INFLUENCIA DEL CONTEXTO EN EL USO E INTERPRETACIÓN DE MEDIDAS DE CENTRALIZACIÓN AFECTADAS POR VALORES ATÍPICOS

Influence of the context in the use and interpretation of measures of centralization affected by outliers

Martínez, M. L.[a] y Huerta, M. P.[b]

[a]Florida Universitària, [b]Universitat de València

Resumen

En esta investigación estamos interesados en analizar hasta qué punto el uso e interpretación de las medidas de centralización están influidas por el contexto en el que se formulan los datos y si la presencia de valores atípicos entre ellos influye en la elección del mejor representante del conjunto de datos proporcionados. Se analizan las respuestas de los estudiantes a un cuestionario de problemas desde el punto de vista de su dificultad y de la influencia que tienen el contexto y los valores atípicos sobre su resolución.

Palabras clave: *didáctica, matemáticas, práctica educativa, contextos, medidas de centralización.*

Abstract

In this research we are interested in analyzing to what extent the use and interpretation of the measures of centralization are influenced by the context in which data are made and if the presence of outliers between them influences the choice of the best representative of a data set. We analyze students' responses to a questionnaire of problems from the point of view of its difficulty and the influence of the context and outliers on its resolution.

Keywords: *didactics, mathematics, educational practice, contexts, measures of centralization.*

INTRODUCCIÓN

Muchos trabajos de investigación siguen mostrando las dificultades que tienen los estudiantes al tratar con tareas de contenido estadístico en el que los significados, y no tanto los cálculos estén presentes. Nos referimos al uso e interpretación de las medidas de centralización, con sus medidas de dispersión correspondientes, como la mejor manera de representar, describir o estimar el comportamiento de un conjunto de datos. No nos referimos, en cambio, a sus cálculos sino a su uso e interpretación, que según se informa puede depender del contexto en el que los datos están expresados (Batanero, Godino y Navas, 1997). Este trabajo está en esa misma línea. Pretende ver hasta qué punto el uso e interpretación de las medidas de centralización están influidas por el contexto en el que se formulan los datos y si la presencia de valores raros, extraños o atípicos (estadísticamente hablando) entre ellos influye en la elección de su mejor representante. Cuanto mejor entendamos estas dificultades mejor podremos hacer aportaciones para mejorar el razonamiento estadístico de los ciudadanos.

En este trabajo nos centramos en el estudio de la resolución de problemas estadísticos ante la presencia de la variabilidad de los datos, por estudiantes de diferentes niveles educativos y formación. En particular, estudiamos la influencia que el contexto y el formato de los datos en el que se formulan los problemas puedan tener sobre las resoluciones de los estudiantes.

Los objetivos específicos en este trabajo son los siguientes:

Martínez, M. L. y Huerta, M. P. (2016). Influencia del contexto en el uso e interpretación de medidas de centralización afectadas por valores atípicos. En J. A. Macías, A. Jiménez, J. L. González, M. T. Sánchez, P. Hernández, C. Fernández, F. J. Ruiz, T. Fernández y A. Berciano (Eds.), *Investigación en Educación Matemática XX* (pp. 335-344). Málaga: SEIEM.

a. Identificar y medir los niveles de dificultad de los problemas considerados en relación con el contexto en el que se ha situado dicha resolución.

b. Evaluar la representatividad de la media aritmética para el resolutor ante la presencia de valores atípicos.

En este sentido, existe un claro consenso en que los problemas estadísticos, en los que la pregunta del problema no se refiere exclusivamente a un cálculo o aplicación de un algoritmo dado, léase el calcular una media, es una tarea difícil para los estudiantes. Pues bien, queremos medir esta dificultad y ver hasta qué punto está relacionada con la tarea propuesta.

La influencia de los valores atípicos, su consideración o no en la determinación de una medida de centralización, puede depender de múltiples factores. Incluso lo que estadísticamente sea calificado como valor atípico, aquel cuyo valor dista más de 1,5 veces el valor f_s, desde el cuarto más cercano (superior o inferior), siendo f_s la cuarta dispersión (Devore, 2005). Lo que estadísticamente se califica como valor atípico puede que para un resolutor no sea más que un valor extraño o raro, producto de un error en la anotación o de un error en la observación. En estos casos, puede ser lícito eliminar esos valores y seguir con el análisis. Pero puede que no sea este su origen, entonces debe considerarse en el análisis global del conjunto de datos al que pertenece (Anderson, Sweeney y Williams, 2010). Pero esto depende del observador y más cuando éste no tiene claro si el dato observado es estadísticamente atípico o por el contrario es extraño o raro y decide considerarlo o no en su análisis.

