



## COMPREENSÃO DE AMOSTRA E AMOSTRAGEM POR ESTUDANTES BRASILEIROS DOS ANOS FINAIS DO ENSINO FUNDAMENTAL

### Comprensión de muestra y muestreo por estudiantes brasileños en los últimos años de la escuela secundaria

### Sample and sampling understanding by brazilian students in the Middle School

Luan Luna<sup>1</sup>

Universidade Federal de Pernambuco (Recife, Brasil)

Gilda Guimarães<sup>2</sup>

Universidade Federal de Pernambuco (Recife, Brasil)

#### Resumo

O uso crescente de dados na sociedade para predição e tomada de decisão torna prioritário o ensino de Estatística para que os estudantes construam o raciocínio inferencial informal. Neste estudo realizamos uma pesquisa com 307 estudantes brasileiros do 6º ao 9º ano do ensino fundamental (11 - 15 anos de idade) buscando investigar o que sabem sobre amostra e amostragem. Os resultados indicaram um baixo desempenho, apesar de uma gradação crescente na média de acerto em função da escolaridade. Entretanto, desde o 6º ano observamos estudantes que responderam de forma adequada a todas as questões, evidenciando a possibilidade da compreensão do conceito desde essa idade. Porém, a maioria deles demonstraram dificuldade em definir amostra apresentando exemplos do cotidiano como resposta e associando população apenas com pessoas. Ao construir amostras priorizavam o tamanho em detrimento da variabilidade e não conseguem perceber amostras tendenciosas. Esses resultados evidenciam dificuldades e não impossibilidades, o que nos deixa com a responsabilidade de planejar boas aulas para a efetiva aprendizagem de todos sobre amostra e amostragem.

**Palavras-chave:** *Educação Estatística, Amostragem, Ensino Fundamental.*

#### Resumen

El creciente uso de datos en la sociedad para la predicción y la toma de decisiones hace que la enseñanza de la Estadística sea una prioridad para que los estudiantes construyan un razonamiento inferencial informal. En este estudio, realizamos una encuesta con 307 estudiantes brasileños del 6º al 9º grado de la escuela primaria (11 - 15 años) buscando investigar lo que saben sobre muestra y muestreo. Los resultados indicaron un bajo rendimiento a pesar de una gradación creciente en el promedio de aciertos en función de la escolaridad. Sin embargo, desde el 6to año observamos estudiantes que respondieron todas las preguntas de manera adecuada, demostrando la posibilidad de comprender el concepto a partir

---

\*Autor de correspondencia: [luancluna@gmail.com](mailto:luancluna@gmail.com) (L. Luna)

<sup>1</sup> <https://orcid.org/0000-0002-2990-253X>; ([luancluna@gmail.com](mailto:luancluna@gmail.com))

<sup>2</sup> <https://orcid.org/0000-0002-1463-1626>; ([gilda.lguimaraes@gmail.com](mailto:gilda.lguimaraes@gmail.com))

de esa edad. Sin embargo, la mayoría de los estudiantes del estudio mostraron dificultad para definir una muestra, presentando ejemplos de la vida cotidiana como respuesta y asociando población solo con personas. Al construir muestras, priorizaron el tamaño sobre la variabilidad y no pudieron percibir muestras sesgadas. Estos resultados muestran dificultades y no imposibilidades, lo que nos deja la responsabilidad de planificar buenas clases para que todos aprendan de manera efectiva sobre la muestra y muestreo.

**Palabras clave:** Educación Estadística, Muestreo, Escuela Primaria.

### **Abstract**

The increasing use of data in society for prediction and decision makes teaching Statistics a priority for students to build informal inferential reasoning. In this study, we carried out a survey with 307 Brazilian students from the 6th to the 9th grade of elementary school (11 - 15 years old) seeking to investigate what they know about sample and sampling. The results indicated a low performance, despite an increasing gradation in the average of correct answers as a function of schooling. However, since the 6th grade we observed students who answered all the questions adequately, demonstrating the possibility of understanding the concept from that age. However, most of the students in the study showed difficulty in defining a sample, presenting examples of everyday life as a response and associating population only with people. When building samples, they prioritized size over variability and failed to perceive biased samples. These results show difficulties and not impossibilities, which leaves us with the responsibility of planning good classes for everyone to effectively learn about the relation sample and sampling.

**Keywords:** Statistical Education, Sampling, Middle School.

*Recibido: 30/05/2022 - Aceptado: 02/09/2022*

## **1. INTRODUÇÃO**

A Estatística é a ciência da coleta, da organização, da representação, da análise e da realização de inferências a partir de dados. É notório o papel que ela proporciona ao conhecimento do mundo físico e social. Deste modo, Moore (1995) frisa que os dados são números com um contexto.

Em qualquer pesquisa estatística, antes da realização da coleta de dados, se faz necessário definir a população/amostra. Como nem sempre é possível fazer um censo, ou seja, com toda a população, utiliza-se a amostragem. Quando se seleciona uma boa amostra, ela é capaz de representar toda a população de interesse. Portanto, utilizamos os dados obtidos de uma amostra para tomarmos decisões. Assim, é fundamental sabermos planejar uma pesquisa amostral e identificarmos quando uma amostra é tendenciosa ou representativa da população.

Embora tais conceitos possuam relevância no nosso dia a dia, diversos pesquisadores (Ben-Zvi et al., 2015) apontam que são poucas as pesquisas realizadas, necessitando de mais investigações na área.

No cenário brasileiro, após mudanças das prescrições curriculares para a Base Nacional Comum Curricular – BNCC (Brasil, 2018), é perceptível uma maior ênfase concernente ao ensino-aprendizagem

de amostra e amostragem, reforçando a necessidade destes conceitos serem compreendidos pelos estudantes.

