



www.ufpe.br/ppgedumatec
e-mail: edumatec@ufpe.br
tel: 21268952

TÂMARA MARQUES DA SILVA GOMES

O TODO É A SOMA DAS PARTES, MAS UMA PARTE REPRESENTA O TODO?

Compreensão de Estudantes do 5º e 9º Ano sobre Amostragem

RECIFE

2013

TÂMARA MARQUES DA SILVA GOMES

O TODO É A SOMA DAS PARTES, MAS UMA PARTE REPRESENTA O TODO?

Compreensão de Estudantes do 5º e 9º Ano sobre Amostragem

Dissertação apresentada ao Programa de Pós-Graduação em Educação Matemática e Tecnológica da Universidade Federal de Pernambuco, como requisito parcial para obtenção do título de Mestre em Educação Matemática e Tecnológica.

Orientadora: Gilda Lisbôa Guimarães

RECIFE

2013

Catálogo na fonte
Bibliotecária Andréia Alcântara, CRB-4/1460

G633t Gomes, Tâmara Marques da Silva.
O todo é a soma das partes, mas uma parte representa o todo? :
compreensão de estudantes do 5º e 9º ano sobre amostragem / Tâmara
Marques da Silva Gomes. – Recife: O autor, 2013.
109 f. : il. ; 30 cm.

Orientadora: Gilda Lisbôa Guimarães.
Dissertação (Mestrado) - Universidade Federal de Pernambuco, CE.
Programa de Pós-graduação em Educação Matemática e Tecnológica,
2013.
Inclui Referências e Apêndices.

1. Estatística educacional. 2. Matemática - Estudo e ensino. 3.
Amostragem. 4. UFPE - Pós-graduação. I. Guimarães, Gilda Lisbôa.
II. Título.

370.21 CDD (22. ed.) UFPE (CE2013-016)



ALUNA

TÂMARA MARQUES DA SILVA GOMES

TÍTULO DA DISSERTAÇÃO

“O TODO É A SOMA DAS PARTES, MAS UMA PARTE REPRESENTA O TODO? COMPREENSÃO DE ESTUDANTES DO 5º E 9º ANO SOBRE AMOSTRAGEM”

COMISSÃO EXAMINADORA:

Presidente e Orientadora
Profa. Dra. Gilda Lisbôa Guimarães

Examinadora Externa
Profa. Dra. Claudia Regina Oliveira de Paiva
Lima

Examinadora Externa
Profa. Dra. Lisbeth Kaiserlian Cordani

Examinadora Interna
Profa. Dra. Verônica Gitirana Gomes Ferreira

Recife, 15 de março de 2013.

À minha mãe, meu maior exemplo de força e perseverança.

Ao meu tio Jocélio (in memoriam) que sempre me estimulou a aprender.

Aos professores que tive ao longo da minha vida, em especial a Gilda, minha orientadora, pela motivação e compreensão durante a pesquisa.

AGRADECIMENTOS

“Escolho meus amigos não pela pele ou outro arquétipo qualquer, mas pela pupila. Tem que ter brilho questionador e tonalidade inquietante. [...] Fico com aqueles que fazem de mim louco e santo. Deles não quero resposta, quero meu avesso. Que me tragam dúvidas e angústias e aguentem o que há de pior em mim. [...] Escolho meus amigos pela cara lavada e pela alma exposta. Não quero só o ombro ou o colo, quero também sua maior alegria. [...] Meus amigos são todos assim: metade bobeira, metade seriedade. [...] Quero amigos sérios, daqueles que fazem da realidade sua fonte de aprendizagem, mas lutam para que a fantasia não desapareça...”

Oscar Wilde

Agradeço, primeiramente, a Deus, pois colocou as pessoas certas ao meu lado durante esse percurso e me deu força e tranquilidade para confiar e saber que tudo daria certo.

À minha mãe, por mesmo sem compreender o que realmente é um mestrado, sempre me motivou e esteve ao meu lado em minhas decisões.

À minha irmã Tatiana e meu cunhado André, pelo apoio e por suportarem todo o meu estresse, tentando diminuí-lo.

Aos meus familiares, por serem meu suporte e estarem sempre presentes.

Agradeço à Gilda, minha orientadora, que me deixou “livre” durante a pesquisa e, quando preciso, “puxou minha orelha”. Foi uma amiga me incentivando e ensinando a refletir sobre minha pesquisa.

Às minhas amigas Érica, Edneri e Natércia que foram excelentes ouvintes nas horas de desespero e nas discussões sobre as nossas pesquisas.

À Taciana Santos e Josué Marques, meus companheiros e psicólogos durante grande parte desses dois anos.

Aos meus companheiros de trabalho, mas especificamente à gestão, pela flexibilidade e compreensão durante o mestrado.

Ao Grupo de Estudos em Educação Estatística no Ensino Fundamental (GREF), que ajudou na inspiração e construção deste estudo.

Aos pesquisadores Verônica Gitirana, Cláudia Lima, Lisbeth Cordani e Antonio Roazzi pela disponibilidade e interesse em contribuir com sugestões e outras análises na pesquisa.

Aos pesquisadores dos Processos de Ensino Aprendizagem da Educação Matemática e Científica, que durante as disciplinas de Seminários também contribuíram bastante com todos os processos dessa pesquisa.

Aos mestrandos do EDUMATEC, especialmente a turma de 2011, sempre dispostos a discutir e aprender mais.

Ao corpo de professores e secretaria da EDUMATEC, cujos ensinamentos e orientações foram importantes na conclusão de tal pesquisa.

Agradeço também à direção das escolas, ao corpo docente e principalmente aos estudantes que participaram desse estudo pela sua disponibilidade e vontade de ajudar.

RESUMO

Ser capaz de analisar, elaborar conclusões e tomar decisões a partir de informações estatísticas são habilidades necessárias atualmente. Para realização dessas atividades, entender conceitos básicos de amostragem é essencial. Embora o ensino da Estatística seja recomendado no currículo escolar brasileiro poucos estudos se preocuparam em pesquisar de modo sistemático e abrangente a compreensão de estudantes do Ensino Fundamental quanto aos conceitos sobre amostragem. Assim, o objetivo do presente estudo foi identificar o que estudantes do 5º e 9º ano do Ensino Fundamental compreendem sobre amostragem. Foram realizadas, individualmente, 40 entrevistas semi-estruturadas com alunos do 5º e 9º ano (vinte de cada ano), contendo treze questões que abordaram diferentes aspectos da amostragem: definição, exemplo, finalidade, seleção, tamanho e representatividade da amostra; definição de população; conceito de aleatoriedade, amostra aleatória e sua utilização; realização de inferências informais a partir de uma amostra. Os resultados evidenciaram que não houve diferença significativa entre o desempenho dos estudantes do 5º e do 9º ano, demonstrando que a escolaridade não foi fator determinante para a adequação das respostas. Os participantes mostraram uma maior facilidade nas questões que abordavam o conceito de população como grupo de pessoas, o tamanho da amostra e a realização de inferências informais a partir de uma amostra. A possibilidade de responder a partir de sua experiência de vida foi um fator importante para os estudantes do 5º ano. Notou-se que os alunos do 9º ano tentaram relacionar seus conhecimentos prévios com o conhecimento escolar, buscando apresentar respostas mais formalizadas e estruturadas para as perguntas que envolviam definição de conceitos. Percebeu-se que as questões que envolviam aspectos referentes à aleatoriedade, representatividade e realização de inferências, as quais, de certa forma, estão ligadas à seleção da amostra, apresentaram maior correlação, o que pode indicar que esses conceitos necessitam de habilidades semelhantes para compreensão. Ao contrário do que era esperado, tanto no 5º como no 9º ano, não foram os mesmos estudantes que responderam correta ou incorretamente as questões sobre a definição do conceito de amostra e população. Acreditava-se que, por requerer habilidades semelhantes, haveria uma correlação entre as mesmas. Entretanto, os resultados sugerem que o contexto da questão é determinante. Assim, essa pesquisa evidenciou que, apesar das grandes dificuldades apresentadas pelos alunos para compreender os conceitos ligados à amostragem, estudantes desde o 5º ano de escolaridade já são capazes de compreendê-los. Essa aprendizagem pode ser potencializada se tais conteúdos forem trabalhados de forma sistemática e contextualizada na escola, a partir de vivências e situações de ensino que os desafiem a analisar e refletir sobre informações estatísticas.

Palavras-chave: Educação Estatística; Educação Matemática; Amostra; População; Ensino Fundamental.

ABSTRACT

You must be able to analyze, draw conclusions and make decisions based on statistical information. To carry out these activities, understand the basics of sampling is essential. Although being recommended the teaching of statistics in the Brazilian curriculum few studies have focused on systematic and comprehensive research on the understanding of elementary school students about the concepts of sampling. Thus, the objective of this study was to identify what students of the 5th and 9th grades of elementary school understand about sampling. 40 semi-structured interviews were conducted individually, with students in the 5th and 9th grade (20 in each grade), containing thirteen questions that addressed different aspects of sampling: definition, example, purpose, selection, size and representativeness of the sample; definition of population; concept of randomness, random sample and its use; informal inferences from a sample. The results showed that there was no significant difference between the performance of students in 5th and 9th grade, demonstrating that schooling was not a determining factor for the adequacy of the answers. Participants showed a greater facility in questions that addressed the concept of population as a group of people, the sample size and informal inferences from a sample. The ability to respond from their life experience was an important factor for students in 5th grade. It was noted that students in 9th grade tried to relate their previous knowledge with school knowledge, seeking to introduce more formalized and structured answers to questions involving definition of concepts. It was perceived that issues involving aspects related to randomness, representativeness and inferences, which in a way, are connected to sample selection, produced the strongest correlations, which may indicate that these concepts require similar skills for understanding. Unlike what was expected, both the 5th and the 9th grade, were not the same students who answered the questions correctly or incorrectly about the definition of sample and population. It was believed that, because it requires similar skills, there would be a correlation there between. However, the results suggest that the context of the question is also decisive. Thus, this research showed that, despite the great difficulties presented by the students to understand the concept of sampling, students from the 5th grade are already capable of understanding aspects about sampling. This learning can be increased if such contents are worked systematically and contextualized in the school, from teaching situations and life experiences that challenge them to analyze and reflect on statistical information.

Keywords: Statistical Education; Mathematics Education; Sample; Population; Elementary School.

SUMÁRIO

INTRODUÇÃO	10
CAPÍTULO 1 – REVISÃO DA LITERATURA	13
Educação Estatística	13
Amostragem	19
<i>Tipos de amostra</i>	21
<i>Amostragem Probabilística</i>	22
Amostra Aleatória Simples (AAS)	22
Amostra Sistemática	24
Amostra Estratificada	24
Amostra por Conglomerado	25
<i>Amostragem Não-probabilística</i>	26
Amostra por Julgamento	27
Amostra por Conveniência	27
Amostra por Cotas	28
<i>Tamanho da Amostra</i>	28
Os Conceitos de Amostra e População na Educação Básica	29
CAPÍTULO 2 – MÉTODO	37
Participantes	37
Procedimentos	38
Categorização e análise dos dados	44
CAPÍTULO 3 – DESCRIÇÃO E ANÁLISE DOS RESULTADOS	46
Análise Multidimensional	85
CAPÍTULO 4 – CONCLUSÕES	93
REFERÊNCIAS	102
APÊNDICE A – Cartões do instrumento de coleta	105
ANEXO A – Tabela de Números Aleatórios	109

INTRODUÇÃO

Nos dias atuais, observa-se que a sociedade sofre constantes transformações ocasionadas, entre outros fatores, pelos impactos da globalização, mudanças essas que atingem todas as esferas sociais, principalmente a educação. Na atual sociedade do conhecimento e da informação, a prática pedagógica deve ser problematizadora e contextualizada, estimulando o aluno a desenvolver um pensamento crítico, capaz de solucionar problemas.

Diariamente, um grande número de informações nos é disponibilizado através das diferentes mídias, sendo necessário conhecimento específico para tratá-las de forma adequada. Percebe-se a propagação e utilização de diferentes linguagens e conceitos estatísticos diariamente, além de afirmações e conclusões baseadas em dados estatísticos.

Segundo Estevam (2010) é essencial saber representar, analisar e questionar os dados apresentados. Diante desse contexto, é fundamental fazer com que o aluno desenvolva procedimentos e estratégias para interpretar as informações que aparecem frequentemente em seu cotidiano.

Vale ressaltar que, embora seja notória a presença da Estatística em nosso dia a dia, muitos não a reconhecem ao analisarem diversas informações e dados veiculados através dos diferentes meios de comunicação. Todos nós estamos nos deparando constantemente com diferentes informações expostas por meio de gráficos e tabelas que, para análise e compreensão crítica, demandam noções básicas de Estatística (CAMPOS, 2005). Sendo assim, observa-se a necessidade de discussões relacionadas à abordagem da Estatística e seus principais conceitos já no início da escolarização.

Tendo como base essa preocupação, foi sendo constituída uma área de estudo denominada Educação Estatística. Essa área estuda os problemas ligados ao ensino e à aprendizagem de conceitos estatísticos e probabilísticos considerando todo o processo de escolarização.

Dessa forma, o ensino da Estatística passa a ser uma recomendação do currículo brasileiro da Educação Básica. Entretanto, ainda são poucos os estudos destinados à compreensão mais aprofundada do desenvolvimento de conceitos essenciais da Estatística pelos estudantes.

Embora a aprendizagem da Estatística envolva a compreensão de diversos conceitos, nesta pesquisa serão abordados conceitos relacionados à amostragem, visto que são fundamentais para o desenvolvimento do pensamento e pesquisas estatísticas e que ainda são poucos os estudos na Educação Básica referentes a estes conteúdos.

A pertinência de estudos referentes ao conceito de amostra é notória, pois o mesmo se faz presente de forma implícita no dia a dia dos estudantes. Por exemplo, quando tomamos uma colher de sopa para saber se esta precisa de mais sal ou quando compramos um perfume após ter recebido uma amostra grátis do mesmo estamos nos baseando na amostragem para tomarmos essas decisões.

Diante da importância e das diversas necessidades de se desenvolver competências ligadas ao raciocínio estatístico, especificamente concepções acerca de amostras e populações, esta pesquisa teve como objetivo responder a seguinte questão: O que estudantes do 5º e 9º ano do Ensino Fundamental compreendem sobre amostragem?

Como objetivos específicos buscou:

- Investigar se os estudantes sabem conceituar amostra;
- Analisar se os estudantes ao opinarem sobre a representatividade de uma amostra consideram o seu tamanho, seleção e variabilidade;
- Analisar se os estudantes compreendem que população é o universo a ser avaliado, envolvendo tanto pessoas como objetos inanimados;
- Verificar se os estudantes realizam inferências informais a partir de uma amostra dada;
- Identificar se há diferenças entre a compreensão de estudantes de diferentes níveis de escolaridade sobre amostragem.

Assim, no Capítulo 1 apresenta-se a revisão da literatura, buscando salientar a importância da Educação Estatística na Educação Básica, bem como uma breve explanação sobre amostragem e as vantagens da exploração de conceitos como amostra e população com estudantes nesse nível de escolaridade.

No Capítulo 2 descreveu-se o método adotado na pesquisa, detalhando participantes, procedimentos e instrumento para coleta de dados.

Os resultados e discussão dos resultados são apresentados no Capítulo 3, que traz uma análise descritiva do desempenho dos estudantes e também uma

análise multidimensional, correlacionando os dados a fim de perceber o comportamento dos sujeitos nas questões de modo mais global.

Por fim, no Capítulo 4 são expostas considerações a respeito de todo o estudo, nas quais ressaltam-se os principais resultados encontrados, bem como suas contribuições para o ensino e para a pesquisa da temática em discussão.

CAPÍTULO 1

REVISÃO DA LITERATURA

Educação Estatística

O ensino da Estatística deve possibilitar ao indivíduo o entendimento dos fenômenos sociais a partir da interpretação de dados e informações, levando os mesmos a serem letrados estatisticamente. Gal (2002) define o letramento estatístico como:

a) competência da pessoa para interpretar e avaliar criticamente a informação estatística, os argumentos relacionados aos dados ou aos fenômenos estocásticos, que podem se apresentar em qualquer contexto e, quando relevante, b) competência da pessoa para discutir ou comunicar suas reações para tais informações estatísticas, tais como seus entendimentos do significado da informação, suas opiniões sobre as implicações desta informação ou suas considerações acerca da aceitação das conclusões fornecidas (GAL, 2002, p. 2-3).

Uma pessoa é considerada letrada estatisticamente, segundo Gal (2002), quando apresenta algumas capacidades: percebe a necessidade de trabalhar com dados (compreendendo que os dados não são unicamente números, mas números inseridos num determinado contexto), conhecendo sua proveniência e a forma de produzi-los; está familiarizado com os termos e ideias básicas da Estatística Descritiva; reconhece os termos e conceitos básicos relacionados às apresentações gráficas e tabulares; compreende noções básicas de probabilidade; entende o mecanismo do processo inferencial ao tomar decisões estatísticas.

Para Garfield e Gal (1999) apud Estevam e Fürkotter (2010) ao construírem conhecimento estatístico os estudantes são capazes de questionar a validade de representações e interpretações de dados e de generalizações realizadas a partir de um único estudo e/ou de amostras pequenas.

O ensino da Probabilidade e da Estatística contribui para a formação dos estudantes, preparando-os para a realidade, pois favorecem o desenvolvimento e a elaboração de questões de uma investigação, de conjecturas e hipóteses, etapas essas que são necessárias à resolução de problemas.

Assim, o currículo escolar brasileiro para os anos iniciais recomenda o ensino da Estatística. Nos Parâmetros Curriculares Nacionais (PCN) (Brasil, 1997), os conceitos e procedimentos relacionados à Probabilidade, Estatística e Combinatória encontram-se agrupados no bloco Tratamento da Informação. Os PCN sugerem que esses conteúdos devem ser trabalhados em todo o processo escolar, desde a Educação Infantil. Segundo este documento, esses conteúdos devem ser abordados em seus aspectos conceituais, procedimentais e atitudinais, sendo exigência na atualidade:

[...] saber ler e interpretar dados apresentados de maneira organizada e construir representações, para formular e resolver problemas que impliquem o recolhimento de dados e a análise das informações. Essa característica da vida contemporânea traz ao currículo de Matemática uma demanda em abordar elementos da estatística, da combinatória e da probabilidade, desde os ciclos iniciais (BRASIL, 1997, p. 132).

Em relação aos conteúdos inseridos no referido bloco, este documento propõe que sejam desenvolvidas pelos alunos as seguintes habilidades:

a) No primeiro ciclo (1º ao 3º Ano): leitura e interpretação de informações contidas em imagens; coleta e organização de informações; criação de registros pessoais para comunicação de informações coletadas; exploração da função do número como código numérico na organização de informações; interpretação e elaboração de listas, tabelas simples, tabelas de dupla entrada e gráficos de barra para comunicar a informação obtida; produção de textos escritos a partir da interpretação de gráficos e tabelas (BRASIL, 1997, p.52).

b) No segundo ciclo (4º e 5º Ano): coleta, organização e descrição de dados; leitura e interpretação de dados apresentados de maneira organizada e construção dessas representações; interpretação de dados apresentados por meio de tabelas e gráficos, para identificação de características previsíveis ou aleatórias de acontecimentos; produção de textos escritos, a partir da interpretação de gráficos e tabelas; construção de gráficos e tabelas com base em informações contidas em textos jornalísticos, científicos ou outros; obtenção e interpretação de média aritmética; exploração da ideia de probabilidade em situações-problema simples, identificando sucessos possíveis, sucessos certos e as situações de "sorte";

utilização de informações dadas para avaliar probabilidades; identificação das possíveis maneiras de se combinar elementos de uma coleção e de contabilizá-las usando estratégias pessoais (BRASIL, 1997, p.61-62).

c) No terceiro ciclo (6^o e 7^o Ano): coleta, organização de dados e utilização de recursos visuais adequados (fluxogramas, tabelas e gráficos) para sintetizá-los, comunicá-los e permitir a elaboração de conclusões; leitura e interpretação de dados expressos em tabelas e gráficos; compreensão do significado da média aritmética como um indicador da tendência de uma pesquisa; representação e contagem dos casos possíveis em situações combinatórias; construção do espaço amostral e indicação da possibilidade de sucesso de um evento pelo uso de uma razão (BRASIL, 1998, p. 74-75).

d) No quarto ciclo (8^o e 9^o Ano): leitura e interpretação de dados expressos em gráficos de colunas, de setores, histogramas e polígonos de frequência; organização de dados e construção de recursos visuais adequados, como gráficos (de colunas, de setores, histogramas e polígonos de frequência) para apresentar globalmente os dados, destacar aspectos relevantes, sintetizar informações e permitir a elaboração de inferências; compreensão de termos como frequência, frequência relativa, amostra de uma população para interpretar informações de uma pesquisa; distribuição das frequências de uma variável de uma pesquisa em classes de modo que resuma os dados com um grau de precisão razoável; obtenção das medidas de tendência central de uma pesquisa (média, moda e mediana), compreendendo seus significados para fazer inferências; construção do espaço amostral, utilizando o princípio multiplicativo e a indicação da probabilidade de um evento por meio de uma razão; elaboração de experimentos e simulações para estimar probabilidades e verificar probabilidades previstas (BRASIL, 1998, p. 90).

Percebe-se que o espírito investigativo deve estar presente no estudo em Estatística. Segundo os PCN (Brasil, 1998) o estudante deve construir procedimentos para coletar, organizar, comunicar e interpretar dados, utilizando representações gráficas, tabelas, entre outras.

Para que os cidadãos possam tomar decisões baseados em dados e informações é necessário que estes possuam habilidades críticas relacionadas à

Estatística. A capacidade de interpretar e avaliar criticamente uma informação estatística é fundamental para a tomada de decisão.

Desse modo, a Educação Estatística deve possibilitar às crianças a observação de situações de incerteza e o desenvolvimento do raciocínio combinatório, permitindo-lhes reconhecer, organizar e representar informações (LOPES, 2003).

Logo, os conhecimentos de Estatística são essenciais para atuarmos na sociedade, pois permitem analisar índices de custo de vida, realizar sondagens, escolher amostras e outras situações cotidianas. Diante destas necessidades, não podemos considerar que o ensino de conceitos estatísticos seja destinado apenas ao Ensino Médio, pois estaríamos privando o estudante da compreensão de situações ocorrentes em sua vida dentro da realidade social.

Vale destacar, que não é suficiente que o cidadão compreenda porcentagens expostas em índices estatísticos como taxas de inflação e desemprego. Faz-se necessário analisar criticamente os dados, ponderando a sua veracidade. Sendo assim, nossos alunos precisam desenvolver tanto a capacidade de organizar e representar dados, como interpretá-los e compará-los para tirada de conclusões.

Lopes (1998) defende que a metodologia da resolução de problemas torna o ensino da Estatística mais significativo, pois ao elaborar uma questão de investigação, o estudante seleciona estratégias para respondê-la, levando-o a coletar, organizar, representar e analisar os dados. Dessa forma, os estudantes estão mais propensos ao desenvolvimento do pensamento crítico.

Salienta-se que o Tratamento da Informação na formação escolar não é abordado apenas no Brasil. Ele também é abordado pelo Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report: a Pré-K-12 Curriculum Framework, aprovado pela Associação Americana de Estatística (American Statistical Association – ASA), o qual tem a Letramento Estatístico como principal objetivo da Educação Estatística, pois consideram que vivenciamos cotidianamente situações que necessitam de um raciocínio estatístico adequado (ASA, 2005).

Alguns objetivos para a Educação Estatística também foram propostos por Garfield e Gal (1999) apud Estevam (2010), entre eles: entender o propósito, a

lógica e o processo de investigações estatísticas, aprimorar habilidades procedimentais, desenvolver habilidades interpretativas e alfabetização estatística¹.

Segundo Gal (2002):

A alfabetização estatística de pessoas envolve estar alfabetizado nos conhecimentos de estatística e científicos, nos quais o conhecimento é composto de cinco elementos cognitivos: habilidades de alfabetização, conhecimento estatístico, conhecimento matemático, conhecimento de contexto e questão crítica; Componente de Disposição composto de dois elementos: posição crítica, e convicção e atitudes. (GAL, 2002, p. 4)

A alfabetização estatística torna o cidadão capaz de interpretar criticamente uma informação estatística em diversos contextos, compreendendo seus significados, elaborando opiniões diante dessas informações, estando sempre atento às informações apresentadas, evitando uma interpretação errônea dos fatos.

Watson (2003) descreve seis níveis de compreensão da alfabetização estatística a partir da capacidade de avaliar estatisticamente situações em diferentes contextos. O nível 1 corresponde ao *entendimento indiossincrático*, no qual as concepções são baseadas em experiências pessoais não relacionadas com conceitos estatísticos importantes como amostragem, probabilidade e inferências; o nível 2, de *entendimento informal*, apresenta compreensão de conceitos simples; o 3º nível, apresenta um *entendimento inconsistente*, pois nesse mais de uma característica do conceito estatístico é levada em consideração em suas respostas; no nível 4 os sujeitos já demonstram um *entendimento consistente*, porém sem crítica. Suas definições possuem mais características dos conceitos estatísticos e levam em consideração o contexto que está inserido; o nível 5 corresponde ao *entendimento crítico*, pois analisam criticamente os contextos e utilizam algumas terminologias estatísticas em suas definições; no nível 6, os estudantes alcançam um *entendimento matemático crítico*.

Para Watson (2003) o nível 6 é o objetivo a ser alcançado pelos estudantes durante a fase escolar, mas para isso é necessário que os conceitos estatísticos sejam trabalhados desde os primeiros níveis. É importante conhecer as dificuldades e compreensões iniciais dos estudantes para que estas sejam desenvolvidas durante a escolaridade.

¹ Entenda-se Alfabetização Estatística como sinônimo para Letramento Estatístico. Optou-se pela utilização do termo por se tratar de citação direta.

Vale salientar, que a concepção de Estatística abordada nesta pesquisa não se caracteriza pela realização de cálculos, exercícios mecânicos ou aplicação de fórmulas e construção de gráficos e tabelas sem significado algum para o estudante, porquanto consideramos que esses, por si só, não possibilitarão o desenvolvimento do pensamento estatístico crítico, o qual visa a utilização desses conceitos para solucionar problemas.

Logo, é primordial que o ensino de conceitos estatísticos esteja fundamentado em problemas significativos para os alunos, envolvendo-os no processo exploratório de investigação, viabilizando o desenvolvimento da interpretação, reflexão e aplicação de conceitos matemáticos no cotidiano, tornando-os mais próximos do aprendiz.

A participação dos estudantes possibilita a familiarização com as etapas pelas quais se passa uma investigação (formular uma pergunta; planejar um estudo; coletar; organizar e analisar dados; interpretar e discutir descobertas e suas implicações), o que facilita o processo de tomada de decisão e retirada de conclusões (LOPES, 2008).

Diante destas exigências, observa-se a necessidade do desenvolvimento da análise de dados a partir de questões que possam ser tratadas por meio da coleta, organização e apresentação de dados, utilizando métodos e ferramentas adequados para construção e avaliação de inferências e previsões, destacando-se a importância da problematização por parte dos próprios estudantes.

Lopes (1998) analisou o ensino da Probabilidade e da Estatística no currículo de Matemática do Ensino Fundamental e destacou a importância desses temas para a formação do estudante, pois possibilitam um trabalho pedagógico interdisciplinar, a realização de experimentos e a exploração da ideia de acaso, o que proporciona uma mudança do olhar determinístico característico da Matemática. A autora ressalta a necessidade de proporcionar situações de aprendizagem nas quais sejam desenvolvidos o pensamento estatístico e probabilístico a fim de contribuir para a formação de uma consciência crítica fundamental para o exercício da cidadania.

Entretanto, nota-se a ausência de literatura que aborde a forma como os estudantes desenvolvem a compreensão de conceitos estatísticos fundamentais, principalmente no que se refere ao Ensino Fundamental.

