CONFLICTOS SEMIÓTICOS DE ESTUDIANTES CON EL CONCEPTO DE MEDIANA
STUDENTS' SEMIOTIC CONFLICTS IN THE CONCEPT OF MEDIAN

SILVIA MAYÉN
Instituto Politécnico Nacional, México
smayen@correo.ugr.es

CARMEN DÍAZ
Universidad de Huelva, España
carmen.diaz@dpsi.uhu.es

CARMEN BATANERO
Universidad de Granada, España
batanero@ugr.es

RESUMEN

El foco de este trabajo es el concepto de mediana, sobre el que la investigación previa es escasa. Analizamos las respuestas abiertas de 518 estudiantes mexicanos de Educación Secundaria y Bachillerato a un problema abierto de cálculo de la mediana. Utilizando las ideas del enfoque onto-semiótico, clasificamos las respuestas según la medida de tendencia central utilizada y describimos los conflictos semióticos detectados. Mediante el test chi-cuadrado estudiamos la relación entre el tipo de respuesta y el grupo de estudiantes. Observamos mejores resultados entre los alumnos de Secundaria en el cálculo y mayor capacidad para elegir la medida de tendencia central más representativa en los alumnos de Bachillerato.

Palabras clave: *Investigación en estadística educativa, Comprensión, Enfoque onto-semiótico, Medidas de tendencia central, Educación Secundaria.*

ABSTRACT

The focus of this research is the concept of median, which has received scarce interest in previous research. We analyse the open responses given by 518 Mexican students from Educación Secundaria (Junior) and Bachillerato (Senior) Secondary Education to a problem involving the computation of median. Using some ideas from the onto-semiotic approach, we classify the responses, taking into account the central tendency measure used, and describe the students' semiotic conflicts. We use the chi-square test to study possible dependence between responses and students' group. We observe better results in computation in Educación Secundaria students but better competence to select the best representative value in Bachillerato students.

Keywords: *Statistics education research, Understanding, Onto-semiotic approach, Measures of center, Secondary education.*

Statistics Education Research Journal, 8(2), 74-93, http://www.stat.auckland.ac.nz/serj

EXTENDED SUMMARY

In this paper we focus on the concept of median, which is fundamental in exploratory data analysis, and is now introduced in Mexico in Educación Secundaria (Junior; 13-15 year-olds) and Bachillerato (Senior; 16-18 year-olds) secondary school. In the first section we introduce the aims of the paper. These aims are (a) to analyse the students' understanding of median and the semiotic conflicts related to this concept, and (b) to assess the students' competence to decide whether the mean or median is preferable to represent a small data set. We also want to compare the two groups of students.

In the second section we summarize the theoretical framework and describe previous research related to the understanding of median. In the onto-semiotic approach (Godino & Batanero, 1997; Godino, Batanero, & Font, 2007; Godino, Batanero, & Roa, 2005) the meaning of mathematical objects emerges from the practices carried out by a person when solving problems related to that object (median in our research). These practices involve different types of objects: concepts (such as median, distribution, variable), properties (for example, the median is resistant to outliers), language (symbols, graphs…), procedures (e.g., computing the median from a table), and arguments. In these practices, mathematical objects are represented in verbal, graphical, or symbolic ways.

The authors take from Eco (1995/1976) the idea of "semiotic function" as a correspondence between an expression and a content, which is fixed by a rule of correspondence, and which suggests that any possible mathematical object (concept, property, argument, procedure, etc.) may play the role of both expression and content in a semiotic function. A *semiotic conflict* appears when the interpretation of mathematical expressions from students does not agree with what is expected by the teacher or the researcher. These semiotic conflicts produce errors in the students' work and these errors are not explained by their lack of knowledge, but because they make an incorrect correspondence between the two terms in a semiotic function. The aim of this work is to determine possible semiotic conflicts related to median in secondary school students.

In the third section we analyse the open responses given by 518 Mexican students from Educación Secundaria (*n* = 162; ages 14-15) and Bachillerato (*n* = 356; 17-18 year-olds) to a problem with three questions. Questions 1 and 2 involve the computation of the median in a small data set with an odd and an even number of elements. In the last part, an outlier is included and students are asked to decide whether the mean would be an adequate central measure to represent the new data set. Using ideas from the onto-semiotic approach, we classify the responses, taking into account the central tendency measure used, and describe the students' semiotic conflicts. Responses in parts 1 and 2 (computation of median) are analysed in Section 4 and responses to question 3 (selecting the best representative value for the data set) in Section 5. A complete semiotic analysis of the correct responses and of one incorrect response is included to exemplify the method. Then, a short description and an example for each of the different categories of responses is presented. In Tables 1 and 2 we summarise the different types of responses.

We used the Chi-square test to study the possible dependence between the responses and the students' group (Educación Secundaria or Bachillerato). We observed better computation of the median by Educación Secundaria students and better competence to select the best representative value in Bachillerato students. The following semiotic conflicts were identified and analysed:

- *Conflicts related to representations:* Confusing the verbal expressions or symbols for mean and median (Cobo, 2003).
- *Conceptual conflicts:* Confusing the concepts of mean and median. Interpreting the median as the centre of the unordered data set (also described by Barr, 1980;

Carvalho, 1998, 2001; and Cobo, 2003), interpreting the median as the geometrical centre of the distribution, confusing the mean or median with a value of the variable, confusing frequencies and values of the variable.

- *Attributing non-existent properties to the median:* Assuming the associative property or assuming the median is an internal operation (a conflict described for the mean by Mevarech, 1983), assuming that the mean is always located in the geometrical centre of the distribution, not perceiving the effect of an outlier on the value of the mean.
- *Procedural conflicts:* Failures in the computation of weighted averages (Mevarech, 1983), not considering the frequencies in computing the median, being unable to solve the case of indeterminacy, using frequencies (instead of the variable values) in computing the median.
- *Argumentative conflicts:* Inability to justify a procedure or wrong argumentation process.

Table 1. Frequency and percentage of responses to questions 1 and 2

Response	Count	%	Count	%
C1.1. Correct computation of median	352	68.0	310	60.0
C1.2. Not ordering the data in computing the median	64	12.4	52	10
C1.3. Median as the midpoint in the range of variation	5	1.0	3	0.6
C2.1. Correct computation of mean	70	13.5	95	18.3
C2.2. Incorrect computation of mean	2	0.4	11	2.1
C3. Solution not related to central tendency measures	3	0.6	7	1.3
C4. No answer	22	4.2	40	7.7
Total	518	100	518	100

Table 2. Frequency and percentage of responses in selecting the best representative value

Response	Count	%
C1. Median is more adequate when there is an outlier	159	30.7
C2.1. Mean is adequate, no justification	32	6.2
C2.2. Mean is adequate because it is a central tendency value	107	20.7
C2.3. Mean is adequate, computation of mean	86	16.6
C2.4. Mean is adequate, error in the computation of mean	11	2.1
C3. No response	123	23.7
Total	518	100.0

All these results suggest that it would be interesting and useful to expand this research in order to provide teachers with more knowledge of students' difficulties in relation to the median.

1. INTRODUCCIÓN

La comprensión de las medidas de tendencia central (media, mediana, y moda) ha sido investigada por autores como Barr (1989), Cai (1995), Gattuso y Mary (1996), Pollatsek, Lima y Well (1981), Watson y Moritz (1999, 2000), y Cobo (2003), quienes describen errores y dificultades entre los estudiantes de diversos niveles educativos. Estas investigaciones se han centrado principalmente en la media aritmética, aunque, en el análisis exploratorio de datos, enfoque recomendado actualmente en el currículo de matemáticas para la Educación Secundaria y Bachillerato en México, se da una gran importancia a la mediana. En la enseñanza en estos niveles educativos en México también se introducen algunas representaciones gráficas basadas en los estadísticos de orden, como el gráfico de la caja, cuya elaboración requiere el cálculo previo de la mediana.

