

Determinação da mediana em contexto tabelar e gráfico

Determination of median in tabular and graphic context

Maria José Carvalho¹

<https://orcid.org/0000-0003-2045-9612>

José António Fernandes²

<https://orcid.org/0000-0003-2015-160X>

Adelaide Freitas³

<https://orcid.org/0000-0002-4685-1615>

Resumo: Neste estudo, cujo objetivo é averiguar a influência do contexto dos dados na determinação da mediana, estudam-se as resoluções de 332 alunos portugueses do 8.º ano, a duas questões, uma com os dados organizados numa tabela de frequências e outra num gráfico de barras. Seguindo uma metodologia mista, quantitativa e qualitativa, realiza-se uma análise de frequências de quatro categorias ordinais de respostas (corretas, parcialmente corretas, incorretas e não respostas) e uma análise semiótica das resoluções aplicando o Enfoque Ontossemiótico do conhecimento e ensino da Matemática na identificação de objetos e processos matemáticos que intervêm nas resoluções. Em geral, os alunos revelaram maior propensão para não responderem à determinação da mediana em contexto gráfico, mas os que responderam tendem a determinar a mediana com menor dificuldade. Os conflitos semióticos identificados aparentam não estar dependentes da forma de apresentação dos dados, embora ocorra maior número de respostas não corretas no contexto tabelar do que no contexto gráfico.

Palavras-chave: Ensino de matemática. Mediana. Representação de dados. Ensino básico.

Abstract: This study analyzes the responses given by 332 8th grade Portuguese students to two questions, one with data set defined by a frequency table and the other defined by a bar graph. The goal is the evaluation of the effect of the data representation on the calculation of the median. Using a combined methodology, a frequency analysis of students' response types and a semiotic analysis of responses, based on onto-semiotic approach of knowledge and mathematics education, were applied. The semiotic analysis' main goal was the identification of objects and mathematical processes, which can characterize semiotic conflicts implied from the student responses. Generally, students revealed greater tendency to not respond to the graphic determination of the median, however, those who do answer, tend to obtain the median with less difficulty in a graphic context. Semiotic conflicts seem not to depend on the presentation of the data set, although there are more answers that are incorrect in the tabular context than in the graphic context.

Keywords: Mathematics teaching. Median. Data representation. Basic education.

¹ Universidade de Aveiro, Departamento de Educação e Psicologia, Aveiro, Portugal.

E-mail: <maria.carvalho@ua.pt>.

² Universidade do Minho, Instituto de Educação, Braga, Portugal.

³ Universidade de Aveiro, Departamento de Matemática, Aveiro, Portugal.

Introdução

O estudo da Estatística, vista como área científica autónoma, com métodos e raciocínios próprios (MOORE, 1991), ocorre em todos os nove anos de escolaridade do ensino básico em Portugal, e visa desenvolver nos alunos capacidades de interpretação, avaliação crítica e comunicação de informação estatística (OLIVEIRA; HENRIQUES, 2014). A sua constante utilização nos variados meios de comunicação social, especialmente de tabelas e gráficos, torna a aprendizagem da Estatística essencial à vida do quotidiano, sendo fundamental que os alunos aprendam a ler e interpretar informação estatística.

Contudo, apesar de a Estatística ser um excelente meio para organizar e comunicar informação, os alunos sentem dificuldades na representação e interpretação estatística, designadamente: na determinação de medidas estatísticas, na elaboração de representações estatísticas e na leitura e interpretação de informação estatística (ESPINEL et al., 2009; MORAIS; FERNANDES, 2011).

Diversos trabalhos de investigação abordam dificuldades dos alunos na representação gráfica e tabelar e na interpretação de enunciados envolvendo a noção de média e mediana. Tais dificuldades estão patentes, por exemplo, nas provas de avaliação nacional (SOUSA et al., 2010), enquanto alguns estudos descrevem o porquê dessas dificuldades (MAYÉN; ORTIZ; DÍAZ, 2009; PONTE; FONSECA, 2001) e outros ainda identificam fatores que influenciam a compreensão da leitura de gráficos e tabelas (FRIEL; CURCIO; BRIGHT, 2001). Relativamente à determinação das medidas de localização central (moda, média e mediana), embora estes conceitos sejam aparentemente simples, os alunos cometem frequentemente diversos erros. Por exemplo, na mediana, não efetuam a ordenação prévia dos dados e confundem-na com a moda ou até com a média (BATANERO, 2000). Batanero (2000) refere que o facto de os alunos tenderem a usar um único procedimento na resolução de problemas e em cálculos dificulta, em particular, a determinação da mediana, porquanto ela depende do número de dados ser par ou ímpar. Para Batanero (2000), a compreensão dos conceitos (conhecimento das definições e propriedades) inclui a resolução dos problemas em que estes devem ser aplicados, reforçando a importância de uma compreensão relacional em detrimento de uma compreensão instrumental.

Neste trabalho, tendo por objetivo compreender se a forma de representação dos dados interfere na resolução dos alunos e conduz a diferentes tipos de conflitos semióticos, estudam-se as respostas e as não respostas de 332 alunos do ensino básico de Portugal a duas questões sobre o cálculo da mediana, em que numa questão os dados são apresentados em tabela e na outra, com enunciado similar, são fornecidos num gráfico de barras, e recorre-se ao modelo teórico do Enfoque Ontossemiótico (EOS) do conhecimento e do ensino da Matemática (GODINO; BATANERO, 1994; GODINO; BATANERO; FONT, 2007) para efetuar análises semióticas das resoluções apresentadas pelos alunos.

Referencial teórico

Antecedentes

Na sociedade atual, a Estatística vem ganhando importância, sendo aplicada nos diversos domínios do saber, na gestão de informação e na vida pessoal e profissional dos cidadãos.

Contudo, essa importância contrasta com dificuldades que os alunos sentem neste âmbito. Garfield e Ahlgren (1988) consideram que os alunos entendem a Estatística à semelhança da Matemática, o que os leva a procurar uma solução única, resultando daí dificuldades dos alunos, pois a natureza da Estatística envolve alguma incerteza.

Vários estudos, desde o ensino básico ao universitário, relatam dificuldades dos alunos neste tema, nomeadamente na leitura, análise, elaboração e interpretação de gráficos ou tabelas e na utilização das medidas de localização (BATANERO; GODINO; NAVAS, 1997; BOA-VENTURA, 2003; MARQUES; GUIMARÃES; GITIRANA, 2011; RIBEIRO; CORREIA; FERNANDES, 2013). Na construção de um gráfico, Li e Shen (1992) verificaram que muitos alunos do ensino secundário não sentem necessidade de selecionar o tipo de gráfico adequado à situação, e, por vezes, omitem as escalas dos eixos (horizontal, vertical ou em ambos), não especificam a origem das coordenadas, não proporcionam divisões suficientes nas escalas dos eixos e não lhes atribuem nomes. No estudo realizado por Arteaga (2008) constata-se que apenas um terço dos futuros professores do ensino primário foi capaz de visualizar a posição central e a dispersão a partir do gráfico, e um número menor de alunos (13%) conseguiu expressar uma conclusão correta sobre o problema proposto.

Segundo Friel, Curcio e Bright (2001), a compreensão gráfica inclui componentes aritméticas e não aritméticas, o que exige conhecimentos numéricos e, sobretudo a compreensão das operações de medição e de classificação. Nessa mesma linha, Arteaga et al. (2011) referem ser prioritário desenvolver o sentido numérico para evitar erros na interpretação de gráficos e tabelas estatísticas. Por outro lado, também têm sido detetados erros conceptuais e dificuldades de aplicação dos conhecimentos sobre as medidas de tendência central (BARROS, 2004; BATANERO; GODINO; NAVAS, 1997). Em Barr (1980) chama-se a atenção para as dificuldades de compreensão de alunos, dos 17 aos 21 anos, sobre a mediana, em dados não organizados, pois no seu estudo 49% dos alunos responderam incorretamente ao não ordenarem previamente os dados.