É a partir da realização de observações, registros e representações de dados, que os estudantes serão capazes de ler e interpretar diferentes informações nos

mais variados contextos. Os conceitos estatísticos são essenciais para a resolução de problemas e auxiliarão os alunos na tomada de decisões.

Atividades investigativas que sejam significativas e contextualizadas para os estudantes, através das quais esses possam lidar com dados reais, contribuem para a realização de todas as etapas que consistem uma investigação estatística: definição do problema, coleta, organização, análise e interpretação de dados. Além disso, situações-problema que partam dos interesses dos educandos, em qualquer nível de ensino, facilitam a atribuição de sentido ao tratamento dos dados.

Para que as habilidades mencionadas anteriormente sejam desenvolvidas é necessária a compreensão de vários conceitos estatísticos, entre eles os de amostra e população. A compreensão desses conceitos e das técnicas de seleção de amostragem é fundamental para a realização de inferências estatísticas (BUSSAB e MORETTIN, 2002). Logo, uma melhor explicitação destes conceitos será apresentada a seguir.

Amostragem

Bussab e Morettin (2002) consideram que a utilização de informações de uma amostra para conclusão de algo referente ao todo faz parte do cotidiano das pessoas. Por exemplo, ao provar uma fruta na feira, decidimos se iremos comprar ou não mais daquela fruta; um cozinheiro prova um pouco da sopa para ver se está "no ponto" ou quando, no supermercado, recebemos amostras de um novo produto para conhecê-lo e decidirmos se iremos comprar. Apesar de não percebermos, essas decisões que tomamos no nosso dia a dia são baseadas em procedimentos amostrais.

Stevenson (1981) argumenta que "a finalidade da amostragem é fazer generalizações sobre todo um grupo, sem precisar examinar cada um de seus elementos" (p.158). A utilização da amostragem é primordial quando desejamos realizar inferência estatística, pois formulamos julgamentos sobre um todo analisando apenas uma parte dele, ou seja, uma amostra.

Segundo Dancey e Reidy (2006) ao produzir e analisar dados estatísticos deve-se levar em consideração as diferenças e relações entre amostras e populações. Quando se fala em população, considera-se um grupo distinto de seres

vivos ou até mesmo objetos inanimados. A amostra é uma seleção de elementos da população.

A amostra é a parte do grupo a ser examinada e o grupo todo é chamado de população ou universo (STEVENSON, 1981). Destaca-se que população não é somente grupo de pessoas, mas qualquer objeto que possa ser contado, organizado ou medido. Da mesma forma, Vieira (2012, p.129) define que "população ou universo é o conjunto de unidades sobre o qual desejamos obter informação e amostra é todo subconjunto de unidades retiradas da população".

Alguns exemplos de população seriam: todos os clientes de uma loja, todos os eleitores do Brasil, a frota de carros produzida por mês em uma determinada fábrica, todas as notas de matemática de uma turma do 5º ano, entre outras. O termo população é referente ao conjunto de elementos que possuem as variáveis de interesse. Se um estudo tem como objetivo avaliar o desempenho nas provas de matemática dos estudantes das turmas de 5º ano de uma escola, as notas de uma única turma não seriam mais a população, mas sim uma amostra.

Tomando-se como exemplo uma pesquisa que queira estudar a aceitação de um determinado programa governamental pelos moradores da cidade X, selecione aleatoriamente 200 pessoas. A população será o número de habitantes e a amostra será os 200 selecionados.

Quando a população a ser estudada é relativamente pequena, pode-se realizar um estudo envolvendo todas as unidades da população, é o que chamamos de censo. Entretanto, na maioria das vezes, não é possível fazer um censo. Em um estudo, no qual a população seja consideravelmente grande não podemos analisar todos os elementos da mesma, sendo necessário selecionar uma amostra que seja representativa da população de interesse, ou seja, que possibilite a realização de inferências estatísticas, permitindo-nos a generalização de maneira mais segura possível das conclusões obtidas por meio da amostra para a população (FARIAS, SOARES e CÉSAR, 2003).

Além do tamanho da população a ser estudada outras razões podem ser apontadas para a utilização de amostras: o custo e tempo gasto para a realização de censos; o valor científico das amostras e a necessidade de reposição das unidades examinadas. Em testes de qualidade, por exemplo, quando os elementos normalmente são destruídos, seria inviável realizar um censo, pois toda a população seria extinta.

Vale salientar que, em determinadas situações, o censo é mais vantajoso do que a amostragem, tais como: quando a população é pequena (o tamanho da amostra será grande em relação ao da população) ou quando se dispõe dos dados da população. Cabe ao pesquisador definir de acordo com a finalidade da sua pesquisa qual técnica irá utilizar.

Ao realizar-se um estudo, deve-se conhecer a população analisada a fim de selecionar adequadamente a amostra, evitando que os resultados alcançados pela amostra possuam um viés de seleção, ou seja, que a amostra seja tendenciosa², pois o objetivo primordial da amostragem é a representatividade³ da população.

Voltando ao exemplo citado anteriormente de uma pesquisa sobre a aceitação de um determinado programa governamental, se analisássemos somente os moradores que são beneficiados pelo programa, teríamos uma amostra tendenciosa, não sendo possível fazer inferências adequadas sobre a população. Para que isso não ocorra, Bussab e Morettin (2002) destacam a utilização da amostragem aleatória simples, na qual os elementos são sorteados, tendo a mesma probabilidade de serem selecionados.

Há várias técnicas para obtenção de uma amostra que serão brevemente apresentadas aqui, pois o objetivo principal deste estudo não é aprofundar-se em técnicas de amostragem.

Tipos de amostragem

Deve-se ter em mente que, para fazer um levantamento amostral, é necessário: especificar os objetivos com bastante clareza, a fim de evitar dúvidas posteriores; definir a população a ser estudada; indicar as variáveis a serem observadas; especificar o grau de precisão desejado, pois os levantamentos são sujeitos a incerteza, devido a erros de medida ou devido ao fato de apenas uma parte da população ser examinada; escolher os instrumentos de medida e a forma de abordagem; eleger a unidade amostral e o plano amostral a ser utilizado; se necessário executar a prova experimental (prova-piloto ou pré-teste), pois é quando

² Considera-se amostra tendenciosa ou viesada aquela viciada no que se refere a uma ou mais características que a impedem de ser considerada representativa da população que foi extraída.

³ Uma amostra para ser boa tem que ser representativa, ou seja, deve conter em proporção tudo o que a população possui.

se verificam potenciais erros e, por fim, selecionar a amostra, após ser decidido qual deve ser o respectivo tamanho.

Após conhecer as características da população analisada e estabelecer quais critérios serão utilizados para selecionar os elementos da amostra, escolhe-se qual técnica de amostragem será usada. De acordo com os propósitos da pesquisa tem-se um plano amostral.

Amostragem Probabilística

Uma amostra é dita probabilística se todos os elementos da população tiverem probabilidade conhecida, que não seja zero, de pertencer a amostra.

Nesse tipo de amostra há a necessidade de listagem de todos os elementos da população, pois os elementos da amostra são selecionados através de alguma forma de sorteio aleatório: sorteio manual, tabelas de números aleatórios⁴ ou números aleatórios gerados por computador. A utilização de sorteio é garantia para que não haja intervenção e influência do pesquisador na obtenção da amostra e a possibilidade que todas as unidades da população têm de pertencer a amostra.

Além dessas características, na amostragem probabilística pode-se estimar o erro amostral, estendendo as informações da amostra para a população com um maior controle sobre o risco de tomar decisões erradas e capacidade de generalização.

Serão apresentados sucintamente quatro tipos de amostragem probabilística: amostra aleatória simples (AAS), amostra sistemática, amostra estratificada e amostra por conglomerado.

- Amostragem Aleatória Simples (AAS)

Dentre as várias maneiras de se selecionar uma amostra probabilística ou aleatória de uma população a mais simples é atribuir a todos os elementos da

⁴ As Tabelas de Números Aleatórios são construídas de modo que os dez algarismos (0 a 9) são distribuídos ao acaso nas linhas e colunas. Na tabela de números aleatórios os dez algarismos 0,1,2,...,7,8,9, podem ser lidos isoladamente ou em grupos; podem ser lidos em qualquer ordem, como por colunas, num sentido ou noutro, por linhas, diagonalmente etc. A opção de leitura, porém, deve ser feita, antes de iniciado o processo. Em anexo, encontra-se um exemplo de Tabela de Números Aleatórios.

população a mesma probabilidade de pertencer a amostra, é a chamada amostra aleatória simples (AAS).

Ao utilizar a AAS, a população deve ser conhecida e cada unidade⁵ deve estar identificada, seja por número ou nome. Os elementos que farão parte da amostra serão escolhidos mediante sorteio, podendo ser feito manualmente com urnas ou em computador com programas específicos, ou ainda utilizando uma tabela de números aleatórios. Na seleção por sorteio, o primeiro passo é listar os elementos da população, numerados de acordo com a quantidade para então, serem sorteados.

Como mencionado anteriormente, na AAS cada unidade da população tem a mesma oportunidade de ser inserido na amostra, por este motivo a AAS tende a produzir amostras representativas.

Assim, considerando N como o tamanho da população, a probabilidade de cada elemento ser selecionado será $1/N$. Dessa forma, a amostragem probabilística garante o acaso na escolha, possibilitando a realização de afirmações sobre a população com base nas amostras.

Um exemplo de utilização da AAS seria para a seleção de uma comissão de 7 pessoas que representassem os funcionários de uma empresa X , que possui 200 empregados. Para que não houvesse reclamações, bastaria numerar todos os funcionários realizando-se um sorteio. Dessa forma, todos teriam a mesma chance de compor a comissão.

Entre as vantagens da utilização da AAS, pode-se destacar sua fácil compreensão e a possibilidade dos resultados amostrais serem projetados para toda a população.

Vale destacar que, em alguns casos, nos quais a população seja consideravelmente grande, é difícil construir um processo amostral que conduza a uma amostra aleatória fácil de selecionar, podendo levar a custos elevados ou a uma amostra pouco representativa da população.

⁵ Refere-se à unidade amostral, através da qual são observadas e medidas as características quantitativas e qualitativas da população. A amostra é composta pelo conjunto de unidades amostrais. Cada unidade amostral gera uma única observação da variável de interesse.

- Amostragem Sistemática

É bastante semelhante à AAS, porém as unidades são selecionadas obedecendo a um sistema preestabelecido. Torna-se conveniente quando a população está ordenada segundo algum critério como fichas ou lista telefônica. É possível colher uma amostra utilizando a ordenação natural dos indivíduos, como prontuários, quarteirões de uma cidade e etc. Pode ser utilizada sem restrições se a ordem dos elementos na população não tiver qualquer relacionamento com a variável de interesse.

Nesse tipo de amostragem os membros da população são determinados a partir de intervalos fixos. Por exemplo, numa população de 200 elementos, da qual se deseja retirar uma amostra de 10, escolhe-se cada k -ésimo item da lista, sendo k o tamanho da população dividido pelo tamanho da amostra, neste caso $k = 200/10 = 20$. Sorteia-se um número de 1 a 20, que será o primeiro número da amostra, logo os próximos itens serão retirados de 20 em 20. Suponhamos que o primeiro número sorteado seja 10, na sequência teríamos o 30, 50, 70 e assim sucessivamente.

A inspeção da qualidade da produção de um objeto Y é um exemplo no qual pode ser utilizada a amostra sistemática. Se num dia são fabricados 5.000 objetos e testa-se cerca de 100, tem-se $5.000/100 = 50$, ou seja, será sorteado um objeto entre os 50 primeiros e os demais serão acrescentados de 50 em 50. Se o primeiro for o 5º, o próximo será o 55º, depois o 105º e assim sucessivamente.

A principal vantagem da amostragem sistemática está na grande facilidade na determinação das unidades da amostra. Contudo, há a possibilidade da existência de ciclos de variação da variável de interesse, especialmente se o período desses ciclos coincidir com o período de retirada dos elementos da amostra. Por exemplo, se são atendidos 20 pacientes por dia em uma clínica e o sorteado for o 20º, corre-se o risco de obter uma amostra tendenciosa, pois será escolhido sempre o último elemento da fila.

- Amostragem Estratificada

Esta técnica de amostragem é comum em populações muito heterogêneas, nas quais uma AAS poderia ser pouco representativa.

Antes de selecionar a amostra, divide-se a população em subgrupos mais homogêneos, chamados de estratos, o que demanda um maior conhecimento (mais informações) da população. Após a divisão, procede-se o sorteio em cada estrato como na AAS, depois unem-se as amostras retiradas de cada estrato para formar uma só.

É importante analisar exaustivamente os estratos de modo que todo elemento da população pertença somente a um estrato e nenhum elemento da população deve ser omitido. Os elementos de cada estrato devem ser o mais homogêneo possível, permitindo a heterogeneidade entre os estratos.

Para dividir a população deve-se estabelecer variáveis de estratificação de acordo com a homogeneidade/heterogeneidade da população; custo e grau de relacionamento entre as variáveis. Além disso, as variáveis de estratificação devem estar estreitamente relacionadas com as características de interesse da pesquisa.

Dependendo da pesquisa, são exemplos de variáveis de estratificação: sexo, faixa etária, escolaridade, nível salarial, entre outras características específicas de cada população.

Este tipo de amostragem favorece o aumento da precisão das informações obtidas por meio da amostra sem aumentar os custos e pode ser associada a outra técnica de amostragem probabilística. A amostra estratificada também garante que todos os subgrupos sejam representados na amostra.

- Amostragem por Conglomerado

Ao contrário da amostragem estratificada, na amostragem por conglomerado os subgrupos são heterogêneos, visando a representatividade da população. Conglomerado (ou cluster) é o conjunto de unidades da população. Cada conglomerado é como uma miniatura da população, ou seja, deve conter um maior número de características da população. Portanto, será melhor quanto maior a heterogeneidade da população. Conglomerados podem ser quarteirões, domicílios e etc.

Sorteia-se um número suficiente de conglomerados, cujos elementos constituirão a amostra. Sendo assim, as unidades de amostragem, sobre as quais é feito o sorteio, passam a ser os conglomerados e não mais as unidades individuais da população.

Um exemplo para a utilização de amostra por conglomerado seria para estimar o rendimento médio familiar em uma cidade, pode-se selecionar bairros (conglomerados) com características diferentes e pesquisar a renda de todas as famílias dos bairros escolhidos.

É recomendado o uso da amostragem por conglomerado quando é preciso fazer entrevistas ou observações em grandes áreas geográficas; quando o custo de obtenção dos dados cresce com o aumento da distância entre os elementos; quando não se tem a lista de todos os elementos da população ou a obtenção desta listagem é dispendiosa.

Amostragem Não-probabilística

Na amostragem não-probabilística, a escolha das unidades da população para composição da amostra depende, ao menos em parte, do julgamento do pesquisador. Neste tipo de amostra não há a seleção aleatória dos elementos, pois a escolha ocorre de forma deliberada, o que pode acarretar na não representatividade da amostra, não permitindo a discussão e generalização dos resultados para a população.

A escolha de um método não-probabilístico, normalmente, encontrará desvantagem frente ao método probabilístico. No entanto, em alguns casos, se faz necessário a opção por este método. Entre as razões está a não disponibilidade para sorteio da população ou inviabilidade de identificação de todos os elementos desta. Além disso, às limitações de tempo, recursos financeiros, materiais e pessoais necessários para a realização de uma pesquisa com amostragem probabilística também são fatores que determinam a escolha pela amostragem não-probabilística (MATTAR, 1996).

Quando não houver a intenção de generalizar os dados obtidos na amostra para a população, a amostragem não-probabilística é usada sem problemas, principalmente em pesquisas exploratórias.

São variadas as formas de seleção na amostragem não-probabilística. Apresentaremos de forma concisa a amostra por julgamento, a amostra por conveniência e a amostra por cotas.

- Amostragem por Julgamento

É uma técnica de amostragem mais rápida e barata porque não é necessário construir uma listagem dos itens da população. Contudo, pode ser tendenciosa o que eliminaria a sua representatividade.

É recomendada quando o tamanho da população é pequeno e esta é muito conhecida, pois o pesquisador pode realmente especificar quais são os itens mais representativos. Sendo assim, a escolha dos elementos da amostra fica a critério do conhecimento e julgamento do pesquisador.

Um exemplo seria a implantação de um novo serviço em uma rede de hotéis. Ao invés de testar em todos os hotéis devido ao aumento de gastos, dificuldades com localização, tempo e outros fatores, pode-se escolher dois hotéis baseando-se no conhecimento do administrador a respeito de suas características.

Vale salientar que, como é uma amostragem não probabilística, não permite a avaliação do erro amostral⁶.

- Amostragem por Conveniência

A amostragem por conveniência ou acidental é apropriada e comumente utilizada para geração de ideias principalmente em pesquisas exploratórias.

Amostras por conveniência podem ser facilmente justificadas em um estágio exploratório da pesquisa, como uma base para geração de hipóteses e *insights* e para estudos conclusivos, nos quais o pesquisador aceita os riscos da imprecisão dos resultados do estudo (KINNEAR e TAYLOR, 1979).

É utilizada quando se deseja obter informações de maneira rápida e barata, sendo possível selecionar sujeitos tais como estudantes em sala de aula, mulheres no shopping, vizinhos, entre outros.

O problema de amostras por conveniência é que não há como dizer se esta é representativa, não sendo possível mensurar os erros desta amostragem e nem generalizar os resultados obtidos (KINNEAR e TAYLOR, 1979).

⁶ Será explicado mais adiante.

- Amostragem por cotas

A amostragem por cotas se assemelha à aleatória estratificada da qual se diferencia por não haver sorteio na seleção dos elementos. É a forma mais comum de amostragem não-probabilística.

Segundo Mattar (1996) é um tipo especial de amostra intencional, no qual o pesquisador procura obter uma amostra que seja similar à população sob algum aspecto. São consideradas várias características da população, como sexo, idade, áreas geográficas e alguma medida de nível econômico, pretendendo-se incluir proporções similares de pessoas com as mesmas características (COCHRAN, 1965).

É muito empregada na pesquisa de mercado e de opinião política por ser menos custosa e rápida de usar. A técnica consiste em uma amostra por julgamento realizada em dois estágios. O primeiro estágio é a elaboração de categorias ou cotas de controle de elementos da população e o segundo é a seleção de elementos da amostra com base no julgamento do pesquisador.

Embora a composição amostral seja um espelho da população com respeito às características de controle, não há garantia de que a amostra seja representativa.

Tamanho da Amostra

Embora o objetivo desta pesquisa não seja aprofundar as técnicas de amostragem é importante explicitar, de forma sucinta, alguns conceitos estatísticos que são utilizados para a definição do tamanho de uma amostra, visto que este será um aspecto abordado nesta investigação.

Conhecendo-se a população de interesse e definido o tipo de amostragem que será utilizado têm-se formas diferentes de se calcular o tamanho da amostra. Em geral, baseando-se nos dados seguintes:

- Parâmetro: são características numéricas da população. Ex: média populacional e desvio padrão populacional
- Estatística: são características numéricas de uma amostra. Ex: média amostral e desvio padrão amostral.

- Estimativa: valor acusado por uma estatística que estima o valor de um parâmetro populacional. A média da amostra é utilizada como estimativa da média da população.
- Erro amostral: diferença entre o valor que a estatística pode acusar e o verdadeiro valor do parâmetro que se deseja estimar. Quando tem-se todas as médias da amostra e compara-se com a média da população, a diferença entre essas duas medidas é chamada de Erro de Amostragem ou Margem de Erro. A margem de erro é a distância máxima permitida entre o valor populacional e o valor amostral.

São muitas as maneiras de se definir o tamanho da amostra, utilizando os valores descritos acima. Contudo, aqui não temos como alvo descrever como calcular cada um. Em geral, aumentando-se o tamanho da amostra, aumenta-se a precisão da amostra. Entretanto, é errôneo pensar que o tamanho da amostra deve ser tomado como um percentual do tamanho da população para ser representativa.

Vale salientar que, neste estudo não buscaremos que os estudantes apresentem os cálculos necessários para estabelecer o tamanho da amostra, mas que o identifiquem como aspecto importante para a representatividade da mesma.

Os Conceitos de Amostra e População na Educação Básica

Para Ben-Zvi, Makar, Bakker e Aridor (2011) o conceito de amostra é central para a Estatística, entretanto tem recebido pouca atenção se comparado a conceitos como média, variabilidade e inferências informais⁷.

A necessidade de trabalhar conceitos fundamentais da Estatística é destacada pelos PCN desde as séries iniciais. Segundo o documento:

Com relação à estatística, a finalidade é fazer com que o aluno venha a construir procedimentos para coletar, organizar, comunicar e interpretar dados, utilizando tabelas, gráficos e representações que aparecem frequentemente em seu dia-a-dia (BRASIL, 1997, p.40).

⁷ Utilizamos o termo inferências informais, pois para realização de inferências estatísticas demanda-se conhecimento mais aprofundado das técnicas de amostragem, o que não é objetivo da Educação Básica.

Os parâmetros continuam a enfatizar o ensino de Estatística ao longo da Educação Básica, orientando um aprofundamento dos conteúdos abordados nas séries finais.

Assim, o estudo, nos terceiro e quarto ciclos, dos conteúdos estabelecidos no Tratamento da Informação justifica-se por possibilitar o desenvolvimento de formas particulares de pensamento e raciocínio para resolver determinadas situações-problema, as que envolvem fenômenos aleatórios, nas quais é necessário coletar, organizar e apresentar dados, interpretar amostras, interpretar e comunicar resultados por meio da linguagem estatística (BRASIL, 1998, p. 134).

Os PCN dos 3^o e 4^o ciclos também destacam a "compreensão de termos como frequência, frequência relativa, amostra de uma população para interpretar informações de uma pesquisa" (BRASIL, 1998, p.90).

Segundo Watson (2003) tópicos como amostragem, gráficos, realização de inferências, entre outros são necessários para a construção de habilidades estatísticas mais complexas e por isso, fazem parte do currículo de matemática nas escolas da Austrália ao longo de vários níveis de escolarização.

Os Parâmetros Curriculares Nacionais (1998) orientam a realização de trabalhos com pesquisas, pois dessa forma os alunos terão oportunidade de construir o conceito de amostra, desenvolver e indicar os critérios de escolha da amostra. Para os PCN, ao levantarem questões sobre a seleção da amostra, será possível para os alunos fazer inferências informais a partir da amostra dada.

É a partir da amostra que reunimos algumas informações sobre a população. Logo, segundo Rubin, Bruce e Tenney (1990), o cuidado na seleção da amostra é de suma importância quando utilizamos a estatística inferencial, já que esta busca analisar e interpretar os dados obtidos a partir da Estatística Descritiva, possibilitando conclusões e inferências informais sobre determinada população.

A característica fundamental que determina a força de uma generalização estatística é a representatividade da amostra. Em outras palavras, em que medida as características da população que nos preocupam são refletidas exatamente nas características da amostra (SALMON, 2002 apud INNABI, 2006).

Em suas pesquisas, Innabi (2006) elenca dois critérios que devem ser considerados para que a amostra seja representativa: a amostra é grande o suficiente; a amostra é bastante variada, ou seja, apresenta as características da

população de interesse. Em alguns casos, uma amostra pequena pode apoiar uma generalização forte, em outros, é necessário uma amostra maior. A verdadeira questão é se a amostra é grande o suficiente para capturar, ou representar, a variedade presente na população. Perceber essa relação é uma das dificuldades enfrentadas pelos estudantes.

Em relação à variabilidade, Watson e Kelly (2002), realizaram um estudo sobre a compreensão de estudantes da grade 3⁸, entre 8 e 9 anos, acerca da variabilidade estatística. Realizaram pré e pós-testes associados a 10 aulas que trataram de conteúdos curriculares envolvendo chance e dados com ênfase na variação. Inicialmente os estudantes não levaram em conta a variedade de possibilidades de resultados em diferentes situações envolvendo chance. No pós-teste eles ampliaram suas previsões, baseando-se nos dados apresentados em cada contexto. Embora não tenham atingido um nível de conhecimento estatístico mais complexo, os estudantes apresentaram respostas mais adequadas após as intervenções. A diferença das respostas do pré para o pós-teste mostrou que alunos nessa faixa etária podem alcançar melhoras significativas quando estimulados utilizando variadas situações de amostragem.

Embora, atualmente, seja exigida a capacidade de raciocinar estatisticamente, estudos (RUBIN, BRUCE e TENNY, 1990; ESTEVAM, 2010; GARFIELD, 2003) apontam diversas dificuldades das pessoas quando enfrentam situações sociais que necessitam de raciocínio estatístico, sendo em alguns casos, parciais e baseando seus julgamentos em suas perspectivas pessoais, isto é justificando suas respostas em suas vivências e não nos dados que lhes são apresentados.

Rubin, Bruce e Tenny, 1990; Garfield, 2003; Innabi, 2006 têm apontado em seus estudos dificuldades dos estudantes e de pessoas leigas em estatística em compreender conceitos básicos de amostragem. Além disso, não levam em conta, ao julgar a validade das amostras, fatores relevantes para representatividade da mesma como tamanho e variabilidade.

Entre essas dificuldades, destaca-se a compreensão dos conceitos de amostra e população. Rubin, Bruce, e Tenny (1990) observaram, ao entrevistar alunos do Ensino Médio apresentando-lhes seis questões abertas relacionadas a

⁸ No Brasil equivale a estudantes do 3º e 4º anos.

inferência e amostragem, que uma das maiores limitações destes estudantes é captar os conceitos básicos de amostragem. Suas pesquisas ainda mostraram que os alunos têm modelos inconsistentes da relação entre amostras e populações, pois ao analisarem os resultados obtidos perceberam que, as respostas dos estudantes ora são embasadas em suas intuições sobre representatividade da amostra, ora na variabilidade da mesma, não sendo o tamanho da amostra relacionado a esses conceitos.

Ainda relacionado à opinião dos estudantes sobre amostra e população, Watson (2002) entrevistou um grupo de estudantes das grades 3, 6 e 9⁹ (entre 8 e 15 anos) a fim de compreender o papel do conflito cognitivo na aprendizagem sobre o conceito de amostra em situações nas quais os alunos tinham a oportunidade de expressar suas ideias iniciais sobre amostragem e depois confrontarem com respostas de outros alunos. Para tal, foram apresentadas as questões expostas no Quadro 1:

Quadro 1: Instrumento utilizado por Watson (2002), tradução nossa

1a. Você já ouviu falar antes sobre amostra? O que significa?

1b. "Uma pessoa na TV disse: Numa pesquisa sobre pesos de alunos da 5^o ano os pesquisadores entrevistaram uma amostra de alunos do 5^o ano." O que a palavra amostra significa nessa sentença?

2a. Porque você acha que os pesquisadores usaram uma amostra de alunos do 5^o ano, em vez de estudar todos os alunos do 5^o ano?

2b. Você acha que eles usaram uma amostra de 10 alunos? Porque sim ou porque não? Quantos alunos eles deveriam usar na amostra? Por quê?

2c. Como deveriam ser escolhidos os alunos para essa amostra?

3a. Os pesquisadores foram para 2 escolas: 1 escola no centro da cidade e outra no campo. Cada escola tinha metade de meninas e metade de meninos. Os pesquisadores pegaram uma amostra aleatória para cada escola: 50 crianças eram da cidade e 20 eram do campo.

Uma dessas amostras foi incomum: tinha mais de 80% de meninos. É mais provável que eles vieram de:

- da maior amostra de 50 da cidade ou
- da menor amostra de 20 que era do campo ou
- as duas amostras poderiam ter sido de uma amostra incomum.

Explique sua resposta.

⁹ Equivalente, no Brasil, a estudantes do 3^o, 6^o e 9^o anos.

Watson (2002) percebeu que, após serem apresentadas as respostas de outros alunos, poucos modificaram suas respostas para um nível de compreensão mais adequado. Entretanto, nenhum modificou sua resposta quando lhes foram apresentados pontos de vista de uma natureza menos apropriada.

A autora também observou que os alunos demonstraram mais confiança nas questões que não envolviam cálculos, nas quais eles podiam expressar suas opiniões com exemplos mais descritivos.