Si queremos formar ciudadanos competentes en la interpretación de la información estadística que se encuentra en la vida profesional y cotidiana, sería importante que la enseñanza de la estadística tuviese en cuenta las posibles dificultades de los alumnos en el concepto de mediana y su cálculo. El propósito de este trabajo es analizar dichas dificultades, cuando se pide a los estudiantes calcular la mediana en un conjunto sencillo de datos, con un número par o impar de elementos. También tratamos de averiguar si los estudiantes perciben que la mediana es la medida de tendencia central que se debe calcular cuando el conjunto de datos contiene un valor atípico (valor muy alejado del resto del conjunto de datos). Con todo ello, continuamos nuestra investigación previa (Mayén, Batanero, & Díaz, 2009), con el fin de proporcionar información sobre la comprensión de los estudiantes mexicanos acerca de estos conceptos. En lo que sigue describimos brevemente el marco teórico e investigaciones previas, y presentamos los resultados del estudio.

2. MARCO TEÓRICO Y ANTECEDENTES

2.1. MARCO TEÓRICO

En este trabajo nos basamos en ideas teóricas del enfoque onto-semiótico (Godino & Batanero, 1997; Godino, Batanero, & Font, 2007; Godino, Batanero, & Roa, 2005). En este enfoque, el significado de los objetos matemáticos o estadísticos (por ejemplo, la mediana) emerge de las prácticas llevadas a cabo por una persona al resolver problemas relacionados con dicho objeto. Dichas prácticas involucran diferentes tipos de objetos (elementos de significado), que serán tomados en cuenta en nuestro trabajo:

- Situaciones-problemas, de donde surge el objeto: Por ejemplo, la comparación de datos medidos en escala ordinal o datos cuantitativos con valores atípicos serían problemas de los que surge la idea de mediana.
- Lenguajes (términos, expresiones, notaciones, gráficos): Por ejemplo, las palabras: “mediana,” “percentil 50,” así como los símbolos y gráficos asociados.
- Conceptos: Por ejemplo, mediana, variable, valor, o percentil.
- Propiedades: Por ejemplo, que el valor de la mediana no se ve afectado por los valores atípicos o que no cumple la propiedad asociativa.
- Procedimientos: Como los diferentes algoritmos de cálculo de la mediana para datos aislados, agrupados en tablas o a partir de gráficos.
- Argumentos: Usados para justificar o explicar a otra persona las proposiciones y procedimientos.

Para un objeto matemático (en este caso la mediana) en el enfoque onto-semiótico se diferencia entre significado institucional y personal. El significado institucional incluye las prácticas matemáticas que, en una institución de enseñanza, se intenta transmitir al estudiante, mientras que el significado personal estaría formado por las prácticas matemáticas adquiridas por el estudiante, alguna de las cuáles podrían no coincidir con las pretendidas en la institución. Para el caso de la mediana, estos elementos son analizados con detalle por Cobo y Batanero (2000), quienes muestran su complejidad, incluso cuando su estudio se aborda únicamente desde el punto de vista de la estadística descriptiva. Godino y Batanero (1997) señalan que, en las prácticas matemáticas, intervienen objetos ostensivos (símbolos, gráficos, etc.) y no ostensivos (que evocamos al hacer matemáticas), que son representados en forma textual, oral, gráfica, o simbólica.

En el trabajo matemático, los símbolos (significantes) remiten a entidades conceptuales (significados). Estas representaciones tienen mucha importancia para facilitar la enseñanza y el aprendizaje, pero a veces causan dificultades en los estudiantes. Godino et al. (2007) toman de Eco (1995/1976) la noción de "función semiótica" como una "correspondencia entre conjuntos," que pone en juego tres componentes:

- Un plano de expresión (objeto inicial, considerado frecuentemente como el signo);
- Un plano de contenido (objeto final, considerado como el significado del signo, esto es, lo representado, lo que se quiere decir, a lo que se refiere un interlocutor);
- Un criterio o regla de correspondencia (esto es un código interpretativo que relaciona los planos de expresión y contenido).

Esta idea de función semiótica destaca el carácter esencialmente relacional de la actividad matemática, pues cualquier posible objeto matemático (concepto, propiedad, argumento, procedimiento, etc.) puede jugar el papel tanto de expresión como de contenido en una función semiótica. La diversidad de funciones semióticas que se ponen en juego al resolver un problema de matemáticas o estadística sirve para explicar algunas dificultades y errores de los estudiantes. Godino et al. (2007) denominan *conflicto semiótico* a las interpretaciones de expresiones matemáticas por parte de los estudiantes que no concuerdan con las pretendidas por el profesor o investigador. Estos errores de interpretación (conflictos semióticos), producen equivocaciones en los estudiantes, que no son debidos a falta de conocimiento, sino a no haber relacionado adecuadamente los dos términos de una función semiótica.

Así, este trabajo se orienta a la determinación de conflictos semióticos de estudiantes de Secundaria y Bachillerato en relación con la mediana, con el fin de establecer condiciones de control de dichos conflictos en los procesos de estudio. Seguiremos el método empleado en otras investigaciones previas que utilizan el enfoque onto-semiótico (Cobo, 2003; Godino et al., 2005; Mayén et al., 2009). Este método consiste en realizar un análisis de las funciones semióticas requeridas durante el proceso de resolución correcta de un problema (solución experta), identificando los diferentes tipos de objetos matemáticos que el estudiante debe aplicar en dicho proceso. Seguidamente se compara esta "solución experta" con las soluciones erróneas obtenidas por los estudiantes. La comparación de la cadena de funciones semióticas previstas en el proceso correcto de resolución con la realmente llevada a cabo por el estudiante, en cada solución errónea, permite identificar los *conflictos semióticos,* es decir los significados que no concuerdan con los considerados correctos desde el punto de vista institucional.

2.2. INVESTIGACIONES SOBRE COMPRENSIÓN DE LA MEDIANA

Las investigaciones previas indican que la definición de mediana no es clara para los estudiantes. Barr (1980) realizó un estudio con alumnos de entre 17 y 21 años, concluyendo que estos estudiantes interpretan la mediana como el centro de "algo," pero no comprenden a qué se refiere este “algo.” También en el estudio de McGatha, Cobb, y McClain (1998) los estudiantes calcularon la media al pedirles encontrar una medida de tendencia central, sin tener en cuenta el contexto. Más aún los estudiantes no perciben que la mediana puede ser mejor representante del conjunto de datos que la media en algunas circunstancias (Zawojewski & Heckman, 1997). Estas dificultades también se han descrito en investigaciones realizadas con futuros profesores de educación primaria (Groth & Bergner, 2006; Jaccobe, 2008).

Algunos estudiantes que son capaces de calcular la mediana cuando los datos se dan listados, tienen dificultad para calcularla a partir de una tabla de frecuencias. Incluso los alumnos universitarios encuentran difícil de aceptar que se puedan emplear dos algoritmos diferentes para el cálculo de la mediana dependiendo del tipo de datos (agrupados o no agrupados) y que puedan obtenerse valores distintos en el cálculo con datos agrupados al variar la amplitud de los intervalos de clase. Tampoco comprenden cómo se pasa de la definición de la mediana a su cálculo (Schuyten, 1991).

Por su parte, Batanero, Estepa, y Godino (1997) sugieren que los alumnos se encuentran con obstáculos para calcular la mediana si parten de las representaciones gráficas de las frecuencias acumuladas, ya que no están acostumbrados a las funciones discontinuas a saltos. En caso de interpolar para hallar el valor de la mediana, cometen errores por fallo de razonamiento proporcional. Los alumnos no tienen tampoco suficiente dominio en el manejo de las desigualdades que aparecen asociadas a la definición de mediana y a su cálculo. Otros errores encontrados por Carvalho (1998, 2001) al analizar el cálculo de la mediana en alumnos de 13-14 años son (a) no ordenar los datos al calcular la mediana, entendiendo que la mediana es el centro de la lista de datos “no ordenada”; (b) calcular el dato central de las frecuencias absolutas ordenadas de forma creciente, es decir, confundir la frecuencia con el valor de la variable; o (c) calcular la moda en vez de la mediana.