INVESTIGACIONES PREVIAS

Numerosas investigaciones se han interesado de algún modo u otro por las medidas de posición central o tendencia central, informándonos de que siempre presentan dificultades en el alumnado. Estas dificultades se extienden incluso a los futuros maestros. Así, Reading y Pegg (1996) ya observaron que los niños de grados 7 a 12, al reducir un conjunto de datos en uno solo que los representara, mostraban dificultad a la hora de dar un argumento o justificar su elección. Únicamente un número pequeño de estudiantes fueron capaces de justificarlo en relación con características del conjunto de datos y no de otra manera. Batanero, Godino y Navas (1997) realizan un estudio sobre las dificultades de comprensión de los promedios con profesores de primaria en formación y observan que éstos presentan dificultades en el tratamiento de los ceros y valores atípicos en el cálculo de los promedios, posiciones relativas de media, mediana y moda en distribuciones asimétricas, elección de la medida de tendencia central más adecuada en una determinada situación y el uso de los promedios en la comparación de distribuciones. Estrada, Batanero y Fortuny (2003), en una muestra con 367 futuros maestros, mediante un cuestionario de respuesta múltiple exploran el efecto que tiene la presencia de un valor atípico en un conjunto de 9 datos (mediciones) al considerar la media como mejor estimación de una cantidad desconocida, en un contexto de medida. Reportan que el 47% de los futuros profesores escogieron la respuesta correcta, lo que implica que más de un 50% de ellos no consideraron la posibilidad de descartar el valor atípico como no pertinente para la obtención del valor real más probable. Mayén, Díaz y Batanero (2009) exploran en tareas de resolución de problemas el uso y la interpretación de la mediana como mejor representante de una distribución dada. La tarea, abierta y formulada en un contexto de medida de los pesos de 9 niños, consiste en dotar de significado a la noción “peso del niño mediano” y comparar la mediana y la media como mejor representante del conjunto de datos al que se le añade un valor “atípico”, tal vez exagerado en el contexto en el que se está planteando la tarea. Se enfrenta así la escasa robustez del cálculo de la media ante la presencia de valores atípicos con la robustez de la mediana, y el sentido común a la hora de considerar como atípico un valor en un contexto dado y reconsiderar su validez para determinar qué usará para calcular su mejor representante: la media, mediana o moda. También, enfrentan la calificación de un valor como “atípico”, estadísticamente hablando, y atípico o extraño por el contexto en el que se está considerando este valor. Los resultados indican que la media es más robusta para el usuario, pues un 45% la prefiere frente al 30% que prefiere la mediana, ante la presencia de un valor atípico entre los datos. Ortiz y Font (2014), en un estudio realizado con estudiantes de primer curso de Magisterio, revelaron en su investigación que siguen dándose importantes dificultades relacionadas con la comprensión de los estudiantes de la media aritmética y sus propiedades. Pero una cosa es calcular medidas de tendencia central o de centralización para un conjunto de valores dados y otra ajustar una medida de centralización a un conjunto de datos en los que perece haber presencia de valores raros, anómalos, atípicos; en cuyo caso, hay que considerar la representatividad de la medidas anteriores para describir la "tipicalidad" de los datos (Groth y Bergner, 2006). Estos autores otorgan el nivel SOLO más sofisticado a aquellos estudiantes que son capaces de observar que la media está altamente afectada por datos atípicos, mientras que la mediana y la moda no lo están (ibid., p. 53).

A la media se le concede muchas veces el significado de valor típico o representativo de una distribución dada (Groth y Bergner, 2006, p. 52), lo que hace que en distribuciones simétricas la media se sitúe en el centro del rango de variación de los datos y sea su mejor representante, pero si esto no es así, entonces deja de serlo a favor de la mediana o de la moda (Batanero, Godino, Vallecillo, Green y Holmes, 1994).