Relativamente aos conceitos de moda, média e mediana, num estudo realizado com alunos do 7.º ano divididos em dois grupos, um submetido a tarefas com dados organizados em tabela e outro submetido a tarefas com dados organizados em gráfico, Carvalho (1996) constatou que, no cálculo da mediana, os alunos revelaram muitas dificuldades em determiná-la com sucesso (apenas 12% o fizeram corretamente), quer em suporte gráfico quer em suporte tabelar. Verificou ainda que os alunos não têm em conta as frequências absolutas e não ordenam previamente os dados, calculam o valor central das frequências absolutas ordenadas por ordem crescente, erro também referido por Mayén et al. (2007), calculam a moda em vez da mediana e equivocam-se ao determinar o valor central. Carvalho e César (2000), num estudo realizado com alunos do 7.º ano, envolvendo os conceitos de média e mediana, com dados não organizados e organizados em gráfico de barras e tabelas de frequências, verificaram que a maioria dos alunos aplica, com sucesso, os procedimentos algorítmicos de cálculo da média e da mediana; no entanto, ao selecionarem a medida que melhor representa a distribuição, os argumentos usados pelos alunos indiciam maiores dificuldades no caso da mediana (56%) do que no caso da média (27%). Estas autoras sugerem que a melhor compreensão da média resulta da maior frequência com que é utilizada nos mais variados contextos sociais, o que já não acontece com a mediana.

Num estudo realizado por Barros (2004), com futuros professores do 1.º ao 6.º ano, a autora constatou que, na identificação da mediana de uma distribuição dada sob a forma de gráfico de barras, os erros detetados são relativos a metade da amplitude dos dados (10,8%),

ao cálculo da moda em vez da mediana (2,7%), à identificação da mediana com 50% dos dados (5,4%), o que significa confundir a mediana com a sua localização, e à consideração do 0 (zero) como elemento neutro (5,4%). Mayén, Díaz e Batanero (2009), num estudo comparativo, com dados não organizados, entre alunos do secundário (idades 14-15 anos) e alunos universitários de um mesmo curso (idades 17-18 anos), verificaram que os alunos do ensino secundário exibiram melhor desempenho na determinação da mediana e os alunos do ensino superior foram mais assertivos na escolha de uma medida de tendência central (mediana ou média) que melhor representava a distribuição. Sobre a escolha de uma medida, estudos realizados com futuros professores (GROTH; BERGNER, 2006) e professores (JACOBBE, 2008) revelaram dificuldades em responder corretamente e em fundamentar as suas opções. Groth e Bergner (2006) e Jacobbe (2008) deduzem que a escolha de uma medida que melhor represente uma dada distribuição requer conhecimentos sedimentados sobre as medidas e a sua adequação a cada situação, enquanto a determinação da mediana apenas se baseia no conhecimento técnico do seu algoritmo de cálculo.

Ribeiro (2011) realizou um estudo sobre as estratégias utilizadas por 19 alunos do 7.º ano na realização de tarefas, envolvendo medidas de tendência central, com dados não organizados, organizados em tabela de frequência e em diagrama de caule-e-folhas. A autora constatou que a mediana foi a medida de tendência central com maior percentagem de respostas incorretas em todas as formas de representação dos dados, sendo essa percentagem maior no diagrama de caule-e-folhas (47,4%), seguindo-se na tabela de frequências (42,1%) e, menos ainda, nos dados não organizados (36,9%).

Em termos globais e no que concerne aos estudos realizados por Barros (2004), Batanero (2000), Boaventura (2003), Carvalho (2004), Ribeiro, Correia e Fernandes (2013), constata-se que a medida de tendência central que origina mais dificuldades aos alunos é a mediana. Destes estudos são ainda identificadas outras dificuldades, tais como determinar um conjunto de dados, conhecidas a média, mediana e/ou moda, confundir a mediana e a moda, aplicar as propriedades da média e da mediana, determinar a média em vez da mediana e atribuir significados às medidas de localização. Estes estudos revelam também que algumas das dificuldades demonstradas pelos alunos e os erros que cometem são comuns a alunos de vários níveis de escolaridade (FERNANDES; CARVALHO; CORREIA, 2011) e que eles ocorrem em vários conteúdos do tema Estatística, o que significa que o progresso no ensino não foi suficiente para vencer tais dificuldades.

Marco teórico

Para analisar as resoluções dos alunos optou-se, neste estudo, pelo modelo teórico do Enfoque Ontossemiótico (EOS) do conhecimento e ensino da Matemática (GODINO; BATANERO, 1994; GODINO; BATANERO; FONT, 2007). Godino, Batanero e Font (2007) consideram a configuração epistémica, conjunto de objetos matemáticos envolvidos na resolução de situações-problema, como uma dimensão importante do modelo. Assim, o ponto de partida do EOS é a formulação de uma ontologia de objetos matemáticos primários, que permitem uma análise detalhada e abrangente das práticas mobilizadas nos processos de resolução de problemas. Segundo Godino e seus colaboradores, na análise da atividade matemática é importante

ter-se em consideração seis tipos de objetos primários: *situações-problema*, que são aplicações extra matemáticas, exercícios, problemas, ações que induzem uma atividade matemática; *linguagens*, que são termos, expressões, notações, gráficos que se utilizam para representar os dados de um problema, as operações que se efetuam, os objetos matemáticos que se usam e a solução encontrada em seus diversos registros; *conceitos e definições*, que se referem a formulações introduzidas mediante definições e descrições, as quais constituem práticas realizadas pelos alunos para resolver um problema matemático com uso implícito ou explícito de objetos matemáticos, que o aluno tem de lembrar e aplicar; *propriedades/proposições*, que são enunciados sobre relações ou propriedades dos conceitos que se utilizam para resolver problemas matemáticos; *procedimentos*, que se referem a algoritmos, operações, técnicas de cálculo que os alunos aplicam para a resolução dos problemas; e *argumentos*, que são enunciados usados nas respostas para validar ou explicar as proposições e procedimentos ou a solução dos problemas, os quais podem ser dedutivos, indutivos, formais ou informais.

Os seis tipos de objetos primários relacionam-se originando configurações de objetos e diferentes processos matemáticos, descritos em facetas duais: *pessoal-institucional*, consoante o significado emerge da prática pessoal ou de um grupo de pessoas que partilham o mesmo tipo de problemas, ou seja, uma instituição; *ostensivo-não ostensivo*, conforme o objeto é público, pode ser mostrado a outro ou materializado de alguma forma (associado a notações, símbolos, gráficos), ou é pensado, imaginado por um sujeito, ou está implícito no discurso matemático; *extensivo-intensivo* (particular-geral), relativo à generalização de resultados numa atividade matemática, permitindo centrar a atenção na dialética entre o particular e o geral; *unitário-sistêmico*, em certas circunstâncias os objetos matemáticos participam como entidades unitárias, enquanto noutras devem ser tomados como sistemas para que se possa proceder ao seu estudo; *expressão-conteúdo* (antecedente e conseqüente de qualquer função semiótica), referente à atividade matemática, aos processos de construção e ao uso dos objetos matemáticos que se caracterizam por serem fundamentalmente relacionais, colocando os distintos objetos em relação uns com os outros. Esta relação estabelece-se por meio de funções semióticas, entendidas como uma relação entre um antecedente (expressão, designação ou nome) e um conseqüente (conteúdo, designado ou ente matemático) instituído por um sujeito (pessoa ou instituição) de acordo com determinado critério ou código de correspondência. Em relação às funções semióticas, elas são entendidas como relações entre conjuntos, envolvendo três componentes: *expressão*, que constitui o objeto inicial ou significante; *conteúdo*, que é o objeto final ou significado; e *regra de correspondência*, que é o código interpretativo que regula a relação entre a expressão e o conteúdo. Esta dualidade *expressão-conteúdo* possibilita análises semióticas *a priori* (etapa prévia de análise didático-matemática) e *a posteriori* (análise das produções escritas dos alunos numa tarefa matemática).