Algunos de estos errores fueron encontrados en un trabajo anterior (Mayén et al., 2009) donde, utilizando la misma metodología de análisis y el mismo marco teórico, se analizó una tarea donde se pedía la comparación de dos conjuntos de datos ordinales. En dicho trabajo se encontró que los alumnos no consideran las medidas de tendencia central al comparar dos distribuciones; que confunden el valor de la variable con las frecuencias a la hora de calcular las medidas de tendencia central y de nuevo, que confunden la media y mediana, intentando calcular la media en datos ordinales. Queremos analizar si los errores encontrados en la comparación de datos ordinales se repiten cuando los estudiantes se enfrentan a conjuntos de datos cuantitativos. En lo que sigue describimos la tarea utilizada.

3. MÉTODO

3.1. MUESTRA

La muestra estuvo compuesta por 518 estudiantes mexicanos de dos niveles educativos: 356 alumnos de Bachillerato, de diferentes Centros de Estudios Científicos y Tecnológicos del Instituto Politécnico Nacional (en total seis centros) y 162 estudiantes de Secundaria (dos centros). Los estudiantes de Secundaria tenían una edad de entre 14 y 15 años, y cursaban el último año de este nivel educativo. Habían estudiado por primera vez las medidas de tendencia central en el mismo curso en el que aplicamos el cuestionario, durante aproximadamente un mes y alrededor de dos meses antes de la administración del cuestionario. La otra parte de la muestra está compuesta por estudiantes del último curso de Bachillerato y sus edades oscilan entre 17 y 18 años. En el mismo curso en que se les aplicó el cuestionario habían estudiado estadística, incluyendo el tema de las medidas de tendencia central. En este nivel se tomó un mayor tamaño de muestra, pues estábamos interesados en los conocimientos de los alumnos que van a acceder a la universidad en el siguiente curso. En los dos niveles, los alumnos son de clase social media y los centros están distribuidos por varias zonas de la Ciudad de México.

3.2. PROBLEMA PROPUESTO Y MÉTODO DE ANÁLISIS

Vamos a analizar a continuación el problema propuesto (Figura 1) que es muy similar a otros ejercicios de cálculo de la mediana que aparecen en los libros de texto mexicanos utilizados por los alumnos de la muestra y similar a los que aparecen también en los libros de texto españoles. En los dos primeros apartados del ítem se pide el cálculo de la mediana con número par e impar de datos, introduciendo en la segunda un valor atípico, para analizar si los alumnos comprenden su efecto sobre el cálculo de media y mediana.

El peso en kilos de 9 niños es 15, 25, 17, 19, 16, 26, 18, 19, 24

1. ¿Cuál es el peso del niño mediano?
2. ¿Cuál es la mediana si incluimos el peso de otro niño que pesa 43 kg?
3. En el segundo caso ¿sería la media aritmética un buen representante de los 10 datos? Razona la respuesta.

Figura 1. Problema propuesto

En el último apartado del ítem se evalúa la competencia para decidir cuál de las medidas de posición central es preferible para representar este conjunto de datos. De manera implícita, los alumnos deben usar en este ítem la definición de mediana como elemento central que divide a la población en dos partes iguales. También han de conocer los algoritmos de cálculo y las propiedades citadas (de la mediana y media para poder contestar el último apartado). En lo que sigue elaboramos una solución "experta," para proceder a su análisis semiótico.

Para resolver el primer apartado, habría que ordenar los datos y tomar el elemento central, aplicando directamente la definición de mediana. Los alumnos han de comprender que la mediana es el centro de la distribución cuando los datos están ordenados, respecto al orden numérico habitual y combinar las ideas de centro y orden. A partir de los datos dados, los alumnos tendrían que producir una ordenación como la siguiente y determinar el centro de la fila. El valor correspondiente sería la mediana.

15, 16, 17, 18, <u>19</u>, 19, 24, 25, 26

En el apartado 2 tenemos un nuevo elemento (43), que es un valor atípico, pues comparado con el resto de los pesos de los niños imaginarios en el problema, es excesivo. Como ahora el número de elementos es par, nos encontramos con un caso de indeterminación, pues al ordenar el conjunto de datos y buscar el valor central, encontramos dos valores. Por tanto, los dos elementos centrales del conjunto de datos cumplen la definición de mediana, al estar situados en el centro de la distribución. Para resolver la indeterminación se introduce un convenio que se enseña a los alumnos, que consiste en obtener la media de los dos valores centrales y tomar dicho valor medio como mediana del conjunto de datos. En nuestro caso, al ser iguales los dos valores centrales se obtiene otra vez el valor 19 como mediana.

15, 16, 17, 18, <u>19, 19</u>, 24, 25, 26, 43

En cuanto al apartado 3, el mejor representante es la mediana, ya que la media se vería afectada por el alto valor 43 y no sería razonable que, al preguntar, por ejemplo, cuál es el peso representativo de este grupo de niños diéramos el valor medio 22.2 cuando el 60% de los casos está por debajo de (o igual a) 19 kilos. Por el contrario, la mediana (19) no se ve afectada por los valores extremos y exactamente el 50% de los valores están por encima y debajo de este peso. Esta solución al problema aparentemente sencilla tiene una gran complejidad, como mostramos en el análisis semiótico de la tarea, que se presenta en la Tabla 1. Para facilitar el análisis, dividimos el ítem en unidades, asignando en la primera columna un código a las unidades de análisis. En la segunda columna presentamos una expresión matemática, que, mediante elementos de lenguaje tales como números, símbolos, o palabras, evoca una serie de objetos matemáticos (que describiremos en la tercera columna, donde incluimos el contenido de la función semiótica). Generalmente esta evocación se hace a través de ciertos convenios explícitos o implícitos, muchos de ellos aprendidos. Estos convenios definirían la regla de correspondencia de la función semiótica.

Tabla 1. Análisis de la solución correcta al problema

Unidad	Expresión	Contenido
U2	Para resolver la parte 1, habría que ordenar los datos y tomar el elemento central, puesto que hay un número impar de elementos. 15, 16, 17, 18, 19, 19, 24, 25, 26 (19 kg)	- El alumno ha de recordar y comprender la definición de mediana (concepto) como valor que divide en dos partes el conjunto de datos (propiedad). - Debe colocar los datos de menor a mayor. Supone la idea de orden (concepto), conocer el orden relativo de los números enteros (propiedad) y ser capaz de ordenar la serie (algoritmo). - Debe aplicar el algoritmo de cálculo de la mediana de un conjunto impar de valores aislados (procedimiento). - Uso de la representación numérica; representación simbólica de la unidad de medida (kg) (proceso representacional). - Atribución del valor 19 a la mediana particular del conjunto de datos (particularización de la idea general de mediana a este caso).
U3	2. ¿Cuál es la mediana si incluimos el peso de otro niño que pesa 43 Kg? El número de elementos es par. Se obtiene otra vez el 19: 15, 16, 17, 18, 19, 19, 24, 25, 26, 43 (19 kg)	- Hay que interpretar la definición de mediana como valor central en el caso de número par de valores. - Aplicar el algoritmo de cálculo de la mediana a un conjunto par de valores aislados. Comprensión de que los dos valores centrales cumplen la definición (aplicación de una propiedad). - Conocimiento sobre cómo resolver el caso de indeterminación (algoritmo). - Obtener la media de los dos valores centrales, una vez ordenados (algoritmo); concepto de media o de punto medio. - Uso de la representación numérica; proceso de representación.
U4	3. En este caso, ¿sería la media aritmética un buen representante de los 10 datos? El alumno indica que el mejor representante es la mediana, ya que la media se vería muy afectada por el alto valor atípico 43.	- El alumno ha de recordar la definición de media (concepto) como algoritmo (sumar y dividir por el total). Ha de recordar que la media no es resistente a valores atípicos (propiedad). - Ha de reconocer la presencia de un dato atípico (concepto); particularización de la idea general de “atípico” al caso particular; se requiere también un conocimiento del contexto. - Debe reconocer que en el cálculo de la media intervienen todos los datos, por tanto, no es un estadístico robusto (propiedad). - Ha de tomar la decisión de que la mediana es preferible (argumento de análisis y síntesis).