Por tais razões, neste texto, temos por objetivo levantar o que estudantes brasileiros dos anos finais do Ensino Fundamental (11 – 15 anos de idade) sabem sobre amostra e amostragem. Tal pesquisa consiste em um recorte de um estudo de doutoramento.

## **2. CENSO E AMOSTRA**

Para que os resultados de pesquisas estatísticas sejam confiáveis é primordial planejar bem as etapas definindo, entre outras coisas, a população a ser investigada ou uma amostra. Em Estatística, população refere-se ao conjunto de elementos (indivíduos, objetos etc.) que possuem pelo menos uma característica em comum, como todas as lâmpadas produzidas por uma fábrica em um dia, todo o sangue que corre no corpo de uma pessoa ou todos os habitantes de uma cidade, estado ou país.

Uma pesquisa estatística quando realizada com toda a população de interesse é chamada de pesquisa censitária, censo ou recenseamento. Uma pesquisa censitária permite obter precisão da população investigada. Entretanto, nem sempre é possível investigar toda uma população. Às vezes, o tamanho da população é grande o que demandaria muito tempo e dinheiro para coletar os dados. Outras vezes, a realização da pesquisa pode resultar na inutilização ou no consumo de um produto (caracterizar-se num processo destrutivo). Assim, recorre-se a realização de pesquisa amostral (ou por amostragem), como se deduz pelo nome, utiliza uma amostra, a qual corresponde a um subconjunto da população. A seleção da amostra é de extrema importância, pois os dados coletados necessitam representar toda a população, ao ponto de permitir fazer generalizações. Caso contrário, teremos resultados imprecisos e enviesados.

Existem diferentes técnicas para seleção de uma amostra e esse processo é chamado de amostragem. Stevenson (1981, p. 158) argumenta que "a finalidade da amostragem é fazer generalizações sobre todo um grupo, sem precisar examinar cada um de seus elementos". A utilização da amostragem é primordial quando desejamos realizar inferência estatística, pois formulamos julgamentos sobre um todo analisando apenas uma amostra. No processo de seleção de uma amostra entram em cena outros conceitos importantes de serem esclarecidos: representatividade, variabilidade e tamanho.

A representatividade é peça chave na generalização de resultados de uma amostra para a população e está associada as características importantes da população (gênero e classe social de pessoas, raça de cães, marcas e modelos de celulares etc.) que precisam estar refletidas na amostra. Em outras palavras, uma amostra representativa é aquela que reflete, com a maior precisão possível, características

importantes da população de interesse, considerando a variável em estudo e respeitando as devidas proporções.

Outro fator fundamental na seleção de uma amostra se diz respeito a variabilidade ou ausência de variabilidade de uma população. Quanto mais homogênea for a população, menor poderá ser o tamanho da amostra para representá-la, uma vez que há pouca variabilidade. De forma oposta, quanto mais heterogênea for a população, maior variabilidade, e o quantitativo da amostra deverá ser maior para atender a todas as características dessa população. Portanto, percebe-se que representatividade, variabilidade e tamanho da amostra, são algumas das ideias centrais e inter-relacionadas para a aprendizagem de amostragem.

Entre os poucos estudos de levantamento de conhecimentos prévios de alunos sobre essa temática estão os de Meletiou-Mavrotheris e Papparistodemou (2015) com estudantes do Chipre (11-12 anos de idade) e Ruiz-Reyes e García-Contreras (2021) com estudantes chilenos (entre 13 e 18 anos de idade), os quais observaram que os estudantes conhecem a palavra amostra, mas apresentam exemplos de amostras homogêneas. É fundamental apresentar exemplos que contemplem amostras heterogêneas, pois elas darão subsídio a discussão de variabilidade. As ideias inerentes a amostras homogêneas não ajudam no entendimento da variabilidade e da necessidade em selecionar amostras grandes ao fazer inferências a partir de dados.

Acrescido a isso, Watson e Moritz (2000) argumentam que estudantes australianos (11-12 anos de idade) apresentaram uma diversidade de crenças a respeito do tamanho da amostra, enquanto os com 14-15 anos de idade mostraram preocupação também com a variabilidade, porém não identificaram os riscos de vieses associado à amostra. Innabi (2007) também afirma que poucos estudantes dos Emirados Árabes Unidos (17 e 18 anos de idade) e Gomes e Guimarães (2018) com estudantes brasileiros (10 a 15 anos de idade) que poucos consideram adequadamente o tamanho da amostra e o viés de seleção. A maioria se baseia em crenças pessoais, independentemente das propriedades da amostra selecionada. Tais equívocos incluem “qualquer amostra representa a população” e “qualquer amostra não representa a população”. Esses resultados indicam indiferença entre os estudantes com graus de escolaridade diferentes, revelando a necessidade de se trabalhar com a análise e construção de uma amostra.

Buscando compreender possíveis causas dessas dificuldades, Luna e Guimarães (2021a; 2021b) analisaram atividades de amostragem propostas em 88 livros didáticos de Matemática dos anos finais do Ensino Fundamental, aprovados pelo Programa Nacional do Livro e do Material Didático (PNLD), edições 2017 e 2020. Os autores perceberam que na maioria das vezes, as atividades apresentavam dados

fictícios, excesso de situações de identificação de população/amostra em contexto de pessoas, inexistência de propostas que explorem a relação entre variabilidade e tamanho da amostra, e, poucas atividades de análise de amostras (identificar vieses e julgar a adequação de métodos de amostragem). Em contrapartida, notou-se a boa qualidade das atividades de realização de pesquisa amostral, as quais possuíam sugestões de temas relevantes e indicação de planilhas eletrônicas para as representações gráficas e tabulares e, orientações para o trabalho do professor, no formato em U.