Generalmente, los trabajos que hemos presentado toman como fuente de datos las respuestas del alumnado a cuestionarios muy conocidos por la investigación. Pero, dichas respuestas no siempre son analizadas teniendo en cuenta el contexto en el que se formulan y el formato de datos que es propio del contexto, factores que pueden ser influyentes en los resolutores (Ben-Zvi y Garfield, 2004). Mooney, Langrall y Nisbet (2006) examinan el rol que juega el conocimiento del contexto al interferir o ayudar a estudiantes de nivel medio en sus respuestas a problemas que plantean una comparación de conjuntos de datos. Demuestran que el uso del conocimiento del contexto varía y que se puede describir y clasificar. Recientemente, Orta, Sánchez y Altamirano (2015), al explorar el razonamiento de profesoras en formación acerca de la dispersión de los datos en problemas que involucran riesgos, concluyen que el contexto y el formato de presentación de los problemas ejercen mayor influencia en la resolución que la estructura del problema.

METODOLOGÍA

Muestra

El cuestionario fue cumplimentado por un total de 188 alumnos y alumnas (ver su distribución en Tabla 1). La muestra no fue seleccionada de forma intencionada. Participó todo el alumnado matriculado en los diferentes grupos que estaba presente en el momento de la administración del cuestionario. El alumnado de secundaria cursaba la opción B de 4º y "simultaneidad" cursaba simultáneamente Educación Infantil y Primaria (cursando 4º de Infantil y 3º de Primaria durante el desarrollo de la prueba).

Tabla 9. Distribución de la muestra por niveles educativos: Educación secundaria obligatoria, 4º curso, (ESO). Estudiantes de 2º grado de Educación Infantil (MI) y Primaria (MP), estudiantes de simultaneidad (MS), en EI y EP, y estudiantes que actualizan su diplomatura a grado (MA).

Nivel	*N*	%
ESO	72	38,30
MI	44	23,40
MP	35	18,62
MS	16	8,51
MA	21	11,17
Total	188	100,00

El alumnado no recibió instrucción específica para esta investigación sobre medidas de centralización y dispersión. Sus conocimientos previos correspondían a la formación recibida anteriormente en función de su etapa educativa. Los estudiantes de secundaria, los de MI y de MP no habían sido instruidos aún en contenidos de Estadística en el momento en el que se les administró el cuestionario. Solamente los cursaron los alumnos de MS dentro de la asignatura de Matemáticas para maestros.

Los datos fueron recogidos mediante la administración de un cuestionario constituido por 4 problemas, cada uno de ellos con dos apartados, completando un total de 8 ítems para los que el estudiante podía disponer de calculadora si lo estimaba oportuno. Se trata de problemas de respuesta abierta, creados por los investigadores. A lo largo de los problemas del cuestionario se planteaba la comparación de tres distribuciones, considerada como una competencia estadística básica. Se repiten las preguntas en contextos y formatos de datos diferentes; se podría decir que los problemas coincidían en el fondo, pero no en la forma. Este planteamiento respondía al diseño intencionado de los investigadores. Los problemas P1 y P2 se formularon en el contexto estadístico-social, descritos por dos variables cuantitativas, una discreta en el P1 y la otra continua en el P2. De igual modo hicimos con los problemas P3 y P4 formulados en el contexto estadístico-salud (ver anexo).

Una forma de investigar posibles influencias de las variables independientes de los problemas, contexto y formatos, sobre los resolutores, consiste en determinar las dificultades asociadas. Para medirlas seguimos la metodología propuesta por Carles, Cerdán, Huerta, Lonjedo y Edo (2009) para los problemas de probabilidad. Las dificultades consideradas, Dificultad apreciada del problema (DAP), Dificultad (global) y restringida del Problema (DP y DPR, respectivamente), Dificultad de la solución (DSP) y de su descripción (DDRES y DDRESC) son conceptualmente las mismas que allí y pretender medir el porcentaje de éxito como, por ejemplo, en: Dificultad del problema (DPR): $DPR = 100 - \left(\frac{resultado}{abordado}\right) \times 100$ fórmulas en las que las expresiones literales hablan del número de resolutores que abordan el problema, que dan un resultado o un número correcto. En todos los casos, la dificultad oscila entre 100, máxima dificultad y 0 nula.

