UNIVERSIDAD DE QUINTANA ROO
DIVISIÓN DE CIENCIAS E INGENIERÍA

Una experiencia de aprendizaje con estudiantes de Telebachillerato para la comprensión de las medidas de tendencia central y de variabilidad

TESIS
Para obtener el grado de

MAESTRO EN ENSEÑANZA DE LAS MATEMÁTICAS

PRESENTA
JUAN MANUEL ESCOBEDO HERNÁNDEZ

DIRECTOR
DR. JOSÉ ANTONIO ORTA AMARO

CO-DIRECTOR
DR. CÉSAR CRISTÓBAL ESCALANTE

ASESORES
DR. LUIS FERNANDO CABRERA CASTELLANOS
MTI. MELISSA BLANQUETO ESTRADA
MEM. JAVIER GARCÍA GUZMÁN

CONACYT
Consejo Nacional de Ciencia y Tecnología

Chetumal, Quintana Roo, México, Mayo de 2019

UNIVERSIDAD DE QUINTANA ROO
DIVISIÓN DE CIENCIAS E INGENIERÍA

TRABAJO DE TESIS BAJO LA SUPERVISIÓN DEL COMITÉ DEL PROGRAMA DE MAESTRÍA Y APROBADA COMO REQUISITO PARA OBTENER EL GRADO DE:

MAESTRO EN ENSEÑANZA DE LAS MATEMÁTICAS

COMITÉ DE TESIS

DIRECTOR: ______________________
DR. JOSÉ ANTONIO ORTA AMARO

CO-DIRECTOR: ______________________
DR. CÉSAR CRISTÓBAL ESCALANTE

ASESOR: ______________________
DR. LUIS FERNANDO CABRERA CASTELLANOS

ASESOR: ______________________
MTI. MELISSA BLANQUETO ESTRADA

ASESOR: Javier García Guzmán
MEM. JAVIER GARCÍA GUZMÁN

Chetumal, Quintana Roo, México, Mayo de 2019.

Agradezco enormemente al consejo Nacional de Ciencia y Tecnología (CONACYT), por haberme otorgado la beca para la realización de los estudios de Maestría.

Un gran descubrimiento resuelve un gran problema, pero en la solución de todo problema, hay un cierto descubrimiento. El problema que se plantea puede ser modesto; pero, si pone a prueba la curiosidad que induce a poner en juego las facultades inventivas, si se resuelve por propios medios, se puede experimentar el encanto del descubrimiento y el goce del triunfo. Experiencias de este tipo, a una edad conveniente, pueden determinar una afición para el trabajo intelectual e imprimirle una huella imperecedera en la mente y en el carácter.

George Pólya

La vida es buena solamente por dos cosas, descubrir matemáticas y enseñar las matemáticas.

Simeon Poisson.

El pensamiento estadístico será algún día tan necesario para el ciudadano competente como la habilidad de leer y escribir.

Herbert George Wells

Si privas a los alumnos de tener la oportunidad de participar en esta actividad de proponer problemas, hacer sus propias conjeturas y descubrimientos, de estar equivocados, de estar creativamente frustrados, de tener una inspiración, y de improvisar sus propias explicaciones y demostraciones, les estás privando de las matemáticas en sí mismas.

Paul Lockart

AGRADECIMIENTOS

A Dios. Por permitirme culminar este proyecto y obtener un logro más en mi vida.

A la Dra. Silvia Azucena Mayén Galicia. No tengo palabras para agradecer el que haya confiado en mí y aceptara este proyecto, le agradezco por su paciencia ante mi inconsistencia, por todo lo que me ha enseñado y que he aprendido de usted, por el tiempo invertido, por sus consejos, en general por planear, dirigir y mantener el desarrollo y conclusión de este trabajo.

Al Dr. José Antonio Orta Amaro. Por haber aceptado este proyecto y tenerme paciencia. Me encuentro en deuda con usted; gracias por la confianza, consejos, observaciones, por su tiempo y el ánimo infundido para concluir este trabajo.

Al Dr. Cesar Cristóbal Escalante. Muchas gracias por todas sus enseñanzas, siempre que platico con usted aprendo cosas nuevas, gracias por todo el apoyo que me ha dado, por guiarme, por su tiempo, por sus acertadas observaciones y por haberme puesto en el camino a la Dra. Mayén.

A mis asesores. Dr. Luis Fernando Cabrera Castellanos, MTI. Melissa Blanqueto Estrada y MEM. Javier García Guzmán, por la atenta lectura y por su participación en la revisión de este trabajo de tesis.

A mi madre Margarita Hernández. Que siempre está apoyándome en todo, gracias por tenerme paciencia y preocuparse siempre por mí y mis hermanos, y por estar siempre a nuestro lado.

A mi hermana, mis hermanos y sobrinos. Por el apoyo brindado, comprensión y cariño.

A toda mi familia. Que de cierta forma siempre están pendientes de mí, dándome fuerzas y ánimos para seguir adelante.

A mis profesores de la maestría. En especial a la *Dra. Verónica Vargas* y al *Mtro. Ezequiel Hernández*, por todo lo que he aprendido de ustedes, por siempre guiarme y aclarar mis dudas, porque gracias a lo que aprendí de ustedes ahora me encuentro cosechando muchos éxitos.

A mis compañeros de trabajo. *MC. Mario Domínguez* y *Lic. Priscila Hernández*, gracias por el apoyo y los ánimos infundidos en este proceso.

A mis alumnos del Telebachillerato Comunitario Tres Garantías. De forma especial, a la generación 2014-2017. Gracias por ser parte de este proyecto, y permitirme trabajar con ustedes para que este trabajo fuera posible, gracias por la paciencia y el esfuerzo que pusieron al dar lo mejor de ustedes.

Al Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional (CINVESTAV-IPN). Por permitirme realizar mis estancias de investigación y haberme brindado todas las facilidades durante mi permanencia en esa institución.

A la Universidad de Quintana Roo. Especialmente a la División de Ciencias e Ingenierías, muchas gracias por darme la oportunidad de estudiar esta maestría y por los apoyos otorgados para asistir a mis estancias de investigación al CINVESTAV-IPN y al Congreso Internacional de Enseñanza de la Probabilidad y Estadística.

DEDICATORIA

A la memoria de mi abuelo Aristeo Hernández, sé que desde donde está se encuentra orgulloso de este nuevo logro y que aunque ya no está con nosotros sigue siendo pieza fundamental de nuestra vida.

RESUMEN

El objetivo de este trabajo fue implementar, en un ambiente con tecnología audiovisual, tareas de medidas de tendencia central y de variabilidad dirigidas a estudiantes de Telebachillerato en un marco de experiencia de aprendizaje que permita conocer sus dificultades y errores cuando se les pide realizar actividades relacionadas con estos conceptos, en que es necesario reconocer su cálculo y su interpretación. Para esto, se analizaron las respuestas del antes y después de aplicar la experiencia de aprendizaje a estudiantes de quinto semestre de un Telebachillerato Comunitario (17-30 años). Los problemas que se eligieron fueron tomados de Mayén (2009) y Orta (2014); estos están orientados a evaluar el concepto de la media aritmética, distribución, media ponderada, cálculos de la media y mediana a partir de gráficos, y finalmente analizar la variabilidad. Las respuestas que dieron los estudiantes se categorizaron y fueron analizadas con el enfoque ontosemiótico de la cognición matemática (EOS) propuesto por Godino y colaboradores (Godino y Batanero, 1994; Godino, 2002; Godino, 2003, Godino, Contreras y Font, 2006, Godino, Batanero y Font, 2007), con esto se pudieron constatar algunas dificultades y errores que mostraron los estudiantes con relación a las medidas de tendencia central y variabilidad. Este trabajo se desarrolló usando la metodología de los experimentos de enseñanza (Steffe y Thompson, 2000), en la que Cobb y Gravemeijer (2008) señalan tres fases: *preparación* del experimento, *experimentación* para promover el aprendizaje y ejecución del *análisis* retrospectivo de los datos.

Los resultados muestran que antes de la experiencia de aprendizaje los estudiantes contaban con conocimientos básicos y deficientes respecto a las medidas de tendencia central y de variabilidad, esto fue posible observar a partir del lenguaje que usaban en respuestas, pues no justificaban algorítmicamente sus resultados, tenían confusiones terminológicas, es decir confundían el concepto de media, mediana y moda. Después de la experiencia de aprendizaje, se observó un cambio en sus respuestas ahora ya las justifican con algoritmos y procedimientos, además incluyen argumentos deductivos, fueron capaces de explicar el significado de la media, tres alumnos aun después de la instrucción no

conciben la idea de la media desde el lenguaje terminológico, pues la siguen confundiendo con la moda. También, persisten algunas dificultades al interpretar los gráficos, un aporte importante de este trabajo es que el alumno además de confundir el valor de la variable con la escala, toma en cuenta el cero como valor y con estos datos calcula la mediana, el error que comete es la lectura incorrecta del gráfico, respecto a la construcción de gráficos solo dos alumnos fueron capaces de construir de forma correcta los gráficos, sin embargo 12 de 17 estudiantes fueron capaces de identificar la dispersión de un conjunto de datos y se dedujo que de forma implícita usan los conceptos de rango y desviación estándar, esto observado a partir de los argumentos que daban.

INDICE

ÍNDICE DE FIGURAS

ÍNDICE DE TABLAS

CAPÍTULO 1
EL PROBLEMA Y FUNDAMENTOS

1.1. Introducción

La presente investigación se centra en el análisis de los conceptos de medidas de tendencia central y de variabilidad y su comprensión en estudiantes de Telebachillerato Comunitario cuando resuelven tareas en un marco de experiencias de aprendizaje y en un ambiente con tecnología audiovisual, los cuales se incluyen en el Programa de Probabilidad y Estadística del Bachillerato. En este primer capítulo se expone el planteamiento del problema y la justificación donde se describe la importancia del aprendizaje de los conceptos de medidas de tendencia central y de variabilidad; así también, se describen los objetivos de la investigación y el marco curricular sobre el que se fundamenta este trabajo.

1.2. Problema

El ser humano en sus diferentes actividades de la vida cotidiana se encuentra influenciado por la estadística, por ejemplo: al leer un periódico necesita de conocimientos básicos de estadística para entender el significado de las tablas de datos y las gráficas, en encuestas de diferentes tipos, noticieros en televisoras, en los procesos económicos, entre otros. Tanto en el ambiente internacional, así como en el mexicano persiste la insistencia de fomentar la educación estadística en los distintos niveles escolares. La creciente inclusión de la estadística en los diferentes programas de estudio considera dos aspectos importantes: el carácter multidisciplinar que guarda y su carácter instrumental, que ayuda a la toma de decisiones en la vida cotidiana y profesional (Batanero, 2000).

Batanero (2000) señala que la estadística como ciencia, está en un periodo de notable expansión, como también las investigaciones relacionadas con los procesos de enseñanza–aprendizaje de la estadística; algunas de ellas se centran en analizar dificultades de los alumnos ante las tareas que involucran conceptos estadísticos, así como sus causas, otros estudios analizan la problemática relacionada con la formación de profesores en esta área.

En mi experiencia como docente de matemáticas en este nuevo servicio educativo de Bachillerato en México, denominado *Telebachillerato Comunitario,* he podido percibir deficientes conocimientos matemáticos y estadísticos de los estudiantes. Dada la importancia y utilidad de la estadística, esta problemática ha despertado en mí el interés de profundizar en la comprensión de las medidas de tendencia central y de variabilidad estadística que tienen los estudiantes, ya que se hacen evidentes diversas dificultades a que se enfrentan, como la confusión de las medidas de tendencia central y las medidas de dispersión, el cálculo incorrecto de la desviación estándar; confusión de la varianza con la desviación estándar y errores al interpretar resultados, es decir, falta de comprensión del significado del resultado que obtienen.

Aunado a esto, la estadística recibe poca atención y forma parte de los últimos cursos de matemáticas, por lo que por falta de tiempo se da más énfasis a la enseñanza de las medidas de tendencia central, dejando de lado la enseñanza de las medidas de dispersión.

Para este trabajo, se tomó en cuenta la investigación previa, en concreto, la que ha desarrollado Orta (2014) con estudiantes de secundaria de una escuela pública de la Ciudad de México, que analiza el razonamiento sobre *variabilidad* que tienen los estudiantes de ese grado escolar. Para este fin, examina los resultados mediante la Taxonomía SOLO[1]. Este trabajo trata de dar continuidad a dicha investigación, en el sentido de analizar la comprensión sobre de las medidas de tendencia central y la variabilidad, pero en este caso, con estudiantes de nivel bachillerato.

1.3. Justificación

Holmes (1980), señala que la enseñanza de la estadística recae en la importancia de preparar ciudadanos capaces de comprender la información que aparece en los medios de comunicación y que le sea útil en su vida posterior, y que su desarrollo personal fomente un razonamiento crítico.

[1] Structured Observed Learning Outcomes

El estudio de la estadística se introduce desde el nivel básico del Sistema Educativo Mexicano, con las medidas de tendencia central y el análisis de gráficas. En lo que respecta a la variabilidad estadística, en el tercer grado de educación secundaria el contenido del currículo se centra en: "Medición de la dispersión de un conjunto de datos mediante el promedio de las distancias de cada dato a la media (desviación media). Análisis de las diferencias de la 'desviación media' con el 'rango' como medidas de la dispersión" (SEP, 2011, p. 50).

Por otro lado, el bachillerato es el nivel clave para constituir en los jóvenes una mentalidad crítica a partir de la enseñanza del lenguaje estadístico y sus conceptos. En Matemáticas II del currículo mexicano del segundo semestre del bachillerato general, establece: "Obtener las medidas de tendencia central y de dispersión de datos agrupados y no agrupados, dentro y fuera de situaciones contextualizadas e interpreta y contrasta los datos con la realidad" (SEP, 2013, p. 47).

Para el mismo nivel educativo, en la asignatura de Probabilidad y Estadística I de dicho currículo del quinto semestre, se enuncia:

> Calcula las medidas de centralización en diversas situaciones a partir del conocimiento de los diferentes tipos de agrupación de datos para interpretarlos y analizarlos a través de las mismas. Calcula las medidas de variabilidad en diversas situaciones a partir del conocimiento de los diferentes tipos de agrupación de datos para interpretarlos y analizarlos a través de las mismas. Interpreta el comportamiento de una población a partir de las medidas de centralización y variabilidad de una muestra. (SEP, 2013, p. 25)

Se puede observar, que el estudio de las medidas de tendencia central y de variabilidad en el sistema educativo mexicano están incluidos en los currículos de educación secundaria y bachillerato (SEP, 2011 y SEP, 2017), en el primero se propone la utilización de las medidas de tendencia central y dispersión para analizar sus diferencias. En el caso del Telebachillerato Comunitario se presenta en dos momentos en el primer año (segundo semestre) el currículo plantea el cálculo, la interpretación y contrastación de las medidas de tendencia central y dispersión para la toma de decisiones, y en el tercer año (quinto semestre) propone el uso de las medidas de tendencia central y de variabilidad en diferentes contextos y tipos de agrupación, para interpretar comportamientos poblacionales mediante muestras y poder estructurar argumentos considerando las medidas provenientes de la misma (SEP, 2013) .

Por ello, en este trabajo se propone el diseño de una experiencia de aprendizaje que permita desarrollar conocimientos y fomentar el razonamiento de los estudiantes, relacionados con las medidas de tendencia central y de variabilidad. Este último es de suma importancia, pues analizar la comprensión del concepto de variabilidad implica uno de los conceptos estadísticos que más se dificulta a los estudiantes de nivel medio superior, que generalmente se encuentran en un rango de edad de 15 a 18 años (Orta, 2014).

La realidad cambiante de nuestro entorno ha obligado a la inserción de temas estadísticos en los currículos escolares, pues son herramientas necesarias para analizar y predecir la forma en que los cambios, ya sean naturales o sociales afectan a las personas y a la sociedad. Moore (1990) puntualizó la omnipresencia de la variabilidad, que se debe describir, cuantificar, explicar y controlar al generar datos. Por su parte, Wild y Pfannkuch (1999) señalaron que la variación está presente en todo y dondequiera, que puede tener efectos prácticos y modelables por estadísticos con los propósitos de predicción, explicación o control.

El proceso de documentación de este trabajo de tesis es importante, toda vez que al diseñar, elaborar y experimentar la experiencia de aprendizaje, se pretende que los estudiantes desarrollen conocimientos, habilidades y actitudes para enfrentar con buen juicio y a su debido tiempo, situaciones de la vida cotidiana.

1.4. Objetivo General

El objetivo general de este trabajo es implementar, en un ambiente con tecnología audiovisual, tareas de medidas de tendencia central y de variabilidad dirigidas a estudiantes de Telebachillerato en un marco de *experiencia de aprendizaje* que permita conocer sus dificultades y errores cuando se les pide realizar actividades relacionadas con estos conceptos, en que es necesario reconocer su cálculo y su interpretación.

Para alcanzar este objetivo, será necesario identificar los *significados personales* que los estudiantes asignan a otros conceptos relacionados con la variabilidad, como las medidas de tendencia central y de dispersión, así como sus respectivas representaciones tabulares y gráficas, utilizando como herramienta de análisis el *Enfoque ontosemiótico* y para lo que se llevará a cabo una *experiencia de aprendizaje* considerando el uso de videos.

La importancia de este estudio obedece, entre otros motivos, a la escasez de investigaciones relacionadas con la comprensión de la variabilidad en el bachillerato, así como a la necesidad de evaluar las dificultades con las que se enfrentan los estudiantes al realizar tareas que involucran estos conceptos. Así mismo, tomando en cuenta la puesta en marcha de las nuevas propuestas curriculares, y más aún, en este nuevo servicio educativo denominado "Telebachillerato Comunitario".

1.4.1. Objetivos específicos

- Explorar la comprensión de los conceptos de medidas de tendencia central: $\bar{x}$, *Me, Mo* y variabilidad que tienen estudiantes de Telebachillerato Comunitario.
- Analizar el contenido de material audiovisual para la enseñanza que sobre medidas de tendencia central y variabilidad reciben los estudiantes de Telebachillerato Comunitario.
- Seleccionar tareas adecuadas enfocadas en medidas de tendencia central y de variabilidad para implementar experiencias de aprendizaje tomando en cuenta la enseñanza con videos.
- Analizar la comprensión que tienen los estudiantes frente a tareas de medidas de tendencia central y variabilidad antes y después de una experiencia de aprendizaje.

1.5. Marco curricular

La estadística ha sido incorporada a los currículos escolares a lo largo del tiempo, desde los niveles básicos hasta los más avanzados de la educación, debido a su aplicación en diversos campos de conocimiento como las ciencias sociales, ciencias de la salud o la psicología entre otras.

El estudio de la estadística trasciende la simple realización de cálculos matemáticos o aritméticos, por medio de la sustitución de valores en fórmulas ya establecidas; implica un análisis de la variación del conjunto de datos, su comportamiento tendencial, los patrones que producen, la interpretación de resultados ante distintas situaciones para la toma de decisiones.

1.5.1. La estadística en el currículo mexicano

La educación estadística en México se imparte desde los niveles básicos establecidos en el artículo 37 de la Ley General de la Educación, particularmente en la primaria se realizan aproximaciones acerca de conjuntos de datos, gráficos, las medidas de posición central, derivados todos del tratamiento de la información, esto comúnmente se da al termino del tercer ciclo escolar (tercer grado) en adelante. Acorde al avance de su formación académica, el conocimiento estadístico tiende a ser más riguroso, por tanto, las aplicaciones en situaciones de la vida cotidiana suelen ser más frecuentes, mientras que al llegar al nivel universitario, la educación estadística es utilizada para formalizar investigaciones reales, a fin de reportar hechos, sucesos y tomar decisiones.

1.5.2. La estadística en la Dirección General de Bachillerato (DGB)

El área de conocimiento matemático se concibe como una ciencia formal, pues a lo largo de su historia han desarrollado métodos, lenguajes y procedimientos sistemáticos que posibilitan la matematización de la realidad. Este campo se conforma por conceptos matemáticos como son: el pensamiento numérico, algebraico, geométrico, estadístico que generan habilidades para realizar razonamientos matemáticos y demostraciones, explorar, comprender, representar, predecir, explicar, plantear, modelar y resolver problemas; haciendo uso de la comunicación para establecer vínculos entre las nociones informales e intuitivas y el lenguaje simbólico propio de esta ciencia (SEP, 2013).

La DGB pone en marcha el enfoque de resolución de problemas en la enseñanza de las matemáticas pues es primordial para el aprendizaje de otras ciencias. Según la SEP (2013), trabajar con este enfoque amplía la visión que deben desarrollar los alumnos de acuerdo a los análisis de temas y problemas que surgen en su contexto.

La Reforma Integral de la Educación Media Superior (RIEMS) fue adoptándola la DGB en su plan de estudios a partir del periodo escolar de 2009-2010, teniendo la finalidad de fortalecer y consolidar la identidad de este nivel educativo, en todas sus modalidades y subsistemas; y así proveer una educación oportuna y relevante al estudiante que le permita trabajar interactuando escuela-entorno de manera que exista una facilidad de tránsito académico entre los subsistemas y las escuelas, que se dará gracias al Marco Curricular Común, que compartirán todas las instituciones de bachillerato, basado en desempeños terminales, el enfoque educativo basado en el desarrollo de competencias, la flexibilidad y los componentes comunes del currículum (SEP, 2013).

La asignatura de Probabilidad y Estadística que establece la DGB, pertenece al componente de formación propedéutica, ubicado dentro del campo disciplinar de las Matemáticas; la cual tiene como objetivo:

> Desarrollar en el alumno habilidades, conocimientos y actitudes en relación con la estadística y sus aplicaciones, las técnicas de recolección de datos, la noción de variabilidad, los tipos de variables, la representación tabular y gráfica, la estadística descriptiva y la teoría de conjuntos. (SEP, 2013, p. 7)

1.5.3. El modelo educativo de Telebachillerato Comunitario

El Telebachillerato Comunitario fue implementado con la finalidad de ampliar la cobertura de la Educación Media Superior (EMS) y por tal, contribuir a elevar el nivel educativo de la población. Atendiendo principalmente a los egresados de secundaria de las comunidades rurales con las que se cuente este centro educativo. Este nuevo servicio educativo se imparte en modalidad escolarizada presencial con docentes especializados por área disciplinar, operando con el plan y programas de estudios vigentes del bachillerato general, además de hacer uso intensivo de materiales de estudio como, libros, cuadernos y programas audiovisuales por asignatura elaborados por la red EDUSAT, (SEP, 2013).

El modelo del Telebachillerato Comunitario está diseñado para acercar la EMS a localidades pequeñas, en el que se deberán atender grupos de 12 a 30 alumnos que podrán hacer uso de las instalaciones de la Telesecundaria y la utilización de materiales educativos son por mediación de los docentes (SEP, 2013).

1.5.3.1. La estadística en el Telebachillerato Comunitario

La asignatura Probabilidad y Estadística se encuentra dentro del componente formación propedéutica del bachillerato general y corresponde al campo disciplinar de Matemáticas. Acorde a la RIEMS y al Marco Curricular Común (MCC), tiene el objetivo de generar el desarrollo de la creatividad y el pensamiento lógico y crítico por medio de procesos de razonamiento, argumentación y construcción de ideas, que propician el desarrollo de diferentes competencias para la resolución de problemas que vayan más allá de las aulas de clase.

En el Telebachillerato Comunitario la asignatura de Probabilidad y Estadística se desarrolla en dos momentos, el primero incluye solo estadística descriptiva y el segundo la probabilidad básica, la primera que es en la que se centra esta tesis se ubica en el quinto semestre del plan de estudios del bachillerato general; para esto, el alumno ya debió haber cursado las asignaturas: Matemáticas I, II, III y IV, además la asignatura de Desarrollo Comunitario. El segundo se ubica en el sexto semestre del mismo plan de estudios y con las mismas condiciones. La finalidad de la estadística en este modelo educativo es que el alumno a lo largo del curso desarrolle una serie de conocimientos, habilidades, actitudes y valores para su desenvolvimiento en la sociedad como ciudadano, y sea capaz de resolver y comprender situaciones de la vida cotidiana a través del lenguaje científico y matemático, así como a construir nuevos conocimientos y compartirlos con quienes lo rodean (SEP, 2013).

1.5.3.2. Programa de Estudios vigente de Probabilidad y Estadística I, del Bachillerato General

El bachillerato general al buscar la consolidación de los aprendizajes, profundiza el desarrollo de competencias relacionadas con el campo disciplinar de las Matemáticas. De esta forma, la asignatura de Probabilidad y Estadística I conserva una relación y trabajo conjunto con el resto de las asignaturas, lo cual permite la diversificación de saberes y un trabajo interdisciplinario.

La DGB en su programa de estudios de Probabilidad y Estadística I, lo integra en cuatro bloques que se describen a continuación (SEP, 2013):

Bloque I. Comprendes y describes la variabilidad estadística y sus aplicaciones. Pretende que el docente haga despertar el interés en el alumno de manera que valore la estadística como una herramienta matemática que lo conduzca en la toma de decisiones para organizar, resumir datos y comunicar resultados, así mismo poder distinguir las principales características teóricas de las ramas de la estadística, reconocer y valorar las técnicas de recolección de datos, además logre identificar las variables como atributos de interés de los datos provenientes de una población o muestra reconociendo su comportamiento y diferencia para facilitar su estudio y análisis posterior.

Bloque II. Describes y representas datos de forma tabular y gráfica. En este bloque se aspira a que el docente desarrolle en el alumno conocimientos para construir la representación tabular y gráfica de los datos en categorías mutuamente excluyentes provenientes de una población o muestra, para obtener una mejor comprensión del comportamiento de las poblaciones de objeto de estudio.

Bloque III. Aplicas la estadística descriptiva. La finalidad de este apartado es promover en el alumnado desempeños que le permiten calcular las medidas de tendencia central en diversas situaciones a partir del conocimiento de los diferentes tipos de agrupación de datos, calcular las medidas de variabilidad en diversas situaciones a partir del conocimiento de los diferentes tipos de agrupación de datos, reconocer las diversas técnicas de muestreo y las ventajas al ponerlas en práctica para la recolección de datos de una población.

Bloque IV. Analizas la teoría de conjuntos y sus aplicaciones. En este bloque el docente promueve en el alumnado desempeños que le permiten identificar los elementos de un conjunto y sus operaciones como base para la probabilidad, y comprender las características de experimento, espacio muestral, punto muestral y evento como elementos de la probabilidad simple.

En este estudio se consideró el Bloque III, enfocado en el cálculo y la interpretación de las medidas de tendencia central y de variabilidad, con la finalidad que el estudiantado pueda reconocer la representatividad y dispersión de la información. Por lo que, se diseñará una experiencia de aprendizaje buscando favorecer estos aprendizajes.

CAPÍTULO 2

MARCO CONCEPTUAL Y ANTECEDENTES

2.1. Introducción

Esta sección se compone de tres apartados: el marco conceptual, el marco teórico y las investigaciones previas. En el marco conceptual, se describe cada una de las medidas de tendencia central y de dispersión; así también, la metodología de experimentos de enseñanza, que será fundamental para el diseño de la experiencia de aprendizaje. En el segundo apartado se describe el Enfoque ontosemiótico, con el que se analizarán las respuestas que los estudiantes den a las actividades que se incluyen en la experiencia de aprendizaje. El último apartado integra los antecedentes que sirven como referencia de las diferentes investigaciones que se han desarrollado entorno a estos temas. Los cuáles serán de apoyo para diseñar la secuencia de actividades propuestas para los estudiantes y poder contrastar algunos de sus resultados con los nuestros.

2.2. Marco conceptual

El marco conceptual, según Miles y Huberman (1994), explica de forma gráfica o narrativa, las ideas principales, conceptos y definiciones que serán estudiados en el trabajo a desarrollar; en este caso, se describen los conceptos, fórmulas y propiedades de las medidas de tendencia central y de variabilidad.

2.2.1. Medidas de tendencia central

La media aritmética, la mediana y la moda, también son conocidas como promedios, medidas de localización, de centralización o posición. Son medidas que proporcionan información acerca de un conjunto de datos obtenidos de fuentes primarias (observación, encuestas, etc.) o secundarias (bases de datos) para su análisis. La información que se obtiene de estas medidas nos facilita la interpretación y permite realizar un análisis con mayor profundidad del conjunto de datos. Las medidas de tendencia central dan una descripción acerca de la ubicación de los datos, es decir, permiten identificar cómo se distribuyen cada uno de los datos alrededor el centro, en otras palabras, son valores que representan un colectivo, alrededor del cual se agrupan los demás datos del conjunto.

2.2.1.1. Media aritmética

La media aritmética es una medida de tendencia central muy útil, su función es representar un conjunto de datos, es decir, se obtiene para conocer el valor medio que describe la serie. Lind, Marchal y Mason (2004), Estrada y Hernández (2015), señalan que la estimación de la media se da sumando todos los valores de la serie dividida por el número total de datos. La fórmula para calcular la media aritmética es:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Significado de los símbolos:

$\bar{x}$: *Símbolo que se utiliza para identificar la media*

$\sum$: *Indica la sumatoria de todos los valores que toma la variable numérica x_i desde el primero hasta el n-ésimo.*

x_i: *Cada uno de los datos obtenidos de la muestra.*

n: *Número total de datos.*

Es importante recalcar que la media aritmética únicamente puede ser utilizada con valores de variables cuantitativas. Al mismo tiempo, cuenta con varias propiedades importantes que a continuación se enlistan según las expuestas en el trabajo de Strauss y Bichler (1988), y Lind *et al*. (2004):

a) La media se encuentra entre los valores extremos.
b) La suma de las desviaciones de cada uno de los valores respecto a la media es cero.
c) La media cambia al agregar un dato diferente de la media.
d) La media puede no pertenecer al conjunto de datos.
e) En el cálculo de la media de un conjunto de valores discretos puede obtenerse una fracción que no corresponde a una realidad física.
f) El cálculo de la media considera a todos los valores, incluido el cero y negativo.

Levin y Rubin (2004), mencionan que la media aritmética contempla importantes ventajas, entre las cuales se puede decir que es de fácil cálculo, por lo que, es la más utilizada en análisis de datos. En un conjunto de datos solo se puede estimar una media, y nos es útil para realizar comparaciones de dos o más conjuntos de datos. Por otro lado, una desventaja es que cuando se presentan valores extremos o atípicos, la media tiende a alejarse del contexto, por tanto se dice que la media no puede ser un buen representante del conjunto de datos.

2.2.1.2. Mediana

Es una medida de tendencia central considerada como valor equidistante o el que se localiza en el centro de todo el conjunto de datos, habiéndose ordenado de menor a mayor, entonces, decimos que la mitad de los datos se encuentran por arriba de la mediana y la otra mitad se localiza por debajo (Levin y Rubin, 2004). Este estadístico se representa con el símbolo: $\tilde{x}$ o la abreviación *Me*. Para obtener el valor de la mediana de un conjunto de datos, en primer lugar se organizan en orden descendente o ascendente. De esta forma, se identifica si el conjunto de datos es un número *impar*, o si hay un número *par* de observaciones, por tanto, se debe considerar dos casos posibles para el cálculo de la mediana:

El primer caso puede darse cuando el número de datos (*n*) es impar, entonces la mediana es el valor que se encuentra en el centro del conjunto de datos ordenados. Para saber cuál es este valor se debe calcular la posición o el lugar donde se encuentra la mediana:

Posición de la mediana: $\left(\frac{n+1}{2}\right)$ *-ésimo término del arreglo de datos*

El segundo caso se presenta si el número de datos (*n*) es par, el valor de la mediana ha de ser el valor medio de los dos valores centrales del conjunto de datos ordenados. Es decir, nos encontramos frente al caso de indeterminación, pues al ordenar previamente el conjunto de datos y querer determinar el valor central, nos encontramos con dos valores, entonces, se hace un convenio, por lo que el valor de la mediana se obtiene como la media aritmética de los dos valores centrales. En este caso, al calcular la posición de la mediana se obtiene un valor fraccionario, en consecuencia, se necesita promediar los dos valores adyacentes.

Cuando en un conjunto de datos existe la presencia de valores atípicos, esto es, datos muy grandes o muy pequeños con respecto a los demás, la media aritmética se desproporciona y pierde su representatividad del conjunto de datos, en esos casos es conveniente utilizar la mediana como medida de tendencia central. Para esto, es necesario conocer las propiedades de esta medida de tendencia central, según Lind *et al.* (2004) y Mayén (2009):

a) Un conjunto de datos solo puede tener una mediana
b) El valor de la mediana está comprendida en el rango del conjunto de datos
c) La mediana puede coincidir o no con los valores del conjunto datos
d) En el cálculo de la mediana no se contemplan todos los valores del conjunto de datos
e) La mediana no cambia si cambiamos un dato con valor menor o superior
f) En el cálculo de la mediana se pueden obtener valores diferentes al sistema numérico original, por lo que se dice que no es una operación interna en el conjunto numérico empleado.
g) La mediana conserva los cambios de origen y de escala
h) Si agregamos un nuevo elemento puede modificarse el cálculo, por lo que decimos que no tiene elemento neutro ni elemento simétrico
i) La mediana no tiene la propiedad asociativa
j) La mediana es un representante de un conjunto de datos
k) La mediana no se puede obtener para datos nominales.

En cuanto a las ventajas y desventajas de la mediana, Levin y Rubin (2004), subrayan que una de las ventajas más importantes destaca en que la mediana no se ve afectada por valores extremos al contrario de la media, entonces puede usarse en aquellos casos en que la media aritmética no es un buen representante del colectivo. Es decir, si la distribución del conjunto de datos es asimétrica la mediana es preferible a la media. Por otro lado, se tienen algunas desventajas, la principal es que la mediana no es útil para hacer cálculos adicionales. Al mismo tiempo, se tiene que la mediana utiliza solo algunos valores del conjunto de datos. Por último, cuando se tiene un conjunto con gran cantidad de datos no es fácil ordenar los datos, por lo que se complicaría el cálculo de la mediana.

2.2.1.3. Moda

La moda, se define como el valor más frecuente en un conjunto de datos (Estrada y Hernández, 2015), se usa el símbolo $\hat{x}$ para representarla o la abreviatura *Mo*. Es fácil obtener esta medida para una serie de datos no agrupados, la estimación se realiza por simple inspección, pues, es el valor de la variable que más se repite. La moda de un conjunto de datos puede ser única, existir dos modas (bimodal), o más de dos modas (multimodal), a diferencia de la media y la mediana que son valores únicos.

Seguidamente, se presentan algunas propiedades que mencionan Cobo (2003) y Mayén (2009):

a) La moda es un representante de un conjunto de datos
b) La moda es un estadístico más resistente que la media
c) Puede existir la moda para variables cuantitativas y para variables cualitativas
d) La moda pertenece al rango de la variable
e) La moda es un valor del conjunto de datos
f) La moda no toma en cuenta todos los valores del conjunto de datos
g) Es conmutativa
h) Conservan cambios de origen y escala
i) Puede no existir moda o existir más de dos modas (bimodal o multimodal)
j) La moda no se ve afectada por valores extremos, muy altos o muy bajos
k) La moda se considera como operación interna, porque conserva su valor en el conjunto de datos

Por lo que se refiere a las ventajas de la moda, se puede decir que es la medida de tendencia central que más fácilmente se determina, puesto que se puede obtener por inspección, identificando el valor que tiene más frecuencia. En cuanto a las desventajas aunque su estimación es muy fácil, es considerada una medida de tendencia central poco confiable. Además, cuando se presenta el caso de una distribución amodal (sin moda), la mediana es el mejor representante del conjunto de datos. Otra desventaja se da cuando existen más de dos modas, pues crece la dificultad de interpretación y comparación del conjunto de datos (Levin y Rubin, 2004).

2.2.2. Medidas de dispersión

En la estadística dentro de las medidas descriptivas, se encuentran las medidas de variabilidad o dispersión, que nos proporcionan un valor que indica si los valores del conjunto de datos están próximos entre sí, o si se encuentran muy alejados respecto a los valores de tendencia central. Los valores que toman estas medidas pueden ser mayores a cero o igual a cero; si se presenta este último entonces se dice que no hay dispersión. Si el valor que toman es muy pequeño (cercano a cero), entonces los datos presentan poca dispersión, es decir, están acumulados o próximos al valor central, por lo que, la medida de tendencia central usada es un buen representante del conjunto de datos y la distribución se considera homogénea. Por el contrario, si la dispersión es grande la medida de tendencia central que representa el conjunto de datos no es muy confiable, en este caso se cuenta con una distribución heterogénea.

En este trabajo se abordan las siguientes medidas de dispersión: el rango, la varianza y la desviación estándar.

2.2.2.1. Rango

También conocida como recorrido, es considerada la medida de variabilidad más sencilla, fácil de entender y de estimar. Tiene una utilidad limitada, porque proporciona poca información. El rango es la diferencia aritmética del valor más bajo y del valor más alto de un conjunto de datos (Estrada y Hernández, 2015), como vemos en el cálculo de esta medida solo intervienen dos valores, por tanto es poco representativa de la dispersión de datos de la serie. Por último, el símbolo utilizado para representarlo es R y la fórmula para su cálculo es:

$$R = x_{máx} - x_{min}$$

Significado de los símbolos:

R: es el rango o recorrido

$x_{máx}$: *Valor máximo*

x_{min}: *Valor mínimo*

Durazo (2012), señala algunas propiedades del rango:

a) Al calcular el rango las unidades son las mismas que las de la variable
b) En la estimación solo intervienen dos observaciones del conjunto de datos

c) Si se agrega un dato extremo a la serie, el rango se ve afectado

Dentro de las ventajas que se conocen del rango son su fácil cálculo, así como su facilidad para ser interpretada. También es útil para la construcción de distribuciones de frecuencias. Respecto a las desventajas, al usar solo los valores extremos ignora los demás datos del conjunto. No es recomendable su uso para muestras grandes, tampoco considera si existe o no existe moda alguna (Levin y Rubin, 2004).

2.2.2.2. Varianza

Es una medida de variabilidad que usa cada uno de los datos de la serie. Definida como el promedio de las diferencias cuadradas de las observaciones con respecto a su media aritmética, es decir, es la media de las desviaciones de la media elevadas al cuadrado (Durazo, 2012). Dicho de otra manera, la varianza muestra que tan lejos o cerca se encuentran los datos respecto a la media.

La diferencia de cada dato (x_i) y la media ($\bar{x}$) es conocida como la desviación respecto a la media, se escribe ($x_i - \bar{x}$) y para realizar el cálculo de la varianza, se elevan al cuadrado las desviaciones respecto la media, se suman todas y se divide por n. La varianza es no negativa, y si todas las observaciones son iguales la varianza es cero.

El símbolo que se usa para identificar la varianza está dado por σ^2, (Durazo, 2012). Para efectuar el cálculo de la varianza de un conjunto de datos no agrupados, se presentan la siguiente formula algebraica:

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}$$

dónde:

$\bar{x}$ *representa la media muestral*

n *representa el tamaño de la muestra*

x_i *representa a cualquiera de los datos en la muestra*

Propiedades de la varianza:

a) La varianza señala el grado con que los datos de la distribución se encuentran dispersos, es decir, a mayor medida, mayor dispersión.
b) Es un valor muy grande con respecto a las observaciones, en consecuencia se vuelve difícil para trabajar.

c) En la varianza las desviaciones son cuadradas y se expresa en términos de los datos originales elevados al cuadrado, es decir, las unidades de la varianza son los cuadrados de las unidades de los datos, por lo que no tiene sentido considerarlas pues no tienen una interpretación lógica.
d) Cuando el tamaño de la muestra aumenta, disminuye la varianza.
e) El valor de la varianza siempre será positivo.

En cuanto a las ventajas y desventajas de esta medida, Gutiérrez (2012), señala que en la varianza un dato está muy cerca de la media, al elevar esa distancia al cuadrado se hará más pequeña, es decir estará más cerca. De lo contrario, si el dato está muy lejos de la media, la distancia al elevarlo al cuadrado respecto a la media se hará mayor. Por esto, se hace visible la dispersión alta o baja en una distribución de datos. Respecto a las desventajas, la varianza no viene expresada en la misma unidad de medida que los datos. Por tanto es difícil su interpretación, por lo que suele utilizarse en su lugar el estadístico que se describe a continuación.

2.2.2.3. Desviación estándar

Es la medida de dispersión que más se utiliza (Gutiérrez, 2012). Se define como la raíz cuadrada positiva de la media de los cuadrados de la diferencia entre cada valor y la media. En otras palabras, es la raíz cuadrada positiva de la varianza. La desviación estándar regresa la medida de variabilidad a las unidades originales del conjunto de datos, es decir, el resultado se presenta en las mismas unidades que los datos originales. Esta medida de variabilidad mide el alejamiento o acercamiento promedio de los datos alrededor de la media. El símbolo que usa es (σ).

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}}$$

Dónde:

$\bar{x}$ *representa la media muestral*

n *representa el tamaño de la muestra*

x_i *representa a cualquiera de los datos en la muestra*

Propiedades de la desviación estándar

a) La desviación estándar muestra el nivel de dispersión de los datos en una distribución. Cuando el valor es mayor, mayor será la dispersión y viceversa.
b) La desviación estándar se expresa en unidades de los datos originales.
c) Cuando el tamaño de la muestra aumenta, disminuye la desviación estándar.
d) Si todos los datos de la distribución son iguales, la desviación estándar es igual a cero.
e) En su cálculo se usan todos los datos de la serie; por tanto, se ve afectada cuando hay un cambio de valor en algún dato.

Levin y Rubin (2004), plantean que la principal ventaja que tiene la desviación estándar es, que es más fácil de interpretar que la varianza debido a que la desviación estándar se mide en las mismas unidades que los datos. La desventaja es que no es recomendable su uso, cuando tampoco lo sea el de la media como medida que representa el conjunto de datos.

2.2.3. Experimentos de enseñanza

Actualmente la metodología de experimentos de enseñanza es considerada parte de la investigación sobre la enseñanza y el aprendizaje de las matemáticas, y de otras ciencias de la educación con enfoques de investigación.

> La metodología de los experimentos de enseñanza ofrece a los investigadores la experiencia de primera mano del aprendizaje y razonamiento matemático de los estudiantes, permitiendo construir modelos de las “matemáticas de los estudiantes” que incluyen las modificaciones que los estudiantes hacen en sus formas de operar. (Steffe, Thompson y Glasersfeld, 2000, p. 269).

Un experimento de enseñanza consta de una secuencia de episodios de enseñanza, que involucra investigador-docente, uno o más estudiantes, uno o más investigadores-observadores (Steffe y Thompson, 2000). Durante el estudio se colecta información de lo que sucede en el aula mediante grabaciones de video y/o toma de notas. El tiempo de duración del experimento puede variar, desde unas horas hasta un ciclo escolar.

El ambiente donde se desarrolla un experimento de enseñanza va desde pequeños espacios (habitaciones-laboratorio) para entrevistas, clases completas o ambientes de aprendizaje mayores. El estudio puede estar centrado en el desarrollo de los alumnos, el de los docentes o de algunas ideas, o bien, en las actividades de enseñanza (Kelly y Lesh, 2000).

En un experimento de enseñanza se espera que el alumno construya conocimiento, que el investigador-docente también construya conocimiento sobre la construcción del conocimiento de los alumnos, además que los investigadores construyan conocimientos sobre ambos y sus interacciones. La distinción de estos planos de acción ha conducido a que en ocasiones se les llame a estos estudios *experimentos multiniveles o multietapas*[2] (Confrey, 2006; Lesh y Kelly, 2000).

Para la elaboración de los experimentos de enseñanza, Cobb y Gravemeijer (2008), distinguen tres fases:

> *Preparación* del experimento, *experimentación* para promover el aprendizaje y ejecución del *análisis* retrospectivo de los datos. En la segunda de estas fases tienen lugar las intervenciones en el aula y las sucesivas iteraciones del ciclo de tres pasos: 1) diseño y formulación de hipótesis; 2) intervención en el aula y recogida de datos; y 3) análisis de los datos y revisión y reformulación de hipótesis. (Molina, Castro, Molina y Castro, 2011, pp. 7-8)

La Tabla 1, tomada de Molina *et al.* (2011), contiene las acciones implicadas en cada una de las fases ya mencionadas, las cuales deben estar fundamentadas con la literatura relevante y relacionada con el problema de investigación. Dichas acciones han sido identificadas y clasificadas por (Castro y Molina, 2007; Molina y Ambrose, 2008; Molina, Castro, y Castro, 2009).

[2] “multi-tiered” traducción del autor.

Tabla 1. Acciones contempladas en cada fase de los experimentos de enseñanza

Fases	Acciones
Preparación del experimento	▪ Definir el problema y los objetivos de investigación. ▪ Identificar los objetivos instruccionales. ▪ Evaluar el conocimiento inicial de los alumnos. ▪ Identificar las metodologías de enseñanza adecuadas para los contenidos elegidos, en función de los objetivos planteados y los conocimientos previos de los alumnos. ▪ Diseñar de forma justificada la secuencia de intervenciones en el aula y su temporalización. ▪ Diseñar la recogida de datos. ▪ Delinear una trayectoria hipotética de aprendizaje que describa el resultado esperado del proceso de aprendizaje y el modo en que se va a promover y alcanzar dicho aprendizaje. ▪ Ubicar el experimento dentro de un contexto teórico más amplio en el que se enmarque el modelo teórico emergente.
Experimentación *Antes de cada intervención*	▪ Obtener información sobre el trabajo previo realizado en el aula, para tenerlo en cuenta en el diseño de la intervención y en la posterior interpretación de los datos. ▪ Identificar los objetivos instruccionales de la intervención. ▪ Ultimar el diseño de la intervención, de forma justificada, a partir de la información empírica y teórica disponible. ▪ Elaborar hipótesis/conjeturas sobre los resultados a obtener en la intervención. ▪ Ultimar la selección de los métodos de recogida de datos. ▪ Registrar las decisiones tomadas en el proceso de ejecución de las acciones descritas en los cinco apartados anteriores y su justificación.
En cada intervención	▪ Si es necesario, modificar sobre la marcha, de manera justificada, el diseño de la intervención de acuerdo con los objetivos de la intervención. ▪ Recoger datos de todo lo que ocurre en el aula, incluyendo las decisiones tomadas durante la intervención.
Después de cada intervención	▪ Analizar los datos recogidos en la intervención. ▪ Revisar, y en su caso reformular, las hipótesis/conjeturas de investigación.
Análisis retrospectivo de los datos	▪ Recopilar y organizar toda la información recogida. ▪ Analizar el conjunto de los datos, lo que implica: - Distanciarse de los resultados del análisis preliminar, de las conjeturas iniciales y de la justificación del diseño de cada intervención, para profundizar en la comprensión de la situación de enseñanza y aprendizaje en su globalidad. - Identificar la ruta conceptual seguida por el grupo y por cada alumno, por medio de los cambios que pueden ser apreciados, atendiendo a las acciones específicas del investigador-docente que contribuyeron a dichos cambios.

2.2.4. Enfoque ontosemiótico (EOS) de la cognición y la instrucción matemática

El "enfoque ontosemiótico" (EOS) de la cognición y la instrucción matemática, ha sido desarrollado por Godino y colaboradores (Godino y Batanero, 1994; Godino, 1996, 1999, 2002, 2003, 2008; Godino, Batanero y Roa, 2005). Dentro de este enfoque se analiza el significado de los objetos matemáticos que tienen un carácter complejo, estos objetos emergentes pueden ser "objetos institucionales", si los sistemas de prácticas son compartidos en el seno de una institución, y "personales" si corresponden a una persona.

Este trabajo tiene la finalidad de conocer los objetos personales que se definen como, "Los objetos matemáticos personales, son emergentes del sistema de prácticas personales significativas asociadas a un campo de problema" (Godino y Batanero 1994, p. 336); del mismo modo, definen el significado personal de un objeto, que consiste en "el sistema de prácticas personales de una persona para resolver el campo de problemas del que emerge el objeto en un momento dado" (Godino y Batanero 1994, p. 343).

En el quehacer matemático se identifican lo que Godino ha llamado entidades primarias de los objetos y que se llamarán elementos de significados, que se han de considerar en el análisis:

1. *Situaciones-problema* (aplicaciones extramatemáticas o intramatemáticas, ejercicios) son las tareas que inducen a la actividad matemática de donde surge el objeto. Por ejemplo, estimar la media y la mediana a partir de un gráfico.

2. *Lenguaje* (términos, expresiones, notaciones, símbolos, tablas, gráficos) dados en forma escrita, grafica, oral o gestual. Por ejemplo, las palabras media o media aritmética y mediana, así como los símbolos $(Mo, Me, \tilde{x}, \bar{x})$ y gráficos asociados.

3. *Conceptos-definición* (son dados mediante definiciones o descripciones). Por ejemplo, la definición de mediana como valor central o como valor que divide una distribución ordenada de datos en dos partes iguales, definición de media como promedio.

4. *Proposiciones, propiedades o atributos de los objetos mencionados* (enunciados sobre conceptos), es decir, propiedades particulares y su relación con otros

conceptos. Por ejemplo, la media puede no pertenecer al conjunto de datos, la mediana puede coincidir o no coincidir con los valores del conjunto de datos.

5. *Procedimientos o acciones del sujeto ante las tareas matemáticas* (algoritmos, operaciones, técnicas de cálculo, procedimientos), ante un problema, al tratar de resolverlo o al comunicar la solución se realizan diferentes algoritmos. Para este trabajo se tiene por ejemplo, el algoritmo de cálculo de la media y la mediana para datos aislados, a partir de gráficos.

6. *Argumentos* (enunciados usados para validar o explicar proposiciones o procedimientos) pueden ser deductivos o de otro tipo.

Godino, Batanero y Font (2007), señalan que estos seis tipos de elementos que se ponen en juego en la actividad matemática, amplían la distinción usual entre conceptos y procedimientos, es decir, se puede hablar de comprensión procedimental y conceptual, además de la argumentativa o lingüística. En la que las situaciones-problemas es donde inicia la actividad matemática; seguido por el lenguaje necesario para representar los problemas, procedimientos, conceptos y proposiciones; también los argumentos que son los que se usan para justificar los procedimientos y las soluciones de los problemas, y por último las proposiciones que relacionan los conceptos entre sí.

En el enfoque ontosemiótico, con la diversidad de objetos primarios se usa la idea de *conflicto semiótico* que se define como "la disparidad o desajuste entre los significados atribuidos a una misma expresión por dos sujetos —personas o instituciones— en interacción comunicativa y pueden explicar las dificultades y limitaciones de los aprendizajes y las enseñanzas implementadas" (Godino, 2002, p. 258), esto es, las interpretaciones de los estudiantes que no coinciden con las pretendidas por el profesor o el investigador. Estos conflictos semióticos generan equivocaciones en los alumnos no por la escasez de los conocimientos, más bien, por no lograr relacionar correctamente los dos términos de una función semiótica.

Los elementos en el significado de un objeto matemático contribuyen a definir la problemática, para el diseño de estrategias didácticas y la evaluación de los conocimientos de los estudiantes, considerando los diversos elementos de significado.

El desarrollo de nuestra investigación está enfocado en identificar los *significados personales* que los estudiantes asignan a las medidas de tendencia central y de variabilidad, y así conocer la comprensión que ellos tienen de estos conceptos.

2.3. Investigaciones previas

En este apartado se describen brevemente estudios centrados en la comprensión de las medidas de tendencia central y de variabilidad que se han realizado en diferentes niveles educativos y contextos sociales diversos.

2.3.1. Investigaciones sobre medidas de tendencia central

Las primeras investigaciones que se enlistan están orientadas a conocer las ideas de los estudiantes sobre las medidas de tendencia central y los principales errores que cometen en la resolución de problemas.

Russell y Mokros (1991), estudiaron las nociones que tienen, sobre las medidas de tendencia central, un grupo de estudiantes de 10 a 14 años de cuarto a octavo de primaria. Los resultados obtenidos muestran que los alumnos interpretan la definición de la media de acuerdo con lo que los autores clasifican en cuatro categorías los significados incorrectos de la media: “valor más frecuente” (confusión entre media y moda), "valor razonable" (significado coloquial del término), "punto medio" (confusión con la mediana) y "algoritmo" (este es un significado restringido, se ve solo como algoritmo de cálculo).

Eisenbach (1994), en un curso universitario introductorio de estadística solicita a sus alumnos que expresen el significado de la frase: "¿Qué quiere decir que el salario medio de un empleado es 3.600 dólares?" y obtuvo las siguientes respuestas: "que la mayoría de los empleados gana cerca de 3.600 dólares", o bien que, "es el salario central; los demás trabajadores ganan más o menos de 3600 dólares", con esto el autor destaca que los estudiantes muestran la confusión terminológica entre las palabras "media", "mediana" y "moda". Dicho de otra forma, existe una confusión con las medidas de centralización, es decir, no existe discriminación entre dichas medidas. Es difícil para los estudiantes comprender la idea de promedio sino se visualiza el conjunto de los datos.

Carvalho (1996), efectuó un estudio con 182 estudiantes de 7° de Educación Básica entre 14 y 15 años, encontró que en el cálculo de la media los estudiantes suman las frecuencias absolutas en vez de los valores de las variables, esto cuando los problemas se presentan en tablas de frecuencias. Por otro lado, en la mediana eligen el valor central del conjunto de frecuencias absolutas. Del mismo modo, al presentar los datos en forma de gráfico, la dificultad de interpretación aumenta y a menudo cometen más errores.

Batanero, Godino y Navas (1997), realizaron un estudio sobre las nociones que tienen los profesores de primaria en formación (1° de Magisterio y de 2° de Pedagogía), acerca de los promedios. Los autores llegaron a la conclusión de que los estudiantes tienen dificultad para entender los promedios cuando en su cálculo se tiene presencia de los ceros y valores atípicos. Gran parte de los estudiantes realizan los cálculos de la media considerando estos valores, sin hacer diferencia de las situaciones en que debe quitar o no un valor atípico antes de realizar la estimación. También, resultan difíciles de comprender las posiciones relativas de la media, moda y mediana, debido a que los estudiantes conciben a todas las distribuciones como simétricas, pues se presentan en conjuntos de datos descontextualizados.

Carvalho (1998), en un análisis hecho con las tareas estadísticas escritas que desarrollan normalmente en clases sus estudiantes, halló los siguientes errores en el cálculo de los promedios:

- Moda: Tomar la mayor frecuencia absoluta. Este error lo comete el estudiante al confundir la frecuencia con el valor de la variable, es decir, el alumno no considerar el valor de la variable que más frecuencia tiene.