### 3. ORIENTAÇÃO CURRICULAR BRASILEIRA E ESTUDOS ANTECEDENTES: O CASO DA AMOSTRA E DA AMOSTRAGEM

No Brasil, a prescrição curricular Base Nacional Comum Curricular – BNCC (Tabela 1) enfatiza a definição da amostra somente a partir do 7º ano do Ensino Fundamental.

**Tabela 1**

*Habilidades de aprendizagem referente a amostragem*

Ano de escolaridade	Habilidades
7º ano (12 – 13 anos de idade)	(EF07MA36) Planejar e realizar pesquisa envolvendo tema da realidade social, identificando a necessidade de ser censitária ou de usar amostra, e interpretar os dados para comunicá-los por meio de relatório escrito, tabelas e gráficos, com o apoio de planilhas eletrônicas.
8º ano (13 – 14 anos de idade)	(EF08MA26) Selecionar razões, de diferentes naturezas (física, ética ou econômica), que justificam a realização de pesquisas amostrais e não censitárias, e reconhecer que a seleção da amostra pode ser feita de diferentes maneiras (amostra casual simples, sistemática e estratificada). (EF08MA27) Planejar e executar pesquisa amostral, selecionando uma técnica de amostragem adequada, e escrever relatório que contenha os gráficos apropriados para representar os conjuntos de dados, destacando aspectos como as medidas de tendência central, a amplitude e as conclusões.
9º ano (14 – 15 anos de idade)	(EF09MA23) Planejar e executar pesquisa amostral envolvendo tema da realidade social e comunicar os resultados por meio de relatório contendo avaliação de medidas de tendência central e da amplitude, tabelas e gráficos adequados, construídos com o apoio de planilhas eletrônicas.

*Fonte:* Elaborado pelos autores a partir de Brasil (2018, p. 300-319)

A partir do Tabela 1, percebe-se que é proposto a percepção e a justificativa da pesquisa ser censitária ou amostral e o uso adequado de métodos probabilísticos na realização de pesquisas estatísticas. Além disso, a realização de pesquisas estatísticas amostrais é indicada em todos os anos escolares (7º ao 9º ano). Porém, conceitos como variabilidade e tamanho da amostra, não foram mencionados. Além do

mais, não há alusão da noção de margem de erro e de possíveis efeitos de vieses na seleção de amostras. Especificamente a respeito dos métodos de amostragem, são apontados apenas os probabilísticos (aleatória simples, estratificada e sistemática).

No contexto espanhol, Ruiz-Reyes (2019) e Pedrosa (2019), destacam que os conceitos inerentes à amostragem estão presentes de forma explícita também a partir do 1º ano da educação secundária (12-13 anos). As autoras frisam que o documento curricular apresenta coerência nas proposições das habilidades referentes à amostragem, apesar de considerarem uma introdução tardia.

Em contrapartida, Ruiz-Reyes (2019) sinaliza que no currículo chileno, as primeiras ideias relacionadas à amostragem são apontadas no 4º ano da educação primária (9 -10 anos), solicitando a realização de pesquisas estatísticas e a comparação de amostras aleatórias usando gráficos e tabelas. Além disso propõe que no 7º ano (12 - 13 anos) a abordagem de predição de características de uma população a partir do resultado de amostras aleatórias, a diferenciação de amostras aleatórias e não aleatórias para sugerir a importância da aleatoriedade no processo amostragem. Assim, a proposta curricular chilena oferece uma introdução gradual dos conteúdos associados ao tema de amostragem e possibilita alcançar o desenvolvimento da inferência informal. No entanto, os autores destacam a ausência do conceito explícito de viés em amostras e variabilidade amostral.

#### 4. MÉTODO

Foi realizado um teste diagnóstico para levantar os conhecimentos prévios sobre amostragem com 307 estudantes (Tabela 2) dos anos finais do Ensino Fundamental (11 – 15 anos de idade) de escolas da cidade do Recife, localizada no estado de Pernambuco, Brasil. A amostra da pesquisa foi por conveniência. Os estudantes dessas escolas são de nível socioeconômico baixo, cujos pais predominantemente não concluíram o Ensino Médio.

**Tabela 2**

*Frequência de estudantes por ano de escolaridade*

Ano de escolaridade	Frequência
6º	102
7º	68
8º	64
9º	73

A pesquisa foi realizada no período normal de aula durante o mês de setembro de 2021. Os estudantes resolveram o teste de forma individual, levando aproximadamente 50 minutos para concluí-lo.

O teste explorava diferentes habilidades sobre amostragem. Nesse artigo, apresentamos os resultados de sete itens, os quais tinham por objetivo conceituar amostra, estabelecer a relação entre população e amostra, justificar a vantagem de realizar uma pesquisa amostral e construir e analisar amostras. Os itens eram abertos, exigindo que os estudantes justificassem suas respostas.

Dado que nenhum estudo prévio sobre amostragem de estudantes de escolas brasileiras dos anos finais do Ensino Fundamental (6º ao 9º ano) havia sido realizado, buscamos determinar se as principais conclusões da literatura internacional seriam válidas para os estudantes que vieram de um contexto cultural e curricular diferente.

No item 1 classificamos as respostas dadas pelos estudantes como correta quando a definição remetesse a ideia parte de um todo. Para os itens 2, 3 e 4 foi considerado correto estabelecer adequadamente a população e amostra. No item 5, atribuímos como correto uma justificativa apropriada para realização de uma pesquisa amostral. No item 6 foi concedido um ponto por atribuir um dos fatores para a construção da amostra (tamanho ou variabilidade) e dois pontos por apresentar os dois fatores. Por fim, no item 7, foi designado um ponto ao escolher a amostra da opção E com uma justificativa apropriada e dois pontos pela opção B com a justificativa correta. Portanto, a pontuação final do teste diagnóstico poderia variar entre 0 e 9 pontos.

