende del total	12	8,1%
Al azar	1	0,7%
Total	148	100%

Si se considera la segunda cuestión planteada en la pregunta 3: *¿Cuántos elegirías tú?*, hay un 22.3% indica que sólo elegiría a 10 estudiantes (ver Tabla 4.31), como en el enunciado de la parte anterior de esta pregunta, el 45. 9% señalan que debería considerarse un número mayor dependiendo ya sea, si se considera el total de alumnos entrevistados, o el total de centros que forman parte del estudio y sólo un 5.4% considera que debería entrevistarse a todos los estudiantes de secundaria chilenos, evidenciando que olvidan las razones entregadas en la pregunta anterior, sobre las factibilidades seleccionar una muestra de la población.

Se señala que un 54.6% de los alumnos presenta de manera más clara el hecho de que debe seleccionarse una muestra mayor a 10 estudiantes, aunque no aluden expresamente al concepto de *representatividad de la muestra*, algunos consideran que debe depender del azar,

como también, del total de alumnos entrevistados.

Ítem 3. Los estudiantes de un colegio realizaron una encuesta para determinar el porcentaje de niños que reciclan en sus casas:

- Andrea preguntó a 80 estudiantes que son miembros del club de medioambiente.
- Elena envió un cuestionario a todos los niños del colegio y tomó los primeros 80 que contestaron.
- Pedro quiso el mismo número de niños y niñas. Así que preguntó a 5 niños y 5 niñas de cada curso para conseguir los 80 estudiantes que constituyen la muestra.
- Luis tenía los nombres de los 800 estudiantes en la escuela; puso cada nombre en un papel en un sombrero y sacó 80.
- Sara no conocía demasiados niños por lo que decidió encuestar a 80 niñas. Pero quería asegurarse de tener niñas de varias edades, por lo que tomó 10 niñas de cada nivel.

a. Completa la última columna donde preguntamos si el método de cada niño es apropiado.

Niño	A quien pregunta	¿Te parece un método apropiado? ¿Por qué?
Andrea	80 estudiantes del club de medioambiente.	
Elena	Envió un cuestionario a todos los niños del colegio y tomó los 80 que contestaron primero.	
Pedro	5 niños y 5 niñas de cada curso	
Luis	Puso el nombre de cada niño de la escuela (800) en un papel en un sombrero y sacó 80	
Sara	10 niñas de cada nivel.	

b. De acuerdo a los datos entregados en el enunciado, ¿Cuál crees que es la mejor manera de escoger a los niños para estimar el porcentaje de ellos que recicla en casa?

En el Ítem 3, se les pregunta a los estudiantes si el método de muestreo descrito es o no apropiado, además, se les pide a los alumnos que den sus respuestas completando la tabla que resume la información entregada en el enunciado.

Tabla 4.32. Respuestas de los estudiantes al Ítem 3, pregunta a

Niño	Respuesta	
	Incorrecta Frecuencia (%)	Correcta Frecuencia (%)
Andrea	51(34.5)	97(65.5)
Elena	89(60.1)	59(39.9)
Pedro	24(16.2)	124(83.8)
Luis	57(38.5)	91(61.5)
Sara	46(31.1)	102(68.9)

En la Tabla 4.32, se presentan los resultados de respuestas correctas e incorrectas por cada uno de los datos expuestos en el enunciado, es decir, si los estudiantes encuentran o no apropiado el método de muestreo señalado por cada niño.

En este ítem, ha sido adaptado del trabajo de Meletiou-Mavrotheris y Papparistodemou (2015) y del de Watson y Kelly (2005), preguntas que han sido adaptadas previamente de la

investigación de Jacobs (1997), y busca conocer la comprensión de los estudiantes sobre el sesgo del muestreo en un contexto cercano a la realidad, como es la toma de encuestas.

Se expone una situación problema empleando encuestas, en un contexto escolar, en el que los métodos de muestreo y los resultados variaron, mientras que el tamaño de la muestra se mantuvo constante. De acuerdo a Jacobs (1997), se incluyen tres métodos de muestreo básicos: método de muestreo aleatorio, restringido y autoseleccionado. El método de *muestreo aleatorio* dio a cada miembro de la población la misma oportunidad de ser seleccionados. Los métodos de *muestreo restringidos* pidieron a grupos particulares de personas que podrían seleccionar una respuesta determinada y, en consecuencia, sesgar los resultados en una dirección particular. Los métodos de *muestreo autoseleccionados* permiten que los participantes de la encuesta seleccionados fueran problemáticos porque no había suficiente evidencia para evaluar si la muestra era representativa.

En esta pregunta se puede apreciar que los estudiantes consideran más apropiado el método de muestreo estratificado, como en el caso de Pedro (elige 5 niños y 5 niñas de cada curso), teniendo el 83.8% de respuestas correctas. Por otro lado, contestan incorrectamente cuando se les presenta un método de muestreo autoseleccionado, como es el caso de Elena, teniendo un 60.1% de respuestas incorrectas.

En cuanto a señalar cuál de los métodos propuestos, es el más adecuado para escoger una muestra de estudiantes que recicla, un 64.9% de los estudiantes contestan correctamente, señalando el método de Pedro o el método de Luis como el más apropiado, es decir, prefieren el muestreo aleatorio simple o el estratificado y un 35.1% responde incorrectamente, mencionando algún método sesgado (Andrea, Elena o Sara) o sugiriendo otro método como entrevistar al 50% de los estudiantes o al total, es decir, a toda la población.

Un análisis más completo y detallado de esta pregunta se incluye en Ruiz y Contreras (en prensa) en donde se propone un método de análisis empleando una adaptación de las categorías propuestas por Watson y Kelly (2005) además de las respuestas dadas por los estudiantes, y se describen las respuestas entregadas por los alumnos a cada uno de los métodos de selección de muestras.

Ítem 4. Supongamos que quieres averiguar el porcentaje de niños que viene a tu escuela en diferentes medios de transporte: caminando, auto, micro, bicicleta, u otro. ¿Cómo tomarías una muestra de 50 niños de modo que los resultados sean representativos de toda la escuela?

Cuando se les solicita a los estudiantes, en el Ítem 4, sugerir un método de muestreo para seleccionar una muestra en un contexto cercano a su realidad, se observa que el 54.7% de ellos, menciona un método de muestreo, pero no justifica su elección, el 2.7% de los estudiantes menciona obtener un método sesgado, es decir, no consideran a todos los sujetos de la población, sino a un grupo con alguna característica en particular, finalmente, solo el 33, 8% menciona un método de muestreo apropiado (ver Tabla 4.33).

En síntesis, en el Ítem 4, se tiene un alto porcentaje de respuestas incorrectas 66.2%, destacando que solo un 33.8% contesta correctamente (ver Tabla 4.26). En esta pregunta, se supone que la mayor dificultad surge por no señalar el número total de estudiantes de la escuela (población), lo que les dificulta comprender la situación planteada en el enunciado. Cabe señalar, que un grupo de estudiantes plantearon dicha cuestión al investigador que aplicó el instrumento.

Tabla 4.33. Respuestas de los estudiantes al Ítem 4

Respuestas	Frecuencia	Porcentaje
No contesta	13	8,8
Menciona un método de muestreo, pero no justifica	81	54,7
Menciona obtener una muestra de un grupo sesgado (alguna característica en particular)	4	2,7
Usan algún método de muestreo y justifica	50	33,8
Total	148	100

Ítem 5. Analiza si las siguientes muestras son representativas de la población de estudio y completa la tabla con tus respuestas.

Muestra	¿Es representativa?	Explica por qué
1. Se seleccionan al azar 100 estudiantes, de un colegio de 300 alumnos, para calcular el promedio de notas de la población.		
2. Se les pregunta sobre la calidad del servicio del metro a las primeras 100 personas que toman el metro en la mañana.		
3. Se escogen los niños más altos y los más bajos de un curso para determinar el promedio de la estatura de los alumnos.		

El Ítem 5, ha sido adaptado de una actividad de un texto escolar de primero medio (Del Valle, Muñoz y Santís, 2013, p. 247), en la cual los estudiantes deben indicar si es o no representativa la muestra dada, además de justificar su respuesta explicando el porqué.

Como se aprecia en la Tabla 4.34, en la primera muestra, sólo el 37.8% de los alumnos contesta correctamente, indicando que es representativa la muestra señalada, en una situación que involucra un muestreo aleatorio simple; en la segunda, sólo el 49.3% contesta que no es representativa, debido a que es un método de muestreo restringido, lo que implica que se eligen sujetos de manera sesgada, y que no representan a la población total, finalmente, en la tercera muestra, solo el 47.3% contesta que no es representativa, debido a que se emplea un método de muestreo autoseleccionado, mostrando que la selección de la muestra da problemas al intentar ver si es o no representativa. Se aprecia que los estudiantes reconocen mayoritariamente las muestras no representativas, pero en el primer caso, no identifican que es un muestreo aleatorio simple, llevándolos a pensar que la manera de seleccionar la muestra no es representativa de la población indicada en el enunciado (62,2%).

Tabla 4.34. Respuestas de los estudiantes al Ítem 5

Muestra	Respuesta Correcta		Respuesta Incorrecta	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
1.	56	37.8	92	62.2
2.	73	49.3	75	50.7
3.	70	47.3	78	52.7

Ítem 6. Determina en cada caso la población y una posible muestra de ella. Completa la tabla con tus respuestas.

Datos	Población	Muestra
1. Una fábrica de yogur quiere investigar sobre la calidad de sus productos.		
2. Diego necesita saber el precio de un kilogramo de carne, para una comida familiar.		
3. Ximena estudia respecto del tamaño de las hormigas que habitan en un insectario.		
4. Daniel desea saber si una ciudad cuenta con suficiente lluvia, para realizar una plantación.		

En el Ítem 6, adaptado de una actividad del texto escolar de segundo medio (Muñoz, Rupin y Jiménez, 2013, p. 275), donde deben indicar en cada caso, la población y una posible muestra de ésta.

Tabla 4.35. Respuestas de los estudiantes al Ítem 6

	Población Respuesta		Muestra Respuesta	
	Incorrecto Frecuencia (%)	Correcto Frecuencia (%)	Incorrecto Frecuencia (%)	Correcto Frecuencia (%)
1.	94(63.5)	54(36.5)	104(70.3)	44(29.7)
2.	99(66.9)	49(33.1)	117(79.1)	31(20.9)
3.	61(41.2)	87(58.8)	96(64.8)	52(35.1)
4.	119(80.4)	29(19.6)	127(85.8)	21(14.2)

La mayoría de los estudiantes contesta incorrectamente (ver Tabla 4.35), tanto para identificar la población, como la muestra, evidenciando que tienen algunos problemas al reconocer estos conceptos en contextos más alejados de la realidad escolar, además, de no identificar la diferencia entre la relación parte-todo de los contextos mencionados en el enunciado. Surge la inquietud también de analizar la complejidad en la redacción de la pregunta, debido a que este ítem se adaptó de una actividad propuesta en un libro de texto.

Ítem 7. Lee el siguiente enunciado y responde la pregunta planteada.

Se tomaron al azar cuatro muestras de estudiantes de un colegio. A continuación, se presentan los resultados.

Alumnos	35	29	24	32
Alumnas	15	21	26	18

¿Podrías estimar el porcentaje de alumnos en el colegio? Explica tu procedimiento.

En el Ítem 7, se pide estimar el porcentaje total de alumnos, dando cuatro muestras extraídas de una población desconocida, pregunta tomada y adaptada del texto escolar de séptimo básico (Merino, Muñoz, Pérez, y Rupin, 2015, p. 289).

Tabla 4.36. Respuestas de los estudiantes al Ítem 7

Respuesta	Frecuencia	Porcentaje
No contesta	68	45.9
Calcula el porcentaje de alumnos y el promedio	25	16.9
Calcula el porcentaje de hombres sin el promedio	13	8.8
Calcula el porcentaje con los totales	42	28.4
Total	148	100

En este caso, a pesar de tener un alto porcentaje de alumnos que no contesta (45.9%), el 45.3% (67) de los estudiantes, contesta correctamente (ver Tabla 4.26). Esta actividad se ha adaptado de una más extensa, presente en el texto escolar. Según lo señalado en el texto escolar de séptimo básico (Merino, Muñoz, Pérez, y Rupin, 2015, p. 289), se pretende que los estudiantes asocien la siguiente idea: *al repetir un experimento muchas veces, los porcentajes se estabilizan, es decir se observa que tienden a acercarse a un valor determinado*. Es decir, los estudiantes, tienen problemas al estimar el porcentaje de hombres,

cuando la población es desconocida, a pesar de tener información de cómo se comportan algunas de las muestras de dicha población, por lo que no emplean el concepto de *representatividad* como *la capacidad de reproducir a pequeña escala las características de la población*.

Ítem 8. En un estanque hay peces, pero su dueño no sabe cuántos. Toma 200 peces y les pone una marca; los devuelve al estanque, donde se mezclan con el resto. Al día siguiente, el dueño toma 250 peces del estanque, y encuentra que 25 de ellos están marcados y el resto no.

1. ¿Qué proporción aproximada de peces del estanque están marcados?
2. ¿Cuál es el número aproximado de peces en el estanque?
3. Si el dueño saca ahora 100 peces más del estanque, ¿cuántos aproximadamente estarán marcados?

El Ítem 8, adaptado de Gómez (2014, p.202), el cual ha sido adaptado previamente de un experimento de enseñanza de Fishbein y Gazit (1984) con niños de 10 a 13 años. Se observa en la Tabla 4.37 que el 51.4% de los estudiantes contesta correctamente la primera cuestión, identificando la proporción de peces marcados en el estanque. En la segunda pregunta, se les pide estimar la cantidad total de peces y sólo un 39.2% contesta correctamente esta pregunta, mostrando un porcentaje mejor a lo reportado por Gómez (2014). En la tercera cuestión, se les da una muestra, y se les pide calcular el porcentaje de peces marcados, sólo un 38.5% calcula correctamente el porcentaje solicitado.

Tabla 4.37. Respuestas de los estudiantes al Ítem 8

Pregunta	Respuestas		
	No contesta Frecuencia (%)	Incorrecto Frecuencia (%)	Correcto Frecuencia (%)
1	23(15.5)	49(33.1)	76(51.4)
2	19(12.8)	71(48.0)	58(39.2)
3	22(14.9)	69(46.6)	57(38.5)

Ítem 9. Piensa que tienes un recipiente con 100 dulces; 60 son de frutilla, y 40 de limón. Los dulces están todos mezclados en el recipiente. Sacas un puñado de 10 dulces y cuentas el número de los dulces de frutilla.

1. ¿Cuántos serán de frutilla?
2. Supongamos que seis de tus compañeros repiten experimento, cada uno de ellos saca 10 dulces. (Después de cada extracción, los dulces se vuelven a colocar y se mezcla nuevamente).

Escribe en cada línea un número probable de dulces de frutilla que obtiene tu compañero.

Niño	Número de dulces de frutilla	Niño	Número de dulces de frutilla
Ana		Pedro	
Camila		María	
José		Juan	

En el Ítem 9, adaptado de Shaughnessy, Ciancetta y Canada (2004), se observa en la Tabla 4.38 que el 63.5% de los alumnos identifican correctamente la cantidad solicitada de

caramelos de frutilla (fresa) en la pregunta 1. Es decir, identifican correctamente la proporción de la muestra.

Tabla 4.38. Respuestas de los estudiantes al Ítem 9, pregunta 1

Respuesta	Frecuencia	Porcentaje
Incorrecta	54	36.5
Correcta	94	63.5
Total	148	100)

En la Tabla 4.39 se exponen las respuestas entregadas por los estudiantes al proponer un listado de muestras para seis extracciones distintas. Las respuestas de “todo igual a 6” sugieren que los estudiantes determinan su respuesta empleando el sesgo de equiprobabilidad, posiblemente debido a la instrucción de probabilidad que han recibido. En las respuestas de “variabilidad normal”, los estudiantes indican una muestra con los valores en un rango, por ejemplo, de 4-8, en el caso de la “variabilidad excesiva” sus muestras consideran valores más extremos en los rangos 1-3 o 9-10 o superior.

Tabla 4.39. Respuestas de los estudiantes al Ítem 9, pregunta 2

Niño	No contesta		Todo igual a 6		Variabilidad Normal		Variabilidad Excesiva	
	Frec	%	Frec	%	Frec	%	Frec	%
Ana	38	25,7	51	34,5	28	18,9	31	20,9
Camila	39	26,4	40	27,0	39	26,4	30	20,3
José	41	27,7	35	23,6	45	30,4	27	18,2
Pedro	39	26,4	40	27,0	40	27,0	29	19,6
María	39	26,4	43	29,1	25	16,9	41	27,7
Juan	41	27,7	35	23,6	37	25,0	35	23,6

Estos resultados son un poco mejor a los reportados por Shaughnessy, Ciancetta, y Canada (2004), ya que se puede apreciar en las respuestas del pilotaje que los estudiantes usan un rango apropiado de valores para identificar la variabilidad normal de la muestra.

Ítem 10. Una industria farmacéutica fabrica un producto llamado “Elección del género”, el cual, según la publicidad permite a las parejas incrementar sus posibilidades de tener una niña. Si 100 parejas usaron el medicamento y 52 tuvieron una niña, ¿Piensas que el medicamento es efectivo? ¿Por qué?

La pregunta del Ítem 10, adaptada de García-Ríos (2013, p.347), adaptado previamente de Triola (2004), se aprecia que en un análisis general a la respuesta un 55.4% de los estudiantes la contesta, indicando que el método es efectivo.

Tabla 4.40. Análisis respuestas. Ítem 10a

Respuesta	Frecuencia	Porcentaje
No contesta	26	17.6
No	40	27
Si	82	55.4
Total	148	100

Si bien dan razones con argumentos (ver Tabla 4.41), es posible observar que de acuerdo a las categorías propuestas por García-Ríos (2013), en este caso, hay algunas que no se han identificado en las respuestas que entregan los estudiantes que fueron parte del piloto, además, no realizan cálculos estadísticos más profundos como la distribución que modela la situación, debido a que es un tema que deben ver durante el transcurso del año escolar, pero que no necesariamente habían cursado al momento de la aplicación del instrumento.

Tabla 4.41. Respuestas de los estudiantes al Ítem 10, pregunta b

Respuestas	Clasificación de respuestas	Frec.	%
Va más allá de los datos	No funciona	23	15.5
	Juzgan baja efectividad	9	6.1
	Asignan porcentaje de efectividad	78	52.7
No va más allá de los datos	Juzgan que puede ser efectivo		
	Pudo influir en las personas		
No hay respuesta	Funcionó	12	8.1
	No contesta	26	17.6

Cabe mencionar que, en el análisis factorial realizado que se describe en el apartado 4.7 de este Capítulo, se deduce que este ítem podría eliminarse del estudio debido a que no muestra relación con ningún otro ítem del cuestionario.

Ítem 11. Tenemos 4 libros numerados 1, 2, 3 y 4. Escribe todas las muestras posibles de dos libros *con* y *sin* reemplazamiento.

En el Ítem 11, sólo el 30.4% contesta correctamente esta pregunta anotando ambas muestras pedidas y un 33.8% la contesta parcialmente correcta, escribiendo el listado completo de elementos de solo una de las muestras solicitadas, ya sea la con reemplazo (28.4%) o la sin reemplazo (5.4%).

Como esta pregunta involucra elementos de combinatoria y permutaciones, se observa que algunos estudiantes contestan incorrectamente porque confunden la fórmula que deben emplear para realizar los cálculos, y no identifican que dicha fórmula sólo le da la cantidad de elementos que debe contener cada muestra.

Tabla 4.42. Respuestas de los estudiantes al Ítem 11

Respuesta	Frecuencia	(%)
No contesta	43	29.1
Anota listado con reemplazo	42	28.4
Anota listado con reemplazo incompleto	7	4.7
Anota listado sin reemplazo	8	5.4
Anota listado sin reemplazo incompleto	3	2
Anota ambos listados	45	30.4
Total	148	100

4.7. APROXIMACIÓN AL ESTUDIO DE VALIDEZ DE CONTENIDO

En este apartado se presentan algunos análisis estadísticos como el índice de dificultad, el índice de discriminación, la fiabilidad mediante el alfa de Cronbach y la validez de nuestro instrumento.

Características Psicométricas del cuestionario

Siguiendo la metodología descrita en Contreras (2011), para comprobar las características psicométricas del cuestionario ECM se estudió la fiabilidad de consistencia interna del test, mediante el cálculo del coeficiente *Alfa de Cronbach*, modelo de consistencia interna basado en la correlación entre elementos promedio que permite determinar el grado en que los elementos del cuestionario se relacionan entre sí, obteniendo un índice global de la replicabilidad o de la consistencia interna de la escala en su conjunto. Para este cuestionario se obtuvo un valor de 0.748, considerado como suficiente para garantizar la fiabilidad de la escala.

En segundo lugar, se ha realizado un análisis factorial del conjunto de respuestas a todos los ítems. Según Muñiz (1994) el análisis factorial permite comprobar las hipótesis teóricas sobre la estructura del constructo. Según Martínez-Arias (1995) la estructura factorial de las respuestas a un cuestionario permite, además, comprobar la dimensionalidad latente e interpretar los factores en términos de las dimensiones supuestas en el mismo. La extracción de factores se llevó a cabo mediante el método de componentes principales; con objeto de obtener factores estadísticamente independientes y de máxima variabilidad al tiempo que no se deforma la estructura de los datos (Cuadras, 1981). Este método parte de una estimación inicial más alta de las comunalidades (Martínez-Arias, 1995). Como método de rotación se usó la rotación Varimax (ver cálculos en la Tabla 4.42), método ortogonal, que

conserva las comunalidades y la suma de porcentajes de varianza explicados por los factores (Afifi y Clark, 1990). Está orientado a maximizar la varianza de los factores. Los ítems se puntuaron como 1=correcto, 0= incorrecto.

Posteriormente, como indica Díaz (2007), se contrastó la hipótesis de que los elementos de la matriz de correlaciones fuera de la diagonal principal son diferentes de cero, mediante la prueba de esfericidad de Barlett. Se obtuvo un valor altamente significativo (P-valor=0, Chi-C=11998,2), lo que indica la existencia de correlaciones suficientemente altas entre las variables para poder realizar el análisis. La prueba de Kaiser-Meyer-Olkin dio un valor 0,61, superior al recomendado para poder llevar a cabo un análisis factorial con garantías. Todo ello garantizó la aplicabilidad del método.

El análisis factorial inicial (Tabla 4.43) produce 5 factores con autovalor mayor que 1, que explicaron el 40,9% de la varianza total. El primer factor explica un 15,01% de la varianza, mientras que los siguientes entre el 5,091% y el 8,078% cada uno. La importancia relativa del primer factor proporciona una evidencia de validez de constructo, puesto que el porcentaje es claramente superior al explicado por los restantes (Morales, 1988).

El primer factor hace referencia a todas las cuestiones del bloque 5 (ítem 6) referentes a la identificación de la muestra y la población de una serie de ejemplos. El segundo factor, que explica el 8,078% de la varianza, está relacionado con las tres cuestiones del ítem 8 del bloque 6 y los ítems relacionados con la cuestión cuarta del bloque 5, en las que se identifican la proporción de la población en función de la muestra y por tanto es necesario identificar estos elementos. El factor 3, que explica el 7,33%, relaciona los bloques 2, 3 y 4, con algunos ítems del bloque 5. En este caso se relacionan la detección del sesgo en el método de muestreo, la sugerencia del método de muestreo y la representatividad de la muestra.

El factor 4 que explica el 5,39%, relaciona los bloques 5, 6, 7 y 8. En este caso se relacionan la identificación de la muestra y la población, con la identificación de la proporción de la población en función de la muestra, con la identificación de que si dada la proporción en una población se puede o no determinar la proporción en varias muestras y la obtención de todas las muestras posibles de una población finita. Por último, el quinto factor, que explica el 5,091% de la varianza, relaciona los bloques 1, 2 y 3, que relacionan la comprensión de la definición del concepto de muestra, la detección de sesgos en el método

de muestreo y la sugerencia de métodos de muestreo.

Tabla 4.43. Matriz de componentes rotado

	1	2	3	4	5
Pregunta_1					0,235
Pregunta_2a			0,358		
Pregunta_2b					0,337
Pregunta_2c			0,580		
Pregunta_3_Método_Andrea			0,443		0,400
Pregunta_3_Método_Elena					-0,524
Pregunta_3_Método_Pedro			0,625		
Pregunta_3_Método_Luis					0,375
Pregunta_3_Método_Sara			0,341		
Pregunta_3b			0,450		
Pregunta_4			0,468		-0,441
Pregunta_5_1					0,516
Pregunta_5_2			0,511		
Pregunta_5_3					0,311
Pregunta_6_P_1	0,737				
Pregunta_6_P_2	0,671				0,314
Pregunta_6_P_3	0,497		0,306		
Pregunta_6_P_4	0,539	0,321		-0,445	
Pregunta_6_M_1	0,685			0,387	
Pregunta_6_M_2	0,646				
Pregunta_6_M_3	0,531		0,364		
Pregunta_6_M_4	0,618	0,305		-0,360	
Pregunta_7				0,477	
Pregunta_8_1		0,869			
Pregunta_8_2		0,868			
Pregunta_8_3		0,795			
Pregunta_9_1		0,488			
Pregunta_9_2				0,562	
Pregunta_10					
Pregunta_11				0,492	

En tercer lugar, se ha medido el índice de dificultad y discriminación (Morales, 1988) entre los grupos de alto (percentil 75) y bajo rendimiento (percentil 25) de la muestra. El índice de dificultad, que se describe como la proporción de aciertos (porcentaje si se

multiplica por 100) del 50% del total (los dos 25% con puntuaciones totales extremas), se representa por:

$$\text{Índice de dificultad} = \frac{R_{P_2} + R_{P_7}}{NP_2 + NP_7}$$

Donde RCP indica el número de respuestas correctas en el ítem y NP el tamaño del conjunto, del percentil correspondiente. Por otro lado, se define el índice de discriminación como la diferencia entre la proporción de aciertos en el grupo superior (RCP75/NP75) menos la proporción de aciertos en el grupo inferior (RCP25/NP25). Expresa hasta qué punto el ítem discrimina, establece diferencias, entre el grupo superior e inferior. A mayor diferencia en número de respuestas correctas entre los grupos superior e inferior, el ítem es más discriminante.

$$\text{Índice de discriminación} = \frac{R_{P_7}}{NP_7} - \frac{R_{P_2}}{NP_2}$$

Los resultados, que se muestran en la Tabla 4.44, indican un índice de dificultad alto, por tanto, la respuesta resulta fácil para los estudiantes, en los bloques 1 y 2. Por el contrario, los bloques 3, 5 y 8 son los que presentan un índice de dificultad más bajo, y por tanto más difíciles de realizar para los estudiantes.

Tabla 4.44. Índices de dificultad y discriminación por rendimiento para P₂₅ - P₇₅

Bloque	I. dificultad	I. discriminación
Bloque 1	0,69	0,31
Bloque 2	0,62	0,31
Bloque 3	0,30	0,39
Bloque 4	0,42	0,19
Bloque 5	0,31	0,46
Bloque 6	0,42	0,62
Bloque 7	0,46	0,29
Bloque 8	0,30	0,48

En el caso del índice de discriminación, los valores más bajos se presentan en el bloque 4, lo que indica una baja discriminación (responden igualmente los estudiantes de alto y bajo rendimiento). Por el contrario, en el bloque 6, es donde más diferencia se encuentran entre las respuestas correctas de los grupos de alto y bajo rendimiento, junto al bloque 5 y 8. Siendo las diferencias menos pronunciadas en el resto de bloques.

4.8. CONCLUSIONES SOBRE EL CUESTIONARIO

Se ha construido un cuestionario que se ha aplicado a una muestra de 148 estudiantes chilenos de cuarto año de secundaria (17-18 años) para su validación, que proporciona información sobre su comprensión acerca del muestreo, como también sobre las principales dificultades relacionadas al entendimiento del concepto, resultados que nos permitirán acercarnos a las nociones intuitivas que poseen estos estudiantes.

De acuerdo a la información proporcionada por el pilotaje, se puede afirmar que el cuestionario cumple con las expectativas previstas, por tanto, se considera correcto no modificar el cuestionario para la versión definitiva.

A modo de conclusión, se puede mencionar que los estudiantes distinguen el concepto de muestra en contextos cercanos a sus experiencias (médico, científico, tiendas, etc.), siendo muy pocos los que dan una respuesta más elaborada y más relacionada a un contexto de las matemáticas escolares, es decir, no usan el concepto de aleatoriedad, y solo cinco estudiantes definen *muestra* como un *subconjunto de la población*, resultados similares se pueden verificar en los trabajos de Watson y Moritz (2000) y Meletiou-Mavrotheris y Paparistodemou (2015), en este último un 85% de los alumnos han oído la palabra *muestra* en contextos externos a la escuela.

Otro aspecto a considerar, es que los estudiantes son susceptibles a la *insensibilidad al tamaño de la muestra* (Tversky y Kahneman, 1974), es decir, asumen que la muestra seleccionada, independientemente de su tamaño, siempre representa a la población a la que pertenece, lo que induce a graves errores de interpretación en los análisis estadísticos. Además, como indican Kahneman, Slovic y Tversky (1982), este tipo de personas, también cree en la *ley de los pequeños números*, es decir, creen que, al seleccionar muestras no suficientemente grandes, la distribución muestral se distribuye de la misma manera que la distribución de la población, independientemente del tamaño de la muestra.

