

PREDICCIÓN DE VALORES EN UNA TABLA DE FRECUENCIAS POR ESTUDIANTES CHILENOS

Predicting values in a frequency table by Chilean students

Pallauta, J. D.^a, Gea, M. M.^b, Valenzuela-Ruiz, S.^b y Batanero, C.^b

^aUniversidad de Los Lagos, ^bUniversidad de Granada

Resumen

La interpretación y predicción de datos son capacidades que tienen gran relevancia en la sociedad actual, especialmente para la toma de decisiones informadas. En este trabajo se analiza la capacidad de estudiantes chilenos de educación secundaria para predecir valores a partir de una tabla de distribución de una variable con frecuencias absolutas y relativas. Para ello, se aplicó una tarea a 103 estudiantes chilenos de 7º curso de educación básica (14 a 15 años) y 88 de 1º curso de educación media (15 a 16 años) en la que debían predecir los posibles resultados de un experimento aleatorio equiprobable repetido un número determinado de veces. Las respuestas fueron evaluadas y analizadas según los niveles de lectura. Los resultados mostraron que la mayoría de los estudiantes alcanzó niveles básicos de lectura, evidenciando dificultades en la resolución de este tipo de tareas.

Palabras clave: educación secundaria, niveles de lectura, tablas estadísticas.

Abstract

Data interpretation and prediction are skills that are highly relevant in today's society, especially for informed decision making. In this paper we analyse the ability of Chilean secondary school students to predict values from a distribution table of a variable with absolute and relative frequencies. For this purpose, a task was given to 103 Chilean students in the 7th grade of basic education (14-15 years old) and 88 students in the 1st grade of secondary education (15-16 years old), in which they had to predict the possible outcomes of an equiprobable random experiment repeated a given number of times. The responses were assessed and analysed according to the reading levels. The results showed that most of the students reached basic reading levels, evidencing difficulties in solving this type of task.

Keywords: reading levels, secondary education, statistical tables.

INTRODUCCIÓN

En la actualidad, los ciudadanos tienen acceso a una amplia variedad de datos sobre ámbitos sociales, políticos o económicos, que es difundida a través de los medios de comunicación (p. ej. prensa, televisión e Internet). Esta abundancia de información demanda que la ciudadanía cuente con las capacidades de interpretar, analizar y realizar predicciones fundamentadas de los datos a partir de valores conocidos, para tomar decisiones adecuadas en diferentes ámbitos (Gal, 2022; Kuntze et al., 2015).

El currículo chileno de educación secundaria (MINEDUC, 2015) propone el estudio de tablas estadísticas como un objeto de aprendizaje por sí mismo y también como una herramienta para tratar otros temas, por ejemplo, el análisis de variables bidimensionales, o los conceptos de probabilidad compuesta y condicionada. Aunque las tablas estadísticas están presentes en el currículo, los estudios sobre su interpretación en educación secundaria son escasos, ya que la mayoría se enfoca en niños o en la interpretación de tablas de doble entrada (Gabucio et al., 2010). A pesar, de su uso frecuente para registrar y organizar los datos, y su amplia presencia en los libros de texto de secundaria (Pallauta

et al., 2021a; 2021b), los profesores desconocen los procesos cognitivos implicados en su interpretación (Friel et al., 2001). En este sentido, es relevante que los estudiantes de secundaria comprendan los diferentes tipos de tablas, su estructura y las relaciones entre los datos. Dicha comprensión favorecerá tanto su lectura e interpretación adecuada (Pallauta et al., 2023), como identificar tendencias, establecer posibles correlaciones y reconocer situaciones en que los datos podrían estar siendo utilizados de manera inapropiada o engañosa (OECD, 2021).

El objetivo de este trabajo es analizar la capacidad de predicción de valores a partir de una tabla de distribución de una variable con frecuencias absolutas y relativas de estudiantes chilenos de dos cursos diferentes de educación secundaria.