- Mediana: No ordenar los datos para calcular la mediana; calcular el dato central de las frecuencias absolutas ordenadas de forma creciente; calcular la moda en vez de la mediana; equivocarse al calcular el valor central; tomar como mediana el valor central de las frecuencias de la tabla. Otro error se muestra, cuando el estudiante concibe la mediana como el centro de una distribución no ordenada. Además de confundir la frecuencia con el valor de la variable.

Watson y Moritz (1999), desarrollaron un trabajo con 2,250 estudiantes entre 8 a 18 años, que cursaban desde 3° de primaria hasta el curso de orientación universitaria. Los resultados muestran que los niños inician usando un modo coloquial y progresan su comprensión de los promedios al avanzar en las etapas escolares. Señalan que algunas propiedades de la media, como la representatividad, que no aparece en los contextos cotidianos, sólo lo ponen de manifiesto los estudiantes con conocimientos más avanzados. La idea de que la representatividad es un concepto básico para la comprensión de las medidas de tendencia central, es defendida por los autores, al permitir entender el significado de cada una de ellas y, a la vez, determinar las características del conjunto de datos.

- Media: Hallar la media de los valores de las frecuencias; no tener en cuenta la frecuencia absoluta de cada valor en el cálculo de la media. Por último, se repite nuevamente la confusión de frecuencia y valor de la variable, al igual que no pondera los datos para el cálculo de la media.

Batanero (2000), analizó los componentes del significado de las medidas de posición central, así como las dificultades para su comprensión. Concluyó que los estudiantes no logran comprender los promedios, particularmente los algoritmos de cálculo según su estructura (agrupados y no agrupados) y más aun con datos descontextualizados y las propiedades que contienen.

Watson y Moritz (2000), realizaron un estudio con 137 entrevistas, 97 de ellas a una muestra inicial de estudiantes entre 3º curso de Primaria y 3º Curso de Secundaria (8 a 15 años). Encuentran que para una gran parte de niños el promedio es simplemente un valor en el centro de la distribución, es decir, usan una idea cercana al concepto de la mediana. Al mismo tiempo, se observa que pocas veces relacionan la palabra “promedio” con moda y menos con media aritmética. Al realizar las entrevistas los autores obtuvieron las siguientes definiciones de promedio: "Significa igual", "que es normal", "no eres realmente bueno, pero tampoco malo". Concluyen sugiriendo que los estudiantes progresan conforme aumentan su edad. Del mismo modo, observan que los estudiantes de cursos más altos revelan una menor evolución en comparación con los de menor nivel, deducen que existe ineficiencia en la enseñanza de la estadística.

Carvalho (2001), observó en 533 alumnos, a lo largo de dos cursos escolares, el efecto que el trabajo cooperativo en parejas (grupo experimental) ocasiona sobre el desarrollo de la capacidad lógica y sobre el aprendizaje de los conceptos estadísticos elementales. Para tal estudio, la autora dota a los estudiantes de "tareas estadísticas no habituales" es el título que ella les da, estas tareas consisten en problemas estadísticos sencillos y abiertos. Los resultados obtenidos muestran que los estudiantes trabajaron de forma cooperativa evidenciando un avance más claro, en el desarrollo lógico y en sus competencias estadísticas que sus compañeros. De igual forma, encontró que formar parejas heterogéneas en cuanto a sus capacidades y conocimientos, traía beneficios para los estudiantes con un nivel inferior, así como para su compañero. Por último, el contenido estadístico genera dificultades a los estudiantes, aun si trabajan en parejas y con tareas no habituales, logran realizar las tareas, pero el aprendizaje que se ve reflejado es memorístico y no alcanzan una comprensión profunda.

Cobo (2003), realizó un estudio teórico-experimental acerca del significado y comprensión de las medidas de posición central en la Educación Secundaria Obligatoria (E.S.O.), para tal efecto analiza los problemas, representaciones, procedimientos, definiciones, propiedades y argumentos, desde las dimensiones institucional y personal. El estudio experimental lo desarrolla con una muestra de 312 alumnos de 1° y 4° curso de E.S.O. de diferentes institutos de la provincia de Granada, España. Los resultados de este estudio señalan, primeramente, que los libros de texto ponen mayor énfasis a la definición de la media que al estudio de sus propiedades, también la falta de problemas que soliciten el cálculo de una medida en diferentes mediciones con presencia de errores y demás situaciones. Por otro lado, la autora describe varios algoritmos de cálculo para cada medida de tendencia central, incluidos el cálculo con datos aislados, a partir de tablas o gráficas, además de la inversión de los algoritmos. Con respecto al algoritmo de la media, la autora señala que no es comprendido del todo, aun así, cuando los estudiantes fueron capaces de dar una definición adecuada, únicamente un 20% de los estudiantes lograron invertir el algoritmo para dar solución a los problemas.

Vermette, Gattuso y Bourdeau (2005), analizan la interpretación que los estudiantes dan a los datos estadísticos en la prensa, los autores identifican diferentes niveles de argumentación: que van desde la falta de argumentación, hasta un razonamiento estadístico completo. Los resultados revelan que solo el 41% de los alumnos consideran todos los valores de los datos en su argumentación, mientras que un 30% realiza una argumentación estadísticamente correcta. Además, se observa que la mayor parte de las argumentaciones están sujetas al contexto del problema, pero sin considerar los datos aportados.

García y Garrett (2006), elaboraron un cuestionario sobre medidas de tendencia central que fue aplicado a un grupo de 94 estudiantes de 17 años, que cursaban el ultimo nivel de la educación secundaria, esto con la finalidad de ver cómo actúan los alumnos cuando se les presentan problemas diseñados con preguntas abiertas comparándolas con preguntas de opción múltiple. Eligieron un ítem buscando observar la interpretación de los estudiantes sobre las distribuciones cuando se presentan en forma de gráfico. En el análisis de resultados observaron que gran cantidad de estudiantes en preguntas de opción múltiple han elegido la respuesta correcta, pero no son capaces de describir métodos razonables cuando se les presentan preguntas abiertas; inconsistencias en las afirmaciones que expresan ante una misma situación, pues usan criterios muy diferentes. Concluyen que, usar este tipo de preguntas abiertas y de opción múltiple revelan dificultades e incoherencias en las estrategias que usan los estudiantes al resolver estos problemas.

Pinho (2006 a y b), desarrolló una investigación con estudiantes de entre 14 y 15 años en la Ciudad de Bahía, Brasil. En su estudio analiza primeramente el contenido de seis libros de texto brasileños, sobre el significado de las medidas de tendencia central, siguiendo la misma metodología que Cobo (2003); describe los elementos de significado para determinar el significado institucional de referencia. Posteriormente aplica un cuestionario centrado en la comprensión de la definición de la media y concluye diciendo que la mayoría de los estudiantes no dominan completamente el concepto, únicamente son capaces de llevar a cabo los procedimientos, sin embargo, no comprenden la definición o propiedades de este parámetro.

Pinzón (2012), realizó su trabajo con 63 estudiantes de estadística para psicólogos de segundo semestre, de la Universidad Cooperativa de Colombia. Encontró que con la implementación de modelos pedagógicos por proyectos ayuda a mejorar la comprensión y representatividad de las medidas de tendencia central, pues los proyectos llevan a los estudiantes a introducirse en la investigación, que además permiten, definir los objetivos, concebir los instrumentos necesarios para el desarrollo de la investigación, seleccionar la muestra, categorizar, elaborar tablas, así como analizar e interpretar los datos. En lo que respecta a las medidas de tendencia central, confirma que, para que los estudiantes construyan la representatividad de las medidas de localización, requieren de un poco más de trabajo y direccionamiento en el proceso, por lo que se ven en la necesidad de apoyarse en similitudes que permitan hacer visible la medida a través de su representación y aplicabilidad. Sin embargo, la contextualización y la utilización del software favoreció la comprensión de las medidas de centralización debido a que los estudiantes confundían el concepto de moda con otras definiciones y al finalizar la investigación esa idea se redujo en un 40%, respecto a la idea de considerar a las medidas de tendencia central como promedios aumento y fue más visible en la media aritmética, la moda no lo consideraban un promedio, hasta después de la propuesta didáctica un pequeño porcentaje de estudiantes lo incluyó como parte de los promedios. También, se percataron que los alumnos fueron capaces de considerar a las medidas de centralización como valores centrales, y en especial a la media y mediana como representantes de un conjunto de datos, pues antes, los estudiantes sólo consideraban a la moda como representante del colectivo.

Saldarriaga (2012), trabajó con estudiantes de octavo grado de educación básica secundaria de la Institución Educativa Entrerríos (Antioquia). Los estudiantes realizaron trabajos contextualizados de acuerdo a su región, estos permiten tener interacción constante con el medio que los rodea, por lo que logran reconocer que la estadística se puede aplicar en diferentes contextos. Respecto a los resultados de las medidas de centralización, la autora señala que para los estudiantes resulta difícil comprender las definiciones de media, mediana y moda que son diferentes, en algunos casos se pueden utilizar las tres, o en otros sólo es conveniente utilizar una. También, mostró que las medidas de tendencia central están por fuera del currículo académico de la secundaria, ya que los estudiantes presentaron dificultad a la hora de dar la respuesta, porque aunque algunos sabían acerca de la media desconocían por completo la mediana y la moda.

Sayritupac (2013), este estudio fue realizado con un grupo de 49 alumnos universitarios de primeros ciclos de la carrera de Humanidades de la Pontificia Universidad católica del Perú, entre las edades de los 16 hasta 19 años. Los resultados muestran, que los libros de texto analizados no son del todo adecuados, aun teniendo gran variedad de problemas contextualizados, pues enfatizan en el algoritmo de cálculo y no en el concepto y sus propiedades. Del mismo modo, se encontró que las mayores dificultades se presentaron, principalmente, con problemas que están relacionados con saber reconocer si la media es o no un buen representante de un conjunto de datos en presencia de valores atípicos; además, resulta difícil para los estudiantes obtener conclusiones a partir de dos grupos de datos cualitativos ordinales; también la interpretación adecuada de la media en un contexto específico. Al mismo tiempo, se evidenció que los alumnos no identifican la presencia de un valor atípico en una distribución, o si lo hacen, no reflexionan sobre el efecto que causa la presencia del mismo sobre la media. Por otro lado, ningún alumno fue capaz de usar la mediana para comparar dos grupos de datos ordinales. Un error frecuente de parte de los alumnos fue considerar que la media debe ser necesariamente uno de los datos de la distribución, por último usan términos como "número central" o "mayoría", mostrando confusión al interpretar la media con la mediana o la moda.

Rodríguez, Maldonado y Sandoval (2016), realizaron un estudio con 142 estudiantes de Pedagogía en Matemática de dos universidades del centro sur del país pertenecientes al Consejo de Rectores de Universidades Chilenas (CRUCH). Encontraron que únicamente un 14.3% de los alumnos de la Universidad 1 son capaces de calcular correctamente el promedio aritmético en presencia de un dato atípico, aumentando este porcentaje a un 20% en la Universidad 2. En cuanto al cálculo del promedio aritmético en presencia de datos agrupados con información explicita, observaron que solo un 45.2% de estudiantes de la Universidad 1, calculan correctamente el promedio aritmético, mientras que en la Universidad 2 el porcentaje de logro es de un 73.3%. Respecto al cálculo de la mediana en relación al número de observaciones, en el caso de un número impar de datos el 83.3% de los alumnos de la Universidad 1 realizan correctamente su cálculo, sin embargo solo un 80% es capaz de hacerlo en la Universidad 2, cuando el número de observaciones es par desciende la capacidad de cálculo aproximadamente en un 7% para cada universidad. La media ponderada resulta de difícil cálculo para los estudiantes de las dos universidades. De manera general, los estudiantes realizan los cálculos numéricos sin profundizar en las propiedades, es decir, sin llegar a la reflexión sobre los significados de las medidas de tendencia central.

También, a lo largo de esta revisión de investigaciones se han encontrado aquellas centradas en el análisis de una sola medida de tendencia central. En lo que sigue se citan algunos trabajos enfocados al estudio de la media aritmética.

Campbell (1974), en su trabajo muestra que a menudo se tiende a situar la media en el centro del recorrido de la distribución, considerándola como propiedad de una distribución simétrica. Esto pocas veces es comprendido por los estudiantes, pues eligen como mejor representante del conjunto de datos a la media. Además, no conciben la idea de que si se presenta el caso de una distribución muy asimétrica hay un efecto en la media, esta se traslada hacia algún extremo, por lo que la moda o la mediana pueden ser la medida representativa del conjunto de datos. Por lo que se observa, que los estudiantes eligen la media sin tomar en cuenta la simetría de la distribución de los datos o la existencia de valores atípicos.

Mevarech (1983), llevó a cabo una investigación con 103 estudiantes universitarios de primer curso. Observa que los estudiantes creen que la media tiene la propiedad asociativa, cuando tienen que hallar la media de un conjunto grande de números, lo dividen en partes hallando primero la media de cada parte y luego promediando el resultado obtenido, además toman en cuenta el cero en la estimación de la media, considerándolo un elemento neutro o igual se piensa que la media debe pertenecer al conjunto de datos, a esto se le conoce como la propiedad de clausura.

Strauss y Bichler (1988), llevaron a cabo en Israel una investigación con 80 estudiantes de 8, 10, 12 y 14 años, distinguiendo un conjunto de propiedades:

a) La media es un valor comprendido entre los extremos de la distribución.
b) La suma de las desviaciones de los datos respecto de la media es cero, lo que hace que sea un estimador insesgado.
c) El valor medio está influenciado por los valores de cada uno de los datos. Por ello, la media no tiene elemento neutro.
d) La media no tiene por qué ser igual a uno de los valores de los datos.

e) El valor obtenido de la media puede ser un número no entero (ello puede no tener sentido para la variable considerada), como cuando se dice que el número medio de hijos es 1.1.
f) Hay que tener en cuenta los valores nulos en el cálculo de la media.
g) La media es un "representante" de los datos a partir de los que ha sido calculada.

Los autores encontraron que la mejora en la comprensión de las propiedades se da de acuerdo a la edad, también las diferencias de dificultad, las propiedades que resultaron ser más sencillas en su comprensión son la a, c, y d, mientras que las b, e, f y g fueron las que destacaron como más difíciles. Por otro lado, enfatizaron que la representatividad es un obstáculo que deben vencer los estudiantes cuando tienen que estudiar la media aritmética.

Goodchild (1988), en su estudio encontró que los estudiantes no relacionan adecuadamente un promedio con una distribución. Además, cuando se le solicita construir distribuciones hipotéticas, y al hacerlas se percató que estas no tenían forma acampanada como la distribución normal, esto debido a la falta de comprensión de la media como medida de centralización. Finaliza concluyendo que los estudiantes entienden mejor las propiedades de la media asociadas a la localización que las asociadas a la distribución.

León y Zawojeski (1991), analizaron el efecto de la edad sobre la comprensión de algunas propiedades. Para ello, trabajaron entrevistas con niños entre 8 y 14 años, y encontraron que existe una gran influencia de la edad sobre la comprensión de la media, además observaron que la contextualización de tareas facilita su resolución. Respecto a las propiedades como son la suma de las desviaciones respecto a la media es cero, que la media es un valor representativo de los valores promediados o que hay que tener en cuenta los valores nulos en el cálculo de la media, siguieron siendo muy abstractas para una proporción importante de estudiantes de 14 años.

Zawojewski y Roth (1990), investigaron el efecto de la edad en la comprensión de las siete propiedades enlistadas por Strauss y Bichler (1988). En sus resultados mencionan haber encontrado una influencia importante de la edad sobre la comprensión de la media, así como la contextualización les facilita la resolución de las tareas, pues permite el uso de otros conocimientos y capacidades. Por otro lado, siguen siendo muy abstractas para una gran parte de alumnos de 14 años las propiedades tales como, la suma de las desviaciones respecto a la media es cero, la media es un valor representativo de los valores promediados o que hay que tener en cuenta los valores nulos en el cálculo de la media.

Li y Shen (1992), analizaron proyectos de estadística realizados por alumnos de secundaria. Encontraron que los estudiantes cuando calculan la media a partir de una tabla de frecuencias con datos agrupados en intervalos, olvidan que cada uno de los grupos debería ponderarse de diferente modo, por lo que, solo calculan la media de las marcas de clase.

Tormo (1993), en un trabajo realizado con estudiantes de entre 11 y 16 años, los resultados muestran que los alumnos identifican de manera inadecuada un valor diferente de la media como valor más probable. Esto debido a que las distribuciones y ejemplos en los libros de texto, enfatizan la representatividad de las muestras, pero no destacan la idea de variabilidad.

Cai (1995), en un estudio realizado con 250 estudiantes de 6° curso entre 11 y 12 años en Estados Unidos, se revela que el 88% de estudiantes de la muestra conocen el algoritmo de cálculo de la media, mientras que el 50% han sido capaces de aplicar este concepto al momento de resolver problemas abiertos. Con estos resultados la autora señala que se le está dando más importancia a la enseñanza de algoritmos dejando en detrimento la comprensión de la media y demás cuestiones importantes. Al mismo tiempo, encontró que gran número de estudiantes de entre 12 y 13 años usaron adecuadamente el algoritmo para calcular la media, pocos alumnos lograron dada la media, encontrar un valor desconocido en un conjunto pequeño de datos, en otras palabras, son capaces de invertir el algoritmo. No obstante, la mayoría de los estudiantes optó por el método de ensayo y error, para poder encontrar la distribución dada la media.

Mokros y Rusell (1995), en una investigación, realizada en Australia y Tasmania con 21 estudiantes de 9, 11 y 13 años. Analiza la idea de la representatividad de la media aritmética para describir un conjunto de datos de manera concisa. Los resultados muestran que los estudiantes tenían una noción de la representatividad de los datos que estudiaron. Los autores señalan que la introducción prematura del algoritmo de cálculo de la media, ocasiona que los alumnos pierdan el significado de representatividad. Además de que poseen estrategias para la resolución de problemas sobre promedios, las olvidan cuando conocen el algoritmo de cálculo. Concluyen señalando que la media aritmética es un objeto matemático muy complejo oculto tras un sencillo algoritmo de cálculo y debería ser enseñado cuando los estudiantes hayan desarrollado la idea de representatividad.

Estrada (2002), trabajó con un grupo de 367 futuros profesores de educación primaria. Encontró que existen dificultades por parte de los estudiantes en conceptos estadísticos elementales, tal es el caso de la media, la interpretación de gráficos, lecturas de tablas, entre otros. En cuanto a las medidas de posición central, una gran mayoría de los alumnos usa la idea de media como mejor estimación de una cantidad desconocida en presencia de errores de medida, y un pequeño porcentaje (7%) entiende que el cálculo de la media es afectado por valores atípicos. Un porcentaje alto de la muestra (73%) fue capaz de calcular la media partiendo de un gráfico que representaba la distribución de datos; sin embargo, el resto no logro interpretar adecuadamente el grafico o no uso las frecuencias en la estimación de la media. Además, se dio cuenta que solo un 6% de los futuros profesores no usa las medidas de tendencia central para hacer una comparación de dos conjuntos de datos. También observó que un 33% de los alumnos logra interpretar adecuadamente la frase "número promedio de hijos es igual a 1.2", con el que se detectaron algunas dificultades al considerar la media como una operación interna o situarlas próximas a la mediana y moda.

Garret y García (2005), en un trabajo sobre promedios con alumnos españoles y angoleños de educación secundaria, los autores encontraron algunas estrategias incorrectas empleadas en el cálculo de la media, de las cuales destacan las siguientes: dividir el resultado obtenido por la suma de los valores de la variable, generando confusión entre frecuencia y valor de la variable; usar adecuadamente el algoritmo pero redondeando el resultado encontrado; seleccionar un dato cualquiera como media, para dar respuesta. Además de estas dificultades, los estudiantes responden con un valor a veces correcto otras incorrecto, sin justificar el procedimiento, lo cual limita el estudio para analizar sobre las estrategias que usa.

Mary y Gattuso (2005), realizaron un estudio con 638 estudiantes de entre 14 y 16 años, de los cursos 2°, 3° y 4° de secundaria. En los resultados encontraron que el contexto de los problemas y tipo de promedio tienen efecto en el comportamiento de los estudiantes y cómo conciben ideas no correctas en algunas situaciones. Además, es fácil la comprensión del efecto que tiene sobre la media al incluir un dato igual a cero, pero no lo es el efecto de retirar un dato igual a cero. También, es fácil de comprender el agregar un dato igual a cero, en comparación de remplazar un dato diferente de cero por otro igual a cero. La primer idea encontrada, el cero neutro, esto es, una aplicación errónea de una propiedad de la estructura del grupo, "quitar un dato nulo, no cambia la media". La segunda idea viene de una abusiva generalización de una representación de la media como resultado de un reparto igual.

Estrella (2016), realiza un estudio con 27 profesores de educación primaria, en formación continua, de la Pontificia Universidad Católica de Valparaíso. Este estudio pretende conocer la comprensión de la media en presencia de valores atípicos y para ello, lo realiza a partir del análisis de las respuestas abiertas. La autora observó que un 92% de los estudiantes no tomaron en cuenta el efecto de los valores atípicos sobre la media, es decir calculan la media sin desechar el valor atípico, y no consideran este valor como un error de medida que influye en la estimación. Antes de finalizar la actividad, los profesores realizan una lectura que aborda el conocimiento de la media como el mejor estimador en presencia de errores de medición y la toma de conciencia de los valores atípicos. Los resultados obtenidos después de la lectura muestran que las respuestas correctas aumentan un 7.4%; al mismo tiempo, aumenta la respuesta incorrecta de considerar todos los valores. Por lo que, se concluye con esto que no hay interpretación después de la lectura individual del concepto de dato atípico (como concepto estadístico) ni del efecto de los valores extremos en el cálculo de la media (desde la algoritmia de la media). No obstante, se desarrolla la idea de que la media es mejor estimador que la moda.

Por otro lado, se han encontrado algunas investigaciones que estudian los niveles de comprensión y dificultades de los estudiantes ante el uso de la media ponderada. A continuación se describen.

Pollatsek, Lima y Well (1981), en un trabajo sobre medias ponderadas encontraron que, estudiantes universitarios no ponderan correctamente los valores al plantearles un problema de media ponderada, y optan por calcularlo con la media simple, debido a que no son capaces de diferenciar las situaciones donde deberían calcular una u otra. Además, indican que los estudiantes tienen la idea de que la media muestral debe ser idéntica a la de la población de donde se ha tomado la muestra, por lo que no ven la variabilidad aleatoria de la media en diferentes muestras.

Gattuso y Mary (1996), menciona las dificultades que ha encontrado en los alumnos en la resolución de problemas con el algoritmo de cálculo de la media ponderada, observó que los alumnos no son capaces de ponderar de forma correcta los valores al momento de resolver problemas donde se debe estimar la media ponderada, también comenten el error de eliminar los valores nulos, y estos errores se siguen viendo igual en alumnos universitarios. Destaca el uso de diferentes estrategias en la resolución de problemas, de acuerdo al nivel de educación y su impacto en los resultados y concluye que a mayor edad y con buena instrucción el estudiante mejora el empleo de algoritmos y cada vez más usan con mayor frecuencia notación y métodos algebraicos. No obstante, aquello problemas que necesitan un conocimiento conceptual más profundo son resueltos con mayor frecuencia por los estudiantes de secundaria, esto porque los de primaria usan estrategias intuitivas y los universitarios hacen uso mayormente de fórmulas algebraicas o definiciones formales.

Gattuso y Mary (1998), estudian la evolución de la comprensión del algoritmo de cálculo de la media ponderada de estudiantes de 14 a 17 años de edad, para lo cual, utilizaron problemas con diferentes contextos y forma de representación. Las actividades presentadas fueron: Cálculo de medias ponderadas, efecto que el cambio de un dato produce sobre la media, hallar un valor que falta en un conjunto de datos para obtener un promedio dado y analizar las desviaciones con respecto a la media. Los resultados del trabajo muestran algunos errores que cometen los alumnos como, por ejemplo, utilizar fórmulas incorrectas para calcular la mediana, o no tener en cuenta el peso de los distintos valores de la variable. También, se percataron de que, la forma de presentar los datos influye en los resultados obtenidos, aun cuando los alumnos sean competentes en la interpretación de datos en diversos formatos, su habilidad en el cálculo de las medidas de tendencia central es mayor cuando los datos se presentan agrupados en tablas que cuando se dan en forma gráfica.

Por último, se presentan algunos trabajos que se centran en el estudio de la mediana, como medida de tendencia central, que a menudo resulta difícil la comprensión e interpretación de los estudiantes.

Barr (1980), llevó a cabo un estudio con alumnos de entre 17 y 21 años de edad, según los resultados que se obtuvieron muestran que los estudiantes interpretan la mediana como el centro de "algo", sin embargo, no logran comprender a qué se refiere ese "algo". Esto debido a que, los alumnos no comprenden que una tabla de frecuencias únicamente es un resumen de los datos, considerada una representación alternativa, y no son capaces de pasar de la tabla a la lista de valores. Además, los alumnos no consideran a la mediana como un estadístico que se refiere al conjunto de datos ordenados, pues si los datos se les dan en forma de lista, no conciben la idea del por qué hay que ordenarlos para calcular la mediana.

Zawojewski (1986) en su estudio encontró que los estudiantes tienen dificultad en tomar la mediana como representante del conjunto de datos y en crear conjuntos que tengan como mediana una fijada de antemano. Gran parte de los estudiantes de esta investigación usaron la mediana como uno de los datos, lo que señala que para ellos no existe diferencia entre las medidas de centralización y los datos del conjunto. Además, los estudiantes se apegan sólo en ciertas reglas aprendidas sobre la definición de las medidas de tendencia central.

Schuyten (1991), afirma que alumnos universitarios encuentran difícil aceptar que pueden usarse dos algoritmos diferentes de cálculo para una misma medida de tendencia central, tal es el caso de la mediana, cuando tiene un número par o impar de datos. También, se les dificulta comprender que pueden obtenerse valores distintos para el mismo parámetro, al variar la amplitud de los intervalos de clase. Además, los estudiantes no son capaces de comprender como pasar de la definición de la mediana “valor de la variable estadística que divide en dos efectivos iguales a los individuos de la población supuestos ordenados por el valor creciente del carácter”, hasta su cálculo, con base en la gráfica de frecuencias acumuladas, donde es necesario usar fórmulas de interpolación, que no siempre son comprendidas por los estudiantes.

Estepa (1993), menciona que cuando los estudiantes realizan los cálculos de la mediana, partiendo de una gráfica de las frecuencias acumuladas, se le presentan diferentes obstáculos. En caso de las variables discretas, al construir la gráfica tomo forma de escalera, esto es, una gráfica discontinua, y los estudiantes no usan a menudo este tipo de gráficos por lo que les genera dudas el valor de la mediana, pues hay más de un valor de la variable que podría ser la definición de la mediana. Por otro lado, con las variables continuas, según la amplitud del intervalo, tiende a variar el valor de la mediana, asimismo con los demás estadísticos. En este caso los estudiantes tienen que interpolar para encontrar el valor de la mediana, lo cual genera errores por fallo de razonamiento proporcional, y el dominio en desigualdades que aparecen asociadas a la definición de mediana y a su cálculo son deficientes.

Cobo y Batanero (2000), realizan un estudio con la finalidad de identificar dificultades específicas que los estudiantes de Enseñanza Secundaria pueden encontrar en su comprensión conceptual y en su uso en la resolución de problemas de análisis exploratorio de datos. Dentro de las principales dificultades que se encuentran son que el algoritmo de cálculo de la mediana no es único, pues depende del tipo de datos, así como de la manera en que se presentan. Entonces, los estudiantes deben aprender varios algoritmos y cuando aplicarlos. Además, el valor obtenido no siempre es único, en consecuencia genera problemas de comprensión a los estudiantes.

Jacobbe (2008), realizó un trabajo sobre la media y la mediana con un grupo de profesores estadounidenses, que participaron en entrevistas por 18 meses y 12 observaciones de clases de cada profesor. Concluye que la gran parte de los profesores no están familiarizados con los conceptos y procedimientos que deben enseñar a sus estudiantes, pues de las preguntas que se les hizo en relación con la media y la mediana, respondieron correctamente entre el 63% y el 75%. De los profesores de la muestra, a dos les resulto difícil explicar el significado de la media y la mediana, pero si conocían el algoritmo de cálculo. Además, uno de ellos no fue capaz de explicar la diferencia entre media y mediana. El autor señala que los profesores estadounidenses no cuentan con conocimientos suficientes para desarrollar en sus alumnos niveles de comprensión adecuados sobre las medidas de tendencia central que fijan los estándares americanos (NTCM, 2000).

Mayén (2009), estudió el significado personal que los estudiantes mexicanos (518) de educación secundaria (162) y bachillerato (356), asignan a las medidas de tendencia central. Los resultados de esta investigación tienen especial énfasis en la mediana, entonces, algunas dificultades encontradas son que los estudiantes confunden las definiciones de media y mediana, mientras que el cálculo y la interpretación de la mediana resulta más complicada que el de la media, pues resulta fácil resolver problemas de media simple. Además, los estudiantes no son capaces de identificar de forma correcta los problemas que involucran el uso de la media ponderada, causando conflictos y errores en su solución. Asimismo, encontró dificultades en el cálculo de la media, mediana y moda de un conjunto de datos agrupados en intervalos y presentados en una tabla frecuencias, así como los presentados en gráficos.

2.3.2. Investigaciones sobre medidas de dispersión

En esta sección se describen algunas investigaciones relacionadas con el aprendizaje de la variabilidad, errores y dificultades de comprensión de este concepto. El orden de aparición de estas investigaciones previas es de acuerdo al año y según el nivel académico.

Las primeras investigaciones que a continuación se describen fueron realizadas para el nivel básico y bachillerato, donde se encontraron más investigaciones en relación al concepto de variabilidad y en diferentes contextos.

Shaughnessy, Watson, Moritz y Reading (1999) desarrollaron un estudio con 324 estudiantes de los grados 4°,5°,6° y 12° (9 a 12 años y 17 a 19 años) que pretendía investigar las ideas sobre variabilidad. Los autores usaron una actividad llamada “actividad de los dulces”. En la que se les proponía a los participantes que extrajeran 10 dulces y predijeran cuántos de ellos serían de color rojo. La predicción se repetía cinco veces usando tres formatos diferentes: en la versión del *rango*, se les solicitó el número máximo y mínimo de caramelos en los cinco experimentos. En cuanto a la versión de *elección*, se les propuso seleccionar el resultado más probable de la lista de resultados, entre cinco listas posibles. Por último, en la versión de *lista*, se les pidió que escribieran una lista posible de resultados de los cinco experimentos.

En el mismo sentido, Torok y Watson (2000), dieron continuidad al trabajo de Shaughnessy *et al*. (1999), para lo cual, trabajaron con estudiantes de 4°, 6°, 8° y 10° grado (entre 9 a 17 años de edad), quienes eran alumnos regulares y elocuentes; estos alumnos fueron entrevistados de forma individual durante 45 minutos. Dicha entrevista tenía la finalidad de explorar su comprensión de la variación así como la manera en que la relacionaban con diferentes problemas.

Al examinar la comprensión de la variación de los estudiantes en contextos reales, y con el problema hipotético de los dulces, los autores encontraron que los estudiantes mayores y con más instrucción en probabilidad y estadística revelaron un alto nivel de comprensión; pues consideran el conjunto entero de resultados, en cambio los más jóvenes tendían a enfatizar resultados individuales. Los estudiantes mayores elaboraron mejores representaciones gráficas. Los autores consideran que la exploración de la variación en el salón de clases puede y debe tomar lugar varios años antes de la introducción de la desviación típica.

Por otro lado, Lehrer y Schauble (2002), exploran ideas de error en la medición y la desviación de la media en estudiantes pre-universitarios, este estudio fue desarrollado con 22 docentes. El estudio se desarrolló en un lapso de 2 años, y tareas de medición para ayudar a los estudiantes a percibir el error de medición de los datos distribuidos alrededor de un valor verdadero desconocido.

Shaughnessy y Ciancetta (2002), realizaron un estudio en el que entrevistaron a estudiantes de 6° a 12° grado (de 11 a 18 años), a quienes se les pidió resolver un problema y justificar su respuesta. El problema tenía la finalidad de que los estudiantes imaginaran una serie de juegos y predijeran el porcentaje de victorias y derrotas, para después recopilar datos jugando con la ruleta. Los autores encontraron que sólo 20% de los estudiantes de 6°

a 8° grado, de la muestra dedujo que la oportunidad de ganar era inferior a $\frac{1}{2}$. Los alumnos de 9° grado tuvieron un mejor desempeño en comparación con los grados anteriores, alrededor de 43% resolvieron el problema de forma correcta; 28% de los estudiantes de 9° a 11° grado dieron respuestas correctas y los estudiantes de 12° grado lo hicieron aún mejor, pues el 90% dio la respuesta correcta.

En el mismo orden de ideas, Watson y Kelly (2002), en un estudio con alumnos de 3º a 9° grado (8 a 15 años), buscan observar el razonamiento acerca de la influencia de la variación al resolver problemas relacionados con el azar y de análisis de datos. Para esto, se desarrollaron 10 lecciones durante 8 semanas, se aplicó una prueba previa y una posterior, que incluían los temas: variación en azar, variación en datos y gráficos, y variación en situaciones de muestreo.

Los resultados obtenidos muestran que los estudiantes de tercer grado no lograron los niveles de éxito en la prueba posterior en comparación con los estudiantes de grados superiores, por lo que sugieren incentivar, tanto el interés en la variación del patrón esperado en resultados aleatorios como en el patrón esperado por sí mismo. Además de que es importante presentar ideas de muestreo en este nivel ya que los estudiantes tienen buenas intuiciones al respecto. Según los autores, los estudiantes de tercer grado pueden aprender variación, pues logran tener un avance significativo en la comprensión de este tema.

Otra investigación importante es la de Petrosino, Lehrer, y Schauble (2003), realizada con estudiantes de 9-10 años de edad (4º grado), propusieron actividades de mediciones, en el salón de clases, con la finalidad de que se involucren en la recolección de datos, y reflexionaran sobre diferentes maneras de hacer esto, de tal forma que llegaran a plantear problemáticas de interés para la elaboración de análisis estadísticos. Por lo que, consideraban la idea de distribución, construida y visualizada mediante el proceso de recolección de mediciones; esto les permitió mostrar y estructurar la variación entre observaciones. Al mismo tiempo, el enfoque de su investigación permitía que los estudiantes visualizaran los errores y las fuentes de dichos errores.

Watson, Kelly, Callingham y Shaughnessy (2003), estudiaron la percepción acerca de la variación, con un grupo de 189 estudiantes de 7° grado y 197 de 9°grado, se les pidió resolvieran el siguiente problema: *"Una pirinola 50-50 (blanco y negro) fue girada 50 veces y el resultado del número de veces que cayó sobre la parte negra fue registrada"* (p. 21). Para esto, se les proporcionó a los alumnos tres resultados hipotéticos de esos experimentos y se les solicitó que eligieran cuál de los tres les parecía adecuado como resultado de ese experimento. Finalmente, los resultados mostraron que 52.6% de los estudiantes de los grados 7° y 9° consideraron correctamente que las dos primeras distribuciones eran inventadas y la última era real.

Watson y Kelly (2003) en una investigación realizada con estudiantes de 3°, 5°, 7° y 9° grados (8 a 15 años), buscaban explorar la comprensión de los términos muestra, aleatoriedad y variación. Solicitaron a los participantes que definieran y dieran un ejemplo de la palabra muestra. En el mismo sentido, se les preguntó sobre las palabras aleatoriedad y variación a 379 estudiantes de 7° y 9° grados. Los resultados de la investigación muestran que a lo largo de los grados escolares existe mejoría en las ideas en torno a la noción del término "muestra". Por otro lado, se reportó poco cambio en el manejo de los términos aleatoriedad y variación del 7° al 9° grados. No obstante, los estudiantes de 9° grado tuvieron un desempeño mejor que los de 7° grado.

Por su parte, Bakker (2004), desarrolló un trabajo encaminado a promover el razonamiento acerca de la variabilidad, muestreo, datos y distribuciones, en el estudio participaron 30 estudiantes de 7° y 9° grados, en él se discute el uso de un lenguaje informal para articular ideas sobre variación. Los estudiantes utilizaban expresiones como "promedio", "rango" y "dispersión" de una forma no convencional durante las actividades, mientras que términos no estadísticos como "mayoría", "semicírculo" o "pirámide" también fueron utilizados para expresar algunas características estadísticas de distribuciones.

Ben-Zvi (2004), investigó sobre el razonamiento de dos estudiantes de secundaria acerca de la variabilidad, para esto con una actividad les pidió que compararán las longitudes de apellidos hebreos y americanos utilizando una hoja de cálculo (Excel). El resultado de la actividad mostró que ambos estudiantes evolucionaron desde utilizar la información de local hasta el punto de vista global de describir y explicar la variabilidad entre los grupos.

En cuanto a, Lehrer, Kim y Schauble (2007) estudiaron el desarrollo del concepto de variación en el uso de medidas con el software TinkerPlots. En esta investigación participaron estudiantes de 5° y 6° grado, en la actividad midieron el mismo objeto y se dieron cuenta que su longitud real se obtiene cuando no exista algún error. Se les pidió a los estudiantes que establecieran un grado de precisión para sus mediciones (de cercanía o aceptación). Aunque su criterio de cercanía era no convencional, comenzaron a ver la variación como una relación entre datos y su distancia a un centro, que a menudo es la media o mediana de la muestra.

Las diferencias en las soluciones de los alumnos apoyaron la idea central de que muchas mediciones (datos), son necesarias para hacer una estimación de una cantidad desconocida. Los autores concluyen que los alumnos en estas etapas pueden aprender diferentes formas de medir la variación estadística como el rango, el rango intercuartil, y la desviación estándar.

Por último en esta clasificación, la investigación más reciente es la de Orta (2014), quien realizó un trabajo con estudiantes de tercer año de educación secundaria (14 y 15 años) de una escuela pública de la Cuidad de México. Este trabajo tenía la finalidad de explorar el razonamiento de los estudiantes sobre la variabilidad estadística cuando resuelven problemas de comparación de conjuntos de datos y, para formular la solución, deben interpretar la dispersión en términos de riesgo.

Los resultados del trabajo señalan que aunque los estudiantes sean capaces de identificar la variabilidad de los conjuntos de datos, la mayoría no la interpreta en términos del contexto propuesto (riesgo), tampoco la consideran para la toma de decisiones. Sin embargo, algunos estudiantes lograron asociar centro, dispersión e interpretar la variación en términos de riesgo; cuando realizan una argumentación adecuada a cerca de la decisión que toman.

Finalmente, Orta (2014) sugiere que en la educación secundaria se trabajen problemas mediante las cuales los estudiantes articulen e interpreten en diferentes contextos los conceptos de centro y variación, y las situaciones de riesgo pueden ser una buena opción para considerarlas.

En el mismo orden de ideas, un estudio realizado con universitarios fue el de Delmas y Liu (2005, 2007), que trabajaron con una serie de investigaciones sobre cómo los estudiantes universitarios aprenden desviación estándar, en sus trabajos sugieren que los conceptos de distribución, media y desviación respecto a la media son fundamentales para construir la noción de desviación estándar.

En ese sentido, Makar y Confrey (2005) entrevistaron a 17 futuros profesores de Educación Secundaria de matemáticas y ciencias, con la finalidad de analizar cómo usaban la dispersión cuando comparaban dos distribuciones. Pudieron observar cómo los futuros profesores expresaban importantes ideas acerca de la dispersión, empleando un lenguaje estándar y no estándar. Algunos ejemplos de expresiones estándar en estadística, usadas por los entrevistados, fueron: proporción, media, máximo/mínimo, tamaño muestral, valores atípicos, rango, forma y desviación típica. Para expresar ideas de variación, surgieron dos categorías de términos no estándar, que correspondían a ideas intuitivas sobre dispersión (agrupado, disperso) y distribución (triadas, agrupaciones modales). El uso de estos términos aumento a lo largo de las entrevistas posteriores.

Para finalizar esta sección, Silva y Coutinho (2008) realizaron un estudio con nueve maestros de matemáticas en servicio sobre su razonamiento acerca de la variación con una distribución univariada. Se les pidió a los profesores recolectar datos a través de una encuesta, y crear una distribución.

Una vez organizados los datos en una tabla de frecuencias y representados en un histograma, se solicitó a los profesores pensar sobre una forma distinta de representar el conjunto de las edades. Los autores analizaron el razonamiento de los maestros acerca de la variación cuando estos analizaron la distribución de las edades. Encontraron diferentes niveles de razonamiento de la variación, sin embargo, ninguno fue capaz de integrar el razonamiento sobre la media con el significado de la desviación respecto a la media, ni estimar la frecuencia en el intervalo de *k* desviaciones típicas desde la media.

Finalmente el estudio muestra que el razonamiento predominante de los maestros sobre la variación es verbal, esto les impide enseñar a sus estudiantes el significado de medidas como la desviación típica y se limitan únicamente a la enseñanza de los algoritmos de cálculo.

A modo de conclusión, se observa que el estudio del concepto de variabilidad está en expansión, y los resultados de estas propuestas en cuanto a la enseñanza de este concepto estadístico, son buenas pero, aún existen líneas para trabajar. Los razonamientos entorno a la variabilidad son bajos, tanto en estudiantes de nivel básico, superior y en profesores en servicio, dado que se convierte en un concepto difícil de interpretar. La omnipresencia de la variabilidad ante nuestra realidad cambiante, exige cada vez describir, cuantificar y explicar cambios producidos por fenómenos naturales o sociales.

CAPÍTULO 3
METODOLOGÍA

3.1. Introducción

Este capítulo describe, la metodología de análisis utilizada para el material audiovisual con la finalidad de determinar el contenido matemático en que se basa el programa de Probabilidad y Estadística.

De esta forma, se logrará determinar el significado institucional tanto de las medidas de tendencia central como de la variabilidad y así mismo, detectar significado personal que tienen los estudiantes respecto a estos conceptos.

3.2. Metodología de análisis del contenido

Para realizar este análisis se tomó el video correspondiente al Programa Tres “La estadística descriptiva” de la asignatura de Probabilidad y Estadística I, para alumnos del Quinto Semestre del Telebachillerato. Este material audiovisual, ha sido elaborado por la Dirección General de Televisión Educativa (DGTE) en el año 2014, para la extensión de los servicios educativos de la Secretaria de Educación Pública.

El análisis de contenido de los videos se realizó adecuando el método utilizado por Cobo (2003) y Mayén (2009), el cual consiste en los siguientes pasos:

1. Observación detallada del programa audiovisual para quinto semestre, para comparar el contenido con la lista de elementos de significado y determinar su presencia en el video.
2. Selección y captura de imágenes para ejemplificar cada elemento de significado que se ha encontrado, además de una trascripción de los diálogos que se desarrollan.
3. Elaboración de un resumen escrito en el que se analizan los campos de problemas, definiciones, propiedades, representaciones, procedimientos y argumentos identificados en el video.
4. Elaboración de conclusiones sobre el significado de referencia en el video utilizado por los alumnos de quinto semestre.

El video analizado cuya duración es de 16.53 minutos tiene el propósito de aplicar las medidas de tendencia central y medidas de variabilidad en diversas situaciones, así como el comportamiento de una población a partir de las mismas. La estructura del video está dividida en dos secciones en las que se van intercalando escenas, durante el tiempo que dura video, la primer sección es sobre el análisis de una situación real que plantea un grupo de estudiantes quienes están realizando una investigación sobre el tipo de música que escuchan las personas y la segunda sección es de un profesor que con uso de imágenes va explicando los contenidos conceptuales y los ejemplifica usando el problema del peso en kilogramos de 8 jugadores de un equipo de futbol.

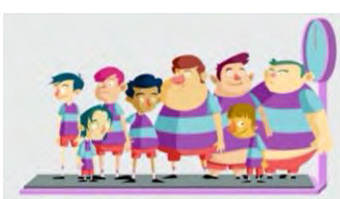
3.3. Campos de problemas

En esta sección se describen los campos de problemas identificados en el video.

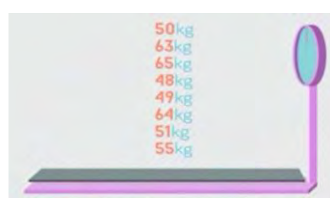
3.3.1. Campos de problemas de la media

1. *Estimar la media de un conjunto de datos.* En este campo de problemas, la idea de la media es el promedio de un conjunto de datos, como una medida que nos proporcione una imagen más clara sobre el valor central de los datos que se está manejando. Se observa en la Figura 1, el ejemplo en el que pide saber el peso promedio que tiene el equipo de fútbol soccer.

“Imaginen a 8 jóvenes, son un equipo de fútbol soccer y nosotros somos los entrenadores, queremos hacer un entrenamiento y una dieta que sirva para todos, por lo tanto, necesitamos saber el peso promedio que tiene el equipo, ¿Cómo obtengo este dato?”



(Escena 1)



(Escena 2)

(Escena 3)

(Escena 4)

Figura 1. Cálculo del peso promedio de un equipo de fútbol soccer

2. *Estimar la media de un conjunto de datos agrupados en intervalos de clase.* La Figura 2, muestra la idea de este campo de problemas, calcular la media de un conjunto de datos agrupados en intervalos cuando se presentan en una tabla de frecuencias absolutas, en el que se involucra otros conceptos como el punto medio.

"Calcular la media aritmética o promedio del conjunto de datos que muestra el tiempo en minutos que las personas escuchan música".

Tiempo en minutos	Número de personas (frecuencia)
0 - 60	19
61 - 120	34
121 - 180	25
181 - 240	12
241 - 300	10
301 - 360	15
361 - 420	5
Total	120

(Escena 1)

Tiempo en minutos	Punto medio del intervalo	Número de personas (frecuencia)
0 - 60	30	19
61 - 120	90	34
121 - 180	150	25
181 - 240	210	12
241 - 300	270	10
301 - 360	330	15
361 - 420	390	5
Total		120

(Escena 2)

Tiempo en minutos	Punto medio del intervalo	Número de personas (frecuencia)	
0 - 60	30	19	30x19=
61 - 120	90	34	90x34=
121 - 180	150	25	150x25=
181 - 240	210	12	210x12=
241 - 300	270	10	270x10=
301 - 360	330	15	330x15=
361 - 420	390	5	390x5=
Total		120	162.5

(Escena 3)

Figura 2. Cálculo del promedio en un conjunto de datos agrupados en intervalos de clase del tiempo que las personas escuchan música

3.3.2. Campos de problemas de la mediana

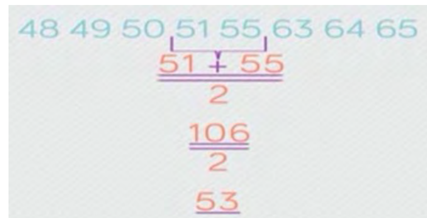
1. *Estimar la mediana, cuando la media no es suficientemente representativa.* Este campo de problemas se encontró cuando la media tiene un valor poco confiable. Se observa en la Figura 3 el ejemplo de los jugadores de fútbol soccer, donde se pide calcular la mediana para el conjunto de datos cuando es par, y seguidamente se incluye un valor y se solicita calcular la mediana para el conjunto de datos impar.

"Volvamos a nuestro ejemplo con los jugadores de fútbol, tenemos 8 jugadores los primeros pesan muy poco y los últimos pesan demasiado, en este caso aplicaremos la mediana"

(Escena 1)

(Escena 2)

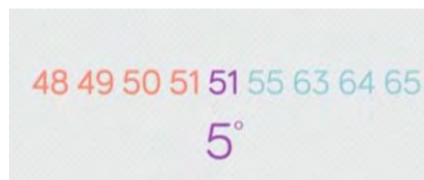
(Escena 3)



(Escena 4)

"si el número de nuestros datos muestrales no fuera par, sino impar, es decir, si en vez de tener 8 jugadores tuviéramos 9, …consideramos al nuevo integrante quien también pesa 51 kilogramos"

(Escena 1)

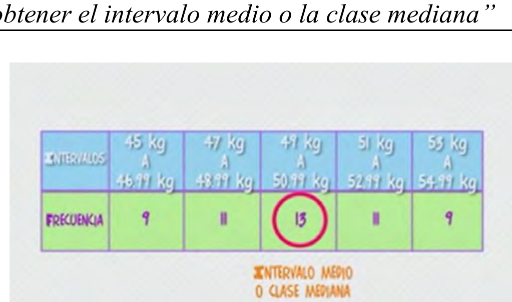


(Escena 2)

Figura 3. Cálculo de la mediana en una distribución de datos impar, del peso de los jugadores del equipo de fútbol soccer al añadir un nuevo integrante

2. *Estimar la mediana para un conjunto de datos agrupados en intervalos de clase.* Este campo de problemas, aparece muy brevemente en el video, en la Figura 4, se señala la clase mediana que se obtuvo en un conjunto de datos ordenados en una tabla de frecuencias acumuladas.

"si los datos están agrupados en intervalos o clases se calculan las frecuencias acumuladas de las clases o intervalos hasta obtener el intervalo medio o la clase mediana"



(Escena 1)

Figura 4 . Estimar la mediana para un conjunto de datos agrupados en intervalos de clase, del peso de un grupo de personas

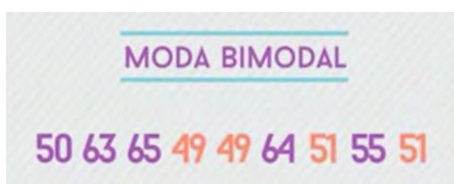
3.3.3. Campos de problemas de la moda

1. *Identificar mediante una simple inspección visual, qué cifra se repite.* Este campo de problemas presenta a la moda como el valor que aparece con más frecuencia en un conjunto de datos y puede estar en distintas representaciones. El ejemplo de la Figura 5, pide encontrar la cifra que se repite más veces y se modifica el ejemplo para determinarla cuando existe una moda, dos modas o ninguna moda.

<table>
<tr><td colspan="2">“veamos de nuevo a nuestros nueve jugadores con su peso anotado de bajo de ellos, vean que cifra se repite”</td></tr>
<tr><td>


(Escena 1)</td><td>


(Escena 2)</td></tr>
<tr><td colspan="2">


(Escena 3)</td></tr>
<tr><td colspan="2">“Si en lugar de pesar 48 kilogramos uno de los muchachos pesara 49”</td></tr>
<tr><td colspan="2">


(Escena 4)</td></tr>
<tr><td colspan="2">“Si no se hubiera integrado al grupo el portero y el jugador número 4 no hubiera subido de peso y todo estuviera como al principio”</td></tr>
<tr><td colspan="2">

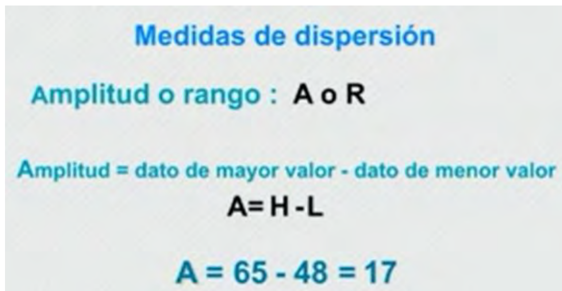

(Escena 5)</td></tr>
</table>

Figura 5. Identificar mediante inspección la moda de un grupo de jugadores de fútbol soccer

3.3.4. Campos de problemas de la amplitud o rango.

1. *Estimar el rango de un conjunto de datos.* En este campo de problemas, el rango nos permite conocer la dispersión de la distribución, sin embargo, sólo considera los datos extremos de un conjunto de datos. La Figura 6, muestra con el ejemplo del equipo de fútbol soccer el cálculo del rango de un conjunto de datos.

"La amplitud o rango del ejemplo del peso de los 8 jugadores"



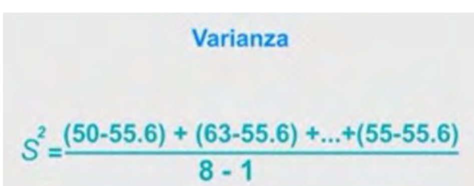
(Escena 1)

Figura 6 . Cálculo del rango de un conjunto de datos, del grupo de jugadores de fútbol soccer

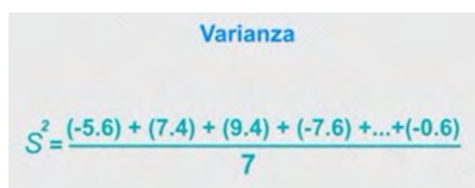
3.3.5. Campos de problemas de la varianza

1. *Estimar la varianza de un conjunto de datos.* Este campo de problemas, reconoce a la varianza de un conjunto de datos como una medida de los valores alrededor de la media. En el ejemplo de la Figura 7, lo considera como el valor indispensable para calcular la desviación estándar.

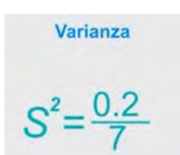
"si sustituimos esos valores utilizando nuestro ejemplo queda algo así, varianza es igual a la suma de cada uno de nuestros 8 datos restados cada uno con la media que calculamos al principio"



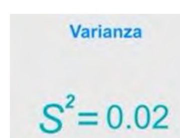
(Escena 1)



(Escena 2)



(Escena 3)



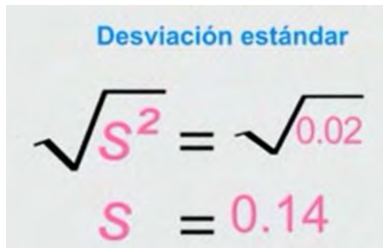
(Escena 4)

Figura 7. Cálculo de la varianza de un conjunto de datos, del grupo de jugadores de fútbol soccer

3.3.6. Campos de problemas de la desviación estándar.

1. *Estimar la desviación estándar de un conjunto de datos.* En la Figura 8, se observa este campo de problemas, en el que presenta a la desviación estándar como medidas de separación con respecto a la media.

"calcular la desviación estándar"



(Escena 1)

Figura 8. Estimar la desviación estándar de un conjunto de datos, del equipo de jugadores de fútbol soccer

3.4. Lenguaje

Como parte de los elementos de significado, el lenguaje de un objeto matemático engloba las representaciones del objeto abstracto (términos, notaciones, expresiones, símbolos, tablas, gráficos).

3.4.1. Términos y expresiones

En el video analizado se detectaron las siguientes expresiones:

1. *Medidas de tendencia central o de posición.* Son aquellas que nos indican los valores centrales alrededor de los cuales se distribuyen los datos. Respecto a estas medidas enlista tres principales; la media aritmética, la mediana y la moda.
2. *Media aritmética o promedio.* En el video utiliza en muchas ocasiones la palabra promedio como sinónimo, y señala que nos sirve para obtener una medida que nos dé una idea más clara sobre el valor central de los datos que se manejan. Y es utilizada en distintas ocasiones como: *"necesitamos saber el peso promedio que tiene el equipo"; "el valor de la media aritmética es 55.6 kilogramos"; "Es equivalente a decir que*

entre mi vecino y yo tenemos dos coches en promedio si él tiene 3 y yo 1 lo cual es falso".