Cuando un resolutor escribe un texto para resolver un problema, cada expresión (segunda columna) evoca un contenido (tercera columna), es decir, se establece una función semiótica a través un proceso de significación. Cuando otra persona lee el texto realiza un proceso de interpretación. Es importante resaltar que, en la actividad matemática, los procesos de significación e interpretación son múltiples y encadenados, como se muestra en los ejemplos que discutimos a continuación. Algunas expresiones (por ejemplo “peso”) remiten a un objeto (magnitud peso) y quizás también a la idea de magnitud en sí misma (como tipo de una clase de diferentes magnitudes). El análisis que describimos a continuación se centra en los razonamientos de los estudiantes al resolver el problema, identificando los objetos estadísticos que usan correcta e incorrectamente. Nos proponemos comprobar si las dificultades encontradas en las investigaciones previas se presentan en los alumnos mexicanos, con qué frecuencia y si varían según el nivel escolar. Queremos también explicar estas dificultades en términos de conflictos semióticos.

4. RESULTADOS EN EL CÁLCULO DE LA MEDIANA

Recogidas las respuestas de los estudiantes, se inició un proceso cíclico de categorización, comparando respuestas similares. Se realizó un análisis semiótico de las respuestas típicas en cada categoría para inferir los objetos y procesos matemáticos que el estudiante usa en su resolución. Se clasificaron las respuestas, considerando en primer lugar, la medida de tendencia central que se usa (media, mediana, moda, o ninguno), y en segundo lugar, las respuestas que corresponden a cada una de ellas, teniendo en cuenta la existencia de conflictos semióticos semejantes. De este modo, se llegó a la lista de categorías de respuestas que se describe a continuación.

C1. Usar la mediana A continuación analizamos las respuestas que utilizan la medida de tendencia central pedida, es decir, los alumnos calculan la mediana. En estas respuestas los estudiantes identifican el término “mediana” con el objeto matemático correspondiente. Dentro de ella hemos encontrado tres variantes.

C1.1. Cálculo correcto de la mediana Son las respuestas semejantes a la solución correcta experta que ya se ha analizado, por lo que no incluimos nuevos ejemplos.

C1.2. El alumno no ordena los valores y considera como mediana el valor central de todos los elementos tal como se le presentan Este error que también es encontrado por Barr (1980), Carvalho (1998, 2001), Cobo (2003) y Mayén et al. (2009). El ejemplo que se muestra en la Figura 2 se analiza en la Tabla 2.

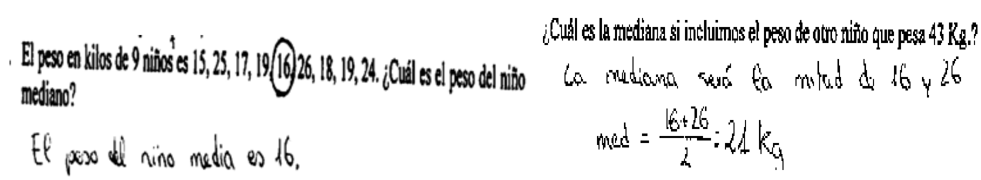

Figura 2. Ejemplo de respuesta en la categoría C1.2

En el ejemplo vemos que el estudiante, al tratar de calcular la mediana, toma el valor central pero no ordena los datos y se produce el conflicto. Es consistente en su respuesta pues emplea el mismo procedimiento en la segunda parte, resolviendo incluso el caso de indeterminación. La explicación que damos es que el estudiante entiende que el orden al que se refiere la definición de mediana es el orden de los datos, es decir, el orden numérico de los niños que ha asignado en la tabla de datos, en lugar del orden numérico natural. Esta es una confusión plausible pues a veces ordenamos a los estudiantes de una clase según diversos criterios tales como altura, apellidos, o colocación en la clase.

Tabla 2. Análisis de ejemplo en la categoría C1.2

Unidad	Expresión	Contenido
U1	*El peso en kilos de 9 niños es 15, 25, 17, 19, 16, 26, 18, 19, 24. ¿Cuál es el peso del niño mediano?* El peso del niño media es 16.	- Asigna a cada numeral (representación numérica) el valor de una cantidad de magnitud peso (concepto); proceso de representación. - Visualiza el conjunto de numerales como un conjunto de datos; construye un objeto (distribución) por composición de objetos simples. - Referencia implícita al peso particular de cada niño (cantidad de medida de la magnitud en cada caso). - Usa la idea de mediana (concepto) como valor central, que corresponde al valor de la variable del elemento central en la serie (propiedad). Usa las ideas de orden y centro (conceptos). - *No ordena* la serie de pesos según el orden habitual en los números enteros. *Conflicto* al olvidar una propiedad esencial en la mediana. El alumno considera el orden en que se dieron los datos, en lugar del orden numérico habitual.
U2	*¿Cuál es la mediana si incluimos el peso de otro niño que pesa 43 kg?* La mediana será la mitad de 16 y 26. med = $\frac{16+26}{2}$ = 21 kg	- El alumno identifica correctamente el caso de indeterminación (propiedad); hay dos valores centrales en la serie (16 y 26). - Supuesto que el conjunto estuviese bien ordenado, el alumno resuelve correctamente el caso de indeterminación (procedimiento). - Calcula correctamente la media (aplica la definición y algoritmo de la media, es decir un concepto y un procedimiento). - El *conflicto* se produce por confusión del orden al que se refiere la definición de mediana. - Usa correctamente símbolos (igualdad, división, suma, kg, etc.) que se refieren a conceptos y procedimientos.

C1.3. Mediana como valor central del rango de datos El estudiante obtiene la media aritmética de los valores extremos de la distribución y al resultado lo considera como mediana, como mostramos en el ejemplo de la Figura 3. En consecuencia, el alumno aplica la idea de mediana como valor central, pero aparece un conflicto en la identificación de cuál es valor del centro que se debe de calcular. De la idea general de "centro" el alumno ha de particularizar a la idea de "mediana como centro"; pero hay un conflicto debido a que la particularización que hace es incorrecta. Es decir, asocia la mediana a la idea de centro, pero confunde "centro estadístico de la distribución," con "centro geométrico del rango de variación." Por otra parte, el alumno ha identificado el máximo y mínimo de la serie de datos y ha calculado el punto medio del segmento marcado por estos dos extremos. Conoce y usa la idea y procedimiento de cálculo del punto medio. Además, el alumno redondea el resultado para obtener un valor entero de la mediana (pues los datos son enteros); atribuyendo una propiedad inexistente a la mediana (ser operación interna). No hemos encontrado este conflicto analizado en la investigación de Cobo (2003), ni en otros antecedentes, por lo que la explicación que hacemos del mismo constituye una aportación de nuestro trabajo.

