



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE EDUCAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM EDUCAÇÃO MATEMÁTICA E
TECNOLÓGICA
CURSO DE DOUTORADO

ERICA MICHELLE SILVA CAVALCANTI

**APRENDIZAGEM DE ESTUDANTES DO ENSINO FUNDAMENTAL SOBRE
LEVANTAMENTO DE HIPÓTESES, ANÁLISE DE DADOS E CONCLUSÕES A
PARTIR DE DADOS ESTATÍSTICOS**

Recife
2019

ERICA MICHELLE SILVA CAVALCANTI

**APRENDIZAGEM DE ESTUDANTES DO ENSINO FUNDAMENTAL SOBRE
LEVANTAMENTO DE HIPÓTESES, ANÁLISE DE DADOS E CONCLUSÕES A
PARTIR DE DADOS ESTATÍSTICOS**

Tese de doutorado apresentada ao Programa de Pós-Graduação em Educação Matemática e Tecnológica (EDUMATEC) da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutora em Educação Matemática e Tecnológica.

Área de concentração: Educação Estatística

Orientadora: Dr.^a Gilda Lisbôa Guimarães

Recife

2019

Catálogo na fonte
Bibliotecária Amanda Nascimento, CRB-4/1806

C376a Cavalcanti, Erica Michelle Silva
Aprendizagem de estudantes do ensino fundamental sobre levantamento de hipóteses, análise de dados e conclusões a partir de dados estatísticos / Erica Michelle Silva Cavalcanti. – Recife, 2019.
185 f. : il.

Orientadora: Gilda Lisbôa Guimarães
Tese (Doutorado) - Universidade Federal de Pernambuco, CE. Programa de Pós-graduação em Educação Matemática e Tecnológica, 2019.

Inclui Referências.

1. Educação - Estatística. 2. Matemática (Ensino fundamental). 3. Matemática – Estudo e ensino - Pesquisa. 4. UFPE - Pós-graduação. I. Guimarães, Gilda Lisbôa (Orientadora). II. Título.

370.21 (22. ed.) UFPE (CE2019-043)

ERICA MICHELLE SILVA CAVALCANTI

**APRENDIZAGEM DE ESTUDANTES DO ENSINO FUNDAMENTAL SOBRE
LEVANTAMENTO DE HIPÓTESES, ANÁLISE DE DADOS E CONCLUSÕES A
PARTIR DE DADOS ESTATÍSTICOS**

Aprovada em: 09 / 04 / 2019.

COMISSÃO EXAMINADORA:

Presidente e orientadora
Prof.^a Dr.^a Gilda Lisbôa Guimarães
Universidade Federal de Pernambuco

Examinadora externa
Prof.^a Dr.^a Ana Cláudia C. Batalha Henriques
Universidade de Lisboa

Examinadora externa
Prof.^a Dr.^a Mônica Maria Lins
Universidade Federal Rural de Pernambuco

Examinadora externa
Prof.^a Dr.^a Alina Galvão Spinillo
Universidade Federal de Pernambuco

Examinador interno
Prof. Dr. Carlos Eduardo Monteiro
Universidade Federal de Pernambuco

À minha amiga Gilda, que me apresentou à Educação Matemática, ainda na graduação, incentivando-me à pesquisa científica como um meio de promoção do professor-pesquisador e da justiça social,

Dedico

AGRADECIMENTOS

À minha mãe, Arlene, que a seu modo sempre me encorajou a fazer o que eu acredito, sendo um exemplo de lealdade e atitude.

Ao meu irmão, Bruno, cujo apoio incondicional à família me tranquilizou tantas vezes, por saber que nas minhas ausências ele estaria presente fazendo o seu melhor.

Ao meu sobrinho, Bruninho, que representa a leveza quando tudo parece pesar mais do que o esperado.

Àqueles familiares que mesmo distantes estão sempre presentes em pensamento, oferecendo apoio nos momentos necessários e torcendo para o progresso um dos outros.

Aos amigos que proporcionaram momentos de descontração e aconselhamentos nas conversas presenciais e virtuais.

À minha orientadora, amiga e exemplo de profissional, Gilda, sempre incentivadora e ativa em todos os âmbitos de sua vida, demonstrando a todos que se pode dedicar com afinco à vida acadêmica, sem abdicar dos prazeres de se estar com amigos e sem prescindir da generosidade em ajudar os demais.

Ao Grupo de Educação Estatística no Ensino Fundamental (GREF), cujo interesse pela área do saber e empenho em ajudar aos demais é bastante agregador.

Ao grupo que faz a linha de Processos de Ensino Aprendizagem em Educação Matemática e Científica. A partir da disciplina de Seminários Avançados a tese construída contou com grandes discussões e amadurecimento.

À Rute Borba, que não apenas na disciplina de Seminários Avançados, como em outras oportunidades da vida acadêmica ajudou muitíssimo na definição metodológica e analítica do estudo.

À secretaria e coordenação da Pós-Graduação em Educação Matemática e Tecnológica, sempre prestativos para com os discentes.

Aos professores Ana Henriques, Mônica Lins, Verônica Gitirana e Carlos Eduardo Monteiro, que constituíram a banca de qualificação do doutorado. Ao lerem o estudo em processo de construção, enriqueceram-no como apenas pesquisadores experientes poderiam fazê-lo.

Aos profissionais das redes públicas de ensino que atuou, os quais compreenderam e viabilizaram o processo de afastamento para os estudos da tese.

Aos professores das turmas investigadas durante a pesquisa, que disponibilizaram tempo de interação com os estudantes para a coleta de dados.

Aos estudantes cujas compreensões e aprendizagens foram investigadas, a recepção deles tornou possível todo o trabalho de pesquisa.

Quando o homem compreende a sua realidade, pode levantar hipóteses sobre o desafio dessa realidade e procurar soluções. Assim, pode transformá-la e o seu trabalho pode criar um mundo próprio: seu eu e as suas circunstâncias. (FREIRE, 1979, p. 30)

RESUMO

A tese contou com o objetivo de analisar aprendizagens apresentadas por estudantes do Ensino Fundamental relacionadas ao levantamento de hipóteses, à análise de dados e às conclusões a partir dos dados, enquanto fases do ciclo investigativo. Objetivo que surgiu da problemática: De que modo o levantamento de hipóteses, análise de dados e habilidades relacionadas a conclusões por estudantes do Ensino Fundamental corroboram para a aprendizagem de justificativas baseadas em evidências? Para tanto, dois estudos foram elaborados. O primeiro investigou os conhecimentos apresentados por estudantes do 5º e do 7º ano de escolas públicas da Região Metropolitana de Recife, diante de habilidades relacionadas ao levantamento de hipóteses, à interpretação de dados reais univariados e bivariados, ao confronto entre hipótese e dados, à avaliação de conclusões e ao uso de linguagem probabilística em predições. Percebeu-se que os estudantes desde o 5º ano já são capazes de levantar hipóteses para questões de pesquisa propostas. Ao interpretar constatou-se que o tipo de distribuição dos dados (uni e bivariada) interferiu bastante nas compreensões dos estudantes, tornando a análise bivariada mais complexa, o que também foi verificado no confronto entre hipótese e dados reais. Quanto à avaliação de conclusões e uso de linguagem probabilística ao prever geraram bastante dificuldades para os estudantes de ambos os anos em todas as atividades. O segundo estudo teve um carácter intervencionista experimental, o qual contou com pré-teste, seguido de três encontros de intervenção e pós-teste, abordando-se as mesmas habilidades do estudo anterior e os mesmos anos de escolaridade. Constatou-se melhoras significativas no desempenho dos estudantes de ambos os anos de escolaridade. Isso considerando-se a média de acerto nos testes (inicial e final) dos sujeitos que passaram pelo processo de intervenção de ensino. Analisando-se a justificativa baseada em evidências dos dados, observou-se que os estudantes do 5º ano partiram de percentuais aquém dos estudantes do 7º ano na maioria das questões. Entretanto, conseguiram avanços significativos tanto em habilidades exploradas a partir de dados univariados quanto a partir de dados bivariados. Já os estudantes do 7º ano melhoraram de modo significativo em suas justificativas apenas na habilidade de interpretação, quando os dados foram bivariados. Salienta-se que a produção escrita pelos estudantes do 7º ano contou com maior resistência, o que interfere na análise da expressão de ideias.

As aulas ministradas apresentaram momentos de intervenção coletiva e outros momentos de interação entre pares com formação de pequenos grupos, buscou-se nesses momentos a exploração da Zona de Desenvolvimento Proximal (ZDP) dos estudantes. Conclui-se a partir dos resultados alcançados que o trabalho com as habilidades investigadas na tese deve ser iniciado logo nos primeiros anos de escolaridade, a fim de se promover o Letramento Estatístico da população, enquanto atitude crítica e investigativa.

Palavras-chave: Educação estatística. Ensino Fundamental. Pesquisa.

ABSTRACT

The aim of this thesis was to analyze learning presented by elementary school students related to the hypothesis analysis, the data analysis and the conclusions from the data, as phases of the research cycle. Objective that emerged from the problematic: How do hypothesis-gathering, data analysis, and conclusion-related skills by elementary school students corroborate the learning of evidence-based justifications? For that, two studies were elaborated. The first one investigated the knowledge presented by 5th and 7th grade students from public schools in the Metropolitan Region of Recife, in relation to skills related to the hypothesis survey, the interpretation of univariate and bivariate real data, the confrontation between hypothesis and data, the evaluation conclusions and the use of probabilistic language in predictions. It was noticed that the students from the 5th grade are already able to raise hypotheses for proposed research questions. When interpreted, it was found that the type of data distribution (univ and bivariate) interfered greatly in the students' understandings, making bivariate analysis more complex, which was also verified in the confrontation between hypothesis and real data. The questions related to the evaluation of conclusions and the use of probabilistic language in predicting generated enough difficulties for students of both years in all activities. The second study had an experimental interventional character, which had a pre-test, followed by three intervention and post-test meetings, addressing the same abilities of the previous study and the same years of schooling. Significant improvements were observed in the performance of students from both years of schooling. This is based on the average accuracy in the tests (initial and final) of the subjects who went through the teaching intervention process. Analyzing the evidence-based justification of the data, it was observed that the students of the 5th grade started with percentages shorter than the 7th year students in most of the questions. However, they have achieved significant advances in both exploited abilities from univariate data and from bivariate data. Already the 7th grade students improved significantly in their justifications only in the ability of interpretation, when the data were bivariate. It should be noted that the written production of the 7th grade students had more resistance, which interferes with the analysis of the expression of ideas. The lecture classes presented moments of collective intervention and other moments of

interaction between peers with small groups, and the exploration of the Proximal Development Zone (ZPZ) of the students was sought in those moments. It is concluded from the results achieved that the work with the skills investigated in the thesis should be started as early as the first years of schooling, in order to promote the Statistical Literacy of the population, as a critical and investigative attitude.

Keywords: Statistical education. Elementary School. Research.

LISTA DE FIGURAS

Figura 2.1 –	Ciclo investigativo (PPDAC) de Wild e Pfannkuch (1999)	30
Figura 2.2 –	Fases do ciclo investigativo de Silva (2013, p. 48)	32
Figura 2.3 –	Diagramas de dispersão (TRIOLA, 2013, p.417)	37
Figura 2.4 –	Atividade utilizada por Cavalcanti (2011, p.48) e adaptada de estudo de Watson e Kelly (2002)	42
Figura 2.5 –	Exemplo de intervenção de professora para realização de conclusão, em Caseiro et al (2014, p.250)	44
Figura 2.6 –	Exemplo de questão (múltipla escolha) de análise de dados em Jacobbe et al (2014, p.5)	45
Figura 2.7 –	Modelos de <i>loopy airplanes</i> (MAKAR, 2013)	47
Figura 2.8 –	Atividade de associação entre variáveis, Nunes (2015)	50
Figura 4.1 –	Estudante (S42, 7º ano) que levanta hipótese na Atividade 1.....	83
Figura 4.2 –	Estudante (S7, 5º ano) que não cita momento algum, refere-se ao uso	83
Figura 4.3 –	Estudante (S35, 7º ano) que interpreta os dados representados no gráfico	84
Figura 4.4 –	Exemplo de resposta que confronta (Q3) hipótese e dados (S16, 5º ano)	85
Figura 4.5 –	Exemplo de resposta que confronta (Q3) hipótese e dados (S33, 7º ano)	85
Figura 4.6 –	Estudante (S28, 7º ano) que retoma hipótese levantada (Q3), mas não interpretou quando solicitado (Q2)	87
Figura 4.7 –	Exemplo de resposta adequada, que justifica pelos dados (S24, 7º ano)	88
Figura 4.8 –	Exemplo de resposta que utiliza linguagem probabilística (S26, 7º ano)	88
Figura 4.9 –	Exemplo de resposta que interpreta os dados (S6, 5º ano)	89
Figura 4.10 –	Exemplo de resposta que retoma hipótese após os dados (S8, 5º ano)	90
Figura 4.11 –	Estudante (S47, 7º ano) que retoma hipótese levantada (Q3),	

	mas não interpretou quando solicitado (Q2)	91
Figura 4.12 –	Exemplo de resposta adequada (Q4), que justifica pelos dados (S16, 5º ano)	92
Figura 4.13 –	Exemplo de resposta (S43, 7º ano) que levanta hipótese na Atividade 3	93
Figura 4.14 –	Exemplo de resposta (S6, 5º ano) que não considera as duas variáveis na hipótese da Atividade 3	93
Figura 4.15 –	Exemplo de resposta (S07, 5º ano) que não faz interpretação dos dados	94
Figura 4.16 –	Exemplo de resposta (S45, 7º ano) que mostra interpretação dos dados	94
Figura 4.17 –	Exemplo de resposta (S35, 7º ano) que sugere percepção de ponto máximo no gráfico	95
Figura 4.18 –	Exemplo de resposta (S10, 5º ano) que retoma hipótese após os dados	95
Figura 4.19 –	Exemplo de resposta (S29, 7º ano) que não relaciona hipótese com dados	95
Figura 4.20 –	Exemplo de resposta (S42, 7º ano) que avalia conclusão a partir dos dados	96
Figura 4.21 –	Exemplo de resposta (S4, 5º ano) que não avalia conclusão a partir dos dados	96
Figura 4.22 –	Exemplo de resposta (S47, 7º ano) que relaciona variáveis, com discordância	97
Figura 4.23 –	Exemplo de resposta (S16, 5º ano) que registra considerações apenas sobre a variável água	98
Figura 4.24 –	Exemplo de resposta (S37, 7º ano) que faz uma interpretação adequada dos dados	98
Figura 4.25 –	Exemplo de resposta (S15, 5º ano) que não faz uma interpretação adequada dos dados	99
Figura 4.26 –	Exemplo de resposta (S47, 7º ano) que retoma hipótese após análise dos dados	99
Figura 4.27 –	Exemplo de resposta (S29, 7º ano) que retoma hipótese e interpreta os dados	100

Figura 4.28 – Extrato de protocolo (S11, 5º ano), em que o estudante retoma hipótese, mas errou ao interpretar os dados	102
Figura 5.1 – Fases do ciclo investigativo com adaptação (SILVA, 2013, p. 48)	111
Figura 5.2 – Sujeito (S06) que não levanta hipótese (Q1) no pré-teste	115
Figura 5.3 – Sujeito (S11) que levanta hipótese (Q1) no pós-teste	115
Figura 5.4 – Sujeito (S21) que não levanta hipótese (CQ1) no pré-teste	116
Figura 5.5 – Sujeito (S03) que levanta hipótese (CQ1) no pré-teste	116
Figura 5.6 – Sujeito (S03) que não interpreta (Q2) no pré-teste	118
Figura 5.7 – Sujeito (S03) que interpreta (Q2) e argumenta no pós-teste	119
Figura 5.8 – Sujeito (S29) que não interpreta (CQ2) no pré-teste	120
Figura 5.9 – Sujeito (S29) que interpreta (CQ2) e argumenta no pós-teste	120
Figura 5.10 – Sujeito (S17) que não confronta a hipótese com os dados (Q3) no pré-teste	122
Figura 5.11 – Sujeito (S11) que confronta a hipótese com os dados (CQ3) no pós-teste	123
Figura 5.12 – Sujeito (S21) não avalia conclusão a partir dos dados no pré-teste	124
Figura 5.13 – Sujeito (S21) avalia conclusão a partir dos dados no pós-teste	124
Figura 5.14 – Sujeito (S08) que avalia conclusão (CQ4) e argumenta no pré-teste	125
Figura 5.15 – Sujeito (S08) que avalia conclusão (CQ4) no pós-teste	125
Figura 5.16 – Sujeito (S25) que não utiliza linguagem probabilística ao prever no pré-teste	126
Figura 5.17 – Sujeito (S29) que utiliza linguagem probabilística ao prever no pós-teste	127
Figura 5.18 – Sujeito (S83) que não levanta hipótese (Q1) no pré-teste	130
Figura 5.19 – Sujeito (S54) que levanta hipótese (Q1) no pré-teste	130
Figura 5.20 – Sujeito (S68) que não levanta hipótese (CQ1) no pós-teste	131
Figura 5.21 – Sujeito (S88) que levanta hipótese (CQ1) no pós-teste	131