O presente estudo incide sobre a determinação da mediana em dois contextos distintos: tabelar e gráfico. O estudo tem por base a produção escrita nos dois contextos por um mesmo conjunto de alunos (significado pessoal), comparando-os com os significados institucionais de referência. Desta comparação, a verificação de discrepâncias entre esses significados (significado institucional *versus* significado pessoal) (GODINO; BATANERO, 1994) leva à identificação de conflitos semióticos em cada um dos dois contextos e, adicionalmente, a reconhecer se um tipo de conflito é produzido por um mesmo sujeito em ambos os contextos ou em apenas um deles. A discordância entre os significados atribuídos a uma expressão por dois sujeitos (pessoas ou instituições) em interação comunicativa, designado por *conflito semiótico*, é útil para a análise

na produção e comunicação matemática. Se a disparidade se produz entre significados institucionais, falamos de conflitos semióticos do tipo epistémico; se ela se produz entre práticas que formam o significado pessoal de um mesmo sujeito, expressa conflitos semióticos do tipo cognitivo (GODINO; BATANERO; FONT, 2007).

Metodologia

O objetivo deste trabalho consistiu em averiguar o efeito da representação dos dados (tabelar e gráfica) no cálculo da mediana, quer em termos dos tipos de respostas, quer na identificação de possíveis conflitos semióticos. Para tal, foi efetuada uma análise semiótica das respostas de 332 alunos do 8.º ano, antes da leção do tema de Estatística desse ano escolar, na resolução de duas questões envolvendo a noção de mediana, uma com os dados apresentados sob a forma de tabela (questão 1) e outra sob a forma de gráfico de barras (questão 2), respetivamente, conforme ilustra o Quadro 1.

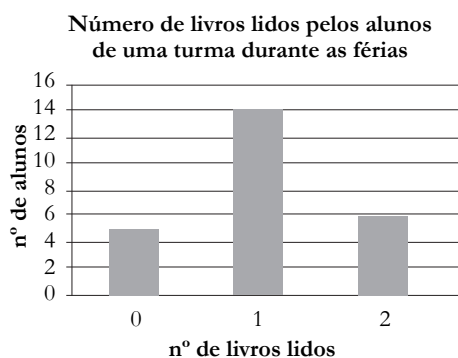
Quadro 1. Enunciado das duas questões propostas aos alunos

1. A tabela seguinte apresenta a distribuição das idades dos alunos de uma turma do 8.º ano:

Idade	13	14	15	16
N.º de alunos	6	4	5	5

Qual a mediana das idades dos alunos da turma? Justifica.

2. No gráfico seguinte está representada a distribuição do número de livros lidos durante as férias pelos alunos de uma turma.



Qual a mediana do número de livros lidos pelos alunos nas férias? Justifica.

Fonte: elaborado pelos autores.

As duas questões propostas enquadram-se nas metas curriculares de Matemática (PORTUGAL, 2013) e fazem parte de um teste diagnóstico, constituído por seis questões sobre Funções e Estatística, que foi aplicado no ano letivo de 2014-2015 a dezoito turmas de dois Agrupamentos de Escolas do distrito do Porto, pelos seus professores titulares, numa aula de Matemática. Dos alunos que participaram no estudo, 189 eram do sexo masculino, tinham idades compreendidas entre os 12 e os 18 anos e com média de 14 anos. As ausências de resposta dos 332 alunos a cada uma das questões foram contabilizadas e categorizadas como “não resposta” (NR). As respostas efetivas foram alvo de análise semiótica, classificadas em corretas (C), parcialmente corretas (PC) e incorretas (I), e foi construída uma tabela de contingência de 4x4 (Tabela 1) destas categorias de respostas e não respostas às duas questões. A partir das respostas não corretas (constituídas pelas respostas PC e I) foram categorizados e quantificados os conflitos semióticos reconhecidos em cada uma das questões, bem como as respostas que exibiram o mesmo conflito nas duas questões simultaneamente.

Assumindo que as respostas dos 332 alunos podem ser representativas do universo de respostas, efetuaram-se testes estatísticos (z-testes) na comparação de duas proporções (quando o número de alunos envolvidos no cálculo de tais proporções é superior a 30) e o teste de McNemar em tabelas 2x2 ou o teste de homogeneidade marginal em tabelas $k \times k$ ($k > 2$) para avaliar se há mudança na categoria das respostas dos alunos entre questões. Considerou-se haver evidência estatisticamente significativa quando o valor de prova p do teste correspondente é inferior ao nível de significância $\alpha=0,05$ (MURTEIRA et al., 2010). Assim, este trabalho insere-se numa perspetiva quantitativa quando se trata de analisar os diferentes tipos de respostas nas duas questões e quando se comparam essas respostas nos dois contextos, tabelar *versus* gráfico, e numa perspetiva qualitativa quando se procede à análise semiótica das respostas, portanto numa metodologia mista em termos globais. A pesquisa quantitativa permitiu comparar percentagens de categorias de respostas (C, PC, I e NR) e percentagens de categorias de conflitos semióticos entre as duas questões, tendo em vista confrontar o desempenho dos alunos na determinação da mediana face a dados sob representação tabelar *versus* representação gráfica. A análise qualitativa focou-se na análise semiótica das respostas dos alunos e permitiu a identificação de conflitos semióticos inerentes às respostas não corretas. Para tal, através de um processo de comparação de respostas semelhantes entre si, com recurso ao EOS do conhecimento e do ensino da Matemática (GODINO; BATANERO; FONT, 2007), desenvolveu-se uma categorização de tipos de conflitos, cujas categorias são apresentadas na próxima secção.

Apresentação de resultados

Tipos de respostas

Na Tabela 1 (tabela de contingência) apresentam-se as contagens e percentagens dos tipos de respostas (C, PC, I, NR) observadas nas questões 1 e 2, onde se conclui pela não existência de mudança significativa de categoria de resposta entre as duas questões ($p=0,132$).

Tabela 1. Frequências absolutas (percentagens) dos tipos de respostas nas questões 1 e 2

Questão 1	Questão 2				Total
	C	PC	I	NR	
C	20 (6,0)	8 (2,4)	0 (0,0)	6 (1,8)	34 (10,2)
PC	13 (3,9)	12 (3,6)	5 (1,5)	6 (1,8)	36 (10,8)
I	6 (1,8)	41 (12,3)	44 (13,3)	42 (12,7)	133 (40,1)
NR	1 (0,3)	18 (5,5)	6 (1,8)	104 (31,3)	129 (38,9)
Total	40 (12,0)	79 (23,8)	55 (16,6)	158 (47,6)	332

Fonte: elaborada pelos autores a partir dos dados da pesquisa.

Globalmente, na questão 1 predominam as respostas incorretas (40,1%) e na questão 2 prevalece a não resposta (47,6%). Apenas 6% determinaram corretamente a mediana em ambas as questões 1 e 2, sendo a percentagem de respostas corretas ligeiramente superior na questão 2 (12,0%) do que na questão 1 (10,2%), não havendo, porém, evidência estatística de diferenças significativas entre estas duas percentagens ($p=0,553$). Por fim, 31,3% dos alunos não responderam a ambas as questões, sendo as propensões para não responder significativamente diferentes nas duas questões (questão 1: 38,9%; questão 2: 47,6%; $p=0,023$), o que reforça a grande dificuldade dos alunos no cálculo da mediana em ambas as situações, maior em contexto gráfico.