#### 4.1. Atividades propostas

**Item 1** – Conceituar amostra - Você já ouviu a palavra amostra antes? O que ela significa?

**Itens 2, 3 e 4** – Estabelecer população e amostra nas seguintes situações:

Pesquisa A	Quem é a população?	Quem é a amostra?
Para descobrir o nível de satisfação dos usuários de transporte público das cidades que compõem a Região Metropolitana do Recife, um instituto entrevistou 2400 pessoas que utilizam esse meio de locomoção.		
Pesquisa B	Quem é a população?	Quem é a amostra?
14 macacos foram submetidos a doses de um medicamento antiviral com a finalidade de testar a eficácia de determinado tratamento.		
Pesquisa C	Quem é a população?	Quem é a amostra?
O PROCON realizou pesquisa em 20 estabelecimentos de Pernambuco, a fim de descobrir o valor da cesta básica no estado.		

**Item 5** – Justificar a realização de uma pesquisa amostral

Um Instituto entrevistou 2.732 pessoas para saber quais atividades físicas os brasileiros praticam. Por que você acha que o Instituto entrevistou uma amostra e não todos os brasileiros?

**Item 6** – Construir amostra

A direção de sua escola quer saber os meios de locomoção que os estudantes mais utilizam para chegar lá: a pé, bicicleta, ônibus, carro, moto etc. Para isso, a direção realizará uma pesquisa estatística com uma amostra de estudantes da escola. Quais critérios a direção poderia utilizar para selecionar uma amostra de estudantes de modo que os resultados representem a realidade de todos? Justifique a escolha.

**Item 7** – Analisar amostra

Um Instituto estava interessado em saber quais gêneros literários os brasileiros costumam ler. Para isso, realizou-se uma pesquisa por amostra. Qual das opções é mais adequada para a realização da pesquisa em questão? Por quê?

Opção A: 5120 brasileiros que frequentam biblioteca.

Opção B: 5120 brasileiros de diferentes faixas etárias.

Opção C: 512 brasileiros que frequentam biblioteca.

Opção D: 512 brasileiros de diferentes faixas etárias.

Opção E: Homens, mulheres, meninos e meninas.

## 5. RESULTADOS

Observamos um desempenho muito fraco dos estudantes de todos os anos, apesar de uma melhora de acordo com a escolaridade (Tabela 3). Lembrando que, a pontuação máxima é de 9 pontos. Assim, parece que conceitos relacionados a amostragem têm sido pouco explorados em sala de aula ou não foram explorados. No Brasil, apenas em 2019, conceitos relacionados à amostragem foram incluídos no currículo nos diferentes anos escolares. Acrescido a isso, no início de 2020 teve início a pandemia da COVID-19 que prejudicou muito a aprendizagem dos estudantes. Portanto, acreditamos que pesquisas futuras podem indicar um melhor desempenho.

**Tabela 3**

*Média de acerto global por ano escolaridade*

Ano de escolaridade	Média de acerto global
6º ano (11 – 12 anos)	0,88
7º ano (12 – 13 anos)	1,40
8º ano (13 – 14 anos)	2,13



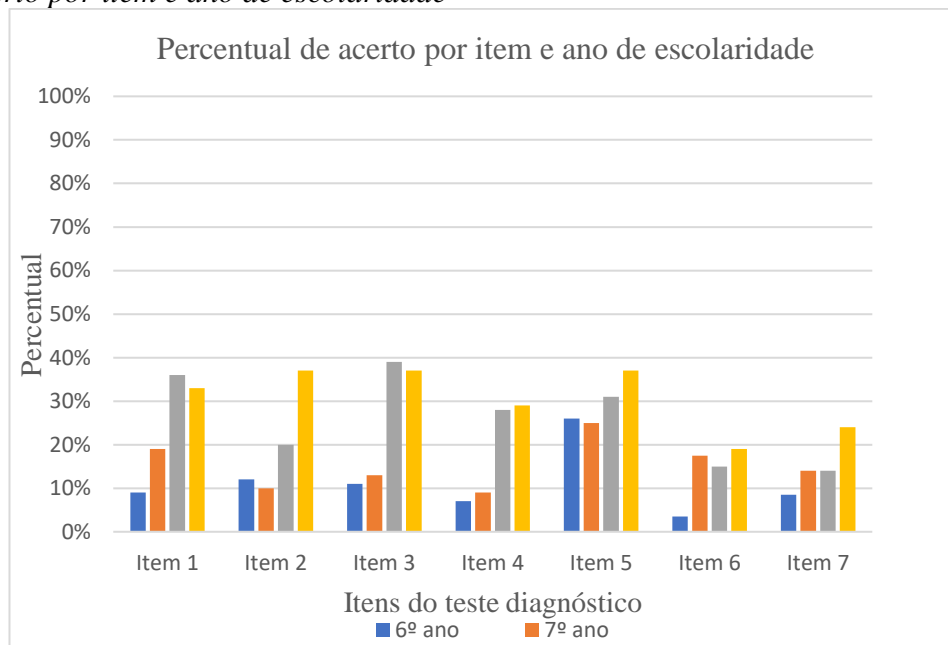
Ano de escolaridade	Média de acerto global
9º ano (14 – 15 anos)	2,59

Para analisar o desempenho em função da escolaridade, foi realizada uma Análise de Variância (ANOVA), a qual constatou diferença significativa em função dos anos escolares [ $F(3,306) = 47,743$ ;  $p = .000$ ]. Segundo o pós-teste Tukey, não foram encontradas diferenças significativas entre o 6º e 7º ano ( $p = .197$ ), entre o 7º e 8º ano ( $p = .059$ ) e entre o 8º e 9º ano ( $p = .0361$ ). Por outro lado, tiveram diferenças significativas ( $p = .000$ ) entre 6º e 8º ano, 6º e 9º ano, e, 7º e 9º ano. Acreditamos, que a experiência escolar e de vida podem ser fatores que influenciam o desempenho dos estudantes, à medida que eles têm acesso a novos conceitos e contextos que os ajudam nessa aprendizagem.