FUNDAMENTOS

Niveles de lectura de tablas y gráficos

Para caracterizar la complejidad que implica la interpretación de tablas y gráficos estadísticos, diversos autores han propuesto niveles de lectura de los mismos. En este trabajo nos centraremos en la jerarquía de Curcio (1989) que consta de tres niveles, ampliada posteriormente por Friel et al. (2001), y que ha sido utilizada en trabajos previos con gráficos estadísticos (p. ej. Díaz-Levicoy, 2018; Gea et al., 2017) y tablas estadísticas (Pallauta et al., 2021). Además, se incorpora un nivel N0 propuesto por Díaz-Levicoy (2018):

- *N0. No leen los datos.* Este nivel considera la ausencia de respuesta o aquellas que carecen de sentido de acuerdo a la pregunta planteada.
- *N1. Leer los datos.* Este nivel se centra en la interpretación de la estructura y los elementos de las tablas estadísticas. Se alcanza respondiendo a preguntas concretas referidas a dicha estructura y a los datos presentados, como indicar la frecuencia de una determinada modalidad o el valor de la variable representada en una tabla. No requiere realizar comparaciones o cálculos con los datos de la tabla.
- *N2. Leer dentro de los datos.* Este nivel involucra establecer relaciones o tendencias de la información presentada en la tabla, a partir de la comparación o la realización de cálculos con los datos. Por ejemplo, determinar la moda de una variable, a partir de los datos de una tabla de frecuencias.
- *N3. Leer más allá de los datos.* Este nivel se centra en extrapolar, predecir o realizar inferencias a partir de la información proporcionada por la tabla. Por ejemplo, interpolar o extrapolar un valor que no se encuentra explícito dentro de una tabla de datos que representa la variación de una variable a lo largo de un periodo de tiempo.

Antecedentes

Algunas investigaciones han evidenciado que los estudiantes tienen dificultades al interpretar datos y realizar predicciones a partir de tablas estadísticas (p. ej. Estepa et al., 1999; Prodromou, 2015). En este sentido, los trabajos que han utilizado los niveles de lectura propuestos por Curcio (1989) o Friel et al. (2001) para analizar las interpretaciones de tablas estadísticas de los estudiantes de educación secundaria, se han centrado en las tablas de doble entrada o se han desarrollado con una muestra muy reducida, limitando la generalización de sus hallazgos. Así Gabucio et al. (2010) analizaron las respuestas de 200 estudiantes españoles, de 5º y 6º de primaria y 1º y 2º de ESO (10 a 14 años) a un cuestionario compuesto de ítems de opción múltiple relacionados con una tabla de doble entrada. Los resultados indicaron dificultad en alcanzar los niveles más altos de lectura y que no hubo incremento de las respuestas correctas conforme se avanzaba de curso. González et al. (2021) analizaron la interpretación de tablas de frecuencias con datos agrupados en intervalos de 10 estudiantes chilenos de 4º de secundaria (15 a 16 años), a partir de una tabla de distribución de una variable en el contexto de la COVID-19. Obtuvieron buenos resultados en el nivel N1, leer los datos,

(Curcio, 1989), mientras que el nivel N3, leer más allá de los datos, fue más complejo, pues los estudiantes se basaron en sus opiniones personales para responder a las preguntas, en lugar de utilizar la información representada en la tabla.

Pallauta et al. (2023) diseñaron y validaron un cuestionario con 128 estudiantes españoles de 3º de ESO (14-16 años), del cual se toma una de las tareas empleadas en este estudio. En la tarea propuesta en este trabajo, el porcentaje de respuestas correctas fue el 52%, mientras el 55,5 % de la muestra se situó en el nivel N2 (leer dentro de los datos) y el 40,6% no alcanza el nivel N1 (Curcio, 1989).

En este trabajo, se analiza la capacidad de predicción de valores a partir de una tabla de distribución de una variable con frecuencias absolutas y relativas basados en los niveles de lectura, con una muestra nueva, de mayor tamaño que incluye estudiantes chilenos de secundaria de dos grupos de edad, a diferencia de Pallauta et al. (2023).

METODOLOGÍA

Muestra y tarea planteada

La muestra fue intencional y se compuso de 191 estudiantes chilenos, 103 de 7º curso de educación básica (EB) (14 a 15 años) y 88 de 1º curso de educación media (EM) (15 a 16 años), que estudiaban en cinco centros educativos de diferentes provincias del país: 1 público, 3 concertados y 1 privado. La distribución de la muestra por género, curso y tipo de centro educativo, se presenta en la Tabla 1.