Watson (2002) salienta a importância da troca entre alunos de turmas diferentes, a fim de desenvolver conceitos relacionados e necessários para o aprofundamento sobre amostragem. Também é indispensável a intervenção do professor com atividades variadas de amostragem, nas quais possam ser percebidas as relações proporcionais entre amostras e populações.

Além das relações existentes entre amostras e populações, há muitos outros conceitos relacionados a amostra que devem ser trabalhados. Estevam e Fürkotter (2010) destacam o princípio de amostragem aleatória, o qual tenta minimizar os erros amostrais, considerando a variabilidade entre indivíduos e a variabilidade entre grupos. Esses autores apontam a importância e dificuldade da compreensão da natureza da variabilidade em cada contexto de análise, entendendo a relevância da amostragem aleatória.

Outro aspecto ressaltado por Innabi (2006) foi a compreensão dos estudantes sobre a validade da amostra. A autora apresentou aos estudantes um teste contendo a questão "Quantas vezes os estudantes da Universidade dos Emirados Árabes frequentavam a biblioteca da universidade por semana durante o semestre letivo?" e fornecia as opções de amostra descritas no Quadro 2:

Quadro 2: Diferentes amostras apresentadas aos estudantes na pesquisa de Innabi (2006), tradução nossa

TIPO DE AMOSTRA	AMOSTRA
grande/ tendenciosa	600 estudantes entre os que entravam na biblioteca
grande/ não tendenciosa	600 estudantes homens e mulheres entre estudantes de diferentes anos e áreas de estudo
pequena/ tendenciosa	6 estudantes entre os que frequentavam a biblioteca
pequena/ não tendenciosa	6 estudantes de diferentes anos e áreas de estudo
sem informações sobre a amostra	nenhuma informação sobre a amostra.

Os estudantes deveriam selecionar entre as cinco amostras propostas e a forma de seleção das mesmas qual(is) era(eram) a(s) representativa(s), não-representativa(s) ou a(s) que não poderia(m) ser julgada(s). Para análise das justificativas dos estudantes a autora, utilizou a categorização a seguir: 1) Resposta adequada com explanação estatística; 2) Resposta adequada com explanação estatística insuficiente; 3) Respostas baseadas em crenças pessoais; 4) Resposta inadequada.

Ao serem questionados sobre a validade de uma amostra, apresentando-se maneiras diferentes de como esta poderia ter sido selecionada, a maioria dos estudantes do ensino secundário não levou em consideração, ao justificarem suas respostas, o tamanho da amostra e nem se a mesma é tendenciosa ou não quando julgaram a validade e representatividade da mesma.

Observa-se que, como analisado por Rubin, Bruce, e Tenny (1990), que a forte presença de opiniões pessoais para embasamento das respostas sobrepõe-se a utilização de critérios estatísticos nas justificativas dos estudantes.

Watson e Kelly (2002) realizaram uma intervenção (10 aulas) com estudantes entre 8 e 9 anos, na Austrália, nas quais foram trabalhados, entre outros, conceitos relativos à amostragem. Observaram que os alunos foram capazes de dar muitos exemplos de situações nas quais se utiliza amostra (degustação de cozinha, supermercado...), foram capazes de levantar o porquê da utilização de amostras e levantaram questões sobre a seleção de uma amostra representativa para o todo.

Vale destacar que, nos estudos de Watson e Kelly (2002) ocorreram mudanças significativas da compreensão de variação da amostragem, do pré para o pós-teste. Os alunos melhoraram sua definição de amostra e a compreensão de amostra aleatória. Respostas comuns para justificar os métodos de seleção, como "Não tenho certeza, porque eu não acho que isso seja muito bom" deram lugar a respostas mais adequadas, como "Este é bom, porque é justo".

Embora tenham melhorado a ideia de amostra, em ambos os testes, poucos estudantes mostraram uma definição mais estruturada do conceito de amostragem aleatória. As crianças apresentaram dificuldade em entender o raciocínio proporcional necessário para fazer inferências a partir de amostras. Entretanto, eles apresentaram boas intuições baseadas em suas experiências extra-escolares. Essas compreensões foram desenvolvidas durante as discussões em aula e percebeu-se uma melhoria no pós-teste, mostrando a importância de se começar a

trabalhar com amostragem neste nível para contribuir na compreensão de outras habilidades matemáticas mais complexas.

Pfannkuch (2008) percebeu o desenvolvimento das concepções acerca dos conceitos referentes à amostra, durante a aplicação de uma sequência didática. As respostas dadas pelos estudantes ao serem questionados sobre as mudanças ocorridas com amostras de tamanhos diferentes sugerem que os alunos passaram a compreender algumas noções sobre variabilidade da amostragem, ligando a amostra à população, utilizando a linguagem associada à Estatística ou embasando suas respostas a partir das imagens e dados fornecidos. Além disso, a autora ressalta a importância da contextualização no ensino de Estatística. Antes de cada atividade realizada, a professora participante da pesquisa envolvia os alunos em um "cenário de história" com discussões e construção de ideias iniciais sobre população, amostra, tamanho da amostra e representatividade desta.

Semelhantemente, Ben-Zvi, Makar, Bakker e Aridor (2011) perceberam que ao trabalharem com uma sequência de atividades, com estudantes de 11 anos, nas quais havia um crescimento do tamanho das amostras os alunos foram estimulados a pensarem sobre as relações população-amostra. A análise do raciocínio inferencial dos sujeitos envolvidos em suas pesquisas sobre amostragem a partir das questões colocadas mostra um desenvolvimento de pontos de vista do que pode ser concluído a partir de uma pequena amostra (inferências contraditórias e cheias de incerteza) e com amostras maiores, as quais favorecem a colocação de inferências informais.

Gil e Ben-Zvi (2010) observaram que, durante uma sequência didática aplicada com crianças de 12 anos, as primeiras explicações dos estudantes, aos serem questionados sobre a validade das informações baseadas na amostra, refletem uma desarmonia entre a percepção de amostragem aleatória e a confiança nas suas inferências. Entretanto, com o passar das atividades e das discussões percebeu-se o desenvolvimento das concepções sobre amostra e amostragem, à medida que tentam inferir a partir de uma amostra aleatória sobre a população. Ideias de aleatoriedade e de amostragem aleatória foram parcialmente compreendidas e utilizadas por alunos nesta idade, os mesmos foram capazes de considerar as implicações da representatividade da amostra e da variabilidade desta, mas não foram capazes de entender as relações entre elas.

Para Innabi (2006) é necessário que os estudantes sejam capazes de definir os conceitos de amostra e população, mas também que compreendam e

possam raciocinar criticamente quando lhes são apresentadas conclusões acerca de uma população a partir de uma amostra analisada.

Além disso, como aponta Ben-Zvi, Makar, Bakker e Aridor (2011), outros conceitos estatísticos também podem ser trabalhados associados ao conceito de amostra, como média, dispersão, probabilidade, aleatoriedade e interpretação de gráficos.

Assim, nesse estudo estamos interessados em identificar o que estudantes do 5º e 9º ano do Ensino Fundamental compreendem sobre amostragem, considerando as diferentes habilidades relacionadas ao conceito de amostra e população elencadas nos estudos citados anteriormente: seleção, variabilidade, tamanho e representatividade da amostra.

CAPÍTULO 2

MÉTODO

Esta pesquisa teve como objetivo identificar o que estudantes do 5º e 9º ano do Ensino Fundamental compreendem sobre amostragem.

Como objetivos específicos buscou:

- Investigar se os estudantes sabem conceituar amostra;
- Analisar se os estudantes ao opinarem sobre a representatividade de uma amostra consideram o seu tamanho, seleção e variabilidade;
- Analisar se os estudantes compreendem que população é o universo a ser avaliado, envolvendo tanto pessoas como objetos inanimados;
- Verificar se os estudantes realizam inferências informais a partir de uma amostra dada;
- Identificar se há diferenças entre a compreensão de estudantes de diferentes níveis de escolaridade sobre amostragem.

Participantes

Por se tratar de uma pesquisa exploratória, optou-se por utilizar a amostragem por conveniência, pois esta é comumente usada para geração de ideias e confirmação ou negação de hipóteses. Além disso, o tempo destinado a este estudo e o número de informações sobre a população de interesse não permitiriam a realização de uma pesquisa com uma amostra representativa. Logo, os resultados obtidos, os quais serão discutidos adiante, não podem ser generalizados para todos os estudantes da Região Metropolitana do Recife. Pode-se afirmar então, que a amostra refletirá apenas as características das 8 (oito) turmas envolvidas na pesquisa (quatro de cada ano de escolaridade pesquisado), destacando que esta serve de base para a realização de pesquisas futuras envolvendo diversos aspectos da amostragem.

Participaram dessa pesquisa 40 estudantes, sendo 20 do 5º (entre 9 e 10 anos) e 20 do 9º ano (entre 13 e 14 anos) do Ensino Fundamental, de 2 (duas) escolas da rede pública municipal de ensino da Região Metropolitana do Recife, uma localizada na cidade do Recife e a outra em Jaboatão dos Guararapes. Em cada uma das escolas havia 2 (duas) turmas de cada ano de ensino, totalizando 8 turmas estudadas (quatro do 5º ano e quatro do 9º ano).

Embora tenha se optado pela amostra de conveniência, para a seleção dos estudantes utilizou-se uma amostra aleatória simples, a fim de evitar que fossem entrevistados somente os “melhores” alunos. Sendo assim foram sorteados 5 (cinco) sujeitos de cada turma, totalizando 20 (vinte) alunos do 5º ano e 20 (vinte) do 9º ano.

Vale ressaltar que, os estudantes não participaram de nenhuma intervenção antes da entrevista. Também questionou-se às professoras polivalentes do 5º ano e aos professores de Matemática do 9º ano se estes já haviam trabalhado conteúdos ligados à Amostragem. No que diz respeito ao ensino de Estatística, duas docentes do 5º ano afirmaram utilizar atividades com gráficos e tabelas; as outras duas declararam propor situações envolvendo coleta de dados, construção e interpretação de gráficos e tabelas e média aritmética, mas sem relacionar essas atividades com a amostragem; os dois professores de matemática do 9º ano elencaram as medidas de tendência central, gráficos e tabelas como conteúdos trabalhados com os estudantes.

Para melhor compreendermos o raciocínio dos estudantes, foi realizada uma aproximação com o método clínico-piagetiano, o qual permite uma interação entre estudante e pesquisador no intuito de compreender o pensamento do outro a partir de suas argumentações ou explicitações. Assim, o interesse de uma pesquisa norteada por este método e baseada na Teoria Piagetiana é o raciocínio e o processo utilizado pelo indivíduo para chegar a sua conclusão e não a resposta isoladamente. Desse modo, os participantes desse estudo participaram de uma entrevista semi-estruturada individual, sendo estimulados a justificar e explicitar os conhecimentos utilizados em sua resposta.

Procedimentos

Por se tratar de uma pesquisa exploratória foi necessário construir um instrumento para coleta de dados. A partir da análise do referencial teórico e estudo

dos conceitos envolvidos nessa pesquisa foram elaboradas questões que abordassem tais conceitos, adequando-os a faixa etária dos participantes. Para a construção e validação do instrumento de pesquisa foram realizados dois estudos pilotos. O primeiro foi aplicado a 4 (quatro) sujeitos (dois de cada ano) o qual continha 17 (dezesete) questões abertas e as mesmas eram apresentadas apenas oralmente aos estudantes. O segundo foi realizado com um estudante do 5º e outro do 9º ano, contendo 15 (quinze) questões apresentadas de forma oral e impressa. Esses estudos foram realizados para adequar as questões propostas, identificando dificuldades na compreensão das mesmas pelos estudantes e a possibilidade de respostas ambíguas ou difíceis de serem analisadas.

Após a realização dos testes iniciais, definiu-se o instrumento para coleta de dados (Apêndice 1), procurando considerar as dificuldades relacionadas ao conceito de amostra e população elencadas por estudos citados anteriormente: seleção, variabilidade, tamanho e representatividade da amostra.

Os entrevistados foram questionados oralmente, mas também recebiam cartões coloridos com as questões impressas. Para cada questão havia uma cor de cartão, pois percebeu-se durante estudos pilotos, que alguns entrevistados confundiam os exemplos das diferentes questões.

Os cartões também serviram de suporte para facilitar a compreensão das perguntas pelos estudantes, já que estes podiam realizar a leitura quantas vezes desejassem. Como o que estava sendo avaliado não era a habilidade de leitura do estudante, em todas as atividades a pesquisadora lia junto com o aluno a questão para garantir que as questões apresentadas fossem compreendidas pelos mesmos.

As entrevistas foram realizadas em apenas uma sessão e tiveram, em média, 30 minutos, sendo áudio-gravadas a fim de contribuir para uma análise mais profunda da compreensão e resolução das atividades pelos estudantes.

Para apresentação das questões organizamos as mesmas em blocos envolvendo os diferentes aspectos e conceitos ligados à amostragem.

Bloco 1 – Definição de amostra

O primeiro bloco tem como objetivo identificar qual o conhecimento que os estudantes têm sobre a definição do conceito de amostra e sua finalidade. As duas primeiras questões foram adaptadas do estudo de Watson (2002).

1. O que você acha que é amostra?
2. Tu podes me dar um exemplo de amostra?
3. Para saber qual candidato a prefeito do Recife tem mais chance de ganhar as eleições, pesquisadores entrevistaram uma amostra de mil eleitores. O que amostra vai significar nesse caso?
4. Para saber qual o candidato a prefeito do Recife tem mais possibilidade de ganhar a eleição, pesquisadores entrevistaram uma amostra de 1000 eleitores. Por que você acha que eles usaram uma amostra e não todos os eleitores do Recife?

Nos dois primeiros itens buscou-se identificar quais conhecimentos prévios os estudantes possuíam sobre a palavra amostra. Na terceira questão, os entrevistados foram estimulados a perceber um conceito mais formalizado de amostra em um contexto de pesquisa estatística. A quarta questão estava ligada a terceira e tinha como objetivo perceber se os estudantes compreendiam a finalidade de uma amostra.

Bloco 2 – Definição de população

Nessa parte da entrevista busca-se identificar o conceito de população tido pelos alunos em situações nas quais a ideia de população pode ou não ser um grupo de pessoas.

5. Nessa pesquisa sobre os candidatos a prefeito do Recife, qual seria a população analisada?
6. Se fosse realizada uma pesquisa para saber quanto tempo duram os computadores da marca X. Qual seria a população analisada nessa pesquisa?

Na questão 5 tem-se a população como grupo de pessoas, exigindo do entrevistado apenas a noção básica do conceito de população. Todavia, na questão 6, esse conceito é ampliado para população como o universo a ser pesquisado, sendo este animado ou inanimado. No caso, a população seria o grupo de objetos (computadores).

Bloco 3 – Seleção da amostra

Na terceira parte serão observados aspectos relacionados à seleção da amostra de uma forma mais geral, levando em conta critérios necessários para a escolha da mesma. Será observado se os alunos compreendem a importância de que a seleção desta não seja tendenciosa ou manipulada e que represente a população analisada

7. Um pesquisador queria saber a merenda preferida dos alunos das escolas públicas de Recife. Como ele não tinha condições de entrevistar todos os alunos resolveu entrevistar apenas 200 alunos. Como ele poderia escolher esses alunos para ter uma ideia melhor da preferência de todos?

8. Cinco amigos queriam saber aproximadamente quantos livros as pessoas que moram no bairro deles liam por ano. Como o bairro tinha uns 10.000 moradores, não dava para entrevistar todo mundo. Cada um teve uma ideia para saber quem podiam entrevistar. Qual dessas ideias você acha que será melhor para saber o que eles querem? Por quê?

Amigo 1	100 moradores que frequentavam a biblioteca da comunidade.
Amigo 2	100 moradores do bairro.
Amigo 3	10 moradores que frequentavam a biblioteca da comunidade.
Amigo 4	10 moradores do bairro.
Amigo 5	Homens, mulheres, meninos e meninas.

Na questão 7 o estudante será estimulado a sugerir formas para seleção da amostra, sendo observado quais critérios serão estabelecidos pelo mesmo para a escolha dos sujeitos. Essa atividade visa perceber se o entrevistado compreende que para a escolha da amostra é necessário estabelecer critérios que levem em conta os propósitos de cada pesquisa e busquem alcançar os objetivos da mesma.

A questão 8 foi adaptada de uma atividade realizada por Innabi (2006), nela serão abordados os conceitos de seleção, tamanho e representatividade da amostra. É importante destacar que as respostas dadas para esta questão servirão para analisar a criticidade dos estudantes em relação à seleção como um todo e se estes relacionam o tamanho da amostra como fator importante para a seleção da mesma.

Nesta questão foram apresentadas amostras de diferentes tamanhos e critérios de seleção, visando observar a capacidade de julgamento dos estudantes

ao serem confrontados com diversas formas de seleção da amostra para uma mesma pesquisa.

Em relação ao tamanho, o estudante deverá levar em conta se o quantitativo apresentado para a amostra é considerado o ideal para o que a suposta pesquisa objetiva. Dessa forma, supõe-se que ele perceberá que o tamanho da amostra é um aspecto que deve ser levado em conta para a seleção da amostra.

Ao escolher qual amostra será mais representativa, o estudante também deverá observar a tendenciosidade na seleção da mesma. No caso desse exemplo, se o objetivo da pesquisa é saber a média de livros lidos por ano em um bairro X, entrevistar apenas frequentadores da biblioteca não garantiria que o resultado da pesquisa represente toda a população, pois a amostra teria um viés de seleção.

Bloco 4 – Aleatoriedade

Embora o conceito de aleatoriedade esteja diretamente relacionado ao processo de seleção de uma amostra, visando uma melhor análise, optou-se por separar as questões envolvendo tal conceito em um bloco específico.

9. Para definir a ordem dos alunos na fila para a merenda, a professora colocou o nome dos alunos em um saquinho e foi retirando aleatoriamente. O que significa aleatório para você?

10. Eu quero saber o que as pessoas acham do carnaval do Recife. Para selecionar as pessoas que participarão da pesquisa eu vou usar uma amostra aleatória. O que significa amostra aleatória?

11. Dê um exemplo de outra situação na qual podemos usar uma amostra aleatória.

Na pergunta 9 abordou-se o conceito de aleatoriedade, visando reconhecer se os estudantes compreendem ou têm alguma ideia inicial acerca deste a partir de uma situação hipotética.

As questões 10 e 11 envolveram o conceito de amostra aleatória. Para a resolução destas esperava-se que os estudantes relacionassem as definições apresentadas anteriormente nas perguntas sobre a definição de amostra e aleatoriedade, expondo o significado de amostra aleatória e aplicando-o em uma determinada situação.

Bloco 5 – Realização de inferências informais

O bloco 5 foi composto de apenas um item, através do qual o aluno deverá retirar informações de uma amostra dada, realizando inferências informais acerca dos dados coletados.

Primeiramente foi apresentado o banco de dados juntamente com a pesquisa hipotética.

12. Em uma escola com 150 alunos foi realizada uma pesquisa sobre o uso do celular. Foi sorteada uma amostra de 20 alunos e realizou-se uma pesquisa, obtendo os dados abaixo.

Aluno	Sexo	Pra que usa mais	Pra que usa menos
Aluno 1	F	Fotos	Ligações
Aluno 2	F	Músicas	Ligações
Aluno 3	F	Fotos	Ligações
Aluno 4	F	Músicas	Mensagem
Aluno 5	F	Músicas	Ligações
Aluno 6	F	Fotos	Ligações
Aluno 7	F	Mensagem	Fotos
Aluno 8	F	Mensagem	Mensagem
Aluno 9	F	Fotos	Ligações
Aluno 10	F	Fotos	Mensagem
Aluno 11	M	Músicas	Ligações
Aluno 12	M	Mensagem	Fotos
Aluno 13	M	Mensagem	Músicas
Aluno 14	M	Mensagem	Fotos
Aluno 15	M	Músicas	Ligações
Aluno 16	M	Mensagem	Fotos
Aluno 17	M	Fotos	Mensagem
Aluno 18	M	Músicas	Fotos
Aluno 19	M	Mensagem	Músicas
Aluno 20	M	Mensagem	Fotos

Após a exposição da tabela foram feitas perguntas sobre a organização do banco de dados a fim de perceber como o sujeito utiliza a tabela e verificar se este consegue elaborar conclusões a partir dos dados apresentados.

- a. Que tipo de informação a tabela traz?
- b. Que conclusões você pode tirar dessa tabela?
- c. O que você pode dizer sobre o uso do celular nessa escola?

É importante ressaltar que a questão “a” foi realizada para verificarmos se os estudantes compreendiam uma representação em tabela (banco de dados). Foram realizadas explicações acerca da organização das informações presentes na mesma para que a incompreensão da tabela não fosse impedimento para a realização de inferências informais, que é o objetivo desta parte da entrevista.

As questões “b” e “c” visaram analisar a capacidade do estudante em extrair o máximo de conclusões a respeito de uma determinada população a partir de uma amostra dada, no caso a população dos 150 estudantes da escola e a amostra de 20 destes. Após a breve familiarização com o banco de dados, foi observado se o estudante entendia o conjunto de dados como uma amostra que representava uma determinada população, utilizando as informações apresentadas para generalização.

Bloco 6 – Representatividade da amostra

A representatividade da amostra já foi analisada de forma implícita nas questões 7 e 8. Contudo, será observada de forma objetiva na última indagação da entrevista, que é uma continuidade da questão 12.

13. E se, nesta pesquisa sobre o uso do celular na escola, forem sorteados outros 20 alunos de uma mesma turma? Iria ser melhor ou pior para representar a escola? Por quê?

Esta questão teve como objetivo estimular o estudante a refletir sobre a importância da amostra ser variada, contendo características de toda a população para que seja representativa. O entrevistado deveria comparar as duas opções de amostra e indicar qual seria mais adequada, justificando sua resposta.

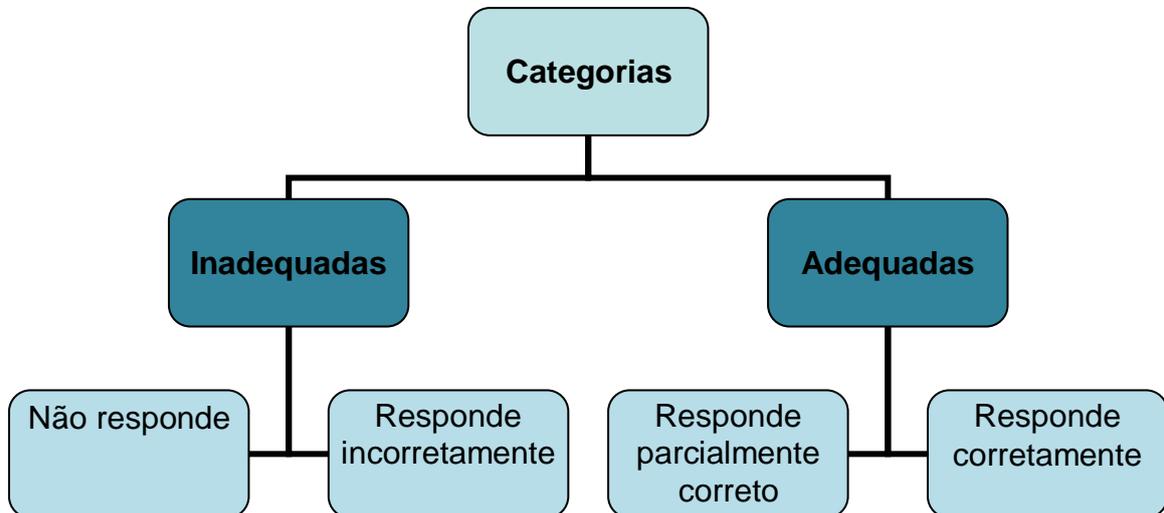
Categorização e análise dos dados

Os dados coletados foram categorizados e analisados com o auxílio do software SPSS (Statistical Package for the Social Sciences).

A partir dos resultados encontrados foram elaboradas, primeiramente, duas categorias (níveis) para as respostas apresentadas. Posteriormente, essas foram divididas em quatro subcategorias, as quais são apresentadas no esquema abaixo

(Figura 1) e serão detalhadas com exemplos para cada variável analisada no capítulo seguinte.

Figura 1: Categorias e subcategorias para classificação das respostas dos estudantes



Optou-se por agrupar inicialmente as respostas em duas categorias para a realização do Teste Qui-quadrado, a fim de identificar se a diferença entre as respostas do 5º e 9º ano para cada questão foi estatisticamente significativa. Essa junção ocorreu devido o tamanho da amostra ser relativamente pequeno, o que poderia interferir nos resultados do Teste, pois para verificação do mesmo não pode haver categorias sem nenhuma ocorrência ou com o quantitativo inferior a cinco.

As quatro subcategorias foram elaboradas com o intuito de enriquecer a análise, observando e valorizando a diversidade das respostas apresentadas pelos estudantes.

CAPÍTULO 3

DESCRIÇÃO E ANÁLISE DOS RESULTADOS

No presente capítulo, será realizada a descrição dos resultados das treze questões utilizadas como instrumento de pesquisa, assim como a discussão dos mesmos. Algumas questões abordam mais de um conceito relacionado à amostragem, os quais são analisados individualmente em cada questão específica. Os resultados encontrados nas questões que envolvem os mesmos conceitos são confrontados com o que já argumentaram estudos anteriores a respeito de tais aspectos.

A opção por essa organização na análise ocorreu principalmente por não haver estudos que englobem todos os conceitos relacionados à amostragem estudados nessa pesquisa. Logo, ela será norteadada pelos objetivos propostos inicialmente, os quais retomamos aqui:

- Identificar o que estudantes do 5º e 9º ano do Ensino Fundamental compreendem sobre amostragem;
- Investigar se os estudantes sabem conceituar amostra;
- Analisar se os estudantes ao opinarem sobre a representatividade de uma amostra consideram o seu tamanho, seleção e variabilidade;
- Analisar se os estudantes compreendem que população é o universo a ser avaliado, envolvendo tanto pessoas como objetos inanimados;
- Verificar se os estudantes realizam inferências informais a partir de uma amostra dada;
- Identificar se há diferenças entre a compreensão de estudantes de diferentes níveis de escolaridade sobre amostragem.

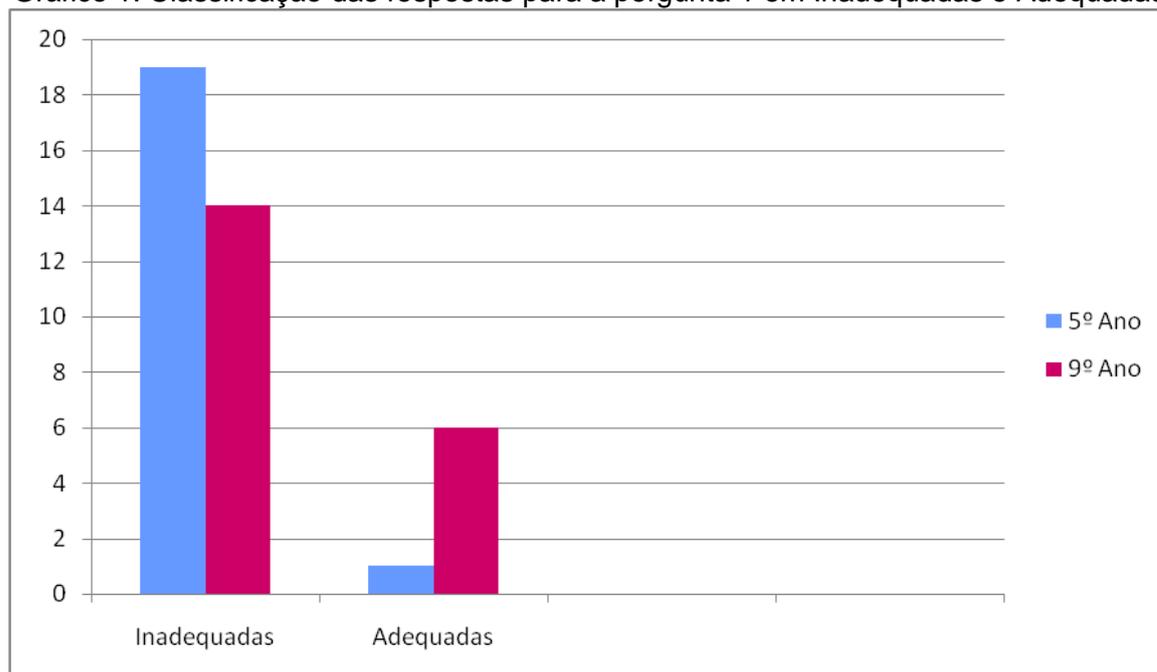
Bloco 1 – Definição de amostra

Inicialmente analisamos as questões referentes ao conceito de amostra, ou seja, como definem e quais exemplos apresentam. Na primeira questão foi solicitado que os alunos respondessem:

O que você acha que é amostra?