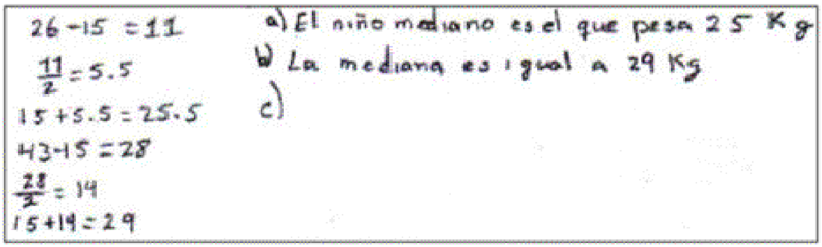

Figura 3. Ejemplo de respuesta en la categoría C1.3

C2. Usar la media, en vez de la mediana Otro grupo de respuestas corresponde a los alumnos que resuelven el problema calculando directamente la media aritmética, sin hacer ningún tipo de referencia a la mediana ni llegar a calcularla. Muestran confusión entre los conceptos o bien entre la terminología de media y mediana, es decir, se trata de un fallo en un proceso de significación, pues atribuyen al término "mediana" un significado no esperado por el profesor. Dentro de esta categoría general se han diferenciado varios casos.

C2.1. Cálculo correcto de la media con datos aislados Son los alumnos que, para resolver el problema, calculan correctamente la media. Estos alumnos conocen y aplican correctamente el algoritmo de cálculo de la media, poniendo en relación el concepto media con el algoritmo correspondiente y son capaces de realizar las operaciones requeridas. También realizan una serie de procesos de particularización (del concepto general media a la media concreta del conjunto de datos; del conocimiento del algoritmo general, a la aplicación del algoritmo al caso particular). Además, se realizan correctamente los correspondientes procesos de representación de datos, algoritmo y resultado. En el siguiente ejemplo el alumno ha asignado un valor numérico a cada numeral, visualizando el conjunto de valores como un conjunto de datos (proceso de composición al ir de los elementos al todo). El proceso de significación de la magnitud peso, la unidad de medida y los valores individuales de los pesos, así como del número de datos han sido correctos. El alumno comprende la idea de distribución y visualiza la media como una propiedad de la distribución (no de cada dato aislado). Por otro lado, el alumno recuerda el algoritmo de cálculo de la media para datos aislados, y lo aplica correctamente. Sin embargo, se produce un conflicto semiótico al no diferenciar entre media y mediana, al menos verbalmente.

$$\frac{15+25+17+19+16+26+18+19+24}{9} = 19.8 \textit{ kilos; 19.8 es el peso del niño mediano;}$$

$$\frac{15+25+17+19+16+26+18+19+24+43}{10} = 22.2 \textit{ kilos; 22.2 es el peso del niño mediano}$$

C2.2. Cálculo incorrecto de la media con datos aislados El alumno utiliza la media, en lugar de la mediana, por lo que se repite el conflicto descrito para el caso anterior, pero además comete algún error de cálculo. En el ejemplo presentado en la Figura 4, el alumno calcula correctamente la media en la primera parte. Al valor obtenido, suma el nuevo valor 43, y divide por dos para responder el segundo apartado. La respuesta contiene diferentes conflictos. Por un lado, hay un fallo en el proceso de significación, ya que a nivel conceptual o terminológico el estudiante confunde media y mediana. Por otro lado, se añade un conflicto procedimental a la hora de calcular la media, porque ésta no tiene la propiedad asociativa (atribuye una propiedad inexistente por generalización abusiva de esta propiedad). En consecuencia, la media calculada sería incorrecta, pues no pondera la primera media por su frecuencia. Este error ya fue descrito por Cobo (2003), Mayén et al. (2009), y Mevarech (1983).

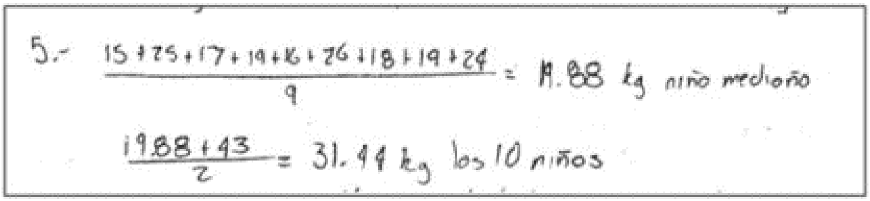

Figura 4. Ejemplo de respuesta en la categoría C2.2

C3. Dar una solución que no tiene relación con las medidas de tendencia central En lugar de calcular una medida de tendencia central, el estudiante da como respuesta un valor de la variable que no corresponde ni al valor central del conjunto de datos ordenados según el orden numérico, ni tampoco al valor central del conjunto de datos original (sin ordenar). Pensamos que el conflicto se debe a que el alumno usa el sentido coloquial del término "mediano" dando un valor cualquiera de la variable diferente del máximo o del mínimo, pero que no corresponde al centro. Este error no lo hemos visto explicado en las investigaciones previas, aunque el número de alumnos que lo utiliza ha sido muy pequeño. Mostramos un ejemplo donde el valor 18 no está en el centro ni se relaciona con la media o moda: *"El peso del niño mediano es de 18 kilos."*

Una vez categorizadas las respuestas, obtuvimos la frecuencia de cada una de ellas, que se presentan en la Tabla 3. Observamos que la mayor parte de los estudiantes usa la mediana, y casi todos la calculan correctamente (68% en la primera parte y 60% en la segunda, pues algunos estudiantes no resuelven el caso de indeterminación), siendo los resultados mejores que en la investigación de Cobo (2003), quien obtuvo solamente el 38% de resultados correctos en el cálculo de la mediana.

Tabla 3. Frecuencia y porcentajes de respuestas en los apartados 1 y 2

	Apartado 1		Apartado 2	
Respuesta	Frecuencia	%	Frecuencia	%
C1.1. Cálculo correcto de la mediana	352	68.0	310	60.0
C1.2. Calcula la mediana sin ordenar los valores	64	12.4	52	10
C1.3. Mediana como valor central del rango de datos	5	1.0	3	0.6
C2.1. Cálculo correcto de la media con datos aislados	70	13.5	95	18.3
C2.2. Cálculo incorrecto de la media con datos aislados	2	0.4	11	2.1
C3. Dar una solución sin relación con las medidas de tendencia central	3	0.6	7	1.3
C4. No contesta	22	4.2	40	7.7
Total	518	100	518	100

Explicamos esta diferencia por la mayor edad de nuestro grupo de estudiantes, en especial en los estudiantes de Bachillerato que también tienen más formación. Un 12.4%

de estudiantes en la primera parte y 10% en la segunda parte calcula la mediana sin ordenar los valores, respuesta que hemos explicado por la existencia de un conflicto semiótico consistente en la confusión del orden referido en la definición de mediana. Este mismo error también fue encontrado por Barr (1980) en su estudio con alumnos de entre 17 y 21 años y por Cobo (2003) en un 26.2% de estudiantes, aunque estos autores no explican la posible causa del error. Por otra parte, un 13.5% y 18.3% de estudiantes de nuestra muestra confunde media y mediana en la primera y segunda parte del ítem, calculando la media correctamente, mientras el porcentaje fue del 25% en la investigación de Cobo (2003). Un pequeño número de estudiantes presenta conflictos no descritos anteriormente, como el de confundir "centro estadístico" con "centro geométrico" del rango de variación de la variable o no utilizar las medidas de tendencia central para resolver el problema.

En la Tabla 4 y Figura 5 clasificamos los resultados en función de la medida de tendencia central empleada, donde se aprecia que la mayor parte de estudiantes reconoce la idea de mediana en este ítem. Observamos que el 81.3% de los alumnos de la muestra conoce o recuerda la definición de mediana, y es capaz de calcularla para un conjunto impar de valores aislados, lo que arroja resultados aceptables. En resumen, podemos decir que el ítem ha sido sencillo para nuestros estudiantes y que hemos obtenido mejores resultados que los de Cobo (2003).