Por otra parte, como los cálculos pueden estar basados o no en la utilización de medidas estadísticas, para que la respuesta del problema se considere correcta, es necesario acompañar el razonamiento de una medida de dispersión que permita asegurar que la reducción de los datos es la adecuada. Las respuestas que no usen un razonamiento basado en parámetros estadísticos se codificarán como NRE (No Razonamiento Estadístico). Las respuestas que usen un razonamiento basado en parámetros estadísticos que no sean centrales se codificarán como RENC (Razonamiento Estadístico No Central). Las respuestas que usen un razonamiento basado en medidas centrales con análisis de la dispersión de los datos se codificarán como RECC (Razonamiento Estadístico Central Completo). Como consecuencia de la experimentación, se ha observado que el único análisis de la dispersión que realiza el alumnado es el cálculo del rango que, en ocasiones, puede estar o no acompañado del reconocimiento de valores atípicos. En consecuencia, contemplaremos dentro de este tipo de respuestas aquellas que calculen la media aritmética con y sin el valor atípico para valorar la representatividad de la media. Daremos por correcta la respuesta en la que el alumnado haga referencia a la variabilidad, sin realizar cálculos de la dispersión de los datos, y acompañe la reflexión con el cálculo de la mediana o de la media aritmética (sin considerar la observación atípica para el caso de la media). Si no se realiza este análisis, se codificarán como RECI (Razonamiento Estadístico Central Incompleto).

RESULTADOS Y DISCUSIÓN

Sobre las dificultades de los problemas

La tabla siguiente (Tabla 2) muestra las dificultades asociadas a los problemas leídas globalmente.

Tabla 2. Dificultades globales de los problemas en %, sin considerar contextos y formatos, máxima 100.

Problema\Dificultad	DAP	DP	DPR	DSP	DDRES	DDRESC
P. 1.1	2,66	4,26	1,64	43,17	11,67	97,48
P. 1.2	4,79	6,38	1,68	88,27	32,39	97,48
P. 2.1	4,26	5,32	1,11	91,67	20,22	97,89
P. 2.2	20,21	20,74	0,67	99,33	44,30	98,80
P. 3.1	8,51	10,64	2,33	22,09	9,52	98,03
P. 3.2	11,17	11,70	0,60	20,96	28,31	97,48
P. 4.1	10,11	11,70	1,78	76,33	21,08	97,71
P. 4.2	28,19	31,38	4,44	97,04	42,64	98,65

El gráfico siguiente (Gráfico 1) permite una comparación de las dificultades dependiendo del problema considerado.

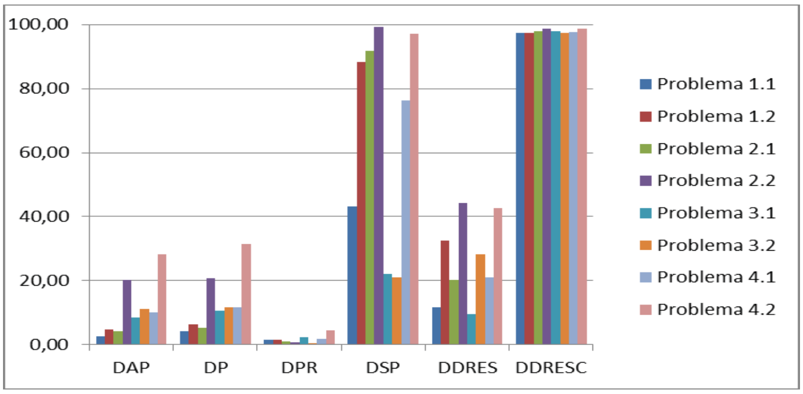

Figura 1. Dificultades globales de los problemas por tipo de dificultad

De él se desprende que, independientemente del problemas del que se trata, justificar una respuesta basándose en razonamientos estadísticos es una tarea casi imposible para la muestra de estudiantes considerada (DDRESC>97%), aunque estos lo intentan (DDRES<45%), intento que depende de lo que se pregunta en el problema (mayor en la cuestión 2 que en la 1). Esta dificultad se corresponde con la dificultad apreciada mayor en la tarea de ordenación de muestras por razones estadísticas que la tarea de escoger el mejor representante, destacando este hecho en el problema P2.2 y P4.2 (DAP>20%), cuyas preguntas hacen referencia al valor esperado. Por otra parte, si observamos aisladamente la capacidad de dar respuestas correctas a los problemas, podemos llegar a concluir, erróneamente, que el problema P3 es el menos difícil (DSP<22% aprox.). Sin embargo, si consideramos la solución numérica junto con la descripción de las respuestas, la percepción cambia. Esto es debido a que en el problema P3 el resolutor puede dar con el tratamiento más efectivo con diferentes cálculos erróneos como, por ejemplo, sumar los metros caminados por los niños y niñas. Es cierto que, en este problema, la media aritmética en el grupo 3 es superior a la media de los grupos 1 y 2. La media de este grupo es representativa y no posee observaciones atípicas. En consecuencia, el resolutor que ha respondido a partir del cálculo de la media aritmética proporciona una respuesta numérica correcta. No obstante, no se da por correcta la respuesta si no está acompañada de una medida de dispersión o de una reflexión respecto a la presencia de observaciones atípicas.