3. *Mediana.* Considerada en el video como el valor de la variable estadística que divide en dos partes iguales al conjunto de datos, una vez que se han ordenado de menor a mayor. Y declaran una dificultad que tiene el cálculo de la mediana que depende de si el total de conjunto de datos es un número par o impar, o si los datos están agrupados o no. se hace énfasis que la mediana es el valor del centro de la lista cuando los datos están ordenados.
4. *Moda.* Establecen que es el valor que se puede identificar mediante una simple inspección visual. En repetidas ocasiones hablan de la moda como: *"el valor que aparece con más frecuencia entre nuestros datos"; "la cifra se repite"; "se repite el número 51 por lo tanto nuestra moda es 51 kilogramos"; "habría dos modas por lo que la moda seria bimodal".*
5. *Medidas de dispersión, variabilidad o variación*. En el video se señala que, estas medidas nos proporcionan una medida del mayor o menor agrupamiento de los datos respecto a los valores de tendencia central. Específicamente se refiere a, el rango, la varianza y la desviación estándar.
6. *Amplitud o rango.* La medida de dispersión más sencilla, y es la diferencia de entre el dato mayor identificado y el de menor valor. En una sola ocasión se menciona: *"la amplitud o rango en el ejemplo del peso de los 8 jugadores se obtiene mediante la diferencia del mayor peso 65 y el menor peso registrado 48, queda 17".*
7. *Varianza.* El valor *indispensable para calcular la desviación estándar. Erróneamente describen su estimación: "se calcula usando la formula sumatoria de x menos el promedio al cuadrado entre N-1".*

   En el que se entiende que hay que sumar primeramente los valores de "x" y seguidamente restar el valor del promedio al cuadrado, entendiendo que hay que elevar al cuadrado únicamente el promedio. Por otro lado, se encontró que la solución del ejercicio propuesto en el video es errónea, la Figura 7, muestra el cálculo, en el que restan a cada valor de la variable la media pero no lo elevan al cuadrado, y así considerando esos resultados realizan la suma de las diferencias de todos los datos y obtienen como resultado 0.2, cuando lo correcto es realizar la suma de las diferencias cuadradas de cada dato y que el resultado es 367.88 y dividido por 7 se obtiene una varianza de 52.55.

8. *Desviación estándar*. Enuncian diciendo que es sencillo: *"sólo nos resta sacar la raíz cuadrada de nuestra varianza"*.

   Describen a la desviación estándar como la: *"separación con respecto a la media"*.

   Señalan que los valores de estas medidas de dispersión serán mayores o iguales a cero donde el valor de cero indica ausencia de dispersión y cuando es mayor que cero indica que los datos están muy separados. Cuando proceden al cálculo de esta medida siguiendo su ejemplo, dicen que *"debemos calcular la raíz cuadrada de 0.02*" (en el video la voz lo llama cero punto dos y en la imagen aparece 0.02), lo que es igual a 0.14, lo cual es erróneo ya que la varianza está mal calculada.

3.4.2. Notaciones y símbolos

Las notaciones son la forma sencilla de expresar de manera simbólica un concepto. En el video se han encontrado algunas que se presentan a continuación:

1. *Media aritmética*: $\bar{x}$
2. *Mediana*: $\tilde{x}, Me$
3. *Moda*: $\hat{x}$
4. *Amplitud o rango*: A, R
5. *Varianza*: s^2
6. *Desviación estándar*: $s, \sqrt{s^2}$

Del mismo modo, se encontraron los siguientes símbolos que aparecen en la descripción de las fórmulas.

1. $\sum_{i=1}^{n}$ Indica la suma de todos los datos desde el primero hasta el *n-ésimo*.
2. x_i Representa cada uno de los datos
3. f_i Significa la frecuencia con que aparece un valor
4. N Representa el número total de datos

5. $\sum_{i=1}^{n}[x_i - \bar{x}]$ Para indicar la suma de las desviaciones de los valores individuales respecto a la media.
6. A Límite superior del conjunto de datos.
7. R Límite inferior del conjunto de datos.

3.4.3. Otras representaciones

En el video se identifica conjuntos de datos sin formato, algunos en tablas de datos, tablas de frecuencias, tablas de frecuencias acumuladas, pictogramas, polígonos de frecuencias acumuladas, graficas de distribución y graficas de barras.

3.5. Definiciones

También llamado ideas matemáticas, abstracciones, generalizaciones (conceptos).

3.5.1. Definiciones de media

1. *Definición de media como promedio aritmético de un conjunto de datos.* Se identifica que se refieren a la media aritmética como el valor central de los datos y lo utilizan como sinónimo de promedio. *"La media aritmética o promedio nos sirve para obtener una medida que nos dé una idea más clara sobre el valor central de los datos que estamos manejando";* el ejemplo de promedio, es sobre la media de datos aislados, pero la definición que da es de la media ponderada: *"El promedio es igual a la suma de todos los datos por su frecuencia desde el primero hasta el último entre el número total de datos".*

3.5.2. Definiciones de la mediana

Al igual que Cobo (2003) y Mayén (2009) se encontró esta definición:

1. *Definición de mediana como centro de la distribución de datos.* En el video se hace énfasis en el orden de los datos de forma descendente o ascendente, y centra su atención en que la mediana será el centro del conjunto de datos que han sido ordenados. Las definiciones señalan: *"La mediana es el valor del centro de la lista cuando los datos*

están ordenados, si el número total de datos es par, la mediana es igual al promedio de los dos valores que se encuentran en el centro de la lista cuando los datos están ordenados"; "Si los datos están presentados en forma de lista, si el número total de datos es impar, la mediana es el valor del centro".

2. *Definición de mediana como valor que divide en dos el conjunto de datos.* Esta definición de mediana señala que el conjunto de datos es dividido en dos partes iguales, y lo define así: *"La mediana es el valor de la variable estadística que divide en dos partes iguales al conjunto de datos, una vez que se han ordenado de menor a mayor".*

3.5.3. Definiciones de la moda

1. *La moda es el valor más frecuente de la variable estadística.* La definición de la moda la plantean de una forma sencilla de comprender, el valor con más frecuencia y fácil de observar: *La moda es el valor que aparece con más frecuencia entre nuestros datos y se simboliza de la siguiente manera $\hat{x}$".*

3.5.4. Definiciones de rango

1. *Definición de amplitud o rango como medida de dispersión más sencilla.* Destacan que el rango es fácil de calcular y que es la diferencia de los límites extremos. Lo definen de la siguiente manera: *"La amplitud o rango es la medida de dispersión más sencilla es la diferencia de entre el dato mayor identificado con la letra H y el de menor valor representado con la letra L".*

3.5.5. Definiciones de varianza

1. *Definición de varianza como medida de los valores alrededor de la media.* El video, no se hace mucho énfasis en la varianza, indican que es un valor necesario para el cálculo de la desviación estándar y que necesita del promedio para su estimación. La siguiente definición que se aprecia en el video, es poco entendible y confusa: *"La varianza es un valor indispensable para calcular la desviación estándar y se calcula usando la formula sumatoria de x menos el promedio al cuadrado entre N-1".*

3.5.6. Definiciones de desviación estándar

1. *Definición de desviación estándar como medida de variación.* Lo señalan como un procedimiento sencillo, obtener la raíz cuadrada de la varianza, y también el alejamiento de los datos respecto al promedio. A continuación se cita la definición: *"La desviación estándar, sólo nos resta sacar la raíz cuadrada de nuestra varianza"; "La varianza y la desviación estándar son medidas de separación con respecto a la media los valores de estas medidas de dispersión serán mayores o iguales a cero donde el valor de cero indica ausencia de dispersión y cuando es mayor que cero indica que los datos están muy separados".*

3.6. Propiedades

Para realizar el análisis de las propiedades numéricas, algebraicas y estadísticas que se encontraron en el video, para lo que se consideró la clasificación que Mayén (2009) realizó para describir las mismas.

3.6.1. Propiedades numéricas

Son las que se deducen cuando se consideran a las medidas de tendencia central como un número, el valor resultado de un cálculo.

1. *En el cálculo de la media se tiene en cuenta todos los valores de los datos.* Se puede observar en la Figura 9, el ejemplo propuesto utiliza todos los datos disponibles para calcularla, ninguno queda fuera.


1. UTILIZAS TODOS LOS DATOS

Jugadores	kg
1	50
2	63
3	65
4	48
5	49
6	64
7	51
8	55

(Escena 1)

Figura 9. Propiedad de la media que considera todos los valores del conjunto de datos

2. *El valor numérico de la media cambia cuando se cambia cualquier dato.* La Figura 10, muestra que cuando utilizas todos los datos y si alguno se llegara a modificar, afectan el resultado, así como los valores extremos afectaría el resultado agrado de obtener un valor poco confiable. Como el siguiente ejemplo:

"sí 3 de mis 8 jugadores pesan 100 kilogramos y dos pesan entre 34 y 37 kilogramos los valores tan extremos afectaría el resultado agrado de obtener un valor poco confiable"

(Escena 1)

Figura 10. Propiedad de la media al considerar que el valor numérico de la media cambia cuando se cambia cualquier dato

3. "*La media, mediana y moda de un conjunto de datos son siempre valores pertenecientes al rango de la variable.* En una distribución de datos, donde hay que hallar las medidas de centralización, no se puede obtener un valor mayor que el máximo de estos ni menor que el mínimo" (Mayén, 2009). Esta propiedad se ve de manera implícita en el video y sus ejemplos
4. De la misma forma que la propiedad anterior, se encontró de forma implícita en los ejemplos: *"Mientras la moda siempre coincide con uno de los valores dados, la media y mediana no tienen por qué coincidir con los valores de los datos e incluso podría ser un número perteneciente a un conjunto numérico más amplio que el dado"* (Mayén, 2009).

3.6.2. Propiedades algebraicas

Son las que se deducen cuando se considera a las medidas de tendencia central como una operación, es decir como el algoritmo u operación con los datos.

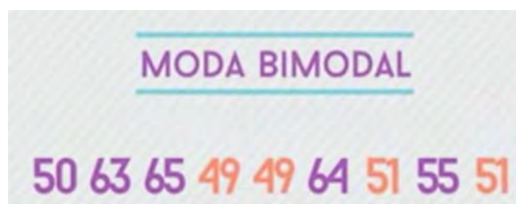
1. *La moda puede no existir o, si existe, no ser única. La media y mediana siempre existen en datos numéricos.* La Figura 11, nos plantea casos en los que no existe moda, existe sólo una o existe más de una; en ejemplos que contiene se pueden observar:

"Veamos de nuevo a nuestros nueve jugadores con su peso anotado de bajo de ellos, vean que cifra se repite, así es se repite el número 51 por lo tanto nuestra moda es 51 kilogramos"

(Escena 1)

"Si en lugar de pesar 48 kilogramos uno de los muchachos pesara 49, habría dos modas por lo que la moda seria bimodal"

(Escena 2)



(Escena 3)

"Si no se hubiera integrado al grupo el portero y el jugador número 4 no hubiera subido de peso y todo estuviera como al principio, no habría moda alguna"

(Escena 4)

Figura 11 . Propiedad de la moda

3.6.3. Propiedades estadísticas

Las propiedades estadísticas son aquellas cuando se consideran como un resumen estadístico o parámetro que caracteriza una distribución.

1. *La media, mediana y moda son representantes de un colectivo.* En el video se menciona que el uso de estas medidas son valores que representan y resumen el conjunto de datos para conocer cómo se concentran y distribuyen entre ellos: *"Las medidas de tendencia central y de posición nos indican los valores centrales alrededor de los cuales se distribuyen los datos"*.
2. La media es un estadístico poco resistente, muy sensible a la variación de los datos, especialmente en los valores atípicos. Se destaca en el video la sensibilidad de la media a valores atípicos. Se puede observar en el ejemplo: *"Es equivalente a decir que entre mi vecino y yo tenemos dos coches en promedio si él tiene 3 y yo 1 lo cual es falso"; "La media puede ser afectada por valores extremos, esto quiere decir que sí 3 de mis 8 jugadores pesan 100 kilogramos y dos pesan entre 34 y 37 kilogramos los valores tan extremos afectaría el resultado agrado de obtener un valor poco confiable"*
3. *La suma de las desviaciones de un conjunto de datos, de su media es cero.* La Figura 12, muestra esta propiedad en la que se ve involucrada la media, y al restárselas a cada valor del conjunto de datos nos da cero. A continuación el ejemplo:

"La suma de las desviaciones de los valores individuales respecto a la media es cero, esto quiere decir que si a cada peso le restamos lo que nos salió del promedio y sumamos el resultado de cada una de estas restas el resultado final será igual a cero"

$$\sum_{i=1}^{n} \left[X_i - \bar{X} \right] = 0$$

$$(50-55.6)+(63-55.6)+(65-55.6)+(48-55.6)+(49-55.6)+(64-55.6)+(51-55.6)+(55-55.6)=0$$

$$(-5.6)+7.4+9.4+(-7.6)+(-6.6)+8.4+(-4.6)+(-.6)=0$$

Figura 12. Propiedad de la media donde la suma de las desviaciones de un conjunto de datos de su media es cero

4. *La varianza y la desviación estándar son estadísticos sensibles a la variación de un valor atípico.* Esta propiedad no se menciona en el video, pero se puede deducir al momento que se calcula la media con un valor atípico, si éste afecta a la media, de igual forma repercutirá en el cálculo de la varianza y por consiguiente en la desviación estándar.
5. *La varianza y la deviación estándar siempre será positiva o cero.* En el video se menciona esta propiedad, al igual que es fácil de observar, puesto que no importa si se obtiene un resultado negativo al restar la media de cada dato ya que, al elevarlo al cuadrado se volverá positivo. El en ejemplo siguiente lo establece el video: *"Calcular la desviación estándar, sólo nos resta sacar la raíz cuadrada de nuestra varianza, es decir debemos calcular la raíz cuadrada de 0.2 lo que es igual a 0.14 los valores de la varianza y la desviación estándar son medidas de separación con respecto a la media los valores de estas medidas de dispersión serán mayores o iguales a cero donde el valor de cero indica ausencia de dispersión y cuando es mayor que cero indica que los datos están muy separados".*

3.7. Procedimientos

En esta sección se describen algunos tipos de algoritmos de cálculo que aparecen en el video que se analiza, siguiendo el método utilizado por Mayén (2009), se presentan los siguientes.

3.7.1. Cálculo de la media

1. *Cálculo de la media de una variable discreta con datos aislados.* La Figura 13, que se presenta a continuación, muestra un ejemplo del algoritmo de la media aritmética en distribuciones de datos aislados.

“Imaginen a 8 jóvenes, son un equipo de fútbol soccer y nosotros somos los entrenadores, queremos hacer un entrenamiento y una dieta que sirva para todos por lo tanto, necesitamos saber el peso promedio que tiene el equipo, ¿Cómo obtengo este dato?”

(Escena 1)

(Escena 2)

Figura 13. Cálculo de la media de una variable discreta con datos aislados

2. *Cálculo de la media de una variable continua o discreta con datos agrupados en intervalos de clase.* En el video se plantea un ejemplo que implica encontrar la media de una distribución en tabla de frecuencia con intervalos en la que es necesario el cálculo del punto medio. En la Figura 14 se presenta el ejemplo.

Calcular la media aritmética o promedio del conjunto de datos que muestra el tiempo en minutos que las personas escuchan música.

Tiempo en minutos	Punto medio del intervalo	Número de personas (frecuencia)	
0 - 60	30	19	30x19=
61 - 120	90	34	90x34=
121 - 180	150	25	150x25=
181 - 240	210	12	210x12=
241 - 300	270	10	270x10=
301 - 360	330	15	330x15=
361 - 420	390	5	390x5=
Total		120	162.5

(Escena 1)

Figura 14. Cálculo de la media de una variable continua o discreta con datos agrupados en intervalos de clase

3.7.2. Cálculo de la mediana

1. *Cálculo de la mediana con datos no agrupados en clases, número de datos impar.* Considerado en el video como un algoritmo sencillo de desarrollar, como se muestra en la Figura 15, mencionan que, si los datos están presentados en forma de lista y si el número total de datos es impar, la mediana es el valor del centro de la lista cuando los datos están ordenados.

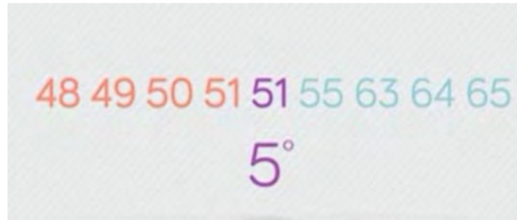
<table>
<tr><td colspan="2">“si el número de nuestros datos muéstrales no fuera par, sino impar, es decir, si en vez de tener 8 jugadores tuviéramos 9, …consideramos al nuevo integrante quién también pesa 51 kilogramos”</td></tr>
<tr><td>
(Escena 1)</td><td>
(Escena 2)</td></tr>
</table>

Figura 15. Cálculo de la mediana con datos no agrupados en clases, número de datos impar

2. *Cálculo de la mediana con datos no agrupados en clases, número par de datos.* Del mismo modo que el anterior, es un ejercicio común, sin embargo, el video lo plantea como un algoritmo con mayor dificultad, por el hecho de estimar el promedio de los dos datos centrales, como se puede observar en la Figura 16.

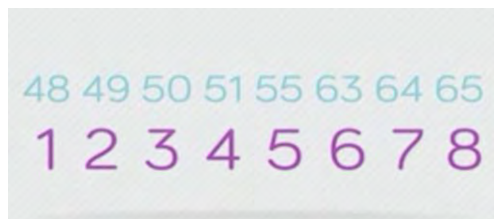
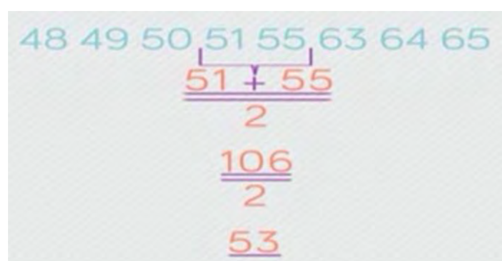
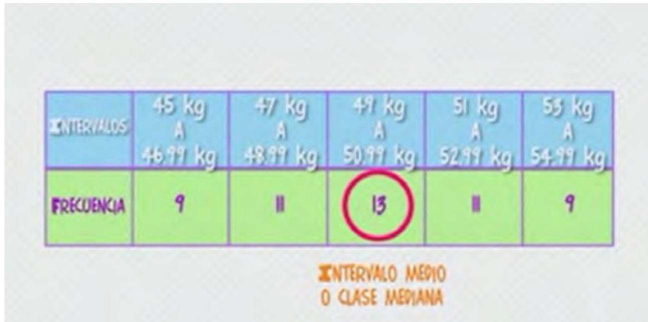
<table>
<tr><td colspan="2">“Volvamos a nuestro ejemplo con los jugadores de fútbol tenemos 8 jugadores los primeros pesan muy poco y los últimos pesan demasiado, en este caso aplicaremos la mediana”</td></tr>
<tr><td>
(Escena 1)</td><td>
(Escena 2)</td></tr>
<tr><td>
(Escena 3)</td><td>
(Escena 4)</td></tr>
</table>

Figura 16. Cálculo de la mediana con datos no agrupados en clases para un número par de datos

3. *Cálculo de la media para datos agrupados en intervalos de clases.* El video señala la clase mediana que se obtuvo en un conjunto de datos ordenados en una tabla de frecuencias acumuladas. Sin embargo, se puede observar en la Figura 17, que se presenta de forma ya resuelta, y el proceso de la obtención de la solución queda implícito y difícil de entender.

“si los datos están agrupados en intervalos o clases se calculan las frecuencias acumuladas de las clases o intervalos hasta obtener el intervalo medio o la clase mediana”



(Escena 1)

Figura 17. Cálculo de la media para datos agrupados en intervalos de clases

3.7.3. Cálculo de la moda

1. *Cálculo de la moda en una variable discreta con datos aislados.* Como se había mencionado antes, el video plantea, encontrar la moda por simple observación. Que es considerada la forma más tradicional en que se da solución a este tipo de situaciones. La Figura 18, muestra el ejemplo, la obtención de la moda.

“Veamos de nuevo a nuestros nueve jugadores con su peso anotado de bajo de ellos, vean que cifra se repite, así es se repite el número 51 por lo tanto nuestra moda es 51 kilogramos”

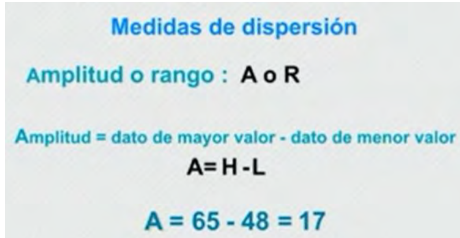
(Escena 1)

Figura 18. Cálculo de la moda en una variable discreta con datos aislados

3.7.4. Cálculo del rango

1. *Cálculo de rango en una variable discreta con datos aislados.* El video nos presenta, una única forma de encontrar el valor del rango que consiste en considerar los datos extremos de un conjunto de datos. El ejemplo presentado en la Figura 19, nos ayuda a comprenderlo.

“La amplitud o rango del ejemplo del peso de los 8 jugadores”



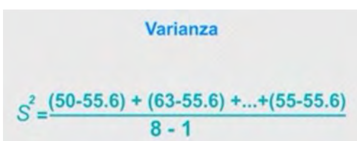
(Escena 1)

Figura 19. Cálculo de rango en una variable discreta con datos aislados

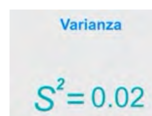
3.7.5. Cálculo de la varianza

1. *Cálculo de la varianza en una variable discreta con datos aislados.* La única propuesta que realiza el video para el cálculo de la varianza. En el ejemplo del equipo de fútbol se puede observar en la Figura 20, que es erróneo su cálculo, pues no elevan las diferencias al cuadrado, sino que solo así las suman.

“si sustituimos esos valores utilizando nuestro ejemplo queda algo así, varianza es igual a la suma de cada uno de nuestros 8 datos restados cada uno con la media que calculamos al principio”



(Escena 1)



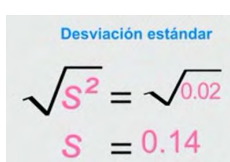
(Escena 2)

Figura 20. Cálculo de la varianza en una variable discreta con datos aislados

3.7.6. Cálculo de la desviación estándar

1. *Cálculo de la desviación estándar en una variable discreta con datos aislados.* El video plantea solo un paso para el cálculo de esta, extraer la raíz cuadrada de la varianza, como se puede ver en la Figura 21. Sin embargo, siendo incorrecto el resultado anterior, del mismo modo este es incorrecto.

“calcular la desviación estándar”



(Escena 1)

Figura 21. Cálculo de la desviación estándar en una variable discreta con datos aislados

3.8. Argumentos

Cobo (2003), señala que los argumentos o razonamientos se utilizan para comprobar las soluciones de los problemas o explicar otra persona la solución y que pueden ser deductivas, o de otro tipo. Mayén (2009), dice que son de utilidad y facilitan la comprensión de los estudiantes, sobre los demás objetos matemáticos.

En esta parte, se enlistan los argumentos que se han encontrado en el video analizado, utilizando las mismas clasificaciones de Mayén (2009).

1. *Justificación de ejemplos o contraejemplos.* En el video se identificaron pocos problemas que expliquen su resolución y seguidos de algún argumento que evidencie su resultado y su aplicación. En este ejemplo se muestra un pequeño argumento: *"Imaginen a 8 jóvenes, son un equipo de fútbol soccer y nosotros somos los entrenadores queremos hacer un entrenamiento y una dieta que sirva para todos por lo tanto, necesitamos saber el peso promedio que tiene el equipo, ¿Cómo obtengo este dato?; "Primero debemos pesar a cada integrante del equipo y apuntar el peso de cada uno, tenemos 8 jugadores; 1, 2, 3 y así sucesivamente hasta llegar 8. Ahora tenemos registro de los 8 pesos distintos; unos pesaron 50 otros 63 el de más peso llegó a 65 y el que menos peso 48 kilogramos, para sacar el promedio solo tendremos que sumar todos los pesos de los jugadores y dividir el resultado entre el número de muchachos que pesamos, en nuestro caso es 8. El resultado de esta operación será el promedio, en nuestro caso el promedio es 55.6 kilogramos, pues si sumamos todos los pesos nos da 445 entre 8 más de 55.6. Estos muchachos pesan 55.6 kilogramos en promedio".*
2. *Uso de gráficos cuando la argumentación verbal o simbólica se apoya en las propiedades visuales de un gráfico*. Se observa en la Figura 22, que para la argumentación de la desviación estándar incluyen una gráfica que muestra los resultados de los cálculos obtenidos respecto a la media, para facilitar su comprensión.

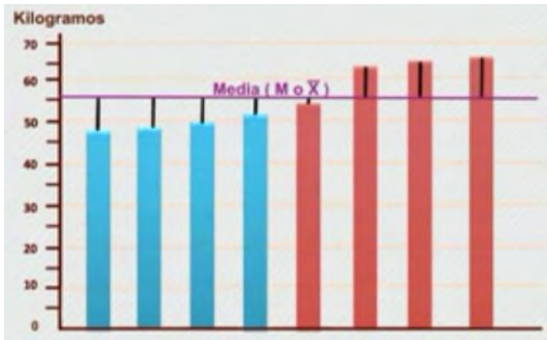
“Vemos graficados algunos pesos en kilogramos de nuestros jugadores, si se fijan bien las columnas no son iguales, unas son más altas y otras más chicas por lo tanto hay medidas que se alejan o se acercan más a la media las medidas nos indican el rango de lejanía o cercanía de las columnas con respecto a la media se llaman medidas de dispersión”
 (Escena 1)

Figura 22. Uso de gráficos cuando la argumentación verbal o simbólica se apoya en las propiedades visuales de un gráfico auxiliar

3. *Razonamientos verbales deductivos.* Este tipo de argumentos a menudo son usados para explicar los resultados obtenidos de alguna estimación. la Figura 23, es un ejemplo de ello.

“Veamos de nuevo a nuestros nueve jugadores con su peso anotado de bajo de ellos, vean que cifra se repite, así es se repite el número 51 por lo tanto nuestra moda es 51 kilogramos”
 (Escena 1)

Figura 23. Razonamientos verbales deductivos para la moda

CAPÍTULO 4

EXPERIENCIA DE APRENDIZAJE

4.1. Introducción

En este capítulo se describen las etapas del diseño de la experiencia de aprendizaje, la descripción de la muestra (los participantes), la forma en la que se aplicó, y los instrumentos utilizados, para conocer el significado y la evolución de conocimientos que sobre medidas de tendencia central y variabilidad tienen los estudiantes del Telebachillerato Comunitario.

4.2. Fases del Trabajo

Para abordar los objetivos de este trabajo, se llevó a cabo una experiencia de aprendizaje tomando en cuenta las tres fases de los Experimentos de Enseñanza (descrito en la sección 2.4), con el fin de describir la comprensión de las medidas de tendencia central y de variabilidad en estudiantes de Telebachillerato. Las fases las se presentan en la Figura 1, tomada de Rubiales (2011) y adaptado a este trabajo:

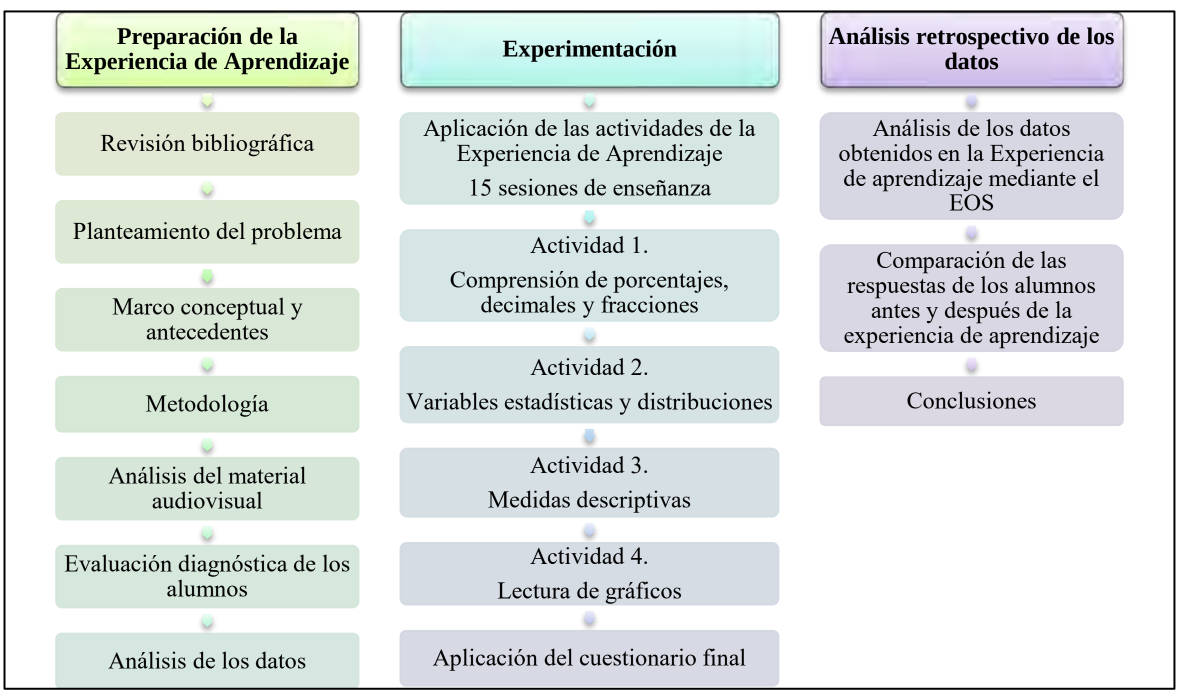


Figura 24. Fases del experimento

Preparación del Experimento: Para el desarrollo de esta fase en este trabajo, se realizó una revisión bibliográfica de investigaciones relacionadas con las medidas de tendencia central y de variabilidad, lo que ayudo a delimitar el problema y el objetivo de la investigación. Seguidamente, permitió identificar los conocimientos previos de los alumnos. En estos se utilizó el cuestionario de medidas de tendencia central para explorar de Mayén (2009): 1) es un instrumento validado en diferentes contextos, y porque se ha aplicado, con estudiantes de este mismo nivel pero de otros subsistemas de bachillerato y mexicano. Del mismo modo, se eligió al modelo EOS para el análisis de las respuestas que den los estudiantes.

Una vez analizadas las respuestas obtenidas en la exploración de conocimientos previos, se diseñó una experiencia de aprendizaje tomando en cuenta las deficiencias detectadas, también se seleccionaron tres problemas del cuestionario de Mayén (2009), y adicionalmente un problema del cuestionario de Orta (2014), con lo que se rediseñó el cuestionario final que se aplicó a los estudiantes, posteriormente a los episodios de enseñanza. Los contenidos de los problemas elegidos son los siguientes:

1) Las definiciones de media y moda
2) La media ponderada
3) La estimación directa de la media a partir de un gráfico y en el cálculo de la mediana a partir de un gráfico
4) Ideas sobre variabilidad.

Experimentación: Para esta fase, se desarrolló una experiencia de aprendizaje dividida en cinco episodios. A su vez, cada uno de tres sesiones realizadas a lo largo de una semana. En la Tabla 2, se describe el esquema de las actividades desarrolladas.

Tabla 2 . Episodios de la fase de experimentación

Distribución Semanal	No. alumnos	Tiempos	No. de sesiones	Actividades	
SEMANA 1	**17**	**150 Min.**	**3**	**Comprensión de porcentajes, decimales y fracciones.**	
Martes	17	50 Min.	1	Aritmética	Cálculo de operaciones básicas (uso de símbolos y signos) Cálculo de fracciones (suma, resta, multiplicación y división)
Miércoles	17	50 Min.	1	Decimales	Uso de decimales (operaciones básicas) Equivalencias (fracciones)
Jueves	17	50 Min.	1	Porcentajes	Regla de tres simple
SEMANA 2	**17**	**150 Min.**	**3**	**Variables estadísticas y distribuciones**	
Martes	17	50 Min.	1	Cualitativas	Variables categóricas nominales Variables categóricas ordinales
Miércoles	17	50 Min.	1	Cuantitativas	Variables numéricas discretas Variables numéricas continuas
Jueves	17	50 Min.	1	Distribuciones	
SEMANA 3	**17**	**300 Min.**	**3**	**Medidas de tendencia central y variabilidad**	
Martes	17	100 Min.	1	Media	Resolución de problemas de tendencia central y de variabilidad
Miércoles	17	100 Min.	1	Mediana y Moda	
Jueves	17	100 Min.	1	Rango, varianza y desviación estándar	
SEMANA 4	**17**	**150 Min.**	**3**	**Lectura de gráficos**	
Martes	17	50 Min.	1	Histograma	Tareas de graficas
Miércoles	17	50 Min.	1	Gráfica de barras	
Jueves	17	50 Min.	1	Ojiva	
SEMANA 5	**17**	**150 Min.**	**3**	**Aplicación de la segunda evaluación**	
Martes	17	50 Min.	1	Aplicación del cuestionario final	
Miércoles	17	50 Min.	1		
Jueves	17	50 Min.	1		

El primer episodio se llevó a cabo en tres sesiones de 50 minutos a lo largo de una semana. Las actividades propuestas estaban orientadas a la comprensión de porcentajes, decimales y fracciones. Con la finalidad de reforzar y clarificar los conocimientos previos de cálculo de operaciones básicas, cálculo de fracciones, el uso de decimales y la regla de tres simple.

El segundo episodio, como el anterior, también se llevó a cabo en tres sesiones de 50 minutos durante una semana. En este caso, las actividades se concentraron en la comprensión de la clasificación de las variables estadísticas y distribuciones. Específicamente las actividades desarrolladas en esta etapa tienen la finalidad de comprender las diferencias entre las variables categóricas nominales y ordinales, variables numéricas discretas y continuas, además del concepto de distribución.

El tercer episodio, se desarrolló en tres sesiones de 100 minutos, por petición de los alumnos, ya que no les quedaban claros los conceptos. Las actividades propuestas en esta etapa consisten en la resolución de problemas de medidas de tendencia central: media, mediana y moda, y medidas de variabilidad: rango, varianza y desviación estándar.

El cuarto episodio, se desarrolló en tres sesiones de 50 minutos, los alumnos realizaron tareas sobre lectura de gráficos que incluye, histograma, gráfica de barras y ojiva. El objetivo de esta actividad es promover en los estudiantes el desarrollo de habilidades y capacidades para interpretar gráficos, estimar datos y realizar cálculos estadísticos como las medidas de tendencia central y de dispersión, además de ubicarlos en los gráficos cuando fuera posible.

En el quinto episodio, realizado en tres sesiones de 50 minutos, se aplicó el cuestionario final que incluye 3 problemas de medidas de tendencia central y uno de variabilidad. El objetivo de esta sesión fue evaluar los conocimientos de los estudiantes sobre las medidas de tendencia central y las medidas de variabilidad después de la experiencia de aprendizaje.

En cada sesión se proporcionó a los estudiantes hojas de trabajo, de donde se obtuvieron los datos para el desarrollo y justificación del análisis realizado.

Análisis Retrospectivo de los Datos: se llevó a cabo el análisis de las respuestas de los alumnos, tomando como base el enfoque ontosemiótico" (EOS). A partir de las respuestas se compararon los elementos de significado del antes y después de la experiencia de aprendizaje, mediante la categorización de las respuestas y con base en los elementos de significado de cada problema propuesto, los resultados se presentan en el capítulo 5.

4.3. Descripción de la muestra

La muestra está integrada por 17 alumnos del Telebachillerato Comunitario, en la comunidad de Tres Garantías, ubicado en el municipio de Othón P. Blanco, del estado de Quintana Roo. Este centro escolar está ubicado a 113km de la ciudad capital Chetumal; con servicios básicos de energía eléctrica, agua potable, no tiene acceso a internet, ni red telefónica.

Los estudiantes son de quinto semestre y tienen estudios previos de estadística básica, que han cursado en segundo semestre.

El motivo de elegir este grupo, es porque son alumnos que están a punto de concluir el bachillerato y se quiere evaluar sus conocimientos antes de acceder a una institución de nivel superior.

A continuación se describen las características de la muestra ilustrando con la Figura 25, la información según el género de la muestra.

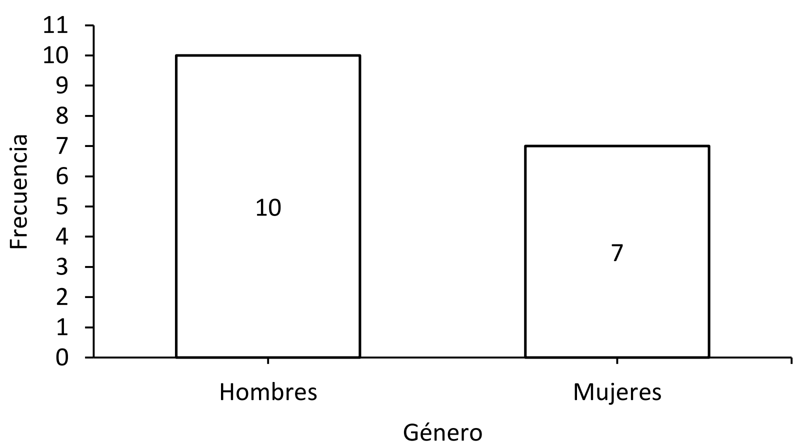


Figura 25 . Distribución de la muestra por género

Las variables independientes en el estudio son: género y edad, pues el centro de estudios es único. En la Figura 25, se puede observar la distribución de la muestra por género, donde la cantidad de hombres es 10, y la cantidad de mujeres es de 7. Una distribución muy común en las instituciones educativas del país. En el caso de Tres Garantías, Quintana Roo, prevalece un contexto cultural muy bajo pero importante, en el que algunos hombres de edad avanzada señalan que la mujer no debería estudiar.

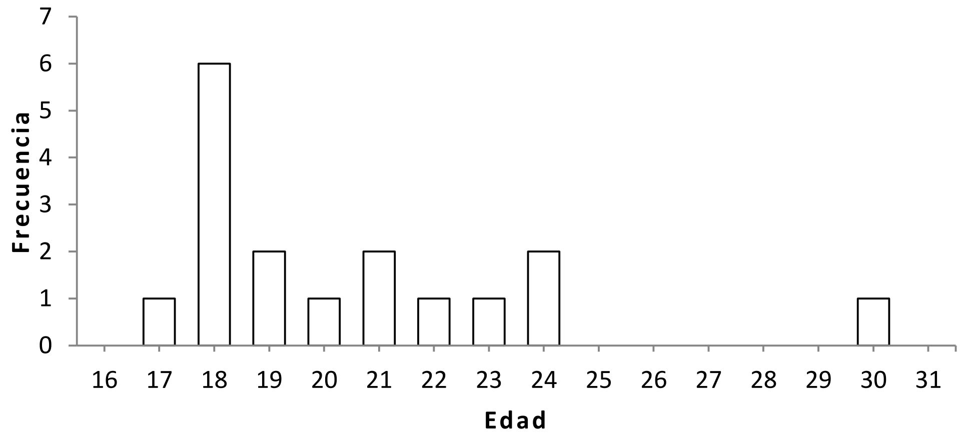


Figura 26. Distribución de la muestra por edad

La distribución de la muestra por edades esta de la siguiente manera: se pueden clasificar en tres grupos, el primero que abarca las edades de 17, 20, 22, 23, y 30 son las que se encuentran en menor frecuencia una persona por cada edad. En el segundo se agrupan 6 alumnos con edades de 19, 21 y 24 años, 2 alumnos por edad. El tercer grupo únicamente lo constituyen alumnos de 18 años, siendo la edad más común en este grado escolar. La variabilidad en las edades obedece a que el centro es de reciente creación y al momento de inscribirse los estudiantes presentaban rezago educativo, es decir, habían abandonado sus estudios de bachillerato en un rango de tiempo que va desde 1 hasta 10 años.

4.4. Instrumentos

El instrumento de medición con el que se pretende obtener una estimación de conocimientos y capacidades de los alumnos, está encaminado a identificar los significados personales que los estudiantes asignan a conceptos relacionados con la variabilidad, como las medidas de tendencia central y de variabilidad.

Los problemas que se han seleccionado son los utilizados por Mayén (2009) y Orta (2014), con validez de contenido, es decir, la adecuación de los problemas de un test como muestra de un universo más amplio de problemas representativos del contenido (Martínez Arias, 1995). Con ello Mayén (2009), comprobó que los problemas de su test son relevantes para el uso que se dará a las puntuaciones y representativos del contenido que se quiere evaluar, representando sus características esenciales.

Los problemas seleccionados tienen como objetivo principal, recopilar información sobre las habilidades matemáticas que los estudiantes utilizan al momento de resolver problemas sobre las medidas de centralización y las medidas de variabilidad, para determinar el desarrollo de conocimiento.

Inferir en el uso de los diversos objetos matemáticos, por medio de las respuestas escritas que nos den los estudiantes y así determinar la utilización y comprensión de los elementos de significado (definiciones, propiedades, argumentos, etc.) e informar sobre la comprensión de ellos. En este sentido, se trata de comprender lo siguiente:

i. El reconocimiento de los problemas que se plantean en diferentes campos.
ii. La utilización de las diferentes definiciones en la solución de los problemas propuestos o en la justificación de la solución.
iii. La comprensión de propiedades básicas, tanto numéricas, como algebraicas y estadísticas, además de su uso adecuado al responder a las preguntas planteadas.
iv. Declaración del lenguaje matemático verbal, numérico y gráfico, asimismo su uso adecuado.
v. Con el cálculo y procedimientos de resolución de problemas, se pretende medir la comprensión de los algoritmos de cálculo frente a su aplicación automática.
vi. Argumentaciones de los alumnos para apoyar sus respuestas y observar hasta qué punto son completas y consistentes.

Al pretender evaluar la comprensión de las medidas de tendencia central y las medidas de variabilidad, se cae en cuenta que no es una variable directamente observable, a la que (León y Montero, 2002) citado en Mayén (2009), la denominan constructo, y que se debe inferir sus características a partir de las respuestas de los alumnos.

Por tanto, la comprensión de los estudiantes sobre un cierto objeto matemático (en este caso las medidas de tendencia central y de variabilidad) es inobservable. Sin embargo, los procedimientos que realizan al resolver problemas, y en específico los problemas presentados como ítems en un cuestionario, sí son observables, siempre y cuando la recogida de datos sea completa y fiable (Godino, 1996).

Los problemas elegidos son de respuesta abierta con la finalidad de recoger con detalle los razonamientos de los estudiantes. A continuación se describen los problemas seleccionados:

Problema 1

Un periódico dice que el número medio de hijos en México es 2.2 hijos por familia.

a) *Explica qué significa para ti esta frase.*
b) *Se han elegido 10 familias mexicanas y el número medio de hijos entre las 10 familias es de 2.2 hijos por familia. Los García tienen 4 hijos y los Pérez tienen 1 hijo. ¿Cuántos hijos podrán tener las otras 8 familias para que la media de hijos en las 10 familias sea 2.2? Justifica tu respuesta.*

El primer problema es de Watson (2000), también fue utilizado en Cobo (2003) en otro contexto y finalmente se toma de Mayén (2009) quien lo adaptó al contexto Mexicano. El contenido, por un lado, permite analizar la definición de la media que el alumno da con sus propias palabras; en la segunda parte del problema el alumno debe dar una distribución de datos considerando los valores dados, de tal forma que dé como respuesta la media ya establecida. Esto implica, el conocimiento del algoritmo de cálculo de la media y la aplicación de su inverso, así como buscar una distribución de datos aislados.

Problema 2

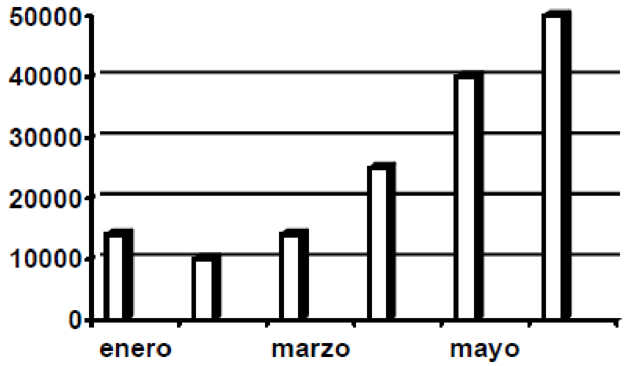
María y Pedro dedican una media de 8 horas cada fin de semana a hacer deporte. Otros 8 estudiantes dedican cada semana una media de 4 horas a hacer deporte.

a) *¿Cuál es el número medio de horas que hacen deporte cada fin de semana los 10 estudiantes?*
b) *María y Pedro dedican además 1 hora cada fin de semana a escuchar música y los otros 8 estudiantes, 3 horas. ¿Cuál sería el número medio de horas que estos 10 estudiantes dedican cada fin de semana, entre las dos actividades: hacer deporte y escuchar música?*

Como el anterior, tomado originalmente de Watson (2000), usado en el estudio de Cobo (2003), y adaptado por Mayén (2009), está compuesto por dos apartados, el primero implica conocer la media ponderada pues solicita a los estudiantes su cálculo, que ha sido un tema con dificultades ya señaladas por otros autores (Li y Shen, 1992; Carvalho, 2001; Cobo, 2003). El segundo apartado también pide el cálculo de la media ponderada, pero además espera que los estudiantes reconozcan la propiedad de que: *la media de la suma de dos o más variables es igual a la suma de las medias de éstas*.

Problema 3

Observa el siguiente diagrama de barras que muestra las ventas de bocadillos de la empresa Bocatta durante los últimos 6 meses del año pasado:



a) *Da un valor aproximado del número medio de bocadillos que se vende al mes.*
b) *Da un valor aproximado de la mediana del número de bocadillos que se vendieron por mes.*

El tercer problema de Zawejowski (1986) lo han usado en sus trabajos Mayén (2009) y Cobo (2003). Tiene la finalidad de identificar la capacidad del alumno al calcular la media y la mediana a partir de un gráfico. Se integra de dos apartados en los que también se pueden analizar algunas propiedades como: "*la mediana y media pueden no coincidir con los datos*" y *"el cálculo de la media y el de la mediana no son operaciones internas"*. La dificultad del problema recae en la lectura del gráfico, requiere que los estudiantes conviertan los gráficos a valores numéricos y con esto realizar los cálculos que se le solicitan.

Problema 4

Una costumbre en los cines es mostrar anuncios comerciales y cortos en la pantalla antes de comenzar la película. El tiempo de espera para una película es la diferencia entre la hora de inicio anunciada y la hora real en la que comienza la película. Un grupo de 10 estudiantes investigó el tiempo de espera de tres cadenas de cines, una de la cadena Cinemex, otra de la cadena Cinépolis y por último una de la cadena Multicinemas. Los estudiantes visitaron los cines en horarios y días diferentes. Cada alumno midió el tiempo de espera en minutos, y lo registró en las tablas que siguen:

Cinemex
12.0
21.0
15.0
15.0
13.0
16.0
16.0
16.0
20.0
18.0

Cinépolis
15.0
15.5
16.0
16.0
16.5
16.5
18.0
16.0
15.5
17.0

Multicinemas
15.5
17.0
18.0
16.5
16.0
16.5
16.5
15.0
15.0
16.0

4.1 Traza una gráfica para cada cadena de cine.
4.2 Calcula el tiempo de espera promedio para cada cadena de cines y señálalo con un segmento de recta en la gráfica que construiste.
Observa con atención las gráficas de la página siguiente.
4.3 Las gráficas que construiste son iguales o diferentes a las que vienen junto con este cuestionario. Iguales______ Diferentes______
En caso de que no sean iguales ¿cuál es la diferencia entre ellas?
El tiempo promedio de espera es de 16.2 minutos en las tres cadenas de cines.
4.4 ¿En cuál de las tres cadenas de cines hay mayor variabilidad en el tiempo de espera? ¿Por qué?
En cuál de las dos cadenas de cines, Cinépolis o Multicinemas, hay mayor variabilidad.
Cinépolis______ Multicinemas_______ Explica tu respuesta.
4.5 Suponiendo que vives a la misma distancia de los tres cines mencionados, ¿a cuál asistirías?
Cinemex________ Cinépolis_________ Multicinemas___________ ¿Por qué?
4.6 ¿Crees que es importante la variabilidad en tiempos de espera que hay en los cines mencionados?

El cuarto problema tomado originalmente de Shaughnessy (2006), y modificado por Orta (2014), pretende analizar los conocimientos sobre el concepto de variabilidad. El problema solicita la construcción de gráficas en las que una vez calculado el promedio de los tres conjuntos de datos, deberá ubicarlos en ellas. El problema tiene los promedios iguales en los tres cines, que implica que el alumno use las medidas de dispersión y, comparando los tres cines, determinen en cual existe mayor variabilidad y haciendo uso de estas medidas pueda tomar una decisión.

Este es un problema bastante complejo, que implica el uso de diferentes objetos matemáticos. Al mismo tiempo, destaca la importancia del uso de las medidas de tendencia central para poder hacer uso de las medidas de dispersión, con las que el estudiante podrá tomar decisiones más acertadas.

4.5. Aplicación del instrumento (procedimientos)

Las actividades de enseñanza se llevaron a cabo en un periodo de 5 semanas, de las cuales en las semanas 1, 2, 4 y 5 fueron impartidas durante 3 sesiones de 50 minutos cada una, en la semana tres, las 3 sesiones tuvieron una duración de 100 minutos cada una, debido a que esta semana se centraba en el estudio de las medidas de tendencia central y de variabilidad se dedicó más tiempo para el aprendizaje de los conceptos y la resolución de los problemas propuestos.

En el transcurso de las actividades se proporcionaron hojas de trabajo a cada estudiante, que contenían los problemas propuestos para resolver en cada sesión. Los estudiantes resolvieron y pasaron al pizarrón para compartir las respuestas a la comunidad de aprendizaje (docente y compañeros alumnos); con la ayuda del profesor se discutieron propiedades y características del problema y finalmente se solucionaron. Durante las sesiones, los alumnos tenían permitido el uso de calculadoras para realizar operaciones básicas.

En la Tabla 2 de la sección 4.2, se presentó la distribución de las actividades, tiempos y sesiones descritas por semanas y a continuación se detalla cada sesión:

Sesión 1: Aritmética. Esta sesión tuvo la finalidad de reforzar conocimientos deficientes que fueron detectados en la aplicación del cuestionario piloto, para lo cual, los alumnos resolvieron problemas de operaciones básicas que involucraba el uso de símbolos y signos matemáticos, que los estudiantes confundían mucho. Del mismo modo, resolvieron problemas que implicaban el cálculo de fracciones de suma, resta, multiplicación y división.

Sesión 2: Decimales. Los problemas resueltos en esta sesión se enfocaron en el uso de los decimales con el cálculo de las operaciones básicas de la aritmética. También, privilegió el uso de las fracciones equivalentes, puesto que en el cuestionario piloto que se aplicó, se detectó que los jóvenes no tenían claro este concepto.

Sesión 3: Porcentajes. Uno de los problemas detectados y que se tomó la decisión de enfrentar en esta secuencia fue el uso de porcentajes, pues los estudiantes confundían los valores calculados de las medidas de tendencia central con los porcentajes. Por tanto, se decidió para esta sesión proponer problemas relacionados principalmente con la regla de tres simple.

Sesión 4: Variables cualitativas. Se consideró de gran importancia que los estudiantes logren diferenciar las variables cualitativas, por lo que en esta sesión las actividades propuestas fueron encaminadas a resolver problemas que implicaban diferenciar variables categóricas nominales y variables categóricas ordinales. Los problemas propuestos eran un listado de diferentes variables de estudio en las que el estudiante tenía que reconocer el tipo de variable y asignar los valores que posiblemente podía tomar la variable.

Sesión 5: Variables cuantitativas. En esta sesión se realizaron actividades que pretendían el uso de las variables numéricas discretas y las variables numéricas continuas. Al igual que en la sesión anterior, se proponen actividades en las que se da un listado de variables y los estudiantes tienen que identificar el tipo de variable (continua o discreta) y asignar los posibles valores que éstas puedan tomar.

Sesión 6: Distribuciones. Las actividades desarrolladas en esta sesión tienen la finalidad de que el alumno conozca y maneje el concepto de distribución, especialmente que tenga en cuenta el reparto equitativo, para poder identificar propiedades de las medidas de tendencia central, además de reconocer y establecer una distribución de datos cuando se da un valor.

Sesión 7: Media. El desarrollo de esta sesión fue el doble del tiempo que las anteriores, pues los estudiantes necesitaron más tiempo para comprender este concepto. Se propusieron varios problemas en los que se plantearon diferentes situaciones con la finalidad que el estudiante comprenda los momentos en que se debe usar y en los que no es posible hacer uso de la media.

Sesión 8: Mediana y moda. En el desarrollo de estos conceptos se propusieron problemas con diferentes contextos, para que el estudiante identifique la utilidad, y las situaciones en las que es adecuado su uso. También, estuvo orientada a que el alumno logre distinguir problemas en los que la moda era el mejor representante de los datos, y que de la misma manera, pueda identificar y explicar cuando exista más de una moda en los diferentes contextos en que se presentan los problemas

Sesión 9: Rango, Varianza y Desviación estándar. Los problemas propuestos en esta sesión fueron orientados a analizar las medidas de variabilidad, en diferentes contextos y que los estudiantes pudieran decir cuando había menor o mayor dispersión de datos.

Sesiones 10, 11 y 12: Histograma, gráfica de barras y ojiva. En estas sesiones se contempla la lectura de gráficos, en la que los estudiantes conozcan las diferencias de estas tres formas de representación de datos. Los problemas planteados permiten que el estudiante identifique cuándo se pueden usar estos gráficos y cómo realizar la lectura de los mismos. Para lograr este objetivo, se platean problemas con preguntas que incitan a analizar los gráficos.

Sesiones 13,14 y 15: Aplicación del cuestionario final. En estas tres últimas sesiones se aplicó el cuestionario final que nos ayudará a observar si los alumnos han desarrollado los conocimientos sobre las medidas de tendencia central y dispersión. Este instrumento se aplicó en tres sesiones en la sesión 13, los estudiantes resolvieron los problemas 1 y 2; en la sesión 14 resolvieron el problema 3, estos primeros tomados de Mayén (2009). En la sesión 15 los estudiantes resolvieron el último problema, tomado de Orta (2014), que se desarrolló en dos momentos.