Figura 5.22 – Sujeito (S80) que não justifica ao interpretar (Q2) no pré-teste	132
Figura 5.23 – Sujeito (S80) que justifica ao interpretar (Q2) no pós-teste	133
Figura 5.24 – Sujeito (S67) que não interpreta no pré-teste	134
Figura 5.25 – Sujeito (S67) que interpreta e justifica no pós-teste	134
Figura 5.26 – Sujeito (S56) que realiza confronto entre hipótese e dados (Q3) no pós-teste	136
Figura 5.27 – Sujeito (S69) que não faz confronto entre hipótese e dados (CQ3) no pós-teste	137
Figura 5.28 – Sujeito (S67) que não avalia conclusão (Q4) no pré-teste	139
Figura 5.29 – Sujeito (S67) que avalia conclusão (Q4) no pós-teste	139
Figura 5.30 – Sujeito (S100) que não avalia conclusão (CQ4) no pré-teste	139
Figura 5.31 – Sujeito (S100) que avalia conclusão (CQ4) no pós-teste	140
Figura 5.32 – Sujeito (S85) que não utiliza linguagem probabilística no pré-teste	141
Figura 5.33 – Sujeito (S85) que utiliza linguagem probabilística no pós-teste	141
Figura 5.34 – Grupo (Turma 7A) que discorda no levantamento de hipótese, Atividade 3.2	161
Figura 5.35 – Grupo (Turma 7B) que concorda no levantamento de hipótese, Atividade 2.2	162
Figura 5.36 – Grupo (Turma 7A) que realiza confronto entre hipótese e dados, Atividade 2.2	164
Figura 5.37 – Grupo (Turma 5A) que faz tomada de decisão, Atividade 2.2	164
Figura 5.38 – Grupo que avalia conclusão (Turma 7A), Atividade 3.2	165
Figura 5.39 – Grupo (Turma 5B) que reanalisa os dados, Atividade 3.2	165
Figura 5.40 – Grupo (Turma 5A) que lança nova hipótese, Atividade 3.2	166

LISTA DE QUADROS

Quadro 2.1 – Modelo de Letramento Estatístico	25
Quadro 3.1 – Tipos de questões propostas no Estudo 1	56
Quadro 3.2 – Objetivos delineados para três encontros de intervenção	71

LISTA DE GRÁFICOS

Gráfico 4.1 –	Análise da relação entre habilidades de interpretar (Q2) e confrontar (Q3) na Atividade 1 para o 5º e 7º ano.....	86
Gráfico 4.2 –	Análise da relação entre habilidades de interpretar (Q2) e confrontar (Q3) na Atividade 2 para o 5º e 7º ano.....	90
Gráfico 4.3 –	Análise da relação entre habilidades de interpretar (Q2) e confrontar (Q3) na Atividade 4 para o 5º e 7º ano.....	101
Gráfico 5.1 –	Média de acerto no pré-teste e pós-teste por grupo de estudantes	108
Gráfico 5.2 –	Percentual de justificativa adequada em cada habilidade no pré-teste e pós-teste do G5I	114
Gráfico 5.3 –	Percentual de justificativa adequada em cada habilidade no pré-teste e pós-teste do G7I	129
Gráfico 5.4 –	Justificativas dos estudantes do 5º ano em questões dos testes, com destaque no pós-teste	143
Gráfico 5.5 –	Justificativas dos estudantes do 7º ano em questões dos testes, com destaque no pós-teste	144
Gráfico 5.6 –	Atividade 1.1 – dados sobre percentual de resíduos de agrotóxicos em amostras de alimentos	149
Gráfico 5.7 –	Atividade 3.1 – dados sobre número de pessoas no Brasil em 2010 por sexo e idade	151
Gráfico 5.8 –	Atividade 1.2 – dados sobre número de mortes no trânsito para cada 100 mil habitantes e limite de velocidade	153
Gráfico 5.9 –	Atividade 2.1 – dados sobre o tempo médio de uso de aparelhos eletrônicos	156
Gráfico 5.10 –	Atividade 3.2 – dados sobre percentual de filhos das mulheres por nível de instrução nas regiões do Brasil	162
Gráfico 5.11 –	Atividade 2.2 – dados sobre as fontes consultadas pelo consumidor ao pesquisar a durabilidade de produtos	163

LISTA DE TABELAS

Tabela 4.1 – Percentual de tipo de resposta para A3Q1 por estudantes do 5º e 7º anos	93
Tabela 4.2 – Percentual de tipo de resposta para A3Q1 por estudantes do 5º e 7º anos	97
Tabela 5.1 – Frequência de estudantes por turma e ano de escolaridade no pré-teste	105
Tabela 5.2 – Média de acerto por ano de escolaridade no pré-teste	105
Tabela 5.3 – Média de acerto por grupo de estudantes no pré-teste e pós-teste	107

SUMÁRIO

1	INTRODUÇÃO	21
2	FUNDAMENTAÇÃO TEÓRICA	24
2.1	LETRAMENTO ESTATÍSTICO	24
2.2	ANÁLISE EXPLORATÓRIA DOS DADOS	27
2.3	O CICLO INVESTIGATIVO	29
2.4	HIPÓTESES, ANÁLISE DE DADOS E CONCLUSÕES	33
2.4.1	Levantamento de hipóteses	33
2.4.2	Análise de dados	34
2.4.3	Conclusões a partir dos dados	36
2.5	DISTRIBUIÇÕES UNIVARIADAS E BIVARIADAS (CORRELAÇÃO)	37
2.6	COMPREENSÕES E APRENDIZAGENS DE ESTUDANTES DO ENSINO FUNDAMENTAL	39
2.6.1	Compreensões e aprendizagens em distribuições univariadas	39
2.6.2	Compreensões e aprendizagens em distribuições bivariadas	49
3	MÉTODO	53
3.1	OBJETIVO GERAL	53
3.2	OBJETIVOS ESPECÍFICOS	53
3.3	HIPÓTESE	53
3.4	SUJEITOS	54
3.5	PROCEDIMENTOS	55
3.5.1	Estudo 1	55
3.5.2	Estudo 2	61
3.5.2.1	Atividades do pré-teste	62
3.5.2.2	Contribuições da Zona de Desenvolvimento Proximal de Vygotsky para o processo de intervenção didática	67
3.5.2.2.1	<i>Objetivos delineados para os encontros de intervenção</i>	<i>71</i>
3.5.2.3	Atividades do pós-teste	81
4	ANÁLISE DAS COMPREENSÕES DOS ESTUDANTES NO	

	ESTUDO 1	82
4.1	ANÁLISE DA ATIVIDADE 1	82
4.2	ANÁLISE DA ATIVIDADE 2	89
4.3	ANÁLISE DA ATIVIDADE 3	92
4.4	ANÁLISE DA ATIVIDADE 4	96
4.5	CONCLUSÕES DO ESTUDO 1	103
5	ANÁLISE DAS APRENDIZAGENS DOS ESTUDANTES NO ESTUDO 2	105
5.1	APRENDIZAGEM DAS HABILIDADES INVESTIGADAS A PARTIR DE FASES DO CICLO INVESTIGATIVO	111
5.1.1	Aprendizagem no 5º ano diante das justificativas	113
5.1.1.1	Levantamento de hipótese	114
5.1.1.2	Interpretação de dados reais	117
5.1.1.3	Confronto entre hipótese e dados reais	121
5.1.1.4	Avaliação de conclusões	124
5.1.1.5	Uso de linguagem probabilística a partir de previsões	126
5.1.2	Aprendizagem no 7º ano diante das justificativas	128
5.1.2.1	Levantamento de hipótese	129
5.1.2.1	Interpretação de dados reais	132
5.1.2.3	Confronto entre hipótese e dados reais	135
5.1.2.4	Avaliação de conclusões	138
5.1.2.5	Uso de linguagem probabilística a partir de previsões	141
5.2	ANÁLISE DO PROCESSO DE INTERVENÇÃO PEDAGÓGICA	145
5.2.1	Encaminhamento pedagógico das habilidades do estudo nos encontros de intervenção	145
5.2.1.1	Atividades coletivas	147
5.2.1.2	Atividades em pequenos grupos	159
5.2.2	Processo de intervenção pedagógica e Zona de Desenvolvimento Proximal (ZDP) de Vygotsky	167
6	CONCLUSÕES	173
	REFERÊNCIAS	180

1 INTRODUÇÃO

Investigações relacionadas aos processos de ensino e de aprendizagem em Estatística acarretaram o início de uma nova área de atuação pedagógica, denominada Educação Estatística (EE), como afirmam Campos, Wodewotzki e Jacobini (2013). Esses autores salientam que a Educação Estatística tem como um de seus objetivos principais:

Valorizar uma postura investigativa, reflexiva e crítica do aluno, em uma sociedade globalizada, marcada pelo acúmulo de informações e pela necessidade de tomada de decisões em situações de incerteza (CAMPOS et al, 2013, p.12).

Na perspectiva da Educação Estatística, o PNAIC (Pacto Nacional pela Alfabetização na Idade Certa), em 2014, apresenta em seu Caderno 7, referente à Educação Estatística, discussões elaboradas por Gitirana (2014), em que a pesquisa aparece como um dos eixos estruturadores da Estatística na escola, desde os anos iniciais. A autora afirma que fazer pesquisa favorece não apenas a formação estatística do cidadão, como também, a formação científica. Alerta, ainda, que a abordagem interdisciplinar da Estatística precisa ser considerada para formação do cidadão em outras áreas do conhecimento.

Mais recentemente, a Base Nacional Comum Curricular (BNCC, 2017) estabelece que a incerteza e o tratamento de dados sejam estudados na unidade temática denominada Probabilidade e Estatística, dentro da Matemática, desde os primeiros anos do Ensino Fundamental. Ressalta, também, que todo cidadão precisa desenvolver habilidades relacionadas à coleta, à organização, à representação, à interpretação e à análise de dados, nos mais variados contextos e de modo a realizar julgamentos fundamentados e tomar decisões adequadas.

A ênfase dada pela BNCC (2017) à consideração da incerteza, ao desenvolvimento de habilidades investigativas e à necessidade de tomada de decisão é importante, uma vez que vai ao encontro não apenas de umas das metas da Educação Estatística, como também das pesquisas mais recentes (WATSON, 2011; MAKAR, 2013; MCPHEE e MAKAR, 2014; ENGLISH, 2014; entre outros) preocupadas com a aprendizagem da estatística desde os primeiros anos de ensino escolar.

Evidenciam-se nos últimos trabalhos voltados para a escolarização básica da população, seja na pesquisa, seja na elaboração de currículos, a defesa e orientação de se considerar os conceitos estatísticos inseridos em processos investigativos que deem sentido aos mesmos. Acredita-se que isso também se justifica pela compreensão crescente de que as fases de qualquer pesquisa estão conectadas e que possuem finalidades, como confirmar ou não hipóteses levantadas, assim como tomar decisões a partir do que se verificou dos dados coletados, sempre tendo em vista a natureza probabilística das decisões de carácter estatístico. Além disso, o conhecimento do ciclo investigativo permite que julgamentos, opiniões e argumentos a respeito da informação estatística sejam explorados com maior seriedade, de modo a desenvolver o senso crítico do cidadão, constantemente imerso em mensagens estatísticas vindas a público das mais diversas formas.

A necessidade de se trabalhar fases do ciclo investigativo está bem explícita nos currículos e pesquisas recentes da área. Entretanto, estudos que abordem de modo sistemático a relação entre levantamento de hipóteses, enquanto fase condutora da pesquisa, análise de dados em distribuições tanto univariadas, como bivariadas e conclusões tomando-se os dados como evidências não foram encontrados nos meios de divulgação de pesquisas em Educação Estatística no Ensino Fundamental. Entender as compreensões dos estudantes e as possíveis aprendizagens advindas de intervenções que visem o desenvolvimento dos mesmos é necessário para a pesquisa e para a prática de sala de aula.

Desse modo, configura-se a problemática de investigação desta tese a partir da seguinte questão: De que modo o levantamento de hipóteses, análise de dados e habilidades relacionadas a conclusões por estudantes do Ensino Fundamental corroboram para a aprendizagem de justificativas baseadas em evidências?

A fim de responder à questão proposta, define-se como objetivo geral da tese analisar aprendizagens apresentadas por estudantes do Ensino Fundamental relacionadas ao levantamento de hipóteses, à análise de dados e a capacidade de realizar conclusões a partir dos dados, enquanto fases do ciclo investigativo.

Na seção seguinte, a fundamentação teórica e devida revisão da literatura são apresentadas. Nesse momento, o Letramento Estatístico da população é defendido e como forma de desenvolvê-lo, aborda-se a perspectiva da Análise Exploratória de Dados, assim como a contemplação de fases do ciclo investigativo.

Descrições e discussões de pesquisas recentes, com estudantes do Ensino Fundamental, a respeito das fases delineadas no objetivo da tese são realizadas.

Na terceira seção se expõe o método de pesquisa delimitado para o objetivo da tese, assim como os objetivos específicos são explicitados. Dois estudos são descritos, o primeiro de carácter diagnóstico, em que compreensões de estudantes do 5º e 7º anos do Ensino Fundamental são investigadas, o segundo traz a proposta de intervenção junto a turmas desses anos. Para o processo de intervenção, cumpre-se discutir o conceito de Vygotsky de Zona de Desenvolvimento Proximal (ZDP), uma vez que a mesma considera a aprendizagem e o posterior desenvolvimento da criança em situações de interação com outras crianças e adultos, o que é extremamente relevante em situações de ensino.

A quarta seção apresenta análises referentes ao primeiro estudo realizado. Comparações entre as compreensões demonstradas por ambos os grupos de estudantes são exploradas, assim como discussões a respeito da influência de cada atividade e suas peculiaridades nesse processo. Conclusões referentes ao primeiro estudo são explicitadas, além de discussões e propostas que levam à implantação do segundo estudo, de carácter interventivo experimental.

A quinta seção trata das análises a respeito do Estudo 2, com as devidas discussões a respeito dos limites e avanços constatados na realização do estudo experimental. O resgate do referencial teórico é realizado. Assim como a discussão da ZDP de Vygotsky em relação aos resultados alcançados é proposta.

Por fim, conclusões são explicitadas na sexta seção, com retomada das evidências que as embasem para a construção da tese aqui apresentada.

2 FUNDAMENTAÇÃO TEÓRICA

O estudo apresenta como principal aporte teórico as discussões e princípios tratados na perspectiva do Letramento Estatístico, especificamente a desenvolvida por Iddo Gal (2002). Outros estudos colaboram na construção da tese, como a Análise Exploratória de Dados (AED), tal qual explorada por Batanero, Estepa e Godino (1991), a partir de Tukey (1977), assim como a própria revisão da literatura.

2.1 LETRAMENTO ESTATÍSTICO

Pensar no ensino de conceitos estatísticos de modo a desenvolver uma postura investigativa, reflexiva e crítica, tal como propõem Campos, Wodewotzki e Jacobini (2013), ao referirem-se à Educação Estatística, é possível dentro de um trabalho escolar que vise o Letramento Estatístico (LE) da população. Assim, o movimento pelo Letramento Estatístico surge como um dos principais avanços na Educação Estatística.

O termo Letramento pode ser utilizado por outros autores como literacia, pela tradução do inglês *literacy*. Tem como um de seus mais atuantes pesquisadores Iddo Gal (2002), referência maior no presente estudo. Letramento Estatístico, para Gal (2002), refere-se à habilidade de interpretar, avaliar criticamente e comunicar mensagens e informações estatísticas. O autor ressalta que LE é necessário a adultos e futuros adultos, pois torna cidadãos e profissionais mais informados. Afirma, ainda, que pesquisadores interessados em processos cognitivos enfatizam a contribuição de julgamentos adequados e do raciocínio probabilístico como habilidades para a tomada de decisões efetivas. O processo de tomada de decisão, aliás, como parte do ciclo investigativo, que será abordado posteriormente, é para Makar (2013) essencial à Estatística, pois de acordo com a autora pesquisa-se para decidir, concluir, fazer inferências.

Em sociedades industrializadas, que contam com uma grande quantidade de conhecimento acumulado, entendimento de processos investigativos é imprescindível. Nesse sentido, Gal (2002) propõe dois componentes inter-relacionados do Letramento Estatístico, os quais atuam de modo conjunto:

- a) Habilidades de pessoas para interpretar e avaliar criticamente informação estatística, argumentos relacionados a dados ou fenômenos estocásticos que eles podem encontrar em diversos contextos e quando relevante;
- b) Habilidades para discutir ou comunicar suas reações a tal informação estatística, assim como seu entendimento do significado da informação, suas opiniões sobre as implicações desta informação ou suas preocupações em relação à aceitabilidade das conclusões dadas. (GAL, 2002, p. 2-3).

O autor estabelece, ainda, que há diferença na habilidade das pessoas em agir, de acordo com o contexto em que os dados se apresentam. Gal (2002) alerta que há o contexto de leitura e o contexto de investigação. No primeiro, a informação estatística se apresenta em situações cotidianas, através de texto, números e símbolos, além de disposição gráfica ou tabular, nesse caso o indivíduo é um consumidor de dados. No segundo contexto, o indivíduo encontra-se engajado em uma investigação empírica e produz ou analisa os próprios dados, tratando-se possivelmente de um estudante, estatístico ou pesquisador.

Detendo-se ao contexto de leitura, foco do presente trabalho, Gal (2002) apresenta um modelo de letramento estatístico, para o adulto (e por implicações para o ensino), que é composto por elementos cognitivos e elementos disposicionais, conforme o Quadro 2.1:

Quadro 2.1 – Modelo de Letramento Estatístico

Letramento Estatístico	
Elementos de conhecimento	Elementos disposicionais
Habilidades de letramento	Crenças e atitudes
Conhecimento estatístico	Postura crítica
Conhecimento matemático	
Conhecimento de contexto	
Questões/habilidades críticas	

Modelo adaptado de Gal (2002, p.4).

Evidencia-se no Quadro 2.1, que entendimento e interpretação de informação estatística requer não apenas conhecimento estatístico, mas a operação conjunta de diferentes conhecimentos e disposições.

Habilidades de letramento referem-se à necessidade de ativação de várias habilidades de processamento de texto para obtenção de significado da informação estatística. Conhecimentos de conceitos e procedimentos estatísticos e probabilísticos, além de conceitos matemáticos relacionados são imperativos. Conhecimento do contexto está relacionado com conhecimento de mundo, nesse caso, familiaridade com o contexto é o principal determinante na interpretação e julgamento da mensagem estatística. Habilidades críticas relacionam-se com a percepção de que informações estatísticas advêm de diferentes fontes, que a depender de seus interesses, nem sempre estarão empenhadas em oferecer um relato objetivo e equilibrado de suas conclusões e implicações. Quanto aos elementos disposicionais do LE, os mesmos mostram-se quando o indivíduo demonstra inclinação para ativar os cinco conhecimentos descritos ou compartilhar com outros suas opiniões, julgamentos ou interpretações alternativas.