Por otro lado, al pedirles que identifiquen la muestra de un estudio en un contexto cotidiano o cercano a su realidad, se pueden apreciar algunas incongruencias en sus respuestas. Por ejemplo, son capaces de identificar las limitantes que pueden llevar a un investigador a seleccionar una muestra y no considerar a todos los sujetos de una determinada población, pero al momento de sugerir una muestra o preguntarles cuál es el tamaño de

muestra que ellos elegirían, un 50% de los estudiantes de nuestro estudio piloto, responde incorrectamente.

En el ítem 3, adaptado de Jacobs (1997), se expone una situación problema empleando encuestas, en un contexto escolar, en el que los métodos de muestreo y los resultados variaron, mientras que el tamaño de la muestra se mantuvo constante. Se incluyen tres métodos de muestreo básicos: método de muestreo aleatorio, restringido y autoseleccionado. El método de muestreo aleatorio dio a cada miembro de la población la misma oportunidad de ser seleccionados. Los métodos de muestreo restringidos pidieron a grupos particulares de personas que podrían seleccionar una respuesta determinada y, en consecuencia, sesgar los resultados en una dirección particular. Los métodos de muestreo autoseleccionados permiten que los participantes de la encuesta seleccionados fueran problemáticos porque no había suficiente evidencia para evaluar si la muestra era representativa. En este sentido, cuando los estudiantes se enfrentan a diferentes métodos de muestreo empleando encuestas, les cuesta identificar las ventajas del muestreo aleatorio, además, son incapaces de detectar los sesgos en los muestreos autoseleccionados.

Al intentar sugerir un método de muestreo, solo el 54.7%, da una respuesta, pero no justifica su elección del método. Cabe señalar que los estudiantes no entregan mayoritariamente un método de muestreo debido a que, en el enunciado del problema, no se menciona la cantidad total de estudiantes (población desconocida).

Al decidir si una muestra es representativa, los estudiantes distinguen mayoritariamente los casos en que la muestra dada no es representativa, pero en el primer ejemplo, no distinguen que se trata de un muestreo aleatorio simple, llevando al error de no creer que dicha muestra es representativa del conjunto total.

Cuando se les solicita identificar la población y la muestra, los estudiantes contestan erróneamente, se cree que esto se debe a que los contextos ejemplificados en esta actividad de un texto escolar, son desconocidos por los estudiantes, y, además, impiden identificar fácilmente la relación parte-todo, en los elementos que conforman la población o la muestra de cada ejemplo.

Otra problemática en los estudiantes aparece al interpretar la representatividad de las muestras dadas de una población, cuando deben calcular el porcentaje total de alumnos,

dando cuatro muestras extraídas de una población desconocida, un alto número de estudiantes no contesta esta pregunta 45.9% y el 45.3%, contesta correctamente (Ítem 7). En cuanto al Ítem 8, en la primera pregunta la mayoría de los estudiantes (51.4%) puede calcular la proporción de la población de acuerdo a los datos entregados, en las otras dos cuestiones se puede apreciar, que un porcentaje cercano al 40% responde correctamente.

Otro aspecto que se puede apreciar es que los estudiantes logran estimar la cantidad solicitada conociendo cómo se comporta la proporción de la población (Ítem 9), obteniendo un 63.5% de respuestas correctas. En cuanto al Ítem 10, se observa que los estudiantes entregan una respuesta con argumentos, no realizando ningún cálculo estadístico, principalmente, debido a que este tema que trata la distribución muestral puede no haber sido tratado aún en su formación, pregunta que en el análisis factorial realizado se sugiere eliminar por no encontrar una relación con los otros ítems del cuestionario.

Finalmente, cuando se les solicita obtener todas las muestras posibles de una población finita, sólo el 30.4% de los estudiantes responde correctamente, ya que otro 33.8% anota solamente uno de los listados solicitados.

Se considera que, con los resultados obtenidos de las características psicométricas del instrumento analizado, sería de gran interés realizar su aplicación en una muestra más grande.

De igual manera, sería interesante realizar una comparación de las respuestas de estos estudiantes con la de cursos inferiores, para observar cómo se presentan las intuiciones de alumnos más pequeños, pues al analizar las respuestas de estudiantes que finalizan su escolaridad obligatoria, es difícil verificar si sus concepciones están o no contaminadas con la instrucción previa.

CAPÍTULO 5:

RESULTADOS DEL ESTUDIO DE EVALUACIÓN DE LA COMPRENSIÓN DEL MUESTREO

- | |
|---|
| <ul style="list-style-type: none">5.1. Introducción5.2. Objetivos5.3. Descripción de la muestra y contexto5.4. Metodología del análisis de los datos5.5. Resultados por ítem5.6. Diferencias entre grupos5.7. Conflictos semióticos detectados en el estudio de evaluación5.8. Conclusiones sobre los resultados del estudio de evaluación |
|---|

5.1. INTRODUCCIÓN

En este Capítulo se describen los principales resultados obtenidos del estudio de la comprensión del muestreo por estudiantes de secundaria, aplicado a una muestra de 1241 alumnos chilenos de octavo año de primaria (13 - 14 años, finalizando la educación primaria), segundo medio (15 - 16 años, a mitad de educación secundaria) y cuarto medio (17 - 18 años, finalizando la educación secundaria), por medio de la utilización de un instrumento validado (Cuestionario) descrito previamente en el Capítulo 4.

El análisis de las respuestas de los estudiantes permitirá una aproximación al significado personal que atribuyen al concepto de muestra y métodos de muestreo, a través de la caracterización de sus conocimientos y los conflictos semióticos potenciales. De igual manera, se podrá constatar en qué medida los objetivos propuestos en el análisis curricular y de los libros de texto son alcanzados por los alumnos que componen la muestra de este estudio.

Se presenta a continuación la formulación de objetivos e hipótesis considerados en este estudio de evaluación. Luego, se describirá la caracterización de los sujetos de la muestra y el contexto, además se especificará la metodología empleada para el análisis de los datos para posteriormente, dar paso a una descripción detallada del análisis de las respuestas por ítem, como también de la comparación de dichas respuestas por grupo y la caracterización de los principales conflictos semióticos detectados. Finalmente, se presentarán las conclusiones generales de este capítulo de acuerdo a los objetivos planteados en cada bloque de preguntas (descritos en la sección 4.3 y resumidos en la Tabla 4.1 del Capítulo 4).

5.2. OBJETIVOS

El objetivo que se detalla en este capítulo corresponde al Objetivo 3 que se ha planteado en el Capítulo 1 de este estudio, es decir:

Aplicar el cuestionario construido para evaluar la comprensión que alcanzan los estudiantes chilenos al comienzo y final de la educación secundaria sobre el muestreo y describir las principales dificultades observadas.

El interés de esta investigación radica en que, al realizar el estudio de evaluación, este permitirá determinar si los elementos del significado del muestreo planteados en las directrices curriculares y libros de texto son alcanzados por los escolares chilenos. Para ello, una vez validado el cuestionario y comprobado los índices de fiabilidad (ver Capítulo 4), se han propuesto los siguientes objetivos específicos:

O3.1. Elegir una muestra de alumnos apropiadamente grande y representativa para que los resultados de la investigación sean fiables y generalizables. Para lograr este objetivo se seleccionó una muestra bastante amplia, la que se describe en el apartado subsiguiente y está formada por 1241 estudiantes chilenos pertenecientes a distintos centros educativos.

O3.2. Describir los resultados obtenidos en cada una de las preguntas del cuestionario. En este capítulo se puntualiza este objetivo, para ello, se presenta el análisis cuantitativo y cualitativo de las respuestas proporcionadas por los alumnos, igualmente, se muestran algunos ejemplos de respuestas para caracterizar su tipología y además se proporciona una interpretación de dichos resultados.

O3.3. Examinar si el conocimiento alcanzado por los alumnos se mantiene al cursar la enseñanza secundaria. Para alcanzar este objetivo, en los análisis se realiza la comparación de los resultados en los tres grupos participantes, como también algunas pruebas estadísticas: ÁNOVA de un factor y la prueba post hoc para determinar si existen o no diferencias entre los grupos de participantes.

El alcance de los objetivos y de la comprobación de las hipótesis iniciales de este estudio se realizará extensamente en el Capítulo 6 de esta memoria.

5.3. DESCRIPCIÓN DE LA MUESTRA Y CONTEXTO

En este estudio, la muestra está constituida por 1241 estudiantes chilenos (ver Tabla 5.1), que se organizan en tres grupos etarios:

Grupo 1: alumnos de octavo año de primaria (13-14 años) que finalizan la educación primaria;

Grupo 2: alumnos de segundo año medio (15-16 años) a la mitad de educación secundaria;

Grupo 3: alumnos de cuarto año medio (17-18 años) que finalizan educación secundaria.

Estos grupos son elegidos para poder analizar si hay algún cambio significativo considerando entre el principio y el fin de la secundaria, es decir, si se observan mejoras a medida que aumenta la edad de los alumnos que participan en el estudio.

Tabla 5.1. Caracterización de la muestra de estudiantes participantes

Centros	Cursos			Total
	8° Básico	2° Medio	4° Medio	
Centro 1	0	212	187	399
Centro 2	79	78	66	223
Centro 3	23	20	26	69
Centro 4	25	16	22	63
Centro 5	77	128	94	299
Centro 6	70	75	43	188
Total	274 (22.1%)	529 (42.6%)	438 (35.3%)	1241

La muestra final de participantes está constituida por 274 (22.1%) estudiantes de octavo año de primaria, 529 (42.6%) estudiantes de segundo año medio y 438 (35.3%) estudiantes

de cuarto año medio, pertenecientes a seis centros educativos de educación secundaria de la ciudad de Osorno, Chile, quienes cursaban el primer semestre del año 2018 (durante los meses de abril y mayo). Este grupo, además, está compuesto por 690 (55.6%) mujeres y 551 (44.4%) hombres (ver Tabla 5.2).

Tabla 5.2. Distribución por sexo de la muestra de estudiantes

Curso Centro	8° Básico		2° Medio		4° Medio		Total
	Mujer	Hombre	Mujer	Hombre	Mujer	Hombre	
Centro 1	0	0	118	94	114	73	399
Centro 2	43	36	44	34	39	27	223
Centro 3	7	16	9	11	11	15	69
Centro 4	14	11	8	8	14	8	63
Centro 5	48	29	64	64	52	42	299
Centro 6	39	31	39	36	27	16	188
Total	151	123	282	247	257	181	1241

La muestra descrita anteriormente es *no probabilística* y ha sido elegida de forma *intencional* (Bisquerra, 2014), debido a que el cuestionario de evaluación fue aplicado en los centros educativos y cursos que expresaron disponibilidad, a los que se accedió por medio de la negociación a través de una entrevista con los directores de los centros educativos y profesores de la asignatura de matemática. No obstante, se pretendió conseguir la representatividad de la misma, considerando además de los factores de acceso, factores tales como el tipo de centro y la edad de los participantes.

En la Tabla 5.3, se observa la distribución de los estudiantes por tipo de centro, así 132 (10.6%) asisten a centros de dependencia particular pagada (privado), 411 (33.1%) a centros particulares subvencionados (concertados) y 698 (56.2%) a centros municipales (públicos).

Tabla 5.3. Distribución de estudiantes por tipo de centro

Tipo de Centro	Cantidad de estudiantes
	Frec. (%)
Municipal (Público)	698 (56.2)
Particular Subvencionado (Concertado)	441 (33.1)
Particular Pagado (Privado)	132 (10.6)
Total	1241

Los estudiantes respondieron los cuestionarios por escrito, completándolos en 1 hora 30 minutos, como una actividad de la clase de matemáticas, con el investigador presente, quien explicaba el objetivo de la evaluación, como también, contestaba las dudas que pudieran surgir sobre la manera de completar el cuestionario.

5.4. METODOLOGÍA DEL ANÁLISIS DE LOS DATOS

Una vez aplicado el cuestionario, se inició el proceso de creación del fichero de datos para posteriormente poder llevar a cabo el análisis de las respuestas entregadas por los participantes de este estudio a cada uno de los ítems planteados.

Para clasificar las respuestas proporcionadas por cada participante y favorecer su codificación e interpretación, a cada estudiante se le asignó el código En , donde n corresponde al número del 1 al 1241, de acuerdo al orden en que se han revisado los cuestionarios.

Posteriormente, para la creación del fichero de datos, cada alumno fue codificado como un elemento (fila), identificándose con un número y en las columnas, se agregaron algunos datos cualitativos necesarios para la caracterización de cada uno de los participantes (centro, tipo de centro, curso, edad, sexo), además, se establecieron las distintas categorías para clasificar las respuestas por ítem, a partir del análisis a priori expuesto en el Capítulo 4 de las respuestas a cada una de las preguntas planteadas en cada ítem.

Para definir las categorías de respuesta para cada ítem, se utilizó el método del análisis de contenido (López, 2002; Porta y Silva, 2003; Zapico, 2007), que se ha descrito en el Capítulo 3. Como se ha mencionado anteriormente, este método posibilita analizar documentos escritos a partir de unidades de análisis que no permiten interpretaciones subjetivas del escrito. Asimismo, con el propósito de clasificar el tipo de respuestas proporcionadas en el cuestionario, esta técnica ha sido utilizada en otras tesis doctorales, entre los cuales se pueden citar: Cobo (2003); Díaz-Levicoy (2018) y Gómez (2014).

Cada una de las categorías de análisis han sido establecidas a partir de las respuestas a priori detalladas previamente en el Capítulo 4, las que fueron codificadas numéricamente para facilitar el análisis cuantitativo de cada ítem, a través del uso del procesador de datos SPSS.

Una vez finalizado el proceso de codificación de datos, teniendo disponible el fichero con todos los datos obtenidos, se procede al análisis de las respuestas emitidas por los estudiantes, para ello los datos se analizan por ítem, se inicia con una descripción del ítem y de su objetivo de evaluación, posteriormente, se presenta el análisis de las respuestas correctas e incorrectas a cada uno de los problemas planteados en el Cuestionario, como también

la caracterización de las respuestas entregadas por los alumnos, empleando una metodología mixta: análisis cualitativo (descriptivo) y análisis cuantitativo (tablas de frecuencias y porcentajes). Además, se incluyen algunos ejemplos de respuestas entregadas por los estudiantes (transcripción literal) en los que se identifican patrones de respuesta.

5.5. RESULTADOS POR ÍTEM

En esta sección se describen los resultados obtenidos en cada uno de los ítems del cuestionario. Se presenta para cada uno de los ítems una descripción de la fuente las respuestas correctas e incorrectas tablas de frecuencias y porcentajes, se presentan las tablas cruzadas de frecuencia y porcentajes por curso al cual pertenecían los participantes la caracterización y transcripción de las respuestas proporcionadas por los estudiantes para describir las categorías de análisis, complementados con un análisis estadístico descriptivo de los resultados, comparando los tres cursos implicados en este estudio.

5.5.1. Resultados Ítem 1

El Ítem 1 del cuestionario (ver Figura 5.1) pertenece al grupo 1 de preguntas cuyo objetivo es *comprender la definición del concepto de muestra* (ver Tabla 4.1, Capítulo 4). Este primer ítem solicita a los estudiantes decir si han escuchado la palabra “*muestra*” anteriormente y explicar con sus palabras ese significado, es decir, mencionando algún contexto en que han oído el término. Esta pregunta se adaptó de la investigación de Watson (2004) acerca de la comprensión del razonamiento de los estudiantes de secundaria sobre el muestreo, en la que realizan entrevistas como parte de un estudio longitudinal.

Ítem 1. ¿Has oído la palabra muestra antes? Explica qué significa para ti.

Figura 5.1. Ítem 1 del cuestionario de evaluación

Al realizar síntesis de los resultados obtenidos en el Ítem 1 (ver Tabla 5.3), 944 estudiantes (76.1%) dan una respuesta correcta indicando que han oído la palabra muestra en algún contexto (salud, supermercado, tiendas, farmacias, etc.), aunque este no sea necesariamente en la clase de matemática.

Tabla 5.3. Respuestas correctas e incorrectas ítem 1

Curso	8° Básico	2° Medio	4° Medio	Total (%)
Respuestas	Frec. (%)	Frec. (%)	Fec. (%)	
Incorrectas	78 (6.3)	104 (8.4)	115 (9.3)	297 (23.9)
Correctas	196 (15.8)	425 (34.2)	323 (26)	944 (76.1)
Total	274	529	438	1241

Al analizar las respuestas de los alumnos, podemos destacar dentro de las respuestas correctas las siguientes: *“para mí muestra significa una pequeña parte de un todo”* (E476); *“Si. Una muestra es sacar un porcentaje o una parte de este para verificar algo”* (E691); *“para mi significa una cantidad pequeña”* (E61); *“Es extraer una parte de un total”* (E73); *“es una parte de algo”* (E 76).

En algunas respuestas correctas más elaboradas, podemos apreciar que los estudiantes evidencian mayor manejo del concepto, indicando términos clave como *población, porción, representativo*, proporcionando así los siguientes argumentos: *“Si. Para mi una muestra es un conjunto de cosas o personas elegidos al azar que se consideran representativos al grupo que pertenecen y con esto, se estudian las características de este grupo”* (E733); *“Sí, he escuchado la plabra “muestra” principalmente en matemática y los medios de comunicación al hablar de estudios, investigaciones y porcentajes. Para mi significa porción o porcentaje, ya que denominan con esta palabra a un extracto de un total”* (E745); *“Sí, significa que esa muestra es un dato sacado de una población, donde no todos son tomados en cuenta, se elige un grupo, al azar, y se cree que este representará la población”* (E725); *“Sí, la muestra es una parte que se toma de un grupo o población, esta muestra contra mayor sea cantidad será más representativo”* (E1108).

Al comparar las respuestas por curso (ver Tabla 5.3), se observa que el porcentaje más alto de respuesta correcta corresponde a los estudiantes de segundo medio (34.2%), seguidos por los alumnos de cuarto medio (26%) y por último los alumnos de octavo año de primaria (15.8%).

Cuando se analizan las respuestas incorrectas (ver Tabla 5.3), sólo 297 estudiantes (23.9%) no contestan o indican ejemplos en contextos poco claros, por ejemplo los alumnos mencionan ideas como: *“Sí. Muestra de sangre”* (E84); *“para mi la palabra muestra es como evidencia, prueba de algo”* (E35); *“Sí, para mi significa mostrar, enseñar, presentar”* (E80).

Además, se observa que el porcentaje de respuestas incorrectas se va incrementando a medida que la edad de los participantes aumenta.

Las respuestas proporcionadas por los alumnos al Ítem 1 (ver Figura 5.1), se presentan en la Tabla 5.4, en donde se observa que un muy bajo porcentaje no contesta esta pregunta (1.5%) o no ha escuchado anteriormente el concepto de muestra (2.1%), mientras que un alto porcentaje contesta de forma afirmativa que ha oído este concepto anteriormente (96.4%). Hay que mencionar además que, cuando indican que *es una parte de un conjunto más grande*, se ve que el porcentaje de respuesta va aumentando paulatinamente a medida que avanza la edad de los participantes. También, se aprecia que los estudiantes mayoritariamente afirman haber oído la palabra muestra, pero un 43.8% sólo es capaz de dar ejemplos en diversos contextos, pero no explica su significado. En este último caso, los alumnos entregan las siguientes respuestas: “*Dar a conocer algo*” (E66); “*mostrar o dar a conocer algún producto*” (E57); “*Sí, es un extracto de algo, de alguna comida o producto*” (estudiante 91)

Tabla 5.4. Análisis respuestas Ítem 1

Curso	8° Básico	2° Medio	4° Medio	Total
Respuestas	Frec. (%)	Frec. (%)	Frec. (%)	
No responde	5 (0.4)	8 (0.6)	5 (0.4)	18 (1.5)
No. Nunca	7 (0.6)	15 (1.2)	4 (0.3)	26 (2.1)
Si. Es una parte de un conjunto más grande	63 (5.1)	134 (10.8)	166 (13.4)	363 (29.3)
Si. Da ejemplos en contextos, pero no explica significado	133 (10.7)	270 (21.8)	141 (11.4)	544 (43.9)
Si. Da una Explicación sin fundamento	66 (5.3)	81 (6.5)	106 (8.5)	253 (20.3)
Si. Cuando dan a conocer algo	0 (0)	21 (1.7)	16 (1.3)	37 (3)
Total	274	529	438	1241

Por otro lado, siguiendo las categorías de análisis empleadas por Watson y Moritz (2000), se pueden clasificar las respuestas de los estudiantes de acuerdo a la Taxonomía SOLO (Biggs y Collis, 1982) como se muestra en la Tabla 5.5. En el *nivel relacional* los estudiantes entregan todos los elementos necesarios para responder a la tarea asignada, en este caso, expresan que una *muestra es una parte de un conjunto más grande o una parte que representa a un todo*; en el *nivel multiestructural* los alumnos mencionan varios aspectos relacionados al concepto, como por ejemplo: “*una prueba*”, “*una muestra de sangre*”; en el *nivel no estructural* mencionan una característica única o un ejemplo, así como “*cuando dan a conocer algo*”; y en el caso del *nivel pre-estructural*, los estudiantes dan una respuesta incompleta o idiosincrática, sin dar un ejemplo concreto de lo que entienden por muestra.

Tabla 5.5. Nivel de respuesta Ítem 1

Curso	8° Básico	2° Medio	4° Medio	Total (%)
Nivel de Respuesta	Frec. (%)	Frec. (%)	Frec. (%)	
Nivel Relacional	63 (5.1)	134 (10.8)	166 (13.4)	363 (29.3)
Nivel Multiestructural	133 (10.7)	270 (21.8)	141 (11.4)	544 (43.9)
Nivel No estructural	0 (0)	21 (1.7)	16 (1.3)	37 (3)
Nivel Pre-estructural	66 (5.3)	81 (6.5)	106 (8.5)	253 (20.3)
No responde / No sabe	12 (1)	23 (1.9)	9 (0.7)	44 (3.5)
Total	274	529	438	1241

Como se observa en los resultados de la Tabla 5.5 la mayoría de los estudiantes se encuentran en el *nivel multiestructural* (43,9%), donde indican varios aspectos relacionados al concepto de muestra, es decir, que reconocieron la relación parcial entre la muestra y la población, por ejemplo, E15 menciona un ejemplo de muestra en un contexto de salud: “*en muestras que piden en el hospital: de sangre; de orina*”; E240 indica un ejemplo de muestra en el contexto de una farmacia: “*muestra de algún perfume*”; E247 alude “*por lo que yo creo, es como un examen*”, es decir, relacionan el concepto de muestra con una prueba. Por otro lado, sólo un 29.3% de los estudiantes entregan todos los elementos necesarios para dar la respuesta a la pregunta, es decir, sus respuestas se encuentran en el *nivel relacional*, pues integraron la relación parte-todo, pues indican que han oído la palabra muestra anteriormente y la definen como *parte de un conjunto más grande*, por ejemplo, E680 explica: “*...de cierta cantidad de individuos o grupo de cosas se extrae una parte de ellos para sacar conclusiones de algún experimento*” y el E25 responde: “*Sí, una muestra es tomar o seleccionar una parte representativa de un conjunto x*”. Además, en el *nivel relacional* se ve que las respuestas mejoran con la edad de los participantes.

Sólo un 3% de los alumnos se clasifican en el *nivel no estructural* proporcionando definiciones que reflejaban la idea de parte sin el todo asociado que caracteriza la relación de la población con la muestra, o mencionan la idea de *dar a conocer algo*, por ejemplo, E237 señala que “*significa mostrar un poco de alguna cosa, tanto como para consumir, observar, analizar, etc.*”. El 20.3% de los estudiantes se clasificaron en el *nivel pre-estructural*, dando una respuesta incompleta o idiosincrática, por ejemplo, E462 menciona lo siguiente: “*Sí, he oído esa palabra “muestra”, para mí varía su significado según el contexto en el cual sea empleado*”.

Estos resultados son similares a los hallazgos de Watson y Moritz (2000), Watson (2004) y Meletiou-Mavrotheris y Papparistodemou (2015), quienes mencionan que los estudiantes habían oído hablar del concepto de muestra utilizando principalmente palabras en contextos fuera de la escuela (médico, científico, tiendas, farmacias, etc.), destacamos que en este estudio muy pocos estudiantes dan una respuesta más elaborada y más relacionada a un contexto de las matemáticas escolares, es decir, no usan el concepto de *aleatoriedad*, y sólo quince estudiantes definen muestra como un *subconjunto de la población*. Similar a los resultados descritos en los antecedentes, se puede observar que las descripciones proporcionadas por los alumnos de este estudio, sobre el significado de la palabra *muestra* también tienden a ser bastante sencillas, es decir, principalmente entregan ejemplos en diversos contextos y en menor medida entregan una definición más elaborada del concepto de muestra.

5.5.2. Resultados Ítem 2

El Ítem 2 (ver Figura 5.2) pertenece al grupo 1 de preguntas cuyo objetivo es *comprender la definición del concepto de muestra* (ver Tabla 4.1, Capítulo 4). Este ítem solicita a los estudiantes que identifiquen la palabra muestra en un contexto cercano a la realidad y que describan su significado en dicho contexto, además, den su opinión acerca de por qué se selecciona una muestra; si están o no de acuerdo a que se seleccione una muestra de tamaño pequeño y que sugieran un tamaño de muestra adecuado para seleccionar una muestra representativa de la población considerada en el enunciado de la situación problema. El ítem 2 ha sido adaptado de la investigación de Meletiou-Mavrotheris y Papparistodemou (2015) que pretende evaluar el razonamiento inferencial informal de estudiantes de secundaria. Se presentará el análisis de cada una de las preguntas que componen este ítem.

Ítem 2. En una investigación sobre los hábitos de ejercicio de los “estudiantes de secundaria chilenos”, algunos investigadores entrevistaron a una *muestra* de ellos.

1. ¿Qué significa la palabra muestra en esta frase?
2. ¿Por qué piensas que los investigadores seleccionan una *muestra* de estudiantes en vez de preguntar a todos?
3. ¿Estás de acuerdo con que los investigadores seleccionen una muestra de 10 estudiantes? ¿Cuántos elegirías tú?

Figura 5.2. Ítem 2 del cuestionario de evaluación

Al analizar las respuestas de la pregunta 1 del Ítem 2 por curso (ver Tabla 5.6), se

observa que un 85% de los estudiantes contesta correctamente a dicha pregunta, señalando principalmente que *una muestra es una parte o una pequeña parte de un conjunto más grande*, por ejemplo, el estudiante 535 menciona que una muestra “*es un subconjunto de individuos de una población*”. Por otro lado, sólo un 15% de los alumnos no es capaz de entregar una respuesta o indicar alguna de las características que permitan definir el concepto. Además, se aprecia que el 16.9% los estudiantes de 8° año de primaria responden correctamente esta pregunta, porcentaje que aumenta al considerar a los estudiantes de 2° Medio (35.8%) y se mantiene similarmente en los estudiantes de 4° Medio (32.2%)

Tabla 5.6. Respuestas correctas e incorrectas Ítem 2, pregunta 1

Curso	8° Básico	2° Medio	4° Medio	Total (%)
Respuestas	Frec. (%)	Frec. (%)	Frec. (%)	
Incorrectas	63 (5.1)	85 (6.8)	38 (3.1)	186 (15)
Correctas	211 (17)	444 (35.8)	400 (32.2)	1055 (85)
Total	274	529	438	1241

Ante la primera pregunta del Ítem 2 (ver Figura 5.2), los alumnos manifiestan que tienen alguna idea bastante clara para entender el concepto de muestra en un contexto cercano a su realidad, ya que sólo 27 (2.2%) no responden, como se ve en la Tabla 5.7. Además, un 56.7% señala que la palabra muestra en esa frase significa *una parte o una pequeña parte de un conjunto más grande*, destacando que dentro de ese porcentaje un 23.3% de los estudiantes de segundo medio manifiesta esa opinión, seguidos por un 21.1% de los estudiantes de cuarto medio y por un 12.3% de los alumnos de octavo año de primaria.

Tabla 5.7. Análisis respuestas Ítem 2, pregunta 1

Curso	8° Básico	2° Medio	4° Medio	Total (%)
Respuestas	Frec. (%)	Frec. (%)	Frec. (%)	
No responde	6 (0.5)	15 (1.2)	6 (0.5)	27 (2.2)
Incorrecto	57 (4.6)	70 (5.6)	32 (2.6)	159 (12.8)
Una parte, una pequeña parte	153 (12.3)	289 (23.3)	262 (21.1)	704 (56.7)
Un porcentaje	16 (1.3)	42 (3.4)	30 (2.4)	88 (7.1)
Una cantidad	39 (3.1)	107 (8.6)	102 (8.2)	248 (20)
Un subconjunto de la población	3 (0.2)	6 (0.5)	6 (0.5)	15 (1.2)
Total	274	529	438	1241

Se observa que un 12.8% (159) responden incorrectamente (ver Tabla 5.7), por ejemplo, E418 alude “*Nos enseña o nos muestra como son los hábitos de ejercicios de los estudiantes de secundaria chilenos*”, es decir, en este tipo de respuesta no identifican alguno de los elementos que forman parte de la definición de muestra que aparecen desglosados en

la Tabla 5.7. Por otro lado, un 83.8% (1040) menciona ideas como *una pequeña parte, un porcentaje, una cantidad*, y solamente quince estudiantes (1.2%) entregan una definición más precisa, entendiendo la palabra *muestra* como un *subconjunto de la población*, resumiendo, como se señaló anteriormente, un 85% de los estudiantes responde correctamente este ítem.