Tabla 4. Frecuencia y porcentaje de medida de tendencia central calculada

Respuesta	Nivel		
	Secundaria	Bachillerato	Total
Mediana			
Frecuencia	138	283	421
% de nivel	85.2	79.5	81.3
Media			
Frecuencia	12	60	72
% de nivel	7.4	16.9	13.9
Otro			
Frecuencia	12	13	25
% de nivel	7.4	3.7	4.8
Total	162	356	518

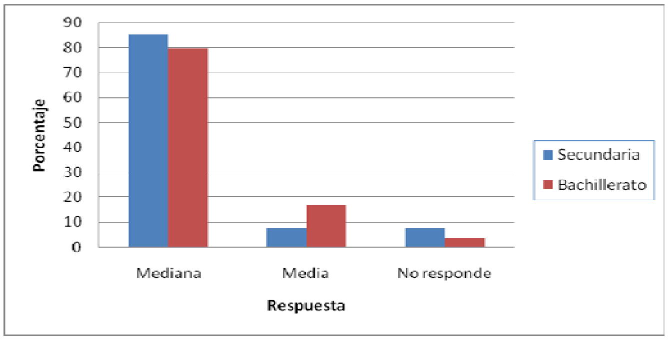

Figura 5. Porcentaje de medida de tendencia central calculada en cada grupo

La principal diferencia que se observa es que los alumnos de Secundaria (que son los menores en edad) dan mayor uso de la mediana, y, por lo tanto, mayor número de respuestas correctas. En este mismo grupo de estudiantes, un bajo porcentaje utiliza la media, coincidiendo la misma proporción de alumnos que no responden al ítem, mientras que aumenta el uso de la media en los estudiantes de Bachillerato. Pensamos que poner mayor énfasis en el estudio de la media durante la enseñanza de estadística en Bachillerato hace que los alumnos traten de resolver todos los problemas usando este estadístico, ya que son pocos los que usan otros procedimientos. Al realizar un contraste de hipótesis de la diferencia en la distribución de respuestas en los dos grupos, se obtuvo un valor Chi cuadrado = 10.845 con 2 grados de libertad y una significación 0.0044. Esto indica que encontramos diferencias significativas de respuestas según el nivel escolar de los estudiantes.

5. ELECCIÓN DE LA MEDIDA DE TENDENCIA CENTRAL MÁS REPRESENTATIVA

En la última parte del ítem se pretende evaluar si los estudiantes son capaces de elegir un buen representante estadístico para una distribución en la que aparece un valor atípico, y su capacidad de argumentar su elección con base en las propiedades de las medidas de tendencia central. Para responder a la pregunta, el alumno debe tener en cuenta las siguientes propiedades: “En el cálculo de la media, intervienen todos los valores (magnitud) de los datos, mientras que en el de la mediana sólo el número de valores,” “la media cambia siempre que cambia algún dato, mientras que la mediana puede no cambiar,” “la media y mediana coinciden únicamente en distribuciones simétricas,” y “la media es menos resistente al efecto de los valores atípicos que la mediana.” Por otro lado, los alumnos deben conocer las definiciones y algoritmos de cálculo de la media y mediana. Las respuestas que hemos encontrado se describen a continuación.

C1. Considera que la media no es representativa y la mediana es preferible Algunos alumnos dan la respuesta correcta, indicando que para este ejemplo particular, la media no sería un buen representante. Estos alumnos detectan el valor atípico en los datos y su efecto sobre el cálculo de la media, convirtiéndola en un representante no adecuado para este problema. El estudiante que da este argumento ha sido capaz de aplicar con éxito las propiedades descritas de la mediana y además sabe elegir un representante adecuado del conjunto de datos. Cobo (2003) también encontró esta respuesta. Un ejemplo de la misma es el siguiente: *“La media aritmética no sería un buen representante porque los valores finales son muy desviados, es decir, tienen mucha diferencia entre sí.”*

C2. Considerar que la media es un representante adecuado En estas respuestas observamos un conflicto sobre la idea de representante, pues la media es muy sensible a los valores extremos. Cobo (2003) también obtuvo esta respuesta, pero no diferencia categorías de respuestas dentro de ella. Nosotros hemos obtenido las siguientes categorías.

C2.1. Considerar que la media aritmética es adecuada como representante del conjunto de datos, sin dar una explicación de cómo llega a esa conclusión En esta respuesta no podemos ver las razones en las que el estudiante apoya su argumento: *“Sí sería un buen representante.”*

C2.2. Suponer la media adecuada en cualquier circunstancia Como en el caso anterior, esta respuesta enmascara un conflicto, al no tener en cuenta la distorsión introducida por el alto valor extremo. Supone un fallo al reconocer que la media sólo se aproxima al centro de la distribución en distribuciones simétricas. Posiblemente esta confusión pueda ser debida a la forma en que algunas veces se definen las medidas de tendencia central, como “estadísticos alrededor de los cuáles se agrupan los datos.” También implica que el alumno piensa que la media es resistente frente a valores atípicos, lo que es razonable si en la enseñanza recibida no se hizo especial énfasis en este concepto. Un ejemplo es el siguiente: “*Sí, porque es un promedio de los 10 pesos, lo que te da una idea general acerca de los niños.”*

C2.3. Además de considerar la media como mejor representante y dar un argumento, el alumno obtiene la media aritmética considerando todos los datos, incluyendo el 43 El alumno no compara el promedio obtenido con la serie de datos, no percibe la diferencia respecto a éstos. Sin embargo, en su respuesta podemos apreciar que conoce el algoritmo de la media y lo aplica sin errores: “*Media aritmética = 22.2; Sí porque nos da una idea de cuánto debe pesar un niño.”*

C2.4. Análogo al caso anterior, pero con error en el cálculo de la media, pues el alumno considera como media aritmética al resultado de dividir el valor nuevo 43 por dos Además de todos los conflictos anteriores, se observa uno más respecto a la definición de media aritmética y al conocimiento de su algoritmo, por no aplicar la media ponderada. No obstante, el valor numérico obtenido representa mejor el conjunto de datos que el que se obtendría al calcular directamente la media, lo que puede explicar el procedimiento usado por este estudiante: “*Sería 21.5 la media aritmética el representante.”*

En la Tabla 5 podemos observar que este apartado fue mucho más difícil que los dos anteriores, al aumentar el número de respuestas en blanco (24%). Fue también más difícil en los estudiantes de Cobo (2003), aunque la autora no informa sobre la frecuencia de estudiantes que no da respuesta. Pensamos que esta dificultad se debe a que la clase de matemáticas se centra mucho en el cálculo de promedios y no se suele incidir en las actividades interpretativas o en la elección del estadístico más adecuado para cada situación.

Tabla 5. Frecuencias y porcentajes de respuestas en el tercer apartado

Respuesta	Frecuencia	%
C1. Considera que la media no es adecuada y la mediana es preferible	159	30.7
C2.1. Considerar la media adecuada, sin justificar la respuesta	32	6.2
C2.2. Indicar que la media es adecuada siempre	107	20.7
C2.3. Considerar la media adecuada y calcularla incluyendo el 43	86	16.6
C2.4. Considerar la media adecuada con error de cálculo en la media	11	2.1
C3. No contesta	123	23.7
Total	518	100.0

De los estudiantes que responden este ítem, la mayoría considera que la media es representativa, sin darse cuenta de que el valor atípico afecta a la media aumentando su valor. En la investigación de Cobo (2003), un 17% de estudiantes de 1º de ESO (12-13 años) y un 11.8% de 4º (15-16 años) aceptan la media como representativa en este ítem, sin embargo, es más frecuente este tipo de respuesta en los alumnos de nuestra muestra. Aproximadamente la tercera parte de estudiantes, 30.7%, responde correctamente e indica que al utilizar la media los pesos medios se distorsionan al incluir un valor tan alejado del resto de los datos. En los estudiantes de Cobo (2003), el 21.4% de 1º de ESO y el 11.1% de 4º, también contestan correctamente a este ítem. Observamos que en esta parte, nuestros resultados son mejores. Hemos diferenciado varias categorías, señalando que el 20% de estudiantes dan esta respuesta sin justificarla y un 18.7% indican esta misma respuesta, incluso después de haber calculado la media.