Ao relacionarmos os itens do teste diagnóstico e os anos escolares (Figura 1) foi possível verificar que desde o 6º ano de escolaridade, em todos os itens, encontramos estudantes sendo capazes de responder de forma correta, indicando a possibilidade de aprendizagem desde esse ano escolar. Esses dados reforçam que a recente inserção no currículo brasileiro e a pandemia podem ter sido fatores determinantes no desempenho.

### Figura 1

Percentual de acerto por item e ano de escolaridade



*Nota.* Item 1: Conceituar amostra; Item 2: Estabelecer população e amostra em contexto de pesquisa com pessoas; Item 3: Estabelecer população e amostra em contexto de pesquisa com animais; Item 4: Estabelecer população e amostra em contexto de pesquisa com estabelecimentos; Item 5: Justificar a realização de uma pesquisa amostral; Item 6: Construir amostra; Item 7: Analisar amostra.

Observamos que no item 5 (justificar a realização de uma pesquisa amostral) o desempenho dos estudantes foi muito próximo entre os anos de escolaridade. O desempenho dos estudantes do 8º e 9º ano são próximos em todos os itens, com exceção do item 7 (avaliar amostra). É importante ressaltar que os itens 6 e 7, os quais envolviam, respectivamente, construir e analisar amostra, foram os que apresentaram menor percentual de acerto para todos os anos.

Procurando compreender melhor os resultados, passamos a analisar de forma mais qualitativa o desempenho dos estudantes em cada item. O item 1 tinha por objetivo compreender o que eles sabiam sobre amostra (Tabela 4).

#### Tabela 4

*Percentual por categoria e ano escolar: Você já ouviu a palavra amostra antes? O que ela significa?*

Categorias de respostas	6º ano	7º ano	8º ano	9º ano
Não sabe/Não responde	11,8	10,3	0	1,5
Resposta não relacionada ao item	19,6	5,9	3,1	0
Responde não e não define amostra	19,6	11,8	12,5	6,8
Responde sim e não define amostra	7,8	17,6	12,5	21,9
Responde sim e define como verbo mostrar	10,8	16,2	17,2	16,4
Responde sim e apresenta exemplos	21,6	19,1	18,8	20,5
Responde sim e define corretamente	8,8	19,1	35,9	32,9

Como podemos observar, o percentual de respostas adequadas – definir amostra como parte de um todo – aumenta conforme avança a escolaridade (com pequena diminuição no 9º ano). Ressaltamos que essa compreensão – parte de um todo – precisa ser mais bem explicitada e compreendida no sentido de ser uma parte que tenha todas as características do todo. Mas, por enquanto, estamos considerando que compreendem amostra quando se referem que é uma parte do todo. Vejamos alguns exemplos:

*Estudante 11 (6º ano):* Sim, é uma coisa pequena (pedaço) de uma coisa grande.

*Estudante 123 (7º ano):* Amostra é uma quantidade pequena de uma população.

*Estudante 179 (8º ano):* Acho que significa que é um fragmento de algo.

*Estudante 229 (9º ano):* Sim, significa uma parcela significativa (ou não) de algo maior que ela.

É perceptível também, que muitos dos estudantes apresentam uma resposta parcialmente adequada quando apresentam exemplos de amostras oriundas do cotidiano (sangue, perfume, comida etc.). Resultados semelhantes foram encontrados em Gomes e Guimarães (2018); Meletiou-Mavrotheris e Papparistodemou (2015); Ruiz-Reyes e Contreras-García (2021) e, Watson e Moritz (2000). Tais estudantes apresentam noções associadas a amostras homogêneas, o que pode contribuir para uma compreensão errônea de que não há variabilidade na população que precisa estar representada na amostra.

*Estudante 4* (6º ano): Sim, significa que pode ser uma amostra de sangue de perfume comida etc.

Encontramos ainda, estudantes que associam amostra com o verbo mostrar.

*Estudante 13*: Sim, significa amostrar algo a alguma pessoa.

Esse tipo de justificativa corrobora com a pesquisa de Gomes (2013) realizada com estudantes dos anos iniciais que evidencia a influência de características linguísticas regionais na construção de conceitos, uma vez que essa palavra “amostrar” é muitas vezes utilizada como sinônimo de mostrar por pernambucanos. No entanto, estudantes chilenos também apresentaram respostas que revelam esse aspecto: “Sí, para mi significa mostrar, enseñar, presentar” (Ruiz-Reyes, 2019, p. 185).

Os itens 2, 3 e 4 do teste diagnóstico objetivavam identificar se os estudantes estabeleciam adequadamente a relação entre população e amostra em contextos de pesquisa que envolviam pessoas, animais e estabelecimentos.

## Tabela 5

*Percentual de acertos nos itens 2, 3 e 4 por ano de escolaridade*

	Identifica a população				Identifica a amostra			
	6º ano	7º ano	8º ano	9º ano	6º ano	7º ano	8º ano	9º ano
Pessoas	25,5	2%	21,9	41,1	29,4	35,3	59,4	65,8
Animais	35,3	42,6	50	43,8	25,5	29,4	51,6	49,3
Estabelecimentos	13,7	13,2	34,4	34,2	25,5	26,5	56,2	53,4

De acordo com a Tabela 5, identificar população foi mais fácil no contexto de animais e identificar amostra teve o melhor desempenho no contexto de pessoas. Esses resultados indicam a necessidade desses conceitos serem trabalhados com diferentes contextos. Ainda assim, tivemos estudantes dos diferentes anos escolares apresentando respostas adequadas:

*Estudante 14* (6º ano): População: Todas as pessoas que utilizam o transporte público da Região Metropolitana do Recife. Amostra: As 2.400 pessoas que participaram da pesquisa.