Cuando los estudiantes explican por qué los investigadores usaron una muestra, la mayoría (65.4%) dio respuestas que reflejaban cuestiones prácticas e ideas sobre el contexto en el que se realiza el muestreo (ver Tabla 5.8): “*porque sería muy difícil y costoso entrevistar a cada uno, por ello se entrevistan a unos pocos*” (E529).

Tabla 5.8. Respuestas correctas e incorrectas Ítem 2, pregunta 2

Curso	8° Básico	2° Medio	4° Medio	Total (%)
Respuestas	Frec. (%)	Frec. (%)	Frec. (%)	
Incorrectas	118 (9.5)	191 (15.4)	120 (9.7)	429 (34.6)
Correctas	156 (12.6)	338 (27.2)	318 (25.6)	812 (65.4)
Total	274	529	438	1241

En el caso de la pregunta 2 (ver Tabla 5.9), el 65.5% (812) tiene una idea más clara de los factores que son importantes al momento de seleccionar una muestra de una población, por ejemplo, mencionan ideas acerca del *excesivo tiempo* que puede llevar el estudio, el *costo de implementación* de la entrevista o sobre el *acceso* a entrevistar a la población total, es decir, un alto porcentaje de alumnos tiene en cuenta los factores que afectan a la selección de muestras representativas y solamente un 34.5% (429 estudiantes) contesta de manera incorrecta.

Tabla 5.9. Análisis respuestas Ítem 2, pregunta 2

Curso	8° Básico	2° Medio	4° Medio	Total (%)
Respuestas	Frec. (%)	Frec. (%)	Frec. (%)	
No responde	14 (1.1)	22 (1.8)	14 (1.1)	50 (4)
Incorrecto	104 (8.4)	169 (13.6)	106 (8.5)	379 (30.5)
Porque es difícil preguntar a todos	79 (6.4)	165 (13.3)	145 (11.6)	389 (31.3)
Porque es costoso	2 (0.2)	4 (0.3)	1 (0.08)	7 (0.6)
Porque lleva mucho tiempo	60 (4.8)	118 (9.5)	85 (6.8)	263 (21.2)
Es más rápido o fácil	3 (0.2)	18 (1.5)	41 (3.3)	62 (5)
Porque obtienen los mismos resultados	12 (1)	33 (2.7)	46 (3.7)	91 (7.3)
Total	274	529	438	1241

La principal razón que mencionan para elegir una muestra es *porque es difícil preguntar a todos* (31.3%) y como segunda elección señalan *porque lleva mucho tiempo*

(21.2%), siendo los alumnos de segundo medio los que más señalan dichas opciones, seguidos por los de cuarto medio y en menor medida los alumnos de octavo año de primaria.

En la pregunta 3 a del ítem 2, sólo el 35.6% contesta correctamente (ver Tabla 5.10), por ejemplo, E529 señala que: “*No, considero que son muy pocas personas para representar la totalidad de un país*”. A pesar que, en las preguntas anteriores, los alumnos muestran una comprensión elemental del concepto de muestra y las limitaciones de realizar un estudio preguntándole a todos los individuos de una determinada población, al momento de cuantificar los sujetos de la muestra sus concepciones no son tan acertadas. Además, cabría señalar que también estas respuestas pueden ser influenciadas en que se trabaja con una población desconocida, la cual es difícil de cuantificar, lo que induce a un error en la interpretación de la pregunta planteada. Nuevamente se observa que los estudiantes de segundo medio son los que responden mejor esta pregunta, seguidos por lo de cuarto medio y el porcentaje más bajo de respuesta correcta lo proporcionan los estudiantes de octavo año de primaria.

Tabla 5.10. Respuestas correctas e incorrectas Ítem 2, pregunta 3 a

Curso	8° Básico	2° Medio	4° Medio	Total (%)
Respuestas	Frec. (%)	Frec. (%)	Frec. (%)	
Incorrectas	195 (15.7)	326 (26.3)	278 (22.4)	799 (64.4)
Correctas	79 (6.4)	203 (16.4)	160 (12.8)	442 (35.6)
Total	274	529	438	1241

En las respuestas a la “pregunta 3 a” del Ítem 2 (ver Tabla 5.11), se observa que un 19.8% no responde; un 35.6% no está de acuerdo con que se elijan solamente 10 estudiantes como muestra, aludiendo razones como: “*No, no estoy de acuerdo, ya que considero que la muestra es muy pequeña comparada con la cantidad de alumnos de secundaria en Chile*” (E514); y un 44.6% si está de acuerdo con la afirmación planteada (seleccionar una muestra de 10 estudiantes), mencionando argumentos como: “*Estoy de acuerdo, ya que es una cifra aceptable*” (E515).

Tabla 5.11. Análisis respuestas Ítem 2, pregunta 3 a

Curso	8° Básico	2° Medio	4° Medio	Total (%)
Respuestas	Frec. (%)	Frec. (%)	Frec. (%)	
No responde	69 (5.6)	91 (7.3)	86 (6.9)	246 (19.8)
No estoy de acuerdo	79 (6.4)	203 (16.3)	160 (12.9)	442 (35.6)
Si estoy de acuerdo	126 (10.2)	235 (19)	192 (15.4)	553 (44.6)
Total	274	529	438	1241

Al analizar el porcentaje de respuestas correctas (ver Tabla 5.12), se observa que el 19.1% de los alumnos de segundo medio, el 18.7% de cuarto medio y el 9.7% de los estudiantes de octavo año de primaria responden correctamente la pregunta 3b del Ítem 2.

Tabla 5.12. Respuestas correctas e incorrectas Ítem 2, pregunta 3 b

Curso	8° Básico	2° Medio	4° Medio	Total (%)
Respuestas	Frec. (%)	Frec. (%)	Frec. (%)	
Incorrectas	154 (12.4)	292 (23.5)	206 (16.6)	652 (52.5)
Correctas	120 (9.7)	237 (19.1)	232 (18.7)	589 (47.5)
Total	274	529	438	1241

Al considerar la segunda interrogante planteada en la pregunta 3: *¿Cuántos elegirías tú?* (ver Tabla 5.13), hay un 34.3% indica que sólo elegiría a *10 estudiantes* como en el enunciado de la parte anterior de esta pregunta; el 36.8% señala que debería considerarse *más alumnos* dependiendo ya sea, si se considera el total de alumnos entrevistados, o el total de centros educativos que forman parte del estudio y sólo un 9.2% considera que debería entrevistarse *a todos los estudiantes* de secundaria chilenos, evidenciando que olvidan las razones entregadas en la pregunta anterior, sobre las factibilidades seleccionar una muestra representativa de la población.

Tabla 5.13. Análisis respuestas Ítem 2, pregunta 3 b

Curso	8° Básico	2° Medio	4° Medio	Total (%)
Respuestas	Frec. (%)	Frec. (%)	Frec. (%)	
No responde	15 (1.2)	29 (2.3)	23 (1.9)	67 (5.4)
10 estudiantes	91 (7.3)	198 (16)	137 (11)	426 (34.3)
Más alumnos	96 (7.7)	187 (15.1)	174 (14)	457 (36.8)
50%	13 (1.1)	27 (2.2)	21 (1.7)	61 (4.9)
Todos	35 (2.8)	45 (3.6)	34 (2.7)	114 (9.2)
Depende del total	11 (0.9)	22 (1.8)	35 (2.8)	68 (5.5)
Al azar	0 (0)	1 (0.1)	2 (0.1)	3 (0.2)
Menos de 10	13 (1.1)	20 (1.6)	12 (1)	45 (3.6)
Total	274	529	438	1241

Se puede señalar que un 56.6% de los alumnos presenta de manera más clara el hecho de que debe seleccionarse una muestra mayor a 10 estudiantes (considerando el conjunto de razones expresadas), aunque no aluden expresamente al concepto de *representatividad de la muestra*, algunos consideran que debe *depender del azar*, como también, *del total de alumnos entrevistados*.

Otro aspecto a considerar, es que los estudiantes son susceptibles a la *insensibilidad al tamaño de la muestra* (Tversky y Kahneman, 1974), es decir, asumen que la muestra seleccionada, independientemente de su tamaño, siempre representa a la población a la que

pertenece, lo que induce a graves errores de interpretación en los análisis estadísticos. Además, como indican Kahneman, Slovic y Tversky (1982), este tipo de personas, también cree en la *ley de los pequeños números*, es decir, creen que, al seleccionar muestras no suficientemente grandes, la distribución muestral se distribuye de la misma manera que la distribución de la población, independientemente del tamaño de la muestra.

Por otro lado, al solicitar que identifiquen la muestra de un estudio en un contexto cercano a su realidad, se aprecian algunas incongruencias en sus respuestas. Por ejemplo, son capaces de identificar las limitaciones que pueden llevar a un investigador a seleccionar una muestra para no considerar a todos los sujetos de una determinada población, pero al momento de sugerir una muestra o preguntarles cuál es el tamaño de muestra que ellos elegirían, un 43.4% de los estudiantes de nuestro estudio, responde incorrectamente.

5.5.3. Resultados Ítem 3

El Ítem 3 (ver Figura 5.3) tiene por objetivo *detectar sesgo en el método de muestreo* y pertenece al segundo grupo de preguntas.

Ítem 3. Los estudiantes de un colegio realizaron una encuesta para determinar el porcentaje de niños que reciclan en sus casas:

- Andrea preguntó a 80 estudiantes que son miembros del club de medioambiente.
- Elena envió un cuestionario a todos los niños del colegio y tomó los primeros 80 que contestaron.
- Pedro quiso el mismo número de niños y niñas. Así que preguntó a 5 niños y 5 niñas de cada curso para conseguir los 80 estudiantes que constituyen la muestra.
- Luis tenía los nombres de los 800 estudiantes en la escuela; puso cada nombre en un papel en un sombrero y sacó 80.
- Sara no conocía demasiados niños por lo que decidió encuestar a 80 niñas. Pero quería asegurarse de tener niñas de varias edades, por lo que tomó 10 niñas de cada nivel.

a. Completa la última columna donde preguntamos si el método de cada niño es apropiado.

Niño	A quien pregunta	¿Te parece un método apropiado? ¿Por qué?
Andrea	80 estudiantes del club de medioambiente.	
Elena	Envió un cuestionario a todos los niños del colegio y tomó los 80 que contestaron primero.	
Pedro	5 niños y 5 niñas de cada curso	
Luis	Puso el nombre de cada niño de la escuela (800) en un papel en un sombrero y sacó 80	
Sara	10 niñas de cada nivel.	

b. De acuerdo a los datos entregados en el enunciado, ¿Cuál crees que es la mejor manera de escoger a los niños para estimar el porcentaje de ellos que recicla en casa?

Figura 5.3. Ítem 3 del cuestionario de evaluación

En el Ítem 3 se les solicita a los estudiantes señalar si cada uno de los métodos de selección de muestras planteados son o no apropiados e indicar su justificación a la respuesta, además, deben señalar cuál de los métodos propuestos es el mejor para la selección de una muestra representativa.

Este ítem ha sido adaptado del trabajo de Meletiou-Mavrotheris y Papparistodemou (2015) y del de Watson y Kelly (2005), enunciado que ha sido adaptado previamente de la investigación de Jacobs (1997), cuyo objetivo es conocer la comprensión de los estudiantes sobre el sesgo del muestreo en un contexto cercano a la realidad, como es la toma de encuestas.

Se incluyen tres métodos de muestreo de primarias en este Ítem: muestreo aleatorio, restringido y autoseleccionado. El método de *muestreo aleatorio* dio a cada miembro de la población la misma oportunidad de ser seleccionados. Los métodos de *muestreo restringidos* preguntaron a grupos particulares de personas que podían ser más capaces de seleccionar una respuesta determinada y, en consecuencia, sesgar los resultados en una dirección particular. Los métodos de *muestreo autoseleccionados* permitieron que los participantes se seleccionaran ellos mismos y fueran problemáticos porque no había suficiente evidencia para evaluar si la muestra era representativa (Jacobs, 1997).

Caracterizando los métodos de muestreos expuestos en el enunciado del problema, se puede apreciar que el método de Andrea y Sara corresponden a un *muestreo restringido*; el método de Elena es un *muestreo autoseleccionado*; el método de Luis y el de Pedro son *muestreos aleatorios*; y más específicamente el método de Pedro corresponde a un *muestreo estratificado*.

Para analizar las respuestas estadísticas apropiadas para cada método señalamos que, en el caso del método de Andrea, en el método de Elena y en el método de Sara, el estudiante debe indicar que cada uno de esos métodos no son apropiados porque su selección de la muestra es sesgada, en cambio, en el método de Pedro y en el método de Luis deben indicar que el método si es apropiado, ya que el primero corresponde a un muestreo estratificado y el segundo a un muestreo aleatorio simple.

En una primera parte, se presentarán los datos resultados de respuestas correctas e incorrectas para cada uno de los métodos propuestos en el enunciado del ítem.

Posteriormente, se presentarán resultados más detallados de las respuestas proporcionadas por los estudiantes las que han sido clasificadas en una serie de categorías, además en dichos análisis, se incluye la comparación por curso.

Tabla 5.14. Respuestas correctas e incorrectas Ítem 3. Método de Andrea

Curso	8° Básico	2° Medio	4° Medio	Total (%)
Respuestas	Frec. (%)	Frec. (%)	Frec. (%)	
Incorrectas	246 (19.8)	482 (38.8)	387 (31.2)	1115 (89.8)
Correctas	28 (2.3)	47 (3.8)	51 (4.1)	126 (10.1)
Total	274	529	438	1241

En la Tabla 5.14 se observa que solamente un 10.1% de los estudiantes contesta correctamente el Método de Andrea, indicando la *respuesta estadística adecuada*: “No es apropiado, porque se excluye a los demás estudiantes” (ver descripción página 195 y Tabla 5.19), por ejemplo, E400 responde que: “No, pues hay gente fuera del club que también puede opinar y responder”. Además, un alto porcentaje (89.8%) de los alumnos entregan una respuesta incorrecta, por ejemplo, E405 señala que: “Sí, porque los del club les gusta reciclar”, las respuestas entregadas por los estudiantes a este método se caracterizan más detalladamente en la descripción de la Tabla 5.20. Principalmente estos resultados se deben a que los estudiantes no identifican que este método es un procedimiento sesgado, ya que Andrea solo encuesta a alumnos que pertenecen al club de medio ambiente, es decir, que poseen una característica particular y corresponde a un muestreo restringido.

Tabla 5.15. Respuestas correctas e incorrectas Ítem 3. Método de Elena

Curso	8° Básico	2° Medio	4° Medio	Total (%)
Respuestas	Frec. (%)	Frec. (%)	Frec. (%)	
Incorrectas	211 (17)	435 (35)	363 (29.3)	1009 (83.3)
Correctas	63 (5.1)	94 (7.6)	75 (6)	232 (18.7)
Total	274	529	438	1241

Para el caso del Método de Elena (ver Tabla 5.15), podemos apreciar que las respuestas mejoran levemente con respecto al método anterior, donde un 18.7% de los estudiantes entrega una respuesta correcta, señalando la *respuesta estadística adecuada*: “No es apropiado, porque se excluye a los demás estudiantes”, en este caso, el estudiante 416 menciona: “No, porque puede que esos 80 niños no reciclen, pero los otros sí”. Un 83.3% de los estudiantes contesta incorrectamente, no identificando que también es un método sesgado, ya que no se eligen los participantes de forma aleatoria y se eligieron a los 80 alumnos que contestaron primero, es decir, corresponde a un método de muestreo

autoseleccionado. Por ejemplo, el estudiante 423 contesta incorrectamente indicando: “*me parece bien, aunque a veces los que contestan primero no lo hacen a conciencia*”

En la Tabla 5.16 para el Método de Pedro, se ve una considerable mejora en la cantidad de respuestas correctas ya que un 56.8% de los estudiantes entrega la *respuesta estadística adecuada*: “*Sí es apropiado, porque es equitativo*”. En este caso podemos mencionar que los estudiantes reconocen la *estratificación de la muestra* y manifiestan que es un *método justo o equitativo*, ya que da la posibilidad de que la misma cantidad de hombres como mujeres den su opinión.

Tabla 5.16. Respuestas correctas e incorrectas Ítem 3. Método de Pedro

Curso	8° Básico	2° Medio	4° Medio	Total (%)
Respuestas	Frec. (%)	Frec. (%)	Frec. (%)	
Incorrectas	130 (10.5)	226 (18.2)	180 (14.5)	536 (43.2)
Correctas	144 (11.6)	303 (24.4)	258 (20.8)	705 (56.8)
Total	274	529	438	1241

En el caso del Método de Luis (ver Tabla 5.17) un 40.8% de los estudiantes responde correctamente, identificando que se trata de un *muestreo aleatorio simple* (puso el nombre de cada niño de la escuela (800) en un papel en un sombrero y sacó 80) y entregan la *respuesta estadísticamente apropiada*: “*Sí es apropiado, porque es un método aleatorio*”. Por otro lado, un 59.2% de los estudiantes contesta incorrectamente, no identificando este proceso como un muestreo aleatorio y señalando principalmente que no es la mejor forma de seleccionar a un grupo porque es un método difícil o que lleva mucho tiempo,

Tabla 5.17. Respuestas correctas e incorrectas Ítem 3. Método de Luis

Curso	8° Básico	2° Medio	4° Medio	Total (%)
Respuestas	Frec. (%)	Frec. (%)	Frec. (%)	
Incorrectas	170 (13.7)	329 (26.5)	236 (19)	735 (59.2)
Correctas	104 (8.4)	200 (16.1)	202 (16.3)	506 (40.8)
Total	274	529	438	1241

Finalmente, para el Método de Sara (ver Tabla 5.18) un 43.8% señala una respuesta correcta, indicando como *respuesta estadística apropiada*: “*No es apropiado, porque deberían elegirse estudiantes de ambos sexos*”. Un 56.2% entrega una respuesta incorrecta, mayoritariamente porque no identifican que se ha seleccionado un grupo sesgado, ya que solo encuesta a 10 niñas de cada nivel, dejando sin participación a los niños y también corresponde a un muestreo restringido.

Tabla 5.18. Respuestas correctas e incorrectas Ítem 3. Método de Sara

Curso	8° Básico	2° Medio	4° Medio	
Respuestas	Frec. (%)	Frec. (%)	Frec. (%)	Total (%)
Incorrectas	158 (12.7)	308 (24.8)	231 (18.6)	697 (56.2)
Correctas	116 (9.3)	221 (17.8)	207 (16.7)	544 (43.8)
Total	274	529	438	1241

Para presentar un análisis más detallado de las respuestas entregadas por los estudiantes a la primera parte del Ítem 3, se proponen en la Tabla 5.19 una serie de categorías de análisis, las cuales se han empleado previamente en Ruiz-Reyes, Ruz, Molina-Portillo y Contreras (enviado), y han sido adaptadas del trabajo de Watson y Kelly (2005) además de considerar las respuestas obtenidas en esta memoria, las que se detallan a continuación.

En la categoría “*respuesta estadística apropiada*”, se describe la respuesta correcta a cada método, es decir, reconocen el método de muestreo y el sesgo en la muestra, además, emplean críticas apropiadas; en la categoría “*ideas no centrales o imparcialidad*” se consideran las respuestas que evalúan cierta parte de la información, más específicamente en relación con la imparcialidad (es decir, que todos los sujetos de la muestra tengan la misma oportunidad de ser seleccionados) y el tamaño de la muestra. Cabe mencionar además que la categoría “*ideas no centrales o imparcialidad*” se subdivide en *dos subcategorías* para algunos métodos, de acuerdo a las respuestas entregadas por los estudiantes, ya sea por el *tamaño de la muestra* o por características del método de selección en donde se destacan ideas como la *justicia* o la *imparcialidad* para seleccionar a los sujetos (ver Tabla 5.19)

En el caso de la categoría “*análisis inapropiado*”, se incluyen las respuestas que entregan una crítica inapropiada al método propuesto, las que se centran en la imprecisión percibida del método aleatorio, la injusticia o la imparcialidad, falta de oportunidad y el tamaño pequeño de la muestra. Además, en el método de Luis y en el de Sara, sólo se ha propuesto una subcategoría en la categoría de “*ideas no centrales o imparcialidad*”, ya que las respuestas de los estudiantes sólo se ajustan a la subcategoría propuesta, no evidenciando incluir otra subcategoría (ver Tabla 5.19). Finalmente, las respuestas de la categoría “*lógica inapropiada*” señalan todos los razonamientos erróneos que realizan los estudiantes.

Posteriormente, al resumir los resultados de cada método, se agrega una quinta categoría: “*No responde*”, para cuantificar cuántos estudiantes no entregan una respuesta al Ítem planteado.

A continuación, para facilitar la visualización de los resultados, analizaremos las respuestas entregadas por los estudiantes a cada uno de los métodos propuestos en el enunciado del Ítem 3 y que se clasifican en cada una de las categorías propuestas en la Tabla 5.19, como también se incluye la diferenciación por curso.

Tabla 5.19. Respuesta estadística adecuada Ítem 3, pregunta a. Adaptadas de Watson y Kelly (2005)

Categorías	Método de Andrea	Método de Elena	Método de Pedro	Método de Luis	Método de Sara
Respuesta estadística apropiada	Detección de sesgo: No es apropiado, porque se excluye a los demás estudiantes.	Crítica apropiada: No es apropiado, porque se excluye a los demás estudiantes.	Muestreo Estratificado. Sí es apropiado, porque es equitativo.	Muestreo aleatorio. Sí es apropiado, porque es un método aleatorio.	Falta de rango y / o variación: No es apropiado, porque deberían elegirse estudiantes de ambos sexos.
Ideas centrales o imparcialidad	Tamaño de muestra adecuado: No es apropiado, porque todos contestaron positivamente. Justicia / Imparcialidad: No es apropiado, porque son un grupo específico.	Tamaño de muestra adecuado: Sí es apropiado, porque tomó la opinión de diferentes alumnos. Justicia / Imparcialidad: No es apropiado, porque la gente que pensaba que era una mala idea no se molestaría en contestar.	Tamaño de la muestra adecuado: Sí es apropiado, porque las respuestas serán variadas. Justicia / Imparcialidad: Sí es apropiado, porque se hace una imagen general.	Tamaño de muestra adecuado: Sí es apropiado, hay mucha gente.	Justicia / Imparcialidad: No es apropiado, porque solo se escogieron niñas.
Análisis inapropiado	Creando sesgo: Sí es apropiado, porque es obvio que reciclan.	Creando sesgo: Sí es apropiado, porque fueron los que quisieron participar.	Creando sesgo: No es apropiado, porque debería preguntarle a todos los niños y niñas	Método demasiado aleatorio: No es apropiado, porque podría escoger a las personas equivocadas	Creando sesgo: Sí es apropiado, porque al tener diferentes edades tendrá diferentes respuestas.
Lógica inapropiada	Malinterpretar la pregunta: Sí es apropiado, porque es una gran cantidad de estudiantes.	Malinterpretar la pregunta: Sí es apropiado, porque son los primeros en contestar	Malinterpretar la pregunta: No es apropiado, son muy pocos estudiantes.	Malinterpretar la pregunta: No es apropiado, es un método difícil.	Malinterpretar la pregunta: Sí es apropiado, porque siempre están de acuerdo en cuidar el medio ambiente.

En la Tabla 5.20 se exponen los resultados obtenidos para el Método de Andrea, en este caso el 10.2% de los estudiantes entrega la respuesta estadística apropiada: “*No es apropiado, porque se excluye a los demás estudiantes*” identificando que es un *método de muestreo restringido* y que se escoge a una muestra sesgada.

Dentro de las “*respuestas estadísticas apropiadas*” para el método de Andrea, por ejemplo, se encuentran afirmaciones como: “*No, me parece porque debería haber preguntado a diferentes cursos*” (E299); “*No, Andrea debió preguntar a niños al azar*” (E508); “*No, no es equitativo*” (E738).

Analizando las respuestas entregadas a la categoría de “*ideas no centrales o imparcialidad*” el 31.6% señala que no es un método apropiado porque “*al pertenecer al club de medio ambiente, todos los participantes contestan positivamente*” (E234). Sólo un 11.2% señala que no es apropiado porque “*es un grupo específico*” (E266).

En la categoría de “*análisis inapropiado*” un 19.6% de los alumnos indican que es un método apropiado “*porque es obvio que reciclan*” (E214), es decir, en este caso contestan de forma incorrecta y no identifican que se está preguntando a un grupo sesgado.

Por lo que se refiere a la categoría “*lógica inapropiada*”, sólo 9 estudiantes sugieren que es un método apropiado porque “*es una gran cantidad de estudiantes*” (E255) y un 26.7% *no responde*.

Tabla 5.20. Frecuencias de las categorías de análisis del Ítem 3, pregunta a. Método de Andrea

Curso	8° Básico	2° Medio	4° Medio	Total
Categorías	Frec. (%)	Frec. (%)	Frec. (%)	Frec. (%)
Respuesta estadística apropiada	28 (2.3)	47 (3.8)	51 (41.1)	126 (10.2)
Ideas no centrales o imparcialidad	66 (5.3)	175 (14.1)	151 (12.2)	392 (31.6)
	29 (2.3)	60 (4.8)	50 (4)	139 (11.2)
Análisis inapropiado	65 (5.2)	97 (7.8)	82 (6.6)	244 (19.6)
Lógica inapropiada	1 (0.1)	2 (0.2)	6 (0.5)	9 (0.7)
No responde	85 (6.9)	148 (11.9)	98 (7.9)	331 (26.7)
Total	274	529	438	1241

Por otro lado, cuando se les presenta a los alumnos un método de muestreo autoseleccionado, como es el caso de Elena se obtiene sólo un 18.7% de respuestas correctas (ver Tabla 5.21). En la categoría *respuesta estadística apropiada*, se encuentran respuestas como: “*No me parece el correcto, pues no está siendo justa con todos los estudiantes*” (E519); “*No es lo mejor porque así puede obtener respuestas de ciertos cursos no más, no te*

asegura que sean niños de todos los cursos” (E674); *“No, ya que no es equitativo porque podrían haber contestado solo niños de un curso menor o mayor”* (E682).

En la categoría *“ideas no centrales o imparcialidad”*, el 2.4% indica que es un método con tamaño de la muestra apropiado porque *“se tomó la opinión a distintos alumnos”* (E287), y un 16% señala que no es apropiado porque *“la gente que piensa que era una mala idea no contestaría”* (E222).

Al analizar las respuestas entregadas en la categoría de *“análisis inapropiado”* el 9.7% de los estudiantes menciona que el método es apropiado *“porque fueron los que quisieron participar”* (E236), en este caso, no logran ver que la representatividad de la muestra se ve comprometida por el método de selección. Además, cabe mencionar que dentro de las respuestas de *“lógica inapropiada”* (1.7%), varios estudiantes mencionan que la muestra ha sido bien seleccionada, ya que consideran que los participantes que respondieron primero y rápidamente son los dan la mejor respuesta. Se puede apreciar que un alto porcentaje de estudiantes no responde este método (51.5%).

Tabla 5.21. Frecuencias de las categorías de análisis del Ítem 3, pregunta a. Método de Elena

Curso	8° Básico	2° Medio	4° Medio	Total
Categorías	Frec. (%)	Frec. (%)	Frec. (%)	Frec. (%)
Respuesta estadística apropiada	63 (5.1)	94 (7.6)	75 (6)	232 (18.7)
Ideas no centrales o imparcialidad	6 (0.5)	9 (0.7)	15 (1.2)	30 (2.4)
	29 (2.3)	109 (8.8)	61 (4.9)	199 (16)
Análisis inapropiado	25 (2)	53 (4.3)	42 (3.4)	120 (9.7)
Lógica inapropiada	1 (0.1)	8 (0.6)	12 (1)	21 (1.7)
No responde	150 (12.1)	256 (20.6)	233 (18.8)	639 (51.5)
Total	274	529	438	1241

En relación al método de Pedro (elige 5 niños y 5 niñas de cada curso), los alumnos consideran más apropiado este *método de muestreo estratificado*, teniendo el 56.8% de *respuestas estadísticas apropiadas*, en este caso en sus argumentos, también mencionan palabras como *“igualdad”* o *“es justo”*. En esta categoría se hallaron las siguientes afirmaciones: *“Sí, porque fue equitativo”* (E251); *“Sí, porque todos los cursos participaron”* (E48); *“Sí, porque es justo”* (E1115). Cabe mencionar que es el método de muestreo que mejor identifican los alumnos de este estudio (ver Tabla 5.22).

En la categoría de *“ideas no centrales o imparcialidad”*, un 3.7% de los estudiantes señalan que es apropiado porque las *“respuestas serán variadas”* (E499) o un 1.1% indica porque *“se hace una imagen general”* (E301).

Al examinar las respuestas proporcionadas en la categoría “*análisis inapropiado*”, solo un 7.1% de los alumnos manifiesta que no es apropiado “*porque debería preguntarles a todos*” (E327), en este caso, no identifican el muestreo estratificado y quieren dar a todos los participantes la oportunidad de ser elegidos.

Por lo que se refiere a la categoría “*lógica inapropiada*”, un 1.5% indica que el método de Pedro no es apropiado, “*son muy pocos estudiantes*” (E214), finalmente, solo un 29.8% de los alumnos no responde.