4.5.1. Tiempos en horas de estadística en Quinto Semestre

Para llevar a cabo la experiencia, es necesario conocer la distribución del tiempo dedicado al tratamiento de los temas que nos competen, que están incluidos en el programa de estudios de Probabilidad y Estadística. Se explican a continuación:

Bloque I. Comprendes y describes la variabilidad estadística y sus aplicaciones. Cuenta con un tiempo asignado de 10 horas distribuido cada uno en 1 hora por sesión, y en total hay 3 sesiones por semana.

Bloque II. Describes y representas datos de forma tabular y gráfica. Este bloque contempla un tiempo asignado de 12 horas para el desarrollo de las sesiones de enseñanza y evaluación.

Bloque III. Aplicas la estadística descriptiva. Este tercer apartado que es el de interés tiene un tiempo asignado de 16 horas de enseñanza aprendizaje, es decir, 16 sesiones de enseñanza con un tiempo de 50 minutos cada una.

CAPÍTULO 5
ANÁLISIS DE RESULTADOS

5.1. Introducción

En este capítulo se realiza un análisis cualitativo de las respuestas que los estudiantes de la muestra dieron a los 4 problemas elegidos, dado que la muestra es pequeña se analizan las respuestas de todos los alumnos. Los problemas que se eligieron contienen los conceptos de media, mediana, lectura de gráficos y variabilidad, de los que se evalúa la comprensión por medio de análisis de los elementos de significado (campos de problemas, procedimientos, lenguaje, definiciones, propiedades y argumentos).

Se consideraron 2 bloques de análisis de acuerdo a la naturaleza de los problemas. En el primer bloque se incluyen los problemas 1, 2, y 3, que incluyen conceptos únicamente de media, mediana y moda. El segundo bloque solo incluye el problema 4, que introduce el concepto de variabilidad respecto a la media, por lo tanto, el análisis a este problema es más amplio, ya que, por un lado, se analizó media aritmética y construcción de gráficos y por el otro, a partir de la distribución de datos, se analiza la variabilidad y el estudiante tiene que interpretar verbalmente los resultados.

La finalidad de este proceso es describir los elementos de significado utilizados en cada problema e incisos, que resuelven los estudiantes, de tal forma que se identifique la comprensión y los conflictos semióticos cuando el significado que le da el estudiante no es el correcto. A continuación se presentan y discuten los resultados, contrastándolos con los de otras investigaciones.

5.2. Análisis del problema 1

El primer problema está compuesto por dos apartados, en el primero se le pide al alumno que con sus propias palabras de una definición de media. El segundo apartado solicita al alumno que de una distribución de datos que, incluyendo los valores dados, resulte una media dada. Como es una medida ya establecida requiere que el alumno conozca el algoritmo de cálculo de la media para datos aislados, así como la aplicación del algoritmo inverso, además implica buscar una distribución. También tiene un lenguaje numérico y verbal, y solicita un argumento. A continuación se presenta la solución del problema y luego el análisis, para esto se nombra al inciso *"a"* como *problema 1.1* y el inciso *"b"* será *problema 1.2.*

Problema 1

Un periódico dice que el número medio de hijos en México es 2.2 hijos por familia.

a) *Explica qué significa para ti esta frase.*
b) *Se han elegido 10 familias mexicanas y el número medio de hijos entre las 10 familias es de 2.2 hijos por familia. Los García tienen 4 hijos y los Pérez tienen 1 hijo. ¿Cuántos hijos podrán tener las otras 8 familias para que la media de hijos en las 10 familias sea 2.2? Justifica tu respuesta.*

En la primera parte se pudieran concebir las siguientes repuestas, (Mayén, 2009):

"Que el número de hijos a los que tocaría cada familia al repartir el total sería 2.2, pero esto es resultado de una operación, pues el número de hijos siempre es un valor entero"

"Las familias tienen alrededor de 2 hijos, pero algunas familias tienen más o menos".

En la segunda parte, se debe de dar una distribución de datos que incluya los otros dos valores que están en la descripción y que resulte la media dada.

Ejemplo 1:

"Considerando que la familia García tienen 4 hijos y los Pérez tienen 1 hijo, las 8 familias restantes tienen 2 hijos cada una, excepto una que tiene 3, y por tanto la media es 2.2".

4	*1*	*2*	*2*	*2*	*2*	*2*	*2*	*2*	*3*
García	*Pérez*	*F3*	*F4*	*F5*	*F6*	*F7*	*F8*	*F9*	*F10*

Entonces:

$$\bar{x} = \frac{4+1+2+2+2+2+2+2+2+3}{10} = \frac{22}{10} = 2.2\ hijos$$

Ejemplo 2:

"Si la media es de 2.2 y son 10 familias, entonces primero se debe calcular el total de hijos, multiplicando la media 2.2 por 10.

$$2.2\ (10) = 22\ hijos\ en\ total$$

Al total de hijos hay que restar los 5 hijos que conjuntamente tienen los García y los Pérez y quedan 17 hijos que se distribuyen entre 8 familias restantes.

$$22 - 5 = 17\ hijos$$

Ahora esos 17 hijos, se pueden distribuir de varias formas, por ejemplo:

Cada familia podría tener 1, y una de ellas 10 hijos.

1	*1*	*1*	*1*	*1*	*1*	*1*	*10*
F1	*F2*	*F3*	*F4*	*F5*	*F6*	*F7*	*F8*

Dos familias podrían tener 1, cuatro familias 2, una de ellas 3 y una 4 hijos.

1	1	2	2	2	2	3	4
F1	F2	F3	F4	F5	F6	F7	F8

En las Tablas 3 y 4, se describen los elementos de significado que incluye el problema.

Tabla 3. Elementos de significado de la solución correcta del problema (inciso a).

Elementos de significado	**Contenido**
Definiciones	▪ Definición de media como promedio aritmético de un conjunto de datos ▪ Definición de la media como algoritmo de cálculo
Propiedades	▪ La media es un valor perteneciente al rango de la variable ▪ La media no tiene por qué ser uno de los valores de los datos ▪ El cálculo de la media no es operación interna ▪ La media es un representante del conjunto de datos
Campos de problemas	▪ Dar una definición de media con sus propias palabras, considerando la idea de media como reparto equitativo en una distribución de datos ▪ Dar una definición de media con sus propias palabras, considerando la idea del valor más probable al tomar un elemento de una población
Algoritmos y procedimientos	*Algoritmos* ▪ Referencia explícita al cálculo correcto de la media de una variable discreta con datos aislados *Procedimientos* ▪ Interpretar el concepto de media según el contexto del problema
Lenguaje	▪ Verbal ▪ Numérico
Argumentos	▪ Razonamientos verbales deductivos

Tabla 4. Elementos de significado de la solución correcta del problema (inciso b).

Elementos de significado	Contenido
Definiciones	▪ Definición de media como promedio aritmético de un conjunto de datos ▪ Definición de la media como algoritmo de cálculo
Propiedades	▪ La media es un valor perteneciente al rango de la variable ▪ La media no tiene por qué ser uno de los valores de los datos ▪ El cálculo de la media no es operación interna ▪ La media es un representante del conjunto de datos
Campos de problemas	▪ Dada la media de un conjunto de datos, obtener los posibles valores de los elementos del conjunto.
Algoritmos y procedimientos	*Algoritmos* ▪ Cálculo de la media de un conjunto de datos aislados *Procedimientos* ▪ Método de ensayo y error ▪ Invertir algoritmo de cálculo de la media. o *El alumno debe calcular el total de hijos de las 10 familias, multiplicando la media 2.2 por 10.* o *Al total, restar los 5 hijos que conjuntamente tienen las dos familias dadas quedan 17 hijos entre 8 familias.* o *Buscar distribución de los 17 hijos de las 8 familias restantes, para lo cual hay varias posibilidades*
Lenguaje	▪ Verbal, numérico, simbólico
Argumentos	▪ Razonamientos verbales deductivos

Categorías de respuestas en el problema 1.1, antes de la experiencia de aprendizaje.

C1. Media como representante de un conjunto de datos y no justifica. Este grupo de alumnos son los que mejor contestaron a este apartado, se puede identificar en la Figura 27, que tienen un conocimiento del lenguaje estadístico al expresar la media como promedio y como representante de un conjunto de datos, es decir expresan que la media es próxima a la moda, además de observar que la media no necesariamente es parte del conjunto de datos.

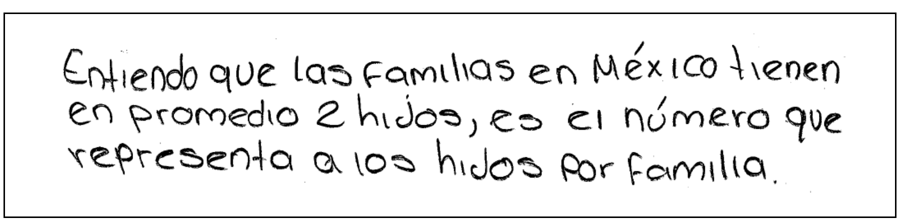


Figura 27 . Ejemplo de respuesta C1.

C2. Error al concebir la idea de la media. Corresponde a los alumnos que confunden la media con la moda, limitan el rango de hijos que pude tener las familias de 2 a 3 y no justifican. También, no logran identificar la propiedad de la media de no ser una operación interna.

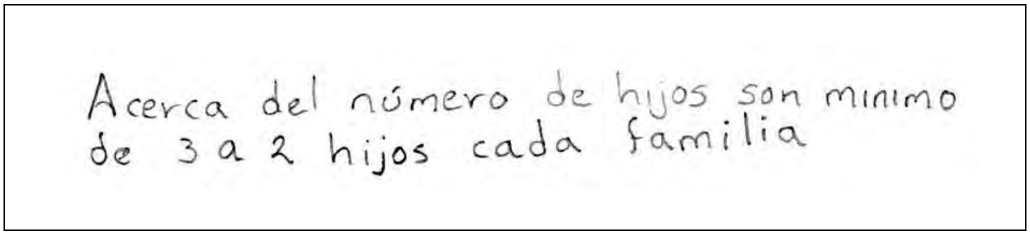

Acerca del número de hijos son minimo de 3 a 2 hijos cada familia

Figura 28 . Ejemplo de respuesta C2.

C3. Respuestas que no tienen relación con el problema. En esta categoría los alumnos entran en conflicto al considerar en algunas ocasiones a la media como la mitad o el medio del conjunto de datos, o porcentajes como se observa en la Figura 29, es decir no conocen el lenguaje estadístico, propiedades y mencionan valores que no se encuentran en la descripción, se detecta falta de concentración y falta de comprensión del problema.

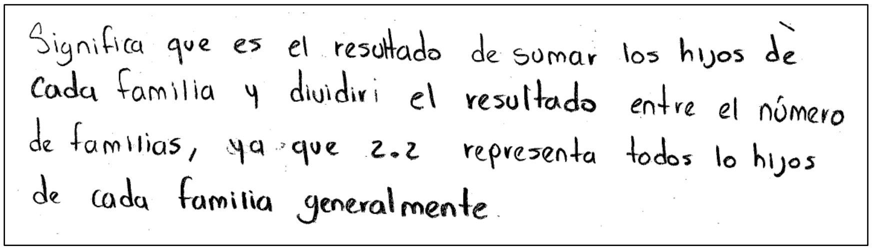

Un porcentaje que corresponde al número de hijos

Figura 29 . Ejemplo de respuesta C3.

C4. No contesta. El grupo de alumnos que ha dejado sin responder este problema.

Categorías de respuestas en el problema 1.1, después de la experiencia de aprendizaje.

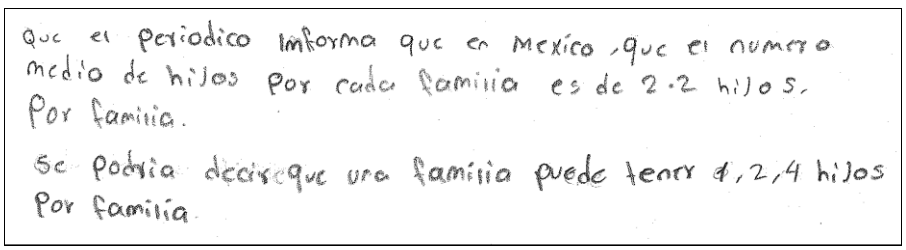
C1. Usa la definición de la media como algoritmo de cálculo. En la Figura 30, se observa que los alumnos explican el algoritmo de cálculo de la media de una variable discreta con datos aislados, por lo que se deduce el uso correcto de la definición de media. Además hacen referencia a que es un representante del colectivo.



Significa que es el resultado de sumar los hijos de cada familia y dividiri el resultado entre el número de familias, ya que 2.2 representa todos lo hijos de cada familia generalmente.

Figura 30. Ejemplo de respuesta C1.

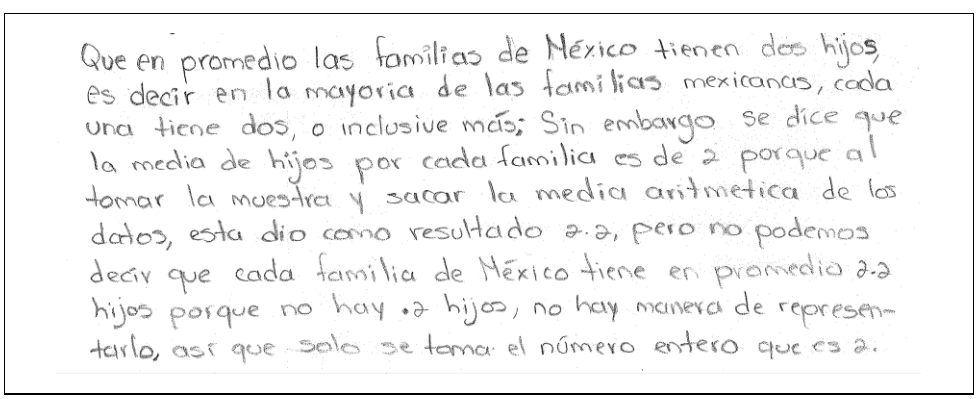
C2. Media como representante de los datos y como valor más probable en una población. Los alumnos de esta categoría reconocen la propiedad de la media de ser un representante de un colectivo, repiten el enunciado del problema y dan una definición de media con sus propias palabras, considerando la idea del valor más probable al tomar un elemento de una población al mencionar que las familias pueden tener 1,2 o 4 hijos, como se observa en la Figura 31.



Que el periodico informa que en Mexico, que el numero medio de hijos por cada familia es de 2.2 hijos. Por familia.
Se podria decir que una familia puede tener 1, 2, 4 hijos Por familia.

Figura 31. Ejemplo de respuesta C2.

C3. Confunde media con moda, pero concluye correctamente. En esta categoría los alumnos comenten un error en la definición de la media como valor representativo, esto es debido a que confunden la media con la moda. Se observa en la Figura 32 que usan la expresión "mayoría" para explicar la media del número de hijos por familia, por lo que se deduce que no se está identificando bien el término relacionado con concepto de media.



Que en promedio las familias de México tienen dos hijos, es decir en la mayoria de las familias mexicanas, cada una tiene dos, o inclusive más; Sin embargo se dice que la media de hijos por cada familia es de 2 porque al tomar la muestra y sacar la media aritmetica de los datos, esta dio como resultado 2.2, pero no podemos decir que cada familia de México tiene en promedio 2.2 hijos porque no hay .2 hijos, no hay manera de represen-tarlo, así que solo se toma el número entero que es 2.

Figura 32. Ejemplo de respuesta C3.

C4. Error en la propiedad de la media de no ser operación interna. Esta categoría la integran las respuestas en la que los alumnos comenten el error en la propiedad de no ser operación interna, pues los alumnos argumentan que no se puede obtener un valor no entero para la media de una variable discreta, se puede observar en la Figura 33.

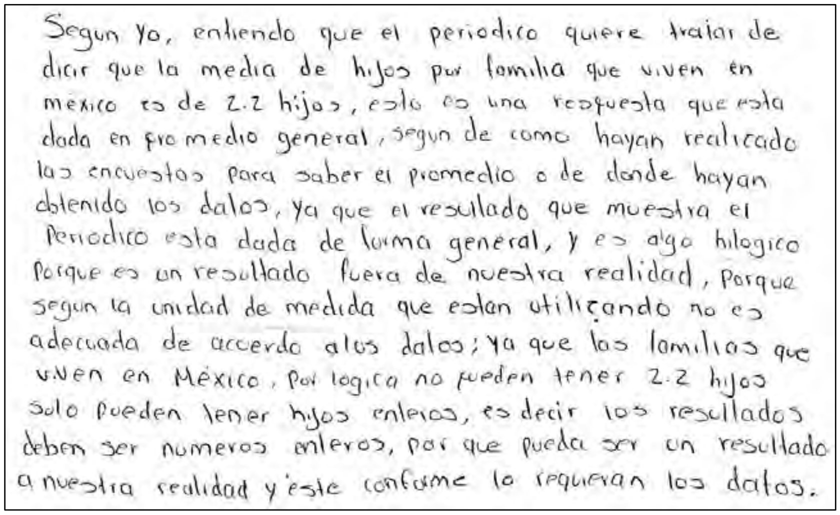

Segun yo, entiendo que el periodico quiere tratar de dicir que la media de hijos por familia que viven en mexico es de 2.2 hijos, esto es una respuesta que esta dada en promedio general, segun de como hayan realizado las encuestas para saber el promedio o de donde hayan obtenido los datos, ya que el resultado que muestra el Periodico esta dada de forma general, y es algo hilogico porque es un resultado fuera de nuestra realidad, porque segun la unidad de medida que estan utilizando no es adecuada de acuerdo alos datos; ya que las familias que viven en México, por logica no pueden tener 2.2 hijos solo pueden tener hijos enteros, es decir los resultados deben ser numeros enteros, por que pueda ser un resultado a nuestra realidad y este conforme lo requieran los datos.

Figura 33. Ejemplo de respuesta C4.

C5. Respuestas que no tienen relación con el problema. Esta categoría concentra las respuestas que dan los estudiantes erróneas o sin sentido, por lo que se consideraron como no relacionadas con el problema.

Seguidamente se agruparon las categorías obtenidas en el análisis, cada una con su respectiva frecuencia con el antes y después de la experiencia de aprendizaje para poder analizar los cambios, para la primera parte del problema 1.

Tabla 5 . Categorías de respuesta del (inciso a), antes y después de la experiencia

Categorías de respuestas del problema 1.1	Antes	Después
C1. Media como representante de un conjunto de datos y no justifica	3	
C2. Error al concebir la idea de la media	2	
C3. Respuestas que no tienen relación con el problema	8	3
C4. No contesta	4	
C1. Usa la definición de la media como algoritmo de cálculo		2
C2. Media como representante de los datos y como valor más probable en una población		3
C3. Confunde media con moda		3
C4. Error en la propiedad de la media de no ser operación interna		6
Total	**17**	**17**

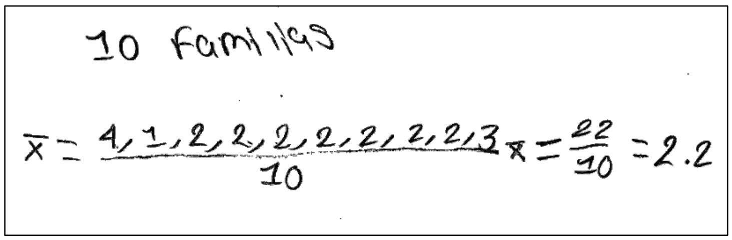
En la Tabla 5, se observa que antes de la experiencia de aprendizaje 8 estudiantes no lograron comprender el campo del problema y dan respuestas que no tienen relación con la media, esta misma categoría se observa después de la experiencia de aprendizaje pero en menos proporción es decir solo 3 estudiantes aun no lograron comprender el campo del problema.

Por otro lado se tienen las categorías siguientes que no se repitieron después de la experiencia, es decir únicamente se observaron antes de la experiencia de aprendizaje, de los cuales 4 estudiantes no contestan a este problema, pudiera ser que no comprenden el campo del problema o no logran entender el lenguaje estadístico. En el mismo sentido, se considera que 2 alumnos conciben de forma errónea la idea de la media esto porque la confunden con la mediana o con porcentajes. En la última categoría que se menciona, 3 estudiantes señalan el valor de la media como un valor que representa al conjunto de datos, pero no argumentan al hacer esta declaración.

Después de la experiencia de aprendizaje surgen nuevas categorías como se puede observar en la Tabla 5, donde se muestra que 6 estudiantes interpretan de manera errónea la propiedad de no ser operación interna, un problema que a pesar de la instrucción recibida no logran comprender los resultados posibles para una variable discreta, pues argumentan que por ser de este tipo no debe tener números decimales. Por otro lado, 3 estudiantes confunden la media con la moda, esto al usar la expresión “mayoría” con la que se nota que los estudiantes no utilizan el término correcto para explicar el concepto que se está usando. Pocos estudiantes, es decir, 5 fueron capaces de dar una respuesta correcta a este apartado, una parte de ellos usaron la definición de media como algoritmo haciendo referencia explícita del proceso de cálculo de la media, el resto destacan la idea de representatividad de la media y consideran la idea del valor más probable, por lo que se deduce que conocen el concepto de la media y su interpretación.

Categorías de respuestas en el problema 1.2, antes de la experiencia de aprendizaje.

C1. Busca distribución dada la media y no justifica. Los alumnos de esta categoría utilizan el método de ensayo y error para encontrar la distribución de datos correspondiente, algunos emplean un procedimiento con el uso del algoritmo para el cálculo de la media, pero no aplican el inverso del algoritmo, sin embargo respecto al lenguaje matemático hay deficiencias en el aritmético, pues se puede ver en la Figura 34, que no asignan el símbolo "+" para indicar la suma de los datos, además de no emitir un argumento de su respuesta.

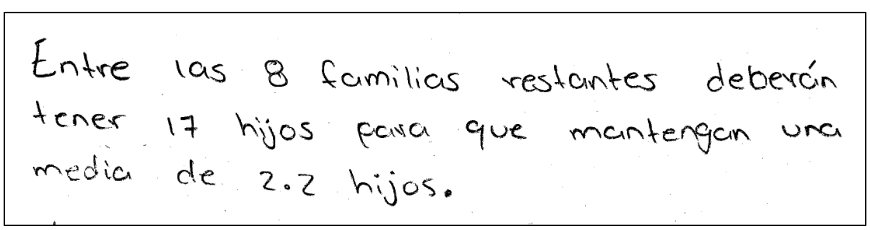


10 famillas

$\bar{x} = \frac{4,1,2,2,2,2,2,2,2,3}{10}$ $\bar{x} = \frac{22}{10} = 2.2$

Figura 34 . Ejemplo de respuesta C1.

C2. Indica el número de hijos en las 8 restantes familias, pero no hace la distribución. Este grupo de alumnos sabe la cantidad de hijos que deben contemplar las familias restantes y toman en cuenta los valores dados en la descripción del problema, pero no logran hacer la distribución de los hijos en las 8 familias restantes. Podría decirse que, no conocen el concepto de distribución, tiene la idea de la media, mas no aplican el inverso del algoritmo para llegar a la respuesta.



Entre las 8 familias restantes deberán tener 17 hijos para que mantengan una media de 2.2 hijos.

Figura 35. Ejemplo de respuestade respuesta C2.

C3. No tiene relación con la distribución de datos de la media dada. Los alumnos en esta categoría no han logrado comprender el problema, aunado a esto se identificaron problemas con el concepto de números decimales y porcentajes. Se observa nulo lenguaje estadístico y desconocen el concepto de distribución.

C4. No contesta. El grupo de alumnos que ha dejado sin responder este problema.

Categorías de respuestas en el problema 1.2, después de la experiencia de aprendizaje.

C1. Busca distribución dada la media. Aquí se integran las respuestas de los estudiantes que se han considerado correctas, pues dan una distribución de los posibles valores para esta parte del problema. A continuación se describen las variantes encontradas de esta categoría.

C1.1. Usa el método de ensayo y error para dar una distribución. En estas respuestas los alumnos tienen nociones del concepto de distribución, consideran los datos de las familias dadas y con base en eso, asignan valores mediante el método de ensayo y error a las 8 familias restantes, de tal manera que al aplicar el algoritmo de la media ésta sea de 2.2, además argumentan que los hijos que pueden tener estas familias se encuentran en un rango de 1 a 4 de la variable, y se deduce que comprenden que el resultado de la media es el valor más probable. También, se observa en la Figura 36, el uso de un lenguaje numérico y simbólico. Con error en la interpretación.

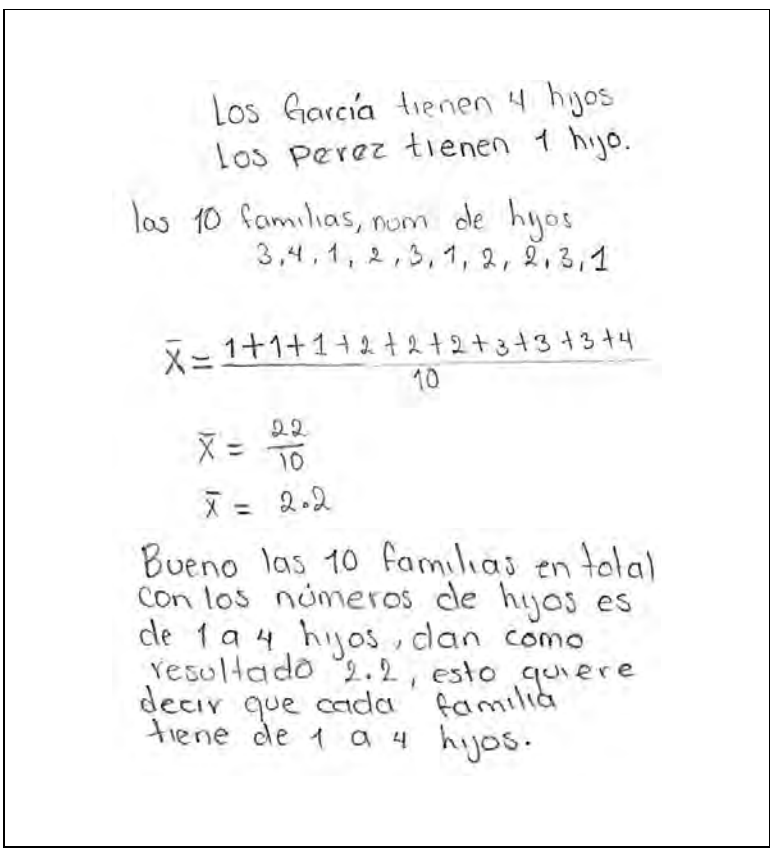
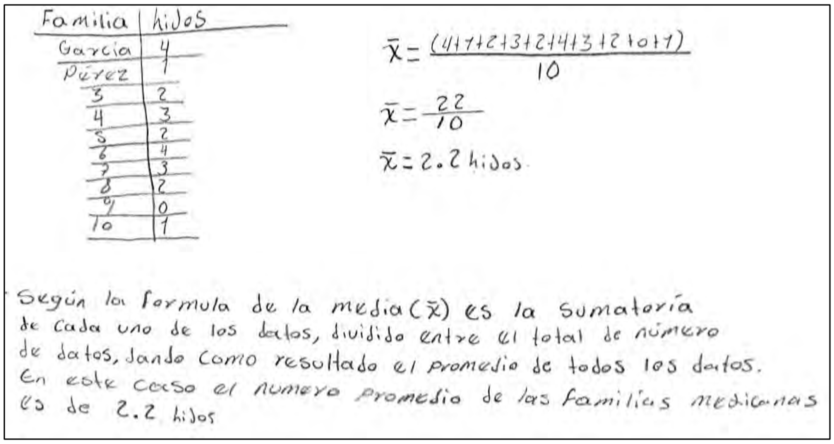


Figura 36. Ejemplo de respuesta C1.1.

C1.2. Da distribución y explica el algoritmo de cálculo de la media. Los alumnos usando el método de ensayo y error, como se observa en la Figura 37, dan una distribución de los posibles datos, seguidamente obtienen la media, con ellos, haciendo uso del algoritmo y por último explican el procedimiento de cálculo de la media



Familia	hidos
García	4
Pérez	1
3	2
4	3
5	2
6	4
7	3
8	2
9	0
10	1

$$\bar{x} = \frac{(4+1+2+3+2+4+3+2+0+1)}{10}$$

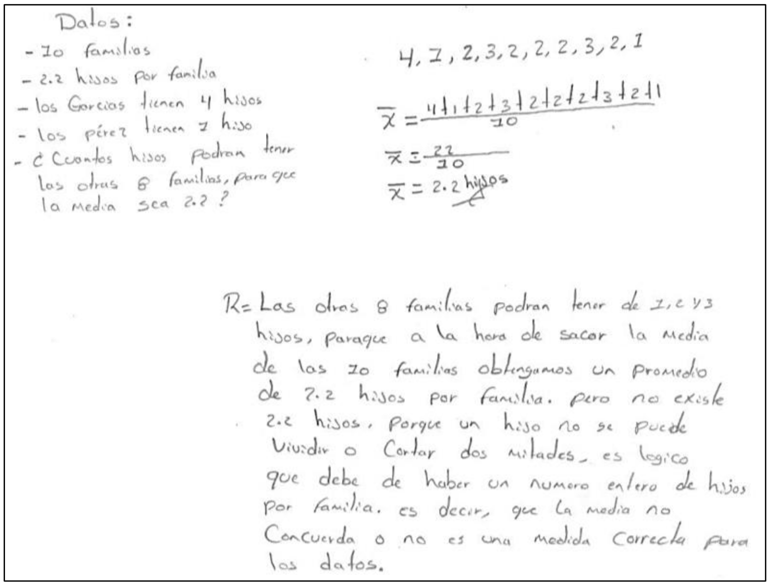
$$\bar{x} = \frac{22}{10}$$

$$\bar{x} = 2.2 \text{ hidos}$$

Según la formula de la media ($\bar{x}$) es la sumatoria de cada uno de los datos, dividido entre el total de número de datos, dando como resultado el promedio de todos los datos. En este caso el numero promedio de las familias mexicanas es de 2.2 hijos

Figura 37. Ejemplo de respuesta C1.2.

C1.3. Da distribución, pero comete error al considerar la media como operación interna. En esta variante se observa en la Figura 38, que los estudiantes también proporcionan una distribución mediante el método de ensayo y error, además de justificar el cálculo de la media con algoritmos. No obstante, argumentan que para un valor entero la media también debería de ser un valor entero. Por lo que señalan que no es una medida correcta para este problema.


Datos:
- 10 familias
- 2.2 hijos por familia
- los Garcias tienen 4 hijos
- los pérez tienen 1 hijo
- ¿Cuantos hijos podran tener los otras 8 familias, para que la media sea 2.2?

4, 1, 2, 3, 2, 2, 2, 3, 2, 1

$\bar{x} = \frac{4+1+2+3+2+2+2+3+2+1}{10}$

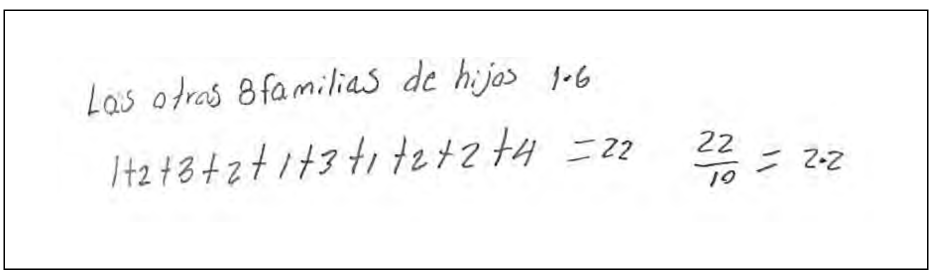
$\bar{x} = \frac{22}{10}$

$\bar{x} = 2.2$ hijos

R= Las otras 8 familias podran tener de 1, 2 y 3 hijos, paraque a la hora de sacar la media de los 10 familias obtengamos un promedio de 2.2 hijos por familia. pero no existe 2.2 hijos, porque un hijo no se puede divirdir o cortar dos mitades, es logico que debe de haber un numero entero de hijos por familia. es decir, que la media no concuerda o no es una medida correcta para los datos.

Figura 38. Ejemplo de respuesta C1.3.

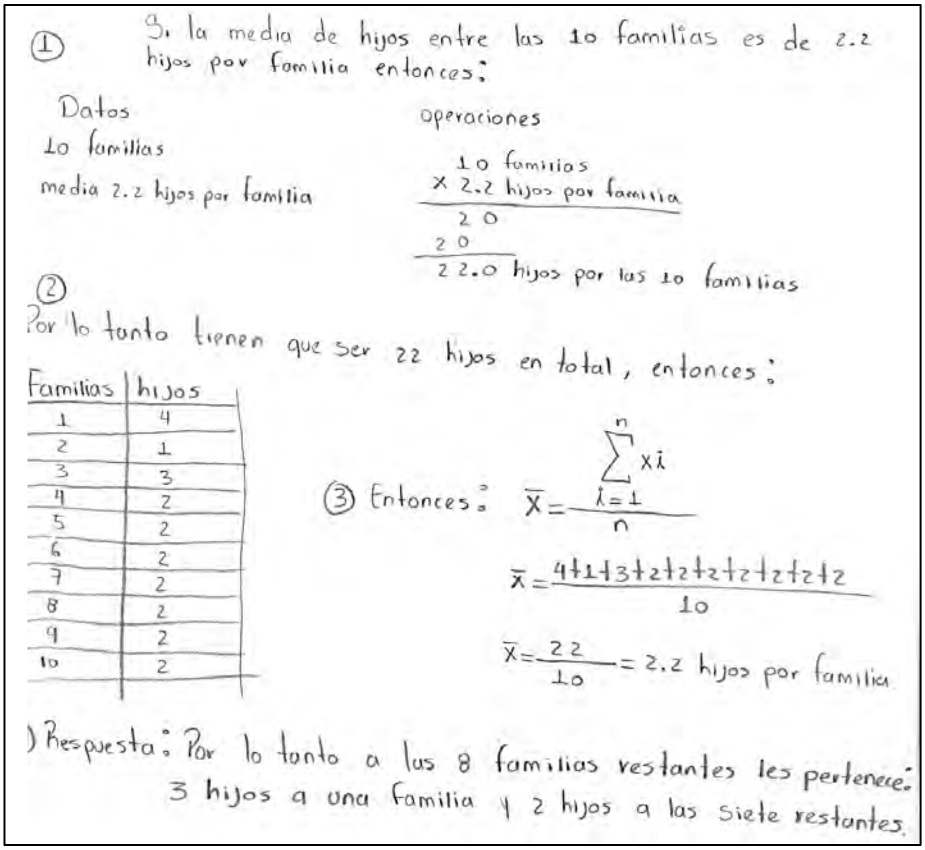
C1.4. Obtiene la distribución con errores de cálculos aritméticos. Esta categoría está integrada por un alumno que concibe la idea de distribución, y mediante el método de ensayo y error asigna valores a la posible distribución. Sin embargo, se observa (Figura 39) que el alumno comete un error en la distribución y en el cálculo aritmético al estimar la media (suma de los datos).


Las otras 8familias de hijos 1-6

$1+2+3+2+1+3+1+2+2+4 = 22 \quad \frac{22}{10} = 2.2$

Figura 39. Ejemplo de respuesta C1.4.

C2. Da distribución dada la media, usando el algoritmo inverso de la media. Los estudiantes conocen el algoritmo de la media y lo invierten para encontrar la distribución. En la Figura 40, se observa que los estudiantes fueron capaces de identificar los elementos de significado al resolver el problema.



(1) Si la media de hijos entre las 10 familias es de 2.2 hijos por familia entonces:

Datos
10 familias
media 2.2 hijos por familia

operaciones

$$\begin{array}{r} 10 \text{ familias} \\ \times\ 2.2 \text{ hijos por familia} \\ \hline 20 \\ 20\ \ \\ \hline 22.0 \end{array}$$ hijos por las 10 familias

(2)
Por lo tanto tienen que ser 22 hijos en total, entonces:

Familias	hijos
1	4
2	1
3	3
4	2
5	2
6	2
7	2
8	2
9	2
10	2

(3) Entonces: $\bar{x} = \frac{\sum_{i=1}^{n} xi}{n}$

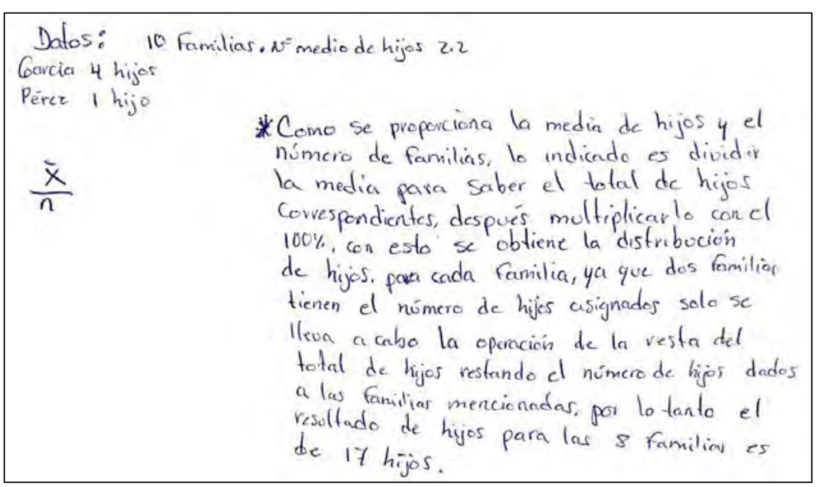
$\bar{x} = \frac{4+1+3+2+2+2+2+2+2+2}{10}$

$\bar{x} = \frac{22}{10} = 2.2$ hijos por familia

) Respuesta: Por lo tanto a las 8 familias restantes les pertenece: 3 hijos a una familia y 2 hijos a las siete restantes.

Figura 40.Ejemplo de respuesta C2.1.

C3. No usa la idea de distribución. En la Figura 41, de esta categoría dos alumnos explican el procedimiento de manera verbal, pudiera decirse que invierten el algoritmo y obtienen la suma de las 8 familias restantes, pero no realizan una distribución de los datos como lo solicita este apartado del problema, por lo que se puede decir que desconocen el concepto de distribución.



Datos: 10 Familias. Nº medio de hijos 2.2
García 4 hijos
Pérez 1 hijo

$\frac{\bar{x}}{n}$

*Como se proporciona la media de hijos y el número de familias, lo indicado es dividir la media para saber el total de hijos correspondientes, después multiplicarlo con el 100%, con esto se obtiene la distribución de hijos. para cada Familia, ya que dos familias tienen el número de hijos asignados solo se lleva a cabo la operación de la resta del total de hijos restando el número de hijos dados a las familias mencionadas, por lo tanto el resultado de hijos para las 8 familias es de 17 hijos.

Figura 41. Ejemplo de respuesta C3

C4. Error al considerar la media como operación interna. El alumno hace referencia explícitamente a un reparto equitativo (Figura 42), pero entiende que este reparto se debe de dar en los hijos que suman la familia García y los Pérez. Del mismo modo, confunde el valor N (número de datos en la muestra), que aparece en el denominador del algoritmo de cálculo de la media, y el número de valores conocidos de dichos datos (número de hijos de los García y Pérez). También, en el aspecto simbólico tiene nociones de cómo se escribe la fórmula, pero comete error de escritura. Por último el argumento verbal no tiene sentido.

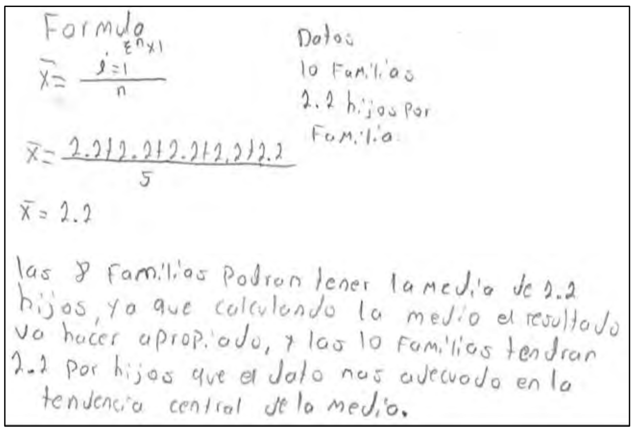


Figura 42. Ejemplo de respuesta C4.

C5. Asignan valores con decimales a los posibles datos de la distribución y confunden media con mediana. La Figura 43, muestra que los alumnos asignan valores a las 8 familias restantes, pero se olvidan de la variable discreta que están trabajando y los valores asignados contienen decimales haciendo probablemente una distribución con reparto equitativo. Por otro lado, obtiene la mediana pero no ordena los datos, explica solo una parte del procedimiento para estimar la mediana.

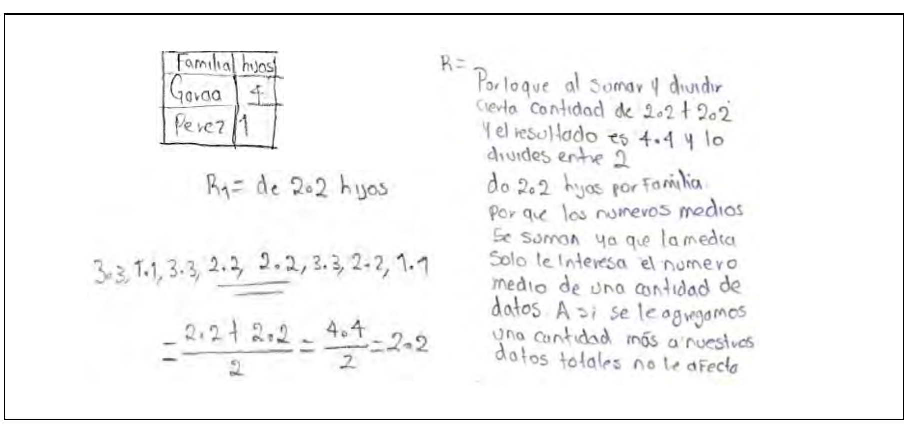


Figura 43. Ejemplo de respuesta C5.

El siguiente punto trata de agrupar las categorías obtenidas en el análisis de la segunda parte del problema 1, y se enlistan en la Tabla 6, cada una con su frecuencia.

Tabla 6 . Categorías de respuesta del (inciso b), antes y después de la experiencia

Categorías de respuestas del problema 1.2	**Antes**	**Después**
C1. Busca distribución dada la media y no justifica	4	
C2. Indica el número de hijos en las 8 restantes familias, pero no hace la distribución	6	
C3. No tiene relación con la distribución de datos de la media dada	6	
C4. No contesta	1	
C1.1. Usa el método de ensayo y error para dar una distribución		4
C1.2. Da distribución y explica el algoritmo de cálculo de la media		2
C1.3. Da distribución, pero comete error al considerar la media como operación interna		3
C1.4. Obtiene la distribución con errores de cálculos aritméticos		1
C2. Da distribución dada la media, usando el algoritmo inverso de la media		2
C3. No usa la idea de distribución		2
C4. Error al considerar la media como operación interna		1
C5. Asignan valores con decimales a los posibles datos de la distribución y confunden media con mediana		2
Total	**17**	**17**

En la Tabla 6 donde se puede observar la frecuencia del antes y después de la experiencia de aprendizaje, no coincidieron categorías que se repitieran, por un lado antes de la experiencia de aprendizaje la mayoría de los estudiantes de la muestra se concentra en dos categorías cada una con 6 de estudiantes, en la primera los estudiantes logran encontrar el número de hijos que deberían sumar las familias restantes, pero no logran hacer la distribución de ese total, es decir a cuantos hijos le corresponde a las familias restantes, por lo que se deduce que los estudiantes tienen nociones básicas del lenguaje aritmético. En la segunda categoría más frecuente con la misma cantidad de alumnos (6), las respuestas que proporcionan no tienen relación con la distribución de datos cuando se da el valor de la media, es decir no tienen sentido en el contexto planteado, probablemente no conocen el lenguaje estadístico, o bien, no comprendieron el campo del problema. Seguidamente 4 estudiantes obtienen una distribución cuando se les proporciona la media, sin embargo no justifica su respuesta que emiten. En la última categoría encontrada antes de la experiencia de aprendizaje, un estudiante no contesta este apartado del problema.

Después de la experiencia de aprendizaje aumentó la variedad de categorías de respuestas no se repitieron pero si surgieron nuevas, por ejemplo: 9 estudiantes fueron capaces de dar una distribución, se deduce que utilizando el método de ensayo y error, 2 jóvenes explican el algoritmo de la media, 3 alumnos cometen el error de argumentar que no es posible ese resultado pues consideran que al trabajar con números enteros sus resultados deben ser del mismo tipo. Solamente, dos alumnos fueron capaces de identificar los elementos de significados que se describen en la Tabla 4 para la solución correcta.

Un aporte importante, es que 2 alumnos han asignado valores con decimales para la posible distribución de datos, haciendo pensar que tienen la idea de reparto equitativo, y además tienen la confusión entre media y mediana.

5.3. Análisis del problema 2

El segundo problema se compone de dos partes, la primera solicita a los estudiantes el cálculo de la media ponderada partiendo de la descripción verbal de un problema donde intervienen unos números muy sencillos (Mayén, 2009), para este apartado se requiere que los estudiantes conozcan las definiciones de media, como algoritmo y como promedio. La segunda parte pide el cálculo de la media ponderada, además de indagar si los alumnos reconocen la propiedad de que *"la media de la suma de dos o más variables es igual a la suma de las medias de éstas"*. En seguida se presenta la solución del problema y posteriormente se hace el análisis, de la misma forma que el problema anterior se nombra al inciso *"a"* como *problema 2.1* y el inciso *"b"* será *problema 2.2.*

Problema 2

María y Pedro dedican una media de 8 horas cada fin de semana a hacer deporte. Otros 8 estudiantes dedican cada semana una media de 4 horas a hacer deporte.

a) *¿Cuál es el número medio de horas que hacen deporte cada fin de semana los 10 estudiantes?*
b) *María y Pedro dedican además 1 hora cada fin de semana a escuchar música y los otros 8 estudiantes, 3 horas. ¿Cuál sería el número medio de horas que estos 10 estudiantes dedican cada fin de semana, entre las dos actividades: hacer deporte y escuchar música?*

En la primera parte, llegar a la solución correcta implica que en la ponderación del cálculo de la media se aplique la propiedad distributiva de la suma y producto y ver que la operación promedio no tiene la propiedad asociativa.

Para resolver el problema se debe multiplicar la frecuencia de alumnos que dedican 8 horas a hacer deporte (2 alumnos) por 8; Luego se multiplica la frecuencia de estudiantes que dedica 4 horas a hacer deporte (8 alumnos) por 4; Seguidamente sumar los valores que resulten y se obtiene la media dividiendo entre el total de datos que es 10.

$$\bar{x} = \frac{2(8) + 8(4)}{10} = \frac{16 + 32}{10} = \frac{48}{10} = 4.8\ horas\ hacen\ deporte$$

En la segunda parte del problema se vuelve aplicar el cálculo de la media ponderada y se debe reconocer la propiedad de que *"la media de la suma de dos o más variables es igual a la suma de las medias de éstas"*.

La suma de las medias:

El número medio que hacen deporte

$$\bar{x} = \frac{2(8) + 8(4)}{10} = \frac{16 + 32}{10} = \frac{48}{10} = 4.8\ horas\ hacen\ deporte$$

El número medio que escuchan música

$$\bar{x} = \frac{2(1) + 8(3)}{10} = \frac{2 + 24}{10} = \frac{26}{10} = 2.6\ horas\ escuchan\ deporte$$

Suma de las medias

$$2.6 + 4.8 = 7.4\ horas\ escuchan\ musica\ y\ hacen\ deporte$$

O bien;

$$\bar{x} = \frac{2(8) + 8(4)}{10} + \frac{2(1) + 8(3)}{10}$$

$$\bar{x} = 2.6 + 4.8$$

$$\bar{x} = 7.4\ horas\ escuchan\ musica\ y\ hacen\ deporte$$

La media de la suma de dos o más variables:

$$\bar{x} = \frac{2(8) + 8(4) + 2(1) + 8(3)}{10} = \frac{74}{10} = 7.4\ horas\ escuchan\ musica\ y\ hacen\ deporte$$

En la Tabla 7 y 8, se describen los elementos de significado que incluye el problema para cada inciso.

Tabla 7. Elementos de significado de la solución correcta del problema (inciso a).

Elementos de significado	Contenido
Definiciones	▪ Definición de media como promedio aritmético de un conjunto de datos.. ▪ Definición de la media como algoritmo de cálculo
Propiedades	▪ La operación promedio no tiene la propiedad asociativa
Campos de problemas	▪ Hallar la media ponderada, dadas las medias
Algoritmos y procedimientos	*Algoritmos* ▪ Cálculo de la media con datos aislados ▪ Cálculo de la media ponderada *Procedimientos* ▪ Multiplicar la frecuencia de alumnos que dedica 8 horas a hacer deporte (2 alumnos) por 8 ▪ Multiplicar la frecuencia de alumnos que dedica 4 horas a hacer deporte (2) por 4 ▪ Sumar los valores resultantes ▪ Estimar la media global dividiendo entre el total de datos (10)
Lenguaje	▪ Verbal, numérico, simbólico
Argumentos	Razonamientos verbales deductivos

Tabla 8. Elementos de significado de la solución correcta del problema (inciso b).

Elementos de significado	Contenido
Definiciones	▪ Definición de media como promedio aritmético de un conjunto de datos. ▪ Definición de la media como algoritmo de cálculo
Propiedades	▪ La operación promedio no tiene la propiedad asociativa ▪ La media de la suma de dos o más variables, es la suma de las medias de éstas.
Campos de problemas	▪ Hallar la media ponderada, dadas las medias
Algoritmos y procedimientos	*Algoritmos* ▪ Cálculo de la media con datos aislados ▪ Cálculo de la media ponderada *Procedimientos* ▪ Multiplicar la frecuencia de alumnos que dedica 8 horas a hacer deporte (2 alumnos) por 8 ▪ Multiplicar la frecuencia de alumnos que dedica 4 horas a hacer deporte (2) por 4 ▪ Sumar los valores resultantes ▪ Estimar la media global dividiendo entre el total de datos (10)
Lenguaje	Verbal, numérico, simbólico
Argumentos	Razonamientos verbales deductivos

Categorías de respuestas en el problema 2.1, antes de la experiencia de aprendizaje.

C1. Calcula correctamente una media ponderada. Los estudiantes de esta categoría han logrado calcular la ponderación correcta de los valores de la variable, por lo que se observa que conocen la propiedad distributiva y comprende que la operación promedio no tiene la propiedad asociativa. En la Figura 44, se observa que hace el uso de algoritmos para justificar sus resultados, la única recomendación que se le daría es que escriba las unidades de medida que se están trabajando para reforzar su respuesta.

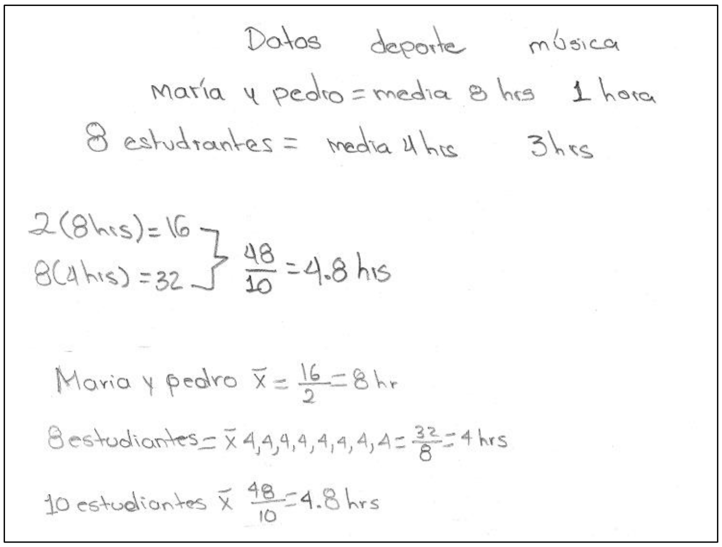


Figura 44 . Ejemplo de respuesta C1

C2. Calcula la media ponderada y no justifica. En esta categoría los alumnos conocen el algoritmo de la media, pues reconocen que hay que trabajar con el inverso y llegan a la respuesta correcta, sin argumentar su resultado, sin embargo persisten algunos problemas del lenguaje aritmético con el símbolo "+" como se puede ver en la Figura 45. Al momento de trabajar en la experiencia de aprendizaje se recomienda que se le haga hincapié al estudiante del uso de las unidades de medida y los argumentos que se deben de redactar, así también se debe reforzar el concepto de media ponderada para que se realice un cálculo más adecuado.

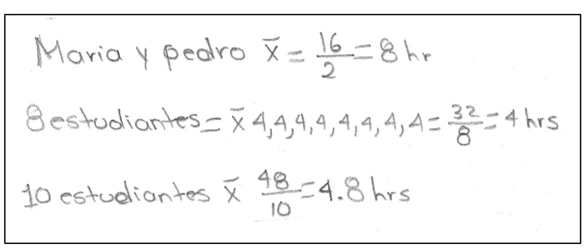


Figura 45 . Ejemplo de respuesta C2.

C3. Respuesta correcta sin justificarla mediante algoritmos. Los estudiantes responden correctamente al problema, sin embargo se observa en la Figura 46 que no justifican mediante algoritmos el cálculo de la media ponderada, ni dan argumento alguno. Por tanto, no se puede obtener algún tipo de información sobre la capacidad de cálculo.

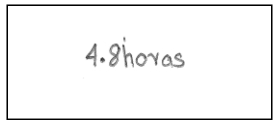


Figura 46 . Ejemplo de respuesta C3.

C4. No tiene relación con la media ponderada. Se observa en esta categoría que los estudiantes no logran comprender el campo de problema, pudiera deberse a deficiencias en el lenguaje estadístico, en algunas respuestas se identificó que confunden la media con la mediana. Es importante inducir el concepto de media ponderada y el algoritmo inverso de la media, para facilitar la resolución de este tipo de problemas.

Categorías de respuestas en el problema 2.1, después de la experiencia de aprendizaje.

C1. Respuestas correctas basadas en la media simple. En esta categoría se encuentran las respuestas que han dado los alumnos de la muestra en la que prefieren resolver el problema haciendo uso de la media simple. En seguida se describe algunas variantes.

C1.1. Pondera correctamente los valores, pero realiza el cálculo de la media con datos aislados. El estudiante es capaz de hacer la ponderación, pero no suma los resultados para aplicar el algoritmo de la media, sino que, hace una distribución equitativa, después obtienen la media simple, de lo que se deduce el uso de la decisión de media como promedio. En la Figura 47, se observa uso de lenguaje numérico y simbólico.