Los resultados muestran también que el problema P1 es el que presenta una menor dificultad apreciada para los resolutores (DAP< 5% para ambos ítems). Este problema está formulado en un contexto en el que se habla de ingresos anuales por familia y, en este tipo de situaciones, parece que todos podemos dar nuestra opinión, ya sea con fundamentos estadísticos o idiosincráticos. El segundo problema con una DAP menor ha sido el P2.1, que también está formulado en el mismo contexto. Sin embargo, la dificultad apreciada ha aumentado cuando hemos utilizado el término verbal "esperas". Este hecho se ha repetido cuando se ha vuelto a usar al formular el problema P4.2. Ya hemos mencionado que en apariencia el problema P3 parecía presentar menos dificultades a los estudiantes, pero si se realizaba un análisis exhaustivo de las dificultades del problema entonces esto no era tan evidente.

Representatividad de la media aritmética para el alumnado ante la presencia de valores atípicos

Los datos se presentan en términos relativos con respecto al alumnado que ha respondido a los diferentes problemas. En la tabla 3 puede observarse como el porcentaje mayor de alumnos que usan un razonamiento estadístico completo corresponde a los maestros de actualización (MA), seguido de primaria (MP) y secundaria (ESO), pero con valores tan bajos que no permiten obtener ninguna conclusión. El alumnado que ha realizado este tipo de razonamiento conoce que es preciso descartar la observación atípica antes de proceder al cálculo de la media aritmética, ya que ésta es muy sensible a los valores extremos. Tal como se afirma en Batanero, Godino y Navas (1997:3), "se trata de discriminar entre el simple conocimiento algorítmico de la fórmula de cálculo,…, de la comprensión relacional del concepto". El porcentaje global de respuestas estadísticamente correctas es muy bajo, 5 alumnos de un total de 188 han sido capaces de dar esta respuesta en el conjunto de la muestra considerada (2,66%). Los futuros maestros de infantil y de simultaneidad no proporcionan ninguna respuesta correcta. A la vez, son los estudiantes de simultaneidad quien mayoritariamente proporcionan razonamientos no estadísticos, después de MI y ESO. El alumnado de simultaneidad proviene del grado de maestro de infantil.

A excepción de un alumno de actualización, que ha utilizado la mediana como medida central, el resto de alumnado ha utilizado la media aritmética. Esto indica que consideran que la media es preferible a la mediana o moda como mejor representante de un conjunto de datos, aunque no han considerado necesario considerar y desechar el valor atípico en todos los casos, tal y como refleja la columna RECI. Por ello, más del 68% del alumnado de MI se ha quedado con el conocimiento algorítmico, siendo este el grupo más numeroso en el que se había elegido la media como mejor representante de los datos.

Tabla 3. Uso de parámetros estadísticos problema P1 (en%)

	NRE	RENC	RECC	RECI
ESO	78,46	0,00	1,54	20,00
MI	48,83	4,65	0,00	46,52
MP	28,57	0,00	2,86	68,57
MS	87,50	0,00	0,00	12,50
MA	42,86	9,52	14,29	33,33

Comparado con el problema P1, la preguntas planteada en P2 da lugar a una disminución en el porcentaje de alumnado que daba una respuesta correcta, incrementándose el número de resolutores que ofrece una respuesta basada en cálculos estadísticos centrales incompletos y disminuyendo el de los que no basaban su respuesta en razonamientos estadísticos (Tabla 4). En este problema, se pedía de forma explícita el tiempo medio por lo que era de esperar que se dieran estos resultados. De nuevo, la formulación de la pregunta influye en la resolución y se manifiesta como una variable de la tarea influyente.

Tabla 4. Uso de parámetros estadísticos problema P2 (en %)

	NRE	RENC	RECC	RECI
ESO	70,31	3,12	0,00	26,57
MI	30,23	0,00	0,00	69,77
MP	14,28	0,00	2,86	82,86
MS	73,33	0,00	0,00	26,67
MA	42,86	0,00	9,52	47,62

En la tabla siguiente (Tabla 5), observamos que siguen siendo contados los casos en los que se realiza una resolución basándose en cálculos estadísticos correctos, ahora en el problema 3.