Tabla 5.22. Frecuencias de las categorías de análisis del Ítem 3, pregunta a. Método de Pedro

Curso	8° Básico	2° Medio	4° Medio	Total
Categorías	Frec. (%)	Frec. (%)	Frec. (%)	Frec. (%)
Respuesta estadística apropiada	144 (11.6)	303 (24.4)	258 (20.8)	705 (56.8)
Ideas no centrales o imparcialidad	11 (0.9)	13 (1.1)	22 (1.7)	46 (3.7)
	3 (0.2)	7 (0.6)	4 (0.3)	14 (1.1)
Análisis inapropiado	26 (2.1)	40 (3.3)	21 (1.7)	87 (7.1)
Lógica inapropiada	2 (0.2)	12 (1)	5 (0.4)	19 (1.5)
No responde	88 (7.1)	154 (12.4)	128 (10.3)	370 (29.8)
Total	274	529	438	1241

En la categoría de “*respuesta estadística apropiada*”, sólo un 40.8%, señala que es apropiado el *método de muestreo aleatorio* propuesto por Luis (ver Tabla 5.23). En Jacobs (1997), se menciona que algunos niños se enfocaron en el sesgo potencial y evaluaron de manera positiva los métodos de muestreo aleatorio porque se preguntó a distintas personas y razonaron que ese aspecto es importante porque diferentes tipos de personas tienen opiniones diferentes. En dicha categoría, aparecieron afirmaciones como: “*me parece apropiado porque de tal manera salen opiniones bastantes variadas*” (E332); “*Sí, es una muestra aleatoria*” (E673).

En el caso de las respuestas a la categoría “*ideas no centrales o imparcialidad*”, un 13.3% indica que es apropiado debido a que “*hay mucha gente*” (E251).

Se debe agregar que en la categoría de “*análisis inapropiado*”, un 3.8% de los alumnos considera que el método no es apropiado “*porque podría escoger a las personas equivocadas*” (E517), es decir, olvidan la importancia de la aleatoriedad en el proceso de selección de la muestra y tienen ideas como “*puede seleccionar a varios del mismo curso*” o “*puede elegir a varios que no reciclen*”.

Por lo que se refiere a la categoría “*lógica inapropiada*”, un 16.1% señala que el método no es apropiado porque “*es un método difícil*” o que “*lleva mucho tiempo*” (E247),

no considerando el factor aleatorio de la elección de cada individuo que forma parte de la muestra. Finalmente, un 27% no responde este apartado.

Tabla 5.23. Frecuencias de las categorías de análisis del Ítem 3, pregunta a. Método de Luis

Curso	8° Básico	2° Medio	4° Medio	Total
Categorías	Frec. (%)	Frec. (%)	Frec. (%)	Frec. (%)
Respuesta estadística apropiada	104 (8.4)	200 (16.1)	202 (16.3)	506 (40.8)
Ideas no centrales o imparcialidad	46 (3.7)	65 (5.2)	42 (3.4)	153 (13.3)
Análisis inapropiado	10 (0.8)	17 (1.4)	20 (1.6)	47 (3.8)
Lógica inapropiada	31 (2.5)	91 (7.3)	78 (6.3)	200 (16.1)
No responde	83 (6.7)	156 (12.6)	96 (7.7)	335 (27)
Total	274	529	438	1241

Cuando los alumnos analizan el método de muestreo restringido de Sara (elegir 10 niñas de cada curso), es uno de los que identifican mejor ya que un 43.8% de estudiantes indica que la muestra está sesgada, es decir, reconocen que no es correcto seleccionar a una muestra donde los sujetos cumplan con una característica en particular y que, para realizar un método de muestreo apropiado, se deberían incluir tanto niños como niñas en dicha muestra (ver Tabla 5.24). Jacobs (1997) en su estudio reporta que los niños que identificaron el sesgo potencial del método de muestreo, evaluaron negativamente los métodos de muestreo restringidos y discutieron que es probable que este tipo de métodos produzcan muestras que contengan individuos con una opinión particular.

Algunos ejemplos de las justificaciones de los estudiantes en la categoría “*respuesta estadística apropiada*” son: “*No, ya que al haber elegido puras niñas puede que no puedan representar las ideas que tienen los varones*” (E278); “*No, porque los niños también deben dar su opinión*” (E308); “*No, debería ser parejo, preguntar a niños y niñas*” (E309); “*No, porque debió tomar a niños y niñas para hacerlo más completo*” (E689).

En relación con la categoría “*ideas no centrales o imparcialidad*”, un 12.4% sugiere que el método no es apropiado porque “*solo se escogieron niñas*” (E221).

En la categoría “*análisis inapropiado*”, el 20.1% basan su opinión en que es un método apropiado debido a que “*selecciona niñas de distintas edades y obtendrá diferentes respuestas*” (E287), olvidándose que sólo corresponde a un grupo en particular y que su respuesta es sesgada.

Por lo que se refiere a la categoría “*lógica inapropiada*”, un 8.8% indica que es un método apropiado porque “*siempre están de acuerdo en cuidar el medio ambiente*” (E378).

Tabla 5.24. Frecuencias de las categorías de análisis del Ítem 3, pregunta a. Método de Sara

Curso	8° Básico	2° Medio	4° Medio	Total
Categorías	Frec. (%)	Frec. (%)	Frec. (%)	Frec. (%)
Respuesta estadística apropiada	116 (9.4)	221 (17.8)	207 (16.7)	544 (43.8)
Ideas no centrales o imparcialidad	25 (2)	75 (6)	54 (4.4)	154 (12.4)
Análisis inapropiado	75 (6)	93 (7.5)	82 (6.6)	250 (20.1)
Lógica inapropiada	15 (1.2)	42 (3.4)	52 (4.2)	109 (8.8)
No responde	43 (3.5)	98 (7.9)	43 (3.5)	184 (14.8)
Total	274	529	438	1241

Los resultados obtenidos al analizar el Ítem 3 se respaldan con los hallazgos descritos en la literatura existente, es decir, en los trabajos de Jacobs (1997), Watson et al. (2003), Watson y Kelly (2005) y los de Meletiou-Mavrotheris y Papanastasiou (2015).

En este estudio se puede mencionar que en uno de los casos de muestreo restringido (método de Sara), un alto porcentaje de los alumnos (58.1%) identifica que es un método sesgado debido a que sólo se selecciona un grupo específico de personas (solo mujeres), es decir, consideran que el método no es apropiado porque deberían elegirse estudiantes de ambos sexos. En cambio, en el método de Andrea, que también corresponde a un muestreo restringido, ignoran la restricción que los alumnos seleccionados en la muestra corresponden a un grupo específico (pertenecen al club de medio ambiente) que no es representativo de la población, y sólo un 25.7% señala que el método no es apropiado, porque se excluye a los demás estudiantes.

Otros alumnos, prefirieron el muestreo estratificado (método de Pedro, 52%), no para aumentar la representatividad del muestreo, sino para aumentar la “imparcialidad” del proceso de selección, asegurando que los niños y las niñas estén representados por igual. Jacobs (1997), encontró que los niños seguían insistiendo en seleccionar un número igual de cada género y concluyó que la preferencia por el muestreo estratificado puede estar vinculada a las concepciones estadísticamente no normativas de la imparcialidad de los niños. En cambio, a pesar de proponerse un método aleatorio (método de Luis), en donde todos los sujetos tienen la misma posibilidad de participar, sólo el 48% de los estudiantes lo considera un método apropiado, pero un 26.4% da una respuesta con una *lógica inapropiada*, centrándose principalmente en cuestiones prácticas, no en la forma de selección de los sujetos de la muestra, aludiendo a justificaciones como “no es apropiado, porque es un método muy difícil” o “no es apropiado, porque lleva mucho tiempo”.

Al enfrentarse los alumnos a diferentes métodos de muestreo empleando encuestas, les cuesta identificar las ventajas del muestreo aleatorio (método de Luis), siendo además incapaces de detectar los sesgos en los muestreos autoseleccionados (método de Elena). No pudieron detectar el sesgo involucrado en la autoselección debido a la preocupación por cuestiones de imparcialidad, en un sentido coloquial, que tanto niñas y niños tuvieran la misma posibilidad de participar; o que la participación voluntaria es para ellos “*más justa*”, ya que a todos se les dio la misma oportunidad de responder el cuestionario, por lo que sólo el 22.3% detecta correctamente que es un método sesgado. Resultados similares se aprecian en las investigaciones de Jacobs (1997) y Meletiou-Mavrotheris y Papparistodemou (2015).

Cuando se les solicita señalar cuál de los métodos propuestos es el más adecuado para escoger una muestra de estudiantes que recicla, un 64.2% de los estudiantes contestan correctamente (ver Tabla 5.25), señalando el método de Pedro o el método de Luis como el más apropiado, es decir, prefieren el *muestreo estratificado* o el *muestreo aleatorio simple*, por sobre los métodos sesgados, es decir, mayoritariamente identifican los métodos aleatorios como más apropiados para la selección de una muestra representativa de una población. Solo un 35.8% de los alumnos contesta incorrectamente.

Al analizar las respuestas por curso, en el caso de las respuestas correctas, se aprecia que los estudiantes de segundo y cuarto medio obtienen resultados similares (25.7% y 25.1% respectivamente) y los alumnos de octavo año de primaria presentan menores resultados ya que solo el 13.4% contesta correctamente.

Tabla 5.25. Respuestas correctas e incorrectas Ítem 3, pregunta b

Curso	8° Básico	2° Medio	4° Medio	Total (%)
Respuesta	Frec. (%)	Frec. (%)	Frec. (%)	
Incorrecta	108 (8.7)	210 (16.9)	126 (10.2)	444 (35.8)
Correcta	166 (13.4)	319 (25.7)	312 (25.1)	797 (64.2)
Total	274	529	438	1241

Para el análisis de las respuestas del “apartado b” del Ítem 3, se detallan en la Tabla 5.26 todas las respuestas entregadas por los estudiantes.

Si se analizan los resultados individuales del método elegido, se observa que 563 estudiantes (45.4%) prefieren el método estratificado de Pedro por sobre el método de muestreo aleatorio de Luis, ya que solo 234 estudiantes (18.9%), consideran que el método

de Luis es el más apropiado (ver Tabla 5.26). Por otro lado, solo el 11.9% no contesta esta pregunta y solo siete estudiantes no señalan ningún método.

Tabla 5.26. Análisis respuestas Ítem 3, pregunta b

Curso	8° Básico	2° Medio	4° Medio	Total (%)
Respuesta	Frec. (%)	Frec. (%)	Frec. (%)	
No contesta	35 (2.8)	80 (6.5)	32 (2.6)	147 (11.9)
Ninguno	6 (0.5)	0 (0)	1 (0.1)	7 (0.6)
Andrea	13 (1.1)	27 (2.2)	14 (1.1)	54 (4.4)
Elena	16 (1.3)	41 (3.3)	36 (2.9)	93 (7.5)
Pedro	111 (8.9)	230 (18.5)	222 (17.9)	563 (45.4)
Luis	55 (4.4)	89 (7.2)	90 (7.3)	234 (18.9)
Sara	13 (1.1)	16 (1.3)	5 (0.4)	34 (2.7)
Otro	25 (2)	46 (3.7)	38 (3.1)	109 (8.8)
Total	274	529	438	1241

Al analizar la categoría “otro” (ver Tabla 5.27) se observa que los estudiantes mencionan otros métodos para seleccionar una muestra representativa, en este caso, entregan solamente dos razones: *preguntar a la mitad del curso* o *preguntar a todos los alumnos*. En el primer caso, suponemos que esta creencia se debe a sus conocimientos de *probabilidad*, ya que atribuyen la *equiprobabilidad* para asegurarse que la muestra sea escogida apropiadamente y en el segundo caso, olvidan las razones señaladas en el ítem 2 para la elección de muestras representativas sobre la factibilidad de escoger una muestra para generalizar las conclusiones sobre una población.

Tabla 5.27. Análisis respuestas Ítem 3, pregunta b, categoría “otro método”

Curso	8° Básico	2° Medio	4° Medio	Total (%)
Respuesta	Frec. (%)	Frec. (%)	Frec. (%)	
Preguntaría a la mitad del curso	5 (0.4)	4 (0.3)	5 (0.4)	14 (1.1)
Preguntaría a todos	20 (1.6)	37 (3)	29 (2.3)	86 (6.9)
Total	25	46	38	109

En este estudio se aprecia que, la mayoría de los estudiantes no optan por señalar el muestreo aleatorio simple, como el método más apropiado para la selección de muestras, ya que solo un 20.9% elige el método de Luis como más apropiado, en cambio, un 52.7% evidencia más aprecio hacia la estratificación de la muestra, debido a que consideran el método de Pedro *más justo* o *equitativo*. Algunos estudiantes parecían estar conscientes del sesgo potencial en el muestreo, pero desestimaron el muestreo aleatorio porque estaban preocupados de que pudiera conducir a resultados extremos: “*No me parece un método apropiado, porque si bien se seleccionan al azar no será posible controlar otras variables*”

(E476), o mencionan argumentos relacionados a la imparcialidad: “*No, porque debería sacar una cantidad de alumnos y alumnas para quedar parejo*” (E1092). Resultados similares se obtienen en Watson et al. (2003) quienes señalan que a los estudiantes no les gusta generalizar a partir de una muestra aleatoria debido a la variabilidad inherente dentro de la población, sino que prefieren un método de muestreo estatificado para representar las características de la población por sobre el método de muestreo aleatorio.

Muy a menudo los estudiantes citaron razones no estadísticas en sus evaluaciones de los métodos presentados; una de esas razones fue la “*imparcialidad*” del proceso, es decir, que los niños y las niñas estén representados por igual. En respuesta al método aleatorio de Luis, algunos alumnos evaluaron el método como “*injusto*” debido al hecho de que algunos niños pudieron haber sido seleccionados y que no quisieran participar, mientras que otros que no fueron seleccionados probablemente quisieran participar: “*No, porque al hacerlo por sorteo puede tener unas complejidades debido a que tal vez no todos estén interesados*” (E278). Este razonamiento se basa en creencias emotivas y personales de lo que constituye una encuesta justa (Watson et al., 2003).

En este sentido, sobre el papel del contexto en el razonamiento del muestreo, Wroughton, McGowan, Weiss y Cope (2013) sugieren que los estudiantes que tienen opiniones firmes sobre un tema, evaluarán la validez de las conclusiones de un estudio basándose en si esa conclusión está o no de acuerdo con su opinión, en lugar de analizar la calidad del método de muestreo utilizado por medio de principios estadísticos.

5.5.4. Resultados Ítem 4

El Ítem 4 (ver Figura 5.4) corresponde a una pregunta de elaboración propia que pertenece al tercer grupo de preguntas del cuestionario, cuyo objetivo es que los estudiantes sean capaces de *sugerir un método de muestreo conocido el tamaño de la muestra, pero con el tamaño de la población desconocida*.

Ítem 4. Supongamos que quieres averiguar el porcentaje de niños que viene a tu escuela en diferentes medios de transporte: caminando, auto, micro, bicicleta, u otro. ¿Cómo tomarías una muestra de 50 niños de modo que los resultados sean representativos de toda la escuela?

Figura 5.4. Ítem 4 del cuestionario de evaluación

En el Ítem 4, se obtiene un alto porcentaje de respuestas incorrectas 59.3% (ver Tabla 5.28), por ejemplo, E1 indica: “*yo los distribuiría en todos los cursos*”, es decir, algunos estudiantes no mencionan método de selección de la muestra. Por otro lado, solo un 40.7% contesta correctamente, es decir, indican el método de muestreo, por ejemplo, E26 menciona: “*seleccionaría aleatoriamente en los distintos cursos*”; E55 indica que: “*enviaría una encuesta a cada curso y después sumaría todas las respuestas por categoría y calcularía porcentajes*”; E217 menciona: “*del total de niños de la escuela sacaría a 50 estudiantes al azar*”. Si se analizan las respuestas correctas, se observa que en algunos casos los estudiantes mencionan la utilización de algún instrumento como una encuesta o cuestionario para obtener la información, además, algunos de ellos mencionan que se debe elegir “*al azar*” a los participantes, es decir, consideran al *muestreo aleatorio* como un buen método para seleccionar una muestra representativa de la población.

Si ordenamos los resultados de este ítem por nivel, se observa que los estudiantes de segundo medio (22.2%) responden correctamente esta pregunta, seguidos por los alumnos de octavo año de primaria (9.8%), siendo el menor porcentaje de respuesta correcta de parte de los alumnos de cuarto medio (8.7%).

Tabla 5.28. Respuestas correctas e incorrectas Ítem 4

Curso	8° Básico	2° Medio	4° Medio	Total (%)
Respuestas	Frec. (%)	Frec. (%)	Frec. (%)	
Incorrectas	152 (12.3)	254 (20.5)	330 (26.6)	736 (59.3)
Correctas	122 (9.8)	275 (22.2)	108 (8.7)	505 (40.7)
Total	274	529	438	1241

Cuando se les solicita a los estudiantes sugerir un método de muestreo para seleccionar una muestra en un contexto cercano a su realidad (ver Tabla 5.29), un 14.7% de los alumnos no responde la pregunta, se observa además que el 38.9% de ellos menciona un método de muestreo, pero no justifica su elección, como E243 que menciona: “*tomaría dos alumnos por curso*”. Por otro lado, un 5.7% alude obtener una muestra de un grupo sesgado, por ejemplo, perteneciente a un nivel escolar u otra característica particular como sexo de los participantes o edad de los participantes, por ejemplo, E236 indica “*lo haría dependiendo de las edades y el transporte*”.

En síntesis, un 85.3% de los estudiantes señala algún método de muestreo, de ese porcentaje el 38.9% indica el método de selección de la muestra, pero no justifica su elección.

Tabla 5.29. Análisis respuestas Ítem 4

Curso	8° Básico	2° Medio	4° Medio	
Respuesta	Frec. (%)	Frec. (%)	Frec. (%)	Total (%)
No contesta	32 (2.6)	101 (8.1)	49 (4)	182 (14.7)
Menciona un método de muestreo, pero no justifica	112 (9)	124 (10)	247 (19.9)	483 (38.9)
Menciona obtener una muestra de un grupo sesgado (alguna característica en particular)	8 (0.7)	29 (2.3)	34 (2.7)	71 (5.7)
Usan algún método de muestreo y justifica	122 (9.8)	275 (22.2)	108 (8.7)	505 (40.7)
Total	274	529	438	1241

Se considera que la mayor dificultad de esta pregunta radica en no señalar el número total de estudiantes de la escuela (población), lo que les impide interpretar correctamente la situación planteada en el enunciado y determinar un método de muestreo apropiado para seleccionar una muestra dada con la población desconocida, además, como los últimos cambios curriculares en el nivel de segundo medio no incluyen los métodos de muestreo, puede ser un factor que influya en el alto porcentaje de respuestas incorrectas en cuarto medio.

5.5.5. Resultados Ítem 5

El Ítem 5 (ver Figura 5.5) ha sido adaptado y seleccionado de las situaciones problemas presentes en el texto escolar de primero medio (Del Valle, Muñoz y Santís, 2013, p. 247). El objetivo de este ítem es que los estudiantes puedan *determinar si la muestra señalada es representativa* o no.

Ítem 5. Analiza si las siguientes muestras son representativas de la población de estudio y completa la tabla con tus respuestas.

Muestra	¿Es representativa?	Explica por qué
1. Se seleccionan al azar 100 estudiantes, de un colegio de 300 alumnos, para calcular el promedio de notas de la población.		
2. Se les pregunta sobre la calidad del servicio del metro a las primeras 100 personas que toman el metro en la mañana.		
3. Se escogen los niños más altos y los más bajos de un curso para determinar el promedio de la estatura de los alumnos.		

Figura 5.5. Ítem 5 del cuestionario de evaluación

Además, para conocer las razones por las que los estudiantes indican si la muestra es o no representativa, se ha decidido agregar una segunda parte a la pregunta, en donde los alumnos deben justificar el porqué de su elección.

En la Tabla 5.30 se exponen los porcentajes de respuesta correcta para facilitar la presentación de los datos al lector.

Tabla 5.30. Porcentaje de respuestas correctas Ítem 5

Muestra	¿Es representativa?			
	Respuesta Correcta			
Curso	8° Básico	2° Medio	4° Medio	Total (%)
	Frec. (%)	Frec. (%)	Frec. (%)	
1.	118 (9.5)	193 (15.6)	165 (13.3)	476 (38.4)
2.	99 (8)	217 (17.4)	198 (16)	514 (41.4)
3.	114 (9.2)	213 (17.2)	200 (16.1)	527 (42.5)

Para la primera muestra, sólo el 38.4% (476 estudiantes) contesta correctamente que la muestra señalada es *representativa* (ver Tabla 5.30), en una situación que involucra un *muestreo aleatorio simple*, por ejemplo, E221 menciona que la muestra es representativa: “Sí, porque es al azar”; en la segunda, sólo el 41.4% (514 alumnos) contesta que *no es representativa* debido a que es un *método de muestreo restringido*, lo que implica que se eligen sujetos de manera sesgada y que no representan a la población total, por ejemplo, E224 menciona: “No es representativa. Porque les pregunta solo a los usuarios que suben en la mañana”; finalmente, en la tercera muestra, solo el 42.5% (527 estudiantes) contesta que *no es representativa*, debido a que se emplea un *método de muestreo autoseleccionado*, mostrando que la selección de la muestra da problemas al intentar ver si es o no representativa, ya que se consideran sujetos con características específicas y no se considera a “*todos*” los sujetos que conforman la población de estudio, por ejemplo, E220 alude que: “No. Porque no se toma en cuenta a los de estatura promedio”.

Se aprecia que los estudiantes reconocen mayoritariamente las muestras *no representativas* (muestras 2 y 3), pero en el primer caso, no identifican que es un *muestreo aleatorio simple*, llevándolos a pensar que la manera de seleccionar la muestra no es representativa de la población indicada en el enunciado, contestando así de forma incorrecta (61,6%, 765 estudiantes).

Dentro de las explicaciones para dar la respuesta a la primera muestra, los alumnos mencionan: porque se eligió a un tercio de la población; porque fue al azar; se debe elegir a la mitad. Para la segunda muestra aluden razones como: deberían ser más personas; porque se entrevista a un solo sector de usuarios; tendrían que entrevistarse en distintos horarios. Finalmente, para la tercera muestra indican explicaciones como: porque se debe considerar la diversidad de alturas; no se considera la variabilidad; se deben medir todos y sacar el promedio.

Por otro lado, se aprecia un alto porcentaje de respuestas incorrectas en cada una de las situaciones propuestas (ver Tabla 5.31). Se observa además que el mayor porcentaje de respuestas incorrectas corresponde a la primera muestra, que involucra una situación de muestreo aleatorio, por ejemplo, E227 señala que la primera muestra *no es representativa*: “*porque ni siquiera es la mitad de los estudiantes*”, en esta respuesta se puede apreciar que el estudiante no considera importante que la muestra ha sido escogida *al azar*, sino que hace referencia a la cantidad de estudiantes y considera que dicha cantidad seleccionada es insuficiente, pensando que el valor correcto es la mitad de estudiantes.

Para el caso de la segunda muestra, se aprecia que el porcentaje de respuestas incorrectas disminuye un poco al 58.6%, por ejemplo, en algunas de las respuestas incorrectas que entregan los estudiantes señalan las siguientes afirmaciones: “*Si. Porque al ser las primeras personas les darán una respuesta más concreta sobre lo que pasa en el metro en las mañanas*” (E260); “*Yo creo que sí. Porque se ve que es una gran cantidad y van a ser casi todas las mismas respuestas*” (E270). Al analizar ese tipo de respuestas, se observa que los estudiantes consideran que al encuestar a las personas que ocupan el metro en las mañanas, pueden generalizar las respuestas debido a que es una gran cantidad de encuestados o porque conocen el funcionamiento de servicio en la hora punta y no identifican que se ha seleccionado una muestra sesgada que no representa a toda la población.

Al revisar las respuestas entregadas a la tercera muestra, se observa que el 57.5% de los estudiantes contesta incorrectamente. En este caso, se encuentran respuestas como: “*No. Se deben seleccionar a todos ya que es un curso*” (E806); “*Si. Porque es un buen método para determinar el promedio de estatura*” (E794). En esta muestra se observa que los estudiantes no reconocen que es un *muestreo autoseleccionado*, en donde se consideran

solamente valores extremos y no valores aleatorios, además, en algunos casos los alumnos estiman que es correcto seleccionar a todos los estudiantes que conforma el curso, para poder obtener la estatura promedio.

Tabla 5.31. Porcentaje de respuestas incorrectas Ítem 5

Curso	¿Es representativa? Respuesta Incorrecta			Total (%)
	8° Básico Frec. (%)	2° Medio Frec. (%)	4° Medio Frec. (%)	
1.	156 (12.6)	336 (27.1)	273 (22)	765 (61.6)
2.	175 (14.1)	312 (25.1)	240 (13.3)	727 (58.6)
3.	160 (12.9)	316 (25.5)	238 (19.2)	714 (57.5)

En síntesis, al decidir si una *muestra es representativa*, los estudiantes distinguen mayoritariamente los casos en que la muestra dada *no es representativa*, pero en el primer ejemplo, no distinguen que se trata de un *muestreo aleatorio simple*, llevando al error de no creer que dicha muestra es representativa del conjunto total.

5.5.6. Resultados Ítem 6

El Ítem 6 (ver Figura 5.6) tiene por objetivo *identificar la muestra y la población*, ha sido adaptado de una actividad de un texto escolar (Muñoz, Rupin y Jimenez, 2013; p. 275), donde los estudiantes deben indicar en cada caso, la población y una posible muestra para cada uno de los enunciados presentados.

En la Tabla 4.11 (ver Capítulo 4) se caracterizan cada una de las respuestas que deben entregar los estudiantes para cada población y para cada muestra de este ítem.

Ítem 6. Determina en cada caso la población y una posible muestra de ella. Completa la tabla con tus respuestas.		
Datos	Población	Muestra
1. Una fábrica de yogur quiere investigar sobre la calidad de sus productos.		
2. Diego necesita saber el precio de un kilogramo de carne, para una comida familiar.		
3. Ximena estudia respecto del tamaño de las hormigas que habitan en un insectario.		
4. Daniel desea saber si una ciudad cuenta con suficiente lluvia, para realizar una plantación.		

Figura 5.6. Ítem 6 del cuestionario de evaluación

La mayoría de los estudiantes contesta incorrectamente, tanto para identificar la población (ver Tabla 5.32 y Tabla 5.33), como la muestra (ver Tabla 5.34 y Tabla 3.35), evidenciando que tienen algunos problemas al reconocer estos conceptos en contextos más alejados de la realidad escolar, además, de no identificar la diferencia entre la relación parte-todo de los contextos mencionados en el enunciado.

Cuando analizamos las respuestas del Ítem 6, para el primer dato, sólo un 31% responde correctamente para la población y un 27.7% para la muestra, por ejemplo, el alumno 417 señala que la población es “*todos los yogures fabricados*” y la muestra “*sólo los yogures de vainilla y frutilla*”, es decir, identifican la relación parte-todo en la información entregada.

Tabla 5.32. Análisis respuestas correctas Ítem 6, Población

Dato	8° Básico Frec. (%)	2° Medio Frec. (%)	4° Medio Frec. (%)	Total Frec. (%)
1.	104 (8.4)	139 (11.2)	142 (11.4)	385 (31)
2.	59 (4.8)	86 (6.9)	120 (9.7)	265 (21.4)
3.	107 (8.6)	133 (10.7)	216 (17.4)	456 (36.7)
4.	20 (1.6)	46 (3.7)	61 (4.9)	127 (10.2)

Al analizar las respuestas incorrectas para el primer dato, un 69% no acierta en identificar la población y un 72.3% no identifica la muestra, se observa que en algunos casos, los estudiantes responden indicando poblaciones y muestras con valores numéricos, por ejemplo, E21, indica que la población es de “*1000 yogures*” y que la muestra es de “*500 yogures*”, en otros casos, mencionan parcialmente la relación parte-todo, pero de forma incorrecta, por ejemplo, E28 menciona que la población es “*la fábrica de yogur*” y la muestra es “*sus productos*”.

Tabla 5.33. Análisis respuestas incorrectas Ítem 6, Población

Dato	8° Básico Frec. (%)	2° Medio Frec. (%)	4° Medio Frec. (%)	Total Frec. (%)
1.	170 (13.7)	390 (31.4)	296 (23.9)	856 (69)
2.	215 (17.3)	443 (35.7)	318 (25.6)	976 (78.6)
3.	167 (13.5)	396 (31.9)	222 (17.9)	785 (63.3)
4.	254 (20.5)	483 (38.9)	377 (30.4)	1114 (89.8)

Para el segundo dato, un 21.4% contesta correctamente la población y un porcentaje similar a la muestra (21.5%), por ejemplo, E430 responde que la población es “*el precio de todos los tipos de carne*” y que la muestra es “*los precios de algunas carnes*”. Por otro lado, en el caso de las respuestas incorrectas un 78.6% comete error al identificar la población y

un porcentaje similar no identifica la muestra (78.5%), por ejemplo, E32 menciona que la población es “*la comida familiar*” y que la muestra es “*la carne*”, es decir, en este caso también identifican incorrectamente la relación parte-todo, impidiéndole identificar correctamente los elementos que conforman la población y la muestra. En otro caso, por ejemplo, E389 menciona que la población es “*la carnicería*” y la muestra es “*la carne*”.