*Estudante 177* (7º ano): População: Macacos em geral. Amostra: 14 macacos.

*Estudante 305* (9º ano): População: Todos os estabelecimentos de Pernambuco. Amostra: 20 estabelecimentos de Pernambuco.

Quando as populações não se referiam a pessoas, os estudantes ainda assim, buscavam relacionar a pessoas, como nos exemplos a seguir:

*Estudante 136* (7º ano): Pessoas que realizaram a pesquisa com os macacos.

*Estudante 201* (8º ano): Pessoas dos estabelecimentos.

*Estudante 268* (9º ano): Todas as pessoas.

Luna e Guimarães (2021b) ao analisarem 88 livros didáticos brasileiros de Matemática do 6º ao 9º ano, constataram uma baixa frequência de atividades de identificação de população e amostra em contextos que não fossem com pessoas, o que certamente leva os estudantes a acreditarem que população se refere sempre a pessoas. Ruiz-Reyes e Contreras-García (2021) também ressaltaram a importância de apresentarmos situações apropriadas quanto a linguagem e o contexto dos respondentes. Ben-Zvi, Makar, Bakker e Aridor (2011) sinalizam que a variedade de situações estimula os estudantes a pensarem sobre as relações população-amostra, acrescentando ser indispensável a intervenção do professor.

Outra justificativa que ressaltamos, é que os estudantes buscam algum número na situação proposta para se referir a população, isto é, eles sentem a necessidade de quantificar quando são solicitados a identificar população.

*Estudante 12 (6º ano): 14 macacos.*

*Estudante 106 (7º ano): 20 estabelecimentos.*

*Estudante 217 (8º ano): 2.400 pessoas que utilizam o transporte público.*

Para os itens 2, 3 e 4, apresentamos, até o momento, o percentual isolado de estudantes que identificaram a população e a amostra corretamente. No entanto, é importante realizarmos uma análise que verifique o percentual de estudantes que estabeleceram adequadamente a relação entre população-amostra, ou seja, aqueles que identificarem simultaneamente de maneira apropriada a população e a amostra (Tabela 6).

## **Tabela 6**

*Percentual de respostas nos itens 2, 3 e 4 por ano de escolaridade*

Ano de escolaridade	Estabelece adequadamente população e amostra		
	Pessoas	Animais	Estabelecimentos
6º ano	11,8	10,8	6,9
7º ano	10,3	13,2	8,8
8º ano	20,3	39,1	28,1
9º ano	36,9	36,9	28,8

Podemos perceber que estabelecer adequadamente a relação entre população e amostra, se mostrou uma habilidade difícil para os estudantes, principalmente para os do 6º e 7º anos. Chamamos atenção que estabelecer adequadamente a relação população/amostra trata-se de uma habilidade fundamental em qualquer pesquisa e certamente subsidiam na formação de um cidadão consumidor e produtor de dados (Gal, 2002).

O item 5 tinha por objetivo levantar quais justificativas os estudantes apresentavam para a realização de uma pesquisa amostral (Tabela 7). Podemos notar um alto percentual de estudantes que indicam uma justificativa inadequada, dentre elas, de que não é possível realizar uma pesquisa censitária. Esse

indicador nos revela um aspecto preocupante, pois, sugere que os estudantes não conhecem, por exemplo, pesquisas do Instituto Brasileiro de Geografia e Estatística (IBGE), o qual realiza um censo a cada 10 anos. Sobretudo, o 8º e 9º anos tiveram maiores percentuais de estudantes que deram esse tipo de justificativa.

### Tabela 7

*Percentual das categorias de respostas no item 5 por ano de escolaridade*

Categorias de respostas	6º ano	7º ano	8º ano	9º ano
Não sabe/não responde	22,6	11,8	6,3	6,8
Impossibilidade de se realizar uma pesquisa censitária	16,7	29,4	34,4	35,6
Outro. Apresenta justificativa sem sentido lógico	34,3	33,8	28,1	20,5
Tamanho da amostra ser suficiente/representativo	10,8	14,7	15,6	23,3
Inviabilidade de realizar a pesquisa devido a fatores de tempo e dinheiro	15,6	10,3	15,6	13,8

Por outro lado, há estudantes que apresentaram respostas adequadas, sejam embasadas na justificativa do tamanho da amostra ser suficiente ou representativo da população ou na inviabilidade de realizar um censo devido a fatores tempo e dinheiro. Todos os 4 anos de escolaridade apresentaram um desempenho semelhante.

*Estudante 101 (6º ano):* Pois são muitos brasileiros e com 2.732 brasileiros, já se tem uma noção.

*Estudante 164 (7º ano):* Porque se fosse entrevistar a todos os brasileiros iria levar muito tempo.

*Estudante 180 (8º ano):* Isso duraria uma eternidade.

*Estudante 262 (8º ano):* Porque entrevistar todos os brasileiros levaria uma quantidade absurda de tempo e dinheiro.

Justificativas similares foram dadas também por estudantes das diferentes etapas da educação básica (Meletiou-Mavrotheris e Papparistodemou, 2015; Ruiz-Reyes e Contreras-García, 2021; Watson e Moritz, 2000). No entanto, essa é uma habilidade que carece de maior atenção em sala de aula, uma vez que os estudantes necessitam ter clareza do papel de utilizar uma amostra, assim como, das vantagens e desvantagens.