Tabla 1.

Distribución por porcentaje de la muestra según género, curso y centro educativo

Tipo de centro educativo	7º EB N=103		1º EM N=88	
	Hombre	Mujer	Hombre	Mujer
Público	7,8	16,5		
Concertado	25,2	27,2	29,5	43,2
Privado	14,6	8,7	17,1	10,2

A los participantes se les propuso resolver, de manera individual, una tarea con dos apartados (Figura 1), tomada de Pallauta et al. (2023) quienes la adaptaron de un libro de texto chileno de 7º EB (Merino et al., 2016). La tarea muestra en una tabla la distribución de una variable (número de una ficha escogida al azar entre 10 numeradas de 1 al 10) con parte de las frecuencias absolutas y relativas obtenidas al realizar 1000 repeticiones de un experimento aleatorio, con resultados equiprobables. Sin embargo, se omitieron los valores de las últimas dos columnas de la tabla, que los estudiantes deben completar a partir de los datos disponibles.

En el primer apartado se pide indicar la frecuencia absoluta del valor 10 en las 1000 repeticiones, para ello, el estudiante debe comprender que, al ser los casos equiprobables, las frecuencias absolutas de los valores 9 y 10 serán similares a los del resto de las columnas y entre todos ellos han de sumar un total de 1000, por tanto, existen varias opciones de respuesta correcta. En el segundo apartado, que sólo se propuso a los estudiantes mayores de 1º EM, se requiere generalizar los datos al incrementar a 10000 las simulaciones del experimento. Puesto que se deben predecir valores que no pueden obtenerse directamente de la tabla, se requiere un nivel de lectura N3 en la categoría de Curcio (1989). Se debe aplicar una proporcionalidad de las frecuencias, pero no debe ser exacta, para reflejar la variabilidad aleatoria que se espera en el experimento.

Figura 1.

Tarea propuesta a los estudiantes

En la siguiente tabla se registraron los resultados al extraer 1.000 veces una ficha desde una urna con 10 fichas numeradas desde el 1 al 10, pero olvidaron completar las dos últimas columnas.

1000 extracciones										
Número de la ficha obtenida										
Número	1	2	3	4	5	6	7	8	9	10
Frecuencia absoluta	98	106	96	97	98	110	102	98		
Frecuencia relativa	0,098	0,106	0,096	0,097	0,098	0,110	0,102	0,098		

Responde

a) ¿Cuántas veces piensas que salió el 10? Justifica tu respuesta.

b) Si se aumentan las extracciones a 10.000 ¿cuántas veces piensas que saldría aproximadamente cada ficha? Justifica tu respuesta.

Categorías de análisis

Recogidas las respuestas escritas individuales de cada estudiante, se utilizó un análisis de contenido (Mayring, 2021) para evaluar las respuestas e identificar el nivel de lectura alcanzado por el estudiante. Este método, típico en estudios cualitativos, permite analizar variables subyacentes en un texto, para extraer información que se pueda tratar estadísticamente. Para asegurar la fiabilidad de la codificación, el 15% de las respuestas fueron analizadas de manera independiente por otro investigador y, en caso de desacuerdo, se discutieron con los demás autores hasta llegar a un acuerdo. A continuación, se describen las categorías utilizadas para estas variables, incluyendo ejemplos de respuestas de los estudiantes que han sido codificados como Ex.

Respuesta correcta. En el primer apartado, se considera correcta cuando se señala un valor alrededor de 100, dado que el total de extracciones son 1000, como indica E66, por tanto, las frecuencias relativas son próximas a 0,1. En el segundo apartado, la respuesta correcta consiste en una estimación basada en 10 veces los valores expuestos en la tabla, usando razonamiento proporcional e indicando cierta variabilidad en los datos. Por ejemplo, E40 indica que cada ficha saldrá alrededor de 1000 por tratarse de un experimento equiprobable.