Tabla 5.34. Análisis respuestas correctas Ítem 6, Muestra

Dato	8° Básico Frec. (%)	2° Medio Frec. (%)	4° Medio Frec. (%)	Total Frec. (%)
1.	87 (7)	157 (12.6)	100 (8.1)	344 (27.7)
2.	64 (5.2)	132 (10.6)	71 (5.7)	267 (21.5)
3.	79 (6.4)	160 (12.9)	123 (9.9)	362 (29.2)
4.	17 (1.4)	35 (2.8)	40 (3.2)	92 (7.4)

Para el tercer dato, se observa que los porcentajes de respuestas correctas mejoran levemente, en este caso para la población el 36.7% responde correctamente y un 29.2% para la muestra. E440 menciona que la población es “*todas las hormigas del insectario*” y la muestra es “*una cantidad de hormigas del insectario*”. Por otro lado, cuando se analizan las respuestas incorrectas, un 63.3% no menciona la población adecuada (ver Tabla 5.33) y un 70.8% no mencionan la muestra apropiada (ver Tabla 5.35), además, se observa nuevamente que algunos estudiantes indican valores numéricos tanto para la población como la muestra, como es el caso de E394 que menciona que la población es de “*150 hormigas*” y que la muestra es de “*100 hormigas*”, sin mencionar la relación parte-todo.

Tabla 5.35. Análisis respuestas incorrectas Ítem 6, Muestra

Dato	8° Básico Frec. (%)	2° Medio Frec. (%)	4° Medio Frec. (%)	Total Frec. (%)
1.	187 (15.1)	372 (30)	338 (27.2)	897 (72.3)
2.	210 (16.9)	397 (32)	367 (29.6)	974 (78.5)
3.	195 (15.7)	369 (29.7)	315 (25.4)	879 (70.8)
4.	257 (20.7)	494 (39.8)	398 (32.1)	1149 (92.6)

Al analizar las respuestas correctas para el cuarto dato, se ve que el porcentaje de respuestas correctas disminuye drásticamente, a sólo un 10.2% para la población y un 7.4% para la muestra. E430 indica que la población es “*todas las ciudades del país*” y la muestra “*algunas ciudades por región*”. En el caso de las respuestas incorrectas (ver Tabla 5.33 y Tabla 5.35), el porcentaje de error aumenta considerablemente y nuevamente algunos estudiantes mencionan un valor numérico para la población y la muestra, por ejemplo, E404

indica que la población es “200 litros de agua” y la muestra es “20 litros de agua”. Por otro lado, el E405 señala que la población es “toda la lluvia en una ciudad” y que la muestra es “sólo el 20% de la lluvia en la ciudad”, en este caso, vemos que el E405 para el caso de la población menciona un “todo” y para la muestra “un porcentaje de ese total”, pero no proporciona exactamente la respuesta señalada en el libro de texto.

Si analizamos las respuestas por curso, se distingue que para el caso de identificar la población los estudiantes de cuarto medio obtienen mejores resultados (ver Tabla 5.32) lo que va disminuyendo en los otros cursos, lo mismo ocurre para la cuarta muestra, en cambio, para las tres primeras muestras los que obtienen más respuestas correctas son los estudiantes de segundo medio, seguidos por los de cuarto medio y los que obtienen resultados más bajos son los de octavo año de primaria (ver Tabla 5.34).

Cuando se les solicita identificar la población y la muestra, los estudiantes contestan erróneamente, creemos que esto se debe a que los contextos expuestos en esta actividad adaptada de un texto escolar (Muñoz, Rupin y Jimenez, 2013, p. 275), son desconocidos por los estudiantes y, además, impiden identificar fácilmente la relación parte-todo en los elementos que conforman la población o la muestra de cada situación. Surge la inquietud también de analizar la complejidad en la redacción de la pregunta, debido a que este ítem se adaptó de una actividad propuesta en un libro de texto.

5.5.7. Resultados Ítem 7

El Ítem 7 (ver Figura 5.7) se ha adaptado de la situación problema planteada en el texto de séptimo año de primaria (Merino, Muñoz, Pérez, y Rupin, 2015, p. 289). Este ítem tiene por objetivo *determinar la proporción en la población, dada una proporción en la muestra*.

<p>Ítem 7. Se tomaron al azar cuatro muestras de estudiantes de un colegio. A continuación, se presentan los resultados.</p>				
Alumnos	35	29	24	32
Alumnas	15	21	26	18
<p>¿Podrías estimar el porcentaje de alumnos en el colegio? Explica tu procedimiento.</p>				

Figura 5.7. Ítem 7 del cuestionario de evaluación

Un 45.6% de los estudiantes contesta correctamente este ítem (ver Tabla 5.36). Si

analizamos los resultados por curso y respuestas correctas, vemos que los estudiantes de segundo medio tienen mejores respuestas (19.8%), seguidos por los alumnos de cuarto medio (15.7%), en cambio, los estudiantes de octavo básico (10.1%) presentan menores resultados.

Tabla 5.36. Respuestas correctas e incorrectas Ítem 7

Curso	8° Básico	2° Medio	4° Medio	Total (%)
Respuestas	Frec. (%)	Frec. (%)	Frec. (%)	
Incorrectas	149 (12)	283 (22.8)	243 (19.6)	675 (54.4)
Correctas	125 (10.1)	246 (19.8)	195 (15.7)	566 (45.6)
Total	274	529	438	1241

En la Tabla 5.37 se presentan las respuestas proporcionadas por los estudiantes al Ítem 7. Se puede observar que un 43.5% de los alumnos no responde esta pregunta. Además, un 15.7% de los estudiantes calcula el porcentaje de alumnos en los cuatro grupos y luego calcula el promedio para estimar el porcentaje de alumnos varones del colegio. De igual modo, un 29.9% calcula el porcentaje directamente con el valor total de alumnos varones. Un 10.9% contesta parcialmente correcta esta pregunta, ya que calcula el porcentaje de alumnos varones por cada grupo, pero no realiza el cálculo del promedio. Las dos formas empleadas para realizar los cálculos y responder a este ítem se describen en la Tabla 4.13 del Capítulo 4.

Tabla 5.37. Respuestas de los estudiantes al Ítem 7.

Curso	8° Básico	2° Medio	4° Medio	Total (%)
Respuesta	Frec. (%)	Frec. (%)	Frec. (%)	
No contesta	110 (8.9)	234 (18.9)	196 (15.8)	540 (43.5)
Calcula el porcentaje de alumnos y el promedio	27 (2.2)	87 (7)	81 (6.5)	195 (15.7)
Calcula el porcentaje de hombres sin el promedio	39 (3.1)	49 (3.9)	47 (3.8)	135 (10.9)
Calcula el porcentaje con los totales	98 (7.9)	159 (12.8)	114 (9.2)	371 (29.9)
Total	274	529	438	1241

A continuación, presentamos algunos ejemplos de las respuestas proporcionadas por los estudiantes. En la Figura 5.8 se presenta el cálculo de los porcentajes realizado por E518 con los valores totales, es decir, calcula el valor total de alumnos de la muestra y luego calcula el porcentaje de alumnos varones por medio de regla de tres. En este caso, el estudiante sólo realiza el cálculo del porcentaje, pero no explica su procedimiento.

Alumnos	35	29	24	32	120
Alumnos	15	21	26	18	80
Total	50	50	50	50	200

$$\frac{100\% \cdot 200}{x} = \frac{120}{50}$$

$$x = 60\%$$

Figura 5.8. Respuesta de E518 al Ítem 7

En la Figura 5.9 se expone la respuesta de E509, en donde éste calcula el porcentaje de varones en cada grupo de muestras y posteriormente calcula el porcentaje de total de alumnos varones por medio del promedio. Además del cálculo respectivo, este estudiante proporciona una explicación de cómo realizó el procedimiento.

Grupo ① 50 → 100 35 → x X = 70% <i>son</i>	Grupo ② 50 → 100 29 → x X = 58%	Grupo ③ 50 → 100 24 → x X = 48%	Grupo ④ 50 → 100 32 → x X = 64%
---	--	--	--

$$\bar{X} = \frac{70 + 58 + 48 + 64}{4} = \frac{240}{4} = 60\%$$

pero se que el porcentaje de cada grupo y luego se que el promedio de aquellos porcentajes para obtener un porcentaje total.

Figura 5.9. Respuesta de E509 al Ítem 7

En la Figura 5.10 se presenta el procedimiento realizado por E543, en este caso, se observa que el alumno realiza el cálculo de porcentajes de varones por cada grupo, pero no calcula el promedio de éstos, entregando una respuesta incompleta, además, se observa en la Tabla 5.34 que el 10.9% de los estudiantes entrega este tipo de respuestas.

$$100 \cdot 35 = \square \div 50$$

$$100 \cdot 24 = \square \div 50$$

$$100 \cdot 29 = \square \div 50$$

$$100 \cdot 32 = \square \div 50$$

Alumnos:

$$100 \cdot 35 = 3500 \div 50 = 70$$

$$100 \cdot 24 = 2400 \div 50 = 48$$

$$100 \cdot 29 = 2900 \div 50 = 58$$

$$100 \cdot 32 = 3200 \div 50 = 64$$

Figura 5.10. Respuesta de E543 al Ítem 7

En síntesis, un 45.6% de los estudiantes logra *determinar la proporción en la población, dada una proporción en la muestra*, además, un 10.9% realiza cálculos parciales, entregando una respuesta incompleta. Se puede considerar que la extensión del cuestionario cómo también el lenguaje utilizado en el enunciado del Ítem 7 pueden haber influido en que un alto porcentaje de estudiantes no diera respuesta a este ítem (43.5%).

5.5.8. Resultados Ítem 8

El Ítem 8 (ver Figura 5.11) ha sido adaptado de Gómez (2014, p.202) en donde el problema está enmarcado en el uso del significado frecuencial de la probabilidad, se ha considerado que puede aportar información relevante para nuestro estudio, pues se trabaja con la muestra de una población finita. Este ítem también pertenece al bloque 6, cuyo objetivo apunta a evaluar lo siguiente: *dada una proporción en la muestra, determinar la proporción en la población*.

Ítem 8. En un estanque hay peces, pero su dueño no sabe cuántos. Toma 200 peces y les pone una marca; los devuelve al estanque, donde se mezclan con el resto. Al día siguiente, el dueño toma 250 peces del estanque, y encuentra que 25 de ellos están marcados y el resto no.

1. ¿Qué proporción aproximada de peces del estanque están marcados?
2. ¿Cuál es el número aproximado de peces en el estanque?
3. Si el dueño saca ahora 100 peces más del estanque, ¿cuántos aproximadamente estarán marcados?

Figura 5.11. Ítem 8 del cuestionario de evaluación

Como se observa en la Tabla 5.38 un alto porcentaje de los estudiantes (52.1%) da la respuesta correcta a la pregunta 1, es decir, realiza correctamente el cálculo de la proporción de peces marcados en el estanque. Como la investigación de Gómez (2014) no incluye esta pregunta no podemos realizar una comparación con dichos resultados.

Tabla 5.38. Respuestas correctas e incorrectas Ítem 8, pregunta 1

Curso	8° Básico	2° Medio	4° Medio	Total (%)
Respuestas	Frec. (%)	Frec. (%)	Frec. (%)	
Incorrectas	147 (11.8)	234 (18.9)	214 (17.2)	595 (47.9)
Correctas	127 (10.2)	295 (23.8)	224 (18.1)	646 (52.1)
Total	274	529	438	1241

En la Figura 5.12 se presenta la respuesta entregada a la pregunta 1 por E495, quien identifica correctamente los valores para calcular la proporción de peces marcados en el estanque y emplea la regla de tres para calcular el valor pedido.

$250 \rightarrow 100\%$
 $25 \rightarrow X\%$
 $X = 10\%$

Resp. = Aprox. el 10% de los peces del estanque están marcados

Figura 5.12. Respuesta E495 pregunta 1 del Ítem 8

En las Figuras 5.13 y 5.14, se presentan algunas de las respuestas incorrectas entregadas en la pregunta 1. En el primer caso, E295 asume que los peces marcados son 200 y calcula el porcentaje que representa dicho valor para la muestra de 250 peces, es decir, no identifica correctamente los valores que deben emplearse para calcular la proporción pedida.

$\frac{200}{250} = 80\%$

$\frac{200 \cdot 100}{250} = 80\%$

dos peces que están marcados son 80% de la población

Figura 5.13. Respuesta E295 pregunta 1 del Ítem 8

Por otro lado, los estudiantes efectúan solamente un procedimiento aditivo, realizando una resta de los valores entregados en el enunciado, deduciendo incorrectamente el valor pedido. Por ejemplo, E103 efectúa la resta “ $250-25=225$ ” para señalar que “*hay 225 peces marcados*” (ver Figura 5.14). Cabe mencionar que el 24.9% de los estudiantes no responde esta pregunta, es decir, sólo el 23% responde un valor incorrecto.

1. ¿Qué proporción aproximada de peces del estanque están marcados?
 225 peces están marcados

Figura 5.14. Respuesta E103 pregunta 1 del Ítem 8

En la respuesta de E486 se observa que intenta usar un pensamiento proporcional, pero no identifica correctamente los valores correctos para realizar la proporción pedida:

$$\frac{25}{250} = \frac{1}{10} = 0.1 = 10\%.$$

Total = $200 + 225 = 425$
 Marcados = $200 + 25 = 225$

$\frac{225}{425} \rightarrow \frac{45}{85} \rightarrow \frac{9}{17}$

Figura 5.15. Respuesta E486 pregunta 1 del Ítem 8

En cuanto a las respuestas de la segunda pregunta del Ítem 8 (ver Tabla 5.39), un 42.3% contesta correctamente acertando en la cantidad de peces en el estanque. Estos resultados son muchos mejores de los reportados por Gómez (2014) con futuros profesores. Se puede considerar que los estudiantes de secundaria chilenos aplican correctamente el razonamiento

proporcional, ya que un buen porcentaje de ellos logra determinar la proporción de la población dada la proporción de la muestra.

Tabla 5.39. Respuestas correctas e incorrectas Ítem 8, pregunta 2

Curso	8° Básico	2° Medio	4° Medio	Total (%)
Respuestas	Frec. (%)	Frec. (%)	Frec. (%)	
Incorrectas	187 (15.1)	273 (22)	256 (20.6)	716 (57.7)
Correctas	87 (7)	256 (20.6)	182 (14.7)	525 (42.3)
Total	274	529	438	1241

En la Figura 5.16 se presenta la respuesta entregada por E558 a la pregunta 2 del Ítem 8, en donde determina la proporción de peces del estanque mediante el uso de la regla de tres y utilizando la proporción de la muestra calculada en la pregunta anterior. Como mencionamos anteriormente, un 42.3% de los estudiantes realiza este procedimiento al responder correctamente esta pregunta.

$$x \rightarrow 100$$

$$200 \rightarrow 20 = \boxed{2000}$$

Figura 5.16. Respuesta E558 pregunta 2 del Ítem 8

Al analizar las respuestas incorrectas, se destacan las siguientes afirmaciones entregadas por los estudiantes, por ejemplo, E103 señala que “*hay 450 peces*”, en este tipo de afirmación se observa que realizan una suma de los valores de las muestras mencionadas en el enunciado ($200 + 250 = 450$) para señalar lo que consideran el valor total de peces en el estanque. Por otro lado, algunos estudiantes mencionan que “*aproximadamente hay 425 peces*” (E529), es decir, suman los valores de las muestras dadas y restan el número de peces marcados ($200 + 250 - 25 = 425$), considerando así que la totalidad de peces en el estanque es de 425. Por otro lado, hay estudiantes que señalan valores sin realizar ningún cálculo, por lo que no es posible distinguir el procedimiento llevado a cabo para llegar a dicha solución, por ejemplo, en la Figura 5.17, E527 da este tipo de respuesta.

300 peces aproximadamente

Figura 5.17. Respuesta E527 pregunta 2 del Ítem 8

En la tercera pregunta del Ítem 8 (ver Tabla 5.40), un 40.2% de los alumnos contesta correctamente. Resultados más bajos que los reportados por Gómez (2014).

Tabla 5.40. Respuestas correctas e incorrectas Ítem 8, pregunta 3

Curso	8° Básico	2° Medio	4° Medio	Total (%)
Respuestas	Frec. (%)	Frec. (%)	Frec. (%)	
Incorrectas	192 (15.5)	295 (23.8)	255 (20.5)	742 (59.8)
Correctas	82 (6.6)	234 (18.9)	183 (14.7)	499 (40.2)
Total	274	529	438	1241

Como este estudio se realiza con estudiantes de secundaria, el procedimiento señalado en la Figura 5.18 es el mayoritariamente empleado por los participantes de nuestro estudio, donde calculan la cantidad de peces marcados al tomar la muestra de 100 peces, empleando la regla de tres.

2. ¿Cuál es el número aproximado de peces en el estanque?

$$\begin{array}{l}
 X \rightarrow 100\% \\
 200 \rightarrow 10 \\
 20000 : 10 = 2000
 \end{array}$$

Figura 5.18. Respuesta E8 pregunta 2 del Ítem 8

En esta pregunta en los resultados presentados por Gómez (2014) con futuros profesores, algunos de los participantes entregan la proporción de la muestra calculada empleando la información de la primera muestra y la información de esta nueva muestra, indicando que son 35 peces marcados. En nuestro caso sólo E295 señala este resultado, en la Figura 5.19 se presenta la respuesta entregada por dicho estudiante, quien indica que hay 35 peces marcados.

$$\begin{array}{l}
 \frac{250}{25} | \frac{100}{10\%} \quad \frac{25 \cdot 100}{250} : 250 = 10\% \\
 \frac{25}{10} | \frac{100}{10\%} \quad \frac{25 \cdot 100}{10} : 10 = 250 \\
 \frac{100 \cdot 10}{1000} : 100 = 10\% \\
 \text{Respuesta: Aproximadamente 35 peces estarán marcados}
 \end{array}$$

Figura 5.19. Respuesta E 295 a la pregunta 3 del Ítem 8

En el caso de la pregunta 3, hay un alto porcentaje de las respuestas incorrectas (59.8%), de ese porcentaje un 28.4% de los alumnos no responde y el 31.4% entrega una respuesta incorrecta, por ejemplo, algunos estudiantes entregan una respuesta numérica, pero no indican el procedimiento por el cual llegan a dicha solución, E103 menciona que “50 están marcados”; E527 señala que “6 están marcados”; E529 indica que “habrá aproximadamente 47 marcados”. Por otro lado, el estudiante 486 realiza un procedimiento proporcional, pero incorrecto, ya que la proporción obtenida en la primera pregunta de este ítem lo lleva a seguir con la proporción errada, como se ve en la Figura 5.20.

$$\frac{100 \cdot 9}{17} = 52 \text{ aprox}$$

Figura 5.20. Respuesta E486 a la pregunta 3 del Ítem 8

5.5.9. Resultados Ítem 9

Este ítem (ver Figura 5.21) ha sido adaptado de la investigación de Shaughnessy, Canada y Ciancetta (2003), sobre la comprensión de los estudiantes de secundaria sobre la variabilidad de la muestra y pretende medir la predicción de resultados de muestras repetidas de una mezcla de caramelos, ítem que se considera pertinente para el cuestionario. Se ha considerado adaptar esta pregunta a nuestro objetivo del bloque 7, que es: *dada la proporción en una población, determinar la proporción en varias muestras.*

Ítem 9. Piensa que tienes un recipiente con 100 dulces; 60 son de frutilla, y 40 de limón. Los dulces están todos mezclados en el recipiente.

1. Sacas un puñado de 10 dulces y cuentas el número de los dulces de frutilla. ¿Cuántos serán de frutilla?
2. Supongamos que seis de tus compañeros repiten experimento, cada uno de ellos saca 10 dulces. (Después de cada extracción, los dulces se vuelven a colocar y se mezcla nuevamente).

Escribe en cada línea un número probable de dulces de frutilla que obtiene tu compañero.

Niño	Número de dulces de frutilla	Niño	Número de dulces de frutilla
Ana		Pedro	
Camila		María	
José		Juan	

Figura 5.21. Ítem 9 del cuestionario de evaluación

En las respuestas de la primera pregunta del Ítem 9 (ver Tabla 5.41), un 51.9% de los estudiantes contesta correctamente, realizando el cálculo de la proporción directa.

Tabla 5.41. Respuestas correctas e incorrectas Ítem 9, pregunta 1

Curso	8° Básico	2° Medio	4° Medio	
Respuestas	Frec. (%)	Frec. (%)	Frec. (%)	Total (%)
Incorrectas	156 (12.6)	264 (21.3)	177 (14.3)	597 (48.1)
Correctas	118 (9.5)	265 (21.4)	261 (21)	644 (51.9)
Total	274	529	438	1241

Por ejemplo, E485 establece la proporción directa y determina que “*serán 6 dulces de frutilla*”, respondiendo así correctamente a la pregunta planteada (ver Figura 5.22).

Handwritten student work for E485 showing a proportion calculation: $100 \rightarrow 60$, $10 \rightarrow x$, and the conclusion "Resp: serán 6 dulces de frutilla".

Figura 5.22. Respuesta estudiante 485 pregunta 1 Ítem 9

En el caso de E485 (ver figura 5.23) hace el cálculo de la proporción empleando los

valores de la muestra y la población, también proporcionando una respuesta correcta a la pregunta.

$$\frac{60}{100} = \frac{6}{10} = \frac{3}{5} \quad * \frac{3}{5} \cdot 10 = \frac{30}{5} = 6 \quad \text{Serán 6 dulces de frutilla.}$$

Figura 5.23. Respuesta E485 pregunta 1 Ítem 9

Además, se observan otros casos en que los estudiantes relacionan la respuesta con las probabilidades y el cálculo de frecuencias absolutas porcentuales, por ejemplo, E529 menciona “*probablemente 6*” (ver Figura 5.24).

60% F
40% L
Probablemente 6

Figura 5.24. Respuesta E485 pregunta 1 Ítem 9

El E487 menciona el “*porcentaje de posibilidades*” y da la respuesta correcta a la primera pregunta del Ítem 9 (ver Figura 5.25).

60 F
40 L → 100 $\frac{60}{100} = 60\%$ Existe un 60% de posibilidades de que sea de frutilla, por lo tanto, en una muestra de 10 deben haber 6 dulces de frutilla.

Figura 5.25. Respuesta E485 pregunta 1 Ítem 9

Del porcentaje de respuestas incorrectas, el 29.8% de los alumnos no responde esta pregunta. El otro 18.3% realiza un cálculo erróneo, como por ejemplo E125 que establece incorrectamente los valores en la regla de tres y efectúa un cálculo incorrecto (ver Figura 5.26).

100 dulces	60 frutilla	40 Limon
100	10	

→ 60
x=96

Figura 5.26. Respuesta estudiante E125 pregunta 1 Ítem 9

En el caso de E496 menciona que “*no podría determinarlo*”, pero a su vez da una respuesta con una variabilidad excesiva “*podrían ser los 10 de frutilla o los 10 de limón*” (ver Figura 5.27), o como E124 dice que “*serán todos*”.

NO podía determinarlo, podían ser los 10 de frutilla o los 10 de limón.

Figura 5.27. Respuesta E496 pregunta 1 Ítem 9

Estos resultados no se pueden comparar con los reportados por Shaughnessy, Canada y Ciancetta (2003) debido a que los investigadores no plantean explícitamente esta pregunta, sino que analizan la variabilidad correspondiente a la pregunta 2 de este ítem.

En cambio, los resultados reportados por Shaughnessy, Ciancetta y Canada (2004) son muchos mejores, quienes obtuvieron un 82.4% de respuestas correctas en esta pregunta y muy pocos alumnos (2.9%) indican un rango de valores para la cantidad de caramelos pedidos. Estos investigadores mencionan que los estudiantes se centran directamente en el valor esperado, como suele ser el caso cuando solo se les pregunta sobre una repetición del experimento.

En la Tabla 5.42 se presentan los resultados de la segunda pregunta del Ítem 9, en donde se han resumido las respuestas de los estudiantes de acuerdo a si emplean correctamente la variabilidad en la extracción de muestras sucesivas.

Tabla 5.42. Respuestas Ítem 9, pregunta 2

Curso	8° Básico	2° Medio	4° Medio	Total (%)
Respuestas	Frec. (%)	Frec. (%)	Frec. (%)	
No contesta	105 (8.5)	233 (18.8)	146 (11.8)	484 (39)
No indican variación	27 (2.2)	81 (6.5)	79 (6.4)	187 (15,1)
Indican variación razonable	65 (5.2)	118 (9.5)	117 (9.4)	300 (24,2)
Indican variación excesiva	77 (6.2)	97 (7.8)	96 (7.7)	270 (21,8)
Total	274	529	438	1241

Un 24.2% de los estudiantes presentan muestras que indican una *variación razonable*, es decir, las muestras toman distintos valores que están dentro de un rango de valores (ver Figura 5.28). Resultados más bajos que los reportados por Shaughnessy, Canada y Ciancetta (2003), quienes mencionan que un 70% de los estudiantes predicen muestras con un rango o margen “razonable”.

Niño	Número de dulces de frutilla	Niño	Número de dulces de frutilla
Ana	5	Pedro	8
Camila	5	María	5
José	7	Juan	6

Figura 5.28. Respuesta E487 pregunta 2 Ítem 9

Además, un 15.1% de los estudiantes ofrece muestras sin *indicar variación*, es decir, anotan todas las muestras iguales a 6 (ver figura 5.29), es decir, sus respuestas se basan en el valor esperado.

Niño	Número de dulces de frutilla	Niño	Número de dulces de frutilla
Ana	6	Pedro	6
Camila	6	María	6
José	6	Juan	6

Figura 5.29. Respuesta E536 pregunta 2 Ítem 9

Por otro lado, un 21.8% de los alumnos indican muestras con una *variabilidad excesiva*, anotando valores muy extremos (ver Figura 5.30). Finalmente, un 39% de los estudiantes no responde esta pregunta.

Niño	Número de dulces de frutilla	Niño	Número de dulces de frutilla
Ana	25	Pedro	37
Camila	30	María	10
José	10	Juan	15

Figura 5.30. Respuesta E103 pregunta 2 Ítem 9

Como señalan Shaughnessy, Canada y Ciancetta (2003), la enseñanza de la probabilidad teórica para eventos de un solo resultado podría interferir con que los estudiantes apliquen correctamente la *variabilidad* en la obtención de los resultados de las extracciones sucesivas de muestras.

5.5.10. Resultados Ítem 10

El Ítem 10 (ver Figura 5.31) ha sido adaptado de García-Ríos (2013, p.347), para explorar el razonamiento inferencial informal empleado por estudiantes de bachillerato. Este ítem también pertenece al bloque 7, cuyo objetivo es: *dada la proporción en una población, determinar la proporción en varias muestras.*

Ítem 10. Una industria farmacéutica fabrica un producto llamado “Elección del género”, el cual, según la publicidad permite a las parejas incrementar sus posibilidades de tener una niña. Si 100 parejas usaron el medicamento y 52 tuvieron una niña, ¿Piensas que el medicamento es efectivo? ¿Por qué?

Figura 5.31. Ítem 10 del cuestionario de evaluación

Como se observa en la Tabla 5.43 un 52.9% responde incorrectamente, dentro de este porcentaje, el 31% (385 alumnos) no contestan el ítem. Por otra parte, el 47.1% de los estudiantes responden correctamente indicando que el medicamento es efectivo, para ello

complementan su respuesta con los argumentos clasificados en la Tabla 5.41.

Si se analizan los resultados por curso, se observa en la Tabla 5.41 que los estudiantes de segundo medio son los que mejor responden esta pregunta (19%), seguidos por los estudiantes de cuarto medio (17.7%), y los alumnos de octavo año de primaria presentan un menor porcentaje de respuesta correcta (10.4%).

Tabla 5.43. Respuestas correctas e incorrectas Ítem 10, pregunta 1

Curso	8° Básico	2° Medio	4° Medio	Total (%)
Respuestas	Frec. (%)	Frec. (%)	Frec. (%)	
Incorrectas	145 (11.7)	293 (23.6)	218 (17.6)	656 (52.9)
Correctas	129 (10.4)	236 (19)	220 (17.7)	585 (47.1)
Total	274	529	438	1241

Al analizar algunas de las respuestas correctas, por ejemplo, E102 señala: “*Sí, pienso que el medicamento es efectivo, porque al menos a más de la mitad de las parejas les resultó*” o E529 menciona: “*Sí, ya que su efectividad supera el 50%*”.

En el caso de las respuestas incorrectas, se indican los siguientes argumentos: “*No es efectivo, ya que antes de tomar las pastillas, las parejas tenían una probabilidad de 50% de que su hija sea niña*” (E536); “*No, porque el porcentaje para que sea efectivo debería ser de un 99.9%*” (E527).

Las categorías de análisis propuestas por García-Ríos (2013) se basan en los componentes de la Inferencia Estadística Informal propuestos en los marcos de Makar y Rubin (2009) y Zieffler, Garfield, delMas, y Reading (2008). García-Ríos (2013) establece que en la categoría *conclusión más allá de los datos* los estudiantes al momento de “hacer juicios, afirmaciones, o predicciones acerca de las poblaciones (generalizar más allá de la muestra) con base en las muestras, no emplean procedimientos y métodos estadísticos formales” (p.346).

Así mismo, García-Ríos (2013) señala que:

Las respuestas de los alumnos que *van más allá de los datos* mencionan la efectividad del producto en la población y no de la muestra en particular; estas respuestas se han clasificado en cuatro categorías: 1) no funciona, 2) juzgan baja efectividad, 3) asignan porcentaje de efectividad y 4) juzgan que puede ser efectivo (p.348). Por otro lado, las respuestas que *no van más allá de los datos*, se dividieron en dos categorías. Cuando concluyen que el producto pudo influir de alguna manera en las personas de la muestra se agruparon en una categoría, y en la otra cuando concluyen que sí funciona el producto (p.349).