Extraída da Tabela 1 a subtabela de contingência de 2x2 relativa às respostas de tipo PC e I, conclui-se pela existência de mudança significativa nas categorias das respostas não corretas, sendo que as percentagens de respostas parcialmente corretas são significativamente diferentes nas duas questões (questão 1: 10,8%; questão 2: 23,8%; $p=0,000$), assim como as percentagens de respostas incorretas (questão 1: 40,1%; questão 2: 16,6%; $p=0,000$). Observa-se que quando os dados estão sob a forma tabelar surge uma menor percentagem de respostas parcialmente corretas e uma maior percentagem de respostas incorretas. Agregando as respostas do tipo PC e I numa só categoria (respostas não corretas), constata-se que as proporções de ocorrência de respostas não corretas são significativamente diferentes nas duas questões (questão 1: 50,9%; questão 2: 40,4%; $p=0,006$), sendo maior na questão 1, onde os dados estão organizados em tabela.

Análise semiótica das respostas

Começou-se por considerar as respostas corretas e, a seguir, as respostas não corretas, para identificar possíveis conflitos semióticos nas respostas às duas questões propostas.

Análise semiótica das respostas corretas

No conjunto das respostas corretas foram identificadas duas categorias de resposta: categoria C1 e categoria C2 (Tabela 2). Na categoria C1 as respostas exibem uma ordenação dos dados antes da identificação do valor correto da mediana, para a série de dados em número par na questão 1 e em número ímpar na questão 2. Na categoria C2 as respostas corretas demonstram a obtenção da mediana sem exibição de ordenação dos dados, mas com justificação adequada.

Tabela 2. Frequências absolutas (percentagens) das categorias de respostas corretas em cada uma das questões 1 e 2 e em ambas as questões simultaneamente

Categorias	Questão	Questão	Questões
	1	2	1 e 2
	n=34	n=40	n=20
C1. Aplicar o algoritmo no cálculo da mediana, com evidência da ordenação dos dados.	31 (91)	39 (98)	20 (100)
C2. Aplicar o algoritmo no cálculo da mediana, sem evidência da ordenação dos dados, mas com justificação adequada.	3 (9)	1 (2)	0 (0)

Fonte: elaborada pelos autores a partir dos dados da pesquisa.

As frequências de alunos que responderam corretamente e de forma similar às duas questões encontram-se sumariadas na Tabela 2. No Quadro 2 apresentam-se exemplos de respostas corretas nas categorias C1 e C2 para ambos os contextos, tabelar (questão 1) e gráfico (questão 2).

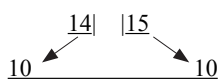
Constata-se que as respostas corretas estão concentradas na categoria C1, para ambas as questões, não havendo diferenças significativas entre as proporções de respostas corretas na categoria C1 nas duas questões (questão 1: 91%; questão 2: 98%; $p=0,231$); mais ainda, todos os alunos que responderam corretamente a ambas as questões incluem-se na categoria C1.

Quadro 2. Análise semiótica de exemplos de categorias C1 e C2 nas questões 1 e 2

	Expressão	Conteúdo
Categoria C1	<p>Questão 1. Resolução do aluno A251 (versão tabelar):</p> <p>13, 13, 13, 13, 13, 13, 14, 14, 14, 14, 15, 15, 15, 15, 15, 16, 16, 16, 16, 16</p> <p>$14 + 15 = 29$</p> <p>$29/2 = 14,5$</p> <p>R: A mediana das idades é 14,5.</p>	<ul style="list-style-type: none"> - Os alunos leem os enunciados (processo de interpretação) e identificam corretamente o(s) valor(es) pretendido(s) (particularização do problema). Os alunos identificam corretamente a questão como sendo uma questão sobre a mediana (reconhecer um campo de problemas). Reconhecem a necessidade de ordenar os dados (aplicar a definição e propriedades) para obter a mediana; - No caso do aluno A251, ele identifica os valores centrais (particularização adequada de um conceito para o caso par) e determina a sua média aritmética (discriminação adequada da definição);
	<p>Questão 2. Resolução do aluno A233 (versão tabelar)</p> <p>0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2</p> <p>R: A mediana do número de livros lidos é 1.</p>	<ul style="list-style-type: none"> - No caso do aluno A233, ele identifica o valor central (particularização adequada de um conceito para o caso ímpar); - Ambos revelam reconhecer que a mediana é apenas um valor central (aplicação adequada da definição).

continua

Quadro 2. continuação

	Expressão	Conteúdo
Categoria C2	<p>Questão 1. Resolução do aluno A61 (versão tabelar):</p>  $\frac{14 + 15}{2} = 14,5.$ <p>R: A mediana das idades é 14,5.</p> <p>Questão 2. Resolução do aluno A13 (versão gráfica): A mediana é 1, pois há tantos dados antes de si como depois de si (14 para cada lado).</p> <p>Questão 2. Resolução do aluno A294 (versão gráfica): Dado que são 5 alunos com 0 livros lidos, 14 alunos com 1 livro lido, num total de 25 alunos, metade destes ficam nos alunos com 1 livro lido. R: A mediana do número de livros lidos é 1.</p>	<ul style="list-style-type: none"> - Os alunos leem os enunciados (processo de interpretação) e identificam corretamente o(s) valor(es) pretendido(s) (particularização do problema). Os alunos identificam corretamente a questão como sendo uma situação-problema sobre a mediana (reconhecer um campo de problemas). Nenhum aluno reconheceu a necessidade de exibir todos os dados ordenados (aplicar a definição) para determinar a mediana; - No caso do aluno A61, ele identifica os valores centrais (particularização adequada de um conceito para o caso par) e determina a sua média aritmética (discriminação adequada da definição); - No caso dos alunos A13 e A294, eles identificam o valor central (particularização adequada de um conceito para o caso ímpar), analisam o gráfico de barras e justificam tendo em conta as frequências absolutas (interpretação adequada do gráfico).

Fonte: elaborado pelos autores a partir dos dados da pesquisa.

Análise semiótica das respostas parcialmente corretas e incorretas

No conjunto das respostas não corretas (categorias PC e I) foram identificados sete conflitos semióticos, comuns às questões 1 e 2, mas não necessariamente comuns a um mesmo aluno para as duas questões, como veremos mais à frente.

Na Tabela 3 apresentam-se as frequências das oito categorias descritas em cada questão e as frequências de alunos que exibiram simultaneamente o mesmo tipo de conflito semiótico nas duas questões, entre as 102 respostas parcialmente corretas ou incorretas.

Na categoria C'1 (Tabela 3) os alunos reconhecem o conjunto correto de dados e o algoritmo de cálculo da mediana, mas não o concluem, identificando os valores centrais no conjunto de dados ordenados sem determinar a sua média (na questão 1, o número de dados é par) ou sem indicar o valor da mediana (na questão 2, o número de dados é ímpar). Na categoria C'2 os alunos ilustram, de alguma forma, o procedimento correto para a determinação da mediana, mas aplicam-no a um conjunto de dados incorreto como, por exemplo, o conjunto dos valores da variável em estudo sem atender às suas frequências absolutas. Na categoria C'3 os alunos indicam valores incorretos para a mediana ou valores corretos determinados de forma incorreta, pois ordenaram os valores da variável, ou apenas uma parte destes, e ainda as respetivas frequências absolutas. Na categoria C'4 os alunos revelam não conhecer o algoritmo de cálculo da mediana e aplicam-no a um conjunto de dados incorreto. Na categoria C'5 os

alunos determinam uma estatística diferente da mediana, tendo sido observado que os alunos tomam a média (8% na questão 1 e 12% na questão 2), a moda (apenas 5% na questão 1), o máximo (1% na questão 1 e 4% na questão 2) e o mínimo (2% na questão 1 e 6% na questão 2), revelando não discriminar tais medidas da mediana. Na categoria C'6 os alunos consideram o número total de alunos como sendo a mediana. Na categoria C'7 os alunos identificam a mediana com o valor 14, que representa um dos valores centrais da variável idade na questão 1, e a frequência absoluta do valor central da variável que corresponde à moda na questão 2. Uma última categoria C'8 foi adicionada para incluir as restantes respostas, constituída por respostas sem sentido, tais como respostas associadas a um qualquer valor da frequência absoluta, sem qualquer justificação ou por tomarem um dado valor da tabela (questão 1) ou um valor dos eixos do gráfico de barras (questão 2), e ainda as respostas com o valor correto da mediana, mas sem qualquer explicação sobre a sua obtenção, podendo resultar de uma atribuição ao acaso.