O gráfico 1 apresenta a distribuição das respostas segundo a adequação das mesmas ao conceito abordado.

Gráfico 1: Classificação das respostas para a pergunta 1 em Inadequadas e Adequadas



Observa-se um maior quantitativo de respostas adequadas entre os estudantes do 9º ano. Ao realizar o Teste Qui-quadrado percebe-se que esta diferença é significativa ($X^2 = 4,329$, gl 1, $p < 0,037$).

Após esta divisão mais geral, as respostas foram classificadas em 4 (quatro) subcategorias a fim de analisarmos de forma mais detalhada as respostas dos participantes.

Tabela 1 – Frequência por ano de escolaridade da pergunta 1.

Categorias	Ano de ensino		Total
	5º Ano	9º Ano	
Não responde	3	3	6
Responde incorretamente	16	11	27
Responde parcialmente correto	0	4	4
Responde corretamente	1	2	3

Observa-se, na Tabela 1, que 3 alunos de cada ano de escolaridade não responderam a questão. A maioria dos estudantes entrevistados respondeu incorretamente, tanto no 5º como no 9º ano.

Para esta pergunta a resposta considerada ideal seria a que se aproximasse o máximo possível da definição estatística para amostra: “*todo subconjunto de unidades retiradas de uma população para obter a informação desejada*” (VIEIRA, 2012, p.129).

Entre os 27 (vinte e sete) alunos que definiram incorretamente, foi observado que, muitas vezes, a palavra “amostra” foi associada ao verbo mostrar. Essa relação evidencia a influencia de características linguísticas regionais na construção de conceitos, uma vez que essa palavra “amostrar” é muitas vezes utilizada como sinônimo de mostrar por pernambucanos, como nos exemplos a seguir:

S01 – Amostrando alguma coisa, um objeto.

S02 – Quando a pessoa pode mostrar alguma coisa.

Quatro alunos do 9º ano foram classificados como tendo definido o conceito de amostra de forma parcialmente correta, uma vez que apresentaram exemplos de amostra utilizados no cotidiano e não a definição do conceito.

S28 - Eu acho que é feito amostra grátis.

S32 - É uma pequena... Um pequeno produto que a pessoa prova pra vê se gosta.

S34 - Rapaz, só o que eu sei é quando eu tiro uma amostra da gasolina com meu pai pra saber se é boa.

P - E o que vai ser essa amostra de gasolina?

S34 - Tá lá o tanque cheio. Ai a gente tira um pouquinho e faz o teste pra saber se presta ou não.

S37 - Uma amostra de sangue que as pessoas tiram.

Apenas três estudantes (um aluno do 5º ano e dois do 9º ano) apresentaram uma definição de amostra classificada por nós como correta uma vez que definem amostra como parte de um todo.

S31 – *É uma pequena quantidade de... é algo menor que o real, que o todo.*

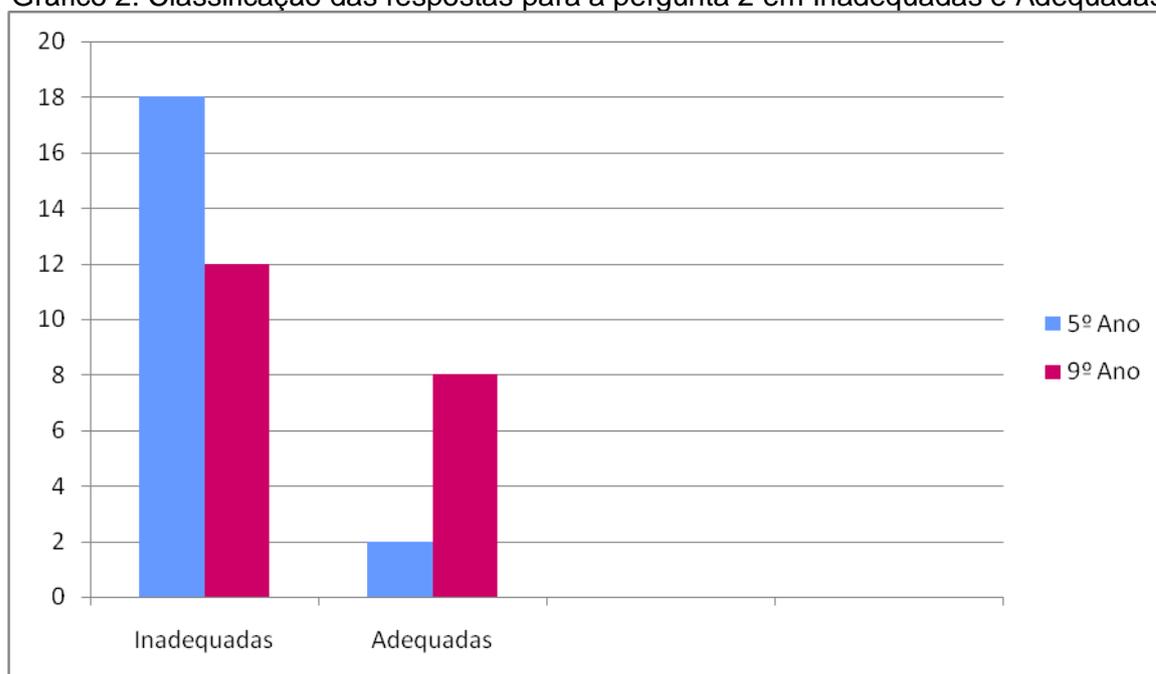
S40 – *Eu acho que é uma pequena parte de tudo.*

Em seguida foi solicitado que os alunos apresentassem um exemplo de amostra.

Tu podes me dar um exemplo de amostra?

De modo geral, percebe-se no Gráfico 2, que o quantitativo de respostas adequadas é maior no 9º ano.

Gráfico 2: Classificação das respostas para a pergunta 2 em Inadequadas e Adequadas



O Qui-quadrado revela que houve diferença estatística entre os níveis de ensino também para esta questão (X^2 4,800, gl 1, $p < 0,028$).

Os resultados para a subdivisão das respostas adequadas e inadequadas estão expostos abaixo na Tabela 2.

Tabela 2 – Frequência por ano de escolaridade da pergunta 2.

Categorias	Ano de ensino		Total
	5º Ano	9º Ano	
Não responde	4	4	8
Responde incorretamente	14	8	22
Responde parcialmente correto	0	0	0
Responde corretamente	2	8	10

Nota-se que, mais uma vez, a maioria dos entrevistados respondeu incorretamente ou não citou nenhum exemplo.

Ao propor essa questão, esperava-se que os estudantes, mesmo sem conhecer o conceito de amostra, partissem de experiências comuns em seu cotidiano para encontrar exemplos do uso de amostra, como o caso das amostras grátis. Contudo, mesmo o quantitativo de respostas corretas, sendo um pouco maior que na questão anterior, observa-se ainda um baixo desempenho dos participantes.

A maior parte dos exemplos dados não apresentou nenhuma relação com a estatística. Os mesmos estabeleciam ligação com o conceito de amostra definido por eles na questão anterior. Logo, grande parte das respostas relacionou amostra ao verbo mostrar, entre outros significados mencionados pelos estudantes.

S01 - Amostra de outros jeitos.

P - Amostrar de outros jeitos? Como assim amostrar de outros jeitos?

S01 - Eu quero lhe amostrar meu caderno, aí eu amostro à senhora. Amostrar!

S14 - Eu pegar esse lápis e lhe amostrar.

S22 - É aquela pessoa que coloca um penico na cabeça.

S39 - Meu teclado.

P - Como assim, teu teclado?

S39- Porque eu não sei tocar, aí ele fica lá de enfeite, só de amostra.

Para esta variável não houve respostas parcialmente corretas, o que era esperado por se tratar de exemplos e não do conceito propriamente dito. Foram consideradas respostas corretas as que utilizaram exemplos para amostra tanto do cotidiano, como as amostras grátis, quanto exemplos mais elaborados, como amostra de sangue ou de outra substância para análise. Dois estudantes do 5º ano e oito do 9º ano conseguiram exemplificar corretamente.

S10 - *Por exemplo, um suco, uma amostra de suco.*

S32 - *Aqueles vidrinhos de perfume ou aqueles papezinhos com o perfume que quando a gente passa e entregam pra gente.*

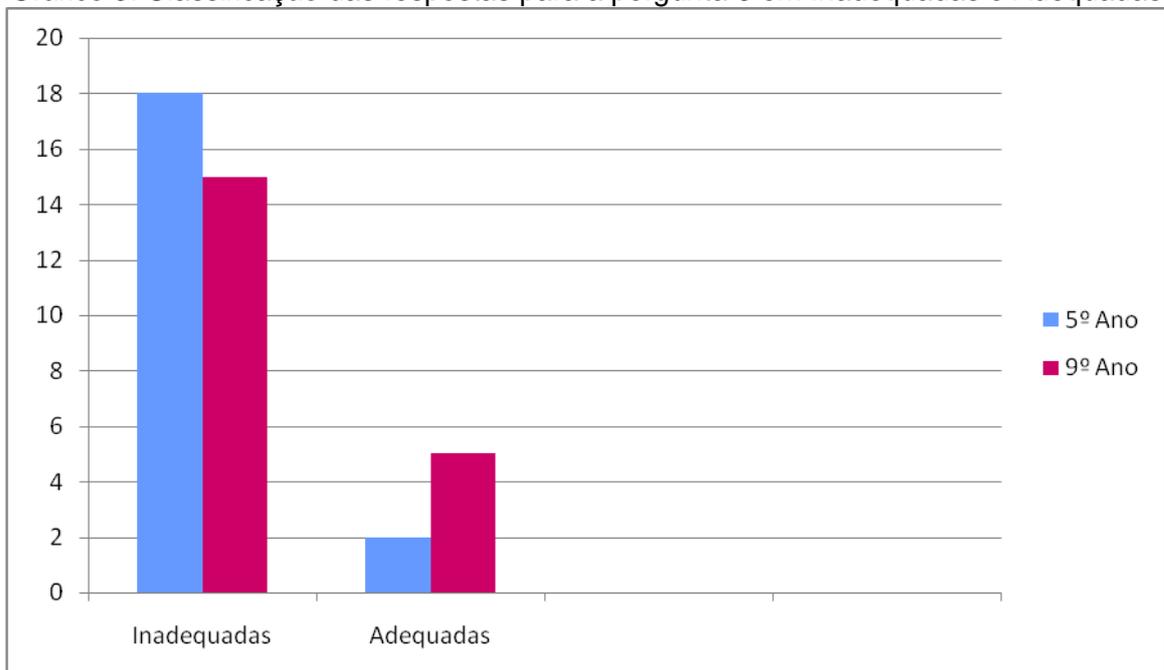
S37 - *De saliva pra saber o DNA.*

A terceira análise também refere-se à definição do conceito de amostra, mas a partir de um contexto situacional. Foi solicitado que os alunos respondessem a questão:

Para saber qual candidato a prefeito do Recife tem mais chance de ganhar as eleições, pesquisadores entrevistaram uma amostra de mil eleitores. O que amostra vai significar nesse caso?

O Gráfico 3 expõe os resultados e a classificação nas categorias gerais (Inadequadas e Adequadas) para esta questão:

Gráfico 3: Classificação das respostas para a pergunta 3 em Inadequadas e Adequadas



Embora o 9º ano tenha apresentado um maior número de respostas adequadas, ao contrário das questões anteriores, não houve diferença estatística entre os anos (X^2 1,558, gl 1, $p < 0,212$).

A divisão dos resultados por subcategorias é exposta na tabela 3.

Tabela 3: Frequência por ano de escolaridade da pergunta 3.

Categorias	Ano de ensino		Total
	5º Ano	9º Ano	
Não responde	3	3	6
Responde incorretamente	15	12	27
Responde parcialmente correto	1	1	2
Responde corretamente	1	4	5

Percebe-se que, mesmo a questão apresentando um exemplo de pesquisa, seis participantes não definiram amostra. Observa-se também que, tanto ao questionar diretamente sobre o conceito de amostra ou utilizando-se de exemplos, como o caso da pesquisa eleitoral, a maior parte dos estudantes responde incorretamente, sem apresentar nenhuma relação com o significado real da palavra amostra.

Esperava-se que ao utilizar exemplos de utilização de amostras os estudantes apresentassem um desempenho melhor em suas respostas, o que não ocorre visto que mais da metade dos sujeitos entrevistados responde incorretamente ou não responde.

Algumas respostas sugerem que os estudantes levaram em conta apenas o contexto da pesquisa citada e não o real conceito de amostra. Essas respostas foram classificadas como incorretas, como os exemplos a seguir:

S03 - Voto.

P - O voto? Então essa amostra de mil eleitores vai significar voto, como assim?

S03 - Em quem ia votar. Pronto, no Conselho Tutelar, votar. A pessoa tem que votar.

P - Então amostra vai ser a mesma coisa que votar?

S03 - É.

S30 - Pra eles “vê” o que os candidatos vão fazer.

P - Essa amostra de mil eleitores vai ser pra eles verem o que os candidatos propõem?

S30 - É.

S38 - Vai ser o que eles pegaram dos eleitores pra poder mostrar a opinião deles.

P - Então, amostra vai ser a mesma coisa que mostrar?

S38 - É.

Observa-se, na Tabela 3, que um estudante do 5º e um do 9º ano foram classificados como responde parcialmente correto uma vez que se referem a

amostra como um “bocado de pessoas” e não a todas as pessoas, demonstrando ter uma ideia de amostra como parte do todo.

S02 - Vai ser uma amostra de um bocado de pessoas.

P - Vai ter que pegar uma amostra de um bocado de pessoas?

S02 - É.

P - No caso, são quantas pessoas essa amostra?

S02 - Mil.

S36 - Vai ser um bocado de gente.

Uma pequena parte da amostra pesquisada (um aluno do 5º ano e quatro do 9º) conseguiu conceituar corretamente amostra, relacionando a definição apresentada com o contexto do exemplo, ou seja, definindo a amostra como sendo uma parte de todos os eleitores do Recife.

S06 - [...] Uma parte de todos (eleitores).

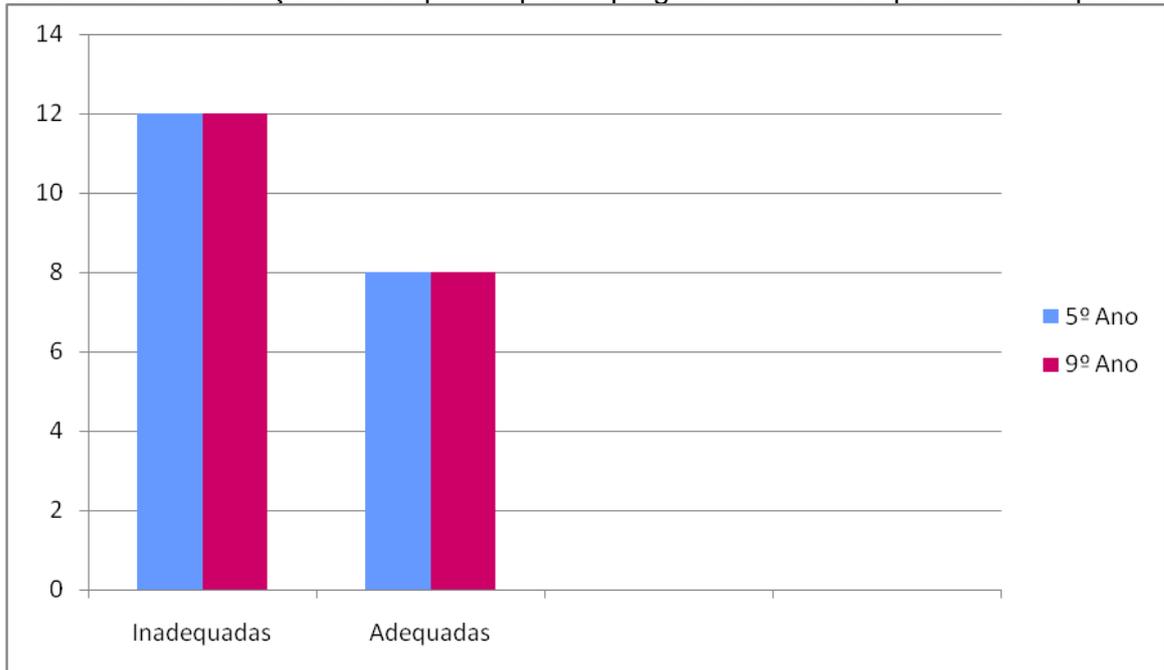
S34 - Eu acho que essa tá significando uma parte dos eleitores pra saber o que eles preferem.

A quarta variável analisada investigava a finalidade da amostra, o porquê utilizar-se da amostragem. Foi requerido que os alunos respondessem a questão:

Para saber qual o candidato a prefeito do Recife tem mais possibilidade de ganhar a eleição, pesquisadores entrevistaram uma amostra de 1000 eleitores. Por que você acha que eles usaram uma amostra e não todos os eleitores do Recife?

A distribuição entre respostas adequadas e inadequadas dos estudantes do 5º e 9º ano é apresentada no Gráfico 4. Para esta questão, não houve diferença estatística entre os níveis de escolaridade (X^2 0,000, gl 1, $p < 1,000$). Vale destacar que, ao contrário do que era esperado, o desempenho dos participantes em relação à adequação das respostas foi o mesmo para esta questão.

Gráfico 4: Classificação das respostas para a pergunta 4 em Inadequadas e Adequadas



Ao classificar nas subcategorias, pode-se observar na Tabela 4, como essas respostas foram apresentadas pelos estudantes.

Tabela 4: Frequência por ano de escolaridade da pergunta 4.

Categorias	Ano de ensino		Total
	5º Ano	9º Ano	
Não responde	8	1	9
Responde incorretamente	4	11	15
Responde parcialmente correto	8	6	14
Responde corretamente	0	2	2

Novamente mais da metade dos entrevistados não respondeu ao questionamento ou respondeu incorretamente. Abaixo estão apresentados exemplos de respostas incorretas:

S08 - Pra saber se eles gostavam dos votos.

S12 - Porque as mil sabiam mais do que os outros.

S25 - Porque ficaram fazendo promessas.

Vale destacar que, os dados da Tabela 4 mostram que o número de respostas parcialmente corretas aumentou em relação às variáveis anteriormente analisadas tanto para o 5º ano (8 alunos) como para o 9º ano (6 alunos). Esses

consideraram o quantitativo de pessoas na população da pesquisa (os eleitores do Recife) e a praticidade de utilizar uma amostra de mil. Esse aspecto é uma das vantagens e finalidade da amostragem.

S04 - Pra não ficar muitas pessoas.

S06 - Se não ia ficar ruim.

P - Ia ficar ruim por quê?

S06 - Porque ia ter muita gente.

S13 - Porque mil já é muito. Imagine se fosse Pernambuco todinho. Num ia dar não. Ia demorar muito.

A resposta mais adequada deveria elencar as vantagens para utilização de uma amostra. Por essa ser uma resposta bastante complexa para o nível de escolaridade aqui estudado, consideraram-se como corretas as respostas dos alunos que estabeleceram uma relação entre a praticidade do quantitativo usado na amostragem com a população analisada e à possibilidade dessa amostra representar o todo.

Dessa forma, apenas dois alunos do 9º ano responderam corretamente com algum aprofundamento estatístico.

S34 - Porque amostra é mais prático.

P - Como assim?

S34 - Se tiver 2 milhões de habitantes no Recife e ele pegou mil já dá pra ter pelo menos uma ideia.

S40 - Porque não era o dia da eleição, porque pra pesquisa só uma parte já dava pra ter uma ideia.

Percebe-se que, os resultados alcançados nessa pesquisa confirmam dificuldades de estudantes em compreender o conceito de amostra apontadas em outros estudos, como os de Rubin, Bruce e Tenny (1990); Estevam (2010) e Garfield (2003). Nesses estudos, observou-se que as definições dadas pelos estudantes são em grande parte baseadas em suas perspectivas pessoais, sem relação com a estatística, o que também pode ser percebido em nossa pesquisa com estudantes do 5º e 9º ano.

Embora a maior parte dos entrevistados não tenha conseguido conceituar e exemplificar amostra de maneira correta, ao expor um exemplo de pesquisa com

uso de amostras e solicitar que explicassem o porquê de utilizá-las (finalidade da amostra), a Tabela 4 mostra que quase metade dos estudantes (16 sujeitos) responderam de forma parcialmente correta. Vale destacar que, 8 (oito) desses estudantes eram alunos do 5º ano, os quais justificaram suas respostas baseando-se na praticidade de se utilizar amostras. Esses resultados indicam que alunos nessa faixa etária já são capazes de compreender o propósito para uso de amostras.

Essas dificuldades apresentadas pelos estudantes apontam a necessidade de outras pesquisas a fim de identificar o porquê dessas limitações. Pode-se hipotetizar que seja apenas uma lacuna do processo didático, pois se eles não foram estimulados a refletir sobre o conceito de amostra, ou seja, se não há o conhecimento estatístico, só resta aos mesmos raciocinar dessa forma.

Como mostra o estudo de Watson (2002), é fundamental colocar os alunos para discutir sobre o conceito. Embora em nosso estudo os entrevistados não tenham sido confrontados com respostas de outros estudantes, percebeu-se que alguns estudantes modificaram sua percepção de amostra da primeira questão, a qual perguntava o que seria amostra de forma direta, para a segunda questão, que contextualizava o uso de amostras. Ao serem questionados de forma diferente sobre o que seria amostra, alguns estudantes permaneceram com suas ideias iniciais, enquanto outros apresentaram respostas mais adequadas, como exemplificado abaixo.

P - S02, o que é que tu acha que é amostra? O que significa essa palavra amostra?

S02 - Quando a pessoa pode mostrar alguma coisa.

P - Quando a pessoa pode mostrar alguma coisa?

S02 - É. Amostra.

P - Mostrar uma coisa a outra pessoa. É a mesma coisa que amostra?

S02 - É.

P - Se eu te der um exemplo, de uma pesquisa que foi feita pra saber qual o candidato a prefeito do Recife ou daqui de Jaboatão. Esse ano não é eleição? Eu quero saber qual o candidato a prefeito de Jaboatão tem mais possibilidade de ganhar a eleição desse ano. Aí os pesquisadores entrevistaram uma amostra de mil eleitores. Nesse caso o que é que a palavra amostra vai significar? Vai ser a mesma coisa que mostrar algo pra alguém?

S02 - Não.

P - Não? O que é que vai ser amostra nesse caso?

S02 - Vai ser uma amostra de um bocado de pessoas.

P - Vai ter que pegar uma amostra de um bocado de pessoas?

S02 - É.

P - No caso, são quantas pessoas essa amostra?

S02 - Mil.

P - S29, o que tu achas que é amostra?

S29 - Apresentar alguma coisa, mostrar alguma coisa.

P - Se eu der um exemplo de uma pesquisa que foi feita pra saber qual candidato a prefeito do Recife está com mais chance de ganhar a eleição. Feito essas que de vez em quando aparecem na televisão. Pra saber isso, os pesquisadores entrevistaram uma amostra de mil pessoas. O que a palavra amostra vai significar nesse caso?

S29 - Eu acho que é uma parte de quem vota pra eles saberem em quem eles iam votar.

Essa possibilidade de passar a mostrar que compreendem a partir somente de questões nos leva a reafirmar se os alunos apresentam tantas dificuldades porque não foram expostos a reflexões sobre o conceito ou por fatores de natureza cognitiva.

Em nossa pesquisa, vale destacar que, mesmo os estudantes do 9º ano tendo demonstrado uma maior facilidade em exemplificar e estruturar uma definição mais adequada para o conceito de amostra, a ocorrência de respostas apropriadas entre os alunos do 5º ano ratifica a hipótese de que há possibilidade de aprendizagem desses conceitos por alunos nessa faixa etária.

Watson e Kelly (2002) nos ajudam a colocar em evidencia que os alunos podem compreender o que seja amostra se forem expostos a atividades que os levem a aprendizagem. No estudo dessas autoras, alunos entre 8 e 9 anos de idade depois de uma pequena intervenção foram capazes de dar exemplos de situações nas quais se utiliza amostra, levantar o porquê da utilização de amostras, levantar questões sobre a seleção de uma amostra representativa para o todo e melhoraram suas definições de amostra. Entretanto, apresentaram, ainda, dificuldades em compreender ideias acerca do tamanho da amostra e sobre amostra aleatória. Estas compreensões foram desenvolvidas durante as discussões em aula mostrando a importância e possibilidade de se começar a trabalhar com amostragem neste nível de escolaridade.

Da mesma forma, Pfannkuch (2008) também percebeu o desenvolvimento das concepções acerca dos conceitos referentes à amostra, durante a aplicação de uma sequência didática.

Ben-Zvi, Makar, Bakker e Aridor (2011) também observaram que ao trabalharem com crianças de 11 anos, uma sequência de atividades nas quais havia um crescimento do tamanho das amostras estimulou os alunos a pensarem sobre as relações população-amostra, levando-os a pensar o que pode ser concluído a partir de uma amostra.

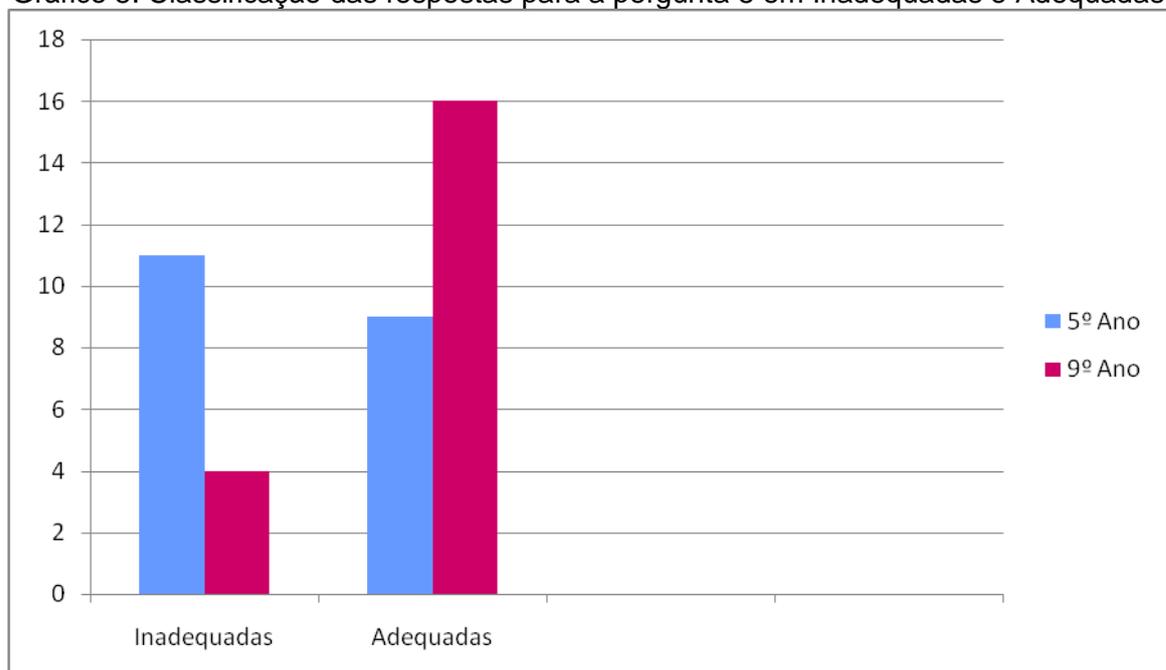
Essa construção de conhecimento ratifica a importância de se propor atividades variadas envolvendo situações de uso da amostragem e a importância da intervenção do professor no processo de aprendizagem.