O modelo proposto por Gal (2002) ressalta a necessidade de se mobilizar diferentes conhecimentos, inclusive de mundo, para se interpretar e julgar informações estatísticas. O autor reconhece que a Estatística encontra-se interligada com outras áreas do saber e que o indivíduo que não produz os dados, mas atua como um receptor, no contexto de leitura, como denomina Gal (2002), necessita desenvolver uma postura crítica diante da confiabilidade ou não do que lhe é apresentado. Isso pressupõe, que a informação estatística que chega ao público, para que se torne familiar, requer detalhamento e clareza quanto ao processo de geração de dados.

Gal (2002) preocupou-se em explicitar as habilidades necessárias a um adulto letrado estatisticamente. Contudo, diferentes pesquisadores, nacionais e internacionais (COUTINHO, ALMOULOU e SILVA, 2012; MONTEIRO, 2016; WATSON, 2011; DÍAZ-LEVICOY, BATANERO, ARTEAGA, LÓPEZ-MARTÍN, 2015; entre outros), que atuam na Educação Estatística de crianças e adolescentes têm se reportado ao Letramento Estatístico, na perspectiva explicitada por esse autor, principalmente no que diz respeito ao desenvolvimento crítico do cidadão diante da informação estatística.

Monteiro (2016) discute a necessidade da Educação Estatística promover o Letramento Estatístico na escola básica, a fim de se favorecer a construção da cidadania crítica pelos estudantes. Como uma das estratégias para o alcance do LE, o autor cita o trabalho com dados reais, pois quando se pensa na realidade

brasileira, principalmente, estatísticas têm sido utilizadas para justificar permanências e mudanças em políticas públicas.

Watson (2011) salienta que o Letramento Estatístico emerge nos primeiros anos de escolaridade, onde muitas ideias básicas são introduzidas através de componentes do currículo de matemática, como manipulação de dados e possibilidades. Preocupada com o desenvolvimento do LE das crianças desde o início da escolarização, a autora criou uma hierarquia que expõe níveis de entendimento, tomando-se o *SOLO* (*Structure of Observed Learning Outcomes*) como base. O modelo *SOLO* foi elaborado por Biggs e Collis (1982) e é empregado para análise de respostas dos estudantes às atividades propostas, o mesmo fundamenta-se em teorias neo-piagetianas para o pensamento concreto-simbólico (CAVALCANTI, 2011).

Na hierarquia apresentada por Watson (2011) e adaptada de Watson e Callingham (2003), ideias que evidenciem análise crítica da atividade são consideradas habilidades com alto grau de compreensão, assim como a utilização de conceitos matemáticos e apreciação da incerteza em situações que a requerem, o que pressupõe raciocínio probabilístico. No modelo proposto por Gal (2002), tais compreensões aparecem como necessárias ao LE da população.

2.2 ANÁLISE EXPLORATÓRIA DOS DADOS

Pensar no Letramento Estatístico da população, tal qual proposto por Gal (2002), que permita interpretação e julgamento da informação estatística, pressupondo uma postura crítica, requer que o ensino, principalmente na Educação Básica, supere a análise estatística convencional, baseada fundamentalmente em procedimentos que conduzem a uma importância exclusiva dos cálculos estatísticos. Batanero, Estepa e Godino (1991) salientam que o desenvolvimento tecnológico possibilitou que a análise de dados assumisse um novo enfoque, denominado Análise Exploratória de Dados (AED), introduzida por Tukey (1977).

A AED vai de encontro com a ênfase em cálculos estatísticos, que se relacionava a um modelo confirmatório, em que o propósito era confirmar ou não uma hipótese pré-estabelecida, tratava-se de colocá-la a prova. Como consequência, diminuía-se a importância visual da representação dos dados e não

se explorava os dados para se extrair dos mesmos qualquer outra informação que se pudesse deduzir.

É importante ressaltar que não se trata de negar o modelo confirmatório. Tukey (1980) estabelece que nem o modelo exploratório, nem o confirmatório é suficiente sozinho, precisa-se de ambos. O autor afirma que encontrar a questão é frequentemente mais importante que encontrar a resposta. Desse modo, compreende-se que a questão define o tipo de análise. Questões importantes podem demandar o mais cuidadoso planejamento de análises confirmatórias. Assim como investigações mais amplas, a partir de questão geral, também são importantes e requerem uma exploração dos dados. Tukey (1980) ainda complementa que a AED não pode ser entendida como um pacote de técnicas. Ela é uma atitude, uma flexibilidade e dependente do exame de representações.

O enfoque da AED, de acordo com Batanero et al (1991), faz uso de tecnologias com capacidade de cálculos e representações gráficas, que permitem a obtenção de uma ampla variedade de gráficos e diferentes estatísticas. Isso modifica a concepção de análise estatística, o foco da análise, que antes estava na busca de um modelo que expressasse a regularidade dos dados, em hipótese, passa a estar na exploração não apenas das regularidades de um conjunto de observações, mas também em seus desvios. A intensão é extrair toda informação possível e gerar novas hipóteses.

Batanero et al (1991) expõem características da AED, possíveis de serem desenvolvidos no ensino secundário, foco dos autores no estudo apresentado.

- Possibilidade de criar situações de aprendizagem, a partir de temas de interesse do estudante;
- Forte apoio em representações gráficas;
- Emprego preferencial das estatísticas de ordem, porque são sensíveis a maior parte dos dados e diminuem os efeitos produzidos pelos valores atípicos, escassos e muito longe das normas;
- Não necessita uma teoria matemática complexa;
- A representação ou o cálculo não são na AED o fim, mas um meio de descobrir a informação oculta dos dados;
- Uso de diferentes escalas ou reformulações.

É interessante observar a importância dada no enfoque da AED, tanto por Tukey (1980), quanto por Batanero et al (1991) à representação dos dados. Curi e Nascimento (2016) afirmam que de todas as maneiras que há de se comunicar informações estatísticas, o gráfico bem produzido é o modo mais simples e mais poderoso para tal realização.

Batanero et al (1991) assumem, ainda, que no enfoque da AED é essencial uma atitude investigativa, sem redução do conhecimento a emprego de técnicas. Diante das recentes pesquisas realizadas com estudantes já desde os anos iniciais e perante as orientações do currículo voltado para esse público, acredita-se que as características expostas por Batanero et al (1991) para o ensino da AED podem ser pensadas em um nível mais elementar, uma vez que a abordagem da Estatística enquanto ciclo investigativo é defendida e orientada já desde os primeiros anos de escolarização.

Ben-Zvi (2016) já defende que crianças a partir do 5º ano do Ensino Fundamental (dez anos de idade) trabalhem com investigações estatísticas em situações de ensino, na perspectiva da AED. Os estudantes podem coletar dados, enquanto amostra, sem que inferências explícitas sejam realizadas. Entende-se, nesse sentido, a preocupação do autor com a instrução de conceitos estatísticos desde os anos iniciais, sem que concepções equivocadas dos mesmos aconteçam. A AED, portanto, permite também esse contato inicial dos estudantes com investigações estatísticas, valorizando a exploração dos dados que se tem, antes de qualquer formalização de carácter mais complexo.

2.3 O CICLO INVESTIGATIVO

Na pesquisa de carácter estatístico, delinear um processo de investigação para um problema é fundamental. Wild e Pfannkuch (1999) apresentaram componentes para o processo investigativo, denominado por eles de estágios. Esses autores adaptaram o modelo PPDAC (Problema, Plano, Dados, Análise e Conclusões) de MacKay e Oldford (1994), caracterizando-o como ciclo investigativo de PPDAC. O esquema (Figura 2.1), a seguir, ilustra a proposta elaborada por Wild e Pfannkuch (1999, p.226).

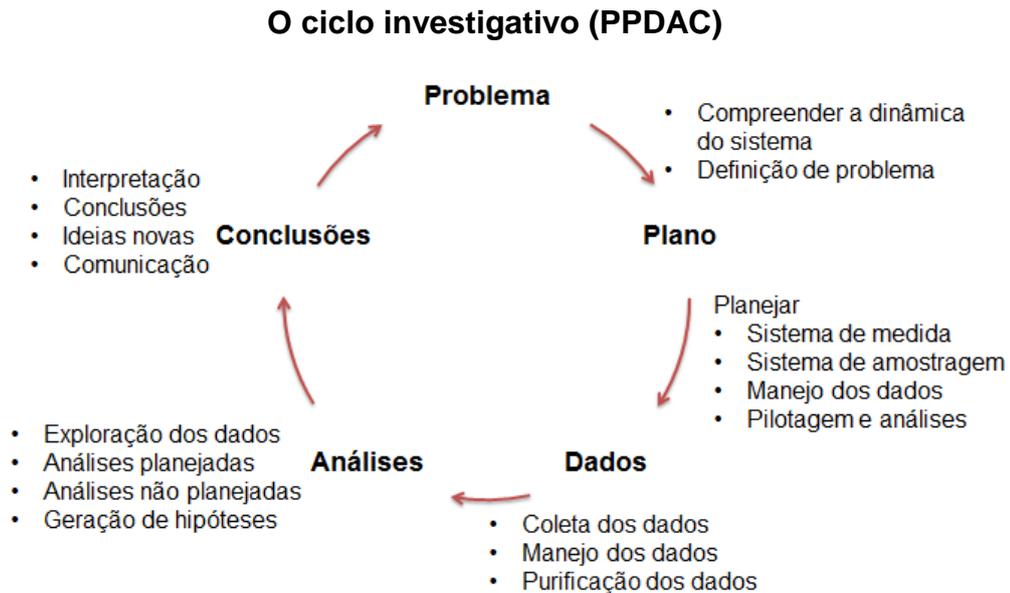


Figura 2.1 – Ciclo investigativo (PPDAC) de Wild e Pfannkuch (1999).

Wild e Pfannkuch (1999) entendem que o desenvolvimento do ciclo pode iniciar futuros ciclos investigativos. Além disso, chamam a atenção para a relação do problema estatístico com um problema real maior. É nesse sentido que se entende que ao se trabalhar estatística em sala de aula sejam utilizados dados reais e de interesse para o estudante (CAMPOS et al, 2013), para que a mesma tenha significado.

Na publicação do estudo de Wild e Pfannkuch (1999), pela *International Statistical Review*, discussões a respeito do mesmo são explicitadas ao término do texto, por diferentes pesquisadores. Dentre os diversos olhares a respeito do estudo realizado, Breslow (1999) concorda com a natureza interativa do ciclo investigativo, como proposta pelos autores, entretanto, salienta que o ciclo investigativo funciona melhor quando a hipótese conduz o estudo, ou seja, apresenta-se desde o início e não após a coleta dos dados, como demonstra o esquema anterior.

Embora eu concorde com W&P que o ciclo investigativo é um processo interativo, eu acredito que ele funciona melhor quando ele é conduzido pela hipótese. Assim, eu colocaria “geração de hipóteses” ou talvez “especificação de hipóteses” no início do ciclo, antes da coleta de dados, ao invés de depois. (BRESLOW, 1999, p.253).

O *GAISE (Guidelines for Assessment and Instruction in Statistics Education)*, em 2007, endossado pela *American Statistical Association (ASA)*, ao apresentar a estrutura curricular da Educação Básica, denominada por eles de Pre-K-12, situam a

necessidade de se considerar a variabilidade na formulação de um problema/questão de pesquisa. Eles estabelecem que a formulação de uma questão estatística requer o entendimento da diferença entre uma questão que antecipa uma resposta determinística e uma questão que antecipa uma resposta apoiada em dados que variam (*GAISE*, 2007, p.15). Compreende-se que em se tratando de pesquisa estatística, considerar a variabilidade é imprescindível.

O *GAISE* (2007) propõe quatro componentes envolvidos no processo investigativo:

- I. Formular questões: esclarecer o problema em mãos; formular uma ou mais questões que possam ser respondidas com dados.
- II. Coletar dados: elaborar um plano para coletar dados apropriados; empregar o plano para coletar os dados.
- III. Analisar os dados: selecionar gráficos e métodos numéricos apropriados; usar esses métodos para analisar os dados.
- IV. Interpretar resultados: interpretar as análises; relacionar a interpretação com a questão original.

Nos componentes expostos no *GAISE* (2007) para a pesquisa estatística, não está claro o papel do levantamento de hipóteses no processo. Nesta tese, acredita-se ser fundamental a explicitação do papel que o levantamento de hipóteses desempenha. Em acordo com a discussão explicitada por Breslow (1999) anteriormente, defende-se aqui que a hipótese tem a importante função de conduzir a pesquisa estatística, o que não significa que novas hipóteses não surjam no decorrer do trabalho investigativo.

Nesse sentido, o presente estudo utiliza o ciclo investigativo elaborado por Guimarães e Gitirana (2013, p. 97), o qual inclui como uma das fases o levantamento de hipóteses, enquanto elemento definidor da amostra. A Figura 2.2 estabelece as fases do ciclo investigativo, conforme proposta das autoras.

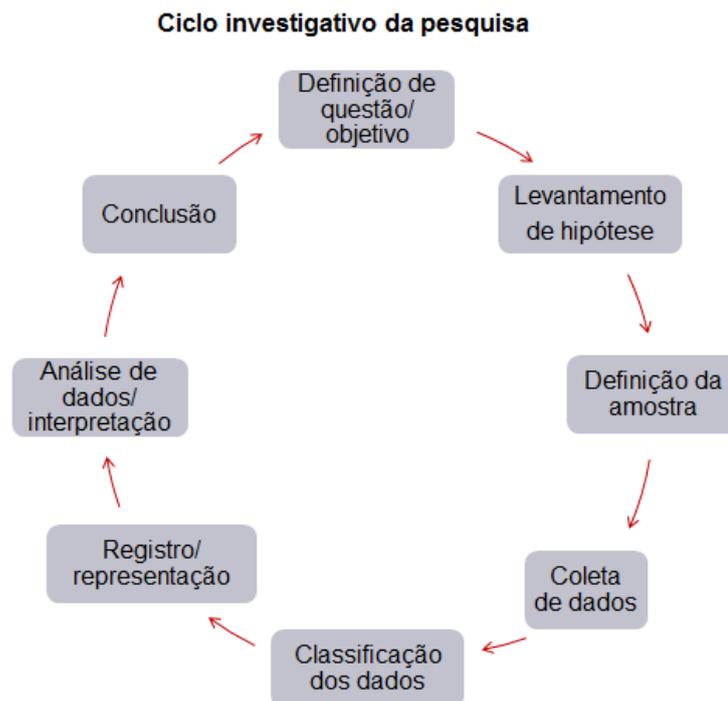


Figura 2.2 – Fases do ciclo investigativo de Guimarães e Gitirana (2013, p. 97).

Na Figura 2.2, as etapas do ciclo investigativo foram delineadas de modo a ressaltar as características diferenciadas de cada um dos momentos, os quais exigem uma abordagem específica ao se pensar em processo de ensino.

Guimarães e Gitirana (2013) salientam que a aprendizagem dos estudantes em relação à pesquisa ocorre trabalhando-se o ciclo investigativo como um todo, assim como explorando-se paralelamente uma ou mais fases do ciclo.

Em acordo com Guimarães e Gitirana (2013), Silva (2013) em seu estudo de mestrado analisou livros didáticos de Matemática e Ciências voltados para os anos iniciais do Ensino Fundamental e constatou que atividades com levantamento de hipóteses foram raras, principalmente no que se refere ao livro de Matemática. Dentre os livros pesquisados, apenas uma atividade de Matemática explorou essa fase, nos livros de Ciências foram 13,4% das atividades.

Em se tratando dos anos finais do Ensino Fundamental, análises em livros didáticos também revelaram baixo número de atividades abordando levantamento de hipóteses em livros de Matemática, de acordo com Mendonça, Gitirana e Lira (2013). Novamente os livros de Ciências exploraram mais essa etapa do ciclo investigativo (37%) do que os livros de Matemática (2%).

Contemplar na prática de sala de aula as diferentes fases do ciclo investigativo é necessário para a aprendizagem da pesquisa estatística. Acredita-se que a utilização de atividades que trabalhe especificamente determinadas fases do ciclo investigativo é essencial, uma vez que umas podem representar maiores dificuldades que outras na aprendizagem, de acordo com pesquisas realizadas na área, o que será discutido com mais detalhes na revisão da literatura.

2.4 HIPÓTESES, ANÁLISE DE DADOS E CONCLUSÕES

Das diferentes etapas da pesquisa, expostas no ciclo investigativo proposto por Guimarães e Gitirana (2013, p. 97), delimitou-se como objeto de estudo na tese a busca da relação entre o levantamento de hipóteses, a análise de dados e as conclusões realizadas a partir da análise. Isso com a finalidade de se obter maior entendimento das compreensões e aprendizagens advindas das fases destacadas.

2.4.1 Levantamento de hipóteses

No campo da metodologia da pesquisa científica, Vergara (2004) afirma que hipóteses ou suposições são a antecipação da resposta ao problema. Se o mesmo foi formulado como uma pergunta, a hipótese apresenta-se como uma afirmação. A investigação, então, acontece de modo que se possa confirmar a hipótese ou refutá-la. A autora expõe que hipóteses estatísticas são elaboradas em formas nula e alternativa. Como exemplo de ambas as formas, Vergara (2004, p.28) apresenta:

- Hipótese nula – Não há relação significativa entre marca e desejo de compra por parte do adolescente.
- Hipóteses um – Há relação significativa entre marca e desejo de compra por parte do adolescente.

Ao dedicarem-se a estudos voltados para a Educação Estatística da população, Guimarães e Gitirana (2013, p.100) mostram-se de acordo com Vergara (2004), quando descrevem hipótese como “*uma afirmativa elaborada como resposta a uma questão, apoiada em uma justificativa e que será colocada à prova, de maneira que poderá ser rejeitada ou não*”. Percebe-se, desse modo, que Guimarães

e Gitirana (2013) corroboram com a ideia de que o levantamento de hipóteses é anterior ao processo de coleta de dados. Ao se definir uma questão ou problema que requeira análise estatística, acredita-se que a elaboração de uma hipótese que possa explicar a questão, de modo justificado, é o que irá direcionar a etapa seguinte do ciclo investigativo, a depender do tipo de pesquisa que se pretenda realizar.