Tabela 3. Frequências absolutas (percentagens) das categorias de conflitos semióticos identificados em cada uma das questões 1 e 2 e em ambas as questões simultaneamente, e valor de prova (p) dos z-testes de comparação de proporções de cada tipo de conflito entre as questões 1 e 2

Categorias (conflitos semióticos)	Questão 1	Questão 2	Valor de prova (p)	Questões 1 e 2
	n=169	n=134		
C'1. Iniciar a aplicação do algoritmo do cálculo da mediana no conjunto de dados, sem concluir.	30 (18)	4 (3)	0,000	2 (2)
C'2. Aplicar o algoritmo do cálculo da mediana, mas sem identificar corretamente o conjunto de dados.	14 (8)	23 (17)	0,019	1 (1)
C'3. Identificar corretamente o conjunto de dados e aplicar incorretamente o algoritmo do cálculo da mediana.	7 (4)	9 (7)	0,320	1 (1)
C'4. Identificar incorretamente o conjunto de dados e aplicar incorretamente o algoritmo do cálculo da mediana.	5 (3)	3 (2)	0,698	0 (0)
C'5. Determinar outra medida estatística que não a mediana.	27 (16)	31 (23)	0,116	19 (19)
C'6. Identificar o número total de dados com a mediana.	23 (14)	3 (2)	0,000	0 (0)
C'7. Identificar o valor da mediana com o valor catorze.	38 (22)	3 (2)	0,000	0 (0)
C'8. Outras respostas.	25 (15)	58 (44)	0,000	3 (3)

Fonte: elaborada pelos autores a partir dos dados da pesquisa.

Da Tabela 3 verifica-se que na questão 1 (contexto tabelar), entre 169 respostas não corretas, o tipo de conflito semiótico mais frequente é C'7 (22%) e os dois tipos menos frequentes correspondem ambos à aplicação incorreta do algoritmo do cálculo da mediana (C'3: 4%, C'4: 3%). Na questão 2 (contexto gráfico), entre as 134 respostas não corretas dessa questão,

para além do conflito pouco informativo C'8, o mais frequente é C'5 (23%). Tomando os 102 alunos com respostas não corretas simultaneamente às duas questões, constata-se que apenas 26% destes apresentaram o mesmo tipo de conflito semiótico na determinação da mediana. Essa percentagem praticamente é absorvida pela categoria C'5 (com 19%, as restantes são residuais), não havendo respostas exibindo os conflitos das categorias C'4, C'6 e C'7 simultaneamente nas duas questões, e pelo mesmo aluno. Comparando as proporções de alunos em cada tipo de conflito semiótico entre as duas questões, constata-se haver diferenças significativas para todos os conflitos, exceto C'3, C'4 e C'5 (onde $p > 0,05$). Saliente-se ainda que o número de alunos incluídos no conflito C'8 é mais frequente na questão 2 (44%) do que na questão 1 (15%).

No Quadro 3 apresentam-se análises semióticas de exemplos dos conflitos C'1 e C'2, onde o algoritmo do cálculo da mediana é aplicado corretamente mas não concluído ou é concluído mas aplicado a um conjunto incorreto de dados.

No Quadro 4 apresentam-se análises semióticas de exemplos das categorias C'3 e C'4, onde o algoritmo do cálculo da mediana não é aplicado corretamente.

Da categorização dos conflitos descritos nos Quadros 3 e 4, reconhece-se que a agregação dos conflitos C'1 e C'2 (designada por C'1+C'2) constitui o conjunto das respostas onde, embora não seja apresentado o valor correto da mediana, não se detetam erros no procedimento algorítmico do cálculo da mediana, sendo que o aluno ou não o conclui ou conclui-o incorretamente. Similarmente, reconhece-se que a agregação dos conflitos C'3 e C'4 (designada por C'3+C'4) corresponde às respostas não corretas onde a aplicação do algoritmo de cálculo da mediana foi incorreta e a agregação dos conflitos C'2 e C'4 (designada por C'2+C'4) corresponde às respostas não corretas onde foi considerado um conjunto de dados incorreto na determinação da mediana em cada questão. Analisando as proporções dessas três categorias agregadas de conflitos, conclui-se que: (i) relativamente à agregação C'1+C'2, não existem diferenças significativas entre as proporções de respostas não corretas com aplicação correta do algoritmo do cálculo da mediana na questão 1 e na questão 2 (questão 1: $44/169=26,0\%$; questão 2: $27/134=20,1\%$; $p=0,230$); (ii) relativamente à agregação C'3+C'4, não existem diferenças significativas entre as proporções de respostas não corretas com aplicação incorreta do algoritmo do cálculo da mediana na questão 1 e na questão 2 (questão 1: $12/169=7,1\%$; questão 2: $12/134=9,0\%$; $p=0,553$); e (iii) relativamente à agregação C'2+C'4, existem diferenças significativas entre as proporções de respostas não corretas provenientes da identificação errada do conjunto de dados a considerar na questão 1 e na questão 2 (questão 1: $19/169=11,2\%$; questão 2: $26/134=19,4\%$; $p=0,047$).

Seguidamente, no Quadro 5 apresentam-se as análises semióticas de cinco exemplos dos conflitos C'5, C'6 e C'7.

Os três tipos de conflitos identificados compreendem as respostas onde a mediana é confundida com outra medida de localização (como a média, o mínimo ou o máximo), com o número de dados ou com frequências absolutas de valores observados da variável. Saliente-se que, numa análise mais detalhada das respostas não corretas incluídas na categoria C'5, constatou-se que 19% dos 102 alunos referidos na Tabela 3, que exibiram simultaneamente o conflito C'5 nas duas questões, optaram pela mesma medida de localização nas duas questões, isto é: todo o aluno que optou, na questão 1, pela determinação da média (13%) ou pela indicação do valor do mínimo (4%) ou do valor máximo (2%), também o fez na questão 2.

Quadro 3. Análise semiótica de exemplos dos conflitos C'1 e C'2, nas questões 1 e 2