Bloco 2 – Definição de população

Uma vez analisada as questões referente à amostra, passamos a analisar as questões referente à população. A quinta análise refere-se à definição do conceito de população. Nesse primeiro caso, a população é um grupo de pessoas. Os resultados para esta questão estão expostos no Gráfico 5.

Nessa pesquisa sobre os candidatos a prefeito do Recife, qual seria a população analisada?

Gráfico 5: Classificação das respostas para a pergunta 5 em Inadequadas e Adequadas



Nota-se que, nesta questão houve um aumento na adequação das respostas, sendo a diferença entre os anos de escolaridade estudados significativa (X^2 5,227, gl 1, $p < 0,022$).

A Tabela 5 traz os dados divididos nas subcategorias, permitindo uma análise mais detalhada das respostas.

Tabela 5: Frequência por ano de escolaridade da pergunta 5.

Categorias	Ano de ensino	5º Ano	9º Ano	Total
Não define		1	1	2
Define incorretamente		10	3	13
Define parcialmente correto		6	9	15
Define corretamente		3	7	10

Observa-se que, nessa variável o desempenho apresentado foi melhor que nas anteriores. Visto que, mais da metade dos estudantes responderam de forma correta ou parcialmente correta. Supõe-se que esse resultado ocorreu devido ao conceito de população como conjunto de pessoas ou habitantes de uma determinada região ser trabalhado em outras disciplinas, como geografia, tanto no início da escolarização como nos anos finais.

Estatisticamente, entende-se por população ou universo como sendo “o conjunto de unidades sobre o qual desejamos obter informação” (VIEIRA, 2012, p.129). A resposta apropriada para o exemplo citado definiria a população da pesquisa como sendo todos os eleitores da cidade em questão, no caso Recife.

Entre as respostas incorretas foi comum haver confusão entre os termos eleitor e candidato, mesmo após explicação e diferenciação dos mesmos. Além disso, respostas baseadas em critérios banais, sem afinidade com o contexto da pesquisa também foram encontradas.

S08 - Os mil eleitores

P - Os mil eleitores seriam a amostra e a população?

S08 - É.

S09 - Humberto Costa...(nome de um dos candidatos a prefeito)

S18 - Eles todinhos.

P - Eles quem?

S18 - Os candidatos.

S20 - Os candidatos.

S33 - Quem trabalha mais.

Percebe-se, na Tabela 5, que seis estudantes do 5º ano e nove do 9º ano apresentaram respostas parcialmente corretas. Classificamos de tal forma, pois estes consideraram como população da pesquisa todos os moradores do Recife e não somente os eleitores da cidade.

S06 - A do Recife.

S10 - Do Recife.

S25 - Recife.

Os participantes que definiram corretamente conseguiram limitar a população ao contexto da pesquisa, nesse caso os eleitores da cidade do Recife, como observa-se nos exemplos abaixo.

S19 - Pode ser a população... dos eleitores.

P - Dos eleitores da onde?

S19 - Do Recife.

S27 - Os eleitores do Recife.

S29 - Os eleitores.

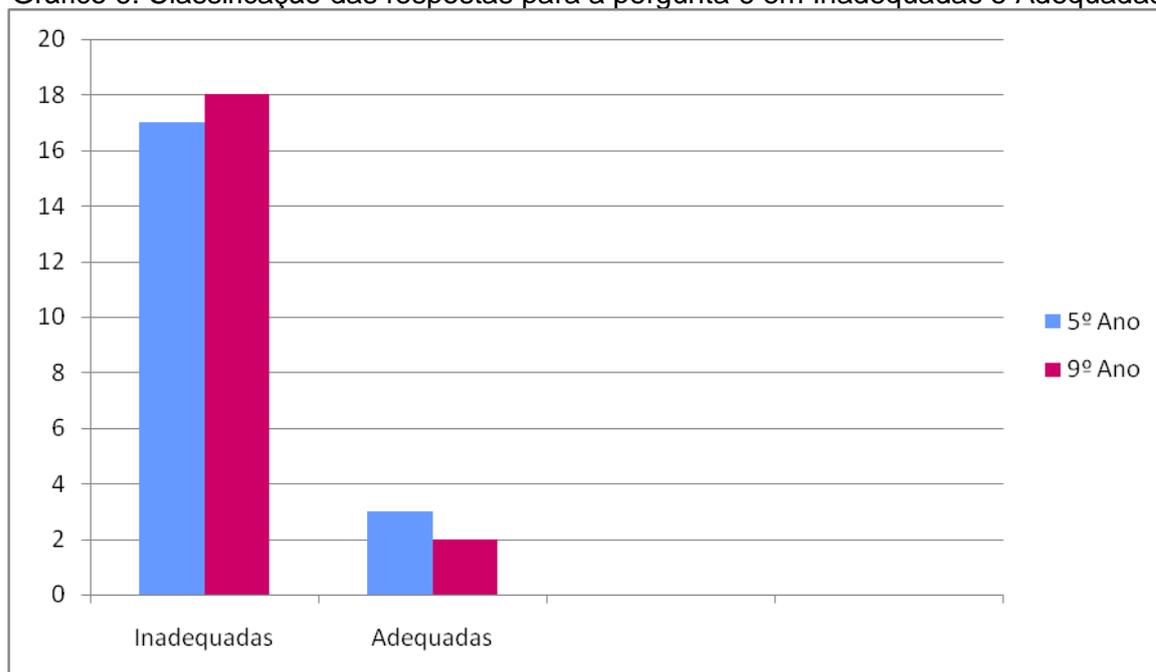
P - Da onde?

S29 - Do Recife.

A sexta análise ainda refere-se ao conceito de população, agora como conjunto de objetos. Esperava-se que os estudantes percebessem que população não se limita apenas a pessoas. Para isso, foi requerido que respondessem a seguinte questão:

Se fosse realizada uma pesquisa para saber quanto tempo duram os computadores da marca X. Qual seria a população analisada nessa pesquisa?

Gráfico 6: Classificação das respostas para a pergunta 6 em Inadequadas e Adequadas



Ao contrário da questão anterior, observa-se no Gráfico 6, que a maioria dos estudantes responde de forma inadequada. Apesar do número de respostas inadequadas do 5º ano ter sido menor que o do 9º, essa diferença não foi significativa (X^2 0,229 gl 1, $p < 0,633$).

Tabela 6: Frequência por ano de escolaridade da pergunta 6.

Categorias	Ano de ensino		Total
	5º Ano	9º Ano	
Não define	1	0	1
Define incorretamente	16	18	34
Define parcialmente correto	1	0	1
Define corretamente	2	2	4

Percebe-se na Tabela 6 que a maioria dos estudantes respondeu incorretamente (trinta e quatro participantes), demonstrando que os mesmos ainda não compreendem população como outro grupo que não de pessoas, tais como animais, espécies de plantas e objetos inanimados.

Esse tipo de pesquisa acerca da durabilidade de alguns objetos é bastante comum principalmente em testes de qualidade. Para essa situação específica a resposta correta para a população analisada seriam os computadores da marca X.

Entre as respostas incorretas foi normal definir a população da pesquisa citada como conjunto de pessoas (usuários da marca, técnicos, entre outros).

S04 - A população quem vai ser, vai ser quem fez o computador.

S12 - A da Vila Arraes (Vila situada próximo a escola).

S22 - O cara que mexe com informática pra saber quanto tempo, porque ele entende mais.

S27 - Quem compra o CCE.

Apenas um estudante do 5º ano apresentou uma resposta que consideramos parcialmente correta, pois o mesmo indicou ter certa compreensão de que, nesse caso, a população não seria de pessoas, mas sim do objeto analisado. A resposta dada pelo aluno está transcrita abaixo para uma melhor análise.

S01 - A fábrica.

P - A fábrica? Qual fábrica?

S01 - Do computador?

P - A fábrica dos computadores dessa marca aí?

S01 - É.

Percebe-se que, mesmo sem definir diretamente qual seria a população, o participante se aproxima da resposta adequada quando afirma que seria a fábrica, estando implícito que os computadores são produto e fazem parte desta.

Dos quarenta participantes, quatro responderam corretamente (dois do 5º ano e dois do 9º ano). Nas respostas apresentadas nota-se uma maior compreensão do contexto da pesquisa.

S02 - Os computadores da HP.

S06 - Os computadores da marca.

S29 - Os computadores da marca.

S40 - Os computadores da marca que eles queriam saber.

Nota-se que dos 5 (cinco) estudantes que apresentaram respostas mais adequadas (parcialmente corretas e corretas), 3 (três) eram do 5º ano. Esses alunos perceberam que população nem sempre são pessoas, mas sim o universo a ser

investigado. Esses dados sugerem que, mesmo conceitos mais complexos, podem ser trabalhados de forma mais simples nos anos iniciais de escolarização.

No que refere-se a definição de uma população, como analisado por Rubin, Bruce e Tenny (1990), a maioria dos estudantes embasam suas respostas em opiniões pessoais. Contudo, observou-se que, para a compreensão do conceito de população como grupo de pessoas, as experiências individuais, embora não envolvam critérios estatísticos, corroboraram para a construção desse conceito, sugerindo que, alguns aspectos relacionados à amostragem são mais facilmente compreendidos pelos estudantes que outros.

Embora em nossa pesquisa tenham sido apresentadas apenas duas situações envolvendo a ideia de população, nota-se que o trabalho com diferentes contextos é importante para que os estudantes entrem em conflito com suas justificativas e construam conceitos baseando-se em critérios estatísticos e não apenas em suas vivências pessoais.

Sendo assim, ressalta-se a importância do que Pfannkuch (2008) expõe em suas pesquisas. Seus estudos apontam a necessidade de contextualizar o ensino de Estatística, envolvendo os alunos com discussões e construção de ideias iniciais sobre população e amostra. Da mesma forma, Ben-Zvi, Makar, Bakker e Aridor (2011) perceberam que a variedade de situações estimula os alunos a pensarem sobre as relações população-amostra, sendo indispensável a intervenção do professor.

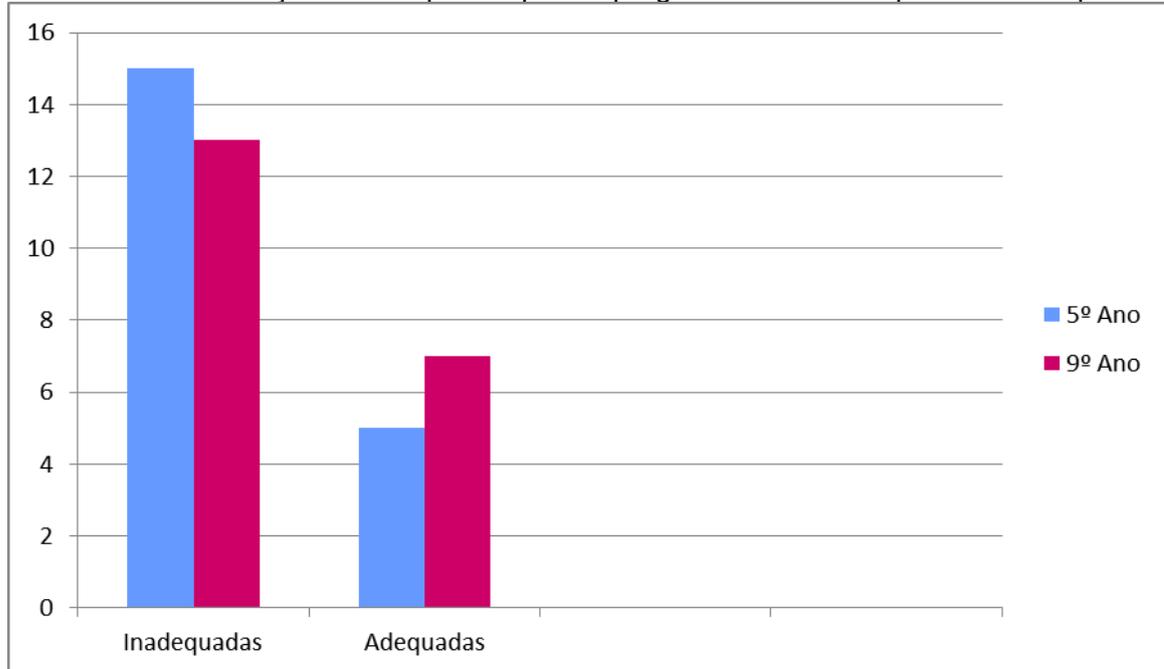
Bloco 3 – Seleção da Amostra

Em relação à seleção de uma amostra, na primeira questão os estudantes eram convidados a propor uma estratégia para a seleção de uma amostra, de forma que esta fosse o mais representativa possível da população analisada.

Um pesquisador queria saber a merenda preferida dos alunos das escolas públicas de Recife. Como ele não tinha condições de entrevistar todos os alunos resolveu entrevistar apenas 200 alunos. Como ele poderia escolher esses alunos para ter uma ideia melhor da preferência de todos?

Percebe-se, no Gráfico 7, que o número de respostas adequadas é maior entre os participantes do 9º ano. Contudo, novamente não houve diferença significativa entre os anos de escolaridade (X^2 0,476, gl 1, $p < 0,490$).

Gráfico 7: Classificação das respostas para a pergunta 7 em Inadequadas e Adequadas



Nessa variável, para distribuição das subcategorias, levou-se em conta os critérios elencados para escolha da amostra. Os resultados para tal classificação são expostos na Tabela 7:

Tabela 7: Frequência por ano de escolaridade da pergunta 7.

Categorias	Ano de ensino		Total
	5º Ano	9º Ano	
Não responde	4	1	5
Responde incorretamente	11	12	23
Responde parcialmente correto	0	0	0
Responde corretamente	5	7	12

Foram consideradas respostas incorretas aquelas que apresentaram características de amostras não-probabilísticas, ou seja, que não poderiam ser generalizadas para toda a população, pois não estabeleciam critérios para sua escolha ou ofereciam viés de seleção, como os exemplos a seguir:

S06 - Pela professora.

P - Como assim?

S06 - É porque ele queria saber qual a merenda preferida.

P - E como ele ia escolher esses alunos pela professora?

S06 - Ele ia pegar a lista das professoras e elas iam escolher os alunos.

P - E como ele ia escolher essas professoras?

S06 - Na reunião ele ia perguntar quem queria.

S11 - Perguntando de um em um.

P - Perguntando de um por um? A todos os alunos?

S11 - É.

S23 - Pegar um aluno de cada escola. Sei lá!

P - Ele ia pegar um aluno de cada escola do Recife?

S23 - É.

P - E como ele ia escolher esse aluno?

S23 - Ia na sala e chamava. Do mesmo jeito que me chamou.

P - Ia chamar ou perguntar quem queria?

S23 - Era. Mais ou menos assim.

Consideraram-se como respostas corretas as que elencavam pelo menos um aspecto referente a uma amostra probabilística, ou seja, que pode ser generalizada para toda a população. Entre esse tipo de resposta foi comum os participantes sugerirem que a amostra fosse escolhida a partir de um sorteio, característica de uma amostra aleatória.

S02 - Sorteando.

P - Sorteando? Mas ia sortear o quê primeiro?

S02 - A merenda.

P - A merenda?

S02 - Não. Sortear os alunos.

P - Os alunos da onde?

S02 - Do colégio. Daqui

P - Então, pra ele saber qual a merenda preferida dos alunos de Jaboatão ele vai sortear os 200 alunos dessa escola?

S02 - Não. Tem que ser mais.

P - Ele ia escolher como as escolas?

S02 - Sorteando.

P - Sorteando também? Então ele iria sortear as escolas e das escolas sorteadas ele ia sortear os alunos?

S02 - (Concorda com a cabeça)

S28 - Ele dividia pelas escolas que tem no Recife e sorteava.

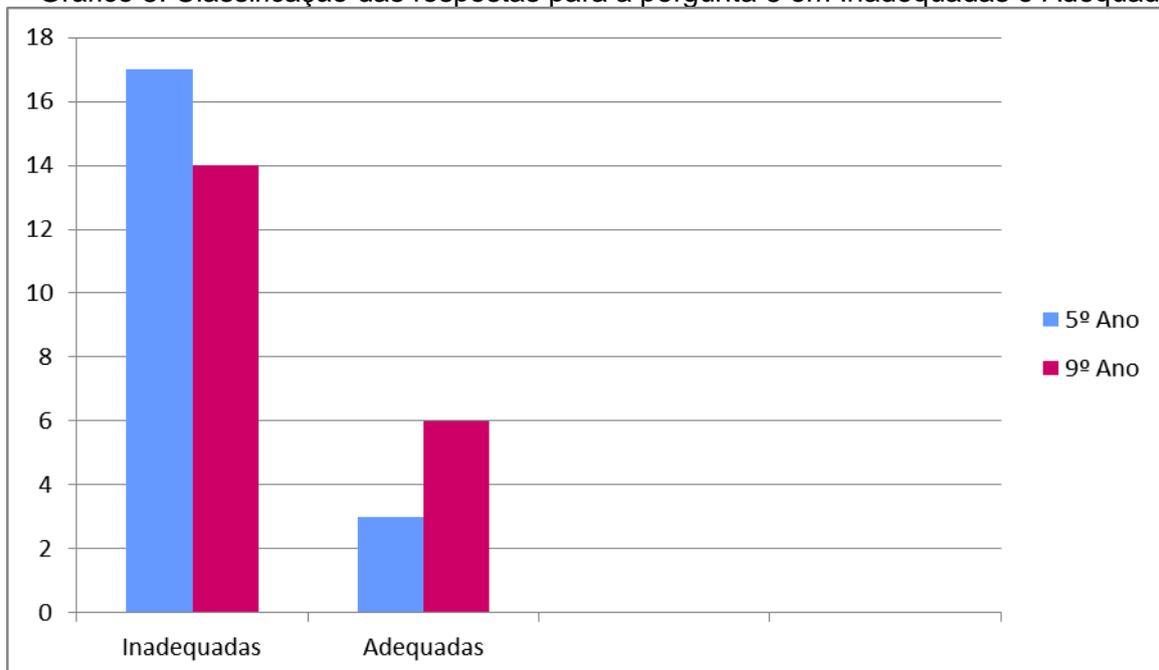
Ainda em relação à seleção de uma amostra, foi apresentada uma situação hipotética, na qual eram listadas cinco opções para a escolha da amostra, e os

participantes eram solicitados a apontar qual seria a mais apropriada. No Gráfico 8, tem-se os resultados para essa questão.

Cinco amigos queriam saber aproximadamente quantos livros as pessoas que moram no bairro deles liam por ano. Como o bairro tinha uns 10.000 moradores, não dava para entrevistar todo mundo. Cada um teve uma ideia para saber quem podiam entrevistar. Qual dessas ideias você acha que será melhor para saber o que eles querem? Por quê?

Amigo 1	100 moradores que frequentavam a biblioteca da comunidade.
Amigo 2	100 moradores do bairro.
Amigo 3	10 moradores que frequentavam a biblioteca da comunidade.
Amigo 4	10 moradores do bairro.
Amigo 5	Homens, mulheres, meninos e meninas.

Gráfico 8: Classificação das respostas para a pergunta 8 em Inadequadas e Adequadas



Mais uma vez não houve diferença estatística entre os anos de escolaridade analisados (X^2 1,290, gl 1, $p < 0,256$). Os resultados desta questão, semelhantemente a atividade anterior, confirmam a dificuldade dos estudantes em selecionar uma amostra o mais representativa possível.

Para uma melhor análise dessas dificuldades e das respostas apresentadas pelos estudantes, a Tabela 8 traz a classificação das respostas nas quatro subcategorias.

Tabela 8: Frequência por ano de escolaridade da pergunta 8.

Categorias	Ano de ensino		Total
	5º Ano	9º Ano	
Não responde	1	0	1
Responde incorretamente	16	14	30
Responde parcialmente correto	2	5	7
Responde corretamente	1	1	2

Nessa variável, estava sendo observado se os participantes seriam tendenciosos na escolha da amostra.

A opção mais adequada seria o Amigo 2, que entrevistou 100 moradores do bairro independente de ir à biblioteca ou não, o que deixa implícito que na amostra existiriam frequentadores e não-frequentadores da biblioteca, que poderiam ser de diferentes gêneros e faixas etárias.

Os estudantes que escolheram o Amigo 5 (homens, mulheres, meninos e meninas), tiveram suas respostas classificadas como parcialmente corretas, pois argumentaram sobre a variabilidade de gênero e idade. Entretanto, não consideraram importante o fato da amostra ser dos moradores do bairro.

Na Tabela 8 nota-se que, mais da metade dos estudantes (dezesseis do 5º ano e quatorze do 9º) responderam incorretamente, apresentando um viés de seleção em suas opções, como nos exemplos abaixo:

S03 - O amigo 1.

P - Por quê?

S03 - Porque foi mais interessante.

P - Por que foi mais interessante?

S03 - Porque frequentam a biblioteca.

S08 - O primeiro.

P - Por que o primeiro?

S08 - Porque pegou gente que frequentava a biblioteca.

S22 - O primeiro.

P - Por que o primeiro?

S22 - Porque pegou gente que frequentava a biblioteca.

Percebe-se que, esses estudantes quando solicitados a justificar a sua escolha argumentavam por ser uma amostra que frequentava a biblioteca, característica que torna a amostra tendenciosa e, portanto, não generalizável para toda a população.

As respostas consideradas parcialmente corretas (duas do 5º ano e cinco do 9º ano) foram classificadas de tal forma, pois, embora optassem por uma amostra não tendenciosa, suas justificativas não abordavam o porquê de uma amostra do bairro todo ser melhor que apenas a dos frequentadores da biblioteca.

S25 - O 2 que entrevistou 100.

[...]

P - Mas o um também entrevistou 100. Por que o 2 e não o 1?

S25 - Porque é do bairro todo e o outro foi da biblioteca.

Apenas dois estudantes deram uma resposta levando em conta que entre a amostra de moradores do bairro haveria pessoas que frequentavam ou não a biblioteca, permitindo com que esta representasse de forma mais significativa a população da pesquisa.

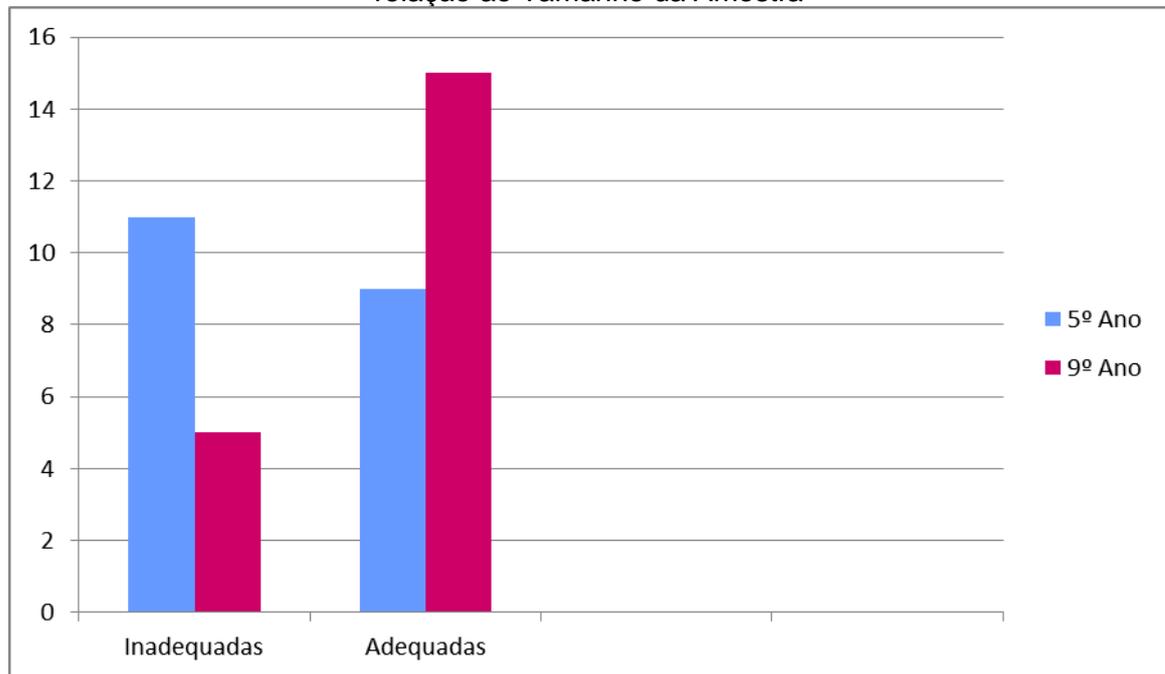
S13 - Pegou quem vai e quem não vai (para a biblioteca).

S40 - Porque era do bairro mesmo sem ir pra biblioteca.

A análise seguinte refere-se à mesma questão, mas analisamos as respostas em função do tamanho da amostra, o qual deve ser definido claramente para a seleção da mesma. A situação apresentada aos participantes para essa observação foi a mesma para a segunda análise sobre seleção da amostra, mas avaliando se os estudantes levavam em conta o tamanho da amostra escolhida em relação ao tamanho da população.

Ao contrário da maioria das questões, essa apresentou um maior quantitativo de respostas adequadas. Embora haja diferença entre os anos de escolaridade a mesma não foi significativa (X^2 3,750, gl 3, $p < 0,053$).

Gráfico 9: Classificação das respostas para a pergunta 8 em Inadequadas e Adequadas em relação ao Tamanho da Amostra



Ao subdividir as respostas, percebe-se que o maior quantitativo de participantes responde parcialmente correto ou não responde, ou seja, não relaciona o tamanho da amostra com a população ou estabelece uma relação sem justificativas claras.

Tabela 9: Frequência por ano de escolaridade da pergunta 8 em relação ao tamanho da amostra.

Categorias	Ano de ensino		Total
	5º Ano	9º Ano	
Não responde	9	4	13
Responde incorretamente	2	1	3
Responde parcialmente correto	9	14	23
Responde corretamente	0	1	1

Nesta variável é interessante perceber Vale ressaltar que, quando os participantes escolheram a amostra que consideravam ideal, os mesmos eram questionados sobre outra opção semelhante ou no quantitativo de pessoas ou no local de pesquisa. Por exemplo, se o estudante optasse pelo Amigo 1 justificando por ser 100 moradores, ele era conflitado com o Amigo 2, que também era uma amostra de 100 pessoas. Caso este argumentasse por serem frequentadores da biblioteca era indagado por que não o Amigo 3, que também tinha uma amostra de pessoas que iam à biblioteca. Essas comparações foram feitas com qualquer

resposta dada. Contudo ainda não foi suficiente para que os participantes relacionassem o tamanho da amostra com a população pesquisada.

Poucos estudantes relacionaram incorretamente (dois do 5º ano e um do 9º), argumentando que cem pessoas era muito ou sem estabelecer ligação entre o número de moradores do bairro e a necessidade da amostra ser representativa deste.

S05 - Porque é melhor tia. O primeiro pegou 100 pessoas. Assim, eu acho que é o 4 porque não tem muitas pessoas.

S39 - Mas é muita gente, pode dar alguma coisa e não dar certo.

As respostas consideradas como parcialmente corretas tiveram essa classificação, pois os estudantes apenas levaram em conta a maior quantidade de moradores, sem relacionar a população com o número de sujeitos da amostra e um dos critérios para a sua representatividade.

S01 - Porque o um tem mais moradores.

S13 – [...] O 2. Porque tem mais pessoas.

S33 - Porque tem mais ideia, tem mais gente.

Os dados da Tabela 9 mostram que somente um participante do 9º ano respondeu corretamente, destacando que o tamanho da amostra é um dos fatores importantes para representatividade desta. E que, embora uma amostra menor também pudesse representar a população no exemplo citado, seria mais adequado um número maior de sujeitos.