En la Tabla 6 resumimos los datos anteriores por grupo sumando las respuestas que hacen referencia a cada medida de tendencia central como más adecuada. En esta parte del ítem, donde se espera una respuesta verbal, notamos que persiste la confusión entre media y mediana, ya que casi la mitad de los estudiantes, 45.6%, contesta que es mejor utilizar la media como representante del conjunto de datos. Otro 30% de alumnos dan la respuesta correcta, y también es alto el porcentaje de estudiantes que no responden al ítem (23.7%). En este apartado, el grupo de Bachillerato obtuvo mejores resultados, con un 36% de respuestas correctas, en comparación con el 19.1% de alumnos de Secundaria respondiendo correctamente. Por otra parte, la mitad de alumnos de Secundaria, así como el 43.5% de los estudiantes de Bachillerato prefieren utilizar la media. Al parecer, resultó difícil para nuestros alumnos dar respuesta a este ítem, ya que el 31% de los estudiantes de Secundaria y el 20% de Bachillerato no lo responde. Se obtuvo un valor Chi cuadrado = 16.312 con 2 grados de libertad y una significación menor que 0.0003, cumpliéndose las condiciones de aplicación. Este valor de Chi-cuadrado indica que las diferencias entre los grupos son estadísticamente significativas.

Tabla 6. Frecuencia y porcentajes de respuestas al tercer apartado por nivel escolar

Respuesta	Nivel		
	Secundaria	Bachillerato	Total
Mediana			
Frecuencia	31	128	159
% de nivel	19.1	36.0	30.7
Media			
Frecuencia	81	155	236
% de nivel	50.0	43.5	45.6
No responde			
Frecuencia	50	73	123
% de nivel	30.9	20.5	23.7
Total	162	356	518

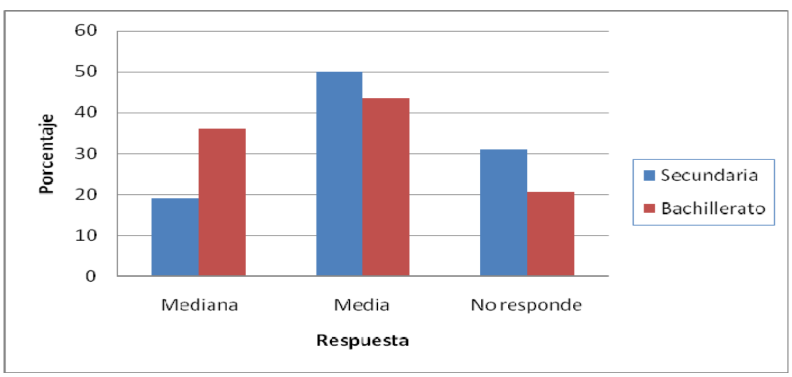

Figura 6. Medida de tendencia central considerada más representative

6. CONCLUSIONES

El estudio indica que el cálculo de la mediana en un conjunto pequeño de valores es sencillo para los estudiantes, pero no así la elección del mejor representante de un conjunto de datos en presencia de valores atípicos. Esta elección sigue siendo difícil para los estudiantes de Bachillerato, aunque estos presentan mejores resultados que sus compañeros. La explicación que damos es que la elección de la medida de tendencia central más adecuada para representar un conjunto de datos requiere un conocimiento estratégico (saber cuándo aplicar un concepto) mientras que el cálculo de la mediana sólo precisa de un conocimiento técnico (aplicar el concepto sin discutir su pertinencia al caso particular). En este sentido, nuestro trabajo confirma los resultados obtenidos en investigaciones previas que sugieren que los estudiantes o incluso profesores no perciben que la mediana puede ser mejor representante del conjunto de datos que la media en algunas circunstancias (Groth y Bergner, 2006; Jaccobe, 2008; Zawojewski y Heckman, 1997).

En el cálculo de la mediana las respuestas son peores en los estudiantes de Bachillerato que en los de Secundaria, mientras que en la elección del mejor representante del conjunto de datos ocurre lo contrario. En consecuencia deducimos que la enseñanza ayuda a desarrollar la comprensión de cuándo se debe utilizar la mediana pero no la competencia en cálculo. Hemos encontrado además los siguientes conflictos que podemos clasificar en relación a los elementos del marco teórico:

1. *Conflictos representacionales:* Consiste en confundir la terminología de media y mediana aunque luego el estudiante use correctamente la definición del concepto. Por ello, calculan la media (incluso correctamente) cuando se les pide calcular la mediana en las dos partes del ítem en coincidencia con los resultados de Cobo (2003).
2. *Conflictos conceptuales:* Cuando los estudiantes no discriminan entre media mediana y moda o entre estos y otros conceptos. Entre los conflictos relacionados con la definición de la mediana, el más frecuente, que aparece en todos los ítems, es interpretarla como centro del conjunto de datos sin ordenar. Este conflicto fue descrito por Barr (1980), Carvalho (1998, 2001), y Cobo (2003), y suponemos es debido a que el estudiante asume que el "orden" que se cita en la definición es el orden en que se han listado los datos y no el orden numérico convencional. También encontramos la interpretación de la mediana como centro geométrico de la distribución, haciendo una particularización incorrecta de la idea de centro que se ha de aplicar en un sentido específico (centro estadístico y no geométrico). Otra confusión ya notada por Carvalho (2001) es la existente entre mediana o media y valor de la variable. En el caso de la mediana el conflicto podría explicarse por un uso del sentido coloquial del término "mediano," dando un valor de la variable diferente del máximo o del mínimo, y que no corresponde al centro. Hacemos notar también que en algunos casos se confunden las frecuencias y los valores de la variable, calculando media o mediana con las frecuencias, error descrito por Carvalho (1998, 2001) y encontrado por Ruiz (2006) en su estudio sobre la variable aleatoria.
3. *Conflicto al atribuir propiedades a un concepto:* Los estudiantes generalizan de manera excesiva propiedades que conocen a conceptos que no conservan estas propiedades. Por ejemplo, hemos encontrado alumnos que suponen una estructura de operación interna a la mediana, conflicto que fue descrito por Mevarech (1983) para el caso de la media, pero no para la mediana. Conflictos de este tipo incluye suponer que la media se sitúa siempre en el centro de la distribución,

independientemente de la forma de la distribución, encontrado también por Navas, Batanero, y Godino (1997). También es patente que los estudiantes no aprecian el efecto del valor atípico en el cálculo de la media, fallando al elegir el mejor representante de un conjunto de datos.

4. *Conflictos procedimentales:* Cuando hay fallo en la aplicación de un procedimiento o se selecciona un procedimiento inadecuado en una situación. Por ejemplo, algunos estudiantes equivocan el total por el cuál hay que dividir en el cálculo de la media. En otros casos no se pondera en el cálculo de la media, error descrito por autores como Cobo (2003) y Mevarech (1983). Otros estudiantes no son capaces de resolver el caso de indeterminación de la mediana dando dos valores diferentes para la mediana, conflicto que suponemos debido a falta de conocimiento del convenio que hay que aplicar en este caso. Una respuesta no descrita en anteriores trabajos que aparece en nuestros estudiantes es no tener en cuenta la frecuencia en el cálculo de la mediana. Este error puede ser parecido al que comenten los estudiantes que no tienen en cuenta las frecuencias en el cálculo de la media. Mevarech indicó que para el caso de la media este error se debe a que los alumnos asumen que la media tiene la propiedad asociativa. En nuestro caso además de la posible atribución incorrecta de la propiedad asociativa a la mediana, pensamos que se produce un conflicto, tanto respecto a la idea de frecuencia, como respecto a la de orden. Al tratar de ordenar de menor a mayor el conjunto de datos asignándole un número, encuentra un valor repetido (19). Los dos valores 19 estarían en el mismo lugar de la ordenación. Para el estudiante es indiferente poner uno antes que el otro. Es decir, puesto que 19 remite al peso de un niño imaginario, si tenemos dos niños del mismo peso, sería indiferente colocar a uno u otro en el quinto lugar en orden de peso. El estudiante resuelve el conflicto omitiendo uno de los valores.
5. *Conflictos argumentativos:* Los estudiantes no son capaces de dar un argumento, de interpretar un argumento o no ven la inconsistencia en una argumentación.