	Expressão	Conteúdo
Conflito C'1	<p>Questão 1. Resolução do aluno A15 (versão tabelar): 13, 13, 13, 13, 13, 13, 14, 14, 14, 14, 15, 15, 15, 15, 15, 16, 16, 16, 16, 16 R: A mediana das idades é 14 e 15.</p> <p>Questão 2. Resolução do aluno A267 (versão gráfica): 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2</p>	<ul style="list-style-type: none"> - Os alunos leem os enunciados (processo de interpretação) e identificam corretamente os valores centrais (particularização do problema). Os alunos identificam corretamente a questão como sendo uma situação-problema de cálculo da mediana (reconhecer um campo de problemas). Reconhecem a necessidade de ordenar os dados (aplicar a definição e propriedades) para a determinação da mediana e a necessidade de identificar os valores centrais (propriedade); - No caso do aluno A15, o conflito surge por ele não determinar a média aritmética dos valores centrais da amostra (particularização inadequada de um conceito) e por não reconhecer que a mediana é apenas um valor central (discriminação inadequada da definição); - No caso do aluno A267, o conflito surge por ele não atender ao número ímpar de dados (discriminação inadequada do conceito) e à mediana depender desse número ser par ou ímpar (expressão e particularização de conceitos); - Ambos os alunos não revelam conhecer que a mediana é um só valor central (aplicação inadequada da definição).
Conflito C'2	<p>Questão 1. Resolução do aluno A209 (versão tabelar): A mediana é 14,5 13 14 15 16 $\frac{14 + 15}{2} = 14,5$</p> <p>Questão 1. Resolução do aluno A73 (versão tabelar): 4, 5, 5, 6, 13, 14, 15, 16 $6 + 13 \div 2 = 12,5$</p> <p>Questão 2. Resolução do aluno A198 (versão gráfica): 0, 1, 2 ↑</p> <p>Questão 2. Resolução do aluno A270 (versão gráfica): 0, 2, 4, 6, 8, 10, 12, 14, 16</p>	<ul style="list-style-type: none"> - Todos os alunos leem o enunciado (processo de interpretação), identificam corretamente o problema de cálculo da mediana (reconhecer um campo de problemas), consideram necessário ordenar os dados (propriedade) para determinar a mediana (definição) e identificam os dois valores centrais no caso par (particularização do problema); - Os alunos A209 e A73 reconhecem a necessidade de determinar a média aritmética dos dois valores centrais (propriedade) e determinam-na (expressão e particularização de conceitos); - O conflito do aluno A209 surge quando não considera as frequências absolutas dos valores da variável (particularizar um conceito). O conflito do aluno A73 surge quando não distingue o valor da variável e a frequência da variável, utilizando indiscriminadamente todos os valores que surgem na tabela de frequências, ordenando-os (definição de mediana) e determinando os valores centrais da mediana desse conjunto de dados (particularizar um conceito). Um outro conflito deste aluno surge quando determina a média aritmética dos valores centrais (propriedade), ao não respeitar a prioridade das regras operatórias (procedimento); - Os alunos A198 e A270 aplicam corretamente o algoritmo do cálculo da mediana (aplicar a definição de mediana ao caso ímpar); - O conflito do aluno A198 surge quando o aluno não considera a frequência absoluta de cada um dos valores da variável (particularização de um conceito). No caso do aluno A270 o conflito surge quando considera os valores do eixo vertical do gráfico de barras em vez dos valores do eixo horizontal para a variável em estudo (confusão entre conceitos), aplicando o algoritmo do cálculo da mediana corretamente (processo de algoritmização).

Fonte: elaborado pelos autores a partir dos dados da pesquisa.

Quadro 4. Análise semiótica de exemplos dos conflitos C’3 e C’4, nas questões 1 e 2

	Expressão	Conteúdo
Conflito C’3	<p>Questão 1. Resolução do aluno A194 (versão tabelar): 13, 13, 13, 13, 13, 13, 14, 14, 14, 14, 15, 15, 15, 15, 15, 16, 16, 16, 16</p> <p>Questão 2. Resolução do aluno A161 (versão gráfica): 1, 2, 2, 2, 2, 2 $2 \div 2 = 1$, mediana = 1</p>	<p>– Os alunos leem o enunciado (processo de interpretação) e ordenam (propriedade) os dados para determinar a mediana (aplicam a definição e reconhecem um campo de problemas), constatando a necessidade de identificar o valor central (propriedade);</p> <p>– No caso do aluno A194, o conflito surge quando ele aplica incorretamente o algoritmo de cálculo da mediana (algoritmo), já que aplica o algoritmo para o caso ímpar (identificação inadequada de propriedades);</p> <p>– No caso do aluno A161, o conflito surge ao ordenar os valores da variável, pois ignora o valor “0” (inadequada particularização de um conceito), obtendo um conjunto de dados par, ao qual aplica corretamente a noção de mediana (aplica a definição de mediana ao caso par).</p>
Conflito C’4	<p>Questão 1. Resolução do aluno A166 (versão tabelar):</p> <p>$4, 5, 5, \phi$</p> <p>$\frac{55}{4} = 13,75$</p> <p>Questão 2. Resolução do aluno A96 (versão gráfica):</p> <p>5, 14, 6</p> <p>$\frac{14}{2} = 7$</p>	<p>– Os alunos leem o enunciado (processo de interpretação), ordenam (propriedade) os dados para determinar a mediana (aplicar a definição e reconhecer um campo de problemas), reconhecendo a necessidade de identificar o valor central (propriedade);</p> <p>– O aluno A166, ao aplicar o algoritmo da mediana (aplicar a definição), exibe o primeiro conflito, já que utiliza os valores das frequências absolutas e não os valores da variável (particularização inadequada de um conceito); o segundo conflito emerge no cálculo esperado da média aritmética dos dois valores centrais (aplicação inadequada do algoritmo) pois para esse cálculo não considera no numerador a adição dos dois valores centrais e no denominador considera o número total de dados (inadequada particularização de conceitos), obtendo um valor incorreto para a mediana dos dados considerados; o terceiro conflito emerge do resultado obtido, pois o aluno revela desconhecer que o valor da mediana deve situar-se entre os valores centrais (propriedade inadequada de conceitos);</p> <p>– O primeiro conflito revelado pelo aluno A96 surge ao considerar o conjunto de frequências absolutas como valores da variável em estudo (particularização inadequada de um conceito) na aplicação do algoritmo de determinação da mediana e emerge ainda um outro conflito ao considerar metade do valor central (aplicação inadequada do algoritmo).</p>

Fonte: elaborado pelos autores a partir dos dados da pesquisa.

Quadro 5. Análise semiótica de exemplos dos conflitos C'5, C'6 e C'7, nas questões 1 e 2

	Expressão	Conteúdo
Conflito C'5	<p>Questão 1. Resolução do aluno A256 (versão tabelar):</p> $\frac{13 \times 6 + 14 \times 4 + 15 \times 5 + 16 \times 6}{4} = 72,25$	<p>– Os alunos leem o enunciado (processo de interpretação), mas não identificam a mediana como a medida de localização central referida no problema (reconhecimento incorreto do campo de problemas);</p> <p>– O aluno A256 apresenta um conflito na sua resolução ao determinar a média em vez da mediana (confusão entre conceitos). Outro conflito emerge ao aplicar incorretamente o algoritmo de cálculo da média (algoritmo), já que divide o valor da soma dos valores da variável pelo número de valores distintos e não pelo número total de alunos (aplicação inadequada de conceito);</p> <p>– O aluno A135 revela um conflito quando indica o valor do máximo do conjunto de dados (particularização incorreta ao problema) em vez de determinar a mediana (confusão entre conceitos);</p> <p>– No caso do aluno A28 o seu primeiro conflito surge quando identifica dois valores, 14 e 12 (produto entre a frequência absoluta e o valor da variável para os valores 1 e 2 da variável) (particulariza ao problema de determinação da média) em vez de determinar a mediana (confusão entre conceitos). Reconhece que o zero é o elemento neutro da adição (propriedade), apesar de não o expressar. Outro conflito emerge da indicação do número total de alunos da turma, já que é 25 e não 26, embora expresse o cálculo da média (algoritmo) em notação adequada (linguagem). Por fim, indica o valor da mediana que, apesar de ser correto, foi determinado de forma incorreta (particularização inadequada de um conceito);</p> <p>– O aluno A171 indica para a mediana o mínimo do conjunto de dados (confusão entre conceitos), originando o primeiro conflito; outro conflito surge da justificação apresentada (aplicação inadequada da definição);</p> <p>– O aluno A272 apresenta um primeiro conflito ao considerar a média e não a mediana (confusão entre conceitos), aplicando incorretamente o algoritmo do cálculo da média (algoritmo); outro conflito emerge ao adicionar os valores do eixo vertical do gráfico de barras como sendo o conjunto de dados correto (confusão entre conceitos e inadequada particularização de um conceito) e assume incorretamente que o “0” não influencia o valor da média (aplicação de propriedade), ao não o incluir.</p>
	<p>Questão 1. Resolução do aluno A135 (versão tabelar):</p> <p>16 pois é o valor mais alto</p>	
	<p>Questão 2. Resolução do aluno A28 (versão gráfica):</p> $14 + 12 = 26$ $26 \div 26 = 1$ <p>A mediana é 1</p>	
	<p>Questão 2. Resolução do aluno A171 (versão gráfica):</p> <p>A mediana do número de livros lidos pelos alunos nas férias é 0, porque é o menor número de livros lidos.</p>	
	<p>Questão 2. Resolução do aluno A272 (versão gráfica):</p> $2 + 4 + 6 + 8 + 10 + 12 + 14 + 16 =$ $= 72 \div 8 = 9$	