E66: Yo creo que salió 98 veces o 97 veces, porque dividí en dos 195 que es lo que falta para completar 1.000 extracciones, y los valores que faltan deben ser enteros y parecidos a los otros (Apartado 1).

E40: Cada ficha saldrá aproximadamente unas 1000 veces, porque al ser al azar cada ficha tiene la misma probabilidad de salir, lo que hace que salgan una cantidad similar de veces (Apartado 2).

Respuesta parcialmente correcta. En el primer apartado serían las respuestas que toman en cuenta el total de extracciones realizadas, sin considerar la tendencia de los datos, como E10, o solo toman en cuenta la tendencia de los datos, desatendiendo la información sobre el total de las extracciones. En el segundo apartado cuando se indica la cantidad de veces que salió alguna ficha, particular, cuya cantidad es producto de la multiplicación del valor de la frecuencia absoluta por 10, no aportando razonamiento sobre la variabilidad, como E5 quien sólo registra la cantidad de veces que podría salir la ficha 1.

E10: Entre 0 y 195, porque no puede haber salido más veces (Apartado 1).

E5: Si el 1 salió 98 veces, saldrá 980 (Apartado 2).

Respuesta incorrecta. En el primer apartado corresponden a las respuestas con valores desajustados a la tendencia de los datos expuestos, además, generalmente, no justifican como E35. En el segundo apartado, las respuestas de este tipo se aferran al carácter aleatorio de la experiencia de extraer una determinada ficha de una urna, por ejemplo, E3 indica que no se puede predecir el resultado.

E35: 6 veces (Apartado 1).

- E3: No se puede saber, ya que a lo mejor el 1 sale 10 veces y el 9, 200 veces, porque al ser una urna no sigue ningún código (Apartado 2).

Adicionalmente, la interpretación tanto de los cálculos realizados como de las justificaciones aportadas en las respuestas permite caracterizar e identificar el nivel de lectura (Curcio, 1989) alcanzado por los estudiantes, que se describe a continuación.

N0. No leen los datos. Este nivel incluye las respuestas incorrectas como E52 en el primer apartado o E18 en el segundo.

- E52: Pienso que son 104, porque a la gente le gustan más los números altos (Apartado 1).

- E18: No lo sé, porque tienes la misma probabilidad de sacar un 3 que un 5 (Apartado 2).

N1. Leer los datos. La respuesta evidencia una lectura literal de los datos presentados en la tabla. Por ejemplo, en el primer apartado algunos estudiantes, como E7, justificaron su respuesta en la frecuencia del valor 6 de la variable, al considerar que 110 es múltiplo de 10. En el segundo, E139 indica los mismos valores presentados en la tabla cuando se realizan 1000 extracciones, añadiendo las veces que podrían haber salido las fichas 9 y 10.

- E7: Una sola vez ya que 110 es múltiplo de 10 (Apartado 1, Resp. Incorrecta).

- E139: 1: 98; 2: 106; 3: 96; 4: 97; 5: 98; 6: 110; 7: 102; 9: 98; 9: 100; 10: 193 (Apartado 2, Resp. Incorrecta).

N2. Leer dentro de los datos. En este nivel se denota una correcta lectura y la realización de cálculos o comparación de los datos como la respuesta de E101. Sin embargo, en el primer apartado, se ignora que la suma de las frecuencias absolutas debe totalizar 1000. En el segundo apartado, se presentan respuestas en que el estudiante identifica que se obtendrá la frecuencia para cada ficha multiplicando el valor de la frecuencia absoluta por 10, para cada valor de la variable, como lo indica E148, pero no se tiene en cuenta la variabilidad aleatoria.

- E101: Yo creo que salió 98 veces, ya que varias veces salió ese valor (Apartado 1, Resp. Correcta).

- E148: $\times 10$, porque se agrega un 0 (Apartado 2, Resp. Parcialmente correcta).

N3. Leer más allá de los datos. En el primer apartado se considera la tendencia de los datos y, además, que la suma de las frecuencias absolutas es 1000, como indica E100. En el segundo apartado, alcanzan este nivel de lectura las respuestas que evidencian una relación de los datos al incrementar a 10000 las extracciones de la urna, como E30, quien señala que cada ficha se obtendrá alrededor de 1000 veces.