En la Tabla 5.44 se han clasificado las respuestas de los estudiantes al ítem 10 siguiendo estas categorías de análisis propuestas por García-Ríos (2013).

Tabla 5.44. Argumentos Ítem 10, pregunta 2

Curso	Clasificación respuestas	8° Básico Frec. (%)	2° Medio Frec. (%)	4° Medio Frec. (%)	Total (%)
No contesta	No responde	94 (7.6)	193 (15.6)	92 (7.4)	379 (30.5)
Va más allá de los datos	No funciona	21 (1.7)	31 (2.5)	48 (3.9)	100 (8.1)
	Juzgan baja efectividad	30 (2.4)	62 (5)	65 (5.2)	157 (12.7)
	Asignan porcentaje de efectividad	103 (8.2)	209 (16.8)	208 (16.8)	520 (41.9)
	Juzgan que puede ser efectivo	11 (0.9)	20 (1.6)	18 (1.5)	49 (3.9)
No va más allá de los datos	Pudo influir en las personas	0 (0)	0 (0)	1 (0.1)	1 (0.1)
	Funcionó	15 (1.2)	14 (1.1)	6 (0.5)	35 (2.8)
	Total	274	529	438	1241

En la Tabla 5.44 se observa que la mayoría de los estudiantes basan sus argumentos en asignar un porcentaje de efectividad del producto (41.9%) y en segundo lugar sus argumentos aluden baja efectividad del producto (12.7). Además, un 30.5% de los estudiantes no contesta este ítem.

Si analizamos los argumentos proporcionados por los estudiantes, para las respuestas que *van más allá de los datos*, en la categoría *no funciona*, E496 alude: “*No, porque no todos cumplieron con las expectativas. El porcentaje de triunfo que tuvo fue casi el mismo que la probabilidad de que hubiera pasado sin el medicamento*”; en la categoría *juzgan baja efectividad*, E627 indica que: “*No, ya que si fuera efectivo las 100 parejas hubieran tenido niñas*”; en la categoría *asignan porcentaje de efectividad*, por ejemplo, E503 menciona: “*En teoría el medicamento sí es efectivo, porque el 50+1 hace la diferencia. Significa que más de la mitad obtuvo el resultado que esperaba*”; en la categoría *juzgan que puede ser efectivo*, E400 indica: “*Sí, pues gran parte obtuvieron resultados correctos*”.

En el caso de las respuestas que *no van más allá de los datos*, en la categoría *pudo influir en las personas*, E322 indica “*Tal vez, aunque la mayoría de los recién nacidos son mujeres sin siquiera haber consumido el producto*”; en la categoría *funcionó*, E26 menciona “*Si, porque es una droga*”.

Considerando los resultados obtenidos en este ítem, también se percibe que las respuestas proporcionadas por los alumnos al Ítem 10, como menciona García-Ríos (2013),

están fundadas “más en sus creencias y en sus conocimientos personales acerca del contexto y no en los datos del problema” (p.343).

5.5.11. Resultados Ítem 11

El Ítem 11 (ver Figura 5.31) corresponde a una pregunta de elaboración propia, en el que se les solicita a los estudiantes señalar las diferentes muestras con y sin reemplazo que pueden determinar de una población finita de cuatro elementos. Esta pregunta pertenece al bloque 8, cuyo objetivo es *obtener todas las muestras posibles de una población finita*.

Ítem 11. Tenemos 4 libros numerados 1, 2, 3 y 4. Escribe todas las muestras posibles de dos libros *con* y *sin* reemplazamiento.

Figura 5.31. Ítem 11 del cuestionario de evaluación

Se considera como respuesta correcta que los estudiantes anoten los listados de libros con y sin reemplazamiento.

Al analizar los resultados por respuesta correcta y curso (ver Tabla 5.45), se percibe que una mayor cantidad de alumnos de segundo medio contesta correctamente (11.4%), seguidos por los alumnos de cuarto medio (7.7%) y los estudiantes de octavo año de primaria (3.3%).

Tabla 5.45. Respuestas correctas e incorrectas Ítem 11

Curso	8° Básico	2° Medio	4° Medio	Total (%)
Respuestas	Frec. (%)	Frec. (%)	Frec. (%)	
Incorrectas	233 (18.8)	387 (31.2)	342 (27.6)	962 (77.5)
Correctas	41 (3.3)	142 (11.4)	96 (7.7)	279 (22.5)
Total	274	529	438	1241

Se observa que en el Ítem 11 (ver Figura 5.31), sólo el 22.4% contesta correctamente esta pregunta anotando ambas muestras pedidas; un 25.8% la contesta parcialmente correcta, escribiendo solo una de las muestras solicitadas, es decir, anotando el listado con o sin reposición (ver Figura 5.32 y 5.33); un 14% escribe el listado con o sin remplazo de forma incompleta, con algunos elementos faltantes; y un 37.8% no responde (ver Tabla 5.46).

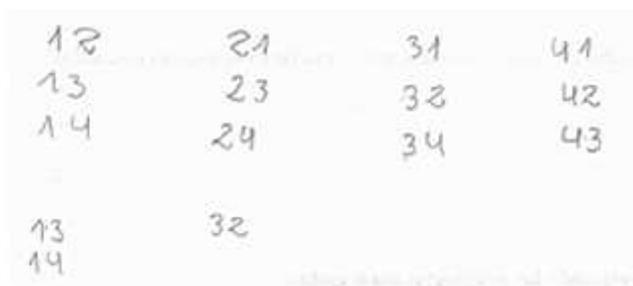


Figura 5.32. Respuesta E233 Ítem 11

En la Figura 5.32 se presenta la respuesta de E233 quien anota el listado con reemplazamiento completo, pero sólo señala tres elementos de la muestra sin reemplazamiento.

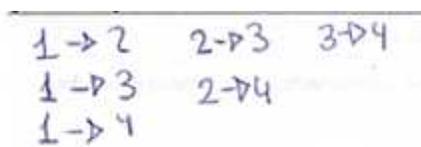


Figura 5.33. Respuesta E6 Ítem 11

En la Figura 5.33 se presenta la respuesta de E6 quien solamente anota el listado sin reemplazamiento.

Tabla 5.46. Análisis respuestas Ítem 11

Curso	8° Básico	2° Medio	4° Medio	Total (%)
Respuesta	Frec. (%)	Frec. (%)	Frec. (%)	
No contesta	104 (8.4)	221 (17.8)	144 (11.6)	469 (37.8)
Anota listado con reemplazo	70 (5.6)	86 (6.9)	116 (9.4)	272 (21.9)
Anota listado con reemplazo incompleto	24 (1.9)	20 (1.6)	29 (2.3)	73 (5.9)
Anota listado sin reemplazo	10 (0.8)	22 (1.8)	16 (1.3)	48 (3.9)
Anota listado sin reemplazo incompleto	25 (2)	38 (3.1)	37 (3)	100 (8.1)
Anota ambos listados	41(3.3)	142 (11.4)	96 (7.7)	279 (22.4)
Total	274	529	438	1241 (100)

Si analizamos las respuestas de los estudiantes, se ve que la información proporcionada por E509 (ver Figura 5.34), quien realiza el cálculo de la combinación y la permutación, además de anotar el listado de todos los elementos de cada muestra respectivamente.

Sin Reemplazo

$$C_2^4 = \frac{4!}{2! \cdot 2!} = \frac{2! \cdot 3!}{2! \cdot 2!} = 6$$

Con Reemplazo

$$P_2^4 = 4 \cdot 3 = 12$$

$\left. \begin{array}{l} 1,2 \\ 2,1 \\ 2,2 \\ 2,3 \\ 2,4 \\ 3,1 \\ 3,2 \\ 3,3 \\ 3,4 \\ 4,1 \\ 4,2 \\ 4,3 \\ 4,4 \end{array} \right\}$

Figura 5.34. Respuesta E509 Ítem 11

En la Figura 5.35 vemos otra forma de respuesta correcta, en que el alumno anota ambos listados de las muestras con y sin reemplazamiento.

con reemplazamiento

1,2 / 1,3 / 1,4
2,3 / 2,1 / 2,4
3,1 / 3,2 / 3,4
4,1 / 4,2 / 4,3

sin reemplazamiento

1,2 / 1,3 / 1,4
2,3 / 2,4
3,4

Figura 5.35. Respuesta E514 Ítem 11

Por otro lado, se encontraron algunas respuestas incorrectas donde los estudiantes no anotan el listado de elementos pedidos (ver Figura 5.36), sino que realizan algunos cálculos combinatorios, donde emplean el principio multiplicativo (para la muestra con reemplazo) y emplean la permutación (para la muestra sin reemplazo), evidenciando problemas al diferenciar que en el caso emplear una combinación o una permutación.

4 · 4 = 16 con reposición

4 · 3 = 12 sin reposición

Figura 5.36. Respuesta E524 Ítem 11

En la Figura 5.37 se observan otros cálculos en donde emplean incorrectamente el cálculo de permutaciones (E541).

Handwritten calculations showing two methods for finding the number of permutations of 4 elements:

- Sin ~~reemplazamiento~~ reemplazamiento:** $4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$
- Con reemplazamiento:** $4 \cdot 4 \cdot 4 \cdot 4 = 16 \cdot 16 = 256$

Figura 5.37. Respuesta E541 Ítem 11

Como esta pregunta involucra elementos de combinatoria y permutaciones, observamos que algunos estudiantes de cuarto medio contestan incorrectamente porque confunden la fórmula que deben emplear para realizar los cálculos, y no identifican que dicha fórmula sólo le da la cantidad de elementos que debe contener cada muestra.

5.6. Diferencias entre niveles educativos

Como en este estudio se analizan los resultados de tres niveles educativos distintos es necesario realizar una comparación de los resultados obtenidos en cada ítem. Para esto se realiza un ANOVA de un factor y una prueba post hoc.

De acuerdo a IMB (s.f): el procedimiento **ANOVA de un factor** genera un análisis de varianza de un factor para una variable dependiente cuantitativa respecto a una única variable de factor (la variable independiente). El análisis de varianza se utiliza para contrastar la hipótesis de que varias medias son iguales. Esta técnica es una extensión de la prueba t para dos muestras.

En la Tabla 5.47 se presenta el ANOVA de un factor por **nivel educativo** (curso) para todos los ítems del cuestionario, se observa que no existen diferencias significativas entre los cursos (octavo año de primaria, segundo medio y cuarto medio) para los Ítems: Ítem 2: tercera pregunta (apartado 2); Ítem 3: Método de Elena, Método de Pedro, Método de Luis y Método de Sara; Ítem 5: preguntas 1, 2 y 3; Ítem 6: muestra 3 y 4; Ítem 7 e Ítem 10. Para todos los otros ítems el p-valor indica que existen diferencias significativas entre grupos.

Tabla 5.47. ANOVA de un factor por nivel educativo (curso)

Preguntas	F	P-valor	Preguntas	F	P-valor
Ítem_1	4,876	0,008	Ítem_6_p_2	8,963	0,000
Ítem_2_preg_a	14,268	0,000	Ítem_6_p_3	32,024	0,000
Ítem_2_preg_b	9,757	0,000	Ítem_6_p_4	5,249	0,005
Ítem_2_preg_c	3,720	0,025	Ítem_6_m_1	4,249	0,014
Ítem_2_preg_c_2	1,234	0,291	Ítem_6_m_2	5,817	0,003
Ítem_3_andrea	4,982	0,007	Ítem_6_m_3	,281	0,755
Ítem_3_elena	,793	0,453	Ítem_6_m_4	1,479	0,228
Ítem_3_pedro	1,141	0,320	Ítem_7	,189	0,827
Ítem_3_luis	2,903	0,055	Ítem_8_preg_1	3,329	0,036
Ítem_3_sara	2,639	0,072	Ítem_8_preg_2	10,468	0,000
Ítem_3_preg_b	7,310	0,001	Ítem_8_preg_3	8,118	0,000
Ítem_4	40,542	0,000	Ítem_9_preg_a	9,948	0,000
Ítem_5_preg_1	1,721	0,179	Ítem_9_preg_b	6,305	0,002
Ítem_5_preg_2	2,897	0,056	Ítem_10	1,517	0,220
Ítem_5_preg_3	1,482	0,228	Ítem_11	7,441	0,001
Ítem_6_p_1	6,108	0,002			

Prueba Post hoc

IBM (s.f) señala que “una vez que se ha determinado que existen diferencias entre las medias, las pruebas de rango post hoc y las comparaciones múltiples por parejas permiten determinar qué medias difieren. Las pruebas de rango identifican subconjuntos homogéneos de medias que no se diferencian entre sí. Las comparaciones múltiples por parejas contrastan la diferencia entre cada pareja de medias y generan una matriz donde los asteriscos indican las medias de grupo significativamente diferentes a un nivel alfa de 0,05”.

Se ha utilizado el método **DMS**, el cual “utiliza pruebas t para realizar todas las comparaciones por pares entre las medias de los grupos. La tasa de error no se corrige para realizar múltiples comparaciones” (IMB, s.f).

En la Tabla 5.48 se presenta la prueba post hoc utilizando el método DMS, para aquellas comparaciones en las que se han encontrado diferencias por nivel educativo (curso), se analizarán los resultados obtenidos por ítem.

Para el Ítem 1 se observan diferencias significativas entre los cursos 8° Año de primaria y 2° medio y 2° medio y 4° medio. Se puede decir que los resultados obtenidos por los alumnos de segundo medio (34.2%) son mejores que los alcanzados por los estudiantes de octavo año de primaria (15.8%), además, lo mismo ocurre si comparamos segundo medio con cuarto medio (26%). Es decir, los estudiantes de segundo medio, son los que entregan

mayor cantidad de definiciones acertadas del concepto de muestra. Por otro lado, no se observan diferencias significativas entre los cursos de 8° año de primaria y 4° medio.

Tabla 5.48. Prueba Post hoc DMS por nivel educativo (curso)

Variable dependiente	(I) Curso	(J) Curso	P-valor	Variable dependiente	(I) Curso	(J) Curso	P-valor
Ítem_1	8° Básico	2° Medio	0,006	Ítem_6_p_4	8° Básico	2° Medio	0,535
	8° Básico	4° Medio	0,500		8° Básico	4° Medio	0,004
	2° Medio	4° Medio	0,017		2° Medio	4° Medio	0,007
Ítem_2_preg_a	8° Básico	2° Medio	0,009	Ítem_6_m_1	8° Básico	2° Medio	0,533
	8° Básico	4° Medio	0,000		8° Básico	4° Medio	0,010
	2° Medio	4° Medio	0,001		2° Medio	4° Medio	0,018
Ítem_2_preg_b	8° Básico	2° Medio	0,048	Ítem_6_m_2	8° Básico	2° Medio	0,601
	8° Básico	4° Medio	0,000		8° Básico	4° Medio	0,024
	2° Medio	4° Medio	0,004		2° Medio	4° Medio	0,001
Ítem_2_preg_c	8° Básico	2° Medio	0,007	Ítem_8_preg_1	8° Básico	2° Medio	0,011
	8° Básico	4° Medio	0,037		8° Básico	4° Medio	0,213
	2° Medio	4° Medio	0,550		2° Medio	4° Medio	0,152
Ítem_3_Andrea	8° Básico	2° Medio	0,112	Ítem_8_preg_2	8° Básico	2° Medio	0,000
	8° Básico	4° Medio	0,002		8° Básico	4° Medio	0,010
	2° Medio	4° Medio	0,060		2° Medio	4° Medio	0,031
Ítem_3_preg_b	8° Básico	2° Medio	0,937	Ítem_8_preg_3	8° Básico	2° Medio	0,000
	8° Básico	4° Medio	0,004		8° Básico	4° Medio	0,002
	2° Medio	4° Medio	0,000		2° Medio	4° Medio	0,436
Ítem_4	8° Básico	2° Medio	0,036	Ítem_9_preg_a	8° Básico	2° Medio	0,057
	8° Básico	4° Medio	0,000		8° Básico	4° Medio	0,000
	2° Medio	4° Medio	0,000		2° Medio	4° Medio	0,003
Ítem_6_p_1	8° Básico	2° Medio	0,001	Ítem_9_preg_b	8° Básico	2° Medio	0,003
	8° Básico	4° Medio	0,119		8° Básico	4° Medio	0,631
	2° Medio	4° Medio	0,039		2° Medio	4° Medio	0,004
Ítem_6_p_2	8° Básico	2° Medio	0,082	Ítem_11	8° Básico	2° Medio	0,000
	8° Básico	4° Medio	0,062		8° Básico	4° Medio	0,030
	2° Medio	4° Medio	0,000		2° Medio	4° Medio	0,067
Ítem_6_p_3	8° Básico	2° Medio	0,000				
	8° Básico	4° Medio	0,005				
	2° Medio	4° Medio	0,000				

En el Ítem 2 para las preguntas a y b se observa que hay diferencias significativas entre los tres cursos. En ambos apartados se observa que los alumnos de segundo medio responden mejor, seguidos por los alumnos de cuarto medio y los de menores resultados son los estudiantes de octavo año de primaria. En el caso de la pregunta c, sólo se observa que no hay diferencias significativas entre los cursos de 2° y 4° medio, ya que los porcentajes de respuesta correcta son muy similares.

Para el Ítem 3 en el Método de Andrea se observa que no hay diferencias significativas entre los alumnos de 8° año de primaria y 2° medio. Lo mismo ocurre para la pregunta b de este ítem.

En el caso del Ítem 4 se observa que hay diferencias significativas en los tres cursos. En este caso, los que entregan una menor cantidad de respuestas correctas son los alumnos de cuarto medio, un porcentaje levemente mejor se ve reflejado en las respuestas de los alumnos de octavo año de primaria, siendo los estudiantes de segundo medio quienes responden de mejor manera esta pregunta.

En el Ítem 6, para la pregunta de identificar la población 1 no existen diferencias significativas entre los cursos de octavo año de primaria y cuarto medio. Para el caso identificar la población 2 se observa que no hay diferencias significativas entre 8° año de primaria y 2° medio como también entre 8° año de primaria y 4° medio. Cuando los estudiantes identifican la población 3 se observa que hay diferencias significativas entre los cursos, en este caso, se observa que los resultados van mejorando paulatinamente con la edad de los participantes. Para identificar la población 4; la muestra 1 y la muestra 2, se observa que no hay diferencias significativas entre los cursos de 8° año de primaria y 2° medio.

Para el Ítem 8 en la pregunta 1 y en la pregunta 3 hay diferencias significativas entre 8° año de primaria y 2° medio y entre 8° año de primaria y 4° medio, además, se ve que no hay diferencias significativas entre los cursos de 2° y 4° medio. En la pregunta 2 se observa que hay diferencias significativas entre los grupos, en este caso los alumnos de 8° año de primaria tienen menor rendimiento, mejora un poco el porcentaje de respuesta correcta en los alumnos de 4° medio y los estudiantes de segundo medio son los que mejor responden esta pregunta.

En el Ítem 9 para la pregunta 1 vemos que no hay diferencias significativas entre los cursos de 8° año de primaria y 2° medio. Para la pregunta 2 se observa que no hay diferencias significativas para los cursos de 8° año de primaria y 4° medio.

Finalmente, para el Ítem 11 no se observan diferencias significativas entre los cursos de 2° y 4° medio, pero si se observan diferencias significativas entre 8° año de primaria y 2° medio y entre 8° año de primaria y 4° medio.

5.7. Diferencias entre sexos

En la Tabla 5.49 se presenta el ANOVA de un factor por **sexo** de los participantes.

Tabla 5.49. ANOVA de un factor por sexo

Preguntas	F	P-valor	Preguntas	F	P-valor
Ítem_1	4,677	0,031	Ítem_6_p_2	3,631	0,057
Ítem_2_preg_a	0,798	0,372	Ítem_6_p_3	5,899	0,015
Ítem_2_preg_b	1,309	0,253	Ítem_6_p_4	4,617	0,032
Ítem_2_preg_c	0,008	0,928	Ítem_6_m_1	14,550	0,000
Ítem_2_preg_c_2	0,003	0,956	Ítem_6_m_2	9,890	0,002
Ítem_3_Andrea	0,197	0,657	Ítem_6_m_3	6,169	0,013
Ítem_3_Elena	0,506	0,477	Ítem_6_m_4	4,622	0,032
Ítem_3_Pedro	0,000	0,998	Ítem_7	4,914	0,027
Ítem_3_Luis	3,879	0,049	Ítem_8_preg_1	9,466	0,002
Ítem_3_Sara	3,211	0,073	Ítem_8_preg_2	13,911	0,000
Ítem_3_preg_b	2,000	0,158	Ítem_8_preg_3	2,878	0,090
Ítem_4	2,019	0,156	Ítem_9_preg_a	1,043	0,307
Ítem_5_preg_1	0,394	0,531	Ítem_9_preg_b	3,513	0,061
Ítem_5_preg_2	3,356	0,067	Ítem_10	16,932	0,000
Ítem_5_preg_3	0,054	0,816	Ítem_11	8,203	0,004
Ítem_6_p_1	10,332	0,001			

En la Tabla 5.49 se observa que no existen diferencias significativas para los Ítems: Ítem 2; Ítem 3 para el método de Andrea, Elena, Pedro, Sara y la pregunta b; Ítem 4, Ítem 5; Ítem 6 población 2; Ítem 8 pregunta 3; Ítem 9. Para todos los otros ítems el p-valor indica que existen diferencias significativas entre los grupos.

5.8. Diferencias entre tipos de centro

En la Tabla 5.50 se presenta el ANOVA de un factor por **tipo de centro**.

Se observa que no existen diferencias significativas para los Ítems: Ítem 2 pregunta a; Ítem 3 para el método de Andrea, método de Pedro, método de Luis; Ítem 5 pregunta 1; Ítem 6 población 1 y muestras 1 y 2; Ítem 7; Ítem 9 pregunta a. Para todos los otros ítems el p-valor indica que existen diferencias significativas entre los grupos.

Tabla 5.50. ANOVA de un factor por Tipo de Centro

Preguntas	F	P-valor	Preguntas	F	P-valor
Ítem_1	17,206	0,000	Ítem_6_p_2	7,254	0,001
Ítem_2_preg_a	2,149	0,117	Ítem_6_p_3	22,481	0,000
Ítem_2_preg_b	7,671	0,000	Ítem_6_p_4	11,364	0,000
Ítem_2_preg_c	6,743	0,001	Ítem_6_m_1	0,207	0,813
Ítem_2_preg_c_2	5,336	0,005	Ítem_6_m_2	0,788	0,455
Ítem_3_Andrea	1,001	0,368	Ítem_6_m_3	16,513	0,000
Ítem_3_Elena	4,818	0,008	Ítem_6_m_4	14,128	0,000
Ítem_3_Pedro	2,484	0,084	Ítem_7	1,151	0,317
Ítem_3_Luis	1,770	0,171	Ítem_8_preg_1	6,908	0,001
Ítem_3_Sara	5,812	0,003	Ítem_8_preg_2	5,190	0,006
Ítem_3_preg_b	3,046	0,048	Ítem_8_preg_3	9,221	0,000
Ítem_4	13,270	0,000	Ítem_9_preg_a	2,661	0,070
Ítem_5_preg_1	2,134	0,119	Ítem_9_preg_b	10,956	0,000
Ítem_5_preg_2	4,824	0,008	Ítem_10	5,576	0,004
Ítem_5_preg_3	1,214	0,297	Ítem_11	5,440	0,004
Ítem_6_p_1	2,278	0,103			

Además, se ha considerado observar más de cerca las diferencias mencionadas por tipo de centro. En la Tabla 5.51 se presenta la prueba post hoc utilizando el método DMS, para aquellas comparaciones en las que se han encontrado diferencias por tipo de centro, se analizarán los resultados obtenidos por ítem.

Tabla 5.51. Prueba Post hoc DMS por Tipo de Centro

Variable dependiente	(I) Tipo de Centro	(J) Tipo de Centro	P-valor	Variable dependiente	(I) Tipo de Centro	(J) Tipo de Centro	P-valor
Ítem_1	M	PS	0,000	Ítem_6_p_3	M	PS	0,000
	M	PP	0,002		M	PP	0,000
	PS	PP	0,626		PS	PP	0,187
Ítem_2_preg_b	M	PS	0,000	Ítem_6_p_4	M	PS	0,632
	M	PP	0,036		M	PP	0,000
	PS	PP	0,754		PS	PP	0,000
Ítem_2_preg_c	M	PS	0,004	Ítem_6_m_3	M	PS	0,005
	M	PP	0,003		M	PP	0,000
	PS	PP	0,307		PS	PP	0,000
Ítem_2_preg_c2	M	PS	0,004	Ítem_6_m_4	M	PS	0,013
	M	PP	0,003		M	PP	0,000
	PS	PP	0,307		PS	PP	0,001
Ítem_3_Elena	M	PS	0,024	Ítem_8_preg_1	M	PS	0,016
	M	PP	0,132		M	PP	0,031
	PS	PP	0,005		PS	PP	0,000
Ítem_3_Sara	M	PS	0,143	Ítem_8_preg_2	M	PS	0,003
	M	PP	0,001		M	PP	0,558
	PS	PP	0,024		PS	PP	0,018
Ítem_3_preg_b	M	PS	0,232	Ítem_8_preg_3	M	PS	0,000
	M	PP	0,017		M	PP	0,499
	PS	PP	0,129		PS	PP	0,002
Ítem_4	M	PS	0,000	Ítem_9_preg_b	M	PS	0,007
	M	PP	0,000		M	PP	0,000
	PS	PP	0,147		PS	PP	0,014

Ítem_5_2	M	PS	0,002	Ítem_10	M	PS	0,002
	M	PP	0,528		M	PP	0,049
	PS	PP	0,184		PS	PP	0,955
Ítem_6_p_2	M	PS	0,029	Ítem_11	M	PS	0,006
	M	PP	0,000		M	PP	0,262
	PS	PP	0,044		PS	PP	0,006

Leyenda: **M**: Centro Municipal (Público); **PS**: Centro Particular Subvencionado (Concertado); **PP**: Centro Particular Pagado (Privado)

Para el Ítem 1 se observa que no existen diferencias significativas entre los centros particular subvencionados y los centros particulares pagados, en los otros dos casos si existen diferencias significativas.

En el caso del Ítem 2 se observa que no existen diferencias significativas entre los centros particulares subvencionados y los centros particulares pagados para la pregunta b y la pregunta c, pero si existen diferencias significativas entre los centros municipales y los particulares subvencionados, como también, entre los centros municipales y particulares pagados.

En el Ítem 3, para el método de Elena no se observan diferencias significativas entre los centros municipales y los particulares pagados. En los otros casos, si hay diferencias significativas. Para el método de Sara y la pregunta b, no se observan diferencias significativas entre los centros municipales y los particulares subvencionados. En los otros casos, si hay diferencias significativas.

En el caso del Ítem 4 no existen diferencias significativas entre los centros particulares subvencionados y los centros particulares pagados, pero sí en los otros casos.

Para el Ítem 6 población 2; muestra 3 y muestra 4 existen diferencias significativas entre los grupos. En el caso de la población 3, no se observan diferencias significativas entre los centros particulares subvencionados y los centros particulares pagados.

En el Ítem 8, pregunta 1 se observan diferencias significativas entre los grupos. Para la pregunta 2 y pregunta 3 no se observan diferencias significativas entre los centros municipales y los centros particulares pagados.

Para el ítem 9 pregunta b, se observan diferencias significativas entre los grupos.

En el caso del Ítem 10 no se observan diferencias significativas entre los centros particulares subvencionados y los centros particulares pagados.

En el Ítem 11 no se observan diferencias significativas entre los centros municipales y los centros particulares pagados.

5.9. Conflictos semióticos detectados en el estudio de evaluación

En este apartado se sintetizan los principales conflictos semióticos (CS) de tipo cognitivo, es decir, todas aquellas disparidades reflejadas en las prácticas matemáticas ejecutadas en las respuestas de los estudiantes y que han sido detectados en cada uno de los ítems que conforman el cuestionario de evaluación.

CS1 No distinguir la relación parte-todo al emplear el concepto de muestra. En este caso, algunos estudiantes no son capaces de explicar que la muestra es una parte de un conjunto más grande y señalan explicaciones incompletas sin fundamentos. Este conflicto aparece en el Ítem 1, como también, en la “pregunta a” del Ítem 2 ya que los estudiantes no son capaces de identificar el concepto de muestra en un contexto particular. Además, en el Ítem 6, la mayoría de los estudiantes no es capaz de distinguir la relación parte- todo para identificar la población y la muestra respectiva de cada uno de los casos expuestos.

CS2 No aplicar la definición de muestra. Los estudiantes en la “pregunta b” del Ítem 2 no son capaces de identificar que para seleccionar una muestra representativa de la población es necesario considerar factores como el excesivo tiempo que puede llevar el estudio; el costo de implementación de la entrevista; o sobre el acceso a entrevistar a la población total, siendo uno de los conceptos básicos del muestreo que mencionan Batanero, Begué, Gea y Roa (2019). Este conflicto también se ve reflejado en las respuestas de la “pregunta 3b” del Ítem 3, en donde deben indicar cuál es el mejor método de muestreo, algunos estudiantes señalan que deben entrevistarse a todos los sujetos que conforman la población. En el Ítem 6, también se aprecia que los estudiantes no son capaces de aplicar esta definición para distinguir entre la población y muestra solicitadas de cada uno de los enunciados dados.

CS3 Insensibilidad al tamaño de la muestra, se observa la aplicación del sesgo asociado a la *heurística de representatividad* (Kahneman, Slovic y Tversky, 1982), es decir, los

estudiantes consideran que una muestra pequeña también es representativa de la población a la cual pertenecen para la “pregunta c” del Ítem 2.