continua

Quadro 5. continuação

	Expressão	Conteúdo
Conflito C'6	Questão 1. Resolução do aluno A218 (versão tabelar): A mediana é 20	– Os alunos leem o enunciado (processo de interpretação), mas não identificam o problema como sendo sobre o cálculo da mediana (reconhecimento incorreto do campo de problemas);
	Questão 2. Resolução do aluno A223 (versão gráfica): A mediana é $5 + 14 + 6 = 25$	– O conflito dos alunos A218 e A223 surge ao considerarem a mediana como sendo o número total de alunos (particularização inadequada de um conceito). Um outro conflito emerge de os alunos não considerarem que a mediana é um valor compreendido entre os valores mínimo e máximo da variável (aplicação de propriedade).
Conflito C'7	Questão 1. Resolução do aluno A30 (versão tabelar): A mediana é 14	– Os alunos leem o enunciado (processo de interpretação), mas não identificam o problema como sendo sobre o cálculo de mediana (reconhecimento incorreto do campo de problemas);
	Questão 2. Resolução do aluno A35 (versão gráfica): A mediana é 14	– O conflito do aluno A30 surge ao considerar a mediana como sendo o primeiro valor central (particularização inadequada de um conceito); – O conflito do aluno A35 surge ao considerar a mediana como sendo a frequência absoluta do valor da mediana (particularização inadequada de um conceito). Um outro conflito emerge de o aluno não considerar que a mediana é um valor compreendido entre os valores mínimo e máximo da variável (aplicação de propriedade).

Fonte: elaborado pelos autores a partir dos dados da pesquisa.

Conclusões

Na presente investigação, a análise das respostas de 332 alunos do 8.º ano, classificadas em “corretas”, “parcialmente corretas” e “incorretas”, além das “não respostas”, permitiu avaliar o conhecimento dos alunos sobre a determinação da mediana em dois contextos de apresentação de dados: tabela de frequências (questão 1) e gráfico de barras (questão 2).

Comparando as três categorias de respostas e as não respostas, entre os contextos tabelar e gráfico, não se registraram mudanças significativas nas respostas de um mesmo aluno. Analisando com mais detalhe as respostas, constata-se que, tendencialmente, um aluno que determina corretamente a mediana, o faz bem nos dois contextos (Tabela 2), evidenciando em ambos os contextos a ordenação dos dados na sua resposta. Nos estudos de Barr (1980), com dados por organizar, e de Carvalho (1996), com dados organizados em tabelas de frequências e gráficos de barras, constata-se que são poucos os alunos que têm em atenção a prévia ordenação dos dados na determinação da mediana. No presente estudo verifica-se que mais alunos efetuaram a ordenação dos dados em contexto tabelar (18%) do que em contexto gráfico (3%) (conflito C'1).

Ao restringir-se a análise às respostas não corretas (parcialmente corretas e incorretas), conclui-se que um aluno que erra em ambos os contextos não tende a manter o tipo de resposta, significando que ele denota dificuldade em ambos os contextos, mas o tipo de dificuldade é diferente entre os contextos. Segundo Arteaga (2008) e Ribeiro (2011), os alunos têm dificuldades na determinação da mediana em diversos contextos, e apresentam mais dificuldades em dados organizados (tabelar e gráfico) do que em dados não organizados.

Nos conflitos semióticos, conclui-se que raramente um aluno revela o mesmo tipo de conflito nos dois contextos, exceto quando confunde o conceito e determina outra medida estatística em vez da mediana (conflito C'5), e isso ocorre com maior frequência quando os dados estão organizados em gráfico de barras (23%) do que quando estão organizados em tabela de frequências (16%). Salienta-se que todo o aluno que optou, na questão 1, pela determinação da média (13%) ou pela indicação do valor do mínimo (4%) ou do valor máximo (2%), também o fez na questão 2. Tal significa que o contexto afeta o desempenho dos alunos com dificuldades na determinação correta da mediana em ambos os contextos, mas por motivos ou causas diferentes. Enquanto determinar a média em vez da mediana é um dos erros mais frequentes em outros estudos (BATANERO, 2000; RIBEIRO; CORREIA; FERNANDES, 2013), já a troca da mediana pela moda não se verificou neste estudo, não corroborando os resultados obtidos por Barros (2004) e Mayén et al. (2007).

Em termos globais, verificou-se uma frequência significativa de “não respostas” na determinação da mediana em contexto gráfico e de respostas não corretas quando os dados estão organizados em tabela. Sobre este facto, Batanero (2000) realça que as dificuldades dos alunos resultam de eles tenderem a usar um único procedimento na determinação da mediana, apesar de ela depender do número de dados ser par ou ímpar.

No presente estudo, das respostas analisadas, foram identificados oito tipos de conflitos semióticos na determinação da mediana, verificando-se em cinco deles (C'1, C'2, C'6, C'7 e C'8) proporções de ocorrência significativamente diferentes entre os contextos tabelar e gráfico. A análise permitiu ainda verificar que os dois contextos em consideração não afetam a aplicação correta ou incorreta do algoritmo do cálculo da mediana por parte dos alunos com respostas não corretas, mas dificulta a identificação correta do conjunto de dados a considerar na determinação da mediana.

Os resultados aqui obtidos, relativamente às dificuldades na determinação da mediana, são também apontados por outros estudos (BOAVENTURA, 2003; RIBEIRO; CORREIA; FERNANDES, 2013), bem como a sua dificuldade de interpretação (ARTEAGA et al., 2011; FERNANDES; CARVALHO; CORREIA, 2011). Embora a recolha dos dados utilizados neste estudo tenha ocorrido antes da leção dos conceitos estatísticos do 8.º ano e depois de uma longa paragem letiva na leção dos conceitos estatísticos, os resultados obtidos neste estudo apontam para a existência de dificuldades dos alunos, geradoras de conflitos semióticos, e para os evitar será necessário dar maior atenção ao significado do conceito de mediana e às suas propriedades, e não apenas à aplicação do seu algoritmo de cálculo.

Em estudos futuros será importante verificar se os conflitos detetados neste estudo prevalecem quando são propostos problemas em outros contextos como, por exemplo, com os dados não organizados ou organizados noutros tipos de representação.

Agradecimentos

Este trabalho foi subsidiado por fundos portugueses através do Centro de Investigação e Desenvolvimento em Matemática e Aplicações (CIDMA) da Universidade de Aveiro e da Fundação para a Ciência e a Tecnologia (FCT), dentro do projeto UID/MAT/04106/2013, e também pelo Centro de Investigação em Educação (CIEd), UID/CED/01661/2014, Instituto de Educação, Universidade do Minho, através de fundos nacionais da FCT/MCTES-PT; EDU2016-74848-P (AEI, FEDER).