- E100: Yo pienso que el 10 salió una cantidad de veces cercana al 100, ya todos los demás lo hicieron un poco más o un poco menos; además todas frecuencias absolutas deben sumar 1000. (Apartado 1, Resp. Correcta).

- E30: Aproximadamente 1000, porque si divido 10000 entre 10 pues sale 1000 (Apartado 2, Resp. Correcta).

RESULTADOS

Los resultados consideran la evaluación de las respuestas junto al nivel de lectura alcanzado, seguido por el índice de dificultad de los dos apartados de la tarea.

En la Tabla 2 se presentan los resultados de la evaluación de las respuestas junto al nivel de lectura de la tarea planteada, recordando que los alumnos de 7º EB solo tuvieron que responder al primer apartado. Un porcentaje importante de estudiantes de 1º EM deja ambos apartados sin contestar, especialmente el segundo apartado donde menos de la mitad lo responde, evidenciando que la tarea fue difícil para los estudiantes, a pesar de estar adaptada de un libro de 7º EB. Las respuestas correctas fueron pocas en el primer apartado en ambos cursos, particularmente en 7º EB, y aumentan algo en

el segundo apartado. Por lo tanto, fue más sencillo para los estudiantes de 1° EM reconocer que al aumentar el número de simulaciones hasta 10000, el número de veces que saldría cada ficha sería aproximadamente 10 veces a los valores obtenidos en 1000 simulaciones. Las respuestas parcialmente correctas son más frecuentes en el primer apartado y en los alumnos de 7° EB, así como las incorrectas.

Tabla 2*Porcentaje de evaluación y nivel de lectura de las respuestas en la Tarea*

	Apartado 1								Apartado 2			
	7° EB				1° EM				1° EM			
	N = 103				N = 88				N = 88			
	N0	N1	N2	N3	N0	N1	N2	N3	N0	N1	N2	N3
Correcto			5,8	10,7		1,1	4,5	18,2		1,1	6,8	22,7
Parcialmente correcto		19,4	21,4			5,7	20,5			1,1	5,7	1,1
Incorrecto	12,6	1,9			3,4	2,3			2,3	6,8		
No responde	28,2				44,3				52,3			

En relación a los niveles de lectura alcanzados por los estudiantes en la tarea, se observa en la tabla 2 que en el primer apartado lo más frecuente en los dos grupos fue alcanzar el nivel N2 (leer dentro de los datos), especialmente en 7° EB, seguido en este curso por el nivel N1 (leer los datos) y N3 (leer más allá de los datos). En el caso de 1° EM lo más frecuente después del nivel N2 es el N3 siendo la mitad el porcentaje de los que quedan en nivel N1. Hay también menos respuestas de estudiantes de este curso que de 7° EB que quedan en nivel N0. En el segundo apartado, planteado únicamente 1° EM, es mayor la proporción de estudiantes que alcanzan el nivel máximo N3 previsto en la tarea, y por tanto disminuyen los porcentajes en los restantes niveles. La asociación entre el nivel de lectura y el curso en el primer apartado no fue estadísticamente significativa en la prueba de Chi-cuadrado ($\chi^2 = 7,045$, g.l = 3, $p = 0,70$).

Por otra parte, en la Tabla 3 se presenta el valor del índice de dificultad de esta tarea, que fue calculado asignando una puntuación igual a 1 a las respuestas correctas o parcialmente correctas y 0 en otro caso, junto a su desviación típica. Se observa en el primer apartado una dificultad moderada, siendo más complejo para los estudiantes de 1° EM, aunque el análisis de las diferencias de las medias de aciertos por curso en este apartado no mostró una diferencia estadísticamente significativa ($t = -1,004$; $p = 0,317$). En el segundo apartado la dificultad se incrementa respecto al anterior en los estudiantes de 1° EM.