Autores interessados em metodologia científica (CERVO; BERVIAN, 2002; FERNANDES; GOMES, 2003; entre outros) observam que em estudos exploratórios ou descritivos não se elaboram hipóteses a serem testadas. Isso porque, no caso da pesquisa exploratória, buscam-se maiores informações sobre determinado assunto de estudo, de modo a se obter maior familiaridade com o problema, podendo-se construir hipóteses durante o estudo ou ao seu término. Já na pesquisa descritiva, Fernandes e Gomes (2003, p.8) afirmam que o objetivo principal da mesma é “descrever, analisar ou verificar as relações entre fatos e fenômenos (variáveis), ou seja, tomar conhecimento do que, com quem, como e qual a intensidade do fenômeno em estudo”. A respeito do levantamento de hipóteses em pesquisas de caráter descritivo, Lakatos e Marconi (1996) defendem que a mesma pode ser um estudo de verificação de hipóteses, ao conter hipóteses explícitas a serem verificadas, derivadas da teoria, em associação de variáveis.

Entende-se, desse modo, que as fases do ciclo investigativo não se constituem momentos rígidos e inflexíveis. Elas acontecem de modo a responder a questão inicial fixada, podendo-se levantar hipóteses desde o princípio, caso seja possível, como também as mesmas podem aparecer em outros momentos da pesquisa. A AED (Análise Exploratória de Dados) já advoga nesse sentido, expondo a necessidade de geração de novas hipóteses numa análise exploratória.

2.4.2 Análise de dados

A Análise Exploratória de Dados, como já explicitada anteriormente, atua nesse momento do ciclo investigativo, respondendo a questões delineadas em seu início e, principalmente, levantando outras. O processo de interpretação dos dados caracteriza essa fase da pesquisa. Nesse sentido, a utilização de dados representados em gráficos tem sido bastante explorada em diferentes âmbitos (sala

de aula e pesquisas acadêmicas) e recursos relacionados ao ensino (Livros Didáticos e sequências de atividades).

Em se tratando de interpretação de gráficos, Friel, Curcio e Bright (2001, p. 129) oferecem uma grande contribuição à discussão, quando propõem que a compreensão da informação em forma escrita ou simbólica envolve três tipos de comportamento:

- Translação – requer mudança na forma de uma comunicação. Ex.: Mudança entre gráficos e tabelas;
- Interpretação – requer reorganização de material e separação dos fatores importantes, daqueles menos importantes;
- Extrapolação e interpolação – consideradas extensões da interpretação, requer expor não apenas a essência da comunicação, mas também identificar consequências. Ex.: Observação de tendências nos dados ou implicações específicas.

Os autores afirmam, ainda, que o gráfico é uma ferramenta que dá sentido à informação. Ao se pensar na compreensão de gráfico, Friel et al (2001, p. 130) destacam que há três níveis de questões a serem consideradas: As de nível elementar; as de nível intermediário; as de nível avançado. As questões do primeiro nível são aquelas que se referem à extração de informação dos dados, as mesmas demandam uma leitura dos dados. No segundo nível, verificam-se questões que exigem interpretação e integração de informações que são apresentadas no gráfico, necessitando leitura entre os dados. No último nível estão as questões que extrapolam os dados, realizando-se predições ou inferências, essas questões são conhecidas também como aquelas que requerem leitura além dos dados. Friel et al (2001) alertam que as primeiras questões não geram tantas dificuldades para os estudantes quando as de nível intermediário e principalmente as do nível avançado.

É importante esclarecer, que os diferentes níveis de questões expostos por Friel et al (2001) envolvem mais de uma fase do ciclo investigativo. A fase da pesquisa denominada por Guimarães e Gitirana (2013) como conclusão requer a compreensão dos dados apresentados em níveis que pressuponham integração de informações (nível intermediário), assim como extrapolações e predições dos dados (nível avançado). O tópico seguinte aborda especificamente essa fase da pesquisa estatística.

2.4.3 Conclusões a partir dos dados

Ao findar um processo investigativo, uma resposta precisa ser dada para a questão (problema levantado) no início do ciclo investigativo, essa resposta é a conclusão do estudo. Vergara (2004) estabelece que só se pode concluir sobre aquilo que se discutiu no decorrer da pesquisa. Em outras palavras, a autora recomenda que nesse momento haja um resgate da pergunta-problema que desencadeou todo o estudo e que toda conclusão advinda nesse momento se atenha aos dados analisados. Afinal, não se pode concluir sobre o que não se estudou.

Tendo-se em vista os diferentes tipos de análise que podem ser realizados em processos de investigação, Garfield e Ben-Zvi (2008) afirmam que conclusões na AED são informais, baseadas naquilo que se vê dos dados e aplicadas apenas para os indivíduos e circunstâncias dos dados que se têm em mãos. Por outro lado, conclusões em inferência estatística são formais e possíveis de serem generalizadas.

Em se tratando de exploração do ciclo investigativo com estudantes da Educação Básica, especificamente do Ensino Fundamental, como é o caso do presente estudo, conclusões baseadas na AED é o foco. Nesse sentido, é de interesse na tese, assim como entre os estudos que abordam a última fase do ciclo investigativo, que os estudantes desenvolvam a capacidade de tomar decisões a partir dos dados. English (2014) ressalta que no mundo real decisões são requeridas onde há incerteza e onde várias alternativas podem ser razoáveis.

Compreende-se que a tomada de decisão a partir dos dados está inclusa no processo de conclusão, do mesmo modo que as predições realizadas com base nos dados. No caso das predições, considerar a incerteza como característica da pesquisa estatística é essencial, para isso, utilização da linguagem probabilística ao se prever é importante.

Percebe-se que para o processo de conclusão, pensar em leitura entre os dados (com a devida relação dos dados), assim como leitura além dos dados (no caso das predições) é necessário. Esclarece-se que estudos recentes vêm denominando de inferência informal a leitura além dos dados na realização de predições (MAKAR, 2013; MCPHEE; MAKAR, 2014; ENGLISH, 2014; entre outros).

2.5 DISTRIBUIÇÕES UNIVARIADAS E BIVARIADAS (CORRELAÇÃO)

A análise de dados, discutida anteriormente, envolve o processo de busca de padrões em dados univariados e bivariados. Acredita-se que o tipo de distribuição (univariada ou bivariada) interfere na análise e compreensão dos dados, assim como na pesquisa como um todo. Distribuições univariadas são amplamente exploradas no Ensino Fundamental, ao contrário daquelas bivariadas, nas quais se analisa a relação entre duas variáveis, o que a literatura vai nomear de correlação.

Triola (2013) expõe que “*existe correlação entre duas variáveis quando os valores de uma variável estão relacionados, de alguma maneira, com os valores da outra variável*” (p.416). Na análise de correlação entre duas variáveis, o autor considera importante explorar visualmente os dados, antes de qualquer análise estatística formal e utilizando-se de diagramas de dispersão (Figura 2.3), distingue os padrões positivo e negativo da correlação linear (os pontos no diagrama assemelham-se a uma reta), assim como a ausência de correlação. Os termos correlação positiva ou negativa também são conhecidos na literatura como correlação direta ou indireta, respectivamente.

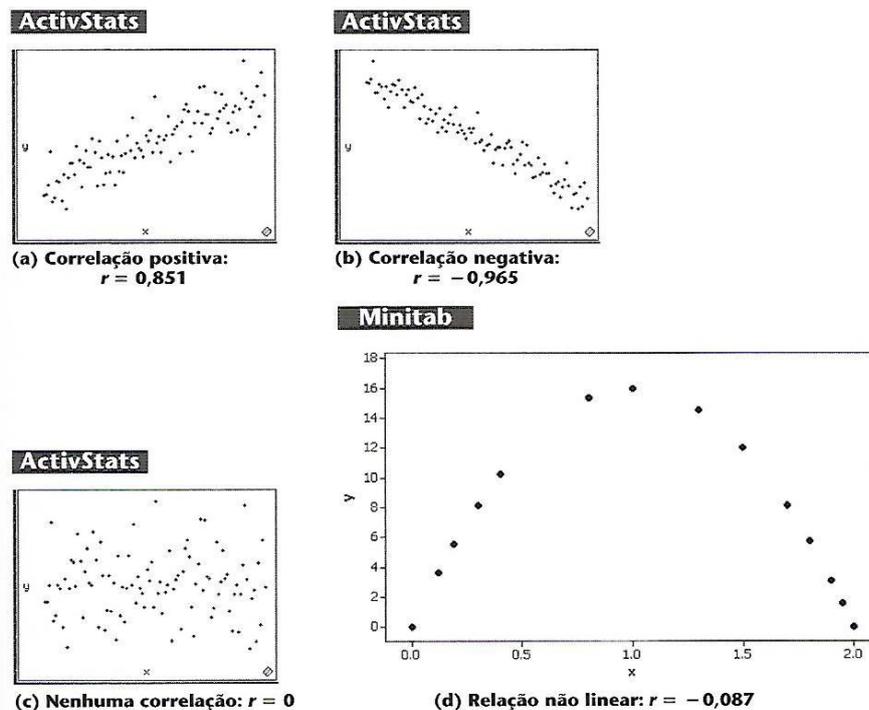


Figura 2.3 – Diagramas de dispersão (TRIOLA, 2013, p.417).

Na correlação positiva, à medida que os valores de uma variável (x) crescem, os valores correspondentes da outra variável (y) também crescem, como exemplo, pode-se citar a relação entre preço da matéria-prima e custo de produção. Já na correlação negativa, quando os valores uma variável (x) crescem, os valores da variável correspondente (y) decrescem, um exemplo pode ser a relação entre número de pessoas comprando determinado produto e baixo custo do mesmo.

Daniels (2017), Triola (2013) e Friel et al (2001) concordam que interpretar correlação de dados requer o cuidado de se evitar o erro comum de impor uma relação causa/efeito. Triola (2013) esclarece que “correlação não implica causalidade” (p. 423).

Friel et al (2001) afirmam que estudos com dados bivariados são importantes no currículo e especifica o trabalho em níveis mais avançados de ensino. Entretanto, pesquisas realizadas com estudantes do Ensino Fundamental, envolvendo relação entre duas variáveis, já revelam avanços no raciocínio correlacional a partir de sequências de ensino apropriadas (NUNES, 2015; FITZALLEN; WATSON, 2014), como será detalhado no tópico seguinte.

Ao discorrerem sobre o raciocínio correlacional, Nunes e Bryant (2011) afirmam que o mesmo é essencial no raciocínio e letramento científico, além disso, é fundamental no sentido de controlar o presente e prever o futuro para maximizar o desejável. Os autores estabelecem também, que o mesmo demanda o reconhecimento que relações entre variáveis não são absolutas, mas existe em graus. Isso faz com que o raciocínio probabilístico esteja presente.

Dentre as demandas cognitivas envolvidas no entendimento do raciocínio correlacional, Nunes e Bryant (2011) propõem a compreensão de: 1) aleatoriedade, na qual o sujeito percebe que eventos sem relação podem ocorrer ao acaso; 2) espaço amostral, habilidade relacionada com a percepção de todos os casos possíveis para a relação; e 3) quantificação da probabilidade, que se refere à capacidade de quantificar, de modo proporcional, a probabilidade de ocorrência de casos.

As demandas cognitivas estabelecidas por Nunes e Bryant (2011) para o raciocínio correlacional surgem do entendimento de que correlação depende de dois esquemas cognitivos: probabilidade e proporcionalidade. Os autores se apoiaram nos trabalhos de Inhelder e Piaget (1958), Karplus, Adi e Lawson, (1980) e Vass, Schiller e Nappi's, (2000) para a proposição das demandas explicitadas.

Nunes e Bryant (2011) discutem, ainda, que a organização dos dados pode ser importante para o raciocínio correlacional, porque alguns estudantes não conseguem começar a fazê-la. Os autores adotam a organização dos dados em tabelas, nesse sentido. Esses autores alertam, também, que correlações diretas são mais facilmente percebidas do que as indiretas.

Por fim, acredita-se importante destacar aqui a sugestão dada por esses autores, de que crianças e adultos são influenciados pela natureza da relação que eles esperam existir, ou seja, baseiam-se em crenças.

Nickerson (1998) utiliza o termo viés de confirmação, amplamente utilizado na psicologia, de acordo com o autor, para se referir à busca ou interpretação de evidências de maneira que sejam parciais, isto é, voltadas para crenças ou expectativas. Em se tratando de informação, Nickerson (1998) afirma que em primeira instância, há a necessidade de se buscar provas, avaliá-las de modo mais objetivo possível e concluir o que as evidências, em conjunto, parecem mostrar. Contudo, uma seletividade não consciente faz com que um peso indevido seja dado a evidências que apoiem ideias preconcebidas, enquanto negligenciam-se evidências que vão de encontro a alguma hipótese existente.

2.6 COMPREENSÕES E APRENDIZAGENS DE ESTUDANTES DO ENSINO FUNDAMENTAL

Recentemente, estudos realizados com estudantes dos anos iniciais e anos finais do Ensino Fundamental têm contemplado o ciclo investigativo como um todo, assim como uma ou mais fases do mesmo. Neste tópico, uma atenção especial é dada às compreensões e às aprendizagens demonstradas pelo público desses anos, especificamente a respeito das fases de levantamento de hipóteses, análise de dados e conclusões tomando-se os dados como evidências. Pesquisas abordando a relação entre variáveis também são apresentadas.

2.6.1 Compreensões e aprendizagens em distribuições univariadas

McPhee e Makar (2014) investigaram crianças australianas (entre 5 e 6 anos de idade), em situações que poderiam emergir práticas de inferência informal

estatística, termo utilizado pelas autoras para a realização de predições. Foi de interesse para as autoras verificar se as crianças tomavam os dados como evidências no processo. O estudo aconteceu em duas fases, de 3 (três) encontros cada, com intervalo de seis meses. Participaram da pesquisa 22 (vinte e duas) crianças, selecionadas de cinco turmas da educação infantil (cinco anos de idade), que no ano escolar seguinte (já ao frequentarem o primeiro ano na Austrália, aos seis anos de idade), retomaram a segunda fase da pesquisa. Na fase 1, a linguagem probabilística ao se fazer predições foi explorada, em situações didáticas comuns da educação infantil, como contação de história, além disso, previsão e registro, com repetições de quantos objetos de mesmo formato caberiam em suas mãos, foram solicitados. As autoras constataram uso de linguagem com consideração da incerteza e verificaram também, nessa fase, que nem todas as crianças melhoraram suas predições. Ressaltam, portanto, que basear-se em evidências pode ser difícil nessa faixa etária.

Na fase 2, predições foram exploradas propondo-se inicialmente a questão de pesquisa “*Quão grande são nossos sapatos?*”. A fase 1, nesse sentido, foi preparatória para a fase 2, na qual estimulou-se predições a partir da pesquisa estatística. Assim, os estudantes na fase 2 coletaram dados sobre o tamanho de seus calçados e foram solicitados a fazer previsões a respeito do tamanho de calçados de outras crianças de mesma idade, esperava-se que os dados coletados fossem utilizados como evidências nas predições. Estudantes em duplas demonstraram habilidade em fazer predições baseadas nos dados e em utilizar linguagem probabilística (como “talvez” e “pode ser”), embora alguns acreditassem que os resultados das previsões seriam idênticos aos coletados. McPhee e Makar (2014) estabelecem que todos os estudantes mostraram capacidade de demonstrar algum aspecto da inferência informal estatística durante o estudo.

English (2014), na Austrália, realizou um estudo longitudinal com crianças da escola elementar, desde o primeiro até o terceiro ano (entre 6 e 8 anos de idade). Três turmas participaram da pesquisa, estudantes juntamente com seus professores. Representações de dados em tabela, com observação da variação e elaboração de inferências informais foram o foco do estudo da autora, nas atividades intituladas *Litterbug Doug* (trabalhada no primeiro ano) e Piquenique de *Baxter Brown* (explorada no segundo ano). Ambas as atividades foram desenvolvidas a partir de histórias de livros infantis. Na primeira história, o personagem recolhia lixo

na cidade, uma tabela de frequência para os três primeiros dias de coleta foi construída. Na atividade, uma coluna para um quarto dia foi deixada em branco, nessa os estudantes (trabalhando sempre em grupos) precisariam prever o número que poderia constar, já que os dados variaram. A segunda atividade foi similar, trata-se da história de um cão e seus cinco amigos, que ao planejarem um piquenique, ficou estabelecido que cada um levaria seis tipos de alimentos. Uma tabela foi construída, partindo-se do número de alimentos para cada tipo diferente observado. Como na primeira atividade descrita, a última coluna foi deixada em branco também, para a previsão de qual número de alimentos uma nova categoria poderia ter.

English (2014) verificou que o contexto das tarefas pode ter interferido nas habilidades dos estudantes em preencher adequadamente a tabela, considerando-se variabilidade e predição. Isso porque na primeira atividade, desenvolvida no primeiro ano, os grupos de crianças (12 de um total de 13) conseguiram distinguir valores prováveis de improváveis, o que se mostrou mais difícil na segunda atividade, realizada no ano seguinte.

No Brasil, Cavalcanti (2011), em estudo diagnóstico com estudantes do 2º ano (entre 7 e 8 anos de idade) e do 5º ano (entre 10 e 11 anos de idade), a respeito da variabilidade de dados, constatou que os sujeitos ao precisarem realizar predição a partir da moda (em situação de interpretação de gráfico de barras) e predição a partir do ponto máximo (em situação de construção e interpretação de gráfico pictórico) apresentaram dificuldades. Na primeira situação, um gráfico com dados sobre o modo que estudantes de uma turma chegaram à escola certo dia (Figura 2.4).