Tabla 5. Uso de parámetros estadísticos problema P3 (en %)

	NRE	RENC	RECC	RECI
ESO	60,35	8,62	1,72	29,31
MI	33,33	5,13	0,00	61,54
MP	14,71	0,00	5,88	79,41
MS	68,75	0,00	0,00	31,25
MA	28,57	28,57	4,76	38,10

El reconocimiento de observaciones atípicas por parte del alumnado es una condición necesaria para que se conciencie de su influencia en el cálculo de la media aritmética. Sin embargo, no es una condición suficiente de cara a la búsqueda de medidas alternativas que representen la reducción de datos de forma correcta. A continuación mostramos dos respuestas de alumnos que involucran la presencia de observaciones atípicas y comentamos cómo evalúan estas observaciones.

Colectivo	Dioptrías										
Doctores	3	0	0	2	1.75	0.5	0	2	2.25	0.5	12
Graduados	0	2	0	0	0	1.5	1,6	1.25	2.25	0.75	9'35
Sin estudios superiores	2	2	0	1	0.5	0.25	1.5	1	2.75	3.5	14'5

Figura 2. Modificación de datos problema P4 del alumno 151

El alumno observa en la tabla el valor 16 (Figura1) y considera que es un valor extraño en relación con el resto de la muestra. No considera la posibilidad de la existencia de observaciones atípicas, directamente decide modificar el valor extraño reconvirtiéndolo en 1.6, un valor mucho más normal dentro del conjunto de valores observados en el contexto de las dioptrías.

1 Público : 1250 : 10 = 125 h de media
Privado : 1340 : 10 = 134 h de media
Concertado : 1250 : 10 = 125 h. " "

En el privado estan más tiempo conectados a internet ya que la media es 134 h. Pero pienso que por que un niño/a sobra pasa las horas.

Figura 3. Respuesta al problema P2.1 del alumno 160

La respuesta refleja que el alumno reconoce que la media se ve afectada por la observación atípica, ya que afirma que la media es 134 h porque un niño/a sobrepasa las horas. Sin embargo, no modifica los cálculos ni su respuesta, que la acepta matizada por el comentario. Puede que la respuesta le acomode mejor de esta manera

Los estudiantes no siempre reconocen la función de las observaciones atípicas ya que, pese a ser reconocidas, no tiene en cuenta cómo afecta su consideración al cálculo de la media aritmética en cualquier caso. Es más, los estudiantes de MI no las han descartado en ningún problema y los de MS únicamente en el problema P4. Destaca los componentes de MA al descartar la observación atípica en todos los problemas, aunque en pequeños porcentajes. En particular, ha habido un alumno de actualización que ha utilizado la mediana y otro medias recortadas para responder a los problemas ante la presencia de observaciones atípicas. Así, el nivel SOLO más abstracto, considerado en Groth y Bergner (2006), puede ser considerado para muy pocos estudiantes de nuestras muestras, pues en general no son capaces de observar el efecto que producen los valores atípicos sobre la media y su representatividad. En comparación con los resultados obtenidos por Estrada, Batanero y Fortuny (2003), obtenemos peores resultados al sobrepasar el 50% de futuros maestros que no descartan el valor atípico para el cálculo de la media aritmética en los problemas que hemos considerado. Llegan a alcanzar, dependiendo del problema y del colectivo, el 100% que ni siquiera reconoce la presencia de una observación atípica, rara o extraña.

CONCLUSIONES

Cuando los problemas se formulan en la manera en la que se han formulado en nuestro cuestionario, conteniendo valores atípicos y tablas con tamaño de colecciones de datos diferentes, la influencia, tanto del formato de los datos como del contexto en el que se formulan, se muestra como un factor que condiciona las dificultades de los problemas. Hemos visto que los problemas que se enunciaban en un contexto estadístico-social presentaban menos dificultad para el alumnado a la hora de abordarlos y dar una respuesta que los enunciados en un contexto estadístico-salud a pesar de que los problemas planteados presentaban la misma complejidad en sus cálculos. Además, los términos en que se planteaba la pregunta también influían en la dificultad del problema y, en particular, en el hecho de tener que proporcionar un valor esperado. El reconocimiento de una observación atípica permitía al alumnado dar una respuesta correcta al problema; este reconocimiento se basaba en su conocimiento del contexto ya que, al no calcular los cuartiles ni realizar gráficos de cajas y bigotes, no tenía herramientas para caracterizar las observaciones atípicas. Creemos que es muy significativo que los estudiantes hayan aceptado el problema y tratado de dar una respuesta al mismo, aunque basándose casi exclusivamente en su conocimiento sobre el contexto en el que estaba formulado, puesto que el proceso de resolución implicaba la identificación de una observación atípica y su influencia en la solución ha sido determinante.