Referências

- ARTEAGA, P. **Análisis de gráficos estadísticos elaborados en un proyecto de análisis de datos**. 2008. 100 f. Disertación (Maestría en Didáctica de las Matemáticas) – Departamento de Didáctica de la Matemática, Universidad de Granada, Granada, 2008.
- ARTEAGA, P. et al. Sentido numérico y gráficos estadísticos en la formación de profesores. **Publicaciones de la Facultad de Educación y Humanidades del Campus de Melilla**, Melilla, v. 41, p. 33-50, 2011.
- BARR, G. V. Some student ideas on the median and the mode. **Teaching Statistics**, Nottingham, v. 2, n. 2, p. 38-41, 1980.
- BARROS, P. M. **Os futuros professores do 2º ciclo e a estocástica: dificuldades sentidas e o ensino do tema**. 2003. 296 f. Dissertação (Mestrado em Supervisão Pedagógica em Ensino da Matemática) – Universidade do Minho, Braga, 2004.
- BATANERO, C. Significado y comprensión de las medidas de tendencia central. **Uno: revista de didáctica de las matemáticas**, Barcelona, n. 25, p. 41-58, 2000.
- BATANERO, C.; GODINO, J.; NAVAS, F. Concepciones de maestros de primaria en formación sobre los promedios: versión ampliada del trabajo publicado. In: SALMERÓN, H. (Coord.). **VII Jornadas LOGSE: evaluación educativa**. Granada: Grupo Editorial Universitario, 1997. p. 310-304. Disponível em: <www.ugr.es/~batanero/pages/ARTICULOS/Logse.pdf>. Acesso em: 6 fev. 2016.
- BOAVENTURA, M. G. **Dificuldades de alunos do ensino secundário em conceitos estatísticos: o caso das medidas de localização**. 2003. 223 f. Dissertação (Mestrado em Supervisão Pedagógica em Ensino da Matemática) – Universidade do Minho, Braga, 2003.
- CARVALHO, C. Algumas questões em torno de tarefas estatísticas com alunos do 7º ano. In: ROQUE, A.; LAGARTO, M. J. (Org.). **Atas do ProfMat 96**. Lisboa: Associação de Professores de Matemática, 1996. p. 165-171.
- CARVALHO, C. Um olhar da psicologia pelas dificuldades dos alunos em conceitos estatísticos. In: FERNANDES, J. A.; SOUSA, M. V.; RIBEIRO, S. A. (Org.). **Ensino e aprendizagem de probabilidades e estatística: atas do I Encontro de Probabilidades e Estatística na Escola**. Braga: Centro de Investigação em Educação da Universidade do Minho, 2004. p. 85-102.

- CARVALHO, C.; CÉSAR, M. As aparências iludem: reflexões em torno do ensino da estatística no ensino básico. In: LOUREIRO, C.; OLIVEIRA, O.; BRUNHEIRA, L. (Org.). **Ensino e aprendizagem da estatística**. Lisboa: Sociedade Portuguesa de Estatística, 2000. p. 212-225.
- ESPINEL, M. C. et al. In: SERRANO, L. (Ed.). **Tendencias actuales de la investigación en educación estocástica**. Málaga: Gráficas San Pancraccio, 2009. p. 133-156.
- FERNANDES, J. A.; CARVALHO, C. F.; CORREIA, P. F. Contributos para a caracterização do ensino da estatística nas escolas. **Bolema**, Rio Claro, v. 24, n. 39, p. 585-606, 2011.
- FRIEL, S.; CURCIO, F.; BRIGHT, G. Making sense of graphs: critical factors influencing comprehension and institutional implications. **Journal for Research in Mathematics Education**, Reston, v. 32, n. 2, p. 124-158, 2001.
- GARFIELD, J.; AHLGREN, A. Difficulties in learning basic concepts in probability and statistics: implications for research. **Journal for Research in Mathematics Education**, Reston, v. 19, n. 1, p. 44-63, 1988.
- GODINO, J. D.; BATANERO, C. Significado institucional y personal de los objetos matemáticos. **Recherches en Didactique des Mathématiques**, Grenoble, v. 14, n. 3, p. 325-35, 1994.
- GODINO, J. D.; BATANERO, C.; FONT, V. The onto-semiotic approach to research in mathematics education. **ZDM: the international journal on mathematics education**, Berlin, v. 39, n. 1-2, p. 127-135, 2007.
- GROTH, R. E.; BERGNER, J. A. Preservice elementary teachers' conceptual and procedural knowledge of mean, median, and mode. **Mathematical Thinking and Learning**, Mahwah, v. 8, n. 1, p. 37-63, nov. 2006.
- JACOBBE, T. Elementary school teachers' understanding of the mean and median. In: BATANERO, R. et al. (Ed.). **Joint ICMI/IASE study: teaching statistics in school mathematics: challenges for teaching and teacher education**. Monterrey: ICMI: IASE, 2008. Disponível em: <http://iase-web.org/documents/papers/rt2008/T2P13_Jacobbe.pdf>. Acesso em: 21 ago. 2018.
- LI, D. Y.; SHEN, S. M. Students' weaknesses in statistical projects. **Teaching Statistics**, Nottingham, v. 14, n. 1, p. 2-8, 1992.
- MARQUES, M.; GUIMARÃES, G.; GITIRANA, V. Compreensões de alunos e professores sobre média aritmética. **Bolema**, Rio Claro, v. 24, n. 40, p. 725-745, 2011.
- MAYÉN, S. et al. Comprensión de las medidas de posición central en estudiantes mexicanos de bachillerato. **Unión**, La Laguna, v. 9, p. 187-201, mar. 2007.
- MAYÉN, S.; DÍAZ, C.; BATANERO, C. Conflictos semióticos de estudiantes con el concepto de mediana. **Statistics Education Research Journal**, Dordrecht, v. 8, n. 2, p. 74-93, 2009.

- MAYÉN, S.; ORTIZ, J. J.; DÍAZ, C. Conflictos semióticos de estudiantes mexicanos en el uso de la mediana. In: SERRANO, L. (Ed.). **Tendencias actuales de la investigación en educación estocástica**. Málaga: Gráficas San Pancrácio, 2009. p. 75-94.
- MOORE, D. S. Teaching statistics as a respectable subject. In: GORDON, F.; GORDON, D. S. (Ed.). **Statistics for the twenty-first century**. Washington: Mathematical Association of America, 1991. p. 14-25.
- MORAIS, P. C.; FERNANDES, J. A. Realização de duas tarefas sobre construção, leitura e interpretação de gráficos estatísticos por alunos do 9.º ano. In: SEMINÁRIO DE INVESTIGAÇÃO EM EDUCAÇÃO MATEMÁTICA, 22., 2011, Lisboa. **Atas...**
- MURTEIRA, B. et al. **Introdução à estatística**. Lisboa: Escolar, 2010.
- OLIVEIRA, H.; HENRIQUES, A. Promover o raciocínio estatístico no ensino básico recorrendo à tecnologia. **Boletim da SPE**, Lisboa, p. 23-31, out. 2014.
- PONTE, J. P.; FONSECA, H. Orientações curriculares para o ensino da estatística: análise comparativa de três países. **Quadrante**, Lisboa, v. 10, n. 1, 93-115, 2001.
- PORTUGAL. Ministério da Educação e Ciência. **Programa de matemática para o ensino básico**. Lisboa, 2013.
- RIBEIRO, A. F. **Estratégias de resolução de tarefas estatísticas por alunos do 7.º ano de escolaridade**. 2011. 96 f. Relatório de Estágio (Mestrado em Ensino de Matemática no 3º Ciclo do Ensino Básico e no Ensino Secundário) – Instituto de Educação, Universidade do Minho, Braga, 2011.
- RIBEIRO, A. F.; CORREIA, P. F.; FERNANDES, J. A. Estratégias usadas por alunos do 7.º ano na resolução de tarefas estatísticas. In: FERNANDES, J. A. et al. (Org.). **Atas do III Encontro de Probabilidades e Estatística na Escola**. Braga: Centro de Investigação em Educação da Universidade do Minho, 2013. p. 193-206.
- SOUSA, H. D. et al. **Relatórios / estatísticas**: 2010. Disponível em: <<http://www.dge.mec.pt/relatoriosestatisticas-0>>. Acesso em: 6 fev. 2016.

Artigo recebido em 05/07/2017. Aceito em 21/03/2018.

Contacto: Universidade de Aveiro, Departamento de Educação e Psicologia, Aveiro, 3810-193, Portugal.