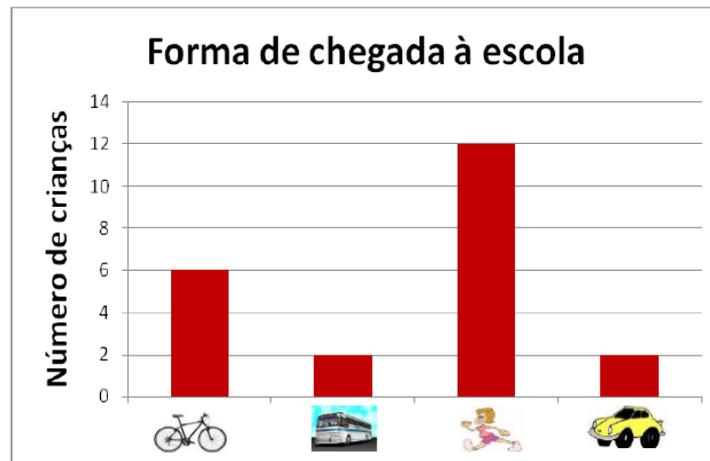


Figura 2.4 – Atividade utilizada por Cavalcanti (2011, p.48) e adaptada de estudo de Watson e Kelly (2002).

Para os dados da Figura 2.4, Cavalcanti (2011) questionou: “*Considerando o gráfico, como você pensa que uma nova criança chegaria à escola?*”. Como resposta adequada, esperava-se que os estudantes reconhecessem que o mais provável seria que a nova criança (que não participou da coleta de dados) fosse andando para a escola, uma vez que é a moda. Entretanto, verificou-se que apenas 16,7% dos estudantes do 5º ano e 4,2% do 2º ano (sem diferença significativa) responderam corretamente, mesmo a maioria dos sujeitos do 5º ano (79,2%) mostrando-se hábeis em reconhecer a moda. Quanto à segunda questão de predição, Cavalcanti (2011) propôs que os estudantes representassem um conjunto de dados com informações sobre livros lidos (Ana leu 4 livros; Daniel leu 1; Luíza leu 6; Bruno leu 3). A partir da construção, questionou: “*Quem você acha que provavelmente gostaria de receber um livro no Natal? Por que você pensa isso?*”. A autora estabelece que o adequado seria os estudantes escolherem aquela que leu mais. Metade dos estudantes do 5º ano assim procederam, com diferença significativa comparada aos estudantes do 2º ano, que obtiveram resultado inferior. Respostas consideradas inadequadas apresentavam escolhas e justificativas baseadas em situações imaginativas e não nos dados, como exemplo de respostas do tipo, tem-se aquelas que afirmavam ser a pessoa que leu menos livros aquela que gostaria de ganhar um livro no Natal, pois iria querer ler mais.

Caseiro, Ponte e Monteiro (2014) realizaram estudo colaborativo, em Portugal, com professora e sua turma, estudantes do 4º ano de escolaridade, no qual fases da pesquisa estatística foram exploradas, desde a elaboração de questão

até as conclusões. A turma já tinha sido instruída a respeito de conceitos estatísticos específicos, como interpretação de representações de conjunto de dados. Para desenvolvimento do estudo, os estudantes formaram grupos de três participantes e escolheram o tema que gostariam de investigar, tais como animais, profissões, *Rally Dakar*, entre outros. Um plano de trabalho foi apresentado aos grupos, constando de: elaboração de questão; como representar graficamente; onde pesquisar; como apresentar. A realização do plano contou com sete momentos, desde a proposta da atividade e motivação até a apresentação das produções dos grupos. Esclarece-se que houve intervenção da pesquisadora, que acompanhava a professora e sua turma, no sentido de explicitar a um grupo específico, a diferença entre gráfico de barras juntas e separadas. Isso porque se observou a utilização de histograma, como representação, apesar do mesmo ser usual apenas em níveis mais avançados de ensino. Na representação em histograma, intervalos de valores referentes a comprimentos de tubarões foram criados. Os estudantes coletaram dados em revistas, livros e internet, assim como a partir de questionários. Constatou-se no estudo que as intervenções da professora concentraram-se na representação dos dados, etapa da pesquisa que a mesma tinha mais familiaridade. A coleta de dados e a formulação de questões (habilidade que gerou mais dificuldades) não contaram com a mesma atenção. A professora reconheceu que foi difícil para ela intervir junto à turma, principalmente na formulação de questões, uma vez que se tratou de uma tarefa nunca realizada.

Caseiro et al (2014) afirmam que para o processo de análise e conclusões, os estudantes utilizaram o *software Excel*, a partir do qual gráficos foram gerados com a inserção de tabelas de frequência pelos grupos. Os autores destacam que essas últimas fases da pesquisa requeriam mais tempo para desenvolvimento. Apesar disso, a professora realizou intervenções com questionamentos ou afirmações, partindo das representações utilizadas pelos grupos, como mostra a Figura 2.5, a seguir:

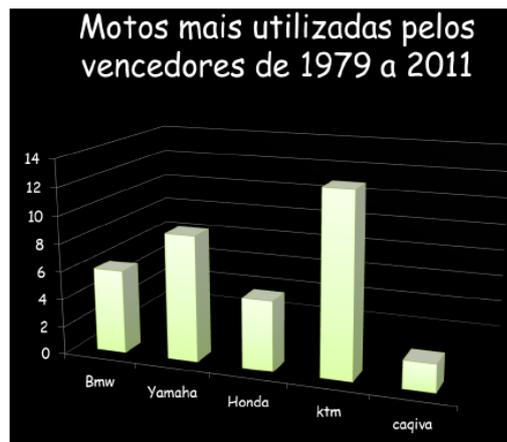


Figura 5 – Gráfico construído pelo grupo cujo estudo era sobre o Dakar

Professora: Então se quisermos ganhar tem de ser com uma ktm?

Miguel: A ktm costuma ser a que quase todos os vencedores usam. É quase sempre a ktm que ganha.

Figura 2.5 – Exemplo de intervenção de professora para realização de conclusão, em Caseiro et al (2014, p.250).

Na Figura 2.5, percebe-se a intervenção da professora, no sentido de estimular a realização de conclusões a partir dos dados. Cabe salientar, nesse momento, que apesar da preocupação no estudo de Caseiro et al (2014) em se trabalhar com o ciclo investigativo proposto por Wild e Pfannkuch (1999), discussões a respeito do levantamento de hipóteses não foram evidenciadas, nem antes da coleta de dados, para confrontação, tampouco depois.

Jacobbe, Foti, Whitaker (2014) utilizaram instrumento de avaliação para medir o conhecimento de estudantes (entre 10 e 13 anos de idade) nos Estados Unidos, a respeito das quatro áreas da pesquisa estatística, propostas pelo *GAISE* (2007): formulação de questão; coleta de dados, análise de dados; e interpretação de resultados. O instrumento de avaliação denominou-se LOCUS e foi desenvolvido a partir do *Evidence Centered Design (ECD)*, de Mislevy e Riconscente (2006), de acordo com os autores. Um teste piloto foi aplicado com 2.075 (dois mil e setenta e cinco), cujas respostas foram niveladas (A, B e C). Jacobbe et al (2014) constataram que, nas questões de múltipla escolha, os estudantes obtiveram um desempenho melhor na formulação de questões estatísticas. O item incluía determinação de natureza do problema, se estatístico ou não, assim como decisão sobre qual dado seria importante na fase de análise e qual inferência poderia ser tirada dos dados. Ainda nas questões de múltipla escolha, o item que gerou maiores dificuldades foi o que requeria análise dos dados. Contudo, os autores ressaltam que o LOCUS não

explorou as típicas questões categorizadas como análise de dados, as quais focam na realização de cálculos, ao contrário, foi exigido dos estudantes o entendimento dos dados, conforme exemplificação (Figura 2.6), a seguir:

Uma escola está planejando uma viagem de estudo para o aquário ou o zoológico por estudantes dos Graus 6 ao 9. Há 100 estudantes em cada nível e cada estudante foi questionado sobre o lugar que ele ou ela preferiria visitar. O gráfico de barras para os quatro níveis é apresentado abaixo.

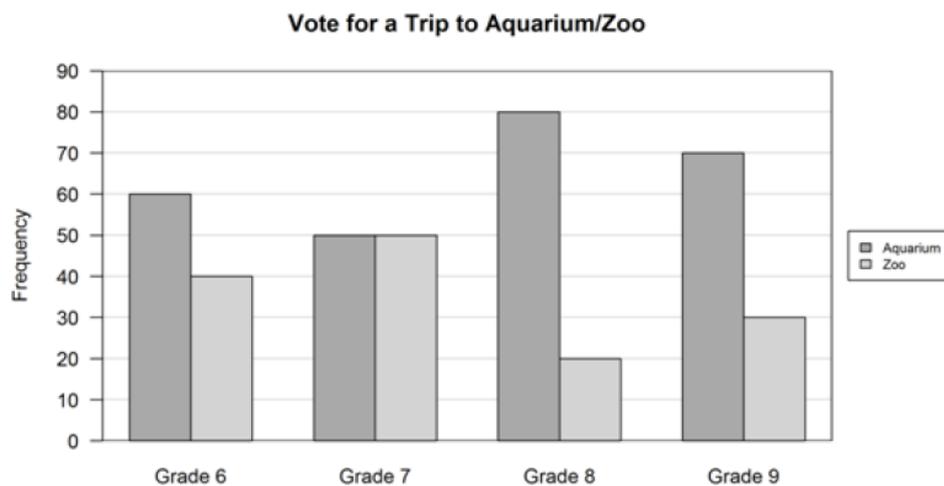


Figura 2.6 – Exemplo de questão (múltipla escolha) de análise de dados em Jacobbe et al (2014, p.5).

Para os dados apresentados na Figura 1.6, questionou-se: “*Em que nível de escolaridade as respostas foram mais consistentes?*”. Como opções de respostas, expôs-se: (A) *Grade 6*; (B) *Grade 7*; (C) *Grade 8*; e (D) *Grade 9*. A resposta “C” é a adequada. Entretanto, a maioria dos estudantes escolheu a alternativa “B”, o que de acordo com os autores revela equívoco na compreensão gráfica. Já nas questões de respostas construídas pelos próprios estudantes, Jacobbe et al (2014) afirmam que mais dificuldades foram verificadas em todas as fases da pesquisa. Entretanto, formulação de questão continuou apresentando um desempenho melhor que as demais fases (coleta, análise de dados e interpretação de resultados), por parte dos estudantes investigados. Interpretação de resultados, nesse momento, foi o item que gerou mais dificuldades, dentre as etapas da pesquisa propostas pelo *GAISE*.

Cabe esclarecer, neste momento, que a análise requerida por Jacobbe et al (2014) no gráfico (Figura 2.7) exigia análise da variabilidade dos dados em situação

de comparação, conforme Cavalcanti (2011). Compreender a variabilidade no tipo de dado explorado (dados qualitativos) demanda perceber que “*respostas mais consistentes*”, de acordo com o comando, significa menor variabilidade na escolha entre aquário e zoológico. Portanto, menor variabilidade se aproxima de ausência de variabilidade, que seria mesma resposta por todos. Nesse sentido, a alternativa “C” é aquela em que o tipo de escolha feita entre as duas opções contou com um número maior de pessoas fazendo a mesma escolha. A alternativa “B”, por outro lado, é aquela em que a variabilidade foi maior, uma vez que metade da turma fez escolha diferente. Assim como Jacobbe et al (2014), Cavalcanti (2011) constatou em seu estudo que esse tipo de atividade foi bastante difícil para estudantes dos anos iniciais. A autora em seu estudo não utilizou gráfico, o que revela que a representação não foi um fator determinante para as dificuldades verificadas entre os estudantes.

Makar (2013) realizou pesquisa com professora e estudantes dos anos finais do Ensino Fundamental, a respeito de previsões advindas de experiências investigativas em sala de aula. A professora solicitou que os sujeitos, trabalhando em pequenos grupos, respondessem à questão: “*Qual é o melhor modelo para um helicóptero de papel e canudo?*”. As estruturas criadas foram denominadas no estudo de *loopy airplanes*. Após discussões com a turma sobre o que o melhor modelo significaria, foram criadas duas estruturas (Figura 2.7), uma menor (fita de 5 cm de comprimento) e outra maior (fita de 8 cm de comprimento), que foram testadas em seguida. Medições das distâncias alcançadas pelos modelos foram feitas, como modo de se coletar dados. Para organização e análise, foi utilizado no estudo o *TinkerPlots Dynamic Data Exploration*. O software *Tinkerplots*, criado por Konold e Miller (2005), fornece um ambiente de aprendizagem em que estudantes do Ensino Fundamental podem desenvolver o entendimento de conceitos estatísticos a partir da análise de dados, de acordo com Konold (2006).



Figura 2.7 – Modelos de *loopy airplanes* (MAKAR, 2013).

Makar (2013) obteve como resultados, que os estudantes não apenas descreveram os dados dos voos, eles realizaram uma conclusão geral, que foi além dos dados. Isso porque inferiram que o voo do helicóptero menor chegou a uma distância maior e que se repetissem o teste chegariam à mesma conclusão. Os estudantes foram solicitados a realizarem novos testes com modelos diferentes, a partir de outras variáveis, como: melhor largura da asa; melhor tamanho da asa (comprimento da fita); melhor localização da asa. A autora ressalta que a estatística descritiva foi utilizada como evidências para as predições realizadas pelos estudantes, os quais perceberam que não poderiam generalizar os resultados alcançados em um caso para outras situações.

Allmond e Makar (2014) utilizaram atividade similar a Makar (2013), em se tratando do uso de helicóptero de papel, para avaliar habilidades de análise de dados e conclusão a partir dos mesmos. Na pesquisa de Allmond e Makar (2014), participaram 27 (vinte e sete) estudantes de 12 anos de idade. Foi solicitado aos sujeitos que investigassem o tempo que levaria um helicóptero de papel para cair. Durante a análise, o *software Tinkerplots* foi empregado e cada sujeito portava um computador. Os estudantes perceberam que quando a altura (de dois metros) foi fixada, a variabilidade de dados foi menor, diferente de quando a altura variou e foi desconhecida. 83% das respostas consideraram a variabilidade dos dados na análise, dentre essas, 40% explicitaram apenas o intervalo de tempo que o objeto levou para cair nos dois contextos (altura fixada e não fixada), 13% citaram, além do intervalo de tempo, a amplitude ser maior ou menor, dependendo da situação e 30% fizeram referência ao intervalo de tempo e expuseram a amplitude em termos quantitativos.

Em se tratando das conclusões elaboradas pelos estudantes, de acordo com os autores, 87% dos sujeitos demonstraram alguma evidência de seu entendimento. Dos que conseguiram tirar conclusões, 48% dos sujeitos focaram no intervalo estimado de tempo que o helicóptero levou para cair e 39% ofereceu um ponto (tempo) estimado de queda.

Henriques e Oliveira (2015) desenvolveram um estudo em que 30 (trinta) estudantes (entre 13 e 14 anos de idade) de Portugal conduziram uma pesquisa estatística, sendo proposta aos mesmos a realização de inferências informais a partir dos dados. Antes da atividade investigativa, os sujeitos contaram com tarefas preliminares de exploração de dados reais no *software Tinkerplots*. O tema de investigação dos estudantes foi o corpo humano, especificamente medidas vitruvianas, como altura, tamanho do pé e envergadura. Os estudantes foram estimulados a utilizarem dados coletados na turma para realizarem predições para a população da escola que faziam parte.

As análises das habilidades dos estudantes por Henriques e Oliveira (2015), foram estruturadas a partir dos três princípios fundamentais da Inferência Estatística Informal, propostos por Makar e Rubin (2009). Para o princípio Generalização além dos dados, Henriques e Oliveira (2015) afirmam que o mesmo emergiu em momentos diferentes das tarefas desenvolvidas pelos estudantes. Assim, o professor iniciou com a questão: “*Como você caracterizaria o estudante do ensino secundário (Grau 7, 8 e 9) de sua escola em relação a algumas das medidas de Vitruvius, como altura, tamanho do pé e envergadura?*”. As autoras relatam que os estudantes demonstraram entender a importância da amostra para a inferência, pois afirmaram que poderiam selecionar aleatoriamente estudantes do 7º ao 9º ano, embora não tenham sido unânimes na forma de se coletar a amostra. Apesar dessa compreensão, todos os dados foram coletados na própria turma, reconhecendo-se a falta de representatividade da amostra para se fazer generalizações. Quanto ao segundo princípio, Dados como evidência, Henriques e Oliveira (2015) constataram que os estudantes se mostraram hábeis em criar uma variedade de configurações gráficas para representar adequadamente os dados coletados. Na comparação entre meninos e meninas, por exemplo, houve utilização de *Boxplots* e afirmações que demonstraram a percepção da diferença nas medidas, considerando-se o gênero. O último princípio, Linguagem probabilística, revelou-se o mais difícil. Apesar disso, foi

possível encontrar inclusão de termos como “talvez”, “provavelmente” ou “tende a ser” nas respostas dos estudantes, o que indica reconhecimento da incerteza.

Os estudos até o momento descritos analisaram compreensões e/ou aprendizagens de estudantes do Ensino Fundamental. Se estudos diagnósticos, como os de Cavalcanti (2011) e Jacobbe et al (2014), constataram dificuldades dos estudantes (em torno de dez anos de idade e mais) com a interpretação dos dados para realização de conclusões, como predições ou inferências informais, outros estudos, preocupados com o desenvolvimento do letramento estatístico, já desde os primeiros anos de escolaridade, conseguiram avanços nas compreensões dos estudantes. Crianças desde os cinco anos de idade demonstraram habilidade em considerar a incerteza, essencial a situações de predição (MCPHEE; MAKAR, 2014), além de mostrarem-se capazes de desenvolver fases do ciclo investigativo. O nível de complexidade das atividades mudou com o aumento da escolaridade, o que é necessário, ainda assim, resultados positivos foram ressaltados pelos diferentes autores quanto às aprendizagens verificadas. Limites também foram destacados, o que requer novos olhares para que os mesmos sejam superados.