S34 – [...] porque o três pegou só 10. Vai dar, mas vai dar muito pouco pra saber do bairro todo.

Nos estudos de Rubin, Bruce, e Tenny (1990) com estudantes do Ensino Médio, os participantes demonstraram modelos inconsistentes em relação à seleção da amostra para que esta represente a população desejada, pois ao analisarem os resultados obtidos perceberam que, as respostas dos estudantes ora são embasadas em suas intuições sobre representatividade da amostra, ora na

variabilidade da mesma, não sendo o tamanho da amostra relacionado a esses conceitos.

Essa compreensão acerca da validade da amostra também foi analisada por Innabi (2006) de forma semelhante ao realizado em nossa pesquisa. Ao serem questionados sobre a validade de uma amostra, apresentando-se maneiras diferentes de como esta poderia ter sido selecionada, a maioria dos estudantes do ensino secundário não levou em consideração ao justificarem suas respostas, o tamanho da amostra e nem se a mesma é tendenciosa ou não quando julgaram a validade e representatividade da mesma.

Nos estudos de Watson e Kelly (2002), estudantes entre 8 e 9 anos levantaram questões sobre a seleção de uma amostra representativa para o todo, entretanto, apresentaram dificuldade em compreender ideias acerca do tamanho da amostra.

O mesmo pode ser percebido ao analisarmos os resultados apresentados aqui, a maioria dos entrevistados não conseguiu elencar aspectos importantes que devem ser levados em consideração ao selecionar uma amostra, tais como variabilidade, tamanho e tendenciosidade da mesma, demonstrando a necessidade de se trabalhar com situações de pesquisa, nas quais os estudantes possam vivenciar as etapas de uma pesquisa estatística desenvolver habilidades para a seleção de uma amostra representativa. Embora em relação ao tamanho da amostra o desempenho dos participantes desta pesquisa tenha sido mais adequado, a maioria das suas respostas apenas levava em conta a amostra maior, sem relacionar a população com o número de sujeitos da amostra e a importância deste para a sua representatividade.

O nosso estudo, bem como pesquisas anteriores, apontou que a elaboração de critérios para a seleção de uma amostra a fim de que esta seja o mais representativa possível é uma tarefa difícil para estudantes de diferentes idades. A partir dos resultados alcançados no que diz respeito ao tamanho da amostra avaliamos se este pode ser um aspecto utilizado como ponto de partida para explicar o cuidado necessário para seleção de amostras. Propor atividades que envolvam os conhecimentos que os estudantes já possuem sobre amostragem ou relacionados pode ser uma alternativa para desenvolver outros conceitos mais complexos.

O trabalho com diferentes atividades contribui para a construção desses conceitos, como observado nos estudos de Pfannkuch (2008) a qual desenvolveu uma sequência didática acerca dos conceitos referentes à amostra com tamanhos diferentes e indicou que os alunos passaram a compreender algumas noções sobre variabilidade da amostragem, ligando a amostra à população, utilizando a linguagem associada à Estatística ou embasando suas respostas a partir das imagens e dados fornecidos.

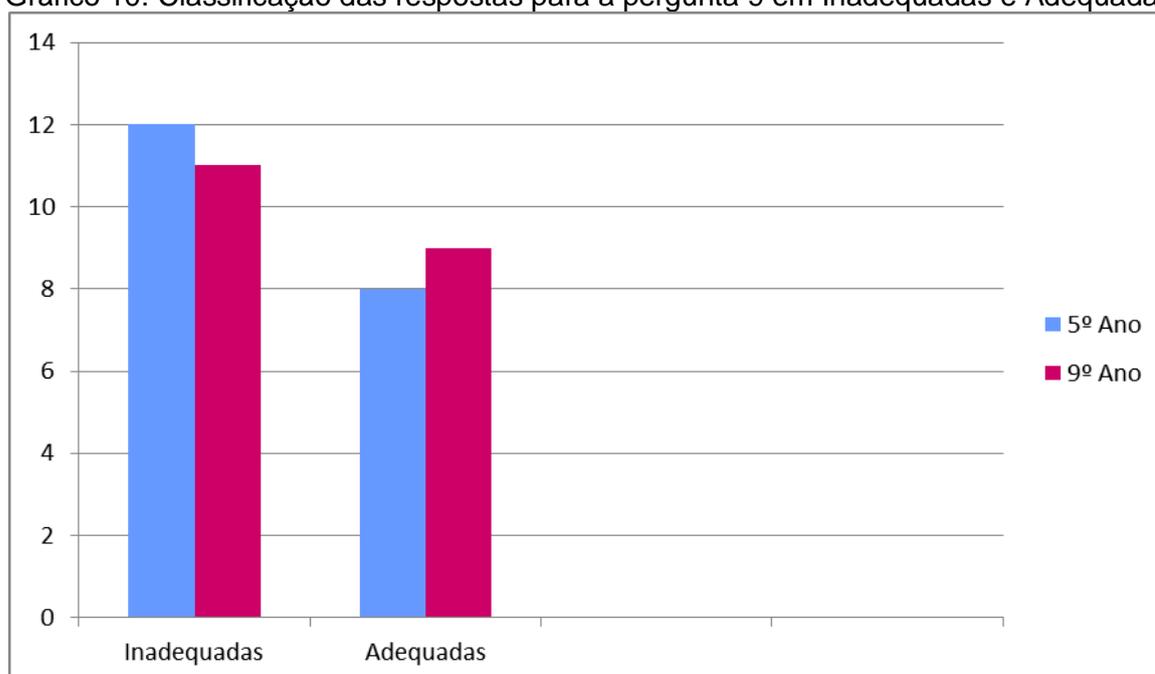
Bloco 4 – Aleatoriedade

Um dos conceitos vinculado ao campo da amostragem é o da aleatoriedade, o qual é objeto da próxima análise. Foi solicitado que os alunos definissem aleatório a partir da seguinte situação proposta:

Para definir a ordem dos alunos na fila para a merenda, a professora colocou o nome dos alunos em um saquinho e foi retirando aleatoriamente. O que significa aleatório para você?

Os dados para esta questão estão expostos no Gráfico 10:

Gráfico 10: Classificação das respostas para a pergunta 9 em Inadequadas e Adequadas



Mais uma vez não houve diferença estatística entre os anos de escolaridade pesquisados ($X^2 0,102$, gl 1, $p < 0,749$).

Ao subdividir as respostas essa similaridade no desempenho dos estudantes também fica facilmente perceptível, como mostra a Tabela 10:

Tabela 10: Frequência por ano de escolaridade da pergunta 9.

Categorias	Ano de ensino	5º	9º	Total
		Ano	Ano	
Não responde		3	2	5
Responde incorretamente		9	9	18
Responde parcialmente correto		8	9	17
Responde corretamente		0	0	0

Embora a ideia de aleatoriedade esteja presente em várias situações do cotidiano, tais como jogos de azar, games e sorteios, a palavra *aleatório*, bem como sua definição não é comum no vocabulário informal. Por isso, mesmo grande parte dos participantes não tendo definido ou definido incorretamente (cinco e dezoito, respectivamente), o quantitativo de 17 (dezessete) estudantes que responderam de forma parcialmente correta deve ser considerado razoável e motivador.

Também deve ser destacado que nenhum participante deu uma resposta correta, a qual abrangeria vários aspectos da aleatoriedade, tais como quebra de ordem, propósito, causa ou imprevisibilidade. A aleatoriedade é um processo repetitivo cujo resultado não descreve um padrão determinístico.

É importante explicar que, o questionamento aos estudantes foi feito após a exposição de um contexto, no qual estava implícito a ideia de um sorteio, cuja principal característica é a de ser aleatório. É provável que isso tenha contribuído para os alunos responderem parcialmente correto.

Entre as respostas incorretas foram comuns erros ligados a finalidade de se usar a aleatoriedade no contexto apresentado e não ao conceito do que seria aleatório realmente.

S09 - Pra se organizar pra fazer a fila direito.

S17 - É uma coisa pra ninguém furar fila mais.

S23 - Organização?

P - É a organização que eu vou colocar na fila? Aleatório vai ser pra organizar?

S23 - *É.*

S29 - *Um atrás do outro.*

Como já mencionado anteriormente, o tipo de definição classificada como parcialmente correta leva a acreditar que a situação apresentada aos estudantes, juntamente com suas vivências facilitou a sua compreensão de aleatoriedade.

S08 - *Um sorteio.*

S14 - *Por sorteio?*

P - *Por sorteio?*

S14 - *É sem saber o que é que vai puxar. O que vai sair.*

S27 - *É tipo um sorteio.*

S33 - *É escolher as pessoas sem ter ideia, sem escolher direto.*

P - *Como assim?*

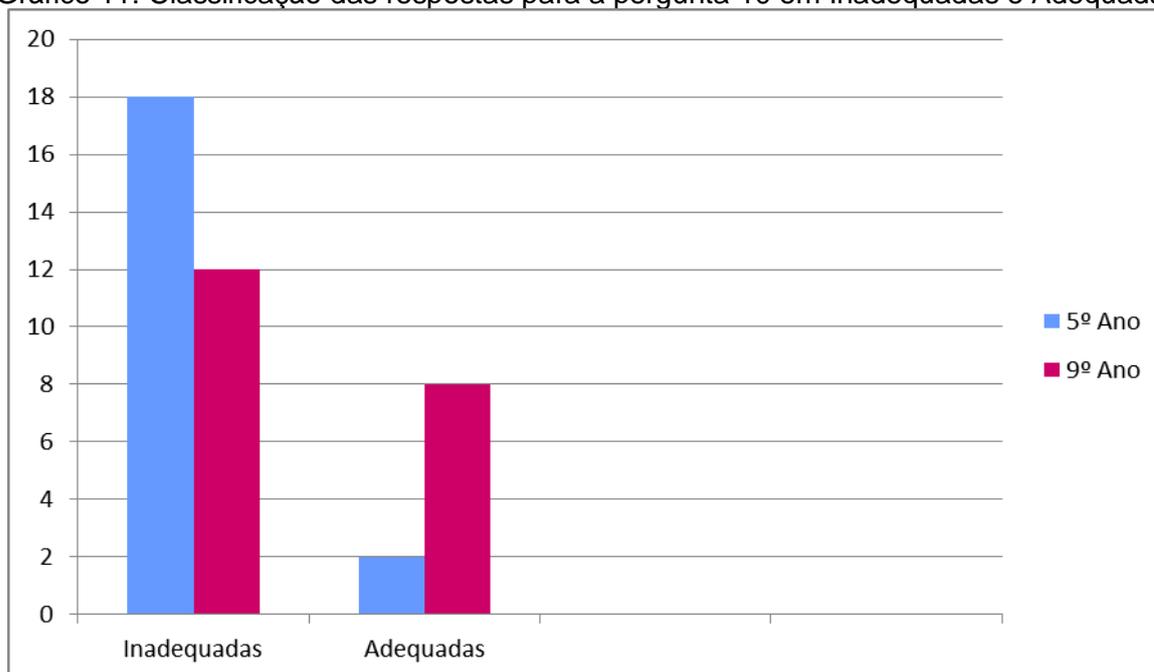
S33 - *Tipo na sorte.*

Após o conceito de aleatoriedade, a próxima análise se destina ao que os estudantes consideram como sendo uma Amostra Aleatória (AA). Esse tipo específico de amostra foi escolhido por ser o mais comumente utilizado em pesquisas probabilísticas. A Tabela 11 traz a classificação das respostas dos estudantes para a questão proposta. Para que os participantes definissem o conceito de AA, era-lhes apresentada a seguinte situação:

Eu quero saber o que as pessoas acham do carnaval do Recife. Para selecionar as pessoas que participarão da pesquisa eu vou usar uma amostra aleatória. O que significa amostra aleatória?

Embora também aborde o conceito de aleatoriedade o desempenho dos estudantes para essa questão é bem distinto, tanto no quantitativo geral de respostas adequadas e inadequadas, quanto entre os anos de escolaridade estudados. Nesse caso, a diferença entre os anos de ensino é estatisticamente significativa (X^2 4,800, gl 1, $p < 0,028$).

Gráfico 11: Classificação das respostas para a pergunta 10 em Inadequadas e Adequadas



A Tabela 11 traz os dados divididos nas quatro subcategorias, a fim de uma análise mais detalhada das respostas.

Tabela 11: Frequência por ano de escolaridade da pergunta 10.

Categorias	Ano de ensino		Total
	5º Ano	9º Ano	
Não define	7	7	14
Define incorretamente	11	5	16
Define parcialmente correto	2	8	10
Define corretamente	0	0	0

Nota-se que a maior parte dos participantes não define (sete do 5º ano e sete do 9º ano) ou define incorretamente (onze do 5º ano e cinco do 9º), o que já era esperado após os resultados das duas questões que envolviam a definição do conceito de amostra, os quais tiveram um maior quantitativo de respostas incorretas (vinte e sete sujeitos em ambas as perguntas).

Entre as respostas incorretas houve tanto respostas que se basearam apenas no contexto da situação apresentada, quanto definições elaboradas a partir do que os estudantes acreditavam ser amostra e aleatório.

*S03 - Botar feito os cantores que tem no Jaboatonense ai bota na parede.
P - E eu vou botar o quê na parede?*

S03 - Carnaval de noite ou de dia.

P - Carnaval de noite ou de dia, isso que vai ser amostra aleatória?

S03 - É. Se escolherem de noite, faz de noite. Se de dia, faz de dia.

S06 - É um carnaval bom.

P - Um carnaval bom vai ser uma amostra aleatória?

S06 - É. Sem briga.

S27 - Sai perguntando pra quem frequenta mais o carnaval.

S37 - Uma sequência de pessoas pra entrevistar.

Apenas dois estudantes do 5º ano e oito do 9º definiram de forma parcialmente correta. Classificamos de tal maneira, pois estas foram as que mais se aproximaram do conceito real de amostra aleatório seja definindo amostra ou aleatório de forma correta ou identificando qual seria a amostra do exemplo dado.

S29 - Eu acho que ele faz um sorteio e usa de amostra.

S33 - Ia juntar o povo que participa das festas de carnaval e ia pegar na sorte como eu te falei.

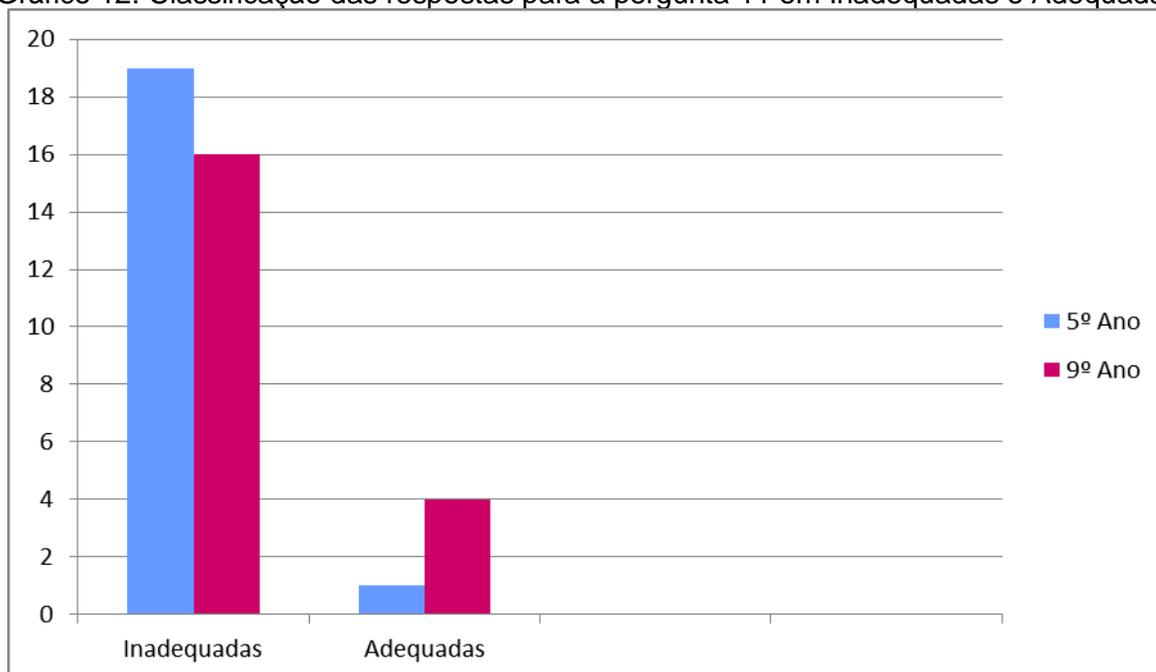
A definição estatística seria: Amostra Aleatória é uma parte da população escolhida ao acaso, na qual, todos têm a mesma chance de pertencer à amostra.

Ainda analisando o que os estudantes compreendem como Amostra Aleatória (AA), foi pedido que os mesmos propusessem uma situação na qual fosse usada uma AA.

Dê um exemplo de outra situação na qual podemos usar uma amostra aleatória.

Após a classificação das respostas em Inadequadas e Adequadas, a realização do Teste Qui-quadrado evidenciou que não houve diferença significativa entre os anos de escolaridade pesquisados (X^2 2,057, gl 1, $p < 0,151$), o que pode ser observado no Gráfico 12.

Gráfico 12: Classificação das respostas para a pergunta 11 em Inadequadas e Adequadas



A distribuição das respostas entre as quatro subcategorias ratifica as dificuldades dos estudantes em compreender o conceito de amostra aleatória, bem como apresentar exemplos para a utilização da mesma.

Tabela 12: Frequência por ano de escolaridade da pergunta 11.

Categorias	Ano de ensino		Total
	5º Ano	9º Ano	
Não responde	9	12	21
Responde incorretamente	10	4	14
Responde parcialmente correto	1	3	4
Responde corretamente	0	1	1

Percebe-se que os resultados, desta variável, explícitos na Tabela 12 são decorrentes da questão anterior, novamente a maior parte dos participantes não exemplifica ou aponta um exemplo incorreto (vinte e um e quatorze, respectivamente).

S18 - O que a pessoa tem pra amostrar. Uma fantasia de carnaval.

S19 - Brincadeiras. Sair mostrando as brincadeiras.

Observa-se na Tabela 12 que quatro participantes (um do 5º ano e três do 9º) apresentaram um exemplo categorizado como parcialmente correto, por trazerem situações que abrangiam de forma simplória ou a ideia de aleatoriedade, o conceito de amostra ou ambos.

S27 - Um concurso aqui na escola. Pra ver quem vai participar ai pega na caderneta e chuta um número.

S28 - Faz um sorteio de novo.

P - Pra quê?

S28 - Pra um campeonato.

P - Como assim?

S28 - Pra ver quem vai contra quem.

Somente um estudante do 9º ano respondeu corretamente, dando um exemplo adequado para o uso de uma amostra aleatória.

S29 - É... Vai ter uma peça de teatro ai só alguns vão. Pede “pros” professores sortearem nas salas os alunos que vão.

Vale destacar que há uma semelhança muito grande entre as respostas dos sujeitos S27 e S29, contudo o que gerou categorizações diferentes foi um aspecto sutil. O S27 indica o “chute” como forma para seleção da amostra, o qual pode conter um viés de seleção. Ao contrário deste o S29 deixa claro que a amostra será escolhida através de sorteios.

A aleatoriedade e o princípio da amostragem aleatória são destacados por Estevam e Fürkotter (2010), pois os mesmos tentam minimizar os erros amostrais, considerando a variabilidade entre indivíduos e a variabilidade entre grupos, ressaltando que a amostra deve ser proporcional ao tamanho dos grupos envolvidos na investigação. Esses autores apontaram a relevância na compreensão da natureza da variabilidade em cada contexto de análise, bem como as dificuldades apresentadas pelos estudantes em apreender esses conceitos.

Essas dificuldades também foram percebidas em nossa pesquisa. A maioria dos estudantes não conseguiu elaborar um conceito adequado para aleatoriedade e amostra aleatória, nem sugerir um exemplo apropriado, no qual fosse possível a utilização desses conceitos.

Resultados como esses levam a pensar na importância da realização de mais intervenções como as realizadas por Gil e Ben-Zvi (2010). Alunos de 12 anos que indicaram inicialmente desconfiança sobre suas inferências, a partir de uma sequência de intervenções sobre a validade das informações baseadas na amostra, passaram a confiar muito mais em suas respostas e a utilizar ideias de aleatoriedade e de amostragem aleatória. Os alunos foram capazes de considerar as implicações da representatividade da amostra e da variabilidade desta, mas ainda não foram capazes de entender as relações entre elas.

Da mesma forma nos estudos de Watson e Kelly (2002) em relação ao pré e pós-teste no qual os alunos melhoraram sua definição de amostra e a compreensão de amostra aleatória, mostrando um avanço em suas percepções acerca desses conceitos e a utilização de termos estatísticos em suas definições.

Bloco 5 – Realização de inferências informais

A análise seguinte refere-se à capacidade dos estudantes de realizarem inferências informais para a população a partir de uma amostra dada.

Em uma escola com 150 alunos foi realizada uma pesquisa sobre o uso do celular. Foi sorteada uma amostra de 20 alunos e realizou-se uma pesquisa, obtendo os dados abaixo.

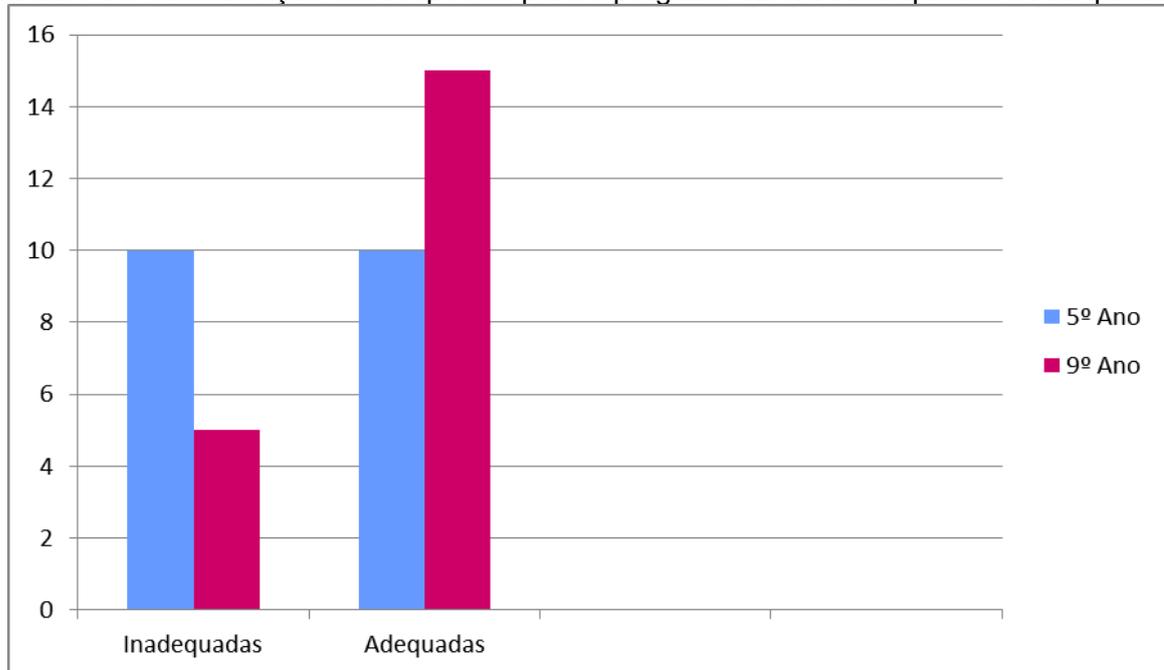
Aluno	Sexo	Para que usa mais	Para que usa menos
Aluno 1	F	Fotos	Ligações
Aluno 2	F	Músicas	Ligações
Aluno 3	F	Fotos	Ligações
Aluno 4	F	Músicas	Mensagem
Aluno 5	F	Músicas	Ligações
Aluno 6	F	Fotos	Ligações
Aluno 7	F	Mensagem	Fotos
Aluno 8	F	Mensagem	Mensagem
Aluno 9	F	Fotos	Ligações
Aluno 10	F	Fotos	Mensagem
Aluno 11	M	Músicas	Ligações
Aluno 12	M	Mensagem	Fotos
Aluno 13	M	Mensagem	Músicas
Aluno 14	M	Mensagem	Fotos
Aluno 15	M	Músicas	Ligações
Aluno 16	M	Mensagem	Fotos
Aluno 17	M	Fotos	Mensagem
Aluno 18	M	Músicas	Fotos
Aluno 19	M	Mensagem	Músicas
Aluno 20	M	Mensagem	Fotos

Após a exposição dos dados eram feitas as perguntas abaixo, a fim de estimular os estudantes a realizarem inferências informais.

- Que conclusões você pode tirar dessa tabela?
- O que você pode dizer sobre o uso do celular nessa escola?

Nota-se, no Gráfico 13, que há um maior quantitativo de respostas adequadas, sugerindo que a realização de inferências informais é uma tarefa mais facilmente realizada pelos estudantes.

Gráfico 13: Classificação das respostas para a pergunta 12 em Inadequadas e Adequadas



Outra vez não houve diferença estatística entre os anos de escolaridade participantes (X^2 2,667, gl 1, $p < 0,102$).

Tabela 13: Frequência por ano de escolaridade da pergunta 12.

Categorias	Ano de ensino		Total
	5º Ano	9º Ano	
Não responde	0	0	0
Responde incorretamente	10	5	15
Responde parcialmente correto	10	15	25
Responde corretamente	0	0	0

A Tabela 13 mostra que nenhum participante deixou de responder aos questionamentos. Contudo, alguns responderam baseados na sua experiência pessoal e não nas informações coletadas da amostra. Essas respostas foram consideradas incorretas, pois o que estava sendo procurado era que conclusões poderiam ser tiradas a partir das informações apresentadas.

S09 - Ligações importantes, mensagens, fotos que os alunos mandam.

S12 - Que é pra ter comunicação se estiver doente ai liga pra mãe e a mãe vem buscar.

S23 - Tipo, agora não que eu não tenho mais celular, mas eu só atendia ligação se fosse da minha mãe, eu podia tá em qualquer aula que atendia. Mensagem eu só respondia se fosse com urgência e só. E música pra escutar pra fazer a tarefa.

P - Essas são as informações que tu podes tirar da tabela ai. Sobre essa amostra?

S23 - Concorda com a cabeça.

S32 - Que é mentira quando dizem que só usam pra ligar pra mãe, que só querem mesmo é escutar música e tirar foto.

As respostas classificadas como parcialmente corretas tiveram essa classificação por se tratarem de inferências muito superficiais, envolvendo apenas uma característica da amostra, como para quem mais usam ou menos usam.

S14 - [...] Só usam mais pra fotos, músicas e mensagem. Não sei pra quê.

S22 - Que usa pra ligar, pra mandar mensagem, que pra o que eles menos usam é pra ligar.

S28 - [...] Eles usam mais pra música e foto e menos pra ligação.

A dificuldade dos estudantes em realizar inferências a partir de uma amostra foi observada também por Rubin, Bruce, e Tenny (1990) ao entrevistarem alunos do Ensino Médio. Eles perceberam que uma das maiores limitações destes estudantes era captar os conceitos básicos de amostragem e inferência.

O mesmo ocorreu nas pesquisas de Watson e Kelly (2002), nas quais as crianças apresentaram dificuldade em entender o raciocínio proporcional necessário para fazer inferências a partir de amostras. Entretanto, elas apresentaram boas

intuições baseadas em suas experiências extra-escolares, ou seja, realizaram inferências informais. O que também podemos perceber ao observar os dados apresentados na Tabela 13. Embora os estudantes não tenham apresentado inferências mais complexas, relacionaram de forma menos sofisticada, os dados da amostra para elaborar conclusões acerca da população.

Mais uma vez, é importante ressaltar o desenvolvimento de atividades que envolvam diversas situações em que seja necessário o raciocínio estatístico a fim de desenvolvê-lo (PFANNKUCH, 2008; BEN-ZVI, MAKAR e ARIDOR, 2011). O trabalho com amostragem desde os anos iniciais é essencial também para contribuir na compreensão de outras habilidades matemáticas mais complexas, como probabilidade, estimação, dispersão, entre outros.