CS4 *Creencia en la ley de los pequeños números*, también se observa que los estudiantes aplican esta creencia para la “pregunta c” del Ítem 2, es decir, que esperan que la muestra pequeña tenga el mismo comportamiento que la población (Kahneman, Slovic y Tversky, 1982).

CS5 *No detectan el sesgo en el método de muestreo*. En los diversos métodos de muestreos propuestos en el Ítem 3 se aprecia que mayoritariamente los estudiantes no identifican las muestras sesgadas, en algunos casos, este conflicto también se ve reflejado en las respuestas entregadas en el Ítem 4, donde algunos estudiantes proporcionan un método de muestreo sesgado para elegir la muestra pedida.

CS6 *Suponer que la aleatoriedad indica máxima variabilidad*. Begué (2019) indica que este conflicto “aparece en aquellas respuestas en las que el valor del rango de la muestra generada se considera excesivo” (p.243), es decir, al ver que la muestra es aleatoria piensan que es un método difícil, o que “cualquier cosa puede suceder”, este conflicto se presenta mayoritariamente en el método de Luis (Ítem 3), que corresponde a un muestreo aleatorio.

CS7 *Confusión entre suceso aleatorio y suceso equiprobable*. De acuerdo a Begué (2019) “el sujeto asigna la *equiprobabilidad* a cada suceso por el carácter aleatorio del fenómeno. Por tanto, tienen una comprensión incorrecta de la aleatoriedad, ya que le asignan una propiedad que no le corresponde. En los argumentos que se sitúan en este conflicto subyace el *sesgo de equiprobabilidad*, descrito por Lecoutre (1992)” (p.242). En nuestro estudio, algunos estudiantes señalan que se debe *escoger a la mitad* de los participantes cuando responden “pregunta 3b” del Ítem 3, en donde deben indicar cuál es el mejor método de muestreo.

CS8 *No reconocer la representatividad de la muestra*. En el caso del Ítem 5, ante la primera muestra presentada, la mayoría de los estudiantes no identifica que la muestra es representativa debido a su proceso de selección, ya que no consideran el muestreo aleatorio como un método apropiado, sino que sus opiniones se basan más en el tamaño de la muestra. Pero si identifican las otras dos muestras no representativas de forma correcta, lo que es un aspecto contradictorio.

CS9 *No aplicar la proporción de la muestra para determinar la proporción de la población.* En el Ítem 7 algunos los estudiantes no son capaces de calcular la proporción de la población a partir de la proporción de cuatro muestras dadas.

CS10 *Aplicación estricta de la proporcionalidad, en relación con el tamaño de la muestra, sin considerar la variabilidad del muestreo.* Begué (2019) indica que “Estas respuestas se caracterizan porque el estudiante identifica la probabilidad en términos de porcentaje y aplica, de manera determinista, dicho porcentaje en el cálculo de las cuatro estimaciones generadas, lo cual conduce a muestras con ninguna variabilidad” (p. 244). En nuestro caso, se les solicita a los estudiantes repetir 6 estimaciones de posibles muestras (Ítem 9, pregunta b), en las cuales varios entregan respuestas iguales a la proporción de datos originales sin aplicar la variabilidad en cada una de las muestras solicitadas.

CS11 *No relacionar la probabilidad con la frecuencia.* Begué (2019) señala que “En este caso el estudiante no comprende el concepto de probabilidad frecuencial ni la convergencia de la frecuencia relativa a la probabilidad teórica” (p. 243). En el caso del Ítem 8, no son capaces de determinar las proporciones de la muestra para poder generalizar a los valores de la población.

CS12 *Inconsistencia entre la probabilidad teórica asumida y la asignación frecuencial de los datos.* Begué (2019) alude que “se corresponde con aquellas respuestas en las que el sujeto no proporciona una muestra desde la probabilidad identificada”. En nuestro caso, en el Ítem 9 pregunta b se aprecian muchas respuestas que no tienen relación con la probabilidad teórica del experimento debido a que entregan valores muy extremos en las muestras con una variabilidad excesiva.

CS13 *Falta de percepción de la independencia de resultados en la repetición de un experimento aleatorio* Begué (2019) menciona que “Los estudiantes señalan que esperan que los resultados obtenidos al repetir el experimento sean similares al obtenido en el enunciado. En este caso, observamos que la respuesta generada se caracterizada por dar valores en los que la proporción es similar al dato, por lo que estos estudiantes razonarían de acuerdo la heurística de la representatividad y, en particular, presentarían el sesgo denominado recencia positiva (Tversky et al. ,1982)” (p. 244). Esto se aprecia en el apartado b del ítem 9, en donde los estudiantes no escriben muestras aplicando distintas asignaciones frecuenciales, sino que,

en algunos casos, solo esperan que en todas las repeticiones del experimento se obtenga el valor esperado, sin aplicar tampoco la variabilidad en dichos resultados.

C14 *Enumeración no sistemática*. Para el caso del Ítem 11 uno de los principales conflictos identificados tiene relación con lo que describen Navarro-Pelayo, Batanero y Godino (1996), quien señalan que “este tipo de error fue descrito por Fischbein y Gazit (1988), y consiste en resolver el problema por enumeración, mediante ensayo y error, sin un procedimiento recursivo que lleve a la formación de todas las posibilidades” (p. 34). En nuestro estudio se aprecia que los estudiantes anotan el listado incompleto de las muestras con o sin reposición solicitadas en el Ítem 11.

C15 *No recordar la fórmula combinatoria correcta*. Navarro-Pelayo, Batanero y Godino (1996) mencionan que este es uno de los errores frecuentes en la resolución de problemas combinatorios, en nuestro caso, en el Ítem 11 los alumnos identifican la combinación o permutación respectiva, pero no son capaces de aplicar la fórmula correctamente cometiendo errores al reemplazar los valores de los parámetros, además, algunos calculan la fórmula combinatoria, pero no anotan el listado de elementos que conforman cada muestra, es decir, no identifican que la fórmula sólo les proporciona la cantidad de elementos de cada muestra.

CS16 *Interpretación errónea del diagrama de árbol*. Aunque los alumnos conocen su aplicación en problemas de probabilidad, no todos emplean el diagrama de árbol para obtener todas las muestras posibles del Ítem 11.

5.10. Conclusiones sobre los resultados del estudio de evaluación

Al concluir este capítulo, se considera que las respuestas obtenidas por la muestra de 1241 estudiantes chilenos de secundaria, proporcionan información suficiente sobre su comprensión acerca del muestreo, como también sobre las principales dificultades relacionadas al entendimiento del concepto, resultados que permiten tener un acercamiento a las ideas que poseen los estudiantes de secundaria sobre las nociones intuitivas de este concepto.

Se presentarán algunas conclusiones generales del capítulo ordenadas por los bloques de preguntas y su respectivo objetivo.

Para el Bloque 1 (Ítems 1 y 2) cuyo objetivo es *comprender la definición del concepto de muestra*, se aprecia que mayoritariamente los estudiantes comprenden el concepto de muestra y son capaces de dar su significado en diversos contextos, mencionando además diferentes aspectos relevantes para definir lo que es una muestra, como por ejemplo, “*es una parte de un conjunto más grande*”, en las respuestas más sencillas se observa que los estudiantes entregan ejemplos en algún contexto para explicar lo que entienden por muestra, estos resultados son similares a los obtenidos por Watson y Moritz (2000); Watson (2004) y Meletiou-Mavrotheris y Paparistodemou (2015).

Cuando los estudiantes intentan explicar por qué se elige una muestra de una determinada población, la mayoría entrega respuestas que reflejan cuestiones prácticas (*es difícil entrevistar a todos los participantes; lleva mucho tiempo*) o mencionan ideas sobre el contexto en el que se realiza el muestreo (*es demasiado lento visitar todas las escuelas*). Además, cuando se les pregunta sobre si están de acuerdo en que se elija una muestra pequeña y sin hacer mención al método de selección, la gran mayoría responde afirmativamente, es decir, responden erróneamente y olvidan que es necesario *seleccionar una muestra representativa* de la población, por otro lado, si se les solicita que indiquen una muestra para un tamaño poblacional desconocido, mencionan diferentes valores y algunos métodos de selección, aunque no se refieren expresamente al concepto de *representatividad de la muestra*. Meletiou-Mavrotheris y Paparistodemou (2015) mencionan que las dificultades que presentan los estudiantes para reconocer el sesgo potencial del método de muestreo podrían atribuirse parcialmente a la naturaleza del Ítem 2, ya que, como señalan Watson y Moritz (2000), el éxito de los estudiantes en identificar el sesgo depende en gran medida de las pistas contextuales proporcionadas en el enunciado de la situación problema.

Si consideramos los errores que cometen los estudiantes, se observa que son susceptibles a la *insensibilidad al tamaño de la muestra* (Tversky y Kahneman, 1974), ya que asumen que la muestra seleccionada, independientemente de su tamaño, siempre es representativa a la población a la que pertenece, lo que les lleva a cometer graves errores de interpretación de la información entregada. Además, se percibe que algunos estudiantes creen en la *ley de los pequeños números* (Kahneman, Slovic y Tversky 1982), esto es, piensan que, al seleccionar muestras no suficientemente grandes, la distribución muestral se distribuye de

la misma manera que la distribución de la población, independientemente del tamaño de la muestra.

En el Bloque 2 (Ítem 3) cuyo objetivo es *detectar sesgo en el método de muestreo*, la mayoría de los estudiantes identifica correctamente *métodos de muestreo estratificados*, pues consideran que este tipo de métodos es *más justo y equitativo*, debido a que participan la misma cantidad de niños y niñas, de acuerdo a Jacobs (1997), la preferencia por el muestreo estratificado puede estar relacionada con las concepciones estadísticas no normativas de equidad que emplean los niños, ya que en su investigación, al incluir muestras en las que niños y niñas existían en proporciones desiguales en la población, descubrió que los niños seguían insistiendo en seleccionar un número igual de cada género.

Seguidamente, aunque en una menor medida, nuestros estudiantes señalan su preferencia por los *métodos de muestreo aleatorio* para elegir una muestra representativa de la población. En este aspecto se muestran resultados mejores que los obtenidos por Jacobs (1997) y Meletiou-Mavrotheris y Paparistodemou (2015), ya que en estas investigaciones mencionan que la mayoría de los niños expresaron su disgusto por el muestreo aleatorio simple y mayor preferencia por la estratificación de la muestra.

Por otro lado, cuando se enfrentan a métodos sesgados, sus creencias impiden que sean capaces de detectar dichos sesgos, es por ello que en los *métodos de muestreo autoseleccionado* tienen problemas de identificar el sesgo respectivo en donde la autoselección impide que los sujetos elegidos sean representativos de la población, esto se debe principalmente a que los estudiantes creen que la participación voluntaria es más justa, de la misma forma, tienen problemas a la hora de identificar el sesgo en muestras que son elegidas con *métodos restringidos* en donde los participantes poseen alguna característica común. Estos resultados son semejantes a lo señalado en Jacobs (1997), Watson et al. (2003), Watson y Kelly (2005) y Meletiou-Mavrotheris y Paparistodemou (2015).

Además, al analizar los argumentos entregados por los alumnos, éstos hacen alusión a razones no estadísticas en sus evaluaciones de los distintos métodos de muestreos expuestos, por ejemplo, una de esas razones fue la “*imparcialidad*” del proceso, es decir, que los niños y las niñas estén representados por igual, Watson et al., (2003) menciona que este tipo de razonamiento se basa principalmente en creencias emotivas y personales de lo que es una

encuesta justa, en este sentido Wroughton, McGowan, Weiss y Cope (2013) sugieren que los estudiantes en lugar de analizar la calidad del método de muestreo utilizado por medio de principios estadísticos, determinan la validez de las conclusiones de un estudio basándose en si esa conclusión está o no de acuerdo con su opinión.

Los resultados del Bloque 3 (Ítem 4) cuyo objetivo es *sugerir un método de muestreo* revelan que la mayoría de los estudiantes que participaron en nuestro estudio es capaz de sugerir un método de muestreo (85% aproximadamente), de ese porcentaje solamente un 5% menciona obtener una muestra de un grupo sesgado que posee alguna característica particular, un 41% señala un método de muestreo y la justificación de su elección, y el otro 39% menciona un método de muestreo sin justificar dicha elección.

Para el Bloque 4 (Ítem 5) cuyo objetivo es *decidir si una muestra es representativa*, la mayoría de los estudiantes identifica correctamente los dos casos en que las muestras *no son representativas*, pero en el caso de la *muestra aleatoria* les dificulta identificar que dicha situación corresponde a un *muestreo aleatorio simple*, por lo que no reconocen la representatividad de la muestra y se centran principalmente en argumentos que aluden al tamaño de la muestra sin considerar la *aleatoriedad* como un componente importante para que todos los participantes tengan la misma posibilidad de ser elegidos.

En el Bloque 5 (Ítem 6) cuyo objetivo es *identificar la población y la muestra* se obtienen muy bajos porcentajes de logro. Se cree que esto se debe principalmente a que el Ítem es adaptado de una actividad propuesta de un texto escolar que emplea contextos demasiado difíciles para poder identificar la relación parte-todo en cada uno de los casos presentados y así deducir la población y la muestra, es decir, usan contextos no tan cercanos a la realidad de los estudiantes, lo que puede haber influido en que se obtuvieran resultados tan bajos.

Para el caso del Bloque 6 (Ítems 7 y 8) cuyo objetivo es *dada una proporción en la muestra determinar la proporción en la población*, se observa que un buen porcentaje de los estudiantes son capaces de determinar la proporción de la población dada la proporción de muestra. Algunos de los resultados obtenidos en el Ítem 8 son mejores que los entregados por futuros profesores en el estudio de Gómez (2014). Por ello se puede valorar que mayoritariamente los estudiantes de secundaria chilenos que participaron en nuestro estudio

aplican correctamente el razonamiento proporcional, ya que un buen porcentaje de ellos logra determinar la proporción de la población dada la proporción de la muestra.

Los resultados del Bloque 7 (Ítems 9 y 10) cuyo objetivo es *dada la proporción en una población determinar la proporción en varias muestras* indican que un buen porcentaje logra este objetivo. En este caso, cuando los alumnos deben indicar el valor posible de diferentes muestras seleccionadas, un número considerable de sus respuestas se centran en el valor esperado, por otro lado, al intentar aplicar la variabilidad en las muestras, anotan muestras con variabilidad excesiva, empleando valores extremos. Shaughnessy, Ciancetta y Canada (2004) señalan que los estudiantes al tener que indicar varias muestras de un experimento, mayoritariamente se centran directamente en el valor esperado, como suele ser el caso cuando solo se les pregunta sobre una repetición del experimento.

Finalmente, en el Bloque 8 (Ítem 11) cuyo objetivo es *obtener todas las muestras posibles de una población finita*, se obtienen bajos resultados de logro, debido a que mayoritariamente los estudiantes contestan parcialmente este ítem, señalando una de las muestras pedidas, en otros casos, responden identificando algunos elementos de las dos muestras y no completando el listado de elementos totales solicitado.

Al considerar las comparaciones realizadas por nivel educativo (curso), se observa a nivel general que mayoritariamente los alumnos de 2º año de secundaria aportan una mayor cantidad de respuestas correctas que los estudiantes de 4º año de secundaria, quienes responden sólo responden mejor el ítem 3, método de Andrea y método de Luis; la parte de identificar las poblaciones y la muestra 4 del Ítem 6. Por otro lado, se puede mencionar que los estudiantes de 8º año de primaria son los que tienen una menor comprensión de los temas de muestreo, pero en el caso del Ítem 4, presentan mejores resultados que los estudiantes de 4º año de secundaria.

CAPÍTULO 6:

CONCLUSIONES

- 6.1. Introducción
- 6.2. Conclusión respecto a los objetivos generales
- 6.3. Conclusión respecto a las hipótesis
- 6.4. Principales aportes
- 6.5. Limitaciones del estudio
- 6.6. Futuras líneas de investigación

6.1. INTRODUCCIÓN

Como se han presentado conclusiones parciales en los capítulos anteriores, en este capítulo se expondrán las conclusiones generales de la investigación realizada sobre la comprensión de los conceptos relativos al muestreo por parte de estudiantes de educación secundaria chilenos para cada uno de los objetivos e hipótesis planteadas en el Capítulo 1.

Así mismo, se presentarán las principales contribuciones derivadas de este estudio, junto con las limitaciones y se proponen algunas posibles futuras líneas de investigación para continuar profundizando en el tema principal de esta memoria.

6.2. CONCLUSIÓN RESPECTO A LOS OBJETIVOS GENERALES

En este apartado, como se ha expuesto en el Capítulo 1, se expondrán las principales reflexiones derivadas de los tres objetivos generales planteados en esta investigación.

Objetivo 1. Analizar la presencia del concepto de muestreo en las directrices curriculares de Educación Secundaria chilena (MINEDUC, 2013, 2015), comparando con las directrices españolas y otras internacionales y analizando su implementación en los libros de textos escolares chilenos.

Se ha considerado pertinente incorporar este objetivo principalmente pues al momento de plantear este estudio, no existían investigaciones previas desarrolladas en el contexto escolar chileno que se enfocaran en la comprensión del muestreo. Además, de presentar un panorama más específico sobre cómo se incorporan los conceptos relativos al muestreo durante la etapa secundaria, tanto en Chile (MINEDUC, 2012; 2013; 2015), como en otros lineamientos internacionales, a considerar los de España (MECD, 2015) y Estados Unidos (CCSSI, 2010).

Este panorama inicial, también contribuye a identificar las principales similitudes y diferencias entre las distintas orientaciones curriculares analizadas, por ejemplo, en cuanto al nivel educativo, se puede señalar que en todas estas directrices se parte introduciendo paulatinamente los temas de estadística de una manera más descriptiva, dejando los temas relacionados con inferencia, en especial los de muestreo para los niveles de enseñanza secundaria, es decir, estas orientaciones siguen en la línea de lo propuesto por distintos investigadores de educación estadística, quienes sugieren que el razonamiento estadístico debe tener un rol más amplio dentro de la matemática escolar (Batanero, 2013; Ben-Zvi, Bakker y Makar, 2015).

Por otro lado, en el currículum chileno, conceptos como el sesgo en el muestreo o la variabilidad no están presentes (ver Tabla 1.14, Capítulo 1), pero si se les da un mayor protagonismo en el currículum español y en las pautas propuestas en el Proyecto Gaise (Franklin et al. 2007). Otro rasgo a destacar es que los lineamientos chilenos son más explícitos a la hora de presentar los objetivos de aprendizaje y mucho más específicos en los criterios de evaluación. Además, como se ha señalado en Ruiz-Reyes, Begué, Batanero y Contreras (2018): “el currículum chileno es más rico en el aspecto procedimental porque considera un número mayor de tareas que den lugar a la necesidad del muestreo” (p. 81).

Se debe agregar que en el Capítulo 1 se ha realizado un análisis epistémico siguiendo las orientaciones del enfoque ontosemiótico del conocimiento y la instrucción matemáticos (EOS, Godino, 2002, Godino, Batanero Y Font, 2007), en donde se presenta la caracterización de los objetos matemáticos presentes en los distintos lineamientos curriculares analizados, para definir el significado de referencia de nuestro estudio. Por otro lado, también se incluye en el Capítulo 3 el análisis epistémico de los objetos presentes en la

colección de textos escolares analizados. Además, en el Capítulo 4, se presenta la descripción de las prácticas matemáticas que los estudiantes deben llevar a cabo al responder cada ítem del cuestionario, como también, de los objetos matemáticos presentes en cada situación-problema.

Los principales resultados de este objetivo se exponen en Ruiz-Reyes, Begué, Batanero y Contreras (2018) y en Ruiz-Reyes, Ruz, Molina-Portillo y Contreras (2018).

Objetivo 2. Construir un instrumento válido y fiable para evaluar la comprensión del muestreo en los alumnos chilenos de Educación Secundaria, siguiendo unos criterios metodológicos rigurosos.

Como se mencionó anteriormente, al no encontrar en la literatura existente un instrumento que englobara los significados de referencias establecidos en el objetivo anterior, se ha creado un cuestionario que comprende 11 ítems que evalúan distintos aspectos relacionados al concepto de muestreo. Para la creación del instrumento se elaboró un banco de ítems a partir de las investigaciones previas analizadas en el Capítulo 2 y de las actividades propuestas en los textos escolares analizados en el Capítulo 3. Por medio de un proceso de revisión iterativa, hasta ajustar las traducciones y/ o adaptaciones de la selección final de preguntas para el cuestionario, se realizó una primera selección considerando el criterio de un grupo de expertos en el área de didáctica de la estadística, para posteriormente realizar la validación de dichos ítems a través de un estudio piloto con una muestra de estudiantes de secundaria chilenos.

Este cuestionario se ha aplicado a una muestra de 148 estudiantes chilenos de cuarto año de secundaria (17-18 años) para su validación, análisis que se presenta en el Capítulo 4 de esta memoria.

Además, de acuerdo a la información proporcionada por el pilotaje con las características psicométricas del instrumento, se puede afirmar que el cuestionario cumple con las expectativas previstas

Objetivo 3. Aplicar el cuestionario construido para evaluar la comprensión que alcanzan los estudiantes chilenos al comienzo y final de la educación secundaria sobre el muestreo y describir las principales dificultades observadas.

Este objetivo se puntualiza en el Capítulo 5 de esta memoria, en donde se presenta el análisis por ítem del cuestionario, aplicado a una muestra de 1241 estudiantes chilenos de educación secundaria, pertenecientes a tres niveles educativos: 8° año de educación primaria (13-14 años), curso que de acuerdo a la actual organización escolar marca el fin de dicha etapa educativa; 2° año medio (15-16 años), curso que está justo a la mitad de educación secundaria y 4° medio (17-18 años), curso que marca el fin de la etapa secundaria.

Se ha decidido incluir el nivel educativo de segundo año medio para tener una mayor cantidad de datos para poder realizar una comparación más completa de la progresión o no de la comprensión de los conceptos de muestreo en toda la etapa escolar que abarca el tema evaluado por el cuestionario.

Al realizar el análisis de las respuestas de los estudiantes (ver Capítulo 5), se considera que sus producciones proporcionan información suficiente sobre su comprensión acerca del muestreo, como también sobre las principales dificultades relacionadas al entendimiento del concepto, resultados que permiten tener un acercamiento a las ideas que poseen los estudiantes de secundaria sobre las nociones intuitivas de este concepto.

Además, se han logrado identificar y caracterizar diversos conflictos semióticos que complementan los aportados por Begué (2019).

6.3. CONCLUSIÓN RESPECTO A LAS HIPÓTESIS

Teniendo en cuenta estos objetivos, como también, las investigaciones previas descritas en el Capítulo 2, se han propuesto en el Capítulo 1 las siguientes hipótesis:

H1. Existen diferencias entre la presentación del muestreo incluidas en una selección de textos escolares y Bases Curriculares de educación secundaria chilena y el expuesto en otras directrices curriculares internacionales (específicamente, las americanas y españolas).

Como se ha señalado anteriormente en las conclusiones del primer objetivo de este estudio, existen muy pocas diferencias entre los objetos matemáticos que están expuestos en

las orientaciones curriculares y los detallados en los textos escolares, al respecto, se puede señalar que en los textos escolares analizados se halla una mayor cantidad de tipos de lenguajes y representaciones gráficas que las indicadas en las directrices curriculares

Si consideramos la comparación con las directrices internacionales, se aprecia que el currículum chileno como los textos escolares, describen una mayor cantidad de procedimientos.

H2. Existen dificultades en la comprensión del concepto de muestreo en los estudiantes chilenos, que van desapareciendo, pero no totalmente al avanzar la edad. En concreto, se esperan los siguientes resultados, que se han encontrado en investigaciones previas en otros contextos:

H2.1: En tareas de generación de muestras (formar muestras, conocida la población), la conjetura es que los estudiantes elaboran muestras que reproducen el valor medio en la población, pero no tienen en cuenta la variabilidad muestral.

Considerando las respuestas proporcionadas por los estudiantes, en el ítem 9, cuyo objetivo es *dada la proporción en una población determinar la proporción en varias muestras*, un buen porcentaje entrega una respuesta apropiada. En este caso, cuando los estudiantes deben indicar el valor posible de diferentes muestras que obtienen al repetir un experimento de extracción de caramelos, un número considerable de sus respuestas se centran en el valor esperado, es decir, las muestras generadas son iguales al valor esperado, observando así que sus respuestas no consideran la variabilidad de las muestras. Shaughnessy, Ciancetta y Canada (2004) señalan que los estudiantes al tener que indicar varias muestras de un experimento, mayoritariamente se centran directamente en el valor esperado, como suele ser el caso cuando solo se les pregunta sobre una repetición del experimento.

H2.2: En tareas de estimación (de la media de una población) a partir de la muestra, una proporción importante de estudiantes aplicarán la equiprobabilidad (dando un valor al azar). Se observarán también dificultades al aplicar la proporcionalidad inversa.

Para el caso del Bloque 6 (Ítems 7 y 8) cuyo objetivo es *dada una proporción en la muestra determinar la proporción en la población*, se observa que un buen porcentaje de los estudiantes son capaces de determinar la proporción de la población dada la proporción de muestra. Algunos de los resultados obtenidos en el Ítem 8 son mejores que los entregados por futuros profesores en el estudio de Gómez (2014). Por ello se puede valorar que mayoritariamente los estudiantes de secundaria chilenos que participaron en nuestro estudio aplican correctamente el razonamiento proporcional, ya que un buen porcentaje de ellos logra determinar la proporción de la población dada la proporción de la muestra.

6.4. PRINCIPALES APORTES

Uno de los principales aportes de este estudio es la construcción y validación de un cuestionario que abarca diferentes bloques de objetivos que permite evaluar la comprensión del muestreo para alumnos de secundaria.

Además, al no existir investigaciones previas en el contexto chileno, los resultados obtenidos permiten caracterizar las ideas intuitivas como las dificultades que presentan los alumnos participantes en este estudio, formando un precedente para la comparación con futuras réplicas del estudio.

Por otro lado, se considera una contribución las comunicaciones enviadas tanto a congresos nacionales como internacionales, de igual modo que los artículos derivados de esta investigación.

6.5. LIMITACIONES DEL ESTUDIO

En relación con el análisis de los textos escolares, se considera que una limitante fue disponer sólo de las ediciones analizadas en el Capítulo 3, ya que eran las que estaban disponibles por su fácil acceso. Sería necesario ampliar el análisis a más editoriales o realizar una comparación con series de textos escolares de otros países.

Una de las principales limitaciones del estudio, podría considerarse la extensión del cuestionario, ya que, en ciertos casos, es uno de los factores que impidió, por ejemplo, que algunos estudiantes dieran respuesta a los dos últimos ítems.

Por otro lado, se podrían considerar otros contextos más cercanos a la realidad de los estudiantes para reformular el Ítem 6, ya que algunos estudiantes no han dado respuesta, debido a que no se distingue fácilmente la relación parte-todo para determinar tanto la población como la muestra. Otro factor a considerar podría ser adaptar los términos del lenguaje empleado en la situación-problema planteada, para contribuir a que un mayor número de estudiantes responda dicho ítem.

Con respecto a las respuestas de los estudiantes, podrían realizarse algunos análisis más específicos en cuanto a los ítems que permiten expresar una mayor cantidad de argumentos.

6.6. FUTURAS LÍNEAS DE INVESTIGACIÓN

Para futuras investigaciones se podría diseñar e implementar una unidad didáctica sobre el muestreo, en donde se incorporen paulatinamente diversos campos de problemas que permitan ir generando un mayor acercamiento a la comprensión de los distintos objetos matemáticos relacionados a dicho concepto, para así promover el Razonamiento Inferencial Informal (RII) empleando los tres componentes sugeridos por Zieffler, Garfield, Delmas, y Reading (2008). Dicha secuencia, podría incorporar una sesión (o más) para la aplicación de este cuestionario y realizar además el análisis de las respuestas en una puesta en común con los estudiantes.

Más aún, sería interesante poder ampliar este estudio con algunas entrevistas a los sujetos que respondan el cuestionario, para estudiar más de cerca las ideas que poseen los alumnos de estos conceptos, pues a veces, no logran expresar todos sus conocimientos en una tarea escrita.

Por otro lado, sería interesante estudiar la comprensión del muestreo con profesores de matemática en formación, ya que al estar presente en el currículo escolar y es su desafío poder contribuir a un buen manejo de los conceptos básicos de la inferencia estadística.