As dificuldades percebidas pelos pesquisadores, em momentos de intervenção, durante a revisão da literatura, fez com que os mesmos levantassem proposições, dentre as quais se destacam: interferência, na aprendizagem, do contexto das atividades desenvolvidas (ENGLISH, 2014); baixa demonstração de predições baseadas em evidências quando as crianças são muito jovens (MCPHEE; MAKAR, 2014); necessidade de se intervir adequadamente junto aos estudantes em todas as fases da pesquisa (CASEIRO et al, 2014); necessidade de investimentos na utilização de linguagem adequada quando se explora situações de incerteza (HENRIQUES; OLIVEIRA, 2015); e reconhecimento da variabilidade na análise dos dados, como fator importante na explicitação de conclusões (ALLMOND; MAKAR, (2014).

2.6.2 Compreensões e aprendizagens em distribuições bivariadas

Os estudos descritos, a seguir, abordam pesquisas que buscaram compreender e desenvolver aspectos do raciocínio correlacional junto a estudantes do Ensino Fundamental.

Nunes (2015) realizou estudo experimental com crianças de 10 anos de idade, utilizando diferentes atividades para o ensino de probabilidade. Entende-se que a compreensão de probabilidade está relacionada com o raciocínio correlacional. Desse modo a autora também abordou em sua pesquisa a relação entre duas variáveis. Na pesquisa definiram-se três grupos: grupo controle, com 29 (vinte e nove) sujeitos; grupo de resolução de problemas, 23 (vinte e três) sujeitos; e grupo de intervenção com ensino de probabilidade, que contou com 24 (vinte e quatro) sujeitos. O grupo de intervenção, com ensino de probabilidade, participou de atividades envolvendo aleatoriedade, espaço amostral, quantificação da probabilidade e associação entre variáveis. A atividade de associação entre variáveis encontra-se em destaque na Figura 2.8.

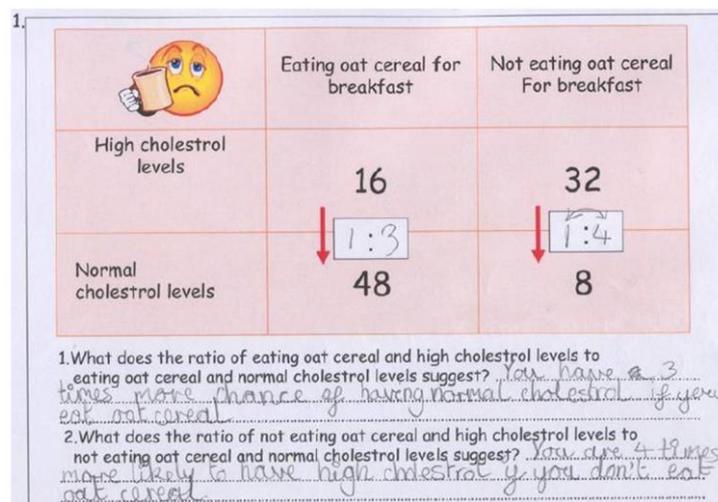


Figura 2.8 – Atividade de associação entre variáveis, Nunes (2015).

Na Figura 2.8, percebe-se que a configuração tabular foi apresentada e os estudantes foram questionados, primeiramente sobre “O que a razão de comer cereais de aveia e níveis elevados de colesterol, assim como comer cereais de aveia e níveis normais de colesterol sugere?”. Depois sobre “O que a razão de não comer cereais de aveia e níveis elevados de colesterol, assim como não comer cereais de aveia e níveis normais de colesterol sugere?”. Após dois pós-testes com os estudantes e devidas análises de covariância, Nunes (2015) constatou melhoria significativa das compreensões dos estudantes em se tratando de associação entre variáveis e probabilidade nas demais situações. A autora percebeu, ainda, que os

estudantes que participaram do grupo de resolução de problemas não apresentaram um efeito positivo em probabilidade.

Fitzallen e Watson (2014) investigaram a habilidade de estudantes, entre 11 e 12 anos de idade, em realizar inferências informais a partir de associações entre variáveis relacionadas com atributos físicos, como altura, tamanho do pé, altura até o umbigo e extensão da mão. O estudo aconteceu em seis semanas, com dois encontros semanais. O *software Tinkerplots* foi utilizado nesse período, sem qualquer instrução prévia, como recurso para representações gráficas por parte dos estudantes, assim como para explicações a respeito das relações entre variáveis encontradas. As autoras constataram que todos os estudantes se mostraram hábeis em descrever os dados em termos de relação entre variáveis e fazer inferências informais sobre os dados a partir das evidências mostradas pelas representações gráficas criadas.

Três níveis de entendimento de covariação foram determinados pela taxonomia *SOLO (Structure of Observed Learning Outcomes)*. No nível uni-estrutural, os estudantes identificaram corretamente quando houve uma tendência evidente em um gráfico, contudo, descreveram uma variável de cada vez e não utilizaram as informações representadas para apoiar as suas conclusões sobre os dados. No nível multi-estrutural, os sujeitos identificaram a tendência dos dados observados e a variação dentro da tendência. No entanto, não foram capazes de explicar como a variação identificada relacionava-se com as suas conjecturas gerais e inferências. No nível relacional, os estudantes descreveram a relação entre dois atributos em ambas as perspectivas, local e global. Tratando-se do nível mais avançado, esses sujeitos conseguiram reunir informações sobre pontos específicos dos dados, variação dentro dos gráficos e da tendência identificada, além de seus conhecimentos do contexto para justificar inferências informais realizadas por eles.

Os estudos de Nunes (2015) e de Fitzallen e Watson (2014) abordaram correlação entre duas variáveis, conceito dificilmente explorado no Ensino Fundamental, mas possível de ser apreendido nessa fase da escolarização, como constatado nessas pesquisas. Nunes e Bryant (2011), discutidos anteriormente, já salientaram que organizar os dados para a análise de correlações poderia ser um dificultador. Nesse sentido, Nunes (2015) explorou associação entre variáveis com dados já organizados em tabela e obteve resultados positivos junto aos estudantes participantes de seu estudo. No estudo de Fitzallen e Watson (2014) os estudantes

organizaram os dados em gráficos. Entretanto, essa pesquisa apresentou características bem distintas da pesquisa de Nunes (2015). Acredita-se que pensar em atributos físicos, a partir da medida de comprimentos do corpo humano (FITZALLEN; WATSON, 2014), já supõe ideias intuitivas de correlação, bem diferente de se pensar em dois acontecimentos de contextualização distinta, como as variáveis alimentação e saúde (NUNES, 2015) e a partir de então analisar dados de associação. Outra diferença é que análises estatísticas do desempenho dos estudantes não foi o foco do estudo de Fitzallen e Watson (2014), o mesmo concentrou-se em revelar níveis de entendimento de relação entre variáveis e conclusões.

3 MÉTODO

3.1 OBJETIVO GERAL

Esta pesquisa tem como objetivo analisar aprendizagens apresentadas por estudantes do Ensino Fundamental relacionadas ao levantamento de hipóteses, à análise de dados e às conclusões a partir dos dados, enquanto fases do ciclo investigativo.

3.2 OBJETIVOS ESPECÍFICOS

- ❖ Analisar os conhecimentos demonstrados por estudantes do Ensino Fundamental (5º e 7º anos), diante de distribuições univariadas e bivariadas, quando levantam hipóteses, interpretam dados reais, confrontam hipóteses e dados, avaliam conclusões e fazem previsões;
- ❖ Analisar as aprendizagens de estudantes do Ensino Fundamental (5º e 7º anos) em estudo experimental de intervenção, diante de distribuições univariadas e bivariadas, quando levantam hipóteses, interpretam dados reais, confrontam hipóteses e dados, avaliam conclusões e fazem previsões.

Os objetivos específicos elencados abordam diferentes habilidades necessárias quando se trabalha com fases da pesquisa estatística, tendo-se em vista o desenvolvimento do Letramento Estatístico nas práticas de sala de aula. Refletir sobre os dados reais divulgados em meios diversos e ser capaz de apresentar uma postura crítica diante dos mesmos são atitudes desejadas.

3.3 HIPÓTESE

O ensino voltado para o desenvolvimento da capacidade de formular hipóteses e analisar dados estatísticos em diferentes distribuições (uni e bivariadas) favorece a tomada de decisões baseadas em evidências, pois permite pensar fases da pesquisa estatística de modo inter-relacionado.

3.4 SUJEITOS

Os participantes do estudo são estudantes do 5º e 7º ano do Ensino Fundamental, de escolas públicas da região metropolitana de Recife. A abordagem de diferentes tipos de distribuição (univariada e bivariada) de dados, que podem interferir na análise que se faz dos mesmos corroborou para a escolha do público alvo da pesquisa.

O estudo de Nunes e Bryant (2011) alerta para as demandas do raciocínio correlacional, que surgem do entendimento de dois esquemas, a probabilidade e a proporcionalidade. No que diz respeito à probabilidade, Silva (2016) ao investigar crianças do 1º, 3º e 5º ano do Ensino Fundamental percebeu que elas, de modo geral, apresentaram compreensões intuitivas e nem sempre coerentes das demandas cognitivas postas por Nunes e Bryant (2011), para a aprendizagem de probabilidade. Contudo, constatou que as crianças mais velhas apresentaram justificativas mais consistentes. É importante esclarecer que a realização de cálculos relacionados com o raciocínio correlacional, a partir de probabilidades, não está entre as habilidades investigadas na tese e não é a intenção verificar se os estudantes do 7º ano compreendem e avançam mais diante do conceito de probabilidade, intrínseco às situações de correlação, o que se deseja analisar é a argumentação utilizada pelos estudantes, tanto do 5º ano quanto do 7º ao responderem a questões pensadas a partir de dados univariados e dados bivariados (correlação).

Em se tratando da linguagem probabilística em situações de incerteza, Henriques e Oliveira (2014) já perceberam que a habilidade requeria maiores investimentos no que se refere à instrução. Quanto à proporcionalidade, Soares e Nehring (2013) afirmam que o conceito se articula com diversas situações do cotidiano e com vários conceitos dentro da própria Matemática. Porém, o ensino restrito a situações de aplicação de procedimento algorítmico (regra de três) e reduzido geralmente ao 7º ano do Ensino Fundamental não prioriza a compreensão.

Nesse sentido, acredita-se que as ideias de probabilidade e proporcionalidade precisam ser consideradas neste estudo, atentando-se para as potencialidades e limites já ressaltados em estudos anteriores, quando se pensa nas compreensões já demonstradas por estudantes do Ensino Fundamental (no caso da probabilidade) e

na abordagem didática estabelecida diante de conceitos matemáticos (no caso da proporcionalidade).

3.5 PROCEDIMENTOS

Dois estudos foram pensados como ponto de partida para a pesquisa. No primeiro momento (Estudo 1), um teste diagnóstico foi estruturado, com a intenção de se obter maiores informações a respeito das compreensões dos estudantes do 5º e 7º anos em relação ao levantamento de hipóteses, à análise dos dados em diferentes distribuições (univariadas e bivariadas) e às conclusões advindas após conhecimento dos dados. Investigou-se também como essas diferentes fases da pesquisa se relacionaram. Já o segundo momento (Estudo 2), caracteriza-se por um estudo intervencionista, no qual as diferentes habilidades relacionadas às fases do ciclo investigativo explicitado são encaminhadas em um processo com intensão de ensino das mesmas.

3.5.1 Estudo 1

Participaram do Estudo 1 duas turmas, uma do 5º ano (22 estudantes) e uma do 7º ano (25 estudantes) do Ensino Fundamental. Ambas as turmas eram de escolas públicas municipais da região metropolitana de Recife. O período em que ocorreu a coleta dos dados foi na segunda metade do último semestre do ano letivo. Os estudantes, portanto, já tinham sido instruídos a respeito da maior parte dos conteúdos propostos para os anos de escolaridade que faziam parte.

O Estudo 1 contou com um teste contendo 4 (quatro) atividades. As atividades foram elaboradas pensando-se no tipo de questão de pesquisa que conduziria o levantamento de hipóteses. Assim, as duas primeiras atividades partiram de questões gerais, que gerariam uma distribuição univariada dos dados e, a depender das experiências de vida ou crenças do sujeito, poderiam suscitar concordância com os dados reais a serem apresentados após registro das hipóteses ou discordância. Já as duas últimas questões foram estruturadas de modo a fazer com que o sujeito levantasse hipóteses a partir da avaliação de existência ou não de relação entre duas variáveis. Os dados reais apresentados após registro de

hipóteses continham na terceira atividade uma relação direta (positiva) e na quarta atividade uma relação indireta (negativa). O Quadro 3.1 exibe as quatro questões discutidas.

Quadro 3.1 – Tipos de questões propostas no Estudo 1

Tipo de questão	Exemplo
Questão geral	<u>Atividade 1</u> : Em quais momentos do dia você acredita que as pessoas mais utilizam a internet dos celulares? Por quê?
	<u>Atividade 2</u> : Para você a água consumida no Brasil é usada em maior quantidade onde (casas, indústrias, plantações, criação de animais)? Explique.
Questão de relação entre variáveis	<u>Atividade 3</u> (geração de relação direta): Para você o que o câncer de pulmão tem a ver com o hábito de fumar cigarros, cachimbos, charutos e outros (produtos com nicotina)? Explique.
	<u>Atividade 4</u> (geração de relação indireta): Você acredita que a quantidade de água no corpo humano tem relação com a idade das pessoas? Por quê?

Após solicitação de registro de hipóteses, dados reais referentes a cada questão foram apresentados em gráficos de barras, tabela ou banco de dados. Todas as atividades exploraram interpretação dos dados, confrontação entre hipóteses levantadas e análise dos dados. Avaliação de conclusão foi requerida após confrontação entre hipóteses e dados nas três primeiras atividades. Apenas as duas primeiras atividades abordaram questões de predição como últimas questões para cada atividade.

Em todas as atividades os estudantes contaram apenas com seus conhecimentos prévios para elaboração das hipóteses, dados não foram apresentados aos mesmos, por isso uma folha foi entregue em separado. É importante salientar que na formulação de uma hipótese o estudante ou outra pessoa que pratique tal ação precisa parar e refletir sobre suas próprias ideias a respeito de um tema, a resposta está no campo da imaginação e incerteza.

Após registro de hipóteses, outra folha contendo dados representados e demais questões foi proposta. As quatro atividades são detalhadas a seguir.

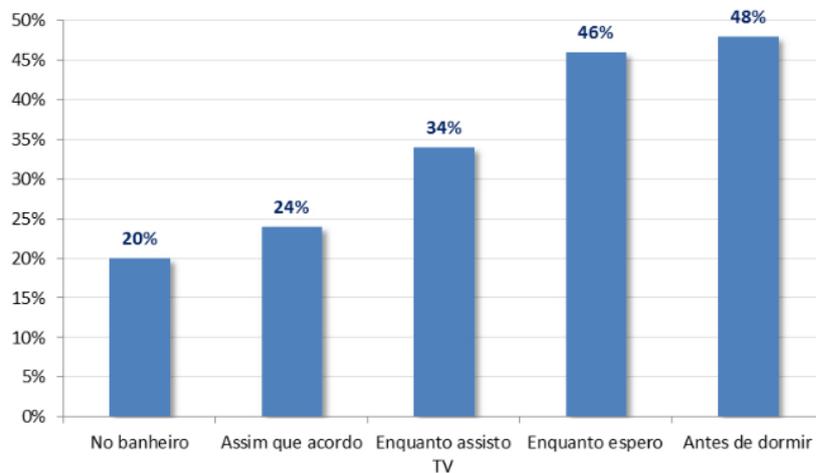
Atividade 1

A Atividade 1, como explicitado no Quadro 3.1, parte de uma questão geral de pesquisa e aborda cinco questões. Para análise dos dados, representação em gráfico de barras é explorada para o conjunto de dados univariados.

Q1: Em quais momentos do dia você acredita que as pessoas mais utilizam a internet dos celulares? Por quê? (*Levantamento de hipótese*)

Figura: Maiores momentos de uso de uso do celular com internet em 2015.

Os maiores momentos de uso do smartphone com internet - Brasil - julho de 2015



Quais os três momentos em que você mais usa a internet do smartphone?

Fonte: Mobile Report- Julho/15 - Nielsen IBOPE

Q2: Como as pessoas têm utilizado a internet de seus celulares de acordo com o gráfico? Explique. (*Interpretação*)

Q3: Os dados do gráfico coincidem com o que você imaginou? Por quê? (*Hipótese em confronto*)

Q4: É possível afirmar que as pessoas usam mais a internet do celular em momentos de trabalho? Por quê? (*Conclusão*)

Q5: Você acredita que os momentos em que as pessoas mais usam a internet no celular podem ser diferentes em 2016? Por quê? (*Predição*)

A primeira questão da atividade tem por objetivo fazer com que os estudantes levantem hipóteses para a questão de pesquisa posta, a partir de seus conhecimentos prévios. A segunda questão requer explicitação da interpretação dos dados expostos. Já a terceira, solicita que os estudantes relacionem as hipóteses registradas com a análise dos dados, a intenção é verificar se os mesmos analisam

suas conjecturas em função dos dados recém-descobertos. A quarta questão objetiva verificar acordos ou desacordos dos estudantes ao avaliarem uma conclusão, os dados precisam ser considerados para isso. A última questão, por fim, investiga a linguagem utilizada pelos estudantes ao fazer previsões sobre os dados. A situação de incerteza demanda linguagem probabilística.

Atividade 2

A Atividade 2 também contou com cinco questões e explorou as mesmas habilidades da atividade anterior. Diferente da Atividade 1, os dados reais para análise e confrontação foram apresentados em tabela.

Q1: Para você a água consumida no Brasil é usada em maior quantidade onde (casas, indústrias, plantações, criação de animais)? Explique. (*Levantamento de hipótese*)

Tabela: Consumo da água em diferentes setores

Uso da água	Percentual (%)
Irrigação de plantações	72
Indústria	7
Criação animal	11
Abastecimento urbano e rural	10

Fonte: Conjuntura de Recursos Hídricos do Brasil, ANA, 2013.