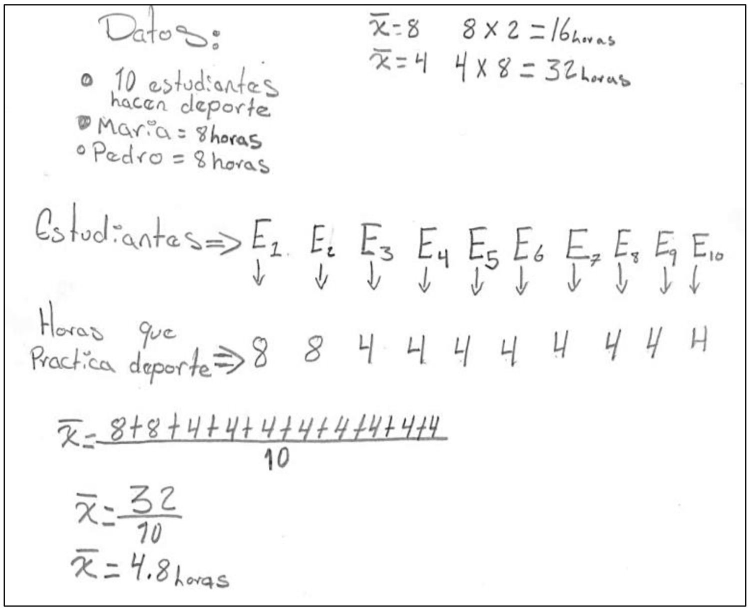


Figura 47. Ejemplo de respuesta C1.1.

C1.2. Correcta, mediante el cálculo de la media de una variable discreta con datos aislados. En esta variante, los estudiantes dan respuesta usando el algoritmo de la media simple, no ponderan, pero se deduce que saben la definición de media como promedio. Además, representan los datos en una tabla de distribución de frecuencias. Se observa (Figura 48) el uso correcto del algoritmo de la media, sus símbolos y procedimientos.

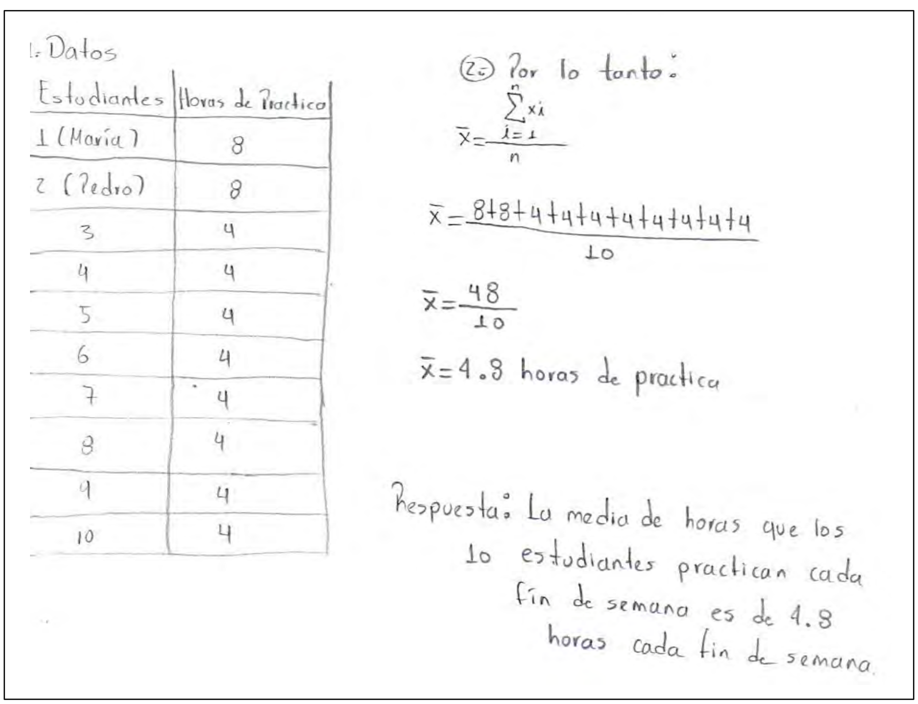
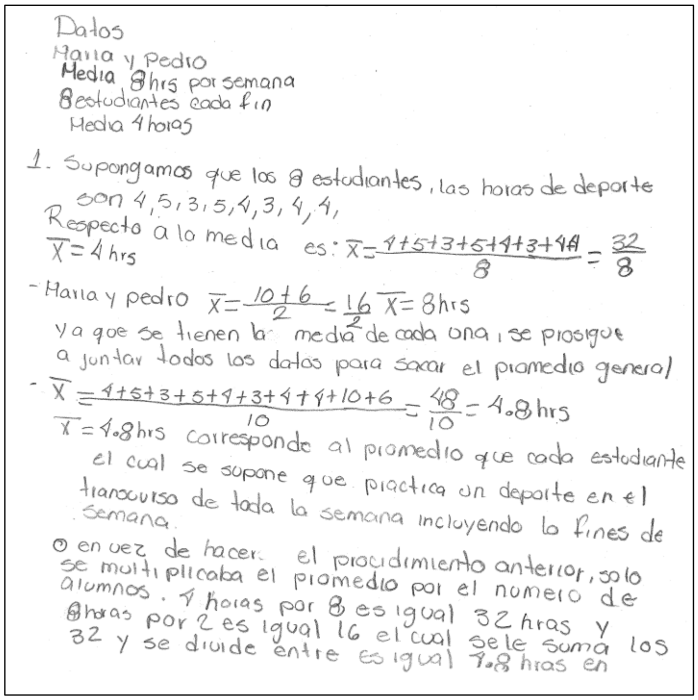


Figura 48. Ejemplo de respuesta C1.2.

C1.3. Correcta, con distribución no equitativa. El estudiante que obtiene esta variante, calcula la media de un conjunto de datos aislados, pero usa una distribución con valores diferentes, mediante el método de ensayo y error encuentra una distribución que al sumar todos los datos y dividirla por el número total de datos de la respuesta correcta. De tal modo, que antes comprueba para cada grupo de datos. Finalmente, se observa (Figura 49) el uso del lenguaje numérico, verbal y simbólico, además de dar argumentos deductivos.



Datos
Maria y Pedro
Media 8hrs por semana
8 estudiantes cada fin
Media 4 horas

1. Supongamos que los 8 estudiantes, las horas de deporte son 4, 5, 3, 5, 4, 3, 4, 4,
Respecto a la media es: $\bar{X} = \frac{4+5+3+5+4+3+4+4}{8} = \frac{32}{8}$
$\bar{X} = 4$ hrs

- Maria y pedro $\bar{X} = \frac{10+6}{2} = \frac{16}{2}$ $\bar{X} = 8$hrs
ya que se tienen la media de cada una, se prosigue a juntar todos los datos para sacar el promedio general
- $\bar{X} = \frac{4+5+3+5+4+3+4+4+10+6}{10} = \frac{48}{10} = 4.8$ hrs
$\bar{X} = 4.8$hrs corresponde al promedio que cada estudiante el cual se supone que practica un deporte en el transcurso de toda la semana incluyendo lo fines de semana.

o en vez de hacer el procedimiento anterior, solo se multiplicaba el promedio por el numero de alumnos. 4 horas por 8 es igual 32 hras y 8horas por 2 es igual 16 el cual se le suma los 32 y se divide entre es igual 4.8 hras en

Figura 49. Ejemplo de respuesta C1.3.

C2. Usar una media simple, sin tener en cuenta las ponderaciones. Categoría compuesta por dos variantes que hacen referencia a que el alumno no ha comprendido la idea de la media ponderada, y en algunos casos comete error en las propiedades y el algoritmo de cálculo de la media simple.

C2.1. Obtiene la media de las dos medias dadas. En esta variante los estudiantes cometen el error de obtener el promedio aritmético de las dos medias dadas. Se infiere que conocen el algoritmo de cálculo de la media, pues lo hacen explicito verbalmente (Figura 50).

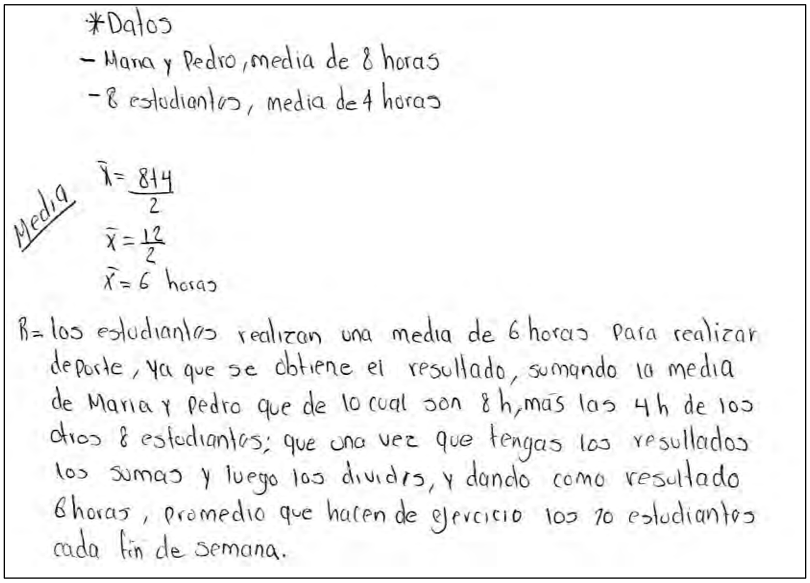


Figura 50. Ejemplo de respuesta C2.1.

C2.2. Error en la propiedad de la media de no ser asociativa. Los alumnos nos dan respuestas en las que se observa (Figura 51) que no consideran la ponderación y el uso incorrecto de la media simple, es decir, cometen el error en la definición de la media como algoritmo de cálculo, debido a que suman las medias dadas y lo dividen por el número total de datos.

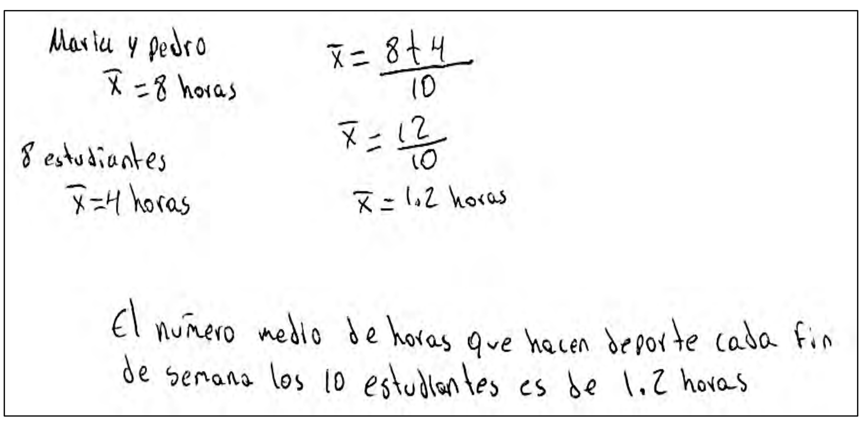


Figura 51. Ejemplo de respuesta C2.2.

C3. Usar la mediana, en vez de la media. En este grupo de respuestas los estudiantes confunden las medidas de tendencia central e incluso las operaciones de una misma medida. En lo que sigue se describen las variantes.

C3.1. Confunde la media con la mediana. El alumno que da esta respuesta obtiene la media, mediana y moda, no obstante, se deduce una confusión terminológica y conceptual. En la Figura 52, se puede ver que busca la mediana de forma correcta, pero usa los símbolos de la media para identificarla y viceversa para la media. En la moda, marca los dos valores centrales de la distribución pero no realiza ninguna estimación. Al final, da los tres resultados de las medidas de tendencia central pero, no específica cual es el resultado del problema.

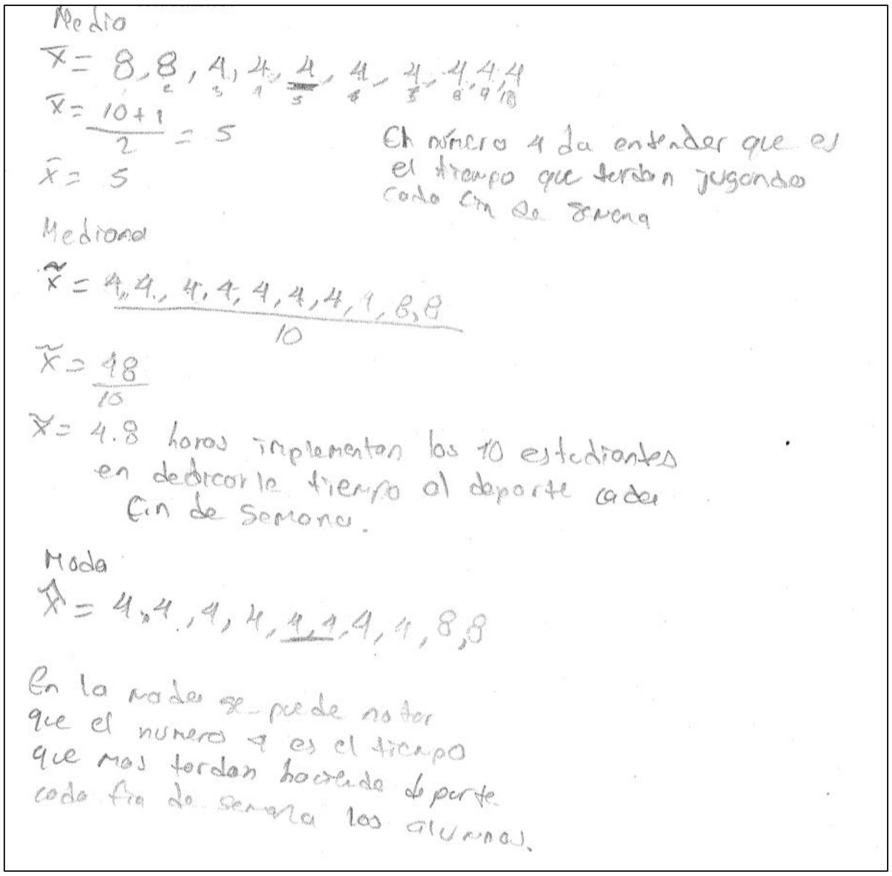


Figura 52. Ejemplo de respuesta C3.1.

C3.2. Obtiene la posición de la mediana, en vez de la media. El alumno confunde la media con la mediana, pero primero intenta buscar la posición de la mediana con una media de 4 y no concluye la operación, después calcula la posición de la mediana sumándole uno a una media de 8 y dividiéndola por dos. El resultado que obtiene de esta última operación lo considera como el número medio. Se observa (Figura 53), una confusión de definiciones entre media y mediana, y además no conoce totalmente la definición de la mediana.

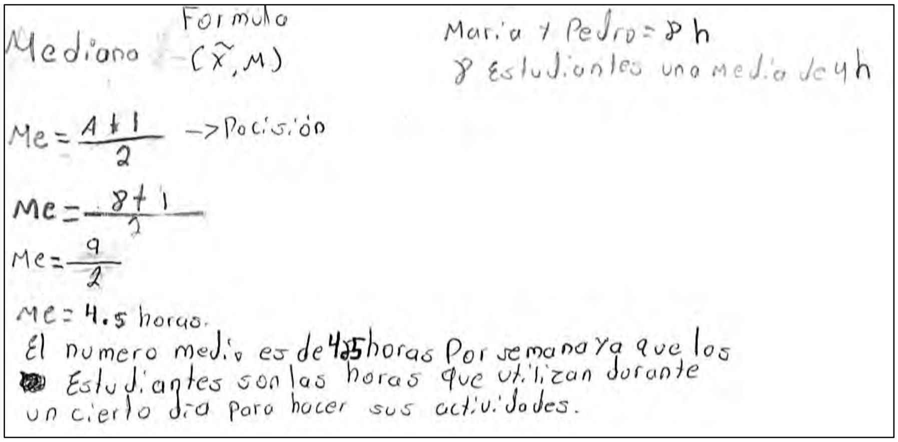


Figura 53. Ejemplo de respuesta C3.2.

C3.3. Obtiene la mediana convirtiendo las unidades de tiempo a minutos. En esta variante el estudiante primero divide la media de María y Pedro en dos y convierte las horas a minutos, la segunda media de los 8 estudiantes las divide por ocho y de la misma forma convierte las horas a minutos, después obtiene la posición de la mediana. Seguidamente vuelve a obtener la mediana de las cuatro horas que corresponden a María, Pedro y los 8 estudiantes. Calcula nuevamente la posición de la mediana y concluye que el número medio de horas es cuatro. Se observa (Figura 54) que el alumno conoce el procedimiento para obtener la mediana con datos impares, pero lo confunde con la media.

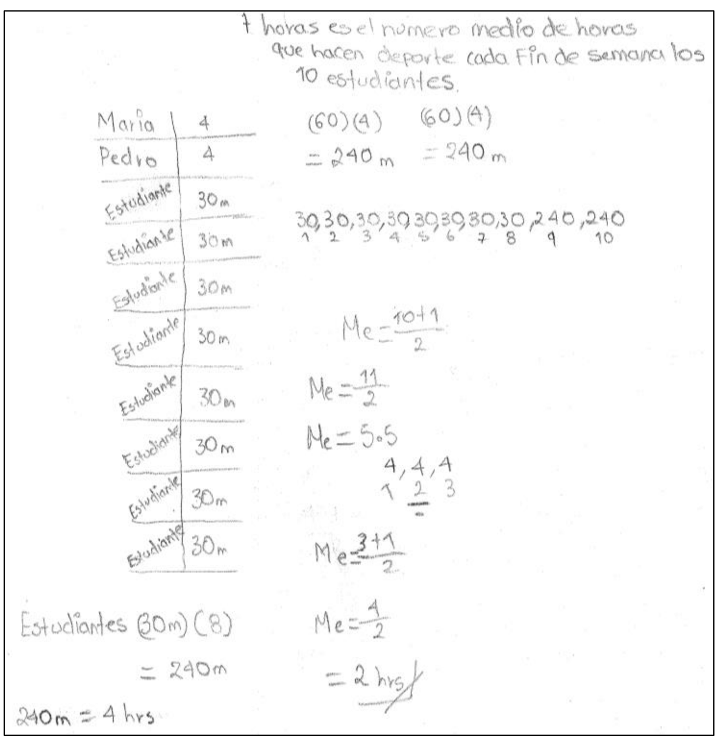


Figura 54. Ejemplo de respuesta C3.3.

C4. No tienen relación con las medidas de tendencia central. Este último grupo, lo integra una respuesta sin sentido que da un alumno, por lo que se ha considerado como una respuesta que no tiene relación con las medidas de tendencia central, pues no hay algoritmos que justifiquen las respuestas que da.

En seguida en la Tabla 9, se agrupan las categorías obtenidas en el análisis, cada una con su respectiva frecuencia y porcentaje, para la primer parte del problema 2.

Tabla 9. Categorías de respuesta del (inciso a), antes y después de la experiencia

Categorías de respuestas del problema 2.1	Antes	Después
C1. Calcula correctamente una media ponderada	1	
C2. Calcula la media ponderada y no justifica	1	
C3. Respuesta correcta sin justificarla mediante algoritmos.	3	
C4. No tiene relación con la media ponderada	12	1
C1.1. Ponderación correcta de los valores, pero realiza el cálculo de la media con datos aislados		1
C1.2. Correcta, mediante el cálculo de la media de una variable discreta con datos aislados		6
C1.3. Correcta, con distribución no equitativa		1
C2.1. Obtiene la media de las dos medias dadas		3
C2.2. Error en la propiedad de la media de no ser asociativa		2
C3.1. Confunde la media con la mediana		1
C3.2. Obtiene la posición de la mediana, en vez de la media		1
C3.3. Obtiene la mediana convirtiendo las unidades de tiempo a minutos		1
Total	**17**	**17**

La Tabla 9 donde se presentan las categorías obtenidas antes y después de la experiencia de aprendizaje, se observar primero que antes de la experiencia de aprendizaje la mayoría de los estudiantes (12) aunque han contestado el problema dan una respuesta que no tiene relación con el concepto de la media ponderada. También se observa que 3 estudiantes respondieron correctamente al problema, no obstante, los estudiantes solo escriben el valor pero no realizan justificación alguna mediante el uso de algoritmos. Por lo que, no se puede indagar sobre la capacidad que tienen estos alumnos a la hora de realizar el cálculo. Por otra parte, solo un 1 de estudiante calcula la media ponderada, pero no da un argumento sobre el resultado. Finalmente, un alumno calcula correctamente la media ponderada siguiendo los pasos propuestos por solución correcta.

Después de la experiencia de aprendizaje, se observa en la Tabla 9 que 8 estudiantes han sido capaces de dar una respuesta correcta, aunque un alumno realizó la ponderación, los 8 estudiantes han optado por usar el algoritmo de la media simple, y ningún estudiante fue capaz de aplicar el algoritmo de la media ponderada. Por otro lado, 9 dieron una respuesta errónea, con gran variedad de errores, 5 usan las medias para obtener el promedio o para sumarlas y dividirlas por el total de los datos, lo que supone un error al concebir la propiedad de la media de no ser asociativa. Es posible darse cuenta que, en este problema persiste la confusión terminológica entre media y mediana.

En comparación con el antes y después, se observa que de 12 personas que no comprendían el campo del problema, se redujo a solo 1 después de la implementación de la experiencia de aprendizaje, al mismo tiempo se ve que las respuestas correctas aumentaron, pasaron de 2 a 8, un avance importante, aunque no suficiente que se ha logrado después de la experiencia de aprendizaje, además que se observan más variedad de errores.

Categorías de respuestas en el problema 2.2, antes de la experiencia de aprendizaje.

C1. Respuesta correcta sin justificarla mediante algoritmos. En la primer categoría que se establece para este problema se encuentran los estudiantes que dan el resultado del problema, pero no justifican mediante algoritmos su respuesta, en algunos casos atienden a la propiedad *"la media de la suma de dos o más variables es igual a la suma de las medias de éstas"*, pues se observa (Figura 55) que suma las dos medias, pero no muestra el procedimiento que siguió. Se identifica pues, un conocimiento básico del concepto de la media ponderada. Es importante recalcarles a los estudiantes a utilizar correctamente los símbolos y signos.

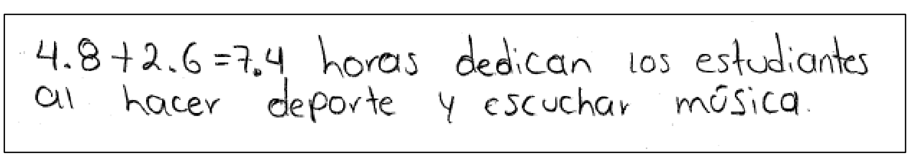


Figura 55 . Ejemplo de respuesta C1.

C2. No tiene relación con la media ponderada, ni propiedades. En esta categoría los estudiantes emiten respuestas sin sentido, es decir sin relación con los conceptos que se están manejando, demuestran que no conocen el lenguaje estadístico. Es primordial iniciar con lo básico de este tipo de conceptos.

C3. No contesta. Se compone de los estudiantes que ha dejado sin responder este problema, por lo tanto, no se puede obtener algún tipo de información.

Categorías de respuestas en el problema 2.2, después de la experiencia de aprendizaje.

C1. La media de la suma de dos o más variables, es la suma de las medias de éstas. Esta categoría contiene las respuestas en la que los estudiantes han logrado identificar una de las propiedades que integran este problema. No obstante, el alumno no pondera invirtiendo los algoritmos, solo hace una distribución de los valores, en la Figura 56, se observa que una vez teniendo la distribución procede a calcular la media simple, como ya tiene una media solo suma a la media que ha obtenido en esa estimación. El alumno tiene conocimiento del concepto, propiedades y algoritmo de la media simple, y opta por esta en vez de usar la media ponderada. También, en su respuesta se da el uso correcto de símbolos, signos y lenguaje verbal y numérico.

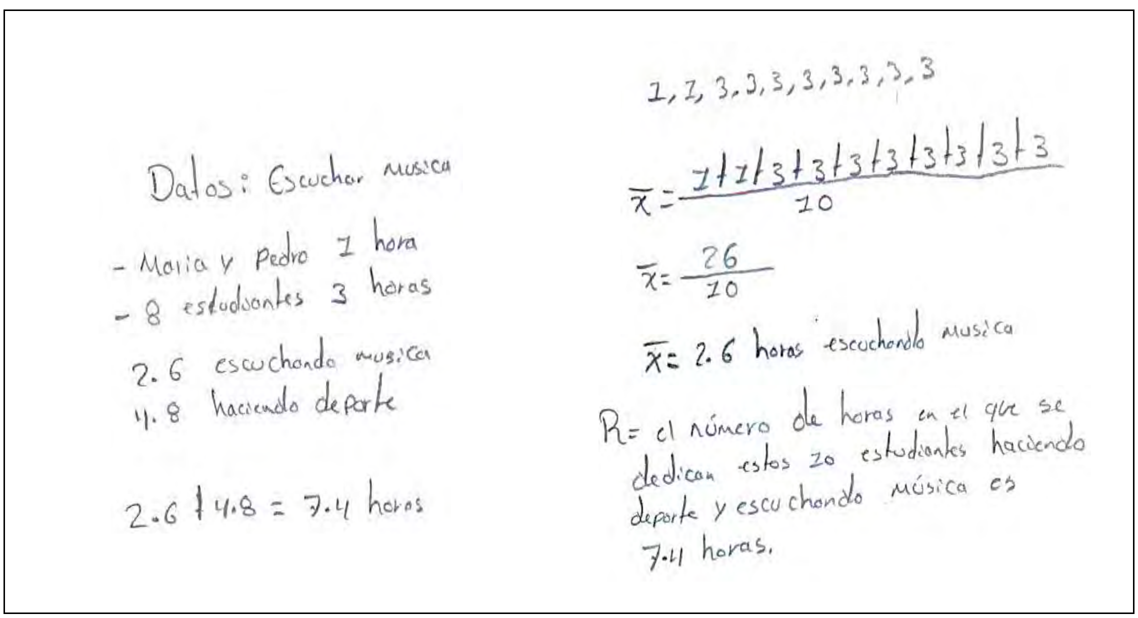


Figura 56. Ejemplo de respuesta C1.

C2.Respuesta correcta, mediante el cálculo de la media de una variable discreta con datos aislados. Los alumnos eligen usar el método más sencillo que es el cálculo de la media simple, para el que obtienen la media para cada conjunto de datos, a forma de comprobar, y seguidamente integran los datos de los dos conjuntos en una sola estimación. De manera implícita le suman un número de horas a los valores anteriores para obtener el tiempo que le dedica un estudiante de forma individual y posteriormente con estos nuevos datos estiman la nueva media de las dos actividades. Se observa (Figura 57) el uso correcto del algoritmo, el procedimiento que elige es claro y bien estructurado, marcando cada uno de los pasos para llegar a una solución, además hace uso del lenguaje simbólico y verbal.

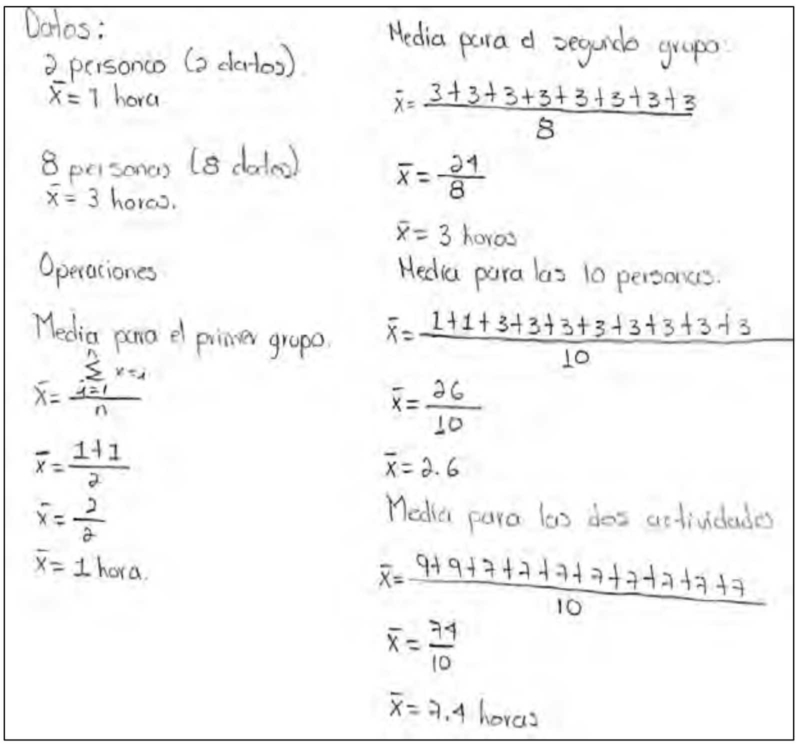
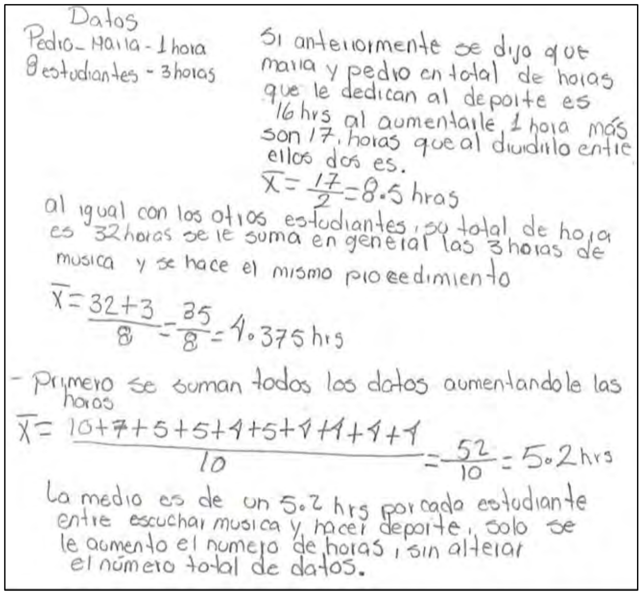


Figura 57. Ejemplo de respuesta C2.

C3. Obtiene media incorrecta con error en la suma total. El alumno da una respuesta en la que primero obtiene la suma total de los dos estudiantes y le adiciona la nueva media, pero lo hace de forma incorrecta pues no hace la distribución entre los datos de la variable, esto lo realiza para las dos medias. Posteriormente, realiza la suma y obtiene la media simple. Se observa en la Figura 58, que el alumno conoce el algoritmo de cálculo de la media y usa los procedimientos de este. Sin embargo, realiza los cálculos incorrectos para obtener la media ponderada.



Datos
Pedro_ Maria - 1 hora
8 estudiantes - 3 horas

Si anteriormente se dijo que maria y pedro en total de horas que le dedican al deporte es 16 hrs al aumentarle 1 hora más son 17. horas que al dividirlo entre ellos dos es.

$\bar{X} = \frac{17}{2} = 8.5$ hras

al igual con los otros estudiantes, su total de hora es 32 horas se le suma en general las 3 horas de musica y se hace el mismo procedimiento

$\bar{X} = \frac{32+3}{8} = \frac{35}{8} = 4.375$ hrs

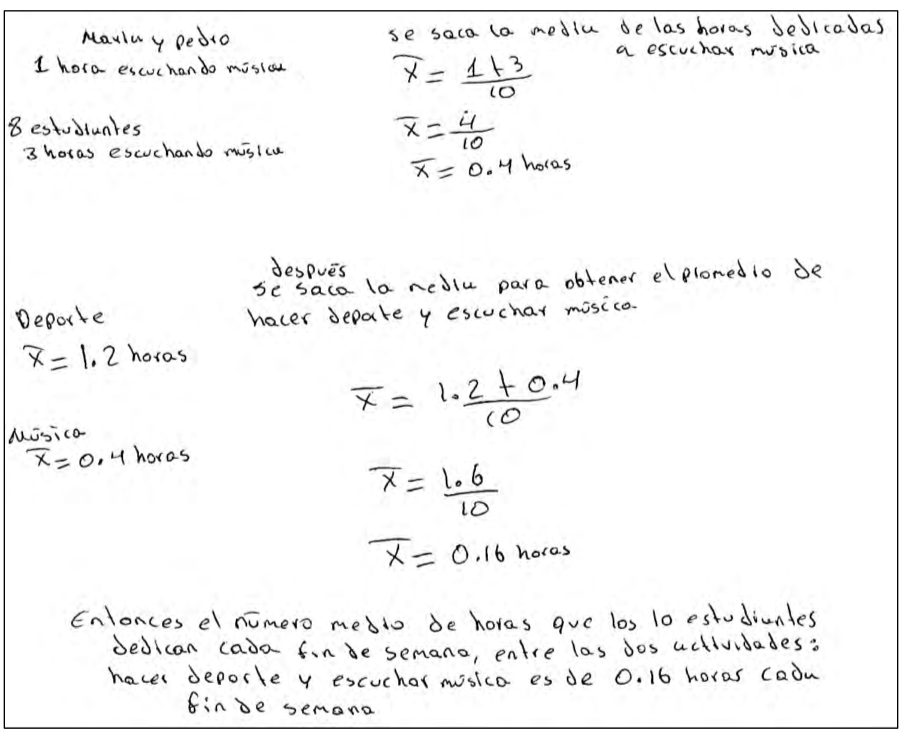
- Primero se suman todos los datos aumentandole las horas

$\bar{X} = \frac{10+7+5+5+4+5+4+4+4+4}{10} = \frac{52}{10} = 5.2$ hrs

La medio es de un 5.2 hrs por cada estudiante entre escuchar musica y hacer deporte, solo se le aumento el numero de horas, sin alterar el número total de datos.

Figura 58. Ejemplo de respuesta C3.

C4. Error en la propiedad de la media de no ser asociativa. Se observa (Figura 59) que los estudiantes cometen el error para las dos medias la de hacer deporte y la de escuchar música. Después, suman los resultados y vuelven a dividir por el total. Además de cometer el error en la propiedad también se puede ver el error en el algoritmo de cálculo de la media con datos aislados.



Maria y pedro
1 hora escuchando música

8 estudiantes
3 horas escuchando música

se saca la media de las horas dedicadas a escuchar musica

$\bar{X} = \frac{1+3}{10}$

$\bar{X} = \frac{4}{10}$

$\bar{X} = 0.4$ horas

Deporte
$\bar{X} = 1.2$ horas

Música
$\bar{X} = 0.4$ horas

después se saca la media para obtener el promedio de hacer depote y escuchar músico.

$\bar{X} = \frac{1.2 + 0.4}{10}$

$\bar{X} = \frac{1.6}{10}$

$\bar{X} = 0.16$ horas

Entonces el número medio de horas que los 10 estudiantes dedican cada fin de semana, entre las dos actividades: hacer deporte y escuchar música es de 0.16 horas cada fin de semana

Figura 59. Ejemplo de respuesta C4.

C5. Error en la propiedad de no ser asociativa y suma con la otra media. Se observa (Figura 60) que el alumno obtiene la media de las dos medias dadas pero comete un error en el algoritmo de cálculo, pues divide por el total de alumnos. Además, el resultado lo suma a la media que había encontrado en el inciso anterior, por lo que concibe de manera errónea la propiedad de la suma de las medias.

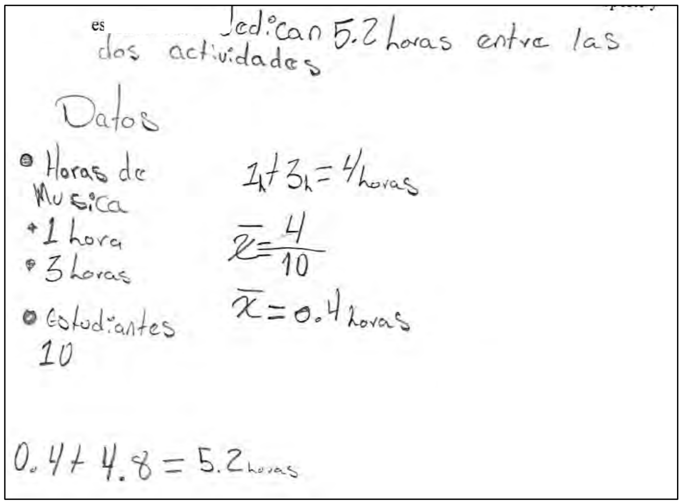


Figura 60. Ejemplo de respuesta C5.

C6. Confunde la media de una actividad y dice que es de las dos actividades. En esta categoría los estudiantes fueron capaces de estimar la media de un conjunto de datos aislados para la actividad de escuchar música, pero al dar respuesta al problema menciona que es el resultado de las dos actividades como lo muestra la Figura 61.

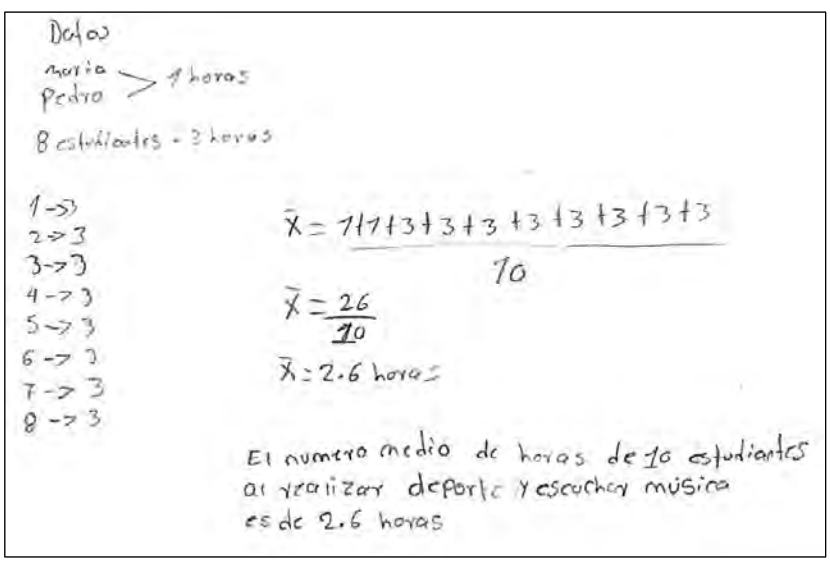
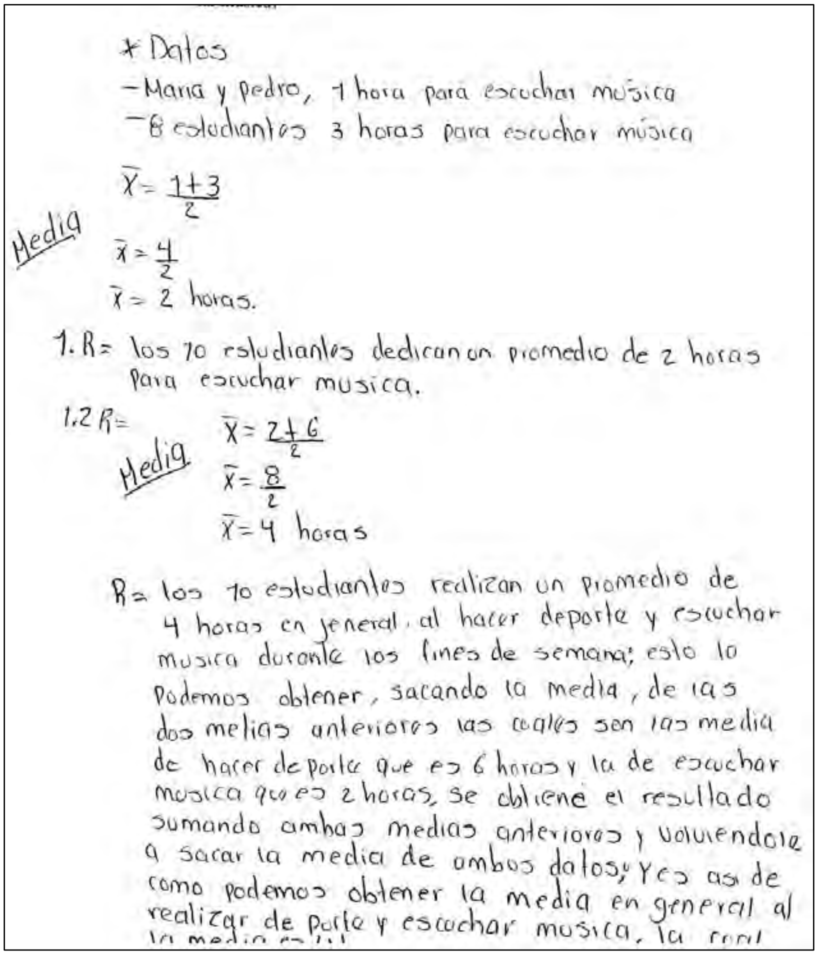


Figura 61. Ejemplo de respuesta C6.

C7. Obtiene la media de las medias dadas. Los alumnos de esta categoría cometen el error de calcular la media de las medias que se le han proporcionado, como se puede ver en la Figura 62. Se piensa que el alumno conoce el algoritmo de la media simple, pero desconoce las propiedades de la media que integran al problema.


* Datos
-Maria y pedro, 1 hora para escuchar música
-8 estudiantes 3 horas para escuchar música

Media
$\bar{x} = \frac{1+3}{2}$
$\bar{x} = \frac{4}{2}$
$\bar{x} = 2$ horas.

1. R= los 10 estudiantes dedican un promedio de 2 horas para escuchar musica.

1.2 R=
Media
$\bar{x} = \frac{2+6}{2}$
$\bar{x} = \frac{8}{2}$
$\bar{x} = 4$ horas

R= los 10 estudiantes realizan un promedio de 4 horas en jeneral. al hacer deporte y escuchar musica durante los fines de semana; esto lo podemos obtener, sacando la media, de las dos melias anteriores las cuales son las media de hacer deporte que es 6 horas y la de escuchar musica que es 2 horas, se obtiene el resultado sumando ambas medias anteriores y volviendole a sacar la media de ambos datos; y es asi de como podemos obtener la media en general al realizar de porte y escuchar musica, la cual

Figura 62. Ejemplo de respuesta C7.

C8. Confunde la media con la mediana. Este alumno muestra un error recurrente en los diferentes problemas. En este caso, se puede observar (Figura 63) que confunde los conceptos de media y mediana, por tanto, comete el error en los algoritmos.

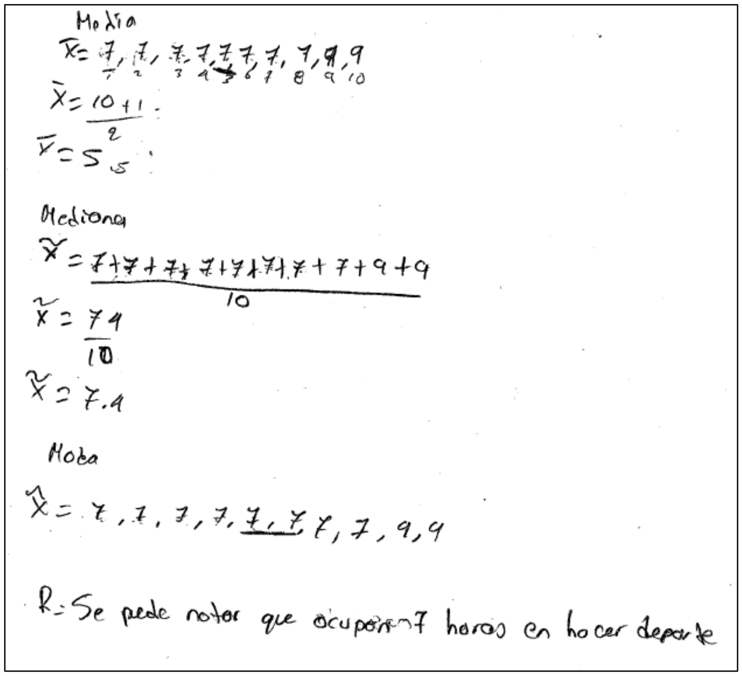


Figura 63. Ejemplo de respuesta C8.

C9. No tiene relación. Esta categoría integrada por las respuestas que no tienen relación con las medidas de tendencia central o no proporcionan información para poder inferir sobre su capacidad de cálculo.

Seguidamente, se agrupan las categorías en una tabla de frecuencias para su análisis.

Tabla 10. Categorías de respuesta del (inciso b), antes y después de la experiencia

Categorías de respuestas del problema 2.2	**Antes**	**Después**
C1. Respuesta correcta sin justificarla mediante algoritmos	2	
C2. No tiene relación con la media ponderada, ni propiedades	14	2
C3. No contesta	1	
C1. La media de la suma de dos o más variables, es la suma de las medias de éstas		2
C2.Respuesta correcta, mediante el cálculo de la media de una variable discreta con datos aislados.		3
C3. Obtiene media incorrecta con error en la suma total.		1
C4. Error en la propiedad de la media de no ser asociativa		2
C5. Error en la propiedad de no ser asociativa y suma con la otra media		1
C6. Confunde la media de una actividad y dice que es de las dos actividades		2
C7. Obtiene la media de las medias dadas		2
C8. Confunde la media con la mediana		2
Total	17	17

Las categorías encontradas y agrupadas en la Tabla 10, muestran que antes de la experiencia de aprendizaje 14 estudiantes contestaron al problema con respuestas que no tiene relación con el concepto de media ponderada, y tampoco han identificado la propiedad que se establece en el problema, probablemente no logran comprender el campo del problema. Por otro lado, 2 estudiantes dan la respuesta correcta, sin embargo, no la justifican por medio de los algoritmos, por lo que no se puede saber más sobre su capacidad de cálculo. Finalmente, un alumno no contestó este problema, por lo que no se puede obtener información acerca de los conocimientos con los que cuenta, entorno a estos conceptos.

Después de la experiencia de aprendizaje se puede ver en la Tabla 10, una gran variedad de respuestas. Esta parte del problema resultó ser muy difícil para los estudiantes de nuestra muestra, solo 5 estudiantes fueron capaces de dar una solución correcta al problema. Sin embargo, esta solución la dieron usando el algoritmo de la media simple, que se les facilita más. Respecto a las demás respuestas, hay una variedad amplia de errores, cada uno integrada por dos o un estudiante, en las que pasan desapercibido el concepto de ponderación, aun cuando fue parte de la instrucción. También en algunos casos, se sigue observando en menor frecuencia la confusión terminológica y conceptual de la media y la mediana.

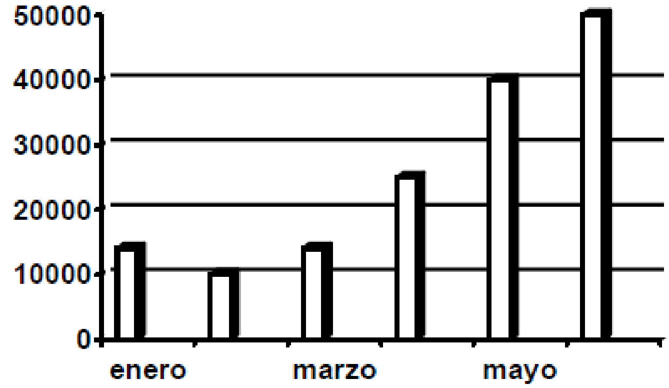
5.4. Análisis del problema 3

Este problema de Zawejowski (1986), ha sido utilizado en los trabajos de Mayén (2009) y Cobo (2003). Tiene como objetivo observar si el alumno es capaz de calcular la media y la mediana a partir de un gráfico. Se divide en dos secciones e incluye las propiedades: "*la mediana y media pueden no coincidir con los datos"* y *"el cálculo de la media y el de la mediana no son operaciones internas"*.

Algunas investigaciones previas como la de Curcio (1989), ha encontrado que la lectura de gráficos resulta difícil a los alumnos, y en este problema se presenta, pues requiere un nivel de lectura de datos, para convertir los gráficos a valores numéricos y así poder hacer los cálculos de la media y la mediana de un conjunto de datos aislados.

Problema 3

Observa el siguiente diagrama de barras que muestra las ventas de bocadillos de la empresa Bocatta durante los últimos 6 meses del año pasado:



a) *Da un valor aproximado del número medio de bocadillos que se vende al mes.*
b) *Da un valor aproximado de la mediana del número de bocadillos que se vendieron por mes.*

Para resolver la primera parte, se debe leer el gráfico y estimar los datos de cada barra para obtener el siguiente conjunto de datos aproximados:

13.000, 10.000, 13.000, 25.000. 40.000, 50.000.

Posteriormente se calcula la media del conjunto de datos:

$$\bar{x} = \frac{13.000 + 10.000 + 13.000 + 25.000 + 40.000 + 50.000}{6} = \frac{151.000}{6}$$

$$\bar{x} = 25.166{,}67\ bocadillos$$

La segunda parte, se pide obtener la mediana, por lo tanto, debe ordenar los datos:

10.000, 13.000, 13.000, 25.000, 40.000, 50.000

Como se trata de un conjunto de datos con número par, se aplica la condición de "indeterminación" y, por tanto, la mediana será el valor de la media de los datos centrales de la distribución, es decir:

$$\frac{(13.000\ +\ 25.000)}{2}\ =\ 19.000\ bocadillos$$

En las Tabla 11 y 12, se describen los elementos de significado del problema.

Tabla 11. Elementos de significado de la solución correcta del problema (inciso a).

Elementos de significado	**Contenido**
Definiciones	▪ Definición de media como promedio aritmético de un conjunto de datos ▪ Definición de la media como algoritmo de cálculo
Propiedades	▪ La media es representante de un conjunto de datos, que no necesariamente coincide con uno de los datos
Campos de problemas	▪ Cálculo de la media a partir de un gráfico
Algoritmos y procedimientos	*Algoritmos* ▪ Cálculo de la media a partir de la lectura de gráficos ▪ Cálculo de la media de una variable discreta con datos aislados, que consiste en la suma de todos los valores, dividiendo por el número de datos *Procedimientos* ▪ Lectura del gráfico ▪ Reconocer la variable que se presenta en el problema ▪ Asignar valores de las ventas de bocadillos que se vendió en los seis meses (Enero-Junio) ▪ *Cálculo de la media a partir de un gráfico:* o *Suma los seis datos.* o *Dividir la suma entre el número total de datos*
Lenguaje	▪ Verbal ▪ Numérico ▪ Gráfico ▪ Simbólico
Argumentos	▪ Razonamientos verbales deductivos

Tabla 12. Elementos de significado de la solución correcta del problema (inciso b).

Elementos de significado	**Contenido**
Definiciones	▪ Concepto de mediana como estadístico de orden ▪ Definición de mediana como valor central de una serie de datos aislados
Propiedades	▪ La mediana no es siempre uno de los valores de los datos ▪ El cálculo de la mediana no es una operación interna
Campos de problemas	▪ Cálculo de la mediana a partir de un gráfico
Algoritmos y procedimientos	*Algoritmos* ▪ Algoritmo de cálculo de la mediana en caso de número par de valores ▪ Algoritmo de cálculo de la media *Procedimientos* ▪ Lectura del gráfico ▪ Reconocer la variable que se presenta en el problema ▪ Asignar valores de las ventas de bocadillos que se vendió en los seis meses (Enero-Junio) ▪ *Cálculo de la mediana a partir de un gráfico con un número par de datos:* o *Ordenar los datos* o *Observar que el número de datos es par (n=6)* o *Identificar la indeterminación* o *Obtener la mediana al calcular la media de los datos que ocupan las posiciones 3 y 4*
Lenguaje	▪ Verbal ▪ Numérico ▪ Gráfico ▪ Simbólico
Argumentos	▪ Razonamientos verbales deductivos.

Categorías de respuestas del problema 3.1, antes de la experiencia de aprendizaje.

C1.1. Cálculo correcto de la media a partir de un gráfico. El estudiante ha sido capaz de obtener los valores numéricos de la variable y de calcular la media a partir de ellos. Por tanto, hace una lectura literal del gráfico, que es el primer nivel de Curcio (1989). Por lo que se puede decir que conoce la definición de la media y su algoritmo. También, se observa el uso del lenguaje numérico y simbólico (Figura 64).

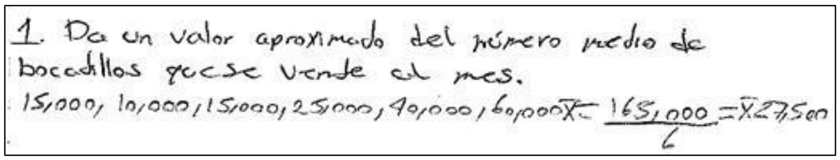

1. Da un valor aproximado del número medio de bocadillos que se vende al mes.
15,000, 10,000, 15,000, 25,000, 40,000, 60,000 $\bar{X} = \frac{165,000}{6} = \bar{X}\, 27500$

Figura 64. Ejemplo de respuesta C1.1.

C1.2. Respuesta correcta sin justificarla mediante algoritmos. Es complicado obtener algún tipo de información sobre su capacidad de cálculo. El valor que obtiene indica que usa la idea de la media, pero se observa (Figura 65) que no escribe el procedimiento para llegar a esta respuesta.

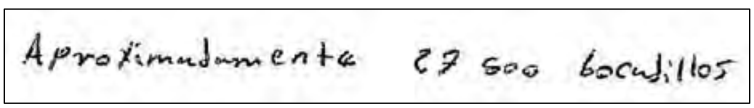

Aproximadamente 27 500 bocadillos

Figura 65. Ejemplo de respuesta C1.2.

C2. Confunde los valores de la variable. Los alumnos realizan la lectura del gráfico de forma incorrecta, ya que confunde los valores de las variables con las marcas de la escala, también fue encontrado por Mayén (2009). En esta categoría se encontraron las siguientes variantes:

C2.1. Calcula la media con las marcas de la escala. Los estudiantes calculan la media, se deduce que conocen la definición de la media como promedio aritmético del conjunto de datos y como algoritmo de cálculo. Sin embargo, cometen el error de confundir los valores de las variables con las marcas de la escala y con estos datos calculan la media como se puede ver en la Figura 66.

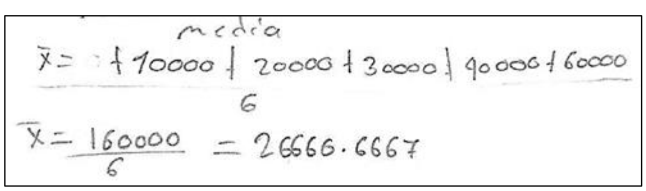

media
$\bar{x} = \frac{10000 + 20000 + 30000 + 40000 + 60000}{6}$
$\bar{x} = \frac{160000}{6} = 26666.6667$

Figura 66. Ejemplo de respuesta C2.1.

C2.2. Cálculo correcto de la media con error sin justificar con algoritmos. La Figura 67 muestra que el alumno obtiene la media de un conjunto de datos sin justificar cómo llegó a este resultado.