Aunque el cálculo de las medidas centrales no presentaba dificultad, la mayoría de los estudiantes se limitaron a calcular la media aritmética sin tener en cuenta si era representativa o no. Salvo contadas excepciones, no calcularon otras medidas centrales y, en caso de hacerlo, correspondió a la moda. Únicamente un alumno calculó la mediana tras observar la presencia de una observación atípica que determinó a través del contexto.

Además, hemos encontrado las mismas dificultades en la comprensión de las medidas centrales que Batanero, Godino y Navas (1997) en relación con el tratamiento de los ceros y valores atípicos en el cálculo de medidas centrales y el uso de promedios en la comparación de distribuciones. El alumnado, en general, demostró no conocer que la media aritmética debe ir acompañada de la desviación típica para valorar si es adecuada o no a la hora de dar respuesta a la comparación de distribuciones ni tampoco se planteó el uso de medidas de tendencia central más allá de la media aritmética. Esto es debido a la falta de uso de la idea de dispersión, ya que aunque algunos calcularon el rango de forma correcta, no llegaron a captar el significado de variabilidad ni su utilidad en la comparación de dos distribuciones. En consonancia con la investigación realizada con futuros maestros de primaria de Ruiz, Arteaga y Batanero (2009), nuestros resultados sugieren que, para el alumnado, las medidas centrales resultan más intuitivas que las medidas de dispersión y no las considera necesarias para comparar distribuciones. Hemos de mencionar que únicamente 14 alumnos de nuestra investigación, de un total de 188, afirmaron conocer el uso del modo estadístico en la calculadora.

Resumiendo, si la esencia del razonamiento estadístico es para Wild y Pfannkuch (1999) el concepto de distribución, y un requisito para comprender la distribución es la variabilidad, los resultados de nuestra investigación indican que ni tan siquiera los maestros en activo poseen este razonamiento. Concluimos que el contexto y el formato de los datos deberían tenerse en cuenta a la hora de diseñar secuencias de enseñanza o trabajos de investigación relacionados con la comparación de conjuntos de datos a tenor de lo expuesto en este trabajo y, rogamos encarecidamente, que la formación de los maestros pase por una atención a estos problemas. El alumnado, en sus primeros años de escolaridad, apenas adquiere una perspectiva algorítmica de las principales medidas estadísticas, manifestando grandes dificultades en interpretar los resultados obtenidos. El recurso excesivo de fórmulas estadísticas y la utilización de problemas descontextualizados tiene como resultado "alumnos mal preparados para el estudio de la estadística, a nivel superior, y adultos estadísticamente analfabetos" (Batanero y Díaz, 2010: 6). Un enfoque alternativo a la enseñanza actual debería estar basado en el análisis de datos como medio para la resolución de problemas como los que hemos utilizado en este estudio.

Referencias

Anderson, D.E.; Sweeney, D.J. y Williams, T. A. (2010). *Estadística para Administración y Economía* (10ª ed.). México: Thomson.

Batanero, C. y Díaz, C. (2010). Training teachers to teach statistics: what can we learn from research? *Statistique et enseignement*, *1*(1), 5-20. Recuperado el 11 de junio, de http://statistique-etenseignement.fr/ojs/

Batanero, C.; Godino, J. D. y Navas, F. (1997). Concepciones de maestros de primaria en formación sobre los promedios. Recuperado el 10 de junio de 2014, de http://www.ugr.es/~batanero/pages/ARTICULOS/Logse.pdf

Batanero, C.; Godino, J. D.; Vallecillos, A.; Green, D.R. y Holmes, P. (1994). Errors and difficulties in understanding elementary statistical concepts. *International Journal of Mathematical Education in Science and Technology*, 25:4, 527-547. DOI: 10.1080/0020739940250406.

Ben-Zvi, D. y Garfield, J. (2004). Statistical Literacy, Reasoning, and Thinking: Goals, Definitions, and Challenges. In D. Ben-Zvi, y J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 3-15). Dordrecht: Kluwer.

Carles, M.; Cerdán, F.; Huerta, M. P.; Lonjedo, M.A. y Edo, P. (2009). Influencia de la estructura y del contexto en las dificultades de los problemas de probabilidad condicional de nivel N0. Un estudio exploratorio con estudiantes sin enseñanza previa. En M.J. González, M.T. González y J. Murillo (Eds.), *Investigación en Educación Matemática XIII* (pp. 173-185). Santander: SEIEM.