Tabla 3*Índice de dificultad de la Tarea*

	Apartado 1		Apartado 2
	7° EB N = 103	1° EM N = 88	1° EM N = 88
Media	0,57	0,50	0,39
Desviación Típica	0,497	0,503	0,490

DISCUSIÓN Y CONCLUSIONES

Este trabajo tuvo como objetivo analizar la capacidad de predicción de valores a partir de una tabla de distribución de una variable con frecuencias absolutas y relativas de estudiantes chilenos de educación secundaria, aportando nuevos resultados sobre el tema en el contexto de la educación secundaria.

Respecto a las respuestas de la tarea, el índice de dificultad fue menor para la muestra considerada en este estudio a diferencia de Pallauta et al. (2023). En el primer apartado, a nivel general ambos grupos tienen resultados similares, sin embargo, se destaca que la mayor parte de los estudiantes de 1° EM no responde, coincidiendo con Pallauta et al. (2023). Además, las respuestas correctas son ligeramente más frecuentes en este curso, mientras que las respuestas parcialmente correctas fueron más frecuentes en 7° EB. En el segundo apartado, planteado solo en 1° EM, se incrementó el porcentaje de estudiantes tanto de los que no responden como los que lo hacen correctamente.

Por otra parte, en este trabajo las respuestas de los estudiantes fueron categorizadas de acuerdo con los diferentes niveles de lectura propuestos por Curcio (1989), destacando la necesidad de incorporar el nivel N0 que evidencia una ausencia de un mínimo de lectura (Díaz-Levicoy, 2018; Díaz-Levicoy et al, 2015). En el primer apartado, la mayoría de las respuestas alcanza el nivel N2 (leer dentro de los datos) especialmente en 7° EB, mientras que el nivel máximo N3 es alcanzado mayoritariamente por los estudiantes de 1° EM.

En el segundo apartado, la mayor parte de los estudiantes de 1° EM alcanzan el nivel N3, aunque más de la mitad no respondió la tarea. De este modo se evidencia, al igual que otros estudios, que es complejo alcanzar niveles más elevados de lectura. También en Gabucio et al. (2010), los ítems de nivel N3 fueron los más complejos para los estudiantes, algo similar ocurrió con González et al. (2021) en que las interpretaciones de los estudiantes se centraron en los niveles básicos de lectura (N1 o N2). Esto se podría explicar porque los currículums de secundaria presentan una limitada incorporación del análisis crítico de la información estadística en distintos contextos, lo que podría justificar esta deficiencia (Callingham y Watson, 2017). Asimismo, los análisis de libros de texto chilenos de educación secundaria evidencian una alta presencia de tareas sobre tablas estadísticas centradas, principalmente, en los niveles más básicos de lectura (Pallauta et al. 2021a; 2021b).

Una implicación del estudio para la enseñanza del tema radica en la importancia de proponer a los estudiantes tareas que promuevan su capacidad de predecir datos a partir de información parcial presentada en tablas estadísticas (Sharma, 2013), que es parte de la cultura estadística (Gal, 2022; Kuntze et al., 2015). Igualmente, se resalta la necesidad de integrar en la enseñanza situaciones que fomenten el uso de justificaciones fundamentadas en los datos, especialmente en contextos de incertidumbre (Batanero y Borovcnik, 2016).

Para finalizar, reconocemos las limitaciones del estudio, al haber utilizado una única tarea. Por ello, se abre una línea de investigación centrada en el análisis de la dificultad para los estudiantes en otras tareas de predicción a partir de tablas estadísticas, y el nivel de lectura que alcanzan en dichas tareas.

Agradecimientos

Proyecto PID2022-139748NB-100 financiado por MICIU/AEI/10.13039/501100011033/ y por FEDER, UE.

Referencias

- Batanero, C. y Borovcnik, M. (2016). *Statistics and probability in high school*. Sense Publishers.
- Callingham, R. y Watson, J. M. (2017). The development of statistical literacy at school. *Statistics Education Research Journal*, 16(1), 181-201. <https://doi.org/10.52041/serj.v16i1.223>