REFERENCIAS

- APA, AERA y NCME (1999). *Standards for educational and psychological testing*. Washington, DC: Autores.
- Arteaga, P. (2011). *Evaluación de conocimientos sobre gráficos estadísticos y conocimientos didácticos de futuros profesores*. Tesis Doctoral. Universidad de Granada.
- Barbero, M. I., Vila, E. y Suárez, J. C. (2003). *Psicometría*. Madrid: UNED.
- Bakker, A. (2004). *Design research in statistics education: On symbolizing and computer tools*. Utrecht, the Netherlands: Beta Press.
- Batanero, C. (2004). Los retos de la cultura estadística. *Yupana*, 1(1), 27-37.
- Batanero, C. (2011). Del análisis de datos a la inferencia: Reflexiones sobre la formación del razonamiento estadístico. *XIII Conferência Interamericana de Educação Matemática*. Belo Horizonte, Brasil. ICME.
- Batanero, C. (2013). Del análisis de datos a la inferencia: Reflexiones sobre la formación del razonamiento estadístico. *Cuadernos de Investigación y Formación en Educación Matemática*, 11(8), 277-291.
- Batanero, C. y Borovcnik. M. (2016). *Statistics and probability in high school*. Rotterdam: Sense Publishers.
- Batanero, C. y Díaz, C. (2005). Análisis del proceso de construcción de un cuestionario sobre probabilidad condicional. Reflexiones desde el marco de la TFS. En A. Contreras, L. Ordóñez, y C. Batanero (Eds.), *Investigación en Didáctica de las Matemáticas. I Congreso Internacional sobre Aplicaciones y Desarrollos de la Teoría de las Funciones Semióticas* (pp. 13-36). Jaén: Universidad de Jaén.
- Batanero, C. y Díaz, C. (2008). Análisis de datos con Statgraphics. Granada: Departamento de Didáctica de la Matemática.
- Batanero, C., Begué, N., Gea, M.M. y Roa. R. (2019). El muestreo: Una idea estocástica fundamental. *Suma*, 90, 41-47.
- Batanero, C., Gea, M.M. y Begué, N. (2019). El sentido del muestreo. *Números*, 100, 121-124.
- Batanero, C., Godino, J. D. Green, D., Holmes, P., y Vallecillos, A. (1994). Errors and difficulties in understanding statistical concepts. *International Journal of Mathematics Education in Science and Technology*, 25 (4), 527-547.
- Batanero, C., Serrano, L. y Garfield, J. (1996). Heuristics and biases in secondary students' reasoning about probability. En L. Puig y A. Gutiérrez (Eds.), *Proceedings of the 20th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 43-50). Valencia: PME Group.
- Begué, N. (2016). *Comprensión de elementos básicos de muestreo en alumnos de educación secundaria obligatoria*. Tesis de Máster. Universidad de Granada.

- Begué, N. (2019). *Comprensión del muestreo y la distribución muestral en estudiantes de educación secundaria obligatoria y bachillerato*. Tesis de Máster. Universidad de Granada.
- Begué, N., Batanero, C., Ruiz, K. y Gea, M.M. (2019). Understanding sampling: a summary of the research. *Beio*, 35(1), 49-78.
- Begué, N., Ruiz-Reyes, K., Gea, M.M. y Batanero C. (2017). El muestreo en el currículo de secundaria: un estudio comparado de los currículos en España y Chile. Trabajo presentado en el IX CIBEM. Madrid.
- Ben-Zvi, D., Bakker, A. y Makar, K. (2015). Learning to reason from samples. *Educational Studies in Mathematics*, 88(3), 291-303.
- Biggs, J. B., y Collis, K. F. (1982). *Evaluating the quality of learning: the SOLO taxonomy*. New York: Academic Press.
- Bisquerra, R. (2004). *Metodología de la investigación educativa*. Madrid: La Muralla S. A.
- Burrill, G. y Biehler, R. (2011). Fundamental statistical ideas in the school curriculum and in training teachers. En *Teaching statistics in school mathematics-Challenges for teaching and teacher education* (pp. 57-69). Springer Netherlands.
- Cañizares, M. J. (1997). *Influencia del razonamiento proporcional y combinatorio y de creencias subjetivas en las intuiciones probabilísticas primarias*. Tesis doctoral. Universidad de Granada.
- Catalán. D., Pérez, B., Prieto, C. y Rupin, P. (2015). *Matemática 8° Básico*. Santiago de Chile: S. M.
- Cazorla, I. (2002). *A relação entre a habilidades viso-pictóricas e o domínio de conceitos estatísticos na leitura de gráficos*. Tesis Doctoral. Universidad de Campinas.
- CCSSI (2010). *Common Core State Standards for Mathematics*. Washington, DC: National Governors Association Center for Best Practices and the Council of Chief State School Officers.
- Chacón, A., García, G., Rupin, P., Setz, J y Villena, M. (2017). *Matemáticas 2° Medio*. Santiago de Chile: S.M.
- Chance, B., delMas, R. C., & Garfield, J. (2004). Reasoning about sampling distributions. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 295-323). Amsterdam: Kluwer.
- Chernoff, E. J. y Russell, G. L. (2012). The fallacy of composition: Prospective mathematics teachers' use of logical fallacies. *Canadian Journal for Science, Mathematics and Technology Education*, 12(3), 259-271. Disponible en: <http://dx.doi.org/10.1080/14926156.2012.704128>.
- Cobo, B. (2003). *Significados de las medidas de posición central para los estudiantes de secundaria*. Tesis doctoral. Universidad de Granada.
- Contreras, J. M. (2011). *Evaluación de conocimientos y recursos didácticos en la formación de profesores sobre probabilidad condicional*. Tesis Doctoral. Universidad de Granada.
- Cordero, F. y Flores, R. (2007). El uso de las gráficas en el discurso matemático escolar. Un

- estudio socioepistemológico en el nivel básico a través de los libros de texto. *Revista Latinoamericana de Matemática Educativa*, 10 (1), 7-38.
- Cuadras, C. M. (1981). *Análisis multivariante*. Barcelona. Eudeba.
- delMas, R. C. (1988). *The effect of activity based instruction and directed evaluation of prediction on misconceptions of probability*. Tesis Doctoral. Universidad de Minnesota.
- del Mas, R. C. (2004). A comparison of mathematical and statistical reasoning. In *The challenge of developing statistical literacy, reasoning and thinking* (pp. 79-95). Springer Netherlands.
- Del Valle, J., Muñoz, G. y Santís, M. A. (2013). *Matemática 1° Medio*. Santiago de Chile: S.M.
- Díaz, C. (2007). *Viabilidad de la enseñanza de la inferencia bayesiana en el análisis de datos en psicología*. Tesis doctoral. Universidad de Granada.
- Díaz-Levicoy, D. (2014). *Un estudio empírico de los gráficos estadísticos en libros de texto de Educación Primaria española*. Tesis de Máster. Universidad de Granada.
- Díaz-Levicoy, D. (2018). *Comprensión de gráficos estadísticos por alumnos chilenos de educación primaria*. Tesis Doctoral. Universidad de Granada.
- Escolano, A. (2009). El manual escolar y la cultura profesional de los docentes. *Tendencias Pedagógicas*, 14, 169-180.
- Ferreira, A. y Mayorga, L. (2010). Propuesta para la evaluación de libros de matemática de todos los niveles educativos. *Revista Ciencias de la Educación*, 20(35), 15-28.
- Fischbein, E. (1975). *The intuitive sources of probabilistic thinking in children*. Dordrecht, Holland: Reidel.
- Fischbein, E., y Gazit, A. (1984). Does the teaching of probability improve probabilistic intuitions? *Educational Studies in Mathematics*, 15(1), 1-24
- Font, V., Godino, J. D., y Gallardo, J. (2013). The emergence of objects from mathematical practices. *Educational Studies in Mathematics*, 82, 97–124.
- Franklin, C. Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., y Scheaffer, R. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report: A preK-12 curriculum framework*. Alexandria, VA: American Statistical Association. Disponible en: www.amstat.org/education/gaise/.
- Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, 70(1), 1-25.
- Galasso, B., Maldonado, L. y Marambio, V. (2016). *Matemáticas 1° Medio*. Santiago de Chile: Santillana.
- García-Ríos, V. N. (2013). Inferencias estadísticas informales en estudiantes mexicanos. JM Contreras, GR Cañadas, MM Gea y P. Arteaga (Eds.), *Actas de las Jornadas Virtuales en Didáctica de la Estadística, Probabilidad y Combinatoria*, 343-357.
- García-Ríos, V.N. y Sánchez, E. (2015). Dificultades en el razonamiento inferencial intuitivo. En J. M. Contreras, C. Batanero, J. D. Godino, G.R. Cañadas, P. Arteaga, E.

- Molina, M.M. Gea y M.M. López (Eds.), *Didáctica de la Estadística, Probabilidad y Combinatoria 2*, (pp. 207-214). Granada: SEIEM.
- Garfield, J. B. y delMas, R. C. (1989). Reasoning about chance events: Assessing and changing students' conceptions of probability. En C. Maher, G. Golding y B. Davis (Eds.), *Proceedings of the XI Meeting of the North American Chapter of the Psychology of Mathematics Education* (Vol.2, pp. 184-195). Rutgers: Rutgers University Press.
- Garfield, J. B. y delMas, R. C. (1990). The use of multiple items to identify misconceptions in probabilistic reasoning. En D. Vere – Jones (Eds.), *Proceedings of the Third International Conference on Teaching Statistics*. Dunedin, Australia: International Statistical Institute.
- Gea, M. M. (2014). *La correlación y regresión en Bachillerato: análisis de libros de texto y del conocimiento de los futuros profesores*. Tesis doctoral. Universidad de Granada.
- Godino, J. D. (1996). Mathematical concepts, their meaning, and understanding. En: L. Puig y A. Gutierrez (Eds.), *Proceedings of the 20th Conference of the International Group for the Psychology of Mathematics Education* (2-417-424), Universidad de Valencia.
- Godino, J. D. (2002) Un enfoque ontológico y semiótico de la cognición matemática. *Recherches en Didactiques des Mathematiques*, 22(2/3), 237-284.
- Godino, J. D. (2012). Origen y aportaciones de la perspectiva ontosemiótica de investigación en Didáctica de la Matemática. En A. Estepa, A. Contreras, J. Deulofeu, M. C. Penalva, F. J. García y L. Ordóñez (Eds.), *Investigación en Educación Matemática XVI* (pp. 49 - 68). Jaén: SEIEM.
- Godino, J. D. (2013). Indicadores de la idoneidad didáctica de procesos de enseñanza y aprendizaje de las matemáticas. *Cuadernos de Investigación y Formación en Educación Matemática 11*, 111-132.
- Godino, J. D. (2017). Construyendo un sistema modular e inclusivo de herramientas teóricas para la educación matemática. En J. M. Contreras, P. Arteaga, G. R. Cañadas, M.M. Gea, B. Giacomone y M. M. López-Martín (Eds.), *Actas del Segundo Congreso Internacional Virtual sobre el Enfoque Ontosemiótico. del Conocimiento y la Instrucción Matemáticos*.
- Godino, J. D., y Batanero, C. (1994). Significado institucional y personal de los objetos matemáticos. *Recherches en Didactique des Mathematiques*, 14(3), 325-355.
- Godino, J. D., y Batanero, C. (1998). Clarifying the meaning of mathematical objects as a priority area of research in Mathematics Education. En A. Sierpinska y J. Kilpatrick (Ed.), *Mathematics education as a research domain: A search for identity* (pp. 177-195). Dordrecht: Kluwer, A. P.
- Godino, J. D., Batanero, C. y Cañizares, M. J. (1988). *Azar y probabilidad. Fundamentos didácticos y propuestas curriculares*. Madrid: Síntesis.
- Godino, J. D., Batanero, C. y Font, V. (2007). The onto-semiotic approach to research in mathematics education. *ZDM. The International Journal on Mathematics Education*, 39(1-2), 127-135.
- Godino, J. D., Batanero, C. y Font, V. (2009). Un enfoque ontosemiótico del conocimiento y la instrucción Matemática. Versión ampliada.

- Godino, J. D., Font, V. y Wilhelmi, M. R. (2008). Análisis didáctico de procesos de estudio matemático basado en el enfoque ontosemiótico. *Publicaciones*, 38, 25-48.
- Godino, J. D., Font, V., Contreras, A. y Wilhelmi, M. R. (2006). Una visión de la didáctica francesa desde el enfoque ontosemiótico de la cognición e instrucción matemática. *Revista Latinoamericana de Investigación en Matemática Educativa*, 9 (1), 117-150.
- Gómez, E. (2014). *Evaluación y desarrollo del conocimiento matemático para enseñar la probabilidad en futuros profesores de educación primaria*. Tesis Doctoral. Universidad de Granada.
- Gómez, E., Batanero, C. y Contreras, J. M. (2014). Conocimiento matemático de futuros profesores para la enseñanza de la probabilidad desde el enfoque frecuencial. *Bolema*, 28(48), 209-229.
- Green, D. R. (1983). From thumbtacks to inference. *School Science and Mathematics* 83 (.7), 541-551.
- Green, D. R. (1991). A longitudinal study of children's probability concepts. En D. Vere Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics* (pp. 320-328). Dunedin: International Statistical Institute.
- Harradine, A., Batanero, C. y Rossman, A. (2011). Students and teachers' knowledge of sampling and inference. En C. Batanero, G. Burrill y C. Reading (Eds.), *Teaching statistics in school-mathematics-challenges for teaching and teacher education: A Joint ICMI/IASE study* (pp. 235- 246). New York: Springer.
- Heitele, D. (1975). An epistemological view on fundamental stochastic ideas. *Educational Studies in Mathematics*, 6, 187-205.
- Herbel, B. A. (2007). From intended curriculum to written curriculum: Examining the "voice" of a mathematics textbook. *Journal for Research in Mathematics Education*, 38(4), 344-369.
- Hernández, R., Fernández, C. y Baptista, P. (2014). *Metodología de la investigación*. México: McGraw Hill.
- IBM. (s.f.). ANOVA de un Factor. Recuperado de: https://www.ibm.com/support/knowledgecenter/es/SSLVMB_sub/statistics_mainhelp_d dita/spss/base/idh_ onew.html
- IBM. (s.f.). ANOVA de un factor: Contrastes post hoc. Recuperado de: https://www.ibm.com/support/knowledgecenter/es/SSLVMB_sub/statistics_mainhelp_d dita/spss/base/idh_ onew_post.html
- Jacobs, V. R. (1997). Children's understanding of sampling in surveys. Trabajo presentado en el *Annual Meeting of the American Educational Research Association*. Chicago, IL: AERA.
- Jones, G. A. y Thornton, C. A. (2005). An overview of research into the teaching and learning of probability. En G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 65-92). Dordrecht, The Netherlands: Kluwer.
- Kahneman, D., Slovic, P. y Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press.

- Kirk, J. y Miller, M. L. (1986). *Reliability and validity in qualitative research*. Newbury Park, CA: Sage University Paper.
- Ko, E.S. (2016). Development of an understanding of a sampling distribution. En D. Ben-Zvi y K. Makar (Eds.), *The Teaching and Learning of Statistics* (pp. 63-70). New York; Springer.
- Konold, C. (1989). Informal conceptions of probability. *Cognition and Instruction*, 6, 59-98.
- Krippendorff, K. (1997). *Metodología de análisis de contenido. Teoría y práctica*. Barcelona: Paidós.
- Langer, E. J. (1982). The illusion of control. En D. Kahneman, P. Slovic, y A. Tversky (Eds.), *Judgment under uncertainty: Heuristic and biases* (pp. 231-238). New York: Cambridge University Press.
- Lecoutre, M. P. (1992). Cognitive models and problem spaces in “purely random” situations. *Educational Studies in Mathematics*, 23(6), 557-568.
- León, O. G. y Montero, I. (2002). *Métodos de Investigación*. Madrid: Mc Graw Hill.
- López, F. (2002). El análisis de contenido como método de investigación. *Revista de Educación*, 4, 167-180.
- Lowe, E. y Pimm, D. (1996). ‘This is so’: a text on texts. En A. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.) *International Handbook of Mathematics Education* (pp. 371-410). Dordrecht: Kluwer.
- Makar, K., y Ben-Zvi, D. (2011). The role of context in developing reasoning about informal statistical inference. *Mathematical Thinking and Learning*. Vol. 13.
- Makar, K. y Rubin, A. (2009). A framework for thinking about informal statistical inference. *Statistics Education Research Journal*. 8(1), 82-105.
- Martínez, M. (2006). La investigación cualitativa (síntesis conceptual). *Revista de Investigación en Psicología* 9(1), 123-146.
- Martínez-Arias, R. (1995). *Psicometría: Teoría de los tests psicológicos y educativos*. Madrid. Síntesis.
- Mayén, S. (2009). *Significados de las medidas de posición central para estudiantes mexicanos de Bachillerato*. Tesis doctoral. Universidad de Granada.
- McDonald, R. P. (2013). *Test theory: A unified treatment*. Sussex, UK: Psychology Press.
- MEC (2007). *Real Decreto 1631/2006, de 29 de diciembre, por el que se establecen las enseñanzas mínimas correspondientes a la Educación Secundaria Obligatoria*. Madrid: Autor.
- MEC (2015). *Real Decreto 1105/2014, de 26 de diciembre, por el que se establece el currículo básico de la Educación Secundaria Obligatoria y del Bachillerato*.
- Meletiou-Mavrotheris, M. y Pappastodemou, E. (2015). Developing students’ reasoning about samples and sampling in the context of informal inferences. *Educational Studies in Mathematics*, 88(3), 385-404.

- Merino, R., Muñoz, V., Pérez, B. y Rupín, P. (2015). *Matemática 7º Básico*. Santiago de Chile: S.M
- Metz, K. (1998). Emergent understanding and attribution of randomness: comparative analysis of the reasoning of primary grade children and undergraduates. *Cognition and Instruction*, 16(3), 285-365.
- MINEDUC (2009). *Ley General de Educación (Ley 20370)*. Valparaíso: Congreso Nacional de Chile
- MINEDUC. (2009). *Objetivos Fundamentales y Contenidos Mínimos Obligatorios de la Enseñanza Básica y Media*. Santiago, Chile. Unidad de Currículum y Evaluación.
- MINEDUC. (2012). *Bases Curriculares: Matemática, Educación Básica*. Santiago, Chile. Unidad de Currículum y Evaluación.
- MINEDUC. (2013). *Matemática Educación Media. Bases Curriculares*. Santiago de Chile: Unidad de Currículum y Evaluación.
- MINEDUC. (2013). Programa de Estudio para Cuarto Año Básico. Matemática. Unidad de Currículum y Evaluación, MINEDUC. Santiago, Chile.
- MINEDUC. (2013). Programa de Estudio para Quinto Año Básico. Matemática. Unidad de Currículum y Evaluación, MINEDUC. Santiago, Chile.
- MINEDUC. (2013). Programa de Estudio para Sexto Año Básico. Matemática. Unidad de Currículum y Evaluación, MINEDUC. Santiago, Chile.
- MINEDUC. (2013). Programa de Estudio de Matemática. Séptimo Básico. Unidad de Currículum y Evaluación, MINEDUC. Santiago, Chile.
- MINEDUC. (2013). Programa de Estudio para Primer Año Medio. Matemática. Unidad de Currículum y Evaluación, MINEDUC. Santiago, Chile.
- MINEDUC. (2013). Programa de Estudio para Segundo Año Medio. Matemática. Unidad de Currículum y Evaluación, MINEDUC. Santiago, Chile.
- MINEDUC. (2015). *Bases Curriculares: Matemática, Educación Media*. Santiago, Chile. Unidad de Currículum y Evaluación.
- MINEDUC. (2018). *Bases Curriculares. Primero a Sexto Básico*. Santiago, Chile. Unidad de Currículum y Evaluación.
- Moore, D. (1999). *Estadística aplicada básica*. Barcelona: Antoni Bosch.
- Moore, D. (2004). *The basic practice of statistics* (3ra ed.). New York: W. H. Freeman.
- Morales, P. (1988). *Medición de actitudes en psicología y educación*. San Sebastián: Universidad de Comillas.
- Moses, L. E. (1992). The reasoning of statistical inference. En D. C. Hoaglin y D. S. Moore (Eds.), *Perspectives on contemporary statistics*. (pp. 107-121). Washington: Mathematical Association of America.
- Muñiz, J. (1994). *Teoría clásica de los tests*. Madrid: Pirámide.

- Muñoz, G., Gutiérrez, V. y Muñoz, S. (2013). *Matemáticas 4° Medio*. Santiago de Chile: Santillana.
- Muñoz, G. y Rupín, P. y Jiménez, L. (2013). *Matemáticas 2° Medio*. Santiago de Chile: S.M. National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston: VA.
- Navarro-Pelayo, V., Batanero, C., Godino, J. D. (1996). Razonamiento Combinatorio en alumnos de secundaria. *Educación matemática*, 8(1): 26-39.
- Olivera, A. y Zúñiga, S. (1987). Distribuciones empíricas. *Editorial Limusa, Serie de probabilidades y estadística*, 4.
- Olivo, E. (2008). *Significado de los intervalos de confianza para los estudiantes de ingeniería en México*. Tesis Doctoral. Universidad de Granada.
- Ortiz, J. J. (2002). *Significado de los conceptos probabilísticos elementales en los textos de Bachillerato*. Universidad de Granada.
- Ortiz, J. J., Batanero, C. y Serrano, L. (2001). El lenguaje probabilístico en los libros de texto. *Suma*, 38, 5-14.
- Pérez Echeverría, M. P. (1990). *Psicología del razonamiento probabilístico*. Madrid: Universidad Autónoma.
- Pollatsek, A., Konold, C. E., Well, A. D. y Lima, S. (1991). Beliefs underlying random sampling. *Memory and Cognition*, 12, 395-401.
- Porta, L. y Silva, M. (2003). *La investigación cualitativa: el análisis de contenido en la investigación educativa*. Córdoba: CENDIE.
- Recio, A. M. (1999). Una aproximación epistemológica a la enseñanza y aprendizaje de la demostración matemática. Tesis doctoral. Universidad de Granada.
- Rittle-Johnson, B., Siegler, R. S. y Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process, *Journal of Educational Psychology*, 93(2), 343-362.
- Rothery, A. (1980). *Children reading mathematics*. Worcester: College of Higher Education.
- Ruiz, K. (2013) *Análisis de recursos en internet para la enseñanza de la probabilidad en la educación primaria*. Trabajo fin de máster, Universidad de Granada.
- Ruiz-Reyes, K., Batanero, C. y Contreras, J.M. (2017). Los tipos de lenguaje del muestreo en los textos escolares chilenos de Educación Secundaria. En J.M. Muñoz-Escolano, A. Arnal-Bailera, P. Beltrán-Pellicer, M.L. Callejo y J. Carrillo (Eds.), *Investigación en Educación Matemática XXI* (pp. 577). Zaragoza: SEIEM.
- Ruiz-Reyes, K., Begué, N., Batanero, C., y Contreras, J. M. (2017). Un estudio comparado de los contenidos de muestreo en la Educación Secundaria Obligatoria en Chile. *Educação Matemática Pesquisa*, 19(3), 67-83.
- Ruiz-Reyes, K., Ruz, F., Molina-Portillo, E. y Contreras, J. M. (2018). El muestreo en el currículo escolar chileno. En L. J. Rodríguez-Muñiz, L. Muñiz-Rodríguez, A. Aguilar-

- González, P. Alonso, F.J. García García, A. Bruno (Eds.), *Investigación en Educación Matemática XXII* (pp. 505). Gijón: SEIEM
- Saiz, O. y Blumenthal, V. (2018). *Matemáticas 3° Medio*. Santiago de Chile: Cal y Canto
- Saldanha, L. y Thompson, P. (2002) Conceptions of sample and their relationship to statistical inference. *Educational Studies in Mathematics*, 51, 257-270.
- Schuyten, G. (1991). Statistical thinking in psychology and education. En D. Vere-Jones (Ed.), *Proceeding of the Third International Conference on Teaching Statistics* (pp. 486-490). Voorburg, The Netherlands: International Statistical Institute.
- Sedlmeier, P. (1999). *Improving statistical reasoning. Theoretical models and practical implications*. Mahwah, NJ: Erlbaum.
- Serrano, L. (1993). *Aproximación frecuencial a la enseñanza de la probabilidad y conceptos elementales sobre procesos estocásticos: Un estudio de concepciones iniciales*. Memoria de Tercer Ciclo. Universidad de Granada.
- Serrano, L. (1996). *Significados institucionales y personales de objetos matemáticos ligados a la aproximación frecuencial de la probabilidad*. Tesis Doctoral. Universidad de Granada.
- Serrano, L., Batanero, C., Ortiz, J. J. y Cañizares, M. J. (1998). Heurísticas y sesgos en el razonamiento probabilístico de los estudiantes de secundaria. *Educación Matemática*, 10(1), 7-25
- Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions. En A. Grouws (Ed.). *Handbook of research in teaching and learning mathematics* (pp. 465-494). New York: McMillan.
- Shaughnessy, J. M., Canada, D., & Ciancetta, M. (2003, July). *Middle school students' thinking about variability in repeated trials: A cross-task comparison*. En N. A. Pateman, B. Dougherty y J. Zilloux (Eds.), *Proceedings of Joint Meeting of PME27 and PMENA XXV* (Vol. 4, pp. 159-166). Honolulu, HI: Program Committee.
- Shaughnessy, J. M., Canada, D. y Ciancetta, M. (2003). *Middle school students' thinking about variability in repeated trials: A cross-task comparison*. En N. A. Pateman, B. Dougherty y J. Zilloux (Eds.), *Proceedings of Joint Meeting of PME27 and PMENA XXV* (Vol. 4, pp. 159-166). Honolulu, HI: Program Committee.
- Shaughnessy, J.M., Ciancetta, M. y Canada, D. (2004). Types of student reasoning on sampling tasks. En M.J. Høines y A.B. Fuglestad (Eds.), *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education* (Vol.4, pp. 177-184). Bergen, Noruega: International Group for the Psychology of Mathematics Education.
- Shaughnessy, J. M., Garfield, J. y Greer, B. (1996). Data handling. En A. J. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International handbook of mathematics education*, (Vol.1, pp. 205–237). Dordrecht, The Netherlands: Kluwer.
- Shaughnessy, J. M., Watson, J., Moritz, J. y Reading, C. (1999). School mathematics students' acknowledgment of statistical variation. Trabajo presentado en el *NCTM*

- Research Preession Symposium: There's More to Life than Centers*. Paper 77th Annual NCTM Conference, San Francisco, California.
- Skemp, R. (1980). *Psicología del aprendizaje de las matemáticas*. Madrid: Morata
- Sthol, H., y Tarr, J. E. (2002). Developing notions of inference using probability simulation tools. *Journal of Mathematical Behavior*, 21, 319–337.
- Tauber, L. (2001) *La construcción del significado de la distribución normal a partir de la actividad de análisis de datos*. Tesis doctoral. Universidad de Granada.
- Triola, Mario F. (2004). *Estadística*. México: Pearson educación.
- Tversky, A. y Kahneman, D. (1974). Judgement under uncertainty: Heuristics and biases. *Science*, 185, 1124-1131.
- Tversky, A. y Kahneman, D. (1982). Judgments of and by representativeness. En D. Kahneman, P.Slovic y A. Tversky (Eds.), *Judgement under uncertainty: Heuristics and biases* (pp. 117-128). New York: Cambridge University Press
- Valdez, J. C. (2016). *Las grandes ideas de probabilidad en el razonamiento informal de estudiantes de bachillerato*. Tesis Doctoral. Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional. México.
- Vanhoof, S., Castro Sotos, A. E., Onghena, P. y Verschaffel, L. (2007). Students' reasoning about sampling distributions before and after the sampling distribution activity. *Proceedings of the 56th Session of the International Statistical Institute*. Lisboa: International Statistical Institute. Recuperado de: http://iase-web.org/documents/papers/isi56/CPM80_Vanhoof.pdf.
- Vera, O.D. (2016). *Comprensión de conceptos elementales del análisis de varianza por estudiantes universitarios*. Tesis Doctoral Universidad de Granada. [<http://hdl.handle.net/10481/41155>]
- Watson, J. M. (1997). Assessing statistical literacy using the media. En I. Gal & J. B. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 107–121). Amsterdam: IOS Press and the International Statistical Institute.
- Watson, J. (2004). Developing reasonig about samples. En D. Ben-Zvi y J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 277–294). Dordrechy: Kluwer.
- Watson, J.M. (2006). *Statistical literacy at school: Growth and goals*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Watson, J. y Kelly, B. (2005). Cognition and instruction: Reasoning about bias in sampling.
- Watson, J., Kelly, B., Callingham, R. y Shaughnessy, M. (2003) The measurement of school students' understanding of statistical variation, *International Journal of Mathematical Education in Science and Technology*, 34:1, 1-29, DOI: 10.1080/0020739021000018791
- Watson, J. M. y Moritz, J. B. (2000a). Developing concepts of sampling. *Journal for Research in Mathematics Education*, 31(1), 44-70.
- Watson, J. M. y Moritz, J. B. (2000b). Development of understanding of sampling for statistical literacy. *The Journal of Mathematical Behavior*, 19(1), 109-136.

- Well, A., Pollatsek, A. y Boyce, S. (1990). Understanding the effects of sample size on the variability of the mean. *Organizational Behavior and Human Decision Processes*, 47, 289-312.
- Wild, C. y Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67(3), 221-248.
- Wroughton, J. R., McGowan, H. M., Weiss, L. V., & Cope, T. M. (2013). Exploring the role of context in students' understanding of sampling. *Statistics Education Research Journal*, 12(2). [Online: [https://iase-web.org/documents/SERJ/SERJ12\(2\)_Wroughton.pdf](https://iase-web.org/documents/SERJ/SERJ12(2)_Wroughton.pdf)]
- Zapico, M. (2007). Interrogantes acerca de análisis de contenido y del discurso en los textos escolares. En MINEDUC (Ed.), *Primer Seminario Internacional de Textos Escolares (SITE 2006)* (pp. 149-155). Santiago: MINEDUC.
- Zieffler, A., Garfield, J., delMas, R., y Reading, C. (2008). A framework to support research on informal inferential reasoning. *Statistics Education Research Journal*, 7(2), 40-58. [Online: [https://iase-web.org/documents/SERJ/SERJ7\(2\)_Zieffler.pdf?1402525008](https://iase-web.org/documents/SERJ/SERJ7(2)_Zieffler.pdf?1402525008)]