Devore, J. (2005). Probabilidad y Estadística para ingeniería y ciencias. México D.F.: Thomson

Estrada, A.; Batanero, C. y Fortuny, J.M. (2003). Dificultades de los profesores en formación en conceptos estadísticos elementales. En E. Castro (ed.) *Investigación en educación matemática: Séptimo Simposio de la Sociedad Española de Investigación en Educación Matemática* (pp. 201-212). Granada: SEIEM.

Groth, R. E. y Bergner, J. A. (2006). Preservice Elementary Teachers' Conceptual and Procedural Knowledge of Mean, Median, and Mode. *Mathematical Thinking and Learning, 8*(1), 37-63, DOI: 10.1207/s15327833mtl0801_3

Mayén, S.; Díaz, C. y Batanero, C. (2009). Conflictos semióticos de estudiantes con el concepto de mediana. *Statistics Education Research Journal, 8*(2), 74-93.

Mooney, E.; Langrall, C. y Nisbet, S. (2006). Developing a model to describe the use of context knowledge in data explorations. Recuperado el 12 de julio de 2015, de http://www.researchgate.net/publication/29462444_Developing_a_Model_to_Describe_the_Use_of_Context_Knowledge_in_Data_Explorations

Orta, J.A.; Sánchez, E. y Altamirano, J. A. (2015). Interpretación de la dispersión de datos en contexto de riesgo por profesoras en formación. En C. Fernández, M. Molina y N. Planas (Eds.) *Investigación en Educación Matemática XIX* (pp. 441-450). Alicante: SEIEM

Ortiz, J.J. y Font, V. (2014). Pre-service teachers' common content knowledge regarding the arithmetic mean. *REDIMAT, 3*(3), 192-219. Recuperado el 10 de junio de 2015, de http://dx.doi.org/10.4471/redimat.2014.51

Reading, C. y Pegg, J. (1996). Exploring understanding of data reduction. En L. Puig y A. Gutiérrez (Eds.), *Proceedings of the 20th Conference of the International group for the Psychology of Mathematics Education,* (pp. 187-194). Universidad de Valencia, España.

Ruiz, B.; Arteaga, P. y Batanero, C. (2009). Competencias de futuros profesores en la comparación de datos. En L. Serrano (Ed.), *Tendencias actuales de la investigación en educación estocástica* (pp. 57-74). Málaga: Gráficas San Pancracio.

Wild, C. y Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review,* 67(3), 223-265.

Anexo. Enunciado de los problemas. Un ejemplo

Problema 1

En un edificio A, de la periferia de Valencia, viven 10 familias cuyos ingresos anuales en miles de euros son: 220; 17.5; 22; 19; 16.5; 20.5; 21; 19.5; 24; 19. En otro edificio B, del centro de Valencia, viven 12 familias cuyos ingresos anuales en miles de euros son: 43; 44.5; 37; 39; 40; 43; 37; 33; 42; 37; 46; 37. En un tercer edificio C, de un pueblo del interior, viven 10 familias cuyos ingresos anuales en euros son: 22000, 19000, 18000, 21000, 21000, 20000, 24000, 24000, 25000, 18000.

1. ¿Qué edificio crees que sería el mejor representante de las familias con mayor poder adquisitivo? ¿Por qué?

2. Si tuvieras que ordenar los edificios anteriores de mayor a menor poder adquisitivo, ¿cómo lo harías? ¿En qué orden quedarían? Justifica la ordenación.

Problema 3

En una clínica infantil para niños y niñas con problemas para caminar, se ha creado tres grupos. A cada grupo se le ha aplicado una técnica diferente de aprendizaje. Al final del tratamiento, se ha realizado una prueba consistente en el número de metros que cada niño anda, seguido y sin caerse. Los resultados para los tres grupos han sido los siguientes:

Grupo	Metros caminados por cada niño/a														
1	1	2	1	2	2	4	1	1	3	2	1	4	2		
2	1	1	2	1	3	2	15	4	1	2	3	4	1	1	
3	2	2	3	1	4	5	2	2	3	3	1	3	5	4	5

1. ¿En qué grupo consideras que el tratamiento ha sido más efectivo? ¿Por qué?

2. Si tuvieras que ordenar los tratamientos de mayor a menor efectividad, ¿cómo lo harías? ¿En qué orden quedarían? Argumenta tu respuesta.