Q2: Como tem sido a distribuição do uso da água em diferentes setores do país, de acordo com a tabela? Explique. (*Interpretação*)

Q3: Os dados da tabela coincidem com o que você imaginou? Por quê? (*Hipótese em confronto*)

Q4: De acordo com a tabela, é possível afirmar que economizar água em casa resolve o problema de falta de água? Por quê? (*Conclusão*)

Q5: Os números da tabela sobre o uso da água podem ser muito diferentes em 2016? (*Predição*)

Na Atividade 2, a questão proposta para avaliar a habilidade dos estudantes de levantar hipótese direciona escolhas que podem ser realizadas pelos sujeitos, diferente da atividade anterior. A segunda questão busca explicitação da

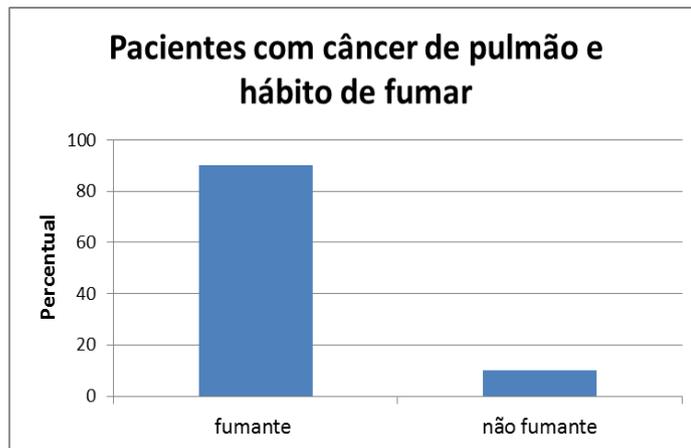
interpretação dos dados expostos na tabela. Já a questão três requer um confronto entre hipótese criada e retomada da mesma após conhecimento dos dados. A questão quatro demanda avaliação de conclusão. Já a última questão solicita predição a partir dos dados, cujo registro será o mote para análise da linguagem (probabilística ou não) utilizada pelo estudante.

Atividade 3

Na Atividade 3, quatro questões são exploradas. A questão de pesquisa proposta busca avaliação por parte do estudante da possibilidade de relação entre duas variáveis. Nesse sentido, as hipóteses levantadas partem de outro contexto, o que possivelmente acarretam compreensões distintas das observadas nas Atividades 1 e 2.

Q1: Para você o que o câncer de pulmão tem a ver com o hábito de fumar cigarros, cachimbos, charutos e outros (produtos com nicotina)? Explique. (*Levantamento de hipóteses*)

Figura: Percentual de pacientes com câncer de pulmão e hábito de fumar tabaco



Fonte: Instituto Nacional de Câncer, INCA.

Q2: Como é a relação entre os pacientes com câncer de pulmão e o hábito de fumar ou não? Explique. (*Interpretação*)

Q3: Os dados do gráfico coincidem com o que você imaginou? Por quê? (*Hipótese em confronto*)

Q4: De acordo com o gráfico, pode-se dizer que o risco de fumar e ter câncer de pulmão é pequeno? (*Conclusão*)

A primeira questão da Atividade 3 avalia a habilidade em levantar hipóteses a partir da possível relação entre duas variáveis. Na segunda questão, a relação entre as variáveis câncer de pulmão e hábito de fumar tabaco é apresentada em um gráfico de barras, numa associação direta entre ambas as variáveis, o que requer interpretação dos dados pelos estudantes. A terceira questão solicita confronto entre hipótese levantada antes do conhecimento dos dados e após análise. A última questão apresenta uma conclusão, que precisa ser avaliada, tomando-se os dados como referência.

Atividade 4

A Atividade 4 aborda três questões, as quais correspondem as três primeiras habilidades exploradas nas atividades anteriores. Na Atividade 4 houve distinção no tipo de representação¹ de dados e no tipo de relação entre variáveis. Nesse caso, a representação apresentada aos estudantes mostra uma relação indireta em um banco de dados.

Q1: Você acredita que a quantidade de água no corpo humano tem relação com a idade das pessoas? Por quê? (*Levantamento de hipótese*)

Tabela: Percentual de água no corpo humano com o aumento da idade

Idade (ser humano)	Percentual de água no corpo (%)
0 a menos de 2 anos	75 a 80
2 a menos de 5 anos	70 a 75
5 a menos de 10 anos	65 a 70
10 a menos de 15 anos	63 a 65
15 a menos de 20 anos	60 a 63
20 a menos de 40 anos	58 a 60

¹ Os dados foram apresentados aos estudantes como representação em tabela. Contudo, já há discussões a respeito da distinção entre tabela, banco de dados e quadro, conforme estudos de Guimarães e Oliveira (2014) e Evangelista e Guimarães (2017), constituindo-se, dessa forma, a representação proposta em um banco de dados.

40 a menos de 60 anos	50 a 58
60 anos acima	menos de 50

Fonte: <http://slideplayer.com.br/slide/2879216/>

Q2: O que acontece com a quantidade de água no corpo com o passar dos anos?

Explique. (*Interpretação*)

Q3: Os dados da tabela confirmam o que você pensou? Por quê? (*Hipótese em confronto*)

A primeira questão aborda o levantamento de hipótese a partir da avaliação da relação entre duas variáveis. Já a segunda questão apresenta dados em uma relação indireta para interpretação pelos estudantes. A terceira questão requer confronto entre hipóteses levantadas antes de conhecimento dos dados e após solicitação de análise dos mesmos.

3.5.2 Estudo 2

O Estudo 2 possui um caráter experimental. Para tal, um teste inicial (pré-teste), seguido de encontros de intervenção e teste final (pós-teste) foram elaborados. Cada um dos anos de escolaridade investigados contou com três turmas, uma que constituiu o grupo controle e outras duas que passaram pelo processo de intervenção, o qual foi o mesmo modelo para todos os estudantes. Isso totalizou 6 (seis) turmas do Ensino Fundamental, 3 (três) de 5º ano e 3 (três) de 7º ano. Esclarecendo-se que se tratam de turmas diferentes daquelas do Estudo 1. As intervenções didáticas junto às turmas contaram com 3 (três) encontros, de aproximadamente duas horas cada. Os encontros foram semanais e as aulas áudio-gravadas.

Importante ressaltar que se optou no Estudo 2 pela exposição de informações estatísticas apenas em gráficos de barras, diferente do que aconteceu no Estudo 1. Acredita-se que as habilidades investigadas e a inter-relação entre as mesmas são muitas, mantendo-se um mesmo tipo de representação, principalmente uma mais explorada em situações escolares, o foco da investigação esteja direcionado para as habilidades elencadas na tese.

3.5.2.1 Atividades do Pré-teste

Para o teste inicial (pré-teste), dois tipos de testes foram construídos. Os quais foram similares em se tratando de habilidades investigadas e tipo de dados, se univariados ou bivariados. A utilização de dois testes atendeu ao desejo de se minimizar o efeito do cansaço durante a realização do teste. Assim, duas atividades constituíram cada tipo de teste, ambas exploraram gráficos de barras nas análises e se distinguiram, principalmente, pelo tipo de dado real explorado: Na primeira atividade (Atividade 1) as questões foram criadas a partir de dados univariados; na segunda atividade (Atividade 2) as questões foram elaboradas a partir de dados bivariados/correlação. Ambos os tipos de testes estiveram presentes em todas as turmas investigadas no Estudo 2. Metade da turma realizou um tipo de teste e a outra metade realizou o outro tipo. Tais testes são denominados aqui de Teste A e Teste B e variaram sobretudo no contexto ou temática sobre o qual foram produzidos, enquanto pesquisa real. A seguir, os mesmos são descritos.

Teste A

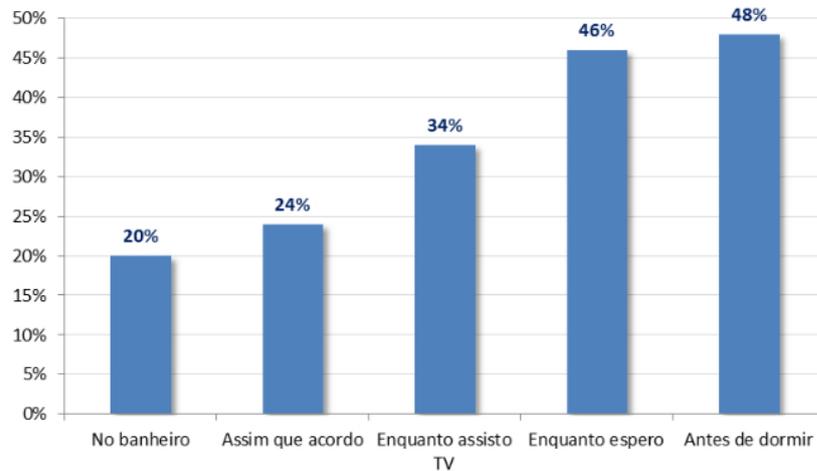
Atividade 1A

A Atividade 1A é similar à primeira atividade do Estudo 1, referente aos momentos de uso da internet dos celulares por brasileiros. Ajustes na linguagem das questões propostas aos estudantes foram realizados, a fim de se garantir a menor interferência possível da pesquisadora durante a aplicação dos testes, com a explicitação de qualquer termo desconhecido, por exemplo.

Q1: Em quais momentos do dia você acredita que as pessoas mais utilizam a internet dos celulares? Por quê? (*Levantamento de hipótese*)

Maiores momentos de uso do celular com internet em 2015

Os maiores momentos de uso do smartphone com internet - Brasil - julho de 2015



Quais os três momentos em que você mais usa a internet do smartphone?

Fonte: Mobile Report- Julho/15 - Nielsen IBOPE

Q2: De acordo com o gráfico, quando as pessoas têm utilizado a internet de seus celulares? Explique. (*Interpretação*)

Q3: Os dados do gráfico apresentam informações parecidas com as que você imaginou na primeira questão? Por quê? (*Confronto entre hipótese e dados*)

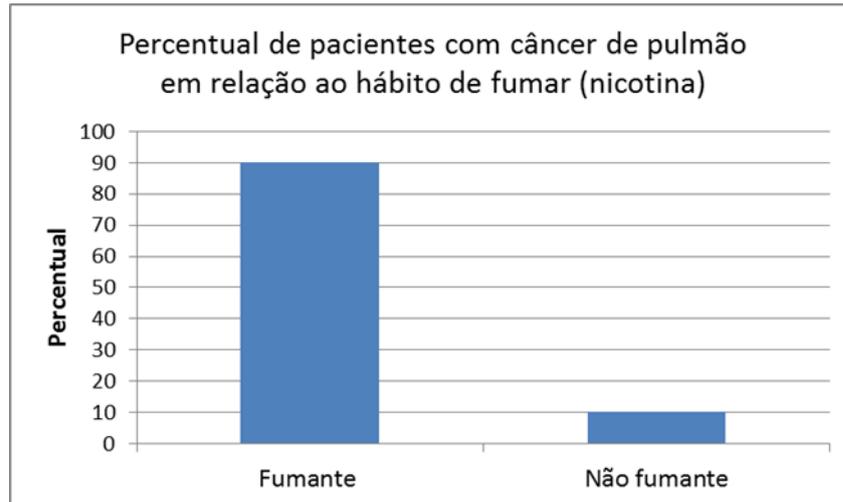
Q4: A partir do gráfico, é possível afirmar que as pessoas usam mais a internet do celular em momentos de trabalho? Por quê? (*Avaliação de conclusão*)

Q5: Você acredita que os momentos em que as pessoas mais usam a internet do celular podem ser diferentes em 2017? Por quê? (*Predição*)

Atividade 2A

A atividade 2A é similar à terceira atividade do Estudo 1, a qual se refere à relação entre ter câncer de pulmão e hábito de fumar produtos a base de tabaco. Ajustes também foram necessários na linguagem de algumas questões propostas.

CQ1: Para você o que o câncer de pulmão tem a ver com o hábito de fumar cigarros, cachimbos, charutos e outros (produtos com nicotina)? Explique. (*Levantamento de hipótese*)



Fonte: Instituto Nacional de Câncer, INCA.

CQ2: De acordo com o gráfico, existe relação entre os pacientes com câncer de pulmão e o hábito de fumar ou não? Explique. (*Interpretação*)

CQ3: Os dados do gráfico apresentam informações parecidas com o que você imaginou? Por quê? (*Confronto entre hipótese e dados*)

CQ4: De acordo com o gráfico, pode-se dizer que o risco de fumar e ter câncer de pulmão é pequeno? Por quê? (*Avaliação de conclusão*)

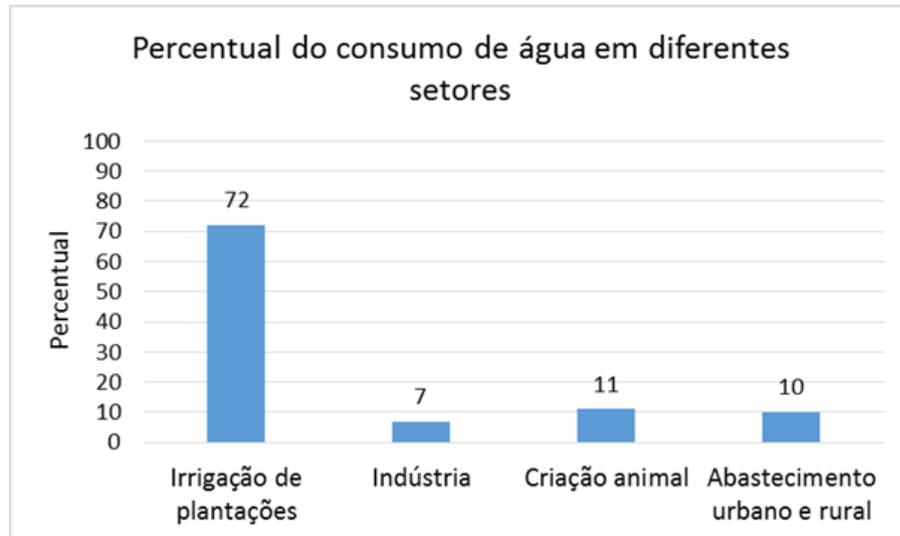
Observa-se que as questões na Atividade 2A receberam um código (CQ) no início, que não colocados aos estudantes, mas que são utilizados na explicitação do método e nas análises posteriores, a fim de diferenciação entre os tipos de dados, se univariados ou em correlação (bivariados). Quando se discute as mesmas habilidades é importante saber em qual das situações as elas foram exploradas.

Teste B

Atividade 1B

A Atividade 1B contou com as mesmas habilidades da primeira atividade do Teste A. Houve uma mudança na temática, que tratou do uso da água em setores do Brasil. Tais dados foram explorados no Estudo 1, contudo a partir da representação em tabela, o que não é o foco do Estudo 2.

Q1: Para você a água consumida no Brasil é usada em maior quantidade onde (casas, indústrias, plantações, criação de animais)? Explique. (*Levantamento de hipótese*)



Fonte: Conjuntura de Recursos Hídricos do Brasil, ANA, 2013.

Q2: De acordo com o gráfico, como tem sido a distribuição do uso de água em diferentes setores do país? Explique. (*Interpretação*)

Q3: Os dados do gráfico apresentam informações parecidas com as que você imaginou na primeira questão? Por quê? (*Confronto entre hipótese e dados*)

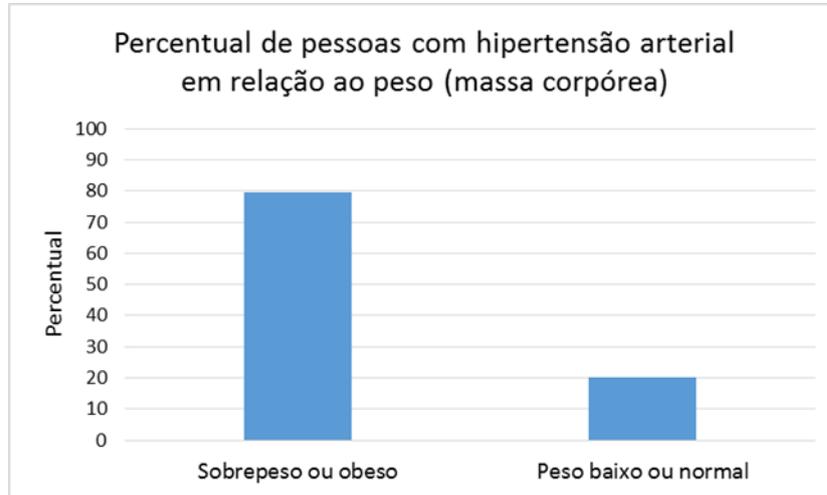
Q4: A partir do gráfico, é possível afirmar que economizar água em casa resolve o problema de falta de água? Por quê? (*Avaliação de conclusão*)

Q5: Depois de conhecer os dados do gráfico, você acredita que o modo que se usa a água no país pode ser muito diferente em 2017? Por quê? (*Predição*)

Atividade 2B

A Atividade 2B contou com dados bivariados, como na segunda atividade do Teste A. Apesar da temática se distinguir da atividade correspondente do Teste A, a Atividade 2B está dentro de um contexto próximo, pois trata de assuntos ligados à saúde.

CQ1: Para você o que a hipertensão arterial (conhecida como pressão alta) tem a ver com o peso (massa corpórea) das pessoas? Explique. (*Levantamento de hipótese*)



Fonte: Revista Brasileira de Medicina do Trabalho, 2009.

CQ2: De acordo com o gráfico, existe relação entre hipertensão e peso (massa corpórea)? Explique. (*Interpretação*)

CQ3: Os dados do gráfico apresentam informações parecidas com o que você imaginou na primeira questão? Por quê? (*Confronto entre hipótese e dados*)

CQ4: De acordo com o gráfico, pode-se dizer que o risco de ser hipertenso e estar acima do peso ideal é pequeno? Por quê? (*Avaliação de Conclusão*)

Salienta-se que a obtenção de dados reais, principalmente em situação de dados bivariados e apropriados para o público do Ensino Fundamental é um processo exaustivo que requer um longo período de pesquisa (CAVALCANTI; GUIMARÃES, 2018). Contudo, o trabalho com dados reais é imprescindível ao objetivo traçado na tese, uma vez que vivências e crenças dos estudantes permeariam todo o estudo, como já alertado na literatura.