No item 6, que contempla construir amostra (Tabela 8), observa-se que essa é uma habilidade muito difícil para os estudantes. Para a definição de uma amostra representativa é preciso considerar a variabilidade da população para definir o tamanho adequado da amostra. Quanto maior a variabilidade da população maior deve ser a amostra. Compreender todos os fatores determinantes de uma população diversificada diante do objetivo da pesquisa é difícil, mas fundamental. Estudos antecedentes já

sinalizavam a dificuldade de estudantes pensarem no planejamento de uma amostra (Gomes e Guimarães, 2018; Meletiou-Mavrotheris e Papparistodemou, 2015, Ruiz-Reyes e Contreras-García, 2021).

### Tabela 8

*Frequência das categorias por ano de escolaridade*

Categorias de respostas	6º ano	7º ano	8º ano	9º ano
Não sabe/não responde	36,3	22,0	32,8	19,2
Responde como vai para a escola	27,5	14,7	9,4	6,8
Apresenta resposta sem relação com a questão	4,9	5,9	1,6	2,7
Responde entrevistando toda a população	5,9	0	7,8	6,8
Apresenta uma forma de realizar a pesquisa (amostragem não probabilística – tendenciosa)	20,5	30,9	26,5	30,1
Considera apenas o tamanho da amostra	0	10,3	7,8	5,5
Considera apenas a variabilidade da amostra	2	5,9	6,2	21,9
Considera o tamanho e a variabilidade da amostra	2,9	10,3	7,9	6,8

Poucos foram os estudantes que deram respostas contendo os dois fatores para construção da amostra (tamanho e variabilidade). Contudo, desde o 6º ano, os estudantes evidenciam capacidade de pensar no planejamento de construção de amostra.

*Estudante 99 (6º ano):* Eu escolheria 5 estudantes de cada sala (meninos e meninas), pois assim eu teria uma estimativa do resultado.

*Estudante 118 (7º ano):* 25% de cada sala de ambos os sexos, seria viável e significativo.

*Estudante 222 (8º ano):* Entrevistaria pelo menos 3 estudantes de cada turma e ambos os gêneros para evitar que a idade e o gênero adulterassem o resultado.

*Estudante 303 (9º ano):* 15 estudantes de cada ano, porque eles viriam de classes sociais diferentes, gêneros diferentes e idades diferentes.

Percebemos também, respostas atreladas apenas a variabilidade da amostra.

*Estudante 143 (7º ano):* Pegar grupo de pessoas de idades diferentes, pois assim teria mais variedade e daria para fazer uma média.

*Estudante 270 (9º ano):* Pegar estudantes de todas as classes e sexos, para não pegar só um grupo de pessoas.

Há aqueles estudantes que apenas fizeram menção ao tamanho da amostra. Nota-se que é um fator apontado por estudantes a partir do 7º ano, que possui o maior percentual, e vai diminuindo a frequência desse tipo de resposta, conforme aumenta a escolaridade.

*Estudante 124 (7º ano):* Selecionar 50% dos estudantes.

*Estudante 203 (8º ano):* Perguntar a 50 estudantes e a partir daí criar uma média.

*Estudante 296 (9º ano):* Chamaria 100 estudantes e faria a porcentagem.

Dentre as respostas inadequadas, tem-se um alto percentual em todos os anos de escolaridade (6º ao 9º ano) de estudantes que sugerem métodos de amostragem não probabilísticos. Resultado similar foi encontrado na pesquisa de Ruiz-Reyes e Contreras-García (2021) com estudantes equivalente no Brasil aos anos finais do Ensino Fundamental (11 -15 anos de idade) e Ensino Médio (15 – 18 anos de idade).

*Estudante 219* (8º ano): Uma votação online, é prática e um local onde podem acessar.

*Estudante 257* (8º ano): Eu colocaria um cartaz no espaço da cantina da escola e quem fosse passando poderia responder a pesquisa.

Esse é um ponto de atenção, pois, acreditamos que se os estudantes propõem construção de amostras por métodos tendenciosos, há uma alta chance de não perceberem vieses de seleção em situações de análise de amostras. Portanto, enfatizamos a importância de intervenções didáticas pelo professor, propostas em livros didáticos e orientações curriculares, contemplarem essas duas grandes frentes para a aprendizagem de amostragem: análise e construção de amostras.

O último item do teste diagnóstico, aqui analisado, versava sobre análise de amostras (Tabela 9). Observa-se um baixo percentual de acerto. Essa não é uma particularidade dos estudantes brasileiros dos anos finais do Ensino Fundamental, pois, Innabi (2007) investigando alunos ainda mais velhos (17 – 18 anos de idade), aponta esse obstáculo, argumentando que eles supervalorizam o tamanho da amostra e não percebem o viés de seleção.

## Tabela 9

*Percentual das categorias de respostas no item 7 por ano de escolaridade*

Categorias de respostas	6º ano	7º ano	8º ano	9º ano
Incorreta (opções a, c e d)	87,3	83,8	79,7	67,1
Parcialmente correta (opção e com justificativa adequada)	8,8	4,4	12,5	17,8
Correta (opção b com justificativa adequada)	3,9	11,8	7,8	15,1

Dentre as respostas parcialmente adequadas, notamos o maior percentual no 9º ano.

*Estudante 30* (6º ano): Opção E. Porque ela está mais definida.

*Estudante 107* (7º ano): Opção E, pois, indica os gêneros.

Quanto às respostas adequadas, observamos variação entre os anos escolares, no entanto, o desempenho é baixo em todos eles.

*Estudante 93* (6º ano): Opção B, pois tem mais pessoas e a diversidade é maior.

*Estudante 133* (7º ano): Opção B, pois com mais pessoas e mais diversidade de idades, vamos ter uma noção maior.