Bloco 6 – Representatividade da amostra

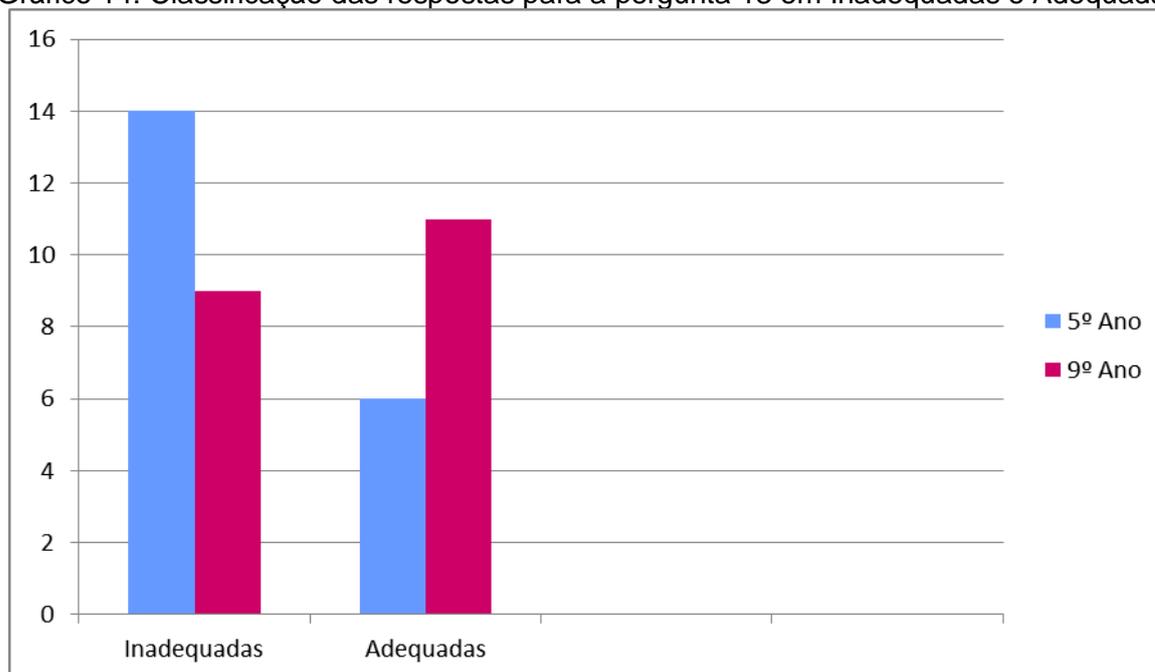
A última análise diz respeito à representatividade da amostra. Embora este aspecto já tenha sido avaliado indiretamente em outras questões, ele é abordado de forma objetiva nesta questão pelo fato da representatividade ser o propósito da amostragem. O questionamento feito aos estudantes era a partir da situação apresentada na questão anterior, lhes sendo indagado:

E se, nesta pesquisa sobre o uso do celular na escola, forem sorteados outros 20 alunos de uma mesma turma? Iria ser melhor ou pior para representar a escola?

Ao classificar as respostas em adequadas e inadequadas e realizar o Teste Qui-quadrado, comprova-se que a diferença existente entre o desempenho do 5º e 9º ano não foi significativa (X^2 2,558, gl 1, $p < 0,110$).

Embora o quantitativo de respostas adequadas tenha sido razoavelmente grande, a maioria dos participantes responde de forma inadequada, o que pode ser observado no Gráfico 14.

Gráfico 14: Classificação das respostas para a pergunta 13 em Inadequadas e Adequadas



A Tabela 14 traz a classificação por subcategorias, permitindo uma melhor visualização e análise da variedade de respostas.

Tabela 14: Frequência por ano de escolaridade da pergunta 13.

Categorias	Ano de ensino		Total
	5º Ano	9º Ano	
Não responde	0	1	1
Responde incorretamente	14	8	22
Responde parcialmente correto	4	9	14
Responde corretamente	2	2	4

Somente um estudante do 9º ano não opinou sobre a representatividade de cada amostra. Mais da metade dos participantes (quatorze do 5º e oito do 9º ano) ao serem solicitados a confrontar as duas amostras responderam incorretamente.

Houve uma variedade nas respostas incorretas, pois alguns estudantes escolheram a amostra que acharam ideal sem expor nenhuma justificativa, outros apresentaram argumentos não condizentes com a representatividade da amostra e outros ainda justificaram com aspectos relacionados à amostragem, mas de forma inadequada, como observa-se nos exemplos seguintes:

S04 - Eu acho que de uma turma. Porque se você sortear da escola toda e pegar uma pessoa que não usa o celular vai dar errado.

S08 - Porque vai ficar mais de uma turma.

S20 - Só de uma turma.

S25 - Melhor (só de uma turma).

P - Por que melhor?

S25 - Porque vai ser mais rápido.

P - Ia ser mais rápido...

S25 - Pra entrevistar eles.

P - E ser mais rápido vai ser melhor?

S25 - É.

As respostas classificadas como parcialmente corretas foram as que escolheram a amostra mais adequada e não justificaram o porquê da sua opção, justificando sem nenhum critério estatístico ou este estava implícito em sua resposta.

S12 - 20 da escola toda.

P - Por quê?

S12 - Porque da sala é pequena e da escola é grande.

P - Da sala é pequena em quê?

S12 - Tem pouca gente e a escola é grande.

S24 - Pior. Porque não vai ter gente de todas as salas.

Como exposto na Tabela 14, um estudante do 5º ano e dois do 9º responderam corretamente levando em conta a variabilidade como fator essencial para a representatividade de uma amostra.

S10 - Não (20 de uma sala só não seria adequado).

P - Por que não?

S10 - Porque cada sala é diferente, as vezes tem alunos pequenos que usam mais pra tirar foto e o primeiro ano pra passar mensagem, pra estudar pra redação...

S14 - Porque se só pegar de uma sala vai ficar muito ruim. Porque só essa sala que vai ter noção e as outras salas podem não concordar com isso. Porque pode ter muita gente que gosta de fazer outra coisa com o celular.

S27 - Pior, professora. Num é melhor escolher da escola toda que tem mais variedade?! Da mesma sala vai ser só daquela.

S29 - Porque de uma só vai ser limitada as opções. Se for da escola toda vai variar mais.

Os resultados apresentados confirmam outras pesquisas (Gil e Ben-Zvi, 2010; Rubin, Bruce e Tenny, 1990; Garfield, 2003; Innabi, 2006), que apontaram as dificuldades dos estudantes em perceber as implicações da representatividade da amostra e, quando compreendem a importância desta, não estabelecem relação com outros fatores como tamanho e variabilidade.

Análise Multidimensional

Após apresentadas e discutidas separadamente cada uma das questões, realizamos uma investigação conjunta das mesmas. Para tal, realizou-se uma análise multidimensional entre cada nível de escolaridade e as questões. Realizamos uma exploração multivariada na qual as questões discutidas anteriormente são retomadas, buscando-se investigar correlações entre os diferentes aspectos da amostragem abordados. Esta análise é importante visto que, nesta pesquisa foram observados diversos conceitos relacionados a amostragem de forma conjunta e não isoladamente.

Procurou-se analisar as correlações possíveis entre as questões que envolviam conceito de amostra, exemplo de amostra, finalidade na utilização de amostras, definição de população, seleção, tamanho e representatividade da amostra e realização de inferências a partir de uma amostra.

Para a realização desta análise multidimensional, utilizou-se o escalonamento multidimensional (Multidimensional Scaling - MDS), que pode ser definido como:

[...] um método baseado na proximidade entre objetos, temas ou estímulos utilizados para produzir uma representação espacial desses itens. Proximidades expressam a semelhança ou não entre os dados. [...] O MDS está preocupado com essa representação em coordenadas euclidianas. As projeções desejadas são encontradas através de uma decomposição espectral apropriada de uma matriz de distâncias (HÄRDLE and SIMAR, 2007, p.331, tradução nossa).

Segundo Oliver (1998) o objetivo principal do MDS é a construção de um espaço métrico com o menor número de dimensões possíveis, que permita representar as proximidades entre os elementos com o maior grau de fidelidade. De acordo com a autora, para compreender essa técnica é necessário o conhecimento

de conceitos como proximidade (valor que assume a similaridade ou distância que existe entre dois objetos) e dimensionalidade (número de dimensões necessárias para representar um conjunto de objetos a partir dos índices de proximidade obtidos).

Aqui, utilizou-se o modelo euclidiano, definido por Oliver (1998) como a raiz quadrada da soma das diferenças entre os elementos ao quadrado. Para isso, o algoritmo ALSCAL (Algorithmic Scaling) foi empregado e incorpora os índices do “ajuste”. Com o ALSCAL, índices como Stress e RSQ (correlação múltipla ao quadrado) são fornecidos. O primeiro indica a qualidade com que a configuração reproduz a informação original. O zero é considerado um “ajuste” perfeito, enquanto que valores superiores a 0,2 correspondem a maus “ajustes” (PORCAR e ESCALANTE, 2009). Já o RSQ é um índice de porcentagem da variância explicada para a configuração obtida. Ele corresponde ao quadrado da correlação entre os dados e as distâncias. Um bom ajuste implica valores de RSQ próximo a 1 (um), isto é, quanto mais próximo de 1 (um) o valor do RSQ mais ajustada está a configuração.

Aplicada a técnica do escalonamento multidimensional com as questões para os sujeitos do 5º ano, em duas dimensões, verificou-se que o índice de Stress (0,15) revelou qualidade no ajuste. Quanto ao índice de RSQ (0,89427), o valor em termos percentuais foi de 89,427% (também utilizado), o que numa escala, pode ser compreendido como um valor bastante alto.

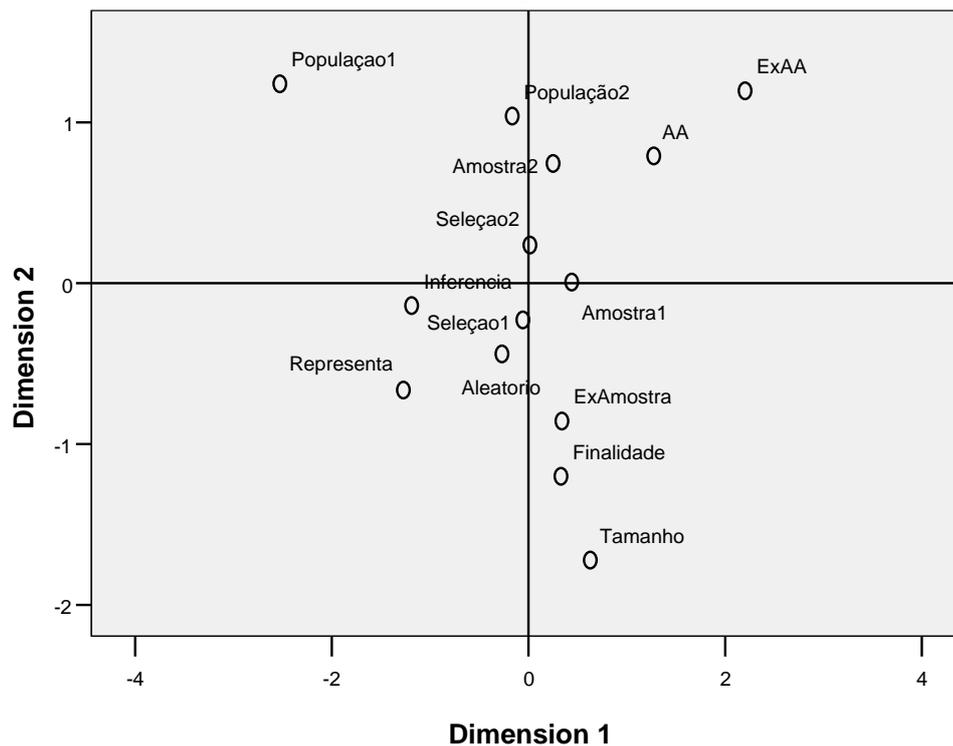
A partir das similaridades e distâncias entre os pontos, encontrou-se os valores para as dimensões 1 e 2, no 5º ano, em função das questões, os quais estão apresentados a seguir, no Quadro 3:

Quadro 3: Valores das dimensões 1 e 2 para o 5º ano

Questões	Dimensão 1	Dimensão 2
Amostra1 (<i>Definição de amostra</i>)	0,4396	0,0066
Amostra2 (<i>Definição de amostra</i>)	0,2498	0,7443
ExAmostra (<i>Exemplo de amostra</i>)	0,3401	- 0,8574
Finalidade (<i>Finalidade da amostra</i>)	0,3309	- 1,2006
População1 (<i>Definição de população</i>)	- 2,5277	1,2391
População2 (<i>Definição de população</i>)	- 0,1653	1,0387
Seleção1 (<i>Seleção da amostra</i>)	- 0,0569	- 0,2288
Seleção2 (<i>Seleção da amostra</i>)	0,0146	0,2368
Aleatório (<i>Definição de aleatoriedade</i>)	- 0,2695	- 0,4397
AA (<i>Definição de amostra aleatória</i>)	1,2742	0,7910
ExAA (<i>Exemplo de amostra aleatória</i>)	2,2020	1,1956
Tamanho (<i>Tamanho da amostra</i>)	0,6287	- 1,7221
Representa (<i>Representatividade da amostra</i>)	- 1,2723	- 0,6644
Inferência (<i>Inferências a partir da amostra</i>)	- 1,1882	- 0,1391

A partir das informações do Quadro 3, o Gráfico 15 foi gerado. O mesmo expõe a distribuição espacial das catorze variáveis em duas dimensões, no modelo euclidiano, entre os sujeitos do 5º ano.

Gráfico 15: Distâncias bidimensionais das variáveis para o 5º ano (Euclidean distance model)



O gráfico apresentado expõe a distribuição espacial das catorze variáveis em duas dimensões, no modelo euclidiano, entre os sujeitos do 5º ano. É importante destacar que a proximidade entre as questões demonstra o comportamento dos sujeitos para aquela questão, isso significa que há uma correlação maior entre o comportamento dos sujeitos do 5º ano para uma determinada questão em relação às outras. Assim, a proximidade entre pontos no gráfico revela que os mesmos sujeitos que responderam adequadamente ou não a uma questão, fizeram o mesmo com aquelas questões correlacionadas. Como é o caso das questões referentes ao conceito de população (ponto população1, no gráfico) e a representatividade da amostra (ponto Representa). Vale esclarecer que a correlação entre as questões não é definida pelo percentual de adequação da resposta para a questão.

Destaca-se que, o agrupamento das questões não significa, em todos os casos, habilidades parecidas para a resolução das mesmas, pois se assim fosse os pontos População1 e População2 estariam bem próximos uma vez que correspondem às questões sobre a definição de uma população. O gráfico sugere que apesar de acharmos que é a mesma habilidade, a situação ser pessoa ou objeto, é determinante. O mesmo acontece com as questões que envolviam o conceito de amostra, representadas pelos pontos Amostra1 e Amostra2. Esses resultados parecem demonstrar que os estudantes não possuem uma definição consolidada do que seria amostra e população.

Já as questões que envolviam aspectos referentes à aleatoriedade, representatividade e realização de inferências, as quais, de certa forma, estão ligadas à seleção da amostra, ficaram bastante próximas, o que pode indicar que esses conceitos necessitam de habilidades semelhantes para compreensão.

Na parte inferior estão concentradas as questões que além de apresentarem uma situação hipotética, os estudantes poderiam responder de forma mais subjetiva, baseando-se, em grande parte, em seus conhecimentos pessoais (pontos Inferência, Seleção1, Finalidade, Representa e Tamanho). Como esses conceitos não são trabalhados sistematicamente na escola, principalmente nesse nível de escolaridade, as definições dadas pelos alunos mostram a forte presença das suas experiências pessoais. Os conceitos que estão mais presentes no cotidiano dos estudantes (população como grupo de pessoas, por exemplo) proporcionaram um maior número de respostas adequadas. Em contrapartida, aqueles que estão presentes em nosso dia a dia, mas não explicitamente, acarretaram respostas

inadequadas. A proximidade dos pontos evidencia que os conceitos apresentam similaridade, quer dizer, que os alunos que erraram uma erraram a outra. O que para nós indica é que devem apresentar habilidades diferentes das demais questões.

Assim como no 5º ano, no 9º buscou-se analisar as catorze questões já destacadas no modelo euclidiano para duas dimensões. Constatou-se que o índice de Stress encontrado (Stress 0,14) foi melhor que o do 5º ano, uma vez que esteve mais próximo de zero, “ajuste perfeito”. O índice de RSQ (0,91957) foi um pouco maior que o do 5º ano, em termos percentuais foi 91,957%, mais próximo ainda de 100%, sendo assim um valor alto.

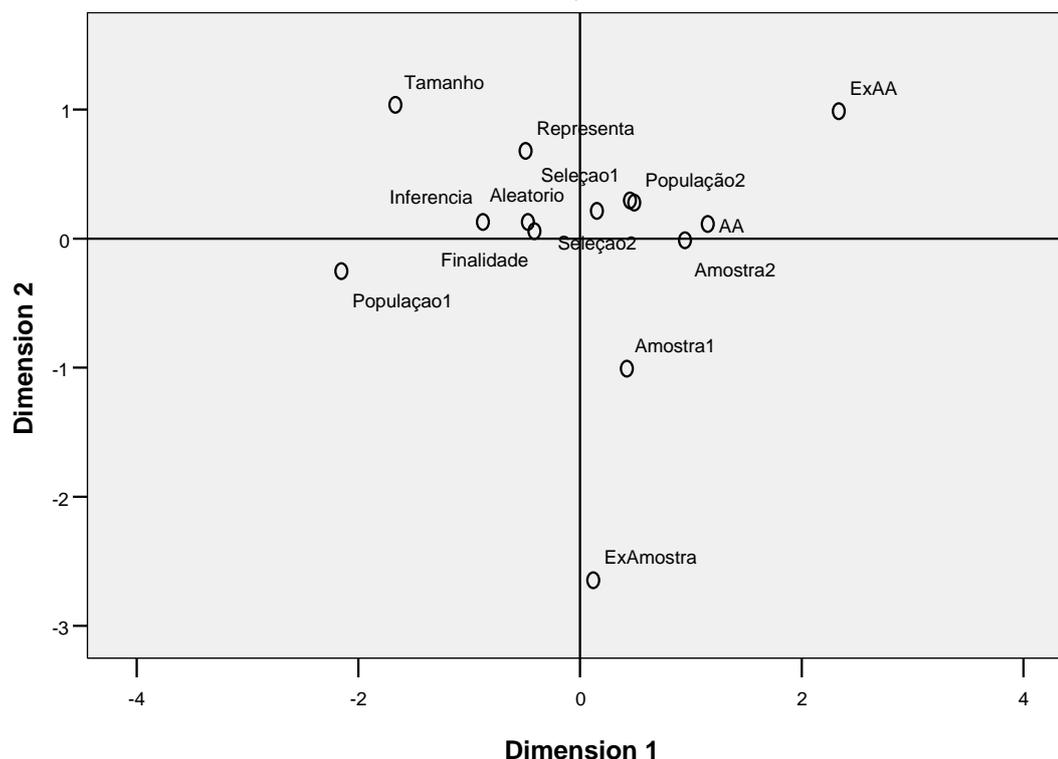
Quanto aos valores para as dimensões 1 e 2 das questões, obteve-se os números apresentados no Quadro 4:

Quadro 4: Valores das dimensões 1 e 2 para o 9º ano

Questões	Dimensão 1	Dimensão 2
Amostra1 (<i>Definição de amostra</i>)	0,4224	- 1,0077
Amostra2 (<i>Definição de amostra</i>)	0,9468	- 0,0128
ExAmostra (<i>Exemplo de amostra</i>)	0,1197	- 2,6468
Finalidade (<i>Finalidade da amostra</i>)	- 0,4115	0,0574
População1 (<i>Definição de população</i>)	- 2,1531	- 0,2526
População2 (<i>Definição de população</i>)	0,4893	0,2789
Seleção1 (<i>Seleção da amostra</i>)	0,4506	0,2954
Seleção2 (<i>Seleção da amostra</i>)	0,1520	0,2145
Aleatório (<i>Definição de aleatoriedade</i>)	- 0,4697	0,1284
AA (<i>Definição de amostra aleatória</i>)	1,1525	0,1135
ExAA (<i>Exemplo de amostra aleatória</i>)	2,3347	0,9873
Tamanho (<i>Tamanho da amostra</i>)	- 1,6661	1,0362
Representa (<i>Representatividade da amostra</i>)	- 0,4913	0,6794
Inferência (<i>Inferências a partir da amostra</i>)	- 0,8763	0,1288

A partir das informações do Quadro 4, o Gráfico 16 foi gerado, o qual expõe a distribuição espacial das catorze variáveis em duas dimensões, no modelo euclidiano, entre os sujeitos do 9º ano.

Gráfico 16: Distâncias bidimensionais das variáveis para o 9º ano (Euclidean distance model)



Percebe-se que, as questões sobre a definição do conceito de amostra (pontos Amostra1 e Amostra2) e amostra aleatória (ponto AA) estão próximas umas das outras. O contrário ocorre quando analisamos as questões que solicitavam aos estudantes que dessem exemplos de amostra (pontos ExAmostra e ExAA), embora requeressem habilidades semelhantes, estes pontos não estão próximos entre si nem dos que representam as questões que envolviam o conceito de amostra. O ponto isolada na extremidade inferior do gráfico (examostra) evidencia que saber apresentar um exemplo (questão respondida corretamente pela maioria dos alunos) não significa os mesmos saibam defini-la. Essa separação é bastante interessante, pois, de certa forma, esperava-se que os estudantes que definissem incorretamente o conceito de amostra seriam os que dariam exemplos também incorretos. Do mesmo modo, aqueles que respondessem adequadamente seriam os mesmos que dariam exemplos adequados, o que não acontece como pode-se observar no gráfico.

Consideramos que este comportamento advém da inexistência de uma definição formalizada dos conceitos investigados. Por não possuírem o conhecimento escolar, o qual é atribuído pelo senso comum como correto, os estudantes mostraram-se duvidosos diante das suas explicações o que acarretou a

discrepância entre suas respostas às questões que abrangiam habilidades semelhantes.

No 9º ano, a distribuição dos pontos que representam as questões que envolviam definições de conceito (População1, Amostra1 e Amostra2) é diferente do exposto no gráfico do 5º ano. No gráfico do 9º ano esses pontos estão localizados na parte inferior e mesmo as variáveis População2 e AA, que embora estejam acima estão bastante próximos a linha horizontal. Isso nos leva a pensar se essa diferença na variedade de respostas ocorre devido ao nível de escolaridade, às experiências pessoais ou ambos os casos. O comportamento dos participantes do 9º ano pode ter sido decorrente da maturidade dos mesmos. Por terem um maior tempo de escolaridade talvez esteja implícita em suas respostas a exigência deles mesmos de uma definição mais formalizada. É provável que a tentativa de conceituar de forma mais estruturada, unindo o conhecimento escolar com seus conhecimentos prévios, tenha gerado essa variedade de respostas, uma vez que as experiências de cada indivíduo são únicas e implicam em diferentes formas de pensar.

Vale destacar que, grande parte dos pontos está centralizada no gráfico, indicando, de certa forma, que o desempenho dos estudantes do 9º ano foi similar. Foram os mesmos estudantes que responderam correta ou incorretamente a maioria das questões, ao contrário do que aconteceu com os estudantes do 5º ano. Essa análise nos mostra que o nível de escolaridade foi um fator importante para o tipo de acerto, mesmo que de forma parcial.

É interessante observar que, ao analisarmos os dados de forma descritiva, na maioria das questões, não há diferença estatística entre os níveis de escolaridade. Contudo, ao correlacionarmos a resposta por cada ano de ensino investigado percebe-se a diferença no comportamento dos estudantes. A forma como estes raciocinam e estruturam suas respostas é diferente. Por isso, enfatizamos a importância de estudos cada vez mais amplos e detalhados sobre essa temática.

Em alguns estudos utilizados como base para essa pesquisa, percebe-se que intervenções e a utilização de situações variadas colaboram com a aprendizagem de conceitos relacionados à amostragem (BEN-ZVI, MAKAR e ARIDOR, 2011; GIL e BEM-ZVI, 2010; PFANNKUCH, 2008; WATSON e KELLY, 2002). Embora tenhamos utilizado o mesmo instrumento, com diversos contextos tanto para o 5º como para o 9º ano e o resultado final tenha sido semelhante, a maneira como os participantes desenvolveram suas respostas foi diferente. Em decorrência disso, vale a pena

pensar na adequação das situações em relação à maturidade cognitiva dos sujeitos, o que exige estudos e intervenções mais aprofundados.

CAPÍTULO 4

Conclusões

A Estatística é uma ciência que busca verificar e mostrar qual o estado ou a situação de determinada população em relação a características previamente estabelecidas. Embora possam ser realizados censos, ou seja, investigar toda a população, a maioria das pesquisas utiliza amostras para alcançar seus objetivos. Assim, é evidente a importância de se compreender o porquê da utilização da amostragem e as diferentes formas de amostras serem selecionadas.

Há inúmeros exemplos de circunstâncias em nosso dia a dia nas quais utilizamos a amostragem. Quando uma dona de casa quer saber se colocou a quantidade adequada de sal na sopa, ela prova uma colherada e tira suas conclusões a partir apenas dessa pequena amostra para saber se a sopa toda está salgada a gosto. Da mesma forma, quando se necessita fazer um exame de sangue, obviamente não será possível retirar todo o sangue de uma pessoa. Uma pequena amostra colhida em um tubo de ensaio é o suficiente para qualquer laboratório realizar os testes. Amostras de solo e de água, amostra de produtos e pesquisas de opinião, entre outras são apenas algumas das situações em que se usa a amostragem e que encontramos em nosso cotidiano.

Porém, pesquisas com alunos de diversos níveis de escolaridade (Rubin, Bruce e Tenny (1990); Estevam (2010); Garfield (2003); Innabi (2006)) apontaram dificuldades dos mesmos em entender conceitos fundamentais como população e amostra e a relação existente entre esses. Mais especificamente, o conceito de variabilidade, representatividade, seleção, tamanho e aleatoriedade são aspectos nos quais os estudantes demonstram limitações em apreender sua finalidade e utilização.

Embora os resultados dessas investigações tenham apresentado tais dificuldades, outros (Ben-Zvi, Makar, Bakker e Aridor, 2011; Pfannkuch, 2008; Watson e Kelly, 2002) destacaram a capacidade dos estudantes de desenvolver essas habilidades quando estimulados.

Ainda que haja variedade de conceitos relacionados à amostragem abordados nas pesquisas anteriores, esses foram analisados, em sua maioria,

isoladamente ou de maneira parcial. Por exemplo, nos estudos de Estevam e Fürkotter (2010), os autores investigam a ideia de aleatoriedade juntamente com a definição de amostra aleatória, nas pesquisas de Watson (2002) ela relaciona o conceito de amostra com sua finalidade.

Diante disso, essa pesquisa teve o intuito de unir diferentes aspectos e conceitos ligados à amostragem a fim de investigar o que alunos da Educação Básica compreendem acerca desses conteúdos.

Assim, o objetivo principal deste estudo foi identificar o que estudantes do 5º e 9º ano do Ensino Fundamental compreendem sobre amostra e população. Mais especificamente buscou-se:

- Investigar se os estudantes sabem conceituar amostra;
- Analisar se os estudantes ao opinarem sobre a representatividade de uma amostra consideram o seu tamanho, seleção e variabilidade;
- Analisar se os estudantes compreendem que população é o universo a ser avaliado, envolvendo tanto pessoas como objetos inanimados;
- Verificar se os estudantes realizam inferências informais a partir de uma amostra dada;
- Identificar se há diferenças entre a compreensão de estudantes de diferentes níveis de escolaridade sobre amostragem.