- Curcio, F. R. (1989). *Developing graph comprehension elementary and middle school activities*. National Council of Teachers of Mathematics.
- Díaz-Levicoy, D. (2018). *Comprensión de gráficos estadísticos por alumnos chilenos de educación primaria*. [Tesis doctoral]. Universidad de Granada.
- Díaz-Levicoy, D., Arteaga, P. y Batanero, C. (2015). Gráficos estadísticos y niveles de lectura propuestos en textos chilenos de Educación Primaria. En C. Fernández, M. Molina y N. Planas (eds.), *Investigación en Educación Matemática XIX* (pp. 229-238). SEIEM.
- Estepa, A., Batanero, C. y Sanchez, F. T. (1999). Students' intuitive strategies in judging association when comparing two samples. *Hiroshima Journal of Mathematics Education*, 7, 17-30.
- Friel, S. N., Curcio, F. R. y Bright, G. W. (2001). Making sense of graphs: critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education*, 32(2), 1-24. <https://doi.org/10.2307/749671>
- Gal, I. (2022). Critical understanding of civic statistics: Engaging with important contexts, texts, and opinion questions. En J. Ridgway (Ed.), *Statistics for Empowerment and Social Engagement* (pp. 323-343). Springer. https://doi.org/10.1007/978-3-031-20748-8_13
- Gabucio, F., Martí, E., Enfedaque, J., Gilabert, S. y Konstantinidou, A. (2010). Niveles de comprensión de las tablas en alumnos de primaria y secundaria. *Cultura y Educación*, 22(2), 183-197. <https://doi.org/10.1174/113564010791304528>
- Gea, M. M., Arteaga, P. y Cañadas, G. R. (2017). Interpretación de gráficos estadísticos por futuros profesores de Educación Secundaria. *Avances de Investigación en Educación Matemática*, 12, 19-37. <https://doi.org/10.35763/aiem.v1i2.189>
- González, C. P., Muñoz, R. y Muñoz, J. (2021). Características argumentativas de la interpretación de tablas de frecuencia en estudiantes chilenos de segundo año medio. *Revista Chilena de Educación Matemática*, 13(1), 17-29. <https://doi.org/10.46219/rechiem.v13i1.39>
- Kuntze, S., Martignon, L., Vargas, F. y Engel, J. (2015). Competencies in understanding statistical information in primary and secondary school levels: An inter-cultural empirical study with German and Colombian students. *Avances de Investigación en Educación Matemática*, 7, 5-25. <https://doi.org/10.35763/aiem.v1i7.103>
- Mayring, P. (2021). *Qualitative content analysis: A step-by-step guide*. Sage.
- MINEDUC. (2015). *Bases curriculares 7° básico a 2° medio*. Unidad de Currículum y Evaluación.
- Merino, R., Muñoz, V., Pérez, B. y Rupin, P. (2016). *Texto del estudiante Matemática 7° básico*. SM.
- OECD. (2021). *21st-century readers: Developing literacy skills in a digital world*. OECD. <https://doi.org/10.1787/a83d84cb-en>
- Pallauta, J. D., Batanero, C. y Gea, M. M. (2023). Un instrumento para evaluar la comprensión de tablas estadísticas en educación secundaria. *Enseñanza de las Ciencias*. 41(3), 89-112. <https://doi.org/10.5565/rev/ensciencias.5926>
- Pallauta, J. D., Batanero, C., Gea, M. M. y Arteaga, P. (2021a). Niveles de lectura y contextos en las actividades sobre tablas estadísticas en libros de texto chilenos y españoles. *Revista Chilena de Educación Matemática*, 13(4), 119-133. <https://doi.org/10.46219/rechiem.v13i4.74>
- Pallauta, J. D., Gea, M. M., Batanero, C. y Arteaga, P. (2021b). Complejidad semiótica de tablas estadísticas y su distribución en textos chilenos de educación básica. En P. D. Diago, D. F. Yáñez, M. T. González-Astudillo y D. Carrillo (Eds.), *Investigación en Educación Matemática XXIV* (pp. 481-488). SEIEM.
- Prodromou, T. (2015). Students' emerging reasoning about data tables of large-scale data. *International Journal of Statistics and Probability*, 4(3), 181-197. <https://doi.org/10.5539/ijsp.v4n3p181>
- Sharma, S. (2013). Assessing students' understanding of tables and graphs: Implications for teaching and research. *International Journal of Educational Research and Technology*, 4(4), 61-69.