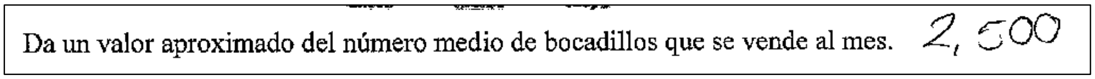

Da un valor aproximado del número medio de bocadillos que se vende al mes. 2, 500

Figura 67. Ejemplo de respuesta 2.2.

C2.3. Elegir como media el valor intermedio del eje de ordenadas. Se identifica primero un error al interpretar el gráfico; el estudiante da como respuesta "30,000" como se observa en la Figura 68, por lo que se piensa que sólo observa los valores de la escala y elige el valor intermedio como media.

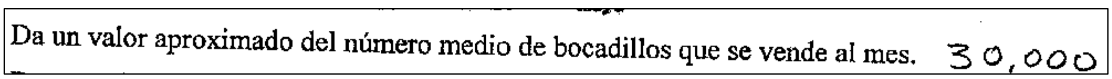

Da un valor aproximado del número medio de bocadillos que se vende al mes. 30,000

Figura 68. Ejemplo de respuesta C2.3.

C3. No tiene relación con el problema. Las respuestas de esta categoría no tienen relación con el problema. En este caso, no se puede explicar por qué elige 10,000 pues no hay suficiente información para poder deducir el resultado que muestra la Figura 69.

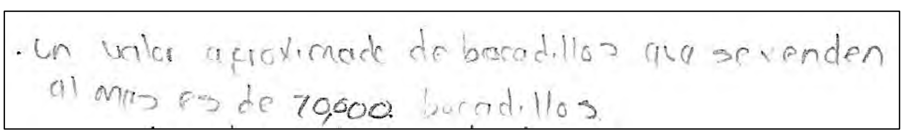

· Un valor aproximado de bocadillos que se venden al mes es de 70,000 bocadillos.

Figura 69. Ejemplo de respuesta C3.

Categorías de respuestas del problema 3.1, después de la experiencia de aprendizaje.

C1. Calcula la media a partir de un gráfico. En esta categoría se analizan las respuestas que utilizan la medida de tendencia central pedida, es decir, el grupo de alumnos de esta categoría calculan la media, pues los estudiantes son capaces de hacer una lectura a nivel literal del gráfico, sin embargo este es el primer nivel de lectura en la clasificación de Curcio (1989). Dentro de esta categoría se han encontrado las siguientes variantes.

C1.1. Cálculo correcto de la media a partir de un gráfico. Los estudiantes que responden correctamente a este apartado han obtenido los valores numéricos de la variable y han sido capaces de realizar el cálculo de la media a partir de ellos. Por tanto, se dice que los estudiantes han hecho una lectura literal del gráfico, que es el primer nivel de Curcio

(1989). Pasa de cada barra, a un valor numérico aproximado, y lo registra en una tabla de frecuencias, se deduce que los alumnos relacionan diferentes tipos de representación del valor de la variable. Seguidamente, calculan correctamente la media de un conjunto de datos aislados, por lo que se constata que conocen la definición de la media y su algoritmo. También, se observa (Figura 70) el uso del lenguaje numérico y simbólico, además de argumentos verbales.

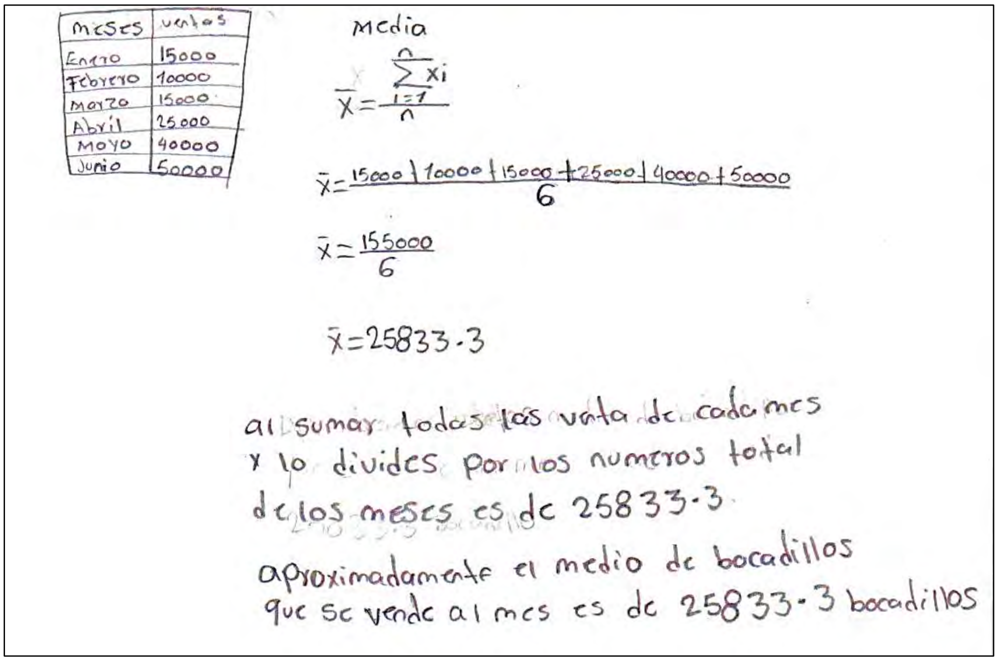


Figura 70. Ejemplo de respuesta C1.1

En este apartado del problema 3, los estudiantes lograron identificar los elementos de significado al resolver el problema, que se describieron en la Tabla 11.

C1.2. Obtiene la media con errores de cálculos aritméticos. En esta variante los alumnos, cometen errores en los cálculos aritméticos cuando estiman la media, es decir, se observa (Figura 71) error en las operaciones aritméticas (suma de todos los datos).

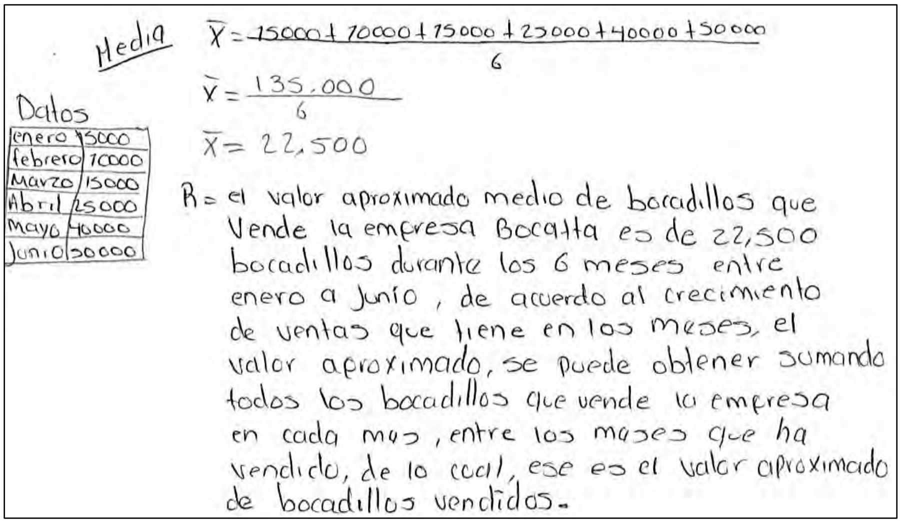


Figura 71. Ejemplo de respuesta C1.3.

C1.3. Cálculo correcto de la media con error al interpretar la gráfica. En el estudio de Mayén (2009) también reporta este error, donde el alumno asigna erróneamente los valores de la variable evidenciando falta de capacidad de lectura de gráficos, aun en el nivel más simple (Curcio, 1989). El alumno falla en la comprensión de la idea de distribución, es decir, divide el conjunto de datos en partes. Asimismo se observa (Figura 72) el uso de símbolos y argumentos, pero tiene un error en el cálculo de las operaciones aritméticas (suma de los datos). No obstante, el cálculo de la media aritmética es correcto.

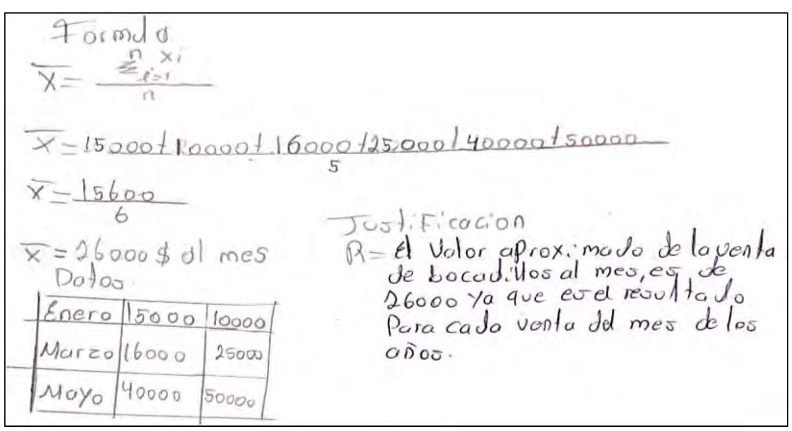
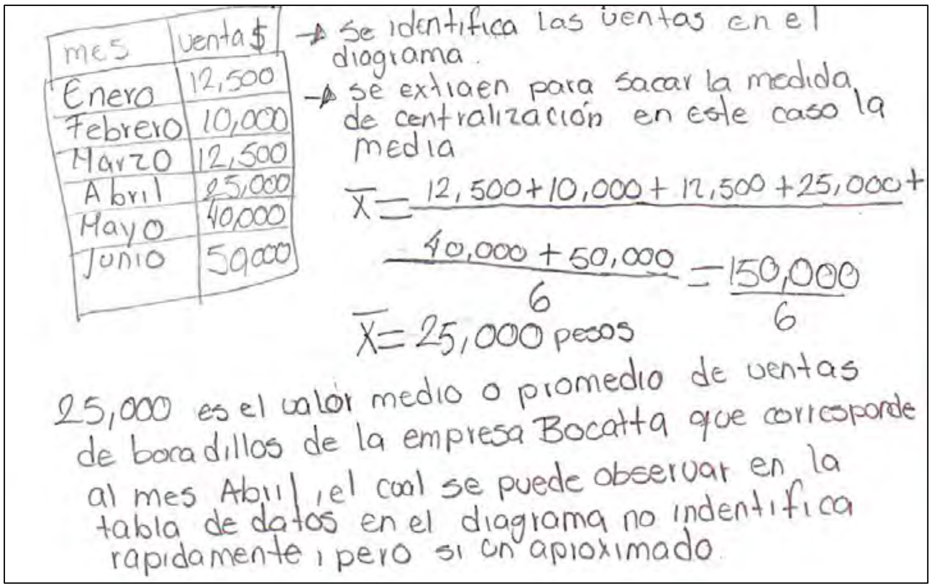


Figura 72. Ejemplo de respuesta C1.4.

C1.4. Confunde el valor central de las variables, con la frecuencia. En esta variante el alumno obtiene la media a partir del gráfico, aunque aplica la idea de media como valor central, se encontró el mismo error de los estudiantes, en el trabajo de Mayén (2009), es decir, asocia la media a la idea de centro, aunque en otro sentido con su argumento se observa (Figura 73) que se genera un conflicto al no saber a qué se refiere este centro, confundiendo "centro estadístico de la distribución", con "centro geométrico del rango de variación".



mes	venta $
Enero	12,500
Febrero	10,000
Marzo	12,500
Abril	25,000
Mayo	40,000
Junio	50,000

→ Se identifica las ventas en el diagrama.
→ se extraen para sacar la medida de centralización en este caso la media

$$\bar{X} = \frac{12,500 + 10,000 + 12,500 + 25,000 + 40,000 + 50,000}{6} = \frac{150,000}{6}$$

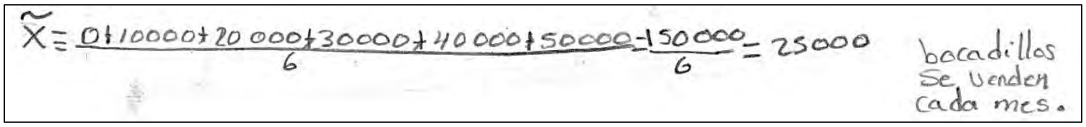
$$\bar{X} = 25,000 \text{ pesos}$$

25,000 es el valor medio o promedio de ventas de boradillos de la empresa Bocatta que corresponde al mes Abril, el cual se puede obseruar en la tabla de datos en el diagrama no indentifica rapidamente, pero si un aproximado

Figura 73. Ejemplo de respuesta C1.5.

C2. Confunde los valores de las variables con las marcas de la escala del eje de las ordenadas. Los alumnos realizan la lectura del gráfico de forma incorrecta, se asume que no es capaz de leer entre los datos según Curcio (1989), pues es el nivel más sencillo de lectura de gráficos. Este error de confundir los valores de las variables con las marcas de la escala del eje de las ordenadas, también fue encontrado por Mayén (2009). Esta categoría integrada por dos variantes que en seguida se describen.

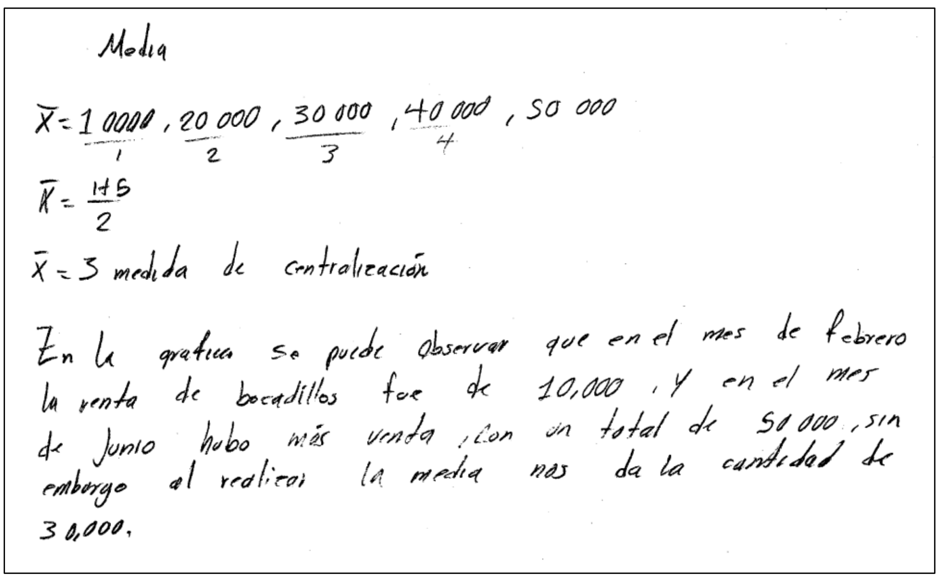
C2.1. Calcula la media con las marcas de la escala. Se observa en la Figura 74, que los estudiantes de esta categoría calculan la media, se deduce que conocen la definición de la media como promedio aritmético del conjunto de datos y como algoritmo de cálculo. Sin embargo, cometen el error de confundir los valores de las variables con las marcas de la escala del eje de las ordenadas, y con estos datos calcular la media. También confunden el valor central, con la frecuencia. No se han encontrado este conflicto en otros estudios, por lo que se considera una aportación de este trabajo.



$\tilde{X} = \frac{0+10000+20\,000+30000+40\,000+50000}{6} = \frac{150\,000}{6} = 25000$ bocadillos se venden cada mes.

Figura 74. Ejemplo de repuesta C2.1.

C2.2. Calcula la mediana en vez de la media con las marcas de la escala. Esta categoría solo está integrada por un estudiante, quien confunde primero los símbolos de la mediana ($Me, \tilde{x}$) con el de la media ($\bar{x}$), con esto también se deduce que existe un conflicto al confundir los conceptos de media con mediana, según se ve en la Figura 75. No obstante, ordena los datos de forma ascendente y encuentra la mediana de un conjunto de datos impar. No se ha encontrado este error en otras investigación, por tanto, también se considera una aportación de este trabajo.



Media

$\bar{X} = \underset{1}{10000}, \underset{2}{20\,000}, \underset{3}{30\,000}, \underset{4}{40\,000}, 50\,000$

$\bar{X} = \frac{1+5}{2}$

$\bar{X} = 3$ medida de centralización

En la grafica se puede observar que en el mes de febrero la venta de bocadillos fue de 10,000, y en el mes de Junio hubo más venta, con un total de 50 000, sin embargo al realizar la media nos da la cantidad de 30,000.

Figura 75. Ejemplo de respuesta C2.2.

C3. La respuesta no tiene relación con las medidas de tendencia central. Este apartado está compuesto por solo una respuesta (Figura 76) en la que el alumno comete el error de dar una solución no relacionada con una de las medidas de tendencia central. El alumno realiza operaciones aritméticas con un valor de 10,000.5 que asigna a algún dato del gráfico, después lo multiplica por 1 y lo divide por dos. Otra operación que realiza es dividir el valor máximo por el valor mínimo de la escala y lo multiplica por dos. Al final realiza un argumento sin sentido sobre el aumento y disminución de las ventas que por cada 10,000.5 bocadillos cada mes aumenta 10 bocadillos.

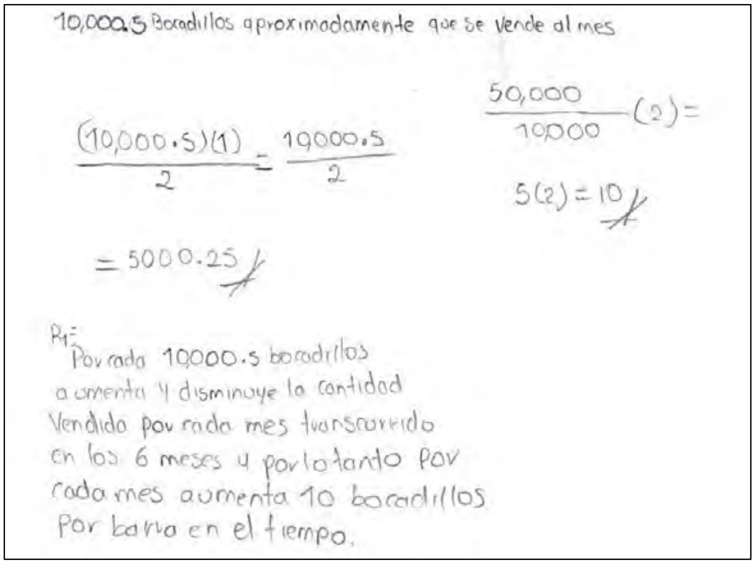


Figura 76. Ejemplo de respuesta C3.

Por otro lado, se presenta en la Tabla 13, el concentrado de las categorías obtenidas en el análisis, con sus frecuencias y porcentajes, donde se han encontrado mejores resultados que en la aplicación del cuestionario piloto, por tanto se deduce que ha sido un problema fácil de resolver, por lo que, los estudiantes son capaces de interpretar al menos este tipo de gráfico.

Tabla 13. Categorías de respuesta del (inciso a), antes y después de la experiencia

Categorías de respuestas	Antes	Después
C1.1. Cálculo correcto de la media a partir de un gráfico	1	9
C1.2. Respuesta correcta sin justificarla mediante algoritmos	5	0
C2.1. Calcula la media con las marcas de la escala	2	2
C2.2. Cálculo correcto de la media con error sin justificar con algoritmos.	1	0
C2.3 Elegir como media el valor intermedio del eje de ordenadas	6	0
C3. No tiene relación con el problema	2	1
C1.3. Obtiene la media con errores de cálculos aritméticos	0	2
C1.4. Cálculo correcto de la media con error al interpretar la gráfica	0	1
C1.5. Confunde el valor central de las variables, con la frecuencia	0	1
C2.2. Calcula la mediana en vez de la media con las marcas de la escala	0	1
Total	17	17

Se observa en la Tabla 13 que 6 estudiantes, antes de la experiencia de aprendizaje dan un valor numérico como respuesta correcta, pero no justifican con algoritmos, por lo que, resulta complicado realizar el análisis, debido a que no se cuenta con más información. A menudo, cometen el error al interpretar de forma incorrecta el gráfico, 8 alumnos confunden las marcas de escala con el valor de la variable, o al convertir el grafico en valores numéricos se genera un error de escritura. Además de estos errores, se han encontrado problemas en el lenguaje verbal, numérico, simbólico y algebraico, así como, errores constantes en operaciones aritméticas. Después de la experiencia de aprendizaje gran parte de ellas se concentran en la resolución correcta de la media, 13 estudiantes. Se observa que un 47% fue capaz de calcular la media a partir de un gráfico. Seguidamente 2 alumnos realizan los cálculos de la media siguiendo los algoritmos y procedimientos, pero al realizar las operaciones aritméticas comete un error. Por otro lado, se coincide con el estudio de Mayén (2009), pues los errores más visibles en este problema, que representan el 18% de los resultados, se da cuando los alumnos confunden los elementos que representa cada eje del gráfico, como es, la variable (número de bocadillos) con la escala, por lo que los estudiantes no alcanzan el primer nivel "leer los datos" definido por Curcio (1989). Por último se observa 1 respuesta que no tiene relación con el problema planteado.

Categorías de respuestas del problema 3.2, antes de la experiencia de aprendizaje

C1. Respuesta incorrecta, no usa la condición de indeterminación. Se observa (Figura 77) que el alumno, además de confundir las variables representadas en cada eje, no realiza ningún algoritmo de mediana, y por tanto, la condición de indeterminación.

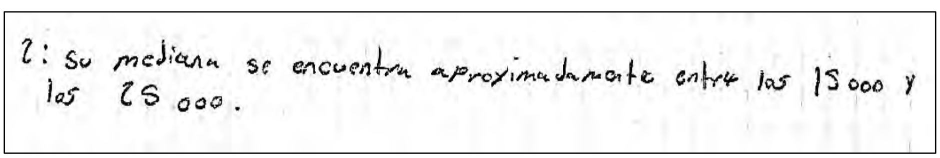
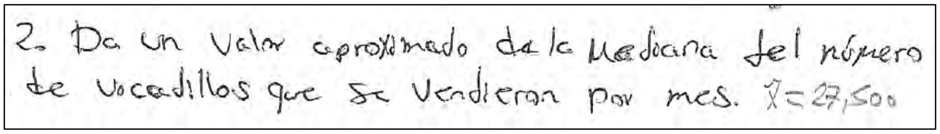

2: Su mediana se encuentra aproximadamente entre los 15 000 y los 25 000.

Figura 77. Ejemplo de respuesta C1.

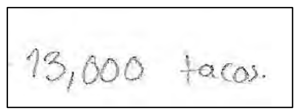
C2. Obtiene la media en vez de la mediana. El alumno confunde los conceptos de media y mediana; hace una lectura del gráfico, asigna valores a las variables y estima la media de un conjunto de datos. En la Figura 78 se puede ver que da como respuesta el valor numérico pero no hace cálculos.



2. Da un Valor aproximado de la Mediana del número de vocadillos que se Vendieron por mes. $\bar{x}$=27,500

Figura 78. Ejemplo de respuesta C2.

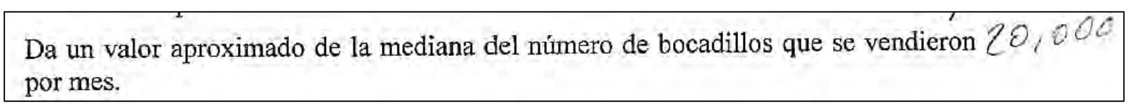
C3. Obtiene la moda en lugar de la mediana. Este error también se ha encontrado en la investigación de Mayén (2009), en el que los estudiantes confunden los conceptos de moda y mediana. Se deduce a partir de la Figura 79, que el estudiante realiza una interpretación correcta del gráfico; asigna valores a la variable y realiza el cálculo correcto de la moda. Sin embargo, lo que se le solicita en esta parte del problema es la mediana.



13,000 tacos.

Figura 79. Ejemplo de respuesta C3.

C4. Cálculo de la mediana sin ordenar los datos. Los alumnos consideran como mediana el valor central de todos los elementos tal y como se le presentan. Este error también fue encontrado en Cobo (2003) y Mayén (2009). No obstante, en el ejemplo (Figura 80), se observa que el alumno interpreta correctamente la gráfica, y resuelve la indeterminación de tener dos valores centrales.

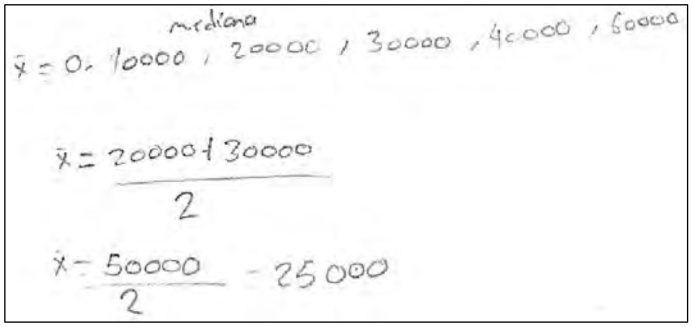


Da un valor aproximado de la mediana del número de bocadillos que se vendieron 20,000 por mes.

Figura 80. Ejemplo de respuesta C4.

C5. Error al interpretar la gráfica. Los alumnos no realizan la lectura del gráfico en el primer nivel que propone Curcio (1989). En algunas variantes de esta categoría los alumnos calculan la mediana del conjunto de datos, estos errores también han sido encontrados en Mayén (2009). En lo que sigue se describen las variantes encontradas.

C5.1. Obtiene la mediana, pero confunde el valor de la variable con la escala y toma en cuenta el 0 como valor. El error que se describe en esta categoría también fue encontrado en Mayén (2009) y Cobo (2003), pero en nuestro caso, los estudiantes además de confundir la variable estadística con la unidad (Figura 81), también toman en cuenta el 0 como valor, y a partir de allí, calculan la mediana siguiendo el procedimiento correcto, ordena los datos, identifica el caso de indeterminación y lo resuelve. Pero, el error cometido es por una lectura incorrecta del gráfico.


mediana

$\bar{x} = 0, 10000, 20000, 30000, 40000, 60000$

$\bar{x} = \frac{20000+30000}{2}$

$\bar{x} = \frac{50000}{2} = 25000$

Figura 81. Ejemplo de respuesta C5.1.

C5.2. Obtiene la mediana a partir del eje de la escala incluyendo el 0, sin justificar con algoritmos. Esta categoría es similar a la anterior, en la que los estudiantes obtienen la mediana a partir del eje de la escala incluyendo el 0 a los valores, sin embargo, no realiza algoritmos de cálculo como se observa (Figura 82) en el ejemplo.

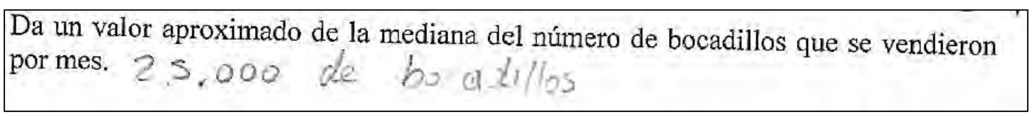

Da un valor aproximado de la mediana del número de bocadillos que se vendieron por mes. 25,000 de bocadillos

Figura 82. Ejemplo de respuesta C5.2.

C5.3. Confunde las marcas de la escala con los valores de la variable. Este error también fue encontrado por Mayén (2009) y Cobo (2003). Los estudiantes cometen el error al interpretar el gráfico, y utilizan como valores de la variable los números que aparecen en la escala del eje de las ordenadas y da como respuesta “30,000”, es decir, lo elige como el valor intermedio del eje vertical, por lo tanto, se podría asumir que confunde variable con unidad estadística. Se observa en la Figura 83 que los alumnos calculan la mediana, pero no justifican su cálculo.

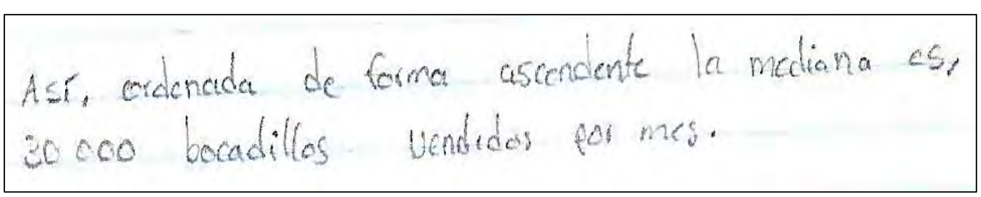

Así, ordenada de forma ascendente la mediana es, 30 000 bocadillos vendidos por mes.

Figura 83. Ejemplo de respuesta C5.3.

C6. La respuesta no tiene relación con la mediana. Se puede ver en la Figura 84 que el alumno da una solución que no está relacionada con el problema, da como valor de la mediana 8,000 que no justifica el cálculo.

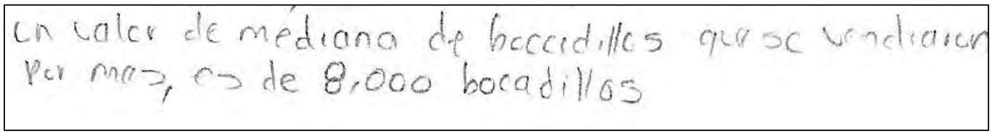


Figura 84. Ejemplo de respuesta C6.

C7. No contesta. No proporciona respuesta al problema planteado.

Categorías de respuestas del problema 3.2, después de la experiencia de aprendizaje

C1. Calcula la mediana a partir de un gráfico. Esta primera categoría que se describe, integra las respuestas que los estudiantes dan cuando se les solicita una medida de tendencia central a partir de un gráfico. Para esto, es necesario que los alumnos sean capaces de hacer una lectura a nivel literal del gráfico, según Curcio (1989). A continuación, se mencionan las variantes que se encontraron en esta categoría.

C1.1. Cálculo correcto de la mediana a partir de un gráfico. Las respuestas que integran esta categoría responden correctamente al problema planteado. Se observa en la Figura 85 que los estudiantes, presentan los elementos de significado que los alumnos pudieron haber deducido al momento de resolver este problema. Además, de que en esta categoría los alumnos son capaces de calcular la mediana de un conjunto de datos par a partir de un gráfico, también, conocen el concepto de la mediana como estadístico de orden y como valor central. De la misma forma, son capaces de leer el grafico, que implica identificar las variables, escala y asignar valores a las barras, pues una vez comprendido esto realiza los procedimientos para obtener la estimación que se le solicita, asimismo, se identifica el uso del lenguaje verbal y numérico que maneja al argumentar.

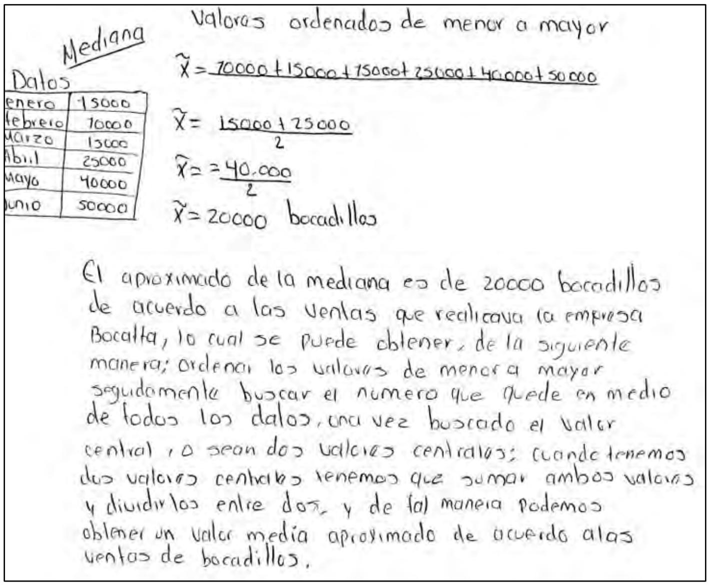


Figura 85. Ejemplo de respuesta C1.1

C1.2. Cálculo correcto de la mediana con error al convertir el gráfico en valores numéricos. Como se puede ver en la Figura 86, el alumno obtiene la mediana de un conjunto de datos par, pero, al realizar la lectura del gráfico y transcribir los valores no asigna un cero, probablemente es un error visual. Del mismo modo, se puede ver que al buscar la posición de la media utiliza el símbolo (*Me*) haciendo referencia a la mediana, cuando el cálculo que realiza es de la posición.

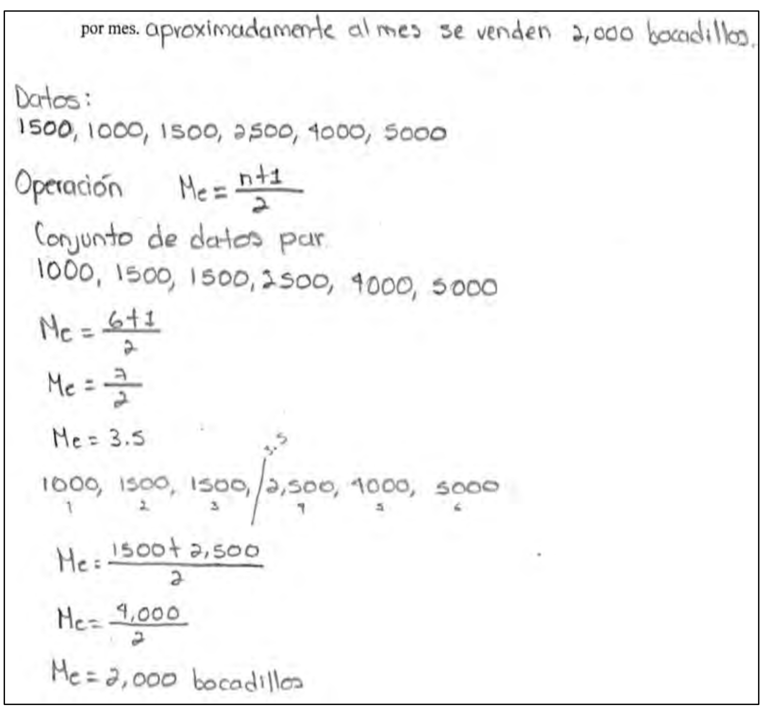
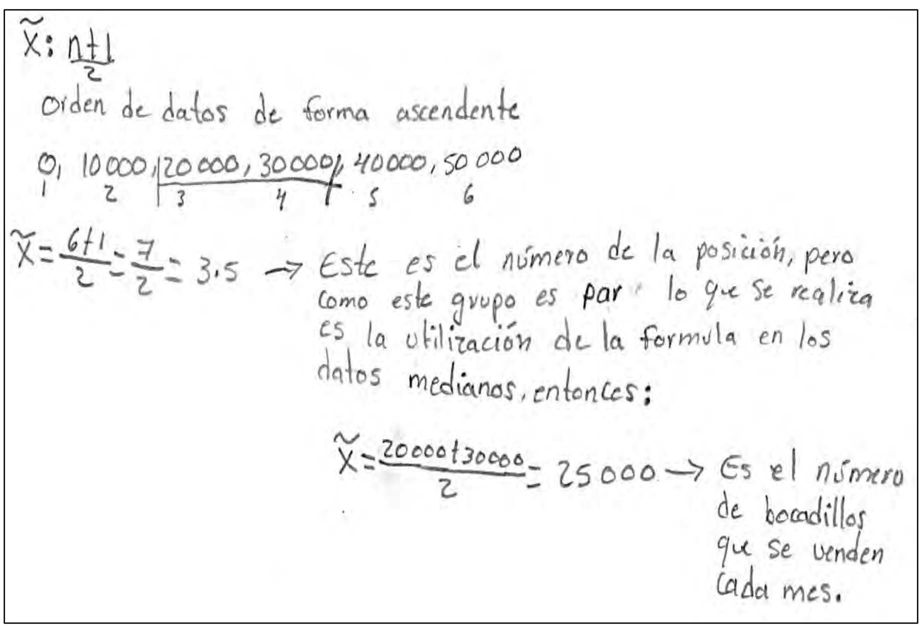


Figura 86. Ejemplo de respuesta C1.2.

C2. Error al interpretar la gráfica. Las respuestas que integran esta categoría, son aquellas en las que los alumnos no logran realiza la lectura del gráfico en el primer nivel que propone Curcio (1989). Sin embargo, algunas variantes de esta categoría muestran que los alumnos calculan la mediana del conjunto de datos, algunos de estos errores también han sido encontrados en Mayén (2009). En lo que sigue se describen las variantes encontradas.

C2.1. Obtiene la mediana, pero confunde el valor de la variable con las marcas de la escala en el eje de las ordenadas. En las investigaciones de Mayén (2009) y Cobo (2003), encuentran el mismo error al interpretar el grafico, en la que los estudiantes usan como valores de la variable la escala del eje de las Y, por que se cree que los estudiantes confunden la variable estadística con la unidad. La Figura 97 muestra que los estudiantes toman estos valores y calculan la mediana siguiendo el procedimiento correcto, ordenan los datos, identifican a la mediana como centro. No obstante, el error cometido es por una lectura incorrecta del gráfico.



$\tilde{X}: \frac{n+1}{2}$

Orden de datos de forma ascendente

$\underset{1}{0},\ \underset{2}{10000},\ \underset{3}{20000},\ \underset{4}{30000},\ \underset{5}{40000},\ \underset{6}{50000}$

$\tilde{X} = \frac{6+1}{2} = \frac{7}{2} = 3.5 \rightarrow$ Este es el número de la posición, pero como este grupo es par lo que se realiza es la utilización de la formula en los datos medianos, entonces:

$\tilde{X} = \frac{20000+30000}{2} = 25000 \rightarrow$ Es el número de bocadillos que se venden cada mes.

Figura 87. Ejemplo de respuesta C2.1.

C2.2. Cálculo incorrecto de la mediana con las marcas de la escala. Este error no lo reportan en las investigaciones de Mayén (2009) y Cobo (2003), por lo que se considera una aportación de este trabajo. En esta categoría se observa (Figura 88) que el alumno no realiza la lectura literal del gráfico, pues confunde la escala con los valores de la variable, también, se deduce que usa la idea de estadístico de orden, por lo que ordena los valores que tomó de la escala de forma ascendente, además, considera el orden y asigna números que indican la posición de los datos. Después, el alumno elige el dato que se encuentra en el centro, usando la idea de valor central, pero termina dividiéndolo por el número de la posición, es decir divide el valor central por la posición.

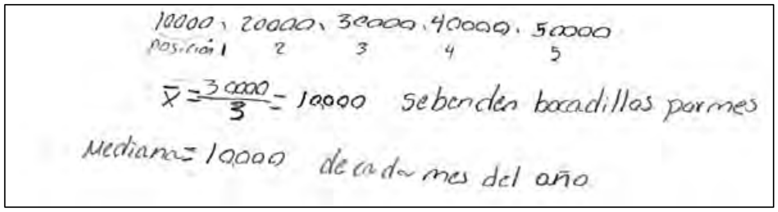


Figura 88. Ejemplo de respuesta C2.2

C2.3. Obtiene la media, y realiza el cálculo incorrecto de la mediana. En las investigaciones de Mayén (2009) y Cobo (2003), no reportan este error, por lo que se considera una aportación. En esta categoría el alumno obtiene la media de datos aislados con los valores de la escala del eje de las Y. Del mismo modo, se puede ver (Figura 89) que al querer calcular la mediana primero ordena los datos, se deduce que tiene nociones que es un estadístico de orden, después calcula la posición de la mediana, pero en lugar de usar el número de datos elige el segundo valor (20,000), el resultado de esta operación lo considera como la respuesta. También, busca el cociente del valor de la cuarta posición con el que ha considerado la mediana y obtiene un valor que no explica a qué se refiere.

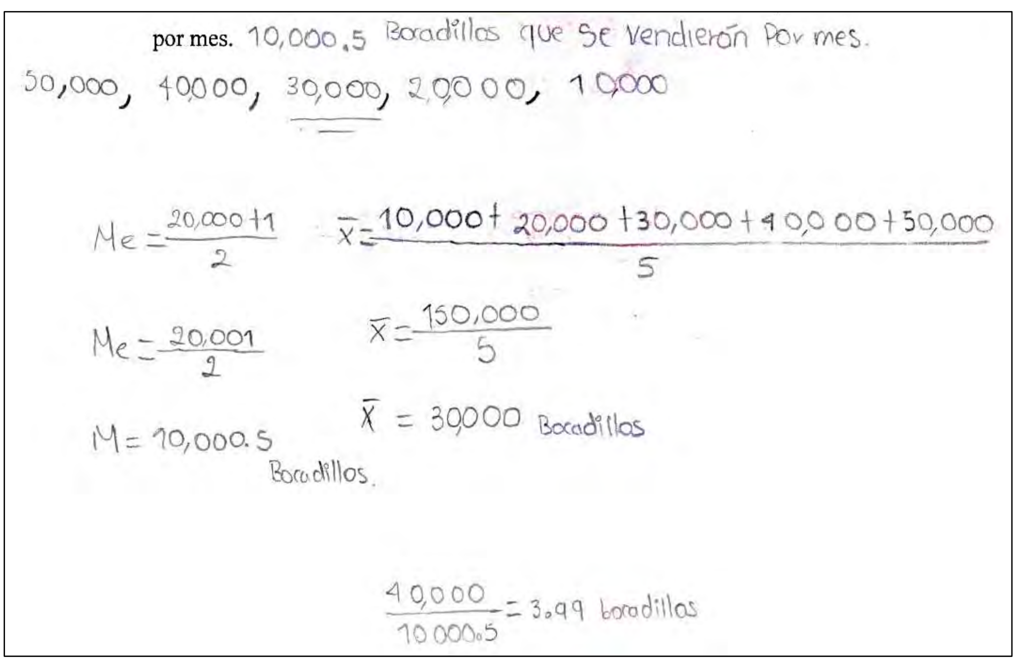


Figura 89. Ejemplo de respuesta C2.3.

C2.4. Confunde la media con la mediana. En esta variante, el alumno confunde los conceptos de media y mediana. Se observa en la Figura 90 que toma los datos de la escala y calcula la media, sin embargo, no agrega los símbolos de suma en los datos y la división la realiza por un número diferente de datos a lo que está sumando.

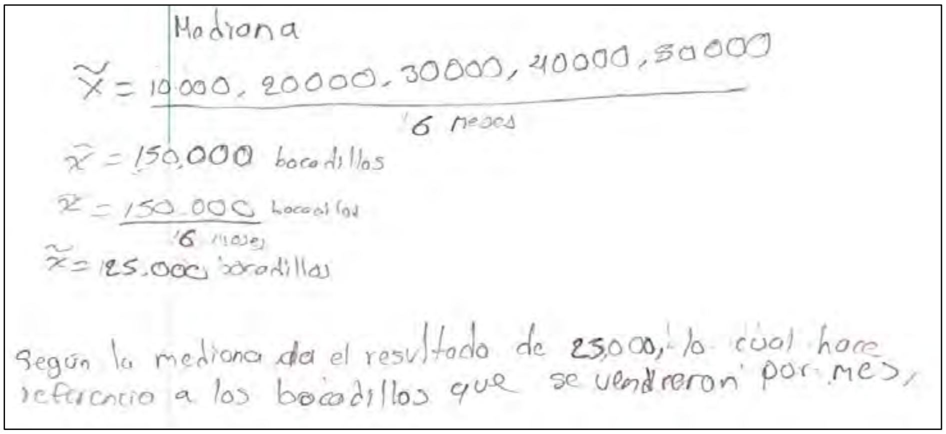


Figura 90. Ejemplo de respuesta C2.4.

C3. Calcula la mitad de cada valor de los datos. Se reporta un error semejante a este en Mayén (2009), pero en el de ella además de calcular la mitad de valor al final obtiene la mediana de los datos. En este caso, como se observa en la Figura 91 el alumno realiza la lectura a nivel literal de la gráfica, le asigna valores a la variable y con ellos calcula la mitad de cada uno, es decir los divide por dos. El resultado de cada valor, lo considera la mediana.

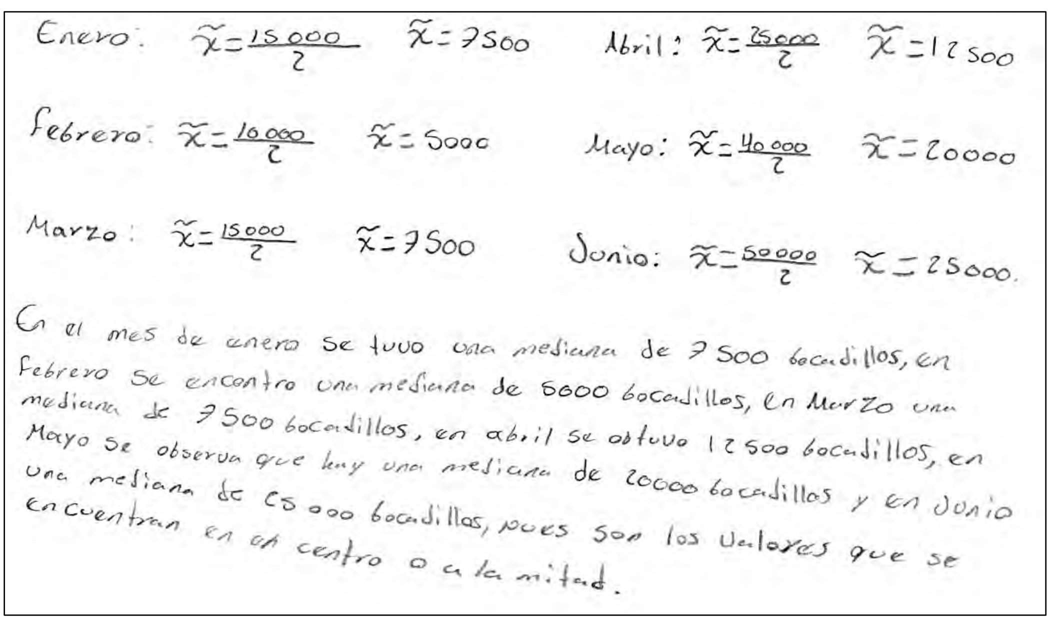


Figura 91. Ejemplo de respuesta C3.

La Tabla 14 muestra las frecuencias de cada categoría encontrada en el análisis de las respuestas del antes y después de la experiencia de aprendizaje.

Tabla 14. Categorías de respuesta del (inciso b), antes y después de la experiencia

Categorías de respuestas de la segunda parte del problema	Antes	Después
C1. Respuesta incorrecta, no usa la condición de indeterminación	1	0
C2. Obtiene la media en vez de la mediana	1	1
C3. Obtiene la moda en lugar de la mediana	2	0
C4. Cálculo de la mediana sin ordenar los datos	3	0
C5.1.Obtiene la mediana, pero confunde el valor de la variable con la escala y toma en cuenta el 0 como valor	2	2
C5.2.Obtiene la mediana a partir del eje de la escala incluyendo el 0, sin justificar con algoritmos	2	0
C5.3. Confunde las marcas de la escala con los valores de la variable	4	0
C6. La respuesta no tiene relación con la mediana	1	0
C7. No contesta	1	0
C1.1.Cálculo correcto de la mediana a partir de un gráfico	0	10
C1.2.Cálculo correcto de la mediana con error al convertir el gráfico en valores numéricos	0	1
C2.2. Cálculo incorrecto de la mediana con las marcas de la escala	0	1
C2.3. Obtiene la media, y realiza el cálculo incorrecto de la mediana	0	1
C3. Calcula la mitad de cada valor de los datos	0	1
Total	17	17

En la Tabla 14, se observa que antes de la experiencia, un estudiante usa la idea de la mediana, cabe destacar, que la gran mayoría tenía problemas al identificar la mediana como estadístico de orden, y otra parte, no logra observar los tipos de representación de los datos, es quizá el apartado con más errores. Asimismo se puede ver, que hubo un alumno que no contesto el problema y otra cantidad mayor no comprendió lo que solicita, y una gran mayoría se concentra en el error al interpretar los gráficos. Después de la experiencia 10 estudiantes son capaces de calcular la mediana a partir de un gráfico, es decir, realizan la lectura del gráfico, asignan valores y estiman la mediana. No obstante, se aprecian errores al momento de leer el grafico como es asignar valores semejantes pero sin un cero, se deduce que es un error de escritura al convertir el gráfico a valores numéricos, este error lo comete un estudiante. Del mismo modo, 5 alumnos cometen el error al interpretar el gráfico, confunden el valor de la variable con las marcas de escala del eje de las Y, sin embargo, de estos además de esto, 3 no son capaces de encontrar la mediana.

5.5. Análisis del problema 4

Este problema tomado de Orta (2014), denominado tiempos de espera en los cines es una modificación del original tomado de Shaughnessy (2006). Tiene la finalidad de indagar las ideas que tienen los estudiantes acerca de la variabilidad. Para lo cual, la primer parte tiene la intención de que los estudiantes construyan gráficas de frecuencias, la segunda parte pide obtener el promedio, además de ubicarlo en las gráficas construidas, y después verifiquen si las gráficas y los promedios son iguales en los tres cines. Con la intención de que los estudiantes tomen a consideración la utilización de las medidas de dispersión, en la tercera y cuarta parte del problema, se pide que comparen los tres cines, después en solo dos cines respectivamente, y que señalen en cual existe mayor variabilidad. Por último, quinta y sexta parte del problema pide tome decisiones, esperando que haga uso las medidas de dispersión.

Este problema solo fue aplicado después de la experiencia de aprendizaje, por lo que no se describe un antes y después. El análisis de este problema, se realiza en tres partes, en la primera parte se analiza la construcción de gráficos, en la segunda parte el cálculo de las medidas de tendencia central, en este caso se pide el cálculo de la media aritmética, y finalmente en la tercera parte, se analiza a partir de la distribución de datos, la variabilidad en la que los estudiantes tienen que interpretar verbalmente los resultados comparando los tres conjuntos de datos.

Ítem 4

Una costumbre en los cines es mostrar anuncios comerciales y cortos en la pantalla antes de comenzar la película. El tiempo de espera para una película es la diferencia entre la hora de inicio anunciada y la hora real en la que comienza la película. Un grupo de 10 estudiantes investigó el tiempo de espera de tres cadenas de cines, una de la cadena Cinemex, otra de la cadena Cinépolis y por último una de la cadena Multicinemas. Los estudiantes visitaron los cines en horarios y días diferentes. Cada alumno midió el tiempo de espera en minutos, y lo registró en las tablas que siguen:

Cinemex
12.0
21.0
15.0
15.0
13.0
16.0
16.0
16.0
20.0
18.0

Cinépolis
15.0
15.5
16.0
16.0
16.5
16.5
18.0
16.0
15.5
17.0

Multicinemas
15.5
17.0
18.0
16.5
16.0
16.5
16.5
15.0
15.0
16.0

4.1. Traza una gráfica para cada cadena de cine.

4.2. Calcula el tiempo de espera promedio para cada cadena de cines y señálalo con un segmento de recta en la gráfica que construiste.

Observa con atención las gráficas de la página siguiente.

Las gráficas que construiste son iguales o diferentes a las que vienen junto con este cuestionario.

Iguales______ Diferentes______

En caso de que no sean iguales ¿cuál es la diferencia entre ellas?

El tiempo promedio de espera es de 16.2 minutos en las tres cadenas de cines.

4.3 ¿En cuál de las tres cadenas de cines hay mayor variabilidad en el tiempo de espera? ¿Por qué?

4.4 En cuál de las dos cadenas de cines, Cinépolis o Multicinemas, hay mayor variabilidad.

Cinépolis______ Multicinemas_______ Explica tu respuesta.

4.5 Suponiendo que vives a la misma distancia de los tres cines mencionados, ¿a cuál asistirías?

Cinemex________ Cinépolis_________ Multicinemas___________ ¿Por qué?

4.6 ¿Crees que es importante la variabilidad en tiempos de espera que hay en los cines mencionados?

En la primera parte, se solicita trazar una gráfica para cada conjunto de datos, y Orta (2014) muestra cómo deberían ser las gráficas ideales para este punto, y que más adelante sirven para comparar con las que elaboran los estudiantes.

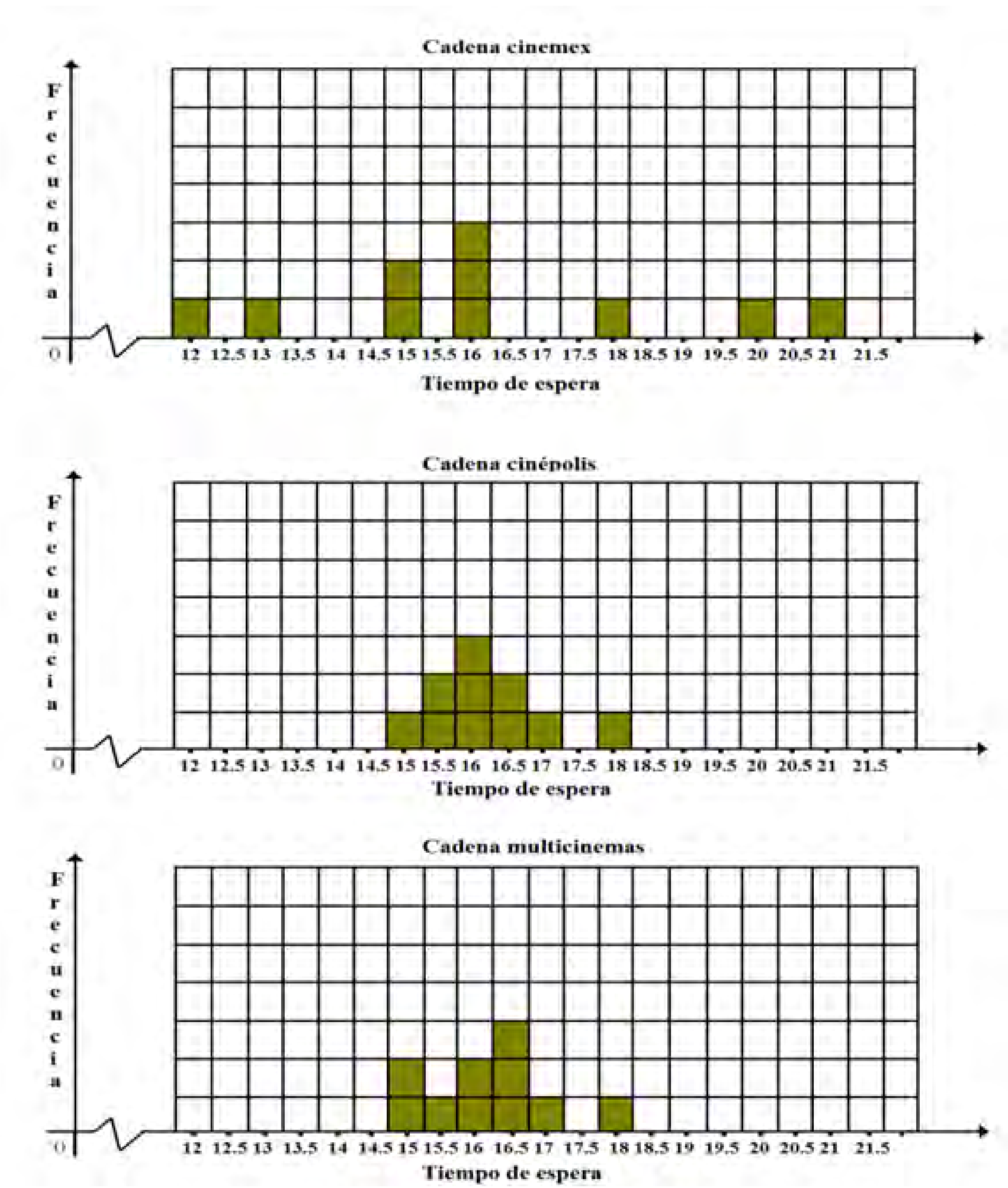


Figura 92. Gráficas sugeridas por Orta (2014), para comparar en el problema.