Para isso, elaborou-se um instrumento contendo treze questões, as quais exploravam conceitos ligados à amostragem: definição, exemplo, finalidade, seleção, tamanho e representatividade da amostra; definição de população; conceito de aleatoriedade, amostra aleatória e sua utilização; realização de inferências informais a partir de uma amostra. A entrevista semi-estruturada foi aplicada a 40 (quarenta) estudantes do 5º (9 e 10 anos) e 9º (13 e 14 anos) anos do Ensino Fundamental, sendo 20 (vinte) de cada ano.

Para analisar as compreensões apresentadas pelos sujeitos, categorizamos as respostas dadas pelos estudantes em cada uma das questões em quatro categorias, as quais consideravam a adequação da resposta ao conceito investigado e também ao analisado em estudos anteriores. Após a categorização das respostas, comparou-se o desempenho dos sujeitos de níveis de escolaridade diferentes a fim de observar se haviam diferença significativa entre os mesmos. Além disso, realizou-se o escalonamento multidimensional (Multidimensional Scaling - MDS), com o

algoritmo ALSCAL (AlgorithmicScaling), correlacionando o comportamento dos estudantes dos anos investigados em função das variáveis analisadas.

No que se refere ao conceito de amostra, a maior parte dos estudantes não respondeu ou respondeu incorretamente as duas perguntas que solicitavam a sua definição (33 sujeitos em ambas). Entretanto, houve diferença estatística significativa entre os anos na primeira questão que abordava tal conceito de forma direta e sem apresentar nenhum contexto. Esses dados sugerem que o mesmo conceito deve ser trabalhado de diferentes formas, buscando uma maior adequação ao nível de escolaridade. Além disso, ratificam a dificuldade em compreender o conceito de amostra por estudantes do Ensino Fundamental e Médio apontadas em outros estudos, como os de Rubin, Bruce e Tenny (1990); Estevam (2010) e Garfield (2003).

Ao serem solicitados a dar um exemplo de amostra, novamente, a maioria dos participantes (30 sujeitos) não exemplificou ou apresentou exemplos incorretos. Mais alunos do 9º ano responderam de forma adequada, essa diferença entre os anos pesquisados foi significativa. Nos estudos de Rubin, Bruce e Tenny (1990) os autores perceberam que várias das respostas dadas pelos estudantes eram baseadas em suas experiências pessoais. Por esse motivo, acreditávamos que essa questão seria razoavelmente fácil por encontrarmos circunstâncias cotidianas nas quais se utilizam amostra, tais como: amostra grátis, amostra de sangue ou de saliva para teste de DNA. Contudo, observou-se que os participantes ou não conseguiram relacionar esses exemplos com a questão proposta ou essas situações não fazem parte do dia a dia dos mesmos.

Percebeu-se que os participantes do 9º ano demonstraram uma maior facilidade em exemplificar e estruturar uma definição mais adequada para o conceito de amostra. Contudo, a presença de respostas apropriadas entre os alunos do 5º ano indica que há possibilidade de aprendizagem desses conceitos por alunos nessa faixa etária. Aqueles que conseguiram conceituar parcialmente correto ou corretamente evidenciaram compreender a amostra como sendo parte de um todo, confirmando nossa hipótese de que são capazes de desenvolver habilidades necessárias para aquisição de conceitos ligados a amostragem.

Quanto à finalidade de utilização de amostras, é interessante destacar que, embora grande parte dos participantes não tenha conseguido conceituar e exemplificar de maneira correta, quando expostos a uma situação de pesquisa com

uso de amostras e solicitados a explicar o porquê de utilizá-las, quase metade dos estudantes (16 sujeitos) responderam de forma mais adequada. Um quantitativo razoável se comparado às outras questões. Entre esses participantes, 8 (oito) eram alunos do 5º ano, os quais justificaram suas respostas baseando-se na praticidade de se utilizar amostras, expressando possuírem competências para ampliar seu conhecimento acerca do propósito para uso de amostras. Os resultados alcançados nessa questão ressaltam a importância do trabalho com diferentes contextos envolvendo o conceito de amostra a fim de possibilitar a aprendizagem. Essa relevância foi evidenciada nas pesquisas de Watson e Kelly (2002) ao realizarem uma intervenção com alunos entre 8 e 9 anos de idade, os quais foram capazes de dar exemplos de situações que se utiliza amostra e levantar o porquê da utilização destas.

As questões que envolviam o conceito de população apresentaram resultados bem diferentes do que era esperado, talvez pela complexidade da segunda atividade. Na primeira questão, que trazia a ideia de população enquanto grupo de pessoas, nota-se que mais da metade dos participantes (25 sujeitos) responde de forma adequada, havendo diferença estatística entre os anos de escolaridade. Porém, na questão que abordava população como conjunto de objetos ocorre o inverso, ou seja, a maioria responde incorretamente (35 sujeitos), sem diferença estatística significativa entre os anos investigados. Assim, é importante que se discuta sobre essas diferenças do desempenho dos estudantes em atividades que envolviam o mesmo conceito. Como afirmam Ben-Zvi, Makar, Bakker e Aridor (2011) a variedade de situações estimula os alunos a pensarem sobre as relações população-amostra.

Ainda no que se refere à ideia de população como grupo de objetos a ser investigado, nota-se que dos 5 (cinco) estudantes que apresentaram respostas mais adequadas, 3 (três) eram do 5º ano. Esses alunos indicaram perceber que população nem sempre são pessoas, mas sim o universo a ser pesquisado. Essa compreensão mostra que, mesmo conceitos mais complexos, podem ser trabalhados de forma mais simples nos anos iniciais de escolarização.

Em relação à seleção da amostra, duas questões foram propostas especificamente para perceber se os estudantes eram capazes de elencar critérios importantes para a escolha de uma amostra. Na primeira atividade os alunos eram convidados a propor uma estratégia para a seleção de uma amostra, de forma que

esta fosse a mais representativa possível da população analisada. O desempenho da maioria dos participantes, mais uma vez, foi abaixo do ideal. Ao sugerir a forma de seleção da amostra esses estudantes apresentaram respostas semelhantes às encontradas no estudo de Rubin, Bruce e Tenny (1990) com estudantes do Ensino Médio, no qual os participantes demonstraram modelos inconsistentes em relação à seleção da amostra para que esta represente a população desejada, pois suas respostas eram embasadas em intuições pessoais.

A segunda questão acerca da seleção da amostra apresentava uma situação hipotética, na qual eram listadas cinco opções de amostra e os participantes eram solicitados a apontar qual delas seria a mais apropriada, justificando a sua escolha. Nessa proposição estava sendo observado se os estudantes eram tendenciosos ao eleger a amostra mais representativa. Dos 40 (quarenta) participantes, 31 (trinta e um) apresentaram respostas inapropriadas. Esse resultado é semelhante ao alcançado no estudo de Innabi (2006) com estudantes do ensino secundário, no qual os participantes eram questionados sobre a validade de uma amostra e ao justificarem suas respostas não levaram em consideração se a amostra continha características da população observada ou se apresentava viés de seleção.

Nessa mesma situação analisamos as respostas considerando, agora, o tamanho em relação à seleção da amostra. Dessa vez, os resultados alcançados, demonstram um desempenho mais adequado dos participantes (23 sujeitos tiveram respostas classificadas como parcialmente corretas e 1 como correta). Apesar disso, a maioria das justificativas dadas apenas levava em conta a amostra maior, sem relacionar a população com o número de sujeitos da amostra e a importância deste para a sua representatividade. Essa dificuldade também pode ser percebida nos estudos de Rubin, Bruce e Tenny (1990), Innabi (2006) e Watson e Kelly (2002), nos quais as respostas dos estudantes sobre a validade da amostra ora eram embasadas em suas experiências pessoais, ora em suas intuições sobre representatividade da amostra, não sendo o tamanho da amostra relacionado a esses conceitos.

Estabelecer critérios para a seleção de uma amostra a fim de que esta possa refletir as características da população de interesse é uma tarefa complicada para estudantes de diferentes idades. No entanto, propor atividades nas quais os alunos possam aplicar seus conhecimentos sobre amostragem mostrou-se uma alternativa para desenvolver conceitos e resolver situações-problema mais complexas. A partir

dos resultados alcançados no que diz respeito ao tamanho da amostra ponderamos se este pode ser um aspecto utilizado como ponto de partida para explicar o cuidado necessário para seleção de amostras. Pfannkuch (2008) observou isso em suas pesquisas ao utilizar amostras com tamanhos diferentes. Percebeu que os alunos passaram a compreender algumas noções sobre variabilidade da amostra, ligando esta à população, utilizando a linguagem associada à Estatística ou embasando suas respostas a partir das imagens e dados fornecidos.

O conceito de aleatoriedade é essencial quando se fala em amostragem. Para analisá-lo, foi solicitado que os alunos definissem aleatório a partir de uma dada situação. Mais da metade dos estudantes respondeu de forma inadequada (23 sujeitos). Vale salientar que, embora a ideia de aleatoriedade esteja presente em várias circunstâncias do nosso cotidiano, a palavra *aleatório*, bem como sua definição não é comum no vocabulário informal. Por isso, o quantitativo de 17 (dezessete) estudantes que responderam de forma parcialmente correta deve ser considerado razoável e motivador. Esse resultado nos faz refletir novamente sobre a importância de abordar tais conceitos nos anos iniciais e proporcionar atividades com diferentes contextos envolvendo aspectos da amostragem. A questão apresentada trazia uma situação na qual estava implícita a ideia de um sorteio, cuja principal propriedade é a aleatoriedade. É possível que isso tenha auxiliado no raciocínio dos alunos que responderam parcialmente correto.

Ainda no âmbito da aleatoriedade, buscou-se saber o que os estudantes consideram como sendo uma Amostra Aleatória (AA). Optou-se por esse tipo de amostra por ser comumente utilizado em pesquisas probabilísticas. Novamente, a maior parte dos entrevistados não define ou define incorretamente (30 sujeitos). Acredita-se que essa atuação seja decorrente principalmente da não compreensão de amostra, como já foi exposto nas três primeiras questões.

Após serem questionados sobre o que seria uma AA, foi pedido que os participantes propusessem uma situação na qual fosse usada uma Amostra Aleatória. Como esperado em função da questão anterior, quase a totalidade dos sujeitos (35) deu exemplos inapropriados.

Os resultados alcançados nas questões que envolviam a aleatoriedade e o princípio da amostragem aleatória ressaltam as dificuldades dos estudantes em compreendê-los destacadas por Estevam e Fürkotter (2010), bem como a necessidade de trabalhá-los na educação básica.

Analisando a capacidade dos estudantes de realizarem inferências informais a partir de uma amostra dada, foram apresentados os dados coletados em uma pesquisa hipotética, através dos quais os participantes deveriam realizar inferências para toda a população. Entre os sujeitos deste estudo, 25 (vinte e cinco) responderam de forma parcialmente correta, apresentando boas intuições baseadas em suas experiências. Esses dados também foram encontrados por Watson e Kelly (2002), nas quais as crianças não conseguiram elaborar inferências mais complexas, mas foram capazes de relacionar os dados da amostra, estruturando conclusões acerca da população.

O último aspecto observado diz respeito à representatividade da amostra. Vale lembrar que, o mesmo foi avaliado indiretamente em outras questões, mas sendo a representatividade o principal propósito da amostragem foi abordada de forma objetiva em uma questão específica. Mais da metade dos participantes responderam incorretamente (22 sujeitos). Alguns estudantes escolheram a amostra, sem justificar o porquê da opção; outros apresentaram justificativas não condizentes com a representatividade da amostra e houve aqueles que utilizaram aspectos relacionados à amostragem, mas de forma inadequada.

Os resultados alcançados aqui confirmaram outras pesquisas (Gil e Ben-Zvi, 2010; Rubin, Bruce e Tenny, 1990; Garfield, 2003; Innabi, 2006) que elencaram dificuldades dos estudantes em perceber a implicações da representatividade da amostra e relacioná-la a outros fatores como tamanho e variabilidade. Ao mesmo tempo, as respostas mais adequadas dão indícios de que, se instigados, estudantes de diferentes idades são capazes de desenvolver habilidades necessárias para elencar critérios necessários para a validade de uma amostra. Nas justificativas parcialmente corretas ou corretas estava implícita a relevância da variabilidade da amostra. Isso demonstra que, embora os alunos não tenham conseguido elaborar uma resposta mais formal e completa, eles percebem que para que uma amostra seja válida ela deve ter algumas características específicas da população.

A partir da análise multidimensional observou-se que as questões sobre aleatoriedade, representatividade e realização de inferências, as quais se relacionam à seleção da amostra, apresentaram correlação, sugerindo que esses conceitos necessitam de habilidades semelhantes para sua compreensão. Essa análise é relevante, pois nos permite pensar em possíveis maneiras de se trabalhar

com amostragem desde os anos iniciais, associando diferentes conceitos a fim de facilitar e contribuir para a aprendizagem dos estudantes.

Finalmente, comparando a compreensão de estudantes dos diferentes níveis de escolaridade sobre amostragem vemos que ao analisarmos os dados de forma descritiva, na maioria das questões, não houve diferença estatística entre os anos de escolaridade investigados. Contudo, ao realizarmos uma análise global dos dados notou-se a diferença no comportamento dos estudantes entre os anos de escolaridade.

Ao examinar os resultados do 5º ano, observou-se que os estudantes apresentaram grandes dificuldades nas questões que envolviam definição de conceitos (amostra, população e amostra aleatória). Já nas questões que apresentavam uma situação problema as quais podiam ser respondidas baseada em seus conhecimentos pessoais, tais como as que envolviam seleção e representatividade da amostra, os mesmos apresentaram um desempenho melhor.

Os alunos do 9º ano também apresentaram dificuldades em definir conceitos. Entretanto, as questões que admitiam uma maior subjetividade levaram a uma separação entre os alunos: os que não sabiam e foram agrupados nas duas categorias de respostas inadequadas e os que sabiam e foram agrupados nas categorias de respostas adequadas. A distribuição também é diferente em relação aos pontos que representam as questões que envolviam definições de conceito. Essa análise nos faz pensar se a diferença no comportamento dos estudantes ocorre devido ao nível de escolaridade, às experiências pessoais ou ambos os casos.

Os resultados deste estudo, bem como o conhecimento da relevância dos conceitos aqui explorados e da necessidade dos mesmos serem trabalhados desde os anos iniciais do Ensino Fundamental apontam algumas implicações desta pesquisa para o ensino da amostragem.

A partir dos resultados expostos, podemos afirmar que alunos desde o 5º ano de escolaridade (entre 9 e 10 anos de idade) são capazes de compreender os conceitos ligados à amostragem, visto que houveram respostas apropriadas entre os participantes. O reconhecimento desse potencial dos estudantes nos leva a argumentar que é imprescindível repensar o que a escola pode e deve propor aos alunos. Acrescido a isso, acreditamos ser fundamental uma abordagem conjunta das diversas habilidades relacionadas à amostragem, como realizado nessa pesquisa.

Sugere-se, finalmente, que sejam realizados outros estudos sobre como o conceito de amostragem pode ser compreendido de forma mais apropriada pelos alunos do Ensino Fundamental.

REFERÊNCIAS

American Statistical Association (ASA). **Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report: A Pre-K-12 Curriculum Framework**. Alexandria, 2005. Disponível em: <http://www.amstat.org/education/gaise/GAISEPreK12_Intro.pdf>. Acesso em 11 set. 2010.

BEN-ZVI, D., MAKAR, K., BAKKER, A. & ARIDOR, K. Children's emergent inferential reasoning about samples in an inquiry-based environment. In: **Seventh Congress of the European Society for Research in Mathematics Education**, Rzeszów, Poland, 9 - 13 February, 2011.

BRASIL. Secretaria de Educação Fundamental. **Parâmetros Curriculares Nacionais: Matemática (1ª a 4ª)**. Secretaria de Educação Fundamental. Brasília, 1997. 142p.

_____. Secretaria de Educação Fundamental. **Parâmetros Curriculares Nacionais : Matemática (5ª a 8ª)**. Secretaria de Educação Fundamental. Brasília, 1998. 148 p.

BUSSAB, W. O. e MORETTIN, P. A. **Estatística básica**. 5. ed. São Paulo: Saraiva, 2002.

CAMPOS, M. A. **Introdução ao tratamento da informação nos ensinos fundamental e médio**. São Paulo: Plêiade, 2005. (Notas em Matemática Aplicada 16 - SBMAC).

COCHRAN, W. G. **Técnicas de amostragem**. Rio de Janeiro: Editora Fundo de Cultura, 1965.

DANCEY, C. P. e REIDY, J. **Estatística sem Matemática para Psicologia: Usando SPSS para Windows**. 3. ed. Porto Alegre: Artmed, 2006.

DONALD, B. O. **Handbook of Statistical Tables**. Reading, Mass: Addison-Wesley, 1962.

ESTEVAM, E. J. G. **(Res)Significando a Educação Estatística no Ensino Fundamental: análise de uma sequência didática apoiada nas Tecnologias de Informação e Comunicação**. Dissertação de Mestrado, Universidade Estadual Paulista, 2010. 211 f.

ESTEVAM, E. J. G. e FÜRKOTTER, M. A variabilidade como fator (res)significante para a Educação Estatística no Ensino Fundamental. **Anais do X Encontro Nacional de Educação Matemática**. Salvador, 2010.

FARIAS, A. A., SOARES, J. F. e CÉSAR, C. C. **Introdução à Estatística**. 2. ed. Rio de Janeiro: LTC, 2003.

GAL, I. Adult's Statistical Literacy: meanings, components, responsibilities. **International Statistical Review**, v. 70, n. 1, p. 1-25, 2002.

GARFIELD, J. Assessing Statistical Reasoning. **Statistics Education Research Journal**. v.2. n. 1. p. 22-38, 2003. Disponível em: <[http://www.stat.auckland.ac.nz/~iase/serj/SERJ2\(1\).pdf](http://www.stat.auckland.ac.nz/~iase/serj/SERJ2(1).pdf)>. Acesso em: 12 set. 2010.

GIL, E., BEN-ZVI, D. **Emergence of reasoning about sampling among young students in the context of informal inferential reasoning**. *Proceedings of the eighth international Conference on Teachings Statistics*, Ljubljana, Slovenia, 2010.

HÄRDLE, W., SIMAR, L. **Applied Multivariate Statistical Analysis**. Second Edition. Ed. Springer, 2007.

INNABI, H. Factors considered by secondary students when judging the validity of a given statistical generalization. **Proceedings of the Seventh International Conference on Teachings Statistics**, Brasil, 2006. Disponível em: <http://www.stat.auckland.ac.nz/~iase/publications/17/2B1_INNA.pdf>. Acesso em: 6 set. 2010.

KINNEAR, T. C. & TAYLOR, J. R. **Marketing research: an applied approach**. Mc Graw Hill, 1979.

LOPES, C. A. E. **A probabilidade e a estatística no ensino fundamental: uma análise curricular**. Dissertação de Mestrado. Faculdade de Educação da UNICAMP, Campinas, 1998. 127 f.

_____. **O conhecimento profissional dos professores e suas relações com estatística e probabilidade na educação infantil**. Tese de Doutorado. Faculdade de Educação da UNICAMP, Campinas, 2003. 289f.

_____. O ensino da estatística e da probabilidade na educação básica e a formação dos professores. **Cadernos Cedes**, vol. 28, n. 74, Campinas, p. 57-73, jan./abr. 2008.

MATTAR, F. **Pesquisa de marketing**. Ed. Atlas, 1996.

OLIVER, Maria del Carmen. **La atención a la diversidad desde los grupamientos flexibles de alumnos**. 1998. 669 f. Tese (Doutorado em La Innovació Educativa i la Formació del Professorat). Facultat de Ciències de L'Educació. Universitat de Barcelona. Barcelona, 1998. (Capítulo 7). Disponível em: <<http://ldei.ugr.es/cddi/uploads/tesis/OliverVera1998.pdf>> Acesso em: 22 de nov. de 2012.

PFANNKUCH, M. Building sampling concepts for statistical inference: A case study. In: **Eleventh International Congress on Mathematics Education**, Monterrey, Mexico, 6 – 13 July, 2008.

PORCAR, M. L.; ESCALANTE, E. E. Límites de la explicación interna (norma de la internalidad): análisis de escalamiento multidimensional. **Revista Electrónica de Investigación y Docencia (REID)**, n. 2, Junho, p.59-77, 2009. Disponível em: <<http://www.revistareid.net/revista/n2/REID2art4.pdf>> Acesso em: 22 jan. 2011.

RUBIN, A. D., BRUCE, B. C. and TENNEY, Y. Learning About Sampling: Trouble at the Core of Statistics. **Proceedings of the Third International Conference on Teachings Statistics**. p. 314-319, New Zeland, 1990. Disponível em: <<http://www.stat.auckland.ac.nz/~iase/publications/18/BOOK1/A9-4.pdf>>. Acesso em: 7 set. 2010.

STEVENSON, W. J. **Estatística aplicada à administração**. São Paulo: Harper & Row do Brasil, 1981.

VIEIRA, S. **Estatística básica**. São Paulo: Cengage Learning, 2012.

WATSON, J. M. (2003): Statistical Literacy at the school level: What should students Know and do? In: **Bulletin of the International Statistical Institute 54th Session Proceedings**, Berlin 2003, (Volume LX, Book 2, Invited Papers, Topic 49, pp. 68-71).

_____. Creating cognitive conflict in a controlled research setting: sampling. **Proceedings of the Sixth international Conference on Teachings Statistics**, South Africa, 2002.

WATSON, J.M. and KELLY, B. A. Can grade 3 students learn about variation? **Proceedings of the Sixth international Conference on Teachings Statistics**, South Africa, 2002.

APÊNDICE A – Cartões do instrumento de coleta

1. O que você acha que é amostra?

2. Tu podes me dar um exemplo de amostra?

3. Para saber qual candidato a prefeito do Recife tem mais chance de ganhar as eleições, pesquisadores entrevistaram uma amostra de mil eleitores. O que amostra vai significar nesse caso?

4. Para saber qual o candidato a prefeito do Recife tem mais possibilidade de ganhar a eleição, pesquisadores entrevistaram uma amostra de 1000 eleitores. *Por que você acha que eles usaram uma amostra e não todos os eleitores do Recife?*

5. Nessa pesquisa sobre os candidatos a prefeito do Recife, qual seria a população analisada?

6. Se fosse realizada uma pesquisa para saber quanto tempo duram os computadores da marca X. Qual seria a população analisada nessa pesquisa?

7. Um pesquisador queria saber a merenda preferida dos alunos das escolas públicas de Recife. Como ele não tinha condições de entrevistar todos os alunos resolveu entrevistar apenas 200 alunos. Como ele poderia escolher esses alunos para ter uma ideia melhor da preferência de todos?

8. Cinco amigos queriam saber aproximadamente quantos livros as pessoas que moram no bairro deles liam por ano. Como o bairro tinha uns 10.000 moradores, não dava para entrevistar todo mundo. Cada um teve uma ideia para saber quem podiam entrevistar. Qual dessas ideias você acha que será melhor para saber o que eles querem? Por quê?

Amigo 1	100 moradores que frequentavam a biblioteca da comunidade.
Amigo 2	100 moradores do bairro.
Amigo 3	10 moradores que frequentavam a biblioteca da comunidade.
Amigo 4	10 moradores do bairro.
Amigo 5	Homens, mulheres, meninos e meninas.

9. Para definir a ordem dos alunos na fila para a merenda, a professora colocou o nome dos alunos em um saquinho e foi retirando aleatoriamente. O que significa aleatório para você?

10. Eu quero saber o que as pessoas acham do carnaval do Recife. Para selecionar as pessoas que participarão da pesquisa eu vou usar uma amostra aleatória. O que significa amostra aleatória?

11. Dê um exemplo de outra situação na qual podemos usar uma amostra aleatória.

12. Em uma escola com 150 alunos foi realizada uma pesquisa sobre o uso do celular. Foi sorteada uma amostra de 20 alunos e realizou-se uma pesquisa, obtendo os dados abaixo.

Aluno	Sexo	Pra que usa mais	Pra que usa menos
Aluno 1	F	Fotos	Ligações
Aluno 2	F	Músicas	Ligações
Aluno 3	F	Fotos	Ligações
Aluno 4	F	Músicas	Mensagem
Aluno 5	F	Músicas	Ligações
Aluno 6	F	Fotos	Ligações
Aluno 7	F	Mensagem	Fotos
Aluno 8	F	Mensagem	Mensagem
Aluno 9	F	Fotos	Ligações
Aluno 10	F	Fotos	Mensagem
Aluno 11	M	Músicas	Ligações
Aluno 12	M	Mensagem	Fotos
Aluno 13	M	Mensagem	Músicas
Aluno 14	M	Mensagem	Fotos
Aluno 15	M	Músicas	Ligações
Aluno 16	M	Mensagem	Fotos
Aluno 17	M	Fotos	Mensagem
Aluno 18	M	Músicas	Fotos
Aluno 19	M	Mensagem	Músicas
Aluno 20	M	Mensagem	Fotos

- Que tipo de informação a tabela traz?
- Que conclusões você pode tirar dessa tabela?
- O que você pode dizer sobre o uso do celular nessa escola?

13. E se, nesta pesquisa sobre o uso do celular na escola, forem sorteados outros 20 alunos de uma mesma turma? Iria ser melhor ou pior para representar a escola?

ANEXO A – Tabela de números aleatórios

3690	2492	7171	7720	6509	7549	2330	5733	4730
0813	6790	6858	1489	2669	3743	1901	4971	8280
6477	5289	4092	4223	6454	7632	7577	2816	9202
0772	2160	7236	0812	4195	5589	0830	8261	9232
5692	9870	3583	8997	1533	6466	8830	7271	3809
2080	3828	7880	0586	8482	7811	6807	3309	2729
1039	3382	7600	1077	4455	8806	1822	1669	7501
7227	0104	4141	1521	9104	5563	1392	8238	4882
8506	6348	4612	8252	1062	1757	0964	2983	2244
5086	0303	7423	3298	3979	2831	2257	1508	7642
0092	1629	0377	3590	2209	4839	6332	1490	3092
0935	5565	2315	8030	7651	5189	0075	9353	1921
2605	3973	8204	4143	2677	0034	8601	3340	8383
7277	9889	0390	5579	4620	5650	0210	2082	4664
5484	3900	3485	0741	9069	5920	4326	7704	6525
6905	7127	5933	1137	7583	6450	5658	7678	3444
8387	5323	3753	1859	6043	0294	5110	6340	9137
4094	4957	0163	9717	4118	4276	9465	8820	4127
4951	3781	5101	1815	7068	6379	7252	1086	8919
9047	0199	5068	7447	1664	9278	1708	3625	2864
7274	9512	0074	6677	8676	0222	3335	1976	1645
9192	4011	0255	5458	6942	8043	6201	1587	0972
0554	1690	6333	1931	9433	2661	8690	2313	6999
9231	5627	1815	7171	8036	1832	2031	6298	6073
3995	9677	7765	3194	3222	4191	2734	4469	8617
2402	6250	9362	7373	4757	1716	1942	0417	5921
5295	7385	5474	2123	7035	9983	5192	1840	6176
5177	1191	2106	3351	5057	0967	4538	1246	3374
7315	3365	7203	1231	0546	6612	1038	1425	2709
5775	7517	8974	3961	2183	5295	3096	8536	9442
5500	2276	6307	2346	1285	7000	5306	0414	3383
3251	8902	8843	2112	8567	8131	8116	5270	5994
4675	1435	2192	0874	2897	0262	5092	5541	4014
3543	6130	4247	4859	2660	7852	9096	0578	0097
3521	8772	6612	0721	3899	2999	1263	7017	8057
5573	9396	3464	1702	9204	3389	5678	2589	0288
7478	7569	7551	3380	2152	5411	2647	7242	2800
3339	2854	9691	9562	3252	9848	6030	8472	2266
5505	8474	3167	8552	5409	1556	4247	4652	2953
6381	2086	5457	7703	2758	2963	8167	6712	9820

Fonte: Donald (1962)