Aunque muchos de los conflictos identificados han sido descritos en investigaciones anteriores, también hemos detectado algunos nuevos a los cuáles no se había prestado suficiente atención. Entre estos conflictos destacamos los siguientes:

- Interpretar la mediana como el centro geométrico del rango de la variable, conflicto que se presenta por un fallo en el proceso de particularización, desde la idea de "centro" a la idea de "centro estadístico" que es la que se espera del estudiante, quien por el contrario, particulariza a la idea de "centro geométrico," que posteriormente aplica y calcula correctamente.
- Suponer estructura de operación interna a la mediana por un proceso indebido de generalización; ya que la propiedad de conservación del conjunto numérico la tienen otras operaciones que el alumno conoce (suma, multiplicación, etc.) pero no las medidas de posición central, exceptuando la moda.
- No tener en cuenta las frecuencias en el cálculo de la mediana, error debido a un fallo al apreciar que la mediana no tiene propiedad asociativa, y en la comprensión de la mediana como estadístico de orden.

Todos estos conflictos enfatizan la necesidad de tener en cuenta la dificultad del concepto de mediana para los estudiantes y de reforzar su enseñanza. También sugiere el interés de continuar la investigación, proponiendo otras tareas relacionadas con la mediana o bien diseñando una enseñanza que permita mejorar la comprensión de los estudiantes. El énfasis actual en el análisis exploratorio de datos, que se incluye actualmente en la Educación Secundaria y Bachillerato, tanto en México como en España, supone prestar más atención a los estadísticos de orden, entre los que se encuentra la mediana, y que consideran la posición relativa de ciertos elementos dentro del conjunto de datos. El éxito de esta propuesta educativa depende de que los profesores conozcan las dificultades y conflictos de sus estudiantes y dispongan de instrumentos que las hagan aflorar para poder corregirlas. La investigación llevada a cabo en este trabajo y los resultados obtenidos permitirán ayudar al profesor en esta tarea.

AGRADECIMIENTO

Trabajo financiado por el proyecto SEJ2004-00789 MEC, Madrid y FEDER.

REFERENCIAS

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

Batanero, C., Estepa, A., & Godino, J. D. (1997). Evolution of students' understanding of statistical association in a computer-based teaching environment. En J. B. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics. Proceedings of the 1996 IASE Round Table Conference* (pp. 183-198). Voorburg, Netherlands: International Statistical Institute.

Batanero, C., Godino, J. D., & Navas, F. (1997). Concepciones de maestros de primaria en formación sobre los promedios [Pre-service primary school teachers' conceptions of averages]. En H. Salmerón (Ed.), *VII Jornadas LOGSE: Evaluación educativa* (pp. 310-304). Granada, Spain: University of Granada.

Cai, J. (1995). Beyond the computational algorithm. Students' understanding of the arithmetic average concept. En L. Meira (Ed.), *Proceedings of the 19th Conference of the International Group for Psychology of Mathematics Education* (Vol. 3, pp. 144-151). Universitye Federal de Pernambuco, Recife, Brazil.

Carvalho, C. (1998). Tarefas estadísticas e estratégias de resposta [Statistics tasks and response strategies]. Trabajo presentado en el *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* [Peer Interaction. Contributions to promoting logic and statistical development in the 7th year of school]. Tesis Doctoral, Universidad de Lisboa.

Cobo, B. (2003). *Significado de las medidas de posición central para los estudiantes de secundaria* [Meanings of central tendency measures for secondary school students]. Tesis Doctoral, Universidad de Granada.

Cobo, B., & Batanero, C. (2000). La mediana en la educación secundaria obligatoria: ¿Un concepto sencillo? [The median in compulsory education: An easy concept?]. *UNO 23*, 85-96.

Eco, U. (1995). *Tratado de semiótica general* [General semiotics]. Barcelona: Lumen. (Trabajo original publicado en 1976)

Gattuso, L., & 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* (Vol. I, pp. 401-408). Universidad de Valencia.

Godino, J. D., & Batanero, C. (1997). Clarifying the meaning of mathematical objects as a priority area of research in mathematics education. En J. Kilpatrick & A. Sierpinska (Eds.), *Mathematics education as a research domain: A search for identity* (pp. 177-196). Dordrecht: Kluwer.

Godino, J. D., Batanero, C., & 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., & 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.

Groth, R. E., & Bergner, J. A. (2006). Preservice elementary teachers' conceptual and procedural knowledge of mean, median, and mode. *Mathematical Thinking and Learning, 8*, 37-63.

Jacobbe, T. (2008). Elementary school teachers' understanding of the mean and median. En C. Batanero, G. Burrill, C. Reading, & A. Rossman (Eds.), *Joint ICMI/IASE study: Statistics in school mathematics. Challenges for teaching and teacher education. Proceedings of the ICMI Study 18 Conference and IASE 2008 Round Table Conference*. Monterrey, Mexico: International Commission on Mathematical Instruction and International Association for Statistical Education.
[Online: www.stat.auckland.ac.nz/~iase/publications.]

Mayén, S., Batanero, C., & Díaz, C. (2009). Conflictos semióticos de estudiantes mexicanos en un problema de comparación de datos ordinales [Semiotic conflicts of Mexican students in comparing ordinal data]. *Revista Latino Americana de Investigación en Matemática Educativa, 12*(2), 151-178.

McGatha, M., Cobb, P., & McClain, K. (1998, April). An analysis of students' statistical understandings. Trabajo presentado en el *Annual Meeting of the American Educational Research Association*, San Diego, CA.

Pollatsek, A., Lima, S., & Well, A. D. (1981). Concept or computation: Students' understanding of the mean. *Educational Studies in Mathematics, 12*, 191-204.

Ruiz, B. (2006). *Un acercamiento cognitivo y epistemológico a la didáctica del concepto de variable aleatoria.* [A cognitive and epistemological approach to the teaching of the concept of random variable] Tesis de Master, CICATA, México.

Schuyten, G. (1991). Statistical thinking in Psychology and Education. En D. Vere-Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics (ICOTS-3),* Dunnedin, (Vol. 1, pp. 486-490). Voorburg, Netherlands: International Statistical Institute.

Watson, J. M., & Moritz, J. B. (1999). The developments of concepts of average. *Focus on Learning Problems in Mathematics, 21*(4), 15-39.

Watson, J. M., & Moritz, J. B. (2000). The longitudinal development of understanding of average. *Mathematical Thinking and Learning, 2*(1 and 2), 11-50.

Zawojewski, J. S., & Heckman, D. J. (1997). What do students know about data analysis, statistics, and probability? En P. A. Kenney & E. A. Silver (Eds.), *Results from the sixth mathematics assessment of the National Assessment of Educational Progress* (pp. 195-223). Reston, VA: National Council of Teachers of Mathematics.

SILVIA MAYÉN
CECYT 6 Miguel Othón de Mendizábal
Instituto Politécnico Nacional
Av. Jardín, calle 4, Colonia del Gas
02950, México, D.F, México.
mayazuc@gmail.com