La segunda parte, pide calcular el tiempo de espera promedio para cada cadena de cines. Entonces, para resolver el problema:

Se deben sumar todos los valores de cada conjunto de datos y dividirlos entre el número total de datos, esto es para cada cadena de cine por lo que queda:

Cinemex

$$\bar{x} = \frac{12 + 21 + 15 + 15 + 13 + 16 + 16 + 16 + 20 + 18}{10}$$

$$\bar{x} = \frac{162}{10}$$

$$\bar{x} = 16.2$$

Cinépolis

$$\bar{x} = \frac{15 + 15.5 + 16 + 16 + 16.5 + 16.5 + 18 + 16 + 15.5 + 17}{10}$$

$$\bar{x} = \frac{162}{10}$$

$$\bar{x} = 16.2$$

Multicinemas

$$\bar{x} = \frac{15.5 + 17 + 18 + 16.5 + 16 + 16.5 + 16.5 + 15 + 15 + 16}{10}$$

$$\bar{x} = \frac{162}{10}$$

$$\bar{x} = 16.2$$

La tercera parte, solicita una situación de comparación de conjuntos de datos, una vez presentada la información en tablas y gráficas, se calcula la media aritmética que resulta ser la misma para los tres conjuntos de datos por lo que los estudiantes deben considerar las medidas de dispersión (rango y desviación estándar) para determinar la variabilidad de cada conjunto de datos, entonces, los cálculos debieron quedar de la siguiente forma:

Tabla 15. Media, rango y desviación estándar de los tres cines

Cines	Media	Rango	Desviación estándar
Cinépolis	16.2	3	0.81
Cinemex	16.2	9	2.68
Multicinemas	16.2	3	0.87

La interpretación correcta es considerar el tiempo de espera en el cine en términos de incertidumbre: es decir, entre mayor sea el rango, mayor incertidumbre en el tiempo de espera, del mismo modo la desviación estándar considerando la media. En el mismo sentido, en Cinemex los tiempos de espera van de 12 a 21 minutos, mientras que en Cinépolis y Multicinemas los tiempos van de 15 a 18 minutos. Por tanto, en términos del rango Cinemex tiene mayor incertidumbre, pues la película puede tardar o no en iniciar, mientras que en Cinépolis y Multicinemas es menor la incertidumbre, pero entre estas dos cadenas es igual. Por otro lado, considerando la desviación estándar entorno a la media, se puede confirmar que en Cinemex los tiempos de espera tienen mayor dispersión y la incertidumbre es mayor que en las otras dos cadenas, pero comparando las desviaciones estándar de Cinépolis y Multicinemas, el que menor dispersión tiene en los tiempos de espera es Cinépolis, por lo que la incertidumbre es menor que en Multicinemas y por tanto, también es menor comparándolo con Cinemex. Por tanto, se esperaba que los estudiantes emitieran un argumento similar al elegir una cadena de cines y justificar su elección.

En este sentido, los estudiantes deben determinar un valor representativo del conjunto de datos, como la media y relacionarla con la de dispersión, lo que les ayudará a elegir uno de los conjuntos de datos.

En las tablas 16, 17 y 18, se describen los elementos de significado para cada apartado.

Tabla 16. Elementos de significado de la primera parte del problema 4

Elementos de significado	Contenido
Definiciones	▪ Gráfico de barras ▪ Histograma ▪ Frecuencias absolutas ▪ Variable estadística
Propiedades	▪ Proporcionalidad (altura de la barra con valor de la variable). ▪ La suma de frecuencias ha de ser igual al tamaño de la muestra
Campos de problemas	▪ Construir un gráfico para cada conjunto de datos
Algoritmos y procedimientos	▪ Recuento de los datos y agruparlos en una tabla de distribución de frecuencias ▪ Seleccionar el grafico a construir (gráfico de barras o histograma) ▪ Representar en el eje X: los tiempos de espera en los cines. ▪ Representar en el eje Y: las frecuencias absolutas. ▪ Representar las barras con una altura proporcional a las frecuencias.
Lenguaje	▪ Gráfico: líneas, puntos, marcas ▪ Verbal: título del gráfico, etiquetas, rótulos en los ejes. ▪ Numérico: para representar las frecuencias y los tiempos de espera
Argumentos	

Tabla 17. Elementos de significado de la segunda parte del problema 4

Elementos de significado	**Contenido**
Definiciones	▪ Definición de media como promedio ▪ Definición de media como algoritmo
Propiedades	▪ La media es un valor perteneciente al rango de la variable ▪ La media es un representante del conjunto de datos
Campos de problemas	▪ Calcular la media para cada conjunto de datos simples
Algoritmos y procedimientos	*Algoritmos* ▪ Cálculo media de un conjunto de datos aislados *Procedimientos* ▪ Sumar todos los datos ▪ Dividir la suma entre en número total de datos
Lenguaje	▪ Verbal ▪ Numérico ▪ Simbólica
Argumentos	▪ Razonamientos verbales deductivos

Tabla 18. Elementos de significado de la tercera parte del problema 4

Elementos de significado	**Contenido**
Definiciones	▪ Definición de media como representante de un conjunto de datos ▪ Definición de rango como medida de dispersión ▪ Definición de desviación estándar como medida de dispersión Definición de promedio ▪ Definición de variabilidad ▪ Concepto de distribución
Propiedades	▪ La media es representante de un colectivo ▪ La desviación estándar muestra el nivel de dispersión de los datos en una distribución. Cuando el valor es mayor, mayor será la dispersión y viceversa.
Campos de problemas	▪ Interpretar la variabilidad verbalmente de los resultados comparando los tres conjuntos de datos.
Algoritmos y procedimientos	▪ Estimar el rango al realizar la diferencia aritmética del valor más bajo y del valor más alto de cada conjunto de datos ▪ Estimar la desviación estándar obteniendo la raíz cuadrada del promedio de los cuadrados de la diferencia entre cada valor y la media para cada conjunto.
Lenguaje	▪ Numérico ▪ Verbal
Argumentos	▪ Razonamientos verbales deductivos

Categorías de respuestas de la primera parte del problema

C1. Gráfico de barras adosado para cada conjunto de datos (Correcto). En esta categoría se encuentran aquellos estudiantes que logran construir un gráfico correcto para cada conjunto de datos, como se muestra en la Figura 93. Reconoce la variable estadística del problema, así como la proporcionalidad (altura de la barra con valor de la variable) y que la suma de frecuencias ha de ser igual al tamaño de la muestra. Sigue un procedimiento adecuado para la construcción del gráfico, primero realiza una distribución de frecuencias, determina las escalas, en el eje X representa los tiempos de espera en los cines, en el eje Y representa las frecuencias absolutas. Finalmente, es deficiente el uso del lenguaje verbal: el título del gráfico, etiquetas y los rótulos en los ejes y utiliza el lenguaje numérico.

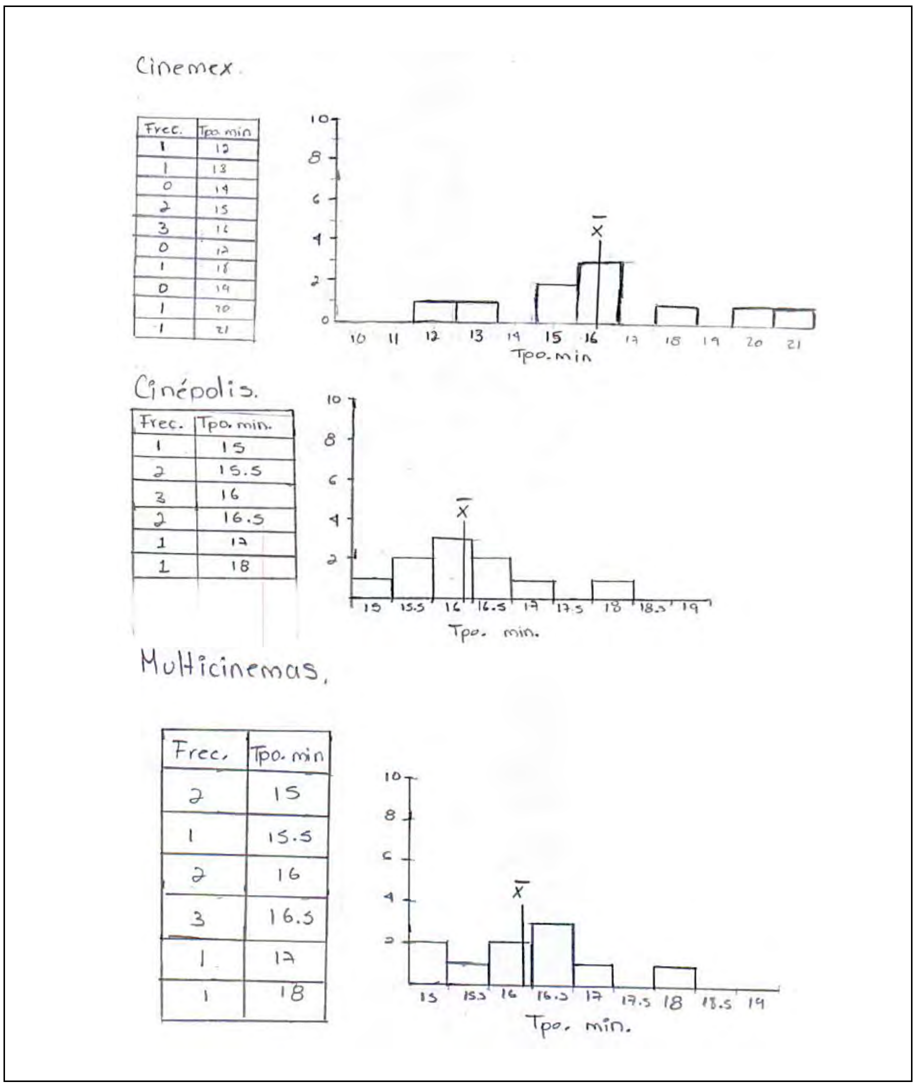


Figura 93. Ejemplo de respuesta C1.

C2. Gráfico de barras con errores en escala (parcialmente correcto). Las partes que conforman un gráfico según Friel, Curcio y Bright (2001) son el marco del mismo, en el que se incluyen las escalas, los ejes, el sistema de coordenadas seleccionado y las etiquetas de las escalas y ejes. Son importantes elementos que nos facilitan la lectura de los gráficos. Por lo que representar las escalas de forma correcta permitirá una correcta interpretación. En este problema se puede ver el uso de una variable continua que, en un conjunto de datos tiene valores enteros y en los dos restantes conjuntos tiene decimales, lo que complica su representación. Esta categoría se compone de las respuestas que dan los estudiantes al realizar un gráfico correcto, pero tienen algún error en la escala del eje X. En lo que sigue se describen sus dos variantes:

C2.1. Gráfico de barras separadas con valores faltantes en el eje X. Este tipo de error también fue encontrado en Bruno y Espinel (2005) y en Arteaga (2011), donde existe una representación errónea de los números naturales en la recta real. Es decir, en esta categoría los alumnos omiten algunos valores de la variable, pues la frecuencia es cero. No obstante, contiene los demás elementos y acompañado de su distribución de frecuencias, como se puede observar (Figura 94).

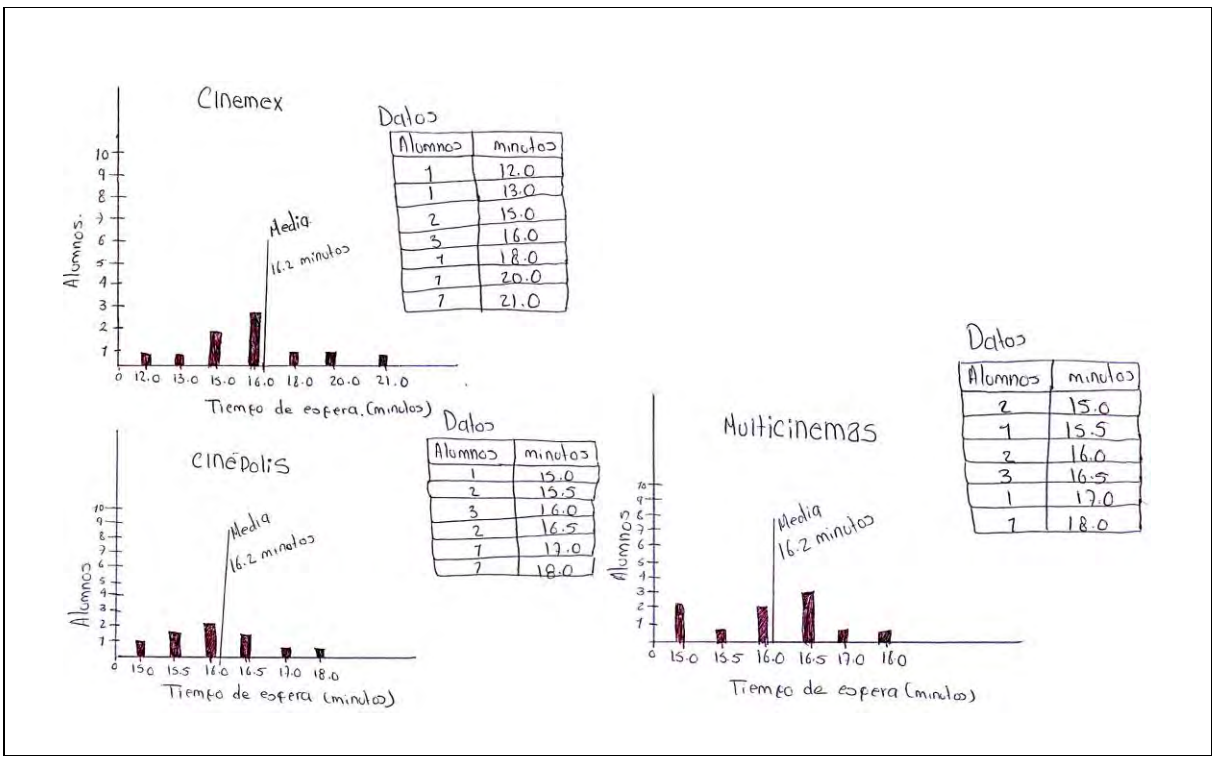


Figura 94. Ejemplo de respuesta C2.1.

C2.2. Gráfico de barras adosado con valores faltantes en el eje X. La Figura 95 es un ejemplo de esta variante, en la que existe el error en la sucesión de los números naturales de la recta real. Algo diferente son los rectángulos adosados del lenguaje gráfico.

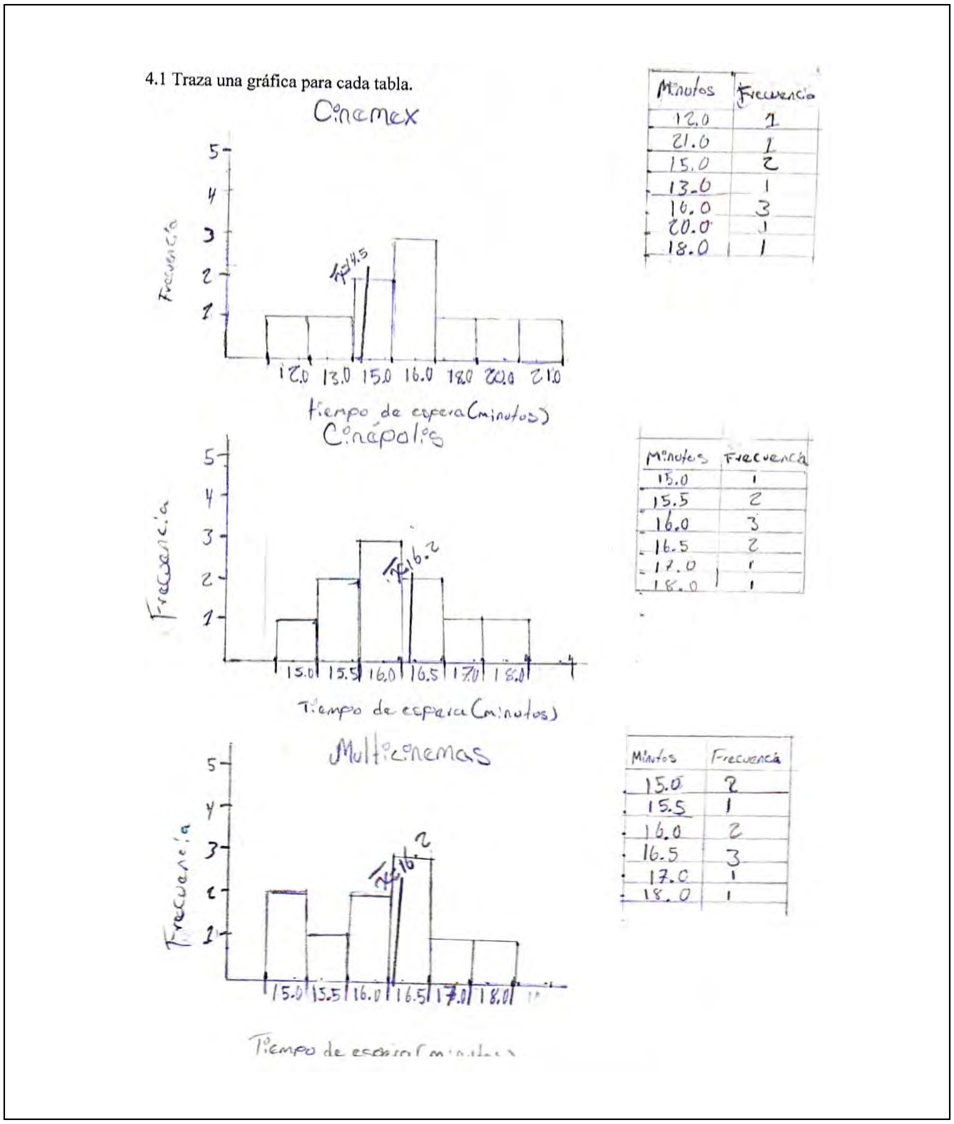


Figura 95. Ejemplo de respuestas C2.2.

C3. Gráficos con barras no centradas (incorrecto). Como la variable es continua, en este problema los jóvenes eligen construir un histograma que es lo más común para este tipo de variable, en la que es necesario agrupar los datos en intervalos de clase. Este error, también descrito en Bruno y Espinel (2005), Espinel (2007) y Arteaga (2011), para variables discretas, en el que tanto en histogramas como en gráficos de barras, los estudiantes no centran las barras en las marcas de clase, siendo éstas los valores enteros, en nuestro caso hay ambas (enteros y decimales), y las barras deberían estar entradas en estos valores. En este trabajo se encontraron dos variantes que se presentan en seguida:

C3.1. Barras no centradas y representación errónea de los números naturales en la recta real. (Incorrecto). En esta categoría se encuentran las respuestas de los estudiantes que cometen dos errores en la construcción del gráfico (Figura 96), el primer error está en no centrar las barras en los valores de la variable, el segundo error está en la recta numérica, pues omiten algunos valores de la variable en el eje X, se intuye que por tener frecuencia nula.

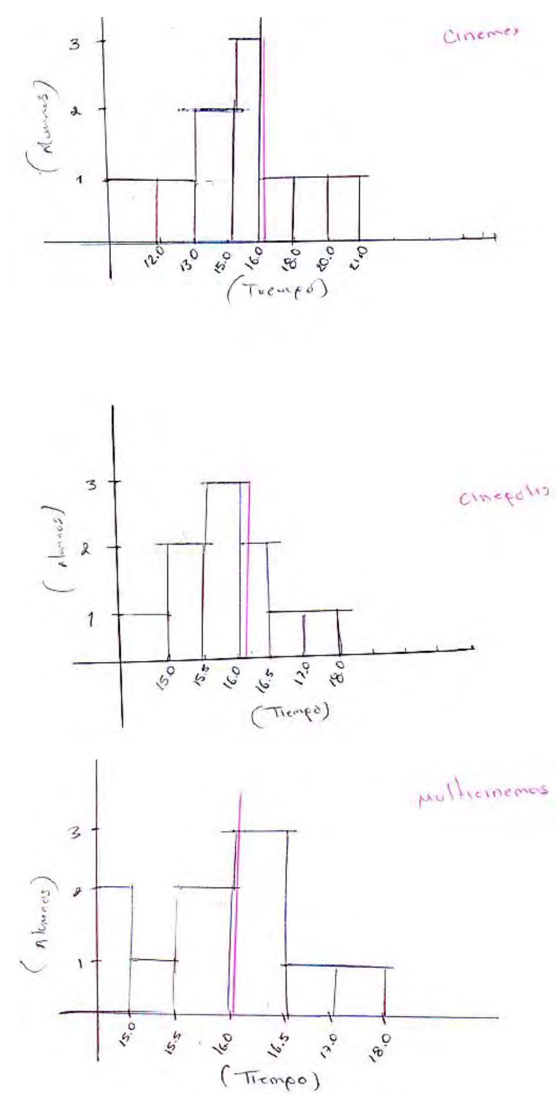


Figura 96. Ejemplo de respuesta C3.1

C3.2. Barras no centradas y error al agrupar los datos (Incorrecto). Los estudiantes de esta categoría no centran las barras en los valores de las variables, como se puede ver en la Figura 97. Además cometen un error al agrupar los datos en la distribución de frecuencias, pues los valores con decimales lo reducen en el entero más próximo. También, presentan deficiencia en los rótulos de los ejes.

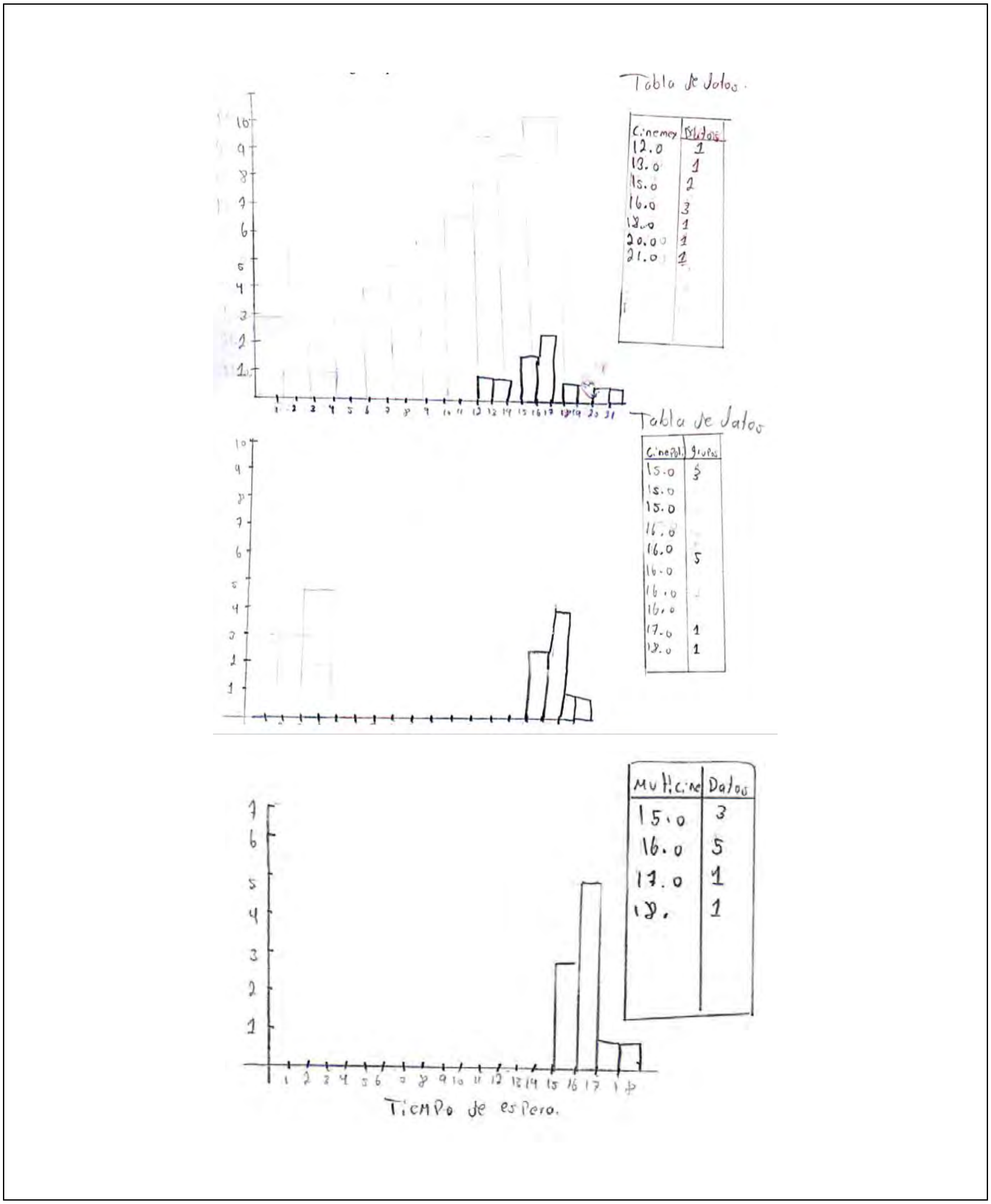


Figura 97. Ejemplo de respuesta C.3.2.

C4. Recuento erróneo de frecuencias y faltan las etiquetas de las escalas del eje X. En esta categoria se puede observar (Figura 98) que el alumno, no asigna los valores numéricos en la recta real del eje X, además contabiliza de manera erronea las frecuencias pues cuenta cuantas veces se repite el valor y se lo asigna a cada dato que ha ordenado de menor a mayor. Por otra parte, los rotulos de los ejes son incorrectos.

Figura 98. Ejemplo de respuesta C4.

C5. Graficos con errores en los ejes. Se encontraron diferentes errores en esta categoria, en la que los estudiantes no identifican correctamente los elementos que deberían ir por ejemplo en los ejes, y en algunas ocasiones los intercambian o escriben valores que no estan en los datos. a continuacion se presentan algunas variantes:

C5.1. Confunde la frecuencia y escribe los días de la semana en lugar de el valor de la varible. La Figura 99 muestra el ejemplo de esta categoria en la que los estudiantes escriben en el eje de las frecuencias los valores que toma la variable (en este caso los tiempos de espera), del mismo modo, que en el eje Y, escribe para cada dato un día de la semana como si hubiera sido el orden de la visita al cine, pues ordena los datos en la distribución de frecuencias en orden ascendente, por lo que, se deduce que no identifica correctamente la variable estadistica.

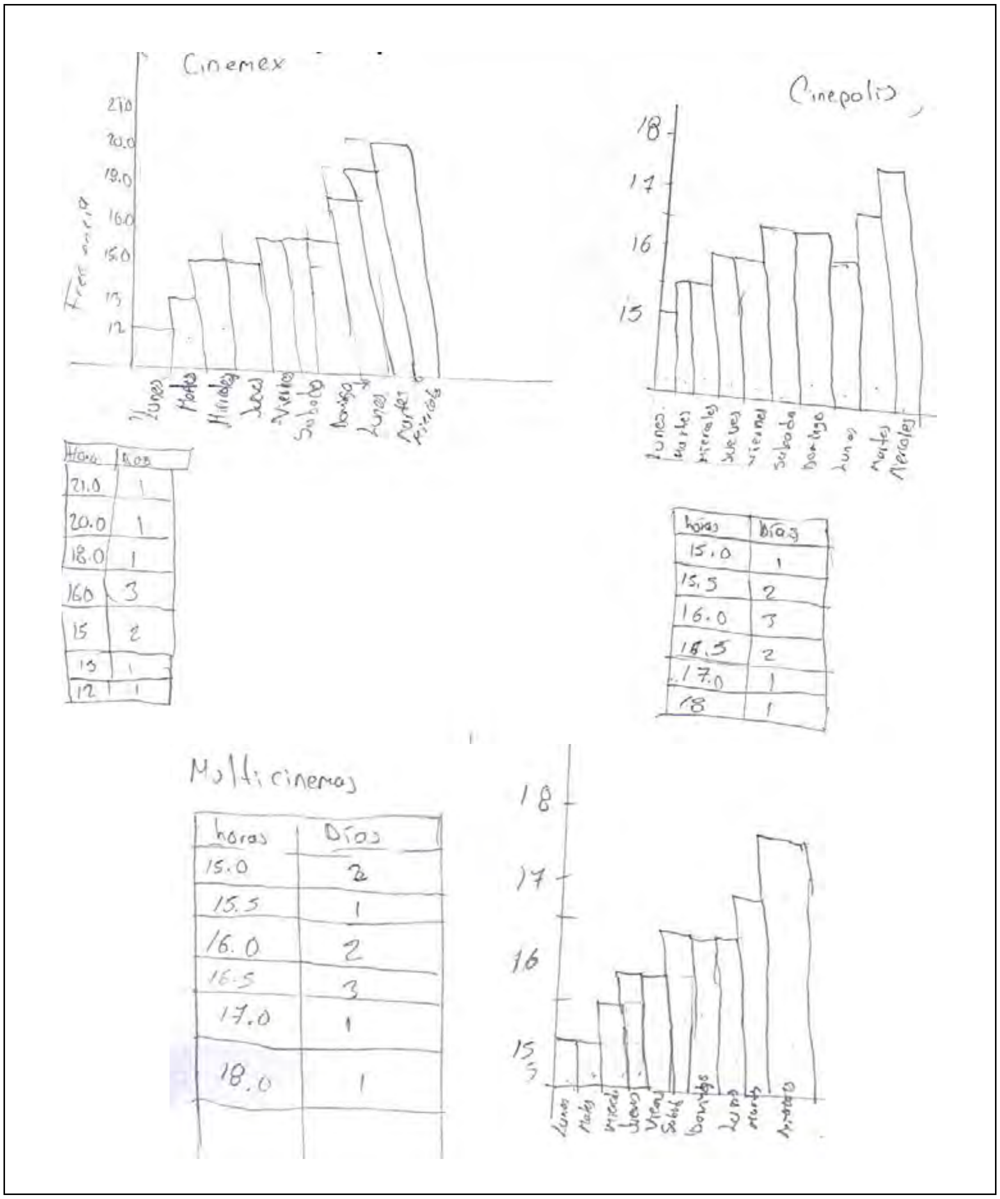


Figura 99. Ejemplo de respuesta C5.1.

C5.2.Intercambia frecuencia y valor de la variable. El error que se describe en esta categoria tambien fue encontrado por Ruiz (2006) y Arteaga (2011). Los estudiantes de esta categoria no toman en cuenta la frecuencia de los tiempos de espera, como se observa en la Figura 100, solo consideran las visitas y los valores de la variable, entonces cometen el error de confundir la distribucion de frecuencias, la variable independiente y la dependiente. Ademas, utilizan escalas diferentes en los dos ejes, al construir un grafico de barras adosadas, en el que intercambian la variable tiempo de espera y las frecuencias absolutas.

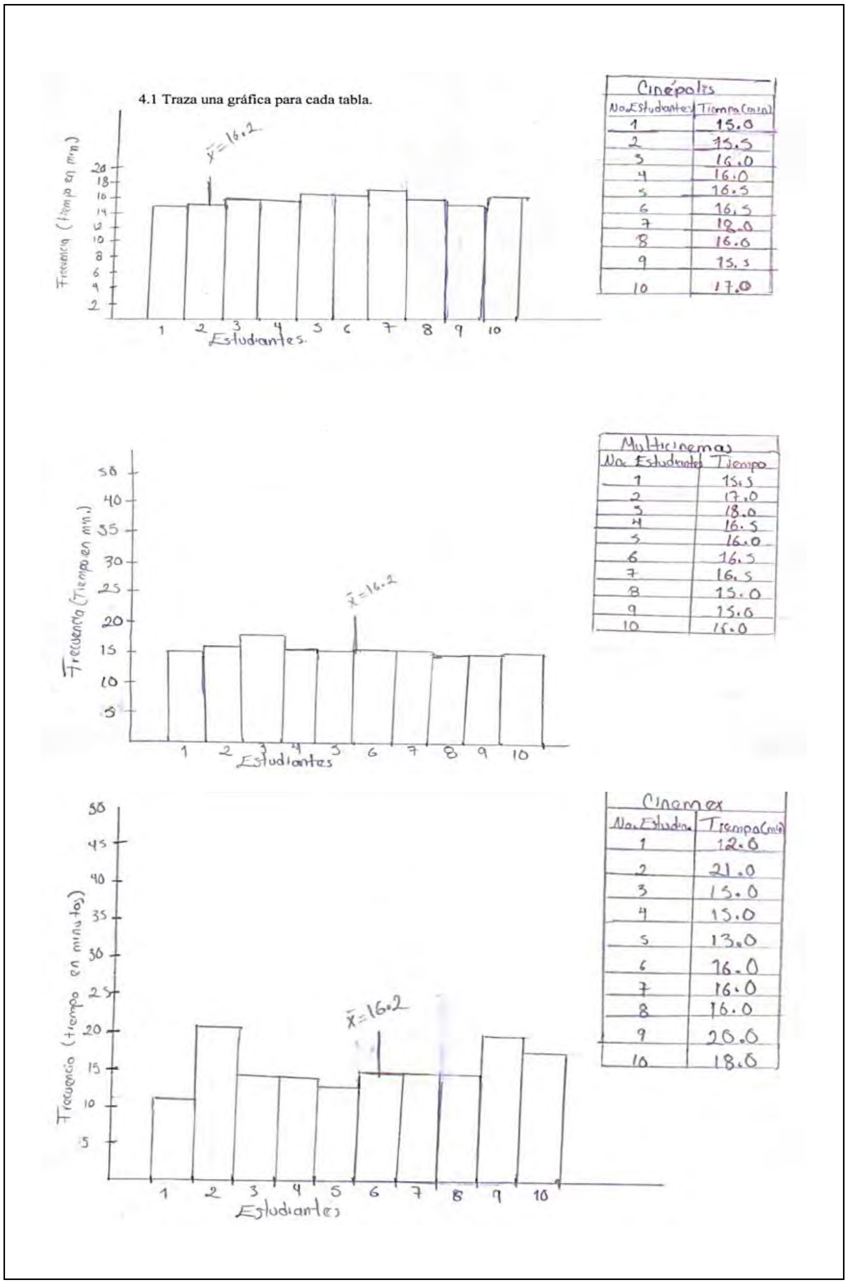


Figura 100. Ejemplo de respuesta C5.2.

Una vez analizado y categorizado las respuestas que dan los estudiantes, se presentan en la Tabla 19, donde se integran las frecuencias, para realizar las conclusiones del ítem 4.1 y determinar la frecuencia de los aciertos y errores de los estudiantes a la hora de construir un gráfico a partir de un conjunto de datos aislados.

Tabla 19. Frecuencias de las categorías de la primera parte del problema

Categorías de respuestas de la primera parte	**Frecuencia**
C1. Gráfico de barras adosado para cada conjunto de datos (Correcto)	2
C2.1. Gráfico de barras separadas con valores faltantes en el eje X	3
C2.2. Gráfico de barras adosado con valores faltantes en el eje X	3
C3.1. Barras no centradas y representación errónea de los números naturales en la recta real. (Incorrecto)	2
C3.2. Barras no centradas y error al agrupar los datos (Incorrecto)	2
C4. Recuento erróneo de frecuencias y faltan las etiquetas de las escalas del eje X	1
C5.1. Confunde la frecuencia y escribe los días de la semana en lugar del valor de la variable	1
C5.2.Intercambia frecuencia y valor de la variable	3
Total	**17**

En los resultados que se muestran en la Tabla 19, se encontró que 2 alumnos construyen de forma correcta un gráfico, también se han considerado parcialmente correctos 6, entonces considerando estas categorías, se ve que 8 estudiantes fueron capaces de construir un gráfico de forma correcta, es decir, la mayoría no logra construir el grafico, lo que sería una desventaja al analizar la variabilidad de 2 o más conjuntos de datos. En este sentido, se ha encontrado que 9 estudiantes del total de nuestra muestra, cometen errores como, no centrar las barras, error que comenten 4 alumnos, un joven no escribe las etiquetas de los datos, solo hace la división de la recta del eje X, por último, 3 estudiantes confunden las frecuencias con los valores de la variable y los intercambian en los ejes.

Categorías de respuestas de la segunda parte del problema

C1. Correcta, mediante el cálculo de la media con datos aislados. Los estudiantes de esta categoría que son la mayoría usan el algoritmo de la media simple, para calcular el promedio de cada conjunto de datos, se deduce que tienen la definición de media como promedio. Se observa en la Figura 101, el uso correcto del algoritmo de la media, sus símbolos y procedimientos, aunque no realizan argumento alguno, sobre cómo lo resolvieron.

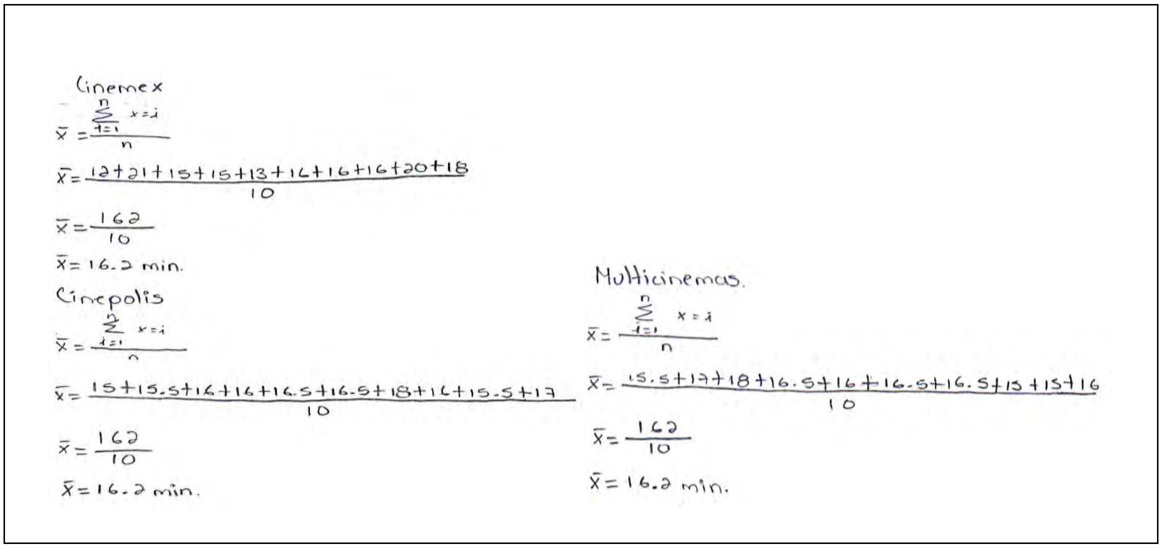


Figura 101. Ejemplo de respuesta C1.

C2. Error en el cálculo de la media. Los estudiantes de esta categoría cometen un error al sumar todos los datos, no obstante, conocen y aplican el algoritmo de la media y sus procedimientos, también hacen uso correcto de los símbolos como se observa en el ejemplo (Figura 102).

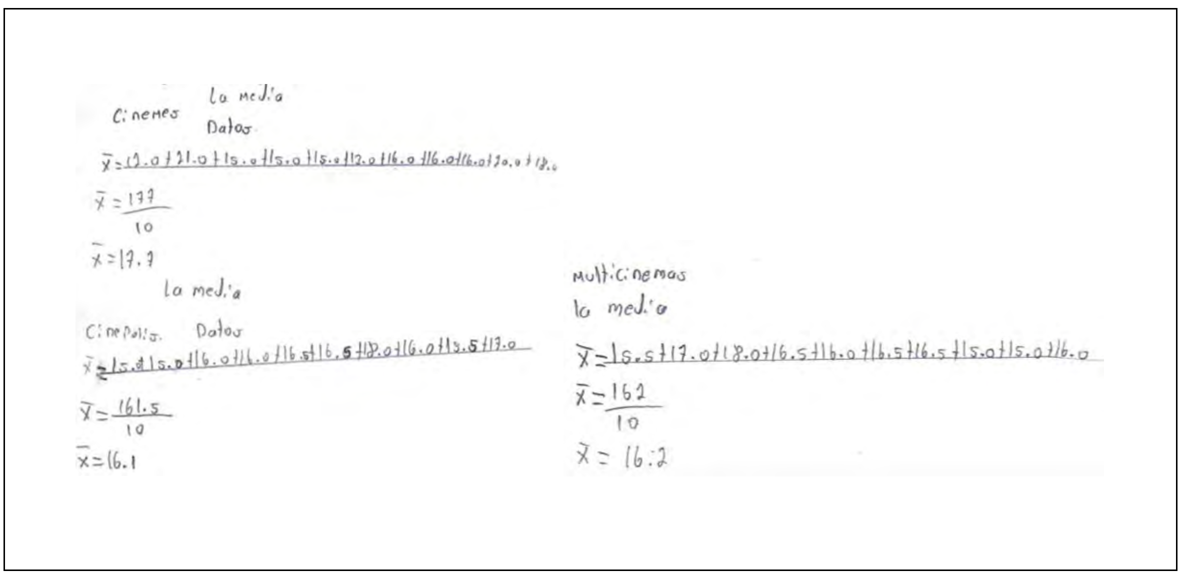


Figura 102. Ejemplo de respuesta C2.

C3. Error al usar la idea de media ponderada. En esta categoría se puede ver (Figura 103) que el estudiante utiliza la idea de media ponderada cuando no es necesario ponderar los datos, sin embargo, considera las 10 observaciones y según el orden en el que se encuentran los datos, los enumera del 1 al 10 y los multiplica por el orden respectivo, una vez haciendo esta multiplicación suma los resultados y los divide por diez. Además, la respuesta la convierte a horas.

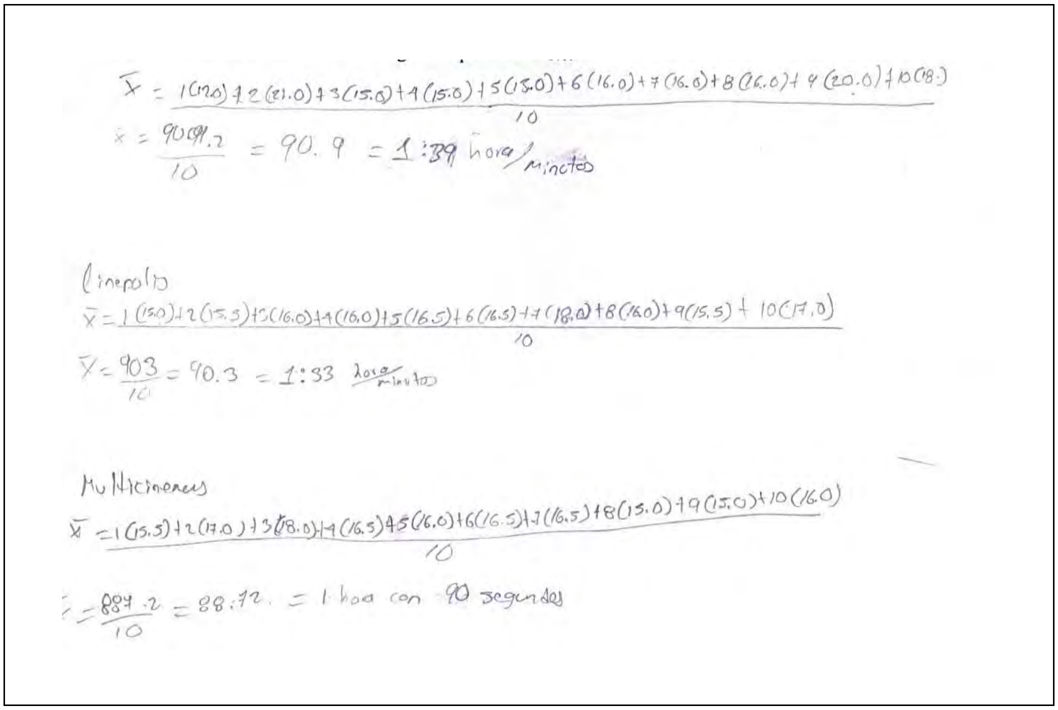


Figura 103. Ejemplo de respuesta C3.

Seguidamente, en la Tabla 20, se han agrupado cada categoría resultante de este ítem, se presenta cada una con su respectiva frecuencia.

Tabla 20. Frecuencias de las categorías de la segunda parte del problema

Categorías de respuestas de la segunda parte	Frecuencia
C1. Correcta, mediante el cálculo de la media con datos aislados	13
C2. Error en el cálculo de la media	3
C3. Error al usar la idea de media ponderada	1
Total	**17**

Los resultados de este problema muestran que la mayoría, 13 estudiantes, logran estimar el promedio de forma correcta; realizan un uso adecuado de los objetos matemáticos que son parte de este problema. No obstante, se observa que 3 personas, a pesar de hacer el uso correcto de los algoritmos y procedimientos, cometen un error al sumar los datos y por ende la media resulta incorrecta. Por último, hay un estudiante que en vez de calcular la media simple usa la idea de media ponderada y comete el error de multiplicarlo por el número de orden del dato.

Categorías de respuestas de la tercera parte del problema

C1. Reconoce la variabilidad de la distribución de un conjunto de datos. Está integrada por una variedad de respuestas, en ella se engloban todas aquellas respuestas que se han considerado correctas y parcialmente correctas. Los estudiantes de esta categoría consideran la dispersión de la distribución de cada conjunto de datos al hacer la comparación, se observa en las variantes, que se presentarán a continuación, que los estudiantes tienen nociones del concepto de variabilidad pues señalan la dispersión o lejanía de los datos para justificar su respuesta, no se observa ninguna operación o comparación entre resultados por lo que puede deducirse que solo usó la inspección de las gráficas y de las tablas con los tres conjuntos de datos para comparar, finalmente unos toman en cuenta la variabilidad para la toma de decisiones otros no.

C1.1. Identifica la dispersión en la distribución y la considera en la toma de decisiones. En esta variante, se observa en la Figura 104, los estudiantes identifican la dispersión a través de la separación entre los datos, con lo que justifica al conjunto de datos con mayor variabilidad. También se observa que considera la variabilidad para la toma de decisiones, pues elige al conjunto de datos con menos variabilidad.

¿En cuál de las tres cadenas de cines hay mayor variabilidad en el tiempo de espera?
¿Por qué? En la Cadena cinemex devido a las distancias que hay de un valor al otro

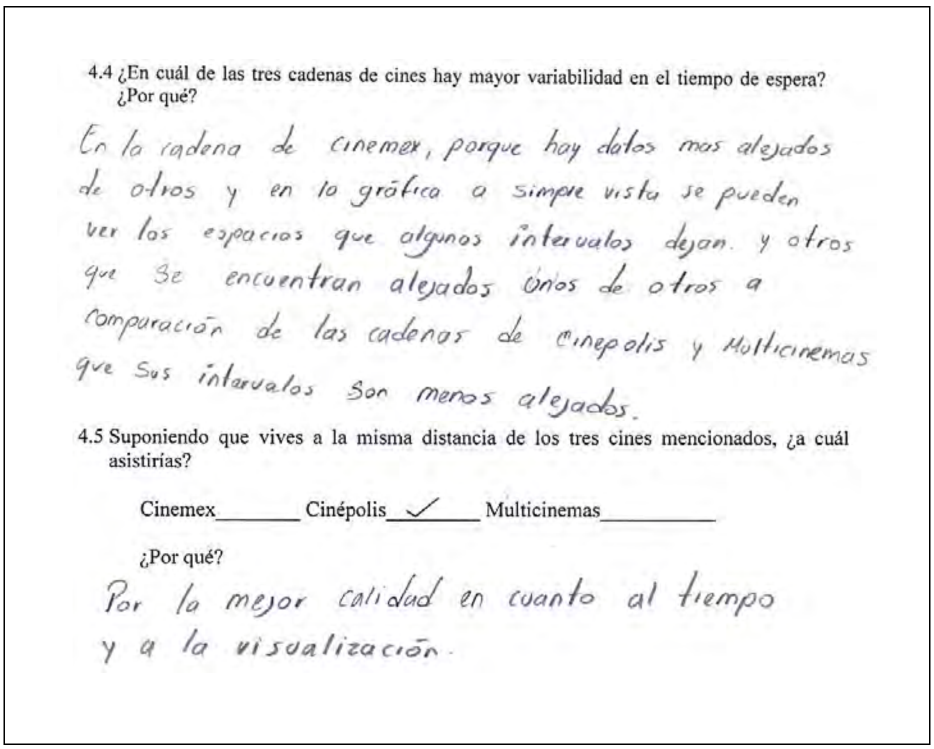
Suponiendo que vives a la misma distancia de los tres cines mencionados, ¿a cuál asistirías?

Cinemex________ Cinépolis___√___ Multicinemas________

¿Por qué? porque en esta no hay mucha variabilidad en los tiempos de espera

Figura 104 . Ejemplo de Respuesta C1.1.

C1.2. Nivel de dispersión de los datos en una distribución, pero no lo considera en la toma de decisiones. En la Figura 105, se deduce que los estudiantes de esta categoría, de forma implícita, describen la propiedad de la desviación estándar que dice: La desviación estándar muestra el nivel de dispersión de los datos en una distribución. Cuando el valor es mayor, mayor será la dispersión y viceversa", pues se observa que el estudiante por inspección en las gráficas, que realizó, identifica la lejanía de los datos para cada conjunto y justifica su elección al considerar el conjunto con mayor separación en sus datos. No obstante, en la toma de decisiones no toma en cuenta la variabilidad pues aunque elige el conjunto de datos menos disperso justifica su elección según la calidad de la cadena de cines.


4.4 ¿En cuál de las tres cadenas de cines hay mayor variabilidad en el tiempo de espera? ¿Por qué?

En la cadena de cinemex, porque hay datos mas alejados de otros y en la gráfica a simple vista se pueden ver los espacios que algunos intervalos dejan. y otros que se encuentran alejados unos de otros a comparación de las cadenas de Cinepolis y Multicinemas que sus intervalos son menos alejados.

4.5 Suponiendo que vives a la misma distancia de los tres cines mencionados, ¿a cuál asistirías?

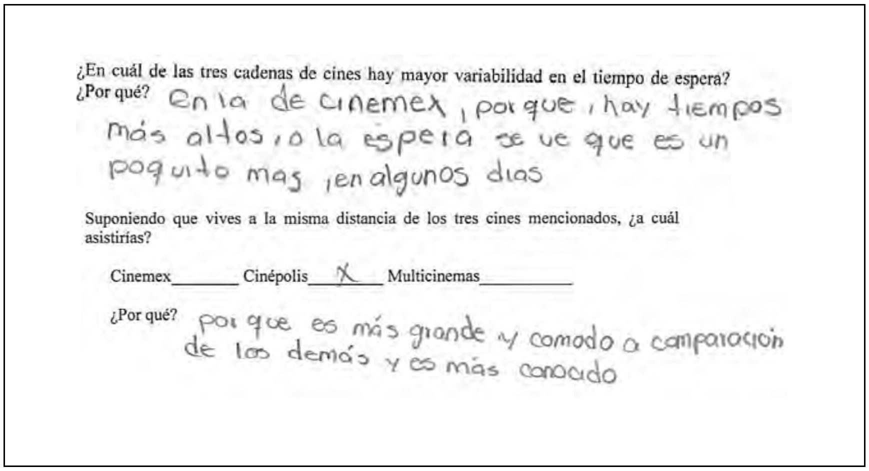
Cinemex_______ Cinépolis__✓__ Multicinemas_______

¿Por qué?

Por la mejor calidad en cuanto al tiempo y a la visualización.

Figura 105. Ejemplo de respuestas C1.2.

C1.3. Dispersión de los datos descrita con base en un máximo. Se observa en la Figura 106, que los estudiantes de esta categoría describen la dispersión tomando en cuenta los tiempos más altos y comparándolos con los de los demás conjuntos, finalmente eligen el conjunto más disperso, y para tomar la decisión, según los tiempos de espera, eligen el conjunto con menos variabilidad, pero justifican su elección según la calidad y reconocimiento de la cadena de cines.


¿En cuál de las tres cadenas de cines hay mayor variabilidad en el tiempo de espera?
¿Por qué? En la de cinemex, porque, hay tiempos más altos, o la espera se ve que es un poquito mas, en algunos dias

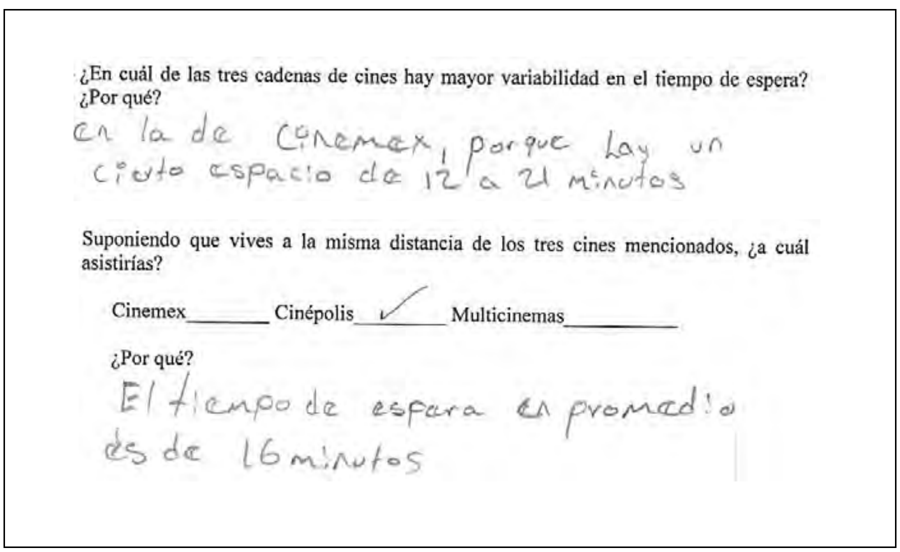
Suponiendo que vives a la misma distancia de los tres cines mencionados, ¿a cuál asistirías?