*Estudante 205* (8º ano): Opção B. O gosto muda com a idade.

*Estudante 307 (9º ano):* Opção B. Pois além de ser mais pessoas, a distribuição vai ser mais fácil, visto que pode haver homens e mulheres, meninos e meninas, com gostos literários diferentes, porém da mesma faixa etária.

As respostas adequadas indicam que estudantes dos diferentes anos de escolaridade, embora em menor percentual, conseguem justificar adequadamente a opção da amostra representativa da população. Porém, conforme destaca Innabi (2007), há diversos equívocos que precisam ser desmistificados, tais como qualquer amostra pode representar a população, uma conclusão é válida (ou não é válida) em função de experiências pessoais, ou ainda, quanto maior a amostra mais válida é a conclusão.

## **6. CONCLUSÕES**

Esse estudo teve como objetivo investigar o que sabem estudantes brasileiros do 6º ao 9º ano do ensino fundamental (11 - 15 anos de idade) sobre amostra e amostragem. Para tal, foi realizado um teste diagnóstico de maneira individual.

Os resultados apontam para a capacidade de estudantes desde o 6º ano pensarem adequadamente sobre amostra e amostragem, evidenciando que é possível trabalhar com tais conceitos em sala de aula desde cedo. Entretanto, estudantes até o 9º ano ainda possuem dificuldades em noções básicas como definir amostra e apresentar exemplos em contextos do dia a dia.

O cenário também é de bastante dificuldade na habilidade identificar população e amostra e estabelecer a relação entre população-amostra. O contexto (pessoas, animais e estabelecimentos) teve influência nesses resultados, pois muitos estudantes associaram população apenas com pessoas e quando o contexto envolvia animais ou estabelecimentos comerciais associaram a pessoas que cuidavam dos animais ou que trabalhavam nos estabelecimentos, buscando sempre relacionar a pessoas.

Quanto a construção de amostras, estudantes de todos os anos de escolaridades, apresentaram predominantemente técnicas de seleção tendenciosos, poucos foram os que levaram em consideração os fatores variabilidade, tamanho e métodos probabilísticos. Sempre o tamanho da amostra foi mais valorizado.

A partir desses resultados chamamos a atenção de professores, autores de prescrições curriculares e de livros didáticos percebam a importância de explorar diferentes situações e contextos para promover o ensino de amostragem e propiciar uma aprendizagem efetiva. Com isso, esperamos que os estudantes sejam capazes de julgar a validade de uma amostra e analisar criticamente generalização estatística apresentadas.



## REFERÊNCIAS

- Ben-Zvi, D., Bakker, A. e Makar, K. (2015). Learning to reason from samples. *Educational Studies in Mathematics*, 88(3), 291-303.
- Brasil (2017). *Base Nacional Curricular Comum*. MEC.
- Gomes, T. M. (2013). *O todo é a soma das partes, mas uma parte representa o todo? Compreensão de Estudantes do 5º e 9º ano sobre Amostragem*. [Dissertação de Mestrado, Centro de Educação - Universidade Federal de Pernambuco].
- Gomes, T. M. e Guimarães, L. G. (2018). Compreensão dos estudantes do ensino fundamental sobre seleção de uma amostra representativa. *Com a Palavra, o Professor*, 3(6), 132-149.
- Innabi, H. A. (2007) Factors Considered by Secondary Students When Judging the Validity of a Given Statistical Generalization. *International Electronic Journal of Mathematics Education*, 2(3), 168-186.
- Luna, L. C. e Guimarães, G. L. (2021a). O que livros didáticos de Matemática propõem para a aprendizagem de amostragem? *Bolema*, 35(70), 815-839.
- Luna, L. C. e Guimarães, G. L. (2021b). Aprendizagem de amostragem nos PCN e na BNCC e a influência nos livros didáticos. *Revista de Educação, Ciências e Matemática*, 11(1), 1-20.
- Meletiou-Mavrotheris, M. e Paparistodemou, E. (2015). Developing students' reasoning about samples and sampling in the context of informal inferences. *Educational Studies in Mathematics*, 88(3), 385-404.
- Moore, D. S. (1995) *A Estatística básica e sua prática* (3ra. ed). W. H. Freeman and Company.
- Pedrosa, N. B. (2019). *Comprensión del muestreo y la distribución muestral en estudiantes de educación secundaria obligatoria y bachillerato*. [Tese Doutorado, Universidad de Granada].
- Ruiz-Reyes, K. (2019). *Comprensión del muestreo por alumnos chilenos de educación secundaria*. [Tesis doctoral, Universidad de Granada].
- Ruiz-Reyes e Contreras-García, J.M. (2021). Understanding sampling by Chilean secondary school students. *Statistics Education Research Journal*, 20, 9-26.
- Ruiz-Reyes, K., Begué, N., Batanero, C. e Contreras-García, J. M. (2017). Un estudio comparado de los contenidos de muestreo en la Educación Secundaria Obligatoria en Chile. *Educación Matemática Pesquisa*, 19(3), 67-83.
- Stevenson, W. J. (1981). *Estatística aplicada à administração*. Harper & Row do Brasil.
- Watson, J. M. e Moritz, J. B. (2000). Developing concepts of sampling. *Journal for Research in Mathematics Education*, 31(1), 44-70.

### Como citar:

Luna, L. y Guimarães, G. (2022). Compreensão de amostra e amostragem por estudantes brasileiros dos anos finais do ensino fundamental. *Revista de Educación Estadística*, 1(1), 1-17. <https://doi.org/10.29035/redes.1.1.7>



Esta obra está bajo una licencia de Creative Commons Reconocimiento-NoComercial-CompartirIgual 4.0 Internacional.