Optou-se no teste inicial (pré-teste) e consequentemente no teste final (pós-teste) pela abordagem exclusiva de dados bivariados em correlação direta, uma vez que as relações indiretas nos dados bivariados são consideradas de maior complexidade pelas pesquisas da área (CARVALHO, 2008) e o presente estudo,

pelo carácter inovador, não tem referências na literatura a respeito da maioria das habilidades investigadas na tese.

3.5.2.2 Contribuições da Zona de Desenvolvimento Proximal de Vygotsky para o processo de intervenção didática

Pensar no processo de intervenção de ensino junto aos estudantes participantes do estudo requer a discussão a respeito da perspectiva teórico-metodológica implementada em todo o processo didático. Acredita-se que a aprendizagem é anterior e extrapola o ambiente escolar. Além disso, relações entre colaboradores precisam ser consideradas quando se almeja a aprendizagem de um grupo. Assim, aliar as contribuições de estudos de Lev Vygotsky referentes à aprendizagem com a exploração de habilidades estatísticas no ensino focalizadas na tese foi uma escolha no sentido de se viabilizar o processo pretendido.

A exploração de informações estatísticas está presente na Educação Básica a partir do currículo de Matemática e relacionam-se com diversas áreas do saber. Gal (2002) já explicitou que interpretação e julgamento de informações de carácter estatístico requerem diferentes conhecimentos, inclusive de mundo. Nesse sentido, é inevitável que crenças ou expectativas, tal como estabelecido por Nickerson (1998) apareçam nas compreensões demonstradas por estudantes, diante do ensino de estatística.

Acredita-se que quando se deseja o Letramento Estatístico da população, investir na aprendizagem de habilidades necessárias a uma postura crítica diante de informações estatísticas é essencial. Nesse sentido, o conhecimento de fases do ciclo investigativo traz grandes possibilidades de potencialização dos conhecimentos cognitivos e disposicionais propostos por Gal (2002). Entende-se que interpretar, avaliar, discutir, comunicar e desconfiar de informações produzidas estatisticamente requer a reflexão sobre diferentes aspectos dos dados reais representados, desde a problemática que levou a elaboração de hipóteses para a mesma até as conclusões acerca dos resultados alcançados. Em se tratando de dados produzidos por outros, em que o leitor os conhece a partir de meios midiáticos, a postura crítica é altamente necessária.

Nesse sentido, investir em situações de ensino organizadas de modo que a aprendizagem de habilidades relacionadas ao Letramento Estatístico da população

se inicie nos primeiros anos de escolaridade é desejável. Para isso, estabelecer as compreensões dos estudantes a respeito de tais habilidades é fundamental e analisar o papel da mediação desempenhada pelo professor nas situações de ensino e a interação do sujeito com seus pares é importante no que diz respeito a aprendizagem. Como o presente estudo tem como foco a aprendizagem dos estudantes acerca das habilidades elencadas, saber que compreensões os sujeitos já apresentam é primordial.

Vygotsky (1896-1934), teórico russo, criador da Psicologia Histórico-Cultural, salienta que a aprendizagem das crianças começa antes delas frequentarem a escola. O autor, interessado na interação entre aprendizagem e desenvolvimento, estabeleceu a chamada Zona de Desenvolvimento Proximal (ZDP):

É a distância entre o nível de desenvolvimento real, determinado pela resolução independente de problemas e o nível de desenvolvimento potencial, determinado através de resolução de problemas sob orientação de adultos ou em colaboração com pares mais capazes. (VYGOTSKY, 1978, p. 86, tradução nossa).

Para Vygotsky (1978), aprendizagem e desenvolvimento caminham juntos, de modo inter-relacionado desde o primeiro dia de vida da criança. Quando o sujeito está inserido na escola, aprendizagens novas despertam uma variedade de processos internos de desenvolvimento. Assim, a aprendizagem adequadamente organizada resulta em desenvolvimento mental para o teórico (VYGOTSKY, 1978, p. 90).

O autor ressalta, ainda, que a única boa aprendizagem é aquela que antecede o desenvolvimento, implicando que quando se adquire uma nova aprendizagem o processo de desenvolvimento está só começando.

Atividades, testes, problemas resolvidos pela criança de modo independente, isto é, sem qualquer colaboração revelam o nível de desenvolvimento real do indivíduo, funções já amadurecidas. Enquanto que fazeres que requerem assistência, seja de adultos (professor), seja de outras crianças (pares) para serem solucionados constituem-se habilidades em estado embrionário. Desse modo, para o autor, o primeiro caso diz respeito ao nível de desenvolvimento real e caracteriza-se por ser desenvolvimento mental retrospectivo. O segundo caso trata-se do nível de desenvolvimento proximal e caracteriza-se por ser desenvolvimento prospectivo (VYGOTSKY, 1978, p. 87).

As ideias do teórico implicam que a zona de desenvolvimento proximal atual, será o nível de desenvolvimento real no futuro, ou seja, o que o indivíduo faz hoje com assistência, será hábil em fazer posteriormente de modo independente.

Assim, a ZDP é uma ferramenta importante para educadores e psicólogos, uma vez que os estudantes de um determinado grupo podem apresentar o mesmo nível de desenvolvimento real em um dado teste, contudo, quando diante de atividades em colaboração, podem estar em níveis diferentes, alguns conseguindo avançar mais que outros em suas habilidades. A ZDP, então, não é a mesma para o grupo citado.

Saxe, Gearhart, Note e Paduano (1994, p.173) expõem a ZDP como “*uma ‘região’ de sensibilidade à instrução*”. Isso significa que o sujeito só é capaz de realizar com ajuda aquilo que está na ZDP. Problemas além da compreensão da criança não são solucionados. Um exemplo de atividade dentro da ZDP é o processo de imitação, a criança só pode realizá-la porque pode entender as relações estabelecidas nesse processo de imitar um adulto ou outra criança. Assim, é importante que as situações de ensino atuem dentro desse campo de sensibilização, para que a mesma seja efetiva.

Percebe-se que, para Vygotsky (1978), a interação entre indivíduos é crucial para a aprendizagem e para o desenvolvimento. Considerar a dimensão social do desenvolvimento humano integra o conceito de ZDP com a teoria histórico-cultural do autor.

Oliveira (2016) expõe que Vygotsky apresentou como um de seus pressupostos básicos a concepção de que o ser humano se constitui enquanto tal, na sua relação com o outro social. Desse modo, “*a cultura torna-se parte da natureza humana num processo histórico que, ao longo do desenvolvimento da espécie e do indivíduo, molda o funcionamento psicológico do homem.*” (p. 24).

A instrução para uma aprendizagem adequadamente organizada requer considerar a inserção do indivíduo em sua cultura, uma vez que isso está intrinsecamente ligada ao funcionamento psicológico. Cada faixa etária conta com atividades que se sobressaem de acordo com Prestes, Tunes e Nascimento (2013), o que precisa ser considerado.

A instrução escolar, nesse sentido, seria o portal da consciência, ou seja, oportunizaria a criança a formação da consciência reflexiva, diante de atividades que a mesma desempenha (PRESTES; TUNES; NASCIMENTO, 2013). Prestes et al

(2013) ressaltam, ainda, que o termo instrução, defendido por Vygotsky, pode se referir tanto ao que ocorre no indivíduo (aprendizagem), quanto ao que externamente (ambiente) produz essa ocorrência.

Percebe-se, assim, que a instrução engloba o aprender, processo pessoal (individual) e considera também os elementos externos ao indivíduo que propicia a aprendizagem, uma vez que forma determinada habilidade.

A partir dos estudos da teoria de Vygotsky, Hedegaard (2002) afirma que conhecimentos científicos considerados importantes para o currículo são guiados pelo ensino. Ao partir desse pressuposto, defende-se que é por meio da instrução que conceitos científicos se relacionam com os conceitos corriqueiros da criança. Para a autora, o papel do professor, então, é dirigir a ação de uma maneira apropriada ao nível atual de desenvolvimento do sujeito, ao contexto cultural e social e às teorias do professor sobre o que é um assunto central.

Em se tratando de aprendizagem da Matemática e processos de grupo, Saxe, Gearhart, Note e Paduano (1994) apresentam elementos da estrutura criada por Saxe, que tem grande influência da obra de Vygotsky. Da estrutura de Saxe, para compreender processos de grupo e aprendizagem de Matemática, na pesquisa e na prática educacional, três componentes foram delimitados: O primeiro refere-se à formação de objetivos (objetivos emergentes) nas práticas culturais cotidianas; o segundo componente diz respeito aos desenvolvimentos cognitivos ligados às atividades das crianças dirigidas para uma finalidade; e o terceiro relaciona-se à influência recíproca da aprendizagem das crianças através de contextos e práticas.

A respeito dos três componentes, Saxe et al (1994) afirmam que as interações das crianças podem estar entrelaçadas com a emergência de objetivos matemáticos. Objetivos influenciados por outras crianças oferecem um contexto para que uma criança construa novas compreensões lógico-matemáticas. Do mesmo modo que interações entre crianças podem estar entrelaçadas com processos de “transferência”. Os autores salientam que nas práticas de sala de aula e outras práticas cotidianas, as crianças se ocupam com problemas recorrentes e frequentemente há apoios sociais para a apropriação do conhecimento construído em uma prática para alcançar os objetivos emergentes em outra prática.

A abordagem de Saxe parte da ZDP de Vygotsky, conforme Saxe et al (1994), uma vez que o conceito de Vygotsky representa um fenômeno complexo: une os processos interacionais sociais às atividades construtivas do indivíduo. Desse modo,

na estrutura de Saxe a ZDP contribui na compreensão de como os objetivos cognitivos das crianças mudam e tomam forma em suas práticas cotidianas.

3.5.2.2.1 *Objetivos delineados para os encontros de intervenção*

O processo de intervenção parte da consideração de que a exploração da habilidade de se levantar hipóteses em diferentes momentos da pesquisa estatística (antes e após conhecimento dos dados) favorece a postura crítica diante dos dados, no sentido de que análises, conclusões e predições sejam realizadas em função da reflexão a partir dos mesmos. A hipótese de intervenção se justifica devido aos resultados alcançados no Estudo 1, no qual estudantes, que a princípio demonstraram não interpretar os dados quando solicitados, retomaram a análise dos dados em momentos posteriores do ciclo investigativo. Além disso, em uma Análise Exploratória dos Dados (AED) geração de novas hipóteses faz parte do processo investigativo e de entendimento dos dados.

O Quadro 3.2 especifica características e objetivos delineados para os três encontros de intervenção, os quais aconteceram com toda a turma.

Quadro 3.2 – Objetivos delineados para três encontros de intervenção

Encontro	Objetivos para encontros de intervenção
1º dia: Intervenções a partir de análise de dados univariados e bivariados - Atividades coletivas	Dados reais univariados (Atividade 1.1): <ul style="list-style-type: none"> ▪ Explorar levantamento de hipóteses antes de conhecimento dos dados e após análise dos mesmos; ▪ Propor dados reais para análise conjunta; ▪ Confrontar dados com hipóteses levantadas; ▪ Propor conclusões para avaliação por parte dos estudantes; ▪ Solicitar predições a respeito dos dados e explorar uso de linguagem probabilística; ▪ Levantar novas hipóteses para estudos futuros.
	Dados reais bivariados (Atividade 1.2): <ul style="list-style-type: none"> ▪ Explorar levantamento de hipóteses antes de conhecimento dos dados e após análise dos mesmos; ▪ Propor dados reais para análise conjunta; ▪ Confrontar dados com hipóteses levantadas; ▪ Propor conclusões para avaliação por parte dos estudantes; ▪ Levantar novas hipóteses para estudos futuros.

2º dia: Intervenções a partir de análise de dados univariados – Atividades coletivas e em pequenos grupos	Dados reais univariados / discussões coletivas (Atividade 2.1): <ul style="list-style-type: none"> ▪ Explorar levantamento de hipóteses antes de conhecimento dos dados e após análise dos mesmos; ▪ Propor dados reais para análise conjunta; ▪ Solicitar conclusões a partir dos dados; ▪ Propor conclusões para avaliação por parte dos estudantes; ▪ Solicitar predições a respeito dos dados e explorar uso de linguagem probabilística; ▪ Levantar hipóteses para estudos futuros.
	Dados reais univariados / discussões em grupos (Atividade 2.2): <ul style="list-style-type: none"> ▪ Explorar levantamento de hipóteses antes de conhecimento dos dados e após análise dos mesmos; ▪ Propor dados reais para análise pelos grupos; ▪ Confrontar dados e hipótese levantada; ▪ Propor conclusões para avaliação por parte dos estudantes; ▪ Solicitar predições a respeito dos dados e explorar uso de linguagem probabilística; ▪ Levantar hipóteses para estudos futuros.
3º dia: Intervenções a partir de análise de dados bivariados – Atividades coletivas e em pequenos grupos	Dados reais bivariados / discussões coletivas (Atividade 3.1): <ul style="list-style-type: none"> ▪ Explorar levantamento de hipóteses antes de conhecimento dos dados e após análise dos mesmos; ▪ Propor dados reais para análise conjunta; ▪ Confrontar dados e hipótese levantada; ▪ Propor conclusões para avaliação por parte dos estudantes; ▪ Solicitar predições a respeito dos dados e explorar uso de linguagem probabilística; ▪ Levantar hipóteses para estudos futuros.
	Dados reais bivariados / discussões em grupos (Atividade 3.2): <ul style="list-style-type: none"> ▪ Explorar levantamento de hipóteses antes de conhecimento dos dados e após análise dos mesmos; ▪ Propor dados reais para análise pelos grupos; ▪ Confrontar dados e hipótese levantada; ▪ Propor conclusões para avaliação por parte dos estudantes; ▪ Levantar hipóteses para estudos futuros.

Para efetivação das intenções de trabalho pensadas para cada encontro, houve momentos de trabalho com direcionamento por parte do professor-pesquisador (adulto) e momentos em que os estudantes atuaram em grupos menores (seus pares). Acredita-se que são momentos que despertam posturas diferenciadas por parte dos estudantes e por parte do pesquisador também. Afinal, em momentos coletivos em que dados e questões são apresentados em um quadro o professor está mais em evidência, direcionando mais as discussões do grande grupo. Enquanto que nos grupos menores o diálogo entre integrantes se sobressai.

Assim, as atividades elaboradas para cada encontro de intervenção estão apresentadas a seguir.

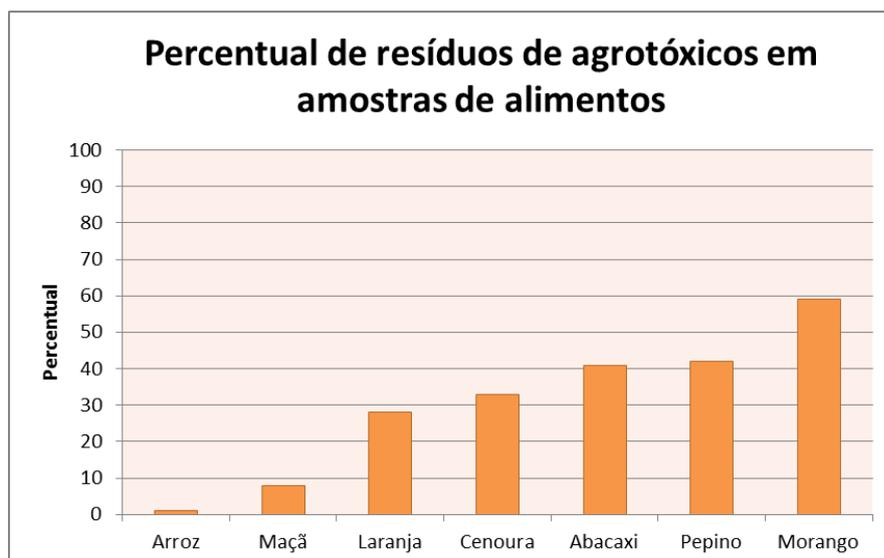
Primeiro dia de intervenção

Ambas as atividades foram trabalhadas de modo coletivo no encontro. Nem todas as habilidades são exploradas nas mesmas. Entende-se que as especificidades de cada tipo de dado conduzem a pertinência das questões.

Atividade 1.1

O tema da Atividade 1.1 foi agrotóxico em amostras de alimentos mostrou-se bastante relevante para ambos os anos de escolaridade, uma vez que a alimentação saudável faz parte das discussões da disciplina de Ciências no Ensino Fundamental. Além disso, nas mídias esse é um tema recorrente. Porém, discutir o termo “agrotóxico” antes do levantamento de hipótese a respeito dos dados que viriam mostrou-se adequado. Alerta-se que o gráfico de barras foi construído a partir de informações disponibilizadas pela Agência Nacional de Vigilância Sanitária (ANVISA).

Levantamento de hipótese: Você sabe o que é agrotóxico? Que alimentos você imagina que têm maiores quantidades de agrotóxicos? Ao se considerar o abacaxi, arroz, cenoura, laranja, maçã, morango e pepino?



Fonte: ANVISA, Resultados de 2012.

Interpretação dos dados reais: O que se percebe dos dados apresentados?

Confronto entre hipótese e dados reais: Os dados reais se parecem com o que vocês pensaram antes?

Avaliação de conclusão: Pensando nas amostras analisadas, podemos afirmar que a maçã é mais “natural” que a cenoura?

Novas hipóteses: Que hipóteses vocês têm para o que esses dados estão mostrando, ou seja, por que vocês imaginam que isso acontece? Por que será que na amostra de morango houve tanto agrotóxico comparada com a de arroz?

Uso de linguagem probabilística ao predizer: Se essa pesquisa fosse repetida neste ano ou no próximo, os