Cinemex_______ Cinépolis___X___ Multicinemas__________

¿Por qué? porque es más grande y comodo a comparación de los demás y es más conocido

Figura 106. Ejemplo de respuesta C1.3

C1.4. Separación de los datos con base en el rango. En esta categoría los estudiantes se refieren al rango sin mencionarlo explícitamente pero se observa en la Figura 107, que considera el valor mínimo y máximo del conjunto de datos más disperso. Además toma la decisión del conjunto de datos menos disperso pero lo justifica con la moda (el valor que más se repite).


¿En cuál de las tres cadenas de cines hay mayor variabilidad en el tiempo de espera?
¿Por qué?
en la de Cinemex, porque hay un cierto espacio de 12 a 21 minutos

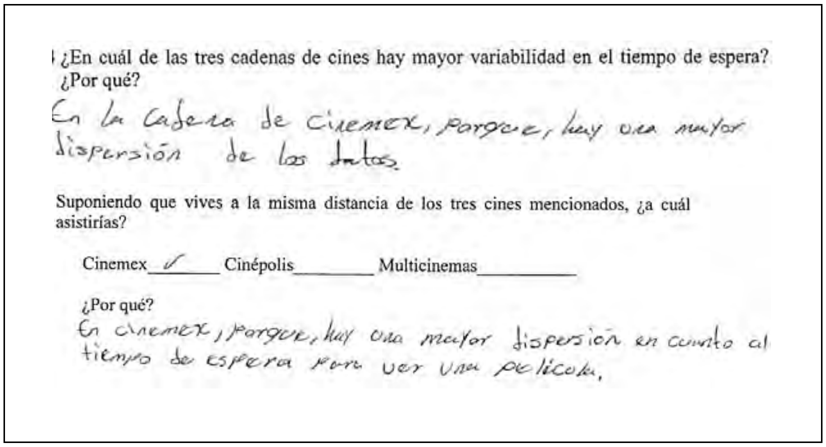
Suponiendo que vives a la misma distancia de los tres cines mencionados, ¿a cuál asistirías?

Cinemex_______ Cinépolis___✓___ Multicinemas__________

¿Por qué?
El tiempo de espera en promedio es de 16 minutos

Figura 107. Ejemplo de respuesta C1.4

C1.5. Identifica la dispersión en la distribución, pero no la considera para la toma de decisiones. La Figura 108, muestra el ejemplo de la categoría donde los estudiantes comparan los tres conjuntos de datos y consideran al que tiene mayor dispersión, pero al tomar una decisión, no eligen al menos disperso sino todo lo contrario, por lo que se deduce que no considera la incertidumbre en el tiempo de espera de la cadena de cines elegida.



¿En cuál de las tres cadenas de cines hay mayor variabilidad en el tiempo de espera? ¿Por qué?

En la cadena de cinemex, porque, hay una mayor dispersión de los datos.

Suponiendo que vives a la misma distancia de los tres cines mencionados, ¿a cuál asistirías?

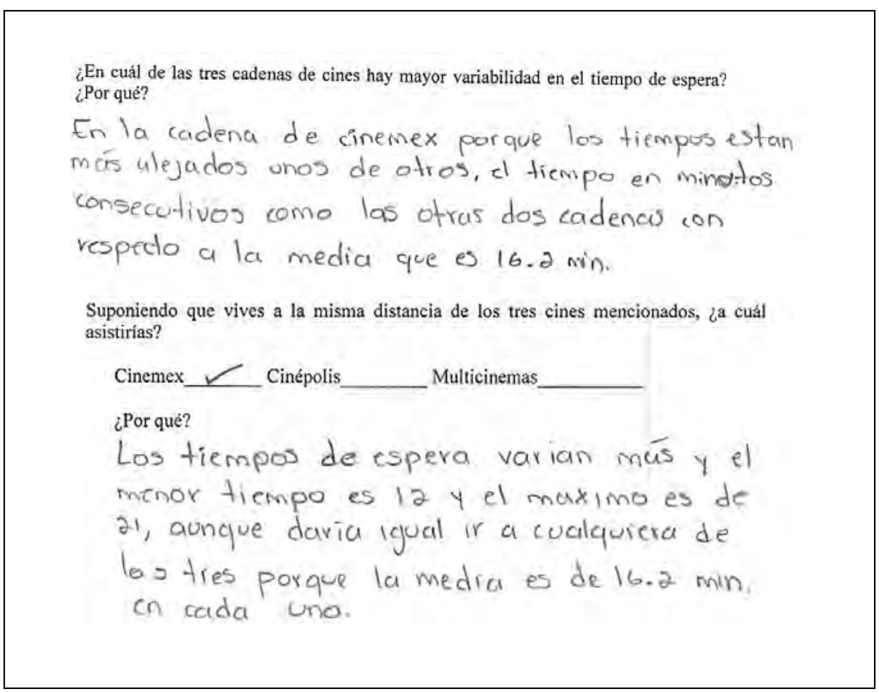
Cinemex ✓ Cinépolis________ Multicinemas__________

¿Por qué?

En cinemex, porque, hay una mayor dispersión en cuanto al tiempo de espera para ver una pelicula.

Figura 108. Ejemplo de C1.5

C1.6. Dispersión del conjunto de datos con base en la media, pero no la consideran en la toma de decisiones. Los estudiantes de esta categoría señalan al conjunto más disperso considerando la distribución de los datos y la lejanía de los mismos respecto a una medida de tendencia central (media), como se ve en la Figura 109. Pero al tomar una decisión siguen eligiendo al conjunto más disperso no lo contrario, aunque considera en este caso el rango, justifica que daría igual ir a cualquier cine, pues el promedio es el mismo para las tres cadenas.



¿En cuál de las tres cadenas de cines hay mayor variabilidad en el tiempo de espera? ¿Por qué?

En la cadena de cinemex porque los tiempos estan más alejados unos de otros, el tiempo en minutos consecutivos como las otras dos cadenas con respecto a la media que es 16.2 min.

Suponiendo que vives a la misma distancia de los tres cines mencionados, ¿a cuál asistirías?

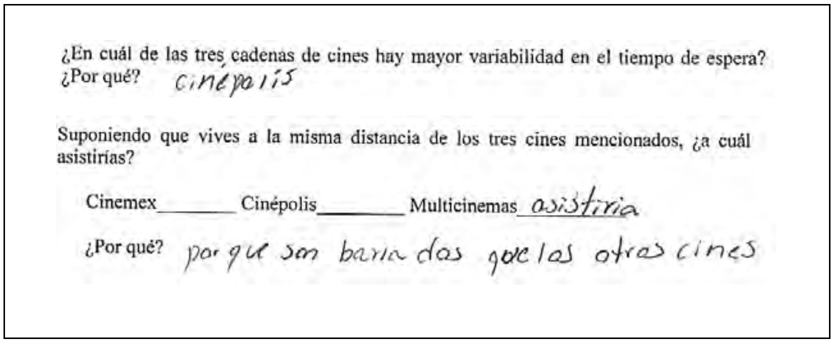
Cinemex ✓ Cinépolis________ Multicinemas__________

¿Por qué?

Los tiempos de espera varian más y el menor tiempo es 12 y el maximo es de 21, aunque daria igual ir a cualquiera de los tres porque la media es de 16.2 min. en cada uno.

Figura 109. Ejemplo de Respuesta C1.6.

C.2. No identifica la variabilidad y considera otros aspectos en la toma de decisiones. La Figura 110, muestra que los estudiantes solo señalan una cadena y no justifican su respuesta, al ver la respuesta que dan se deduce que no identifica la variabilidad de los conjuntos. Finalmente elige una cadena de cines por la variedad que tiene el cine.


¿En cuál de las tres cadenas de cines hay mayor variabilidad en el tiempo de espera? ¿Por qué? cinépolis

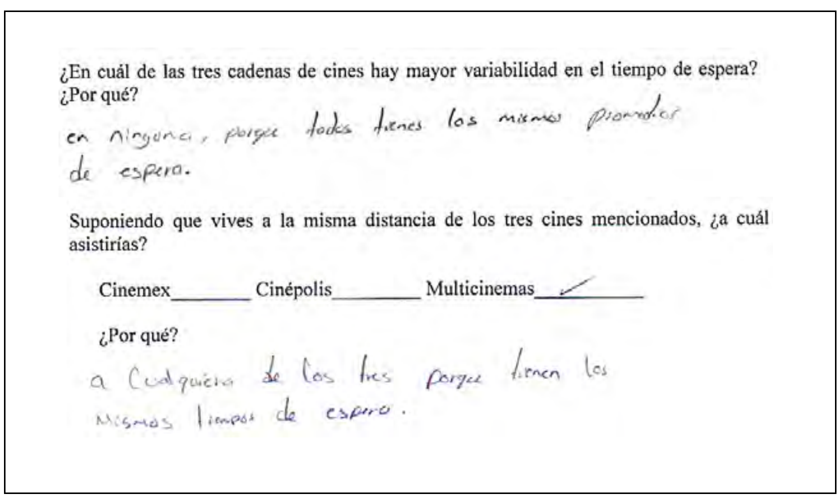
Suponiendo que vives a la misma distancia de los tres cines mencionados, ¿a cuál asistirías?

Cinemex______ Cinépolis______ Multicinemas asistiria

¿Por qué? porque son baria dos que los otras cines

Figura 110. Ejemplo de respuesta C2.

C3. Error al considerar el promedio como medida de dispersión. La Figura 111, muestra el ejemplo de la categoría en la que los estudiantes consideran al promedio como una medida de dispersión por lo que no encuentran variabilidad en los tres conjuntos dado que señalan que el promedio es el mismo. Finalmente, eligen una cadena de cines que no tiene la menor variabilidad y expresan que los tiempos de espera son los mismos, por lo que, elegirían cualquier cine.


¿En cuál de las tres cadenas de cines hay mayor variabilidad en el tiempo de espera? ¿Por qué?

en ninguna, porque todos tienes los mismos promedios de espera.

Suponiendo que vives a la misma distancia de los tres cines mencionados, ¿a cuál asistirías?

Cinemex______ Cinépolis______ Multicinemas ✓

¿Por qué?

a Cualquiera de los tres porque tienen los mismos tiempos de espera.

Figura 111. Ejemplo de respuesta C3.

En seguida se presenta la Tabla 21, que contiene la distribución de frecuencias de la variedad de categorías que se han encontrado al analizar las respuestas de los estudiantes.

Tabla 21. Frecuencias de las categorías de la tercera parte del problema

Categorías de respuestas de la tercera parte	Frecuencia
C1.1. Identifica la dispersión en la distribución y la considera en la toma de decisiones.	4
C1.2. Nivel de dispersión de los datos en una distribución, pero no lo considera en la toma de decisiones	2
C1.3. Dispersión de los datos descrita con base en un máximo, pero no la considera en la toma de decisiones	1
C1.4. Separación de los datos con base en el rango, toma decisiones con base en la moda	1
C1.5. Identifica la dispersión en la distribución, pero no la considera para la toma de decisiones	3
C1.6. Dispersión del conjunto de datos con base en la media, pero no la consideran en la toma de decisiones	1
C.2. No identifica la variabilidad y considera otros aspectos en la toma de decisiones	3
C3. Error al considerar el promedio como medida de dispersión	2
Total	**17**

Las categorías presentadas en la Tabla 21, muestran que 12 estudiantes de 17 fueron capaces de identificar la dispersión al comparar los tres conjuntos de datos, sin embargo, hay variedad en estas categorías en las que: 4 estudiantes usan la idea de dispersión cuando hacen referencia a la separación de los dados, además se dan cuenta que un conjunto es menos disperso y que es el que deben elegir, se puede deducir que de forma implícita considera la incertidumbre en los tiempos de espera de las otras dos cadenas de cine. Seguidamente, 2 alumnos consideran el nivel de dispersión de los datos en una distribución, pues señalan que comparando los tres conjuntos de datos uno tiene información más dispersa y los dos restantes menos dispersa, por lo que se puede pensar que aplican la propiedad de la desviación estándar (aunque no la mencionan), estos alumnos finalmente no consideran la dispersión identificada para tomar decisiones pues toman en cuenta otros aspectos como la calidad de la cadena de cine. Por otro lado, un estudiante señala la dispersión de los datos a partir de identificar un valor máximo en un conjunto de datos, no obstante, a pesar de haber elegido el conjunto con menos variabilidad basa su toma de decisiones en otros aspectos no relevantes para el problema. En el mismo sentido, un estudiante identifica la variabilidad considerando al rango como medida de dispersión, hace alusión a los valores máximo y mínimo del conjunto de datos y a partir de ello determina la lejanía de los datos lo que lo hace deducir que es el conjunto más disperso en comparación con los 2 restantes, pero para elegir al conjunto menos disperso considera el valor con mayor frecuencia haciendo uso de manera implícita de la moda como medida de tendencia central. En el mismo orden de ideas, 3 estudiantes de forma explícita usan la idea de la variabilidad para justificar al conjunto de datos más disperso, no obstante al tomar una decisión eligen ese mismo conjunto por el hecho de tener mayor dispersión por lo que su elección es incorrecta, debido a que tendrá mayor incertidumbre en los tiempos de espera. Una estudiante fue capaz de identificar la dispersión del conjunto de datos respecto a la media, pero al tomar la decisión es incorrecta porque señala que la media es la misma y que le daría igual ir a cualquier cine. Por otro lado, 3 estudiantes no identifican la variabilidad y considera otros aspectos en la toma de decisiones, finalmente dos alumnos cometen el error al considerar el promedio como medida de dispersión.

CAPÍTULO 6
CONCLUSIONES

6.1. Introducción

En este capítulo se resumen los resultados obtenidos a partir del diseño, la implementación y la evaluación de una experiencia de aprendizaje con estudiantes de Telebachillerato para la comprensión de las medidas de tendencia central y de variabilidad. El marco teórico que se ha utilizado, permitió analizar el currículo para este nivel y el material audiovisual (video), que a menudo se usa para la enseñanza de estos conceptos y que corresponde al significado de referencia. Seguidamente, se analizaron las respuestas que dieron los estudiantes para conocer su comprensión a partir del significado institucional que se evalúa con los problemas propuestos y se obtuvo el significado personal del antes y después de la experiencia de aprendizaje que fue deducido de las respuestas dadas.

6.2. Conclusiones del análisis del material audiovisual (significado de referencia)

El análisis del material audiovisual tuvo como objetivo determinar el contenido matemático en que se base el programa de probabilidad y estadística, para identificar el significado institucional que le dan a las medidas de tendencia central y de variabilidad, con lo que se pretendía identificas el significado personal que los estudiantes tienen sobre estos conceptos.

Con el análisis realizado, se encontró que la enseñanza es muy básica en cuestión de información matemática, se presentan las medidas de tendencia central (media, mediana y moda) y las medidas de variabilidad (rango, varianza y desviación estándar) a partir de dos ejemplos la primera una investigación por parte de dos estudiantes, sobre los tipos de música que escuchan las personas y la segunda de un equipo de futbol, de cierta forma, los ejemplos buscan analizar los diferentes tipos de variables aunque no las mencionan verbalmente se perciben de forma implícita.

En términos de contenido matemático, el video tiene diferentes errores de definiciones, cálculos aritméticos, lenguaje verbal, estadístico y algebraico. Por ejemplo: pide estimar la media de un conjunto de datos agrupados en intervalos de clase, pero en ningún momento presenta la fórmula para poder llegar a este resultado y el problema se queda sin resolver y de forma incompleto el ejemplo, lo cual se considera que puede generar muchas dudas a los estudiantes, esto mismo se repite al solicitar el cálculo de las demás medidas descriptivas. Por otro lado, al realizar el ejemplo de calcular la varianza de los pesos en kilogramos de los integrantes del equipo de futbol, no elevan las diferencias al cuadrado respecto a la media, el ejemplo no es muy claro en su procedimiento y al final el resultado que obtienen es incorrecto, por ende la desviación estándar calculada es incorrecta. En cuanto al lenguaje, usa expresiones para definir verbalmente los conceptos implicados, propiedades, entre otros aspectos, un error encontrado es como expresan la varianza pues señalan *"se calcula usando la formula sumatoria de x menos el promedio al cuadrado entre N-1"*, por lo que no se logra comprender cuál es ese procedimiento que indican. En cuanto a las propiedades, se pueden identificar de forma implícita en sus ejemplos, pues no es declarada verbalmente o ni por medio de imágenes.

Los resultados obtenidos a partir de este análisis, permitieron que la información que se les proporcionaría a los estudiantes fuera modificada y complementada con otros libros de texto, de tal forma que se pudiera tener un significado referencial adecuado y más completo.

6.3. Conclusiones respecto a los objetivos (significados personales)

El objetivo fundamental de este trabajo fue *implementar, en un ambiente con tecnología audiovisual, tareas de medidas de tendencia central y de variabilidad dirigidas a estudiantes de Telebachillerato, en un marco de experiencia de aprendizaje que permita conocer sus dificultades y errores cuando se les pide realizar actividades relacionadas con estos conceptos, en que es necesario reconocer su cálculo y su interpretación*.

Es decir, a partir de ello se buscaba identificar, a través de las respuestas que dan a los problemas, los *significados personales* que los estudiantes asignan a las medidas de tendencia central y de dispersión, así como sus correspondientes representaciones tabulares y gráficas.

El análisis de las respuestas permite describir los elementos de significado utilizados en cada problema e incisos, al resolverlos antes y después de la experiencia de aprendizaje, de tal forma que se identifique la comprensión y los conflictos semióticos cuando el significado que le da el estudiante no es el correcto. En lo que sigue se presentan y discuten los resultados obtenidos.

Respecto al problema 1 (descrito en la sección 5.2), que tiene la finalidad que los estudiantes expresen con sus propias palabras la definición de la media y que dada una media establezcan una distribución se encontró que, antes de la experiencia de aprendizaje, los estudiantes contaban con conocimientos básicos pero deficientes en estadística, esto fue observable a partir del lenguaje que utilizaban en sus respuestas, sin embargo un cambio fue, que los estudiantes solo escribían el valor numérico y no justificaban con algoritmos el resultado que daban antes de la experiencia, por lo que se tomó en consideración para el diseño de la experiencia puntualizar el lenguaje aritmético en este apartado principalmente decimales y porcentajes, después de la experiencia todas sus respuestas están justificadas pues se pudo ver los algoritmos y procedimientos empleados para poder llegar a la solución, además que agregan argumentos deductivos para mayor soporte a su respuesta. En cuanto al contenido conceptual, pocos alumnos (5) fueron capaces de explicar el significado de la media a partir de leer una nota periodística o en un contexto, en contraparte 6, alumnos todavía no reconocen la propiedad de la media de no ser operación interna e interpretan de forma incorrecta los valores decimales en variables discretas. En este mismo sentido, tres alumnos al buscar expresar el significado de la media utilizaron palabras que hacen alusión al uso de la moda, por lo que se deduce que no concibieron la idea de la media desde el lenguaje terminológico; hicieron alusión a "la mayoría" para explicar la media y esta palabra nos indica la moda, por lo que se concluye que los estudiantes confunden la media con moda.

En la segunda parte del primer problema, antes de la experiencia de aprendizaje se encontró que era necesario retomar el concepto de distribución, reparto, reparto equitativo, y que además puedan usar los algoritmos para el cálculo de la media y también el algoritmo inverso, del mismo modo se vio un lenguaje aritmético deficiente, por lo que se decidió que durante las sesiones de enseñanza se hiciera hincapié del uso y la importancia de los símbolos al hacer un cálculo de cualquier operación aritmética. Después de la experiencia de aprendizaje los resultados mostraron mayor variedad de categorías, las respuestas correctas aumentaron de 10 a 14 estudiantes, unas consideradas correctas o parcialmente correctas.

En cuanto al problema 2 (descrito en la sección 5.3), cuyo fin es el cálculo de la media ponderada y reconocer la propiedad de que *"la media de la suma de dos o más variables es igual a la suma de las medias de éstas"*, en los resultados obtenidos antes de la experiencia de aprendizaje se observó que solo 2 estudiantes fueron capaces de dar respuesta al problema, uno lo realizo de la forma correcta y el otro solo escribió el valor sin justificación algorítmica, por otro lado, la mayoría dio respuestas que no tenían relación, por lo que se decidió que en los episodios de enseñanza se desarrollaran problemas con esta naturaleza. Después de la experiencia de aprendizaje, la variedad de categoría de respuestas aumento, a 11 estudiantes, con diferencias en sus respuestas, sin embargo, la mayoría, 6 estudiantes, no usan la ponderación para calcular la media, sino que extraen los valores y aplican el cálculo de la media para datos aislados y no como se pretendía en el problema. Los errores reportados en este inciso son que un estudiante sigue teniendo ese conflicto al confundir la media con la mediana, por lo que no concibe la terminología de estas medidas de tendencia central.

El problema 3 (descrito en la sección 5.4), tenía el objetivo de observar si el alumno es capaz de calcular la media y la mediana a partir de un gráfico y que además incluye las propiedades: "*la mediana y media pueden no coincidir con los datos"* y *"el cálculo de la media y el de la mediana no son operaciones internas"*. Al comparar los resultados de la comprensión antes y después de la experiencia de aprendizaje se obtuvo que, antes de la experiencia de aprendizaje los estudiantes tenían nociones de los conceptos de media y mediana, y que en algunas ocasiones hacen una lectura literal del gráfico, según Curcio (1989). Es importante destacar que la mayoría de las respuestas dadas en esta etapa no las justifican con algoritmos de cálculo como los demás problemas, es decir, los estudiantes sólo dan un valor numérico como respuesta. En cuanto a las categorías, se pudieron identificar seis donde se clasificaron las respuestas obtenidas.

Respecto al análisis después de la experiencia de aprendizaje, se encontró que la mayoría de los estudiantes usó los algoritmos de cálculo para justificar sus respuestas, y además dan argumentos deductivos. No obstante, persisten algunos problemas al interpretar los gráficos. Por otro lado, se identificaron más variedad de respuestas ocho categorías.

Los resultados muestran, que en el primer acercamiento, esto es, antes de la experiencia de aprendizaje únicamente 9 estudiantes usaron el concepto de media, pero, sólo un alumno realizó justificación con algoritmos. Por el contrario, después de la experiencia de aprendizaje, 15 alumnos usaron el concepto de la media casi el doble de los estudiantes, por lo que se puede observar un avance significativo.

El segundo apartado del problema, resulto ser más difícil para los estudiantes, se observó que sólo un estudiante fue capaz de obtener la mediana, pero sin ordenar los datos que es la idea principal al momento de la estimación. Al mismo tiempo, se encontraron varias confusiones en los estudiantes sobre las medidas de tendencia central.

Un aporte importante de este trabajo, es que el alumno además de confundir el valor de la variable con la escala, toma en cuenta el cero como valor y con estos datos calcula la mediana, el error que comete es la lectura incorrecta del gráfico, y que debe ser considerada en futuras investigaciones.

Después de la aplicación de la experiencia de aprendizaje, 13 estudiantes fueron capaces de calcular la mediana a partir de un gráfico, por lo que, se puede decir que se ha generado una mejor comprensión en el concepto de la mediana. Otro punto importante, es mencionar que la categoría *"confunde la media con la mediana"* se repite en el antes y el después, y con la misma frecuencia.

Los tipos de respuestas encontradas antes y después de la experiencia de aprendizaje en esta segunda parte del problema es muy variada, pero se han encontrado resultados positivos en la aplicación después de la experiencia.

Este trabajo señala que la estimación de la media y mediana a partir de un gráfico resulta ser difícil para los estudiantes. Por lo que, al desarrollar una experiencia de aprendizaje pertinente, ayuda a los estudiantes en la interpretación de gráficos y cálculo de dichas medidas. Los resultados han sido positivos, ya que los estudiantes reconocen la utilización de la media y la mediana, pero además de eso realizan el algoritmo de cálculo.

El problema 4 (descrito en la sección 5.5), pretendía indagar sobre las ideas de los estudiantes acerca de la variabilidad, para lo que se necesitaba que los estudiantes construyeran gráficos, calcularan el promedio de los tres conjuntos, y finalmente identifiquen la dispersión de los datos, considerando el rango y/o la desviación estándar para tomar un decisión ante la incertidumbre. Los resultados obtenidos muestran que sólo 2 jóvenes fueron capaces de construir un gráfico de forma correcta, y el resto de la muestra aún tiene errores al momento de la construcción, este aspecto se consideró primordial, porque toda vez que los estudiantes sepan construir e interpretar un gráfico será de mucha utilidad para la identificación de los conjuntos de datos más o menos dispersos. La mayoría de los resultados ya fueron reportados por Arteaga (2011). En la segunda parte del problema, se observó que los estudiantes, 13 de 17, calcularon de forma correcta la media aritmética de cada conjunto, sólo 4 estudiantes cometieron errores; 3 aritméticos al calcularla la media y uno de terminología, pues cometió un error al concebir la idea de media ponderada, las medidas de tendencia central eran un aspecto importante para la comparación de los tres conjuntos e identificar la separación de los datos respecto a un centro.

Los resultados obtenidos en la tercera parte del problema, permitieron conocer que 12 de 17 alumnos fueron capaces de identificar la dispersión de un conjunto de datos, comparar entre tres conjuntos y señalar el conjunto más disperso, se pudo ver que los estudiantes a menudo usan de manera implícita los conceptos de rango y desviación estándar esto observado a partir de sus argumentos, es decir, los jóvenes tienen nociones de la variabilidad. De los 12 jóvenes que identificaron la variabilidad en los datos, solo 4 tomaron una decisión a partir de la incertidumbre que causaba tener un conjunto de datos con mayor dispersión, por el contrario, 8 estudiantes consideraron otros aspectos como la calidad y no se centraron en la variabilidad de los datos para elegir al conjunto menos disperso que era el que daba menos incertidumbre en la espera. Finalmente, 5 estudiantes no fueron capaces de concebir la idea de variabilidad, de los cuales 2 confunden las medidas de tendencia central con las medidas de dispersión.

En resumen, los resultados obtenidos al implementar esta experiencia de aprendizaje fueron positivos, pues se observan a partir de ello, respuestas más estructuradas, donde es fácil identificar el uso de los objetos matemáticos (elementos de significado) y que se pretende que el alumno pueda resolver en un futuro problemas similares, en el que implique análisis de gráficos, medidas de tendencia central y de dispersión.

REFERENCIAS

Arteaga, J.P. (2011). *Evaluación de conocimientos sobre gráficos estadísticos y conocimientos didácticos de futuros profesores* (Tesis doctoral). Universidad de Granada. España.

Bakker, A. (2004). Reasoning about shape as a pattern in variability. *Statistics Education Research Journal, 3*(2), 64-83.

Barr, G. V. (1980). Some student's ideas on the median and the mode. *Teaching Statistics, 2*, 38-41.

Batanero, C. (2000). Significado y comprensión de las medidas de posición central. *UNO. Revista de Didáctica de las Matemáticas, 25*, 41-58

Batanero, C., Godino, J. D. y Navas, F. (1997). Evaluación de concepciones sobre la noción de promedio en maestros de primaria en formación. Implicaciones para la formación estadística de los futuros profesores. En H. Salmerón (Ed.). *Actas VII Jornadas LOGSE: Evaluación Educativa,* 301-304. Universidad de Granada.

Ben-Zvi, D. (2004). Reasoning about variability in comparing distributions. *Statistics Education Research Journal 3*(2), 42-63.

Bruno, A. y Espinel, M. C. (2005). Recta numérica, escalas y gráficas estadísticas: un estudio con estudiantes para profesores. *Formación del Profesorado e Investigación en Educación Matemáticas VII*, 57-85.

Cai, J. (1995). Beyond the computational algorithm. Students' understanding of the arithmetic average concept. En L. Meira (Ed.). *Proceeding of the 19th PME Conference, 3*, 144-151. Universidade Federal de Pernambuco, Recife, Brasil.

Campbell, S. K. (1974). *Flaws and fallacies in statistical thinking*. New Jersey: Prentice Hall.

Carvalho, C. (1996). Algumas questoes em torno de tarefas estatísticas com alunos do 7º ano. *Actas do ProfMat, 96,* 165-171. Almada: Associaçao de Profesores de Matemática.

Carvalho, C. (1998). Tarefas estadísticas e estratégias de resposta. *VI Encuentro en Educación Matemática de la Sociedad Portuguesa de Ciencias de la Educación*. Castelo de Vide, Portugal.

Carvalho, C. (2001). Interaçao entre pares. *Contributos para a promoçao do desenvolvimiento lógico e do desempenho estatístico no 7º ano de escolaridade* (Tesis Doctoral). Universidad de Lisboa.

Castro, E. y Molina, M. (2007). *Desarrollo de Pensamiento Relacional mediante trabajo con igualdades numéricas en aritmética básica. Educación Matemática, 19*(2), 67-94.

Cobb, P. y Gravemeijer, K. (2008). Experimenting to support and understand learning processes. En A.E. Kelly, R.A. Lesh y J. Y. Baek (Eds.), *Handbook of design research methods in education. Innovations in Science, Technology, Engineering and Mathematics Learning and Teaching,* 68-95. Mahwah, NJ: Lawrence Erlbaum Associates.

Cobo, B. (2003). *Significado de las medidas de posición central para estudiantes de secundaria* (Tesis doctoral). Universidad de Granada.

Cobo, B. y Batanero, C. (2000). La mediana en la educación secundaria obligatoria: ¿un concepto sencillo?. *UNO. Revista de Didáctica de las Matemáticas, 23*, 85-96.

Confrey, J. (2006). The evolution of design studies as methodology. En R. Keith Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences,* 135-152. New York, ny: cambridge university press.

Curcio, F. R. (1989). *Developing graph comprehension*. Reston, VA: N.C.T.M.

delMas, R. C., & Liu, Y. (2007). Students' conceptual understanding of the standard deviation. In M. C. Lovett & P. Shah (eds.), *Thinking with data* 87-116. New York: Lawrence Erlbaum.

delMas, R., & Liu, Y. (2005). Exploring students' conceptions of the standard deviation. *Statistics Education Research Journal, 4*(1), 55-82. Recuperado de www.stat.auckland.ac.nz/~iase/serj/SERJ4(1)_delMas_Liu.pdf

Durazo, A. (2012). *Probabilidad y estadística 1*. México: Colegio de Bachilleres del Estado de Sonora.

Eisenbach, R. (1994). What does de mean mean?. *Fourth International Conference on Teaching Statistics.* Marrakesh, Marruecos.

Espinel, C. (2007). Construcción y razonamiento de gráficos estadísticos en la formación de profesores. *Investigación en Educación Matemática 11,* 99-119.

Estepa, A. (1993). *Concepciones iniciales sobre la asociación estadística y su evolución como consecuencia de una enseñanza basada en el uso de ordenadores* (Tesis Doctoral). Universidad de Granada.

Estrada, A. (2002). *Análisis de las actitudes y conocimientos estadísticos elementales en la formación del profesorado* (Tesis Doctoral). Universidad Autónoma de Barcelona.

Estrada, M., y Hernández, X. (2015). *Probabilidad y estadística I*. México: FLACSO

Estrella, S. (2016).Comprensión de la media por profesores de educación primaria en formación continua. *Revista Electrónica de Investigación Educativa, 18*(1), 1-22. Recuperado de http://redie.uabc.mx/redie/article/view/635

Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making sense of graphs: Critical factors influencing graph comprehension and instructional implications. *Journal for Research in Mathematics Education*.

García, C. y Garret, A. (2006). On average and open-ended questions. En A. Rossman y B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching*

Statistics. Salvador (Bahia), Brasil: International Association for Statistical Education. Recuperado de www.stat.auckland.ac.nz/~iase/publications/.

Garret, A. y García, J. A. (2005). Un cuestionario y estrategias sobre los promedios. *Formación del Profesorado e Investigación en Educación Matemática, 7*, 197-217.

Gattuso, L. y Mary, C. (1996). Development of concepts of the arithmetic average from high school to University. *Proceedings of the 20th Conference of the International Group for the Psychology of Mathematics Education, 1*, 401-408. Universidad de Valencia.

Gattuso, L. y Mary, C. (1998). Development of the concept of weighted average among highschool students. En L. Pereira-Mendoza, C. Seu Keu, T. Wee Kee y W.K. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching Statistics*, 685-691. Singapur: International Association for Statistical Education.

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. (1999) Análisis epistémico, semiótico y didáctico de procesos de instrucción matemática. Trabajo presentado en el grupo de trabajo “La didáctica de la matemática como disciplina científica” en el *III Simposio de la Sociedad Española de Investigación en Educación Matemática (SEIEM)*. Valladolid, España.

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. (2003). *Teoría de las funciones semióticas. Un enfoque ontológico-semiótico de la cognición e instrucción matemática.* Departamento de Didáctica de la Matemática. Universidad de Granada. Disponible en Internet: URL: http://www.ugr.es/local/jgodino/indice_tfs.htm.

Godino, J. D. y Batanero, C. (1994). Significado institucional y personal de los objetos matemáticos. *Recherches en Didactique des Mathématiques,14 (3),* 325-355.

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 Roa, R. (2005). An onto-semiotic analysis of combinatorial problems and the solving processes by university students. *Educational Studies in Mathematics, 60* (1), 3-36.

Goodchild, S. (1988). School pupils'unsderstanding of average. *Teaching Statistics, 10*(3), 77-81.

Gutiérrez, A. (2012). *Probabilidad y estadística. Enfoque por competencias*. México: McGRAW-HILL

Holmes, P. (1980). Teaching Statistics 11-16. *Sloug: Foulsham Educational*. U.K.

Jacobbe, T. (2008). Elementary school teachers' understanding of the mean and median. En C. Batanero, G. Burrill, C. Reading y A. Rossman (Eds.), Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education. *Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference.* Monterrey: ICMI and IASE. Recuperado de www.ugr.es/~icmi/iase_study/.

Kelly, A. y Lesh, R. (2000). *Handbook of research design in mathematics and science education*. New Jersey, NJ: Lawrence Erlbaum Associates.

Kelly, B. A., y Watson, J. M. (2002). Variation in a chance sampling setting: The lollies task. In B. Barton, K. C. Irwin, M. Pfannkuch, y M. O. J. Thomas (Eds.), Mathematics education in the South Pacific. *Proceedings of the 26th anual Conference of the Mathematics Education Research Group of Australasia, Auckland, NZ, 2*, 366-373. Sydney: MERGA.

Lehrer, R., & Schauble, L. (Eds.). (2002). *Investigating real data in the classroom: Expanding children's understanding of math and science*. New York: Teachers College Press.

Lehrer, R., Kim, M., & Schauble, L. (2007). Supporting the development of conceptions of 10 statistics by engaging students in modeling and measuring variability. *International Journal of Computers for Mathematics Learning*, 12, 195-216.

León, M. R., y Zawokeswski, F. S. (1991). Use of the arithemetic mean: An investigation of four properties. Issues and preliminary resuls. En D. Vere-Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics*, 302-306. Voorburg, Holanda: International Statistical Institute. Recuperado de www.stat.auckland.ac.nz/~iase/publications/.

León, O. G. y Montero, I. (2002). *Métodos de investigación en psicología y educación.* Madrid: McGraw-Hill.

Levin, R. y Rubin, D. (2004). *Estadística para administración y economía*. México; Pearson.

Li, D. Y. y Shen, S. M. (1992). Students'weaknesses in statistical projects. *Teaching Statistics, 14*(1), 2-8.

Lind, D., Marchal, W. y Mason, R. (2004). *Estadística para administración y economía*. México: Alfaomega.

Makar, K. y Confrey, J. (2005). "Variation-talk": articulating meaning in statistics. *Statistics Education Research Journal, 4*(1), 27-54.

Mary, C. y Gattuso, L. (2005). Trois problèmes semblables de moyenne pas si semblables que ça ! L'influence de la structure d'un problème sur les résponses des eleves. *Statistics Education Research Journal, 4*(2), 82-102.

Mayén, S. (2009). *Comprensión de las medidas de tendencia central por estudiantes mexicanos de Educación Secundaria y Bachillerato* (Tesis doctoral). Universidad de Granada. España.

Mevarech, Z. R. (1983). A deep structure model of students' statistical misconceptions. *Educational Studies in Mathematics, 14*, 415-429.

Miles, MB. y Huberman, AM. (1994). *Qualitative Data Analysis (2nd edition).* Thousand Oaks, CA: Sage Publications.

Mokros, J. R. y Russell, S. J. (1995). Children's concepts of averages and representativeness. *Journal for Research in Mathematics Education, 26*(1), 20-39.

Molina, M. y Ambrose, R. (2008). *From an operational to a relational conception of the equal sign. Thirds graders' developing algebraic thinking. Focus on Learning Problems in Mathematics, 30*(1), 61-80.}

Molina, M., Castro, E. y Castro. E. (2009). Elementary students´ understanding of the equal sign in number sentences. *Electronic Journal of Research in Educational Psychology, 17, 7*(1), 341-368.

Molina, M., Castro, E., Molina, J. L., & Castro, E. (2011). Un acercamiento a la investigación de diseño a través de los experimentos de enseñanza. En *Enseñanza de las Ciencias, 29(1),* 75-88.

Moore, D. (1990). Uncertainty. En L. A. Steen (Ed.), *On the shoulders of giants: New approaches to numeracy*, 95–137. Washington, DC: National Academy Press.

NCTM (2000). *Principles and Standards for School Mathematics.* Reston, VA: NCTM.

Orta, J.A. (2014). *Estudio exploratorio sobre la noción de variabilidad estadística asociada al riesgo* (Tesis doctoral). Departamento de Matemática Educativa, Centro de Investigación y de Estudios Avanzados del IPN. México.

Petrosino, A.J, Lehrer, R., y Schauble, L. (2003). Error de estructura y variación experimental como distribución en cuarto grado. *Pensamiento matemático y aprendizaje, 5* (2 y 3), 131-156.

Pinho, M. (2006a). Semiotics function and learning of arithmetics mean. En A. Rossman y B. Chance (Eds.). *Proceedings of the Seventh International Conference on Teaching*

Statistics. Salvador (Bahia), Brasil: International Association for Statistical Education. Recuperado de www.stat.auckland.ac.nz/~iase/publications/.

Pinho, M. (2006b). Interpreting the concept of arithmetic mean from a semiotic vision. En A. Rossman y B. Chance (Eds.). *Proceedings of the Seventh International Conference on Teaching Statistics*. Salvador (Bahia), Brasil: International Association for Statistical Education. Recuperado de: www.stat.auckland.ac.nz/~iase/publications/

Pinzón, L. (2012). *Propuesta didáctica para el aprendizaje de la media aritmética, la mediana y la moda, para estudiantes del programa de psicología* (Tesis de maestría), Universidad Nacional de Colombia.

Pollatsek, A., Lima, S. y Well, A. D. (1981). Concept or computation: Students'understanding of the mean. *Educational Studies in Mathematics, 12*, 191-204.

Rodríguez, F., Maldonado, A. y Sandoval, P. (2016). Comprensión de las medidas de tendencia central: un estudio comparativo en estudiantes de pedagogía en matemática en dos instituciones formadoras chilenas. *Avaliação, Campinas; Sorocaba, SP, 21*(3), 929-952. Doi: http://dx.doi.org/10.1590/S1414-40772016000300013

Rubiales, F.S. (2011). *Niveles de razonamiento del concepto empírico de distribución binomial de estudiantes de bachillerato frente a tareas de simulación física y computacional* (Tesis de maestría). Departamento de Matemática Educativa, Centro de Investigación y de Estudios Avanzados del IPN. México.

Ruiz, B. (2006). *Un acercamiento cognitivo y epistemológico a la didáctica del concepto de variable aleatoria*. Tesis de Maestría. CICATA. México.

Russell, S. J. y Mokros, J. R. (1991). What's typical?: children's ideas about average. En D. Vere-Jones (Eds.) *Proceedings of the Third International Conference on Teaching Statistics,* 307-313. Voorburg, Holanda: International Statistical Institute.

Saldarriaga, A. (2012). *Propuesta didáctica fundamentada en la solución de situaciones problema utilizando las medidas de centralización y su interpretación* (Tesis de maestría). Universidad Nacional de Colombia.

Sayritupac, J. (2013). *Significados de las medidas de tendencia central. Un estudio con alumnos universitarios de carreras de Humanidades* (Tesis de maestría). Pontificia Universidad Católica del Perú.

Schuyten, G. (1991). Statistical thinking in Psychology and Education. En Vere-Jones (Eds.), *Proceedings of the Third International Conference on Teaching Statistics* 486-490. Voorburg, Holanda: International Statistical Institute. Recuperado de http://www.stat.auckland.ac.nz/~iase/publications/-

SEP. (2011). *Programas de Estudio. Educación Básica. Secundaria. Matemáticas*. México, D. F.: Secretaría de Educación Pública.

SEP. (2013). *El modelo de Telebachillerato Comunitario. Documento base para el servicio educativo de Telebachillerato Comunitario.* México, D.F.: Dirección General de Bachillerato.

SEP. (2013). *Programas de estudio. Bachillerato. Matemáticas II.* México, D.F.: Dirección General de Bachillerato.

SEP. (2013). *Programas de estudio. Bachillerato. Probabilidad y Estadística I*. México, D.F.: Dirección General de Bachillerato.

Shaughnessy, J. M., y Ciancetta, M. (2002). Students' understanding of variability in a probability environment. In B. Phillips (Ed.). *CD of the Proceedings of the Sixth International Conference on Teaching Statistics: Developing a statistically literate society, Cape Town, South Africa*. Voorburg, The Netherlands: International Statistics Institute.

Shaughnessy, J.M., Watson, J., Moritz, J., & Reading, C. (1999). School mathematics students' acknowledgement of statistical variation. *Paper presented at the research* *pre sessions of 77th annual meeting of the National Council of Teachers of Mathematics*, San Francisco.

Shaughnessy, M. J. (2006). Research on student´s understanding of some big concepts in statistics. En G. F. Burrill & P. C. Elliot, *Thinking and Reasoning with Data and Chance*, 77-99. United States of America: National Council of Teachers of Mathematics, Inc.

Silva, C. y Coutinho, C. (2008). Reasoning about variation of a univariate distribution: a study with secondary mathematics teachers. En C. Batanero, G. Burrill, C. Reading y A. Rossman (Eds.), *Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education. Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference*. Monterrey, México: International Commission on Mathematical Instruction and International Association for Statistical Education.

Steffe, L. P., Thompson, P. W., y Glasersfeld, E. V. (2000). Teaching Experiment Metodology: Underliying Principles and Essential Elements. En A. E. Kelly, y L. R. A., *Handbook of research design in mathematics and science education,* 267-306. Mahwah: NJ: LAE.

Steffe, L. y Thompson, P. (2000). Teaching experiment methodology: Underlying principles and essential elements. En Anthony Kelly y Richard Lesh (Eds.), *Handbook of research design in mathematics and science education,* 267-306. Mahwah: NJ: Lawrence Erlbaum Associates.

Strauss, S. y Bichler, E. (1988). The development of children's concepts of the arithmetic average. *Journal for Research in Mathematics Education, 19*(1), 64-80.

Tormo, C. (1993). *Estudio sobre cuatro propiedades de la media aritmética en alumnos de 12 a 15 años*. Memoria de Tercer Ciclo. Universidad de Valencia.

Torok, R., & Watson, J.M. (2000). Development of the concept of statistical variation: An exploratory study. *Mathematics Education Research Journal, 12*, 147-169.

Vermette, S., Gattuso L. y Bourdeau, M. (2005). Data analysis or how high school students "read" statistics. *Proceedings of the IASE Satellite Conference Communication of Satistics.* Nueva Zelanda: International Association for Statistical Education. Recuperado de www.stat.auckland.ac.nz/~iase/publications/.

Watson, J. M. y Moritz, J. B. (1999). The developments of concepts of average. *Focus on Learning Problems in Mathematics, 21*(4), 15-39.

Watson, J. M. y Moritz, J. B. (2000). The longitudinal development of understanding of average. *Mathematical Thinking and Learning, 2*(1 y 2), 11-50.

Watson, J., Kelly, B. A., Callingham, R. A., y Shaughnessy, J. M. (2003). The measurement of school students' understanding of statistical variation. *International Journal of Mathematical Education in Science and Technology, 34*(1), 1-29.

Watson, J.M., & Kelly, B.A. (2003). The vocabulary of statistical literacy. In Educational Research, Risks, & Dilemmas: *Proceedings of the joint conferences of the New Zealand Association for Research in Education and the Australian Association for Research in Education [~CD-ROM]*. Auckland, New Zealand.

Wild, D.J. & Pfannkuch, M. (1999), Statistical thinking in empirical enquiry. *International Statistical Review, 67*, 223-265.

Zawojewski, J. (1986). *The teaching and learning processes of junior high school students under alternative modes of instruction in the measures of central tendency* (Tesis Doctoral). University Northwestern. Evanston, Illinois.

Zawojewski, J.S. y Roth L. M. (1990) Use of the arithmetic mean: An Investigation of four properties issues and preliminary results. En D. Vere-Jones (Eds.) *Proceedings of the Third International Conference on Teaching Statistics*, 302-306. Voorburg, The Netherlands: International Statistical Institute.

ANEXOS

ANEXO 1.CUESTIONARIO PILOTO

Tomado de Mayén (2009) y aplicado antes de implementar la experiencia de aprendizaje.

CUESTIONARIO SOBRE MEDIDAS DE POSICIÓN CENTRAL

I. INSTRUCCIONES:

1. Lee con atención cada pregunta antes de contestar.
2. No dejes sin contestar ningún problema y escribe tu razonamiento en aquellos que se te indique.
3. Escribe tu nombre completo, escuela y grupo al que perteneces en la parte superior de todas las hojas que utilices.

II. DATOS PERSONALES

Escuela: ____

Grupo: _______ Edad: ___________ Sexo: ___________ Fecha: ______________

Apellidos: ____________________________ Nombre: __________________________

III. CUESTIONARIO

Item 1. Un periódico dice que el número medio de hijos por familia en México es 2.2 hijos por familia.

1. Explica qué significa para ti esta frase.
2. Se han elegido 10 familias mexicanas y el número medio de hijos entre las 10 familias es de 2.2 hijos por familia. Los García tienen 4 hijos y los Pérez tienen 1 hijo, ¿Cuántos hijos podrán tener las otras 8 familias para que la media de hijos en las 10 familias sea 2.2? Justifica tu respuesta.

Item 2. María y Pedro dedican una media de 8 horas cada fin de semana a hacer deporte. Otros 8 estudiantes dedican cada fin de semana una media de 4 horas a hacer deporte.

1. ¿Cuál es el número medio de horas que hacen deporte cada fin de semana los 10 estudiantes?
2. María y Pedro dedican además 1 hora cada fin de semana a escuchar música y los otros 8 estudiantes, 3 horas. ¿Cuál sería el número medio de horas que estos 10 estudiantes dedican cada fin de semana, entre las dos actividades: hacer deporte y escuchar música?

Item 3. Cuatro amigos se reúnen para preparar una cena. Cada uno de ellos trajo harina para hacer la masa de las pizzas. Como querían hacer cuatro pizzas del mismo tamaño, los que habían traído más harina regalaron a los que llevaban menos. ¿La cantidad de harina regalada por los que habían traído mucha fue *mayor, menor o igual* a la recibida por los que habían traído poca? ¿Por qué piensas eso?

Item 4. Tenemos *seis números* y el más grande es el 5. Sumamos estos números y dividimos la suma entre *seis*. El resultado es 4. ¿Te parece posible? ¿Por qué?

Item 5. El peso en kilos de 9 niños es 15, 25, 17, 19, 16, 26, 18, 19, 24. ¿Cuál es el peso del niño mediano? ¿Cuál es la mediana si incluimos el peso de otro niño que pesa 43 Kg? En este caso, ¿Sería la media aritmética un buen representante de los 10 datos? ¿Por qué?

Item 6. Un profesor califica a sus alumnos del siguiente modo: I=Insuficiente, A=Aprobado, N=Notable, S=Sobresaliente. En la siguiente tabla tenemos las notas que ha puesto a dos grupos de alumnos:

Grupo 1	I A A N N S S I I I A A A N S S I A A S S S S
Grupo 2	S S I I A N A N I I S N A S I N N

1. ¿Qué grupo ha obtenido mejores notas?
2. ¿Cuál sería el promedio (medida de centralización) más apropiado para representar estos datos? Explica tu respuesta.

Item 7. Lucía, Juan y Pablo van a una fiesta. Cada uno lleva cierto número de caramelos. Entre todos llevan una media de 11 caramelos por persona.

1. ¿Cuántos caramelos ha llevado cada uno?

   Lucía ________________ Juan __________________ Pablo _______________

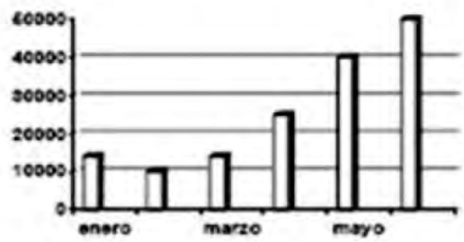
2. ¿Es la única posibilidad? Si____________No____________ Explica cómo has obtenido tus resultados.
3. Un cuarto chico llega a la fiesta y no lleva ningún caramelo. ¿Cuál es ahora la media de caramelos por chico? Explica tu resultado.

Item 8. Nueve estudiantes han pesado un objeto en la clase de ciencias, usando la misma escala. Los pesos registrados por cada estudiante (en gramos) se muestran a continuación:

6.2, 6.3, 6.0, 15.2, 6.2, 6.1, 6.5, 6.2, 6.1, 6.2.

Los estudiantes quieren determinar con la mayor precisión posible el peso real del objeto. ¿Qué harías para calcularlo?

Item 9. Observa el siguiente diagrama de barras que muestra las ventas de bocadillos de la empresa *Bocatta* durante los últimos 6 meses del año pasado:



1. Da un valor aproximado del número medio de bocadillos que se vende al mes.
2. Da un valor aproximado de la mediana del número de bocadillos que se vendieron por mes.

ANEXO 2.CUESTIONARIO FINAL

Integrado por 3 problemas tomados de Mayén (2009) y 1 problema de Orta (2009) aplicado después de implementar la experiencia de aprendizaje.

TELEBACHILERATO COMUNITARIO DEL ESTADO DE QUINTANA ROO
PLANTEL: "TRES GARANTIAS", C.C.T.:23ETK0031T
PROBABILIDAD Y ESTADISTICA I

TELEBACHILLERATO COMUNITARIO Quintana Roo

CUESTIONARIO SOBRE COMPRENSION DE LAS MEDIDAS DE TENDENCIA CENTRAL Y DE VARIABILIDAD

I. INSTRUCCIONES:

1. Lee con atención cada pregunta antes de contestar.
2. No dejes sin contestar ningún problema y escribe tu razonamiento en aquellos que se te indique.
3. Escribe tu nombre completo, escuela y grupo al que perteneces en la parte superior de todas las hojas que utilices.

II. DATOS PERSONALES

Nombre Completo: ____

Edad: ____________________ Fecha: ____________________

III. CUESTIONARIO

Ítem 1. Un periódico dice que el número medio de hijos por familia en México es 2.2 hijos por familia.

1.1. Explica qué significa para ti esta frase.

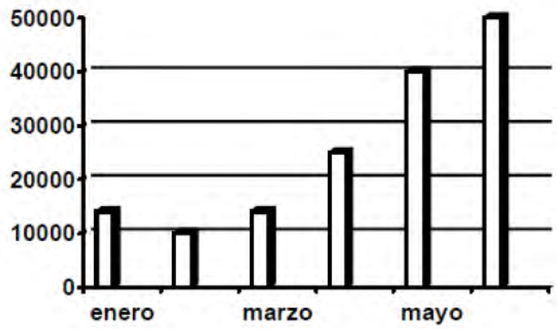
1.2. Se han elegido 10 familias mexicanas y el número medio de hijos entre las 10 familias es de 2.2 hijos por familia. Los García tienen 4 hijos y los Pérez tienen 1 hijo, ¿Cuántos hijos podrán tener las otras 8 familias para que la media de hijos en las 10 familias sea 2.2? Justifica tu respuesta.

Ítem 2. María y Pedro dedican una media de 8 horas cada fin de semana a hacer deporte. Otros 8 estudiantes dedican cada fin de semana una media de 4 horas a hacer deporte.

1.1. ¿Cuál es el número medio de horas que hacen deporte cada fin de semana los 10 estudiantes?

1.2. María y Pedro dedican además 1 hora cada fin de semana a escuchar música y los otros 8 estudiantes, 3 horas. ¿Cuál sería el número medio de horas que estos 10 estudiantes dedican cada fin de semana, entre las dos actividades: hacer deporte y escuchar música?

Ítem 3. Observa el siguiente diagrama de barras que muestra las ventas de bocadillos de la empresa *Bocatta* durante los últimos 6 meses del año pasado:



1.1. Da un valor aproximado del número medio de bocadillos que se vende al mes.
1.2. Da un valor aproximado de la mediana del número de bocadillos que se vendieron por mes.

Ítem 4. Una costumbre en los cines es mostrar anuncios comerciales y cortos en la pantalla antes de comenzar la película. El tiempo de espera para una película es la diferencia entre la hora de inicio anunciada y la hora real en la que comienza la película. Un grupo de 10 estudiantes investigó el tiempo de espera de tres cadenas de cines, una de la cadena Cinemex, otra de la cadena Cinépolis y por último una de la cadena Multicinemas. Los estudiantes visitaron los cines en horarios y días diferentes. Cada alumno midió el tiempo de espera en minutos, y lo registró en las tablas que siguen:

Cinemex
12.0
21.0
15.0
15.0
13.0
16.0
16.0
16.0
20.0
18.0

Cinépolis
15.0
15.5
16.0
16.0
16.5
16.5
18.0
16.0
15.5
17.0

Multicinemas
15.5
17.0
18.0
16.5
16.0
16.5
16.5
15.0
15.0
16.0

a) En la hoja anexa traza una gráfica para cada tabla.

b) Calcula el tiempo de espera promedio para cada cadena de cines y señálalo con un segmento de recta en la gráfica que construiste.

Observa con atención las gráficas de la página siguiente.

a) Las gráficas que construiste son iguales o diferentes a las que vienen junto con este cuestionario.

   Iguales______ Diferentes______

   En caso de que no sean iguales ¿cuál es la diferencia entre ellas?

   El tiempo promedio de espera es de 16.2 minutos en las tres cadenas de cines.

   b) ¿En cuál de las tres cadenas de cines hay mayor variabilidad en el tiempo de espera? ¿Por qué?

      En cuál de las dos cadenas de cines, Cinépolis o Multicinemas, hay mayor variabilidad.

      Cinépolis______ Multicinemas_______

      Explica tu respuesta.

   c) Suponiendo que vives a la misma distancia de los tres cines mencionados, ¿a cuál asistirías?

      Cinemex________ Cinépolis_________ Multicinemas___________

      ¿Por qué?

   d) ¿Crees que es importante la variabilidad en tiempos de espera que hay en los cines mencionados?

Gráficas del problema de los tiempos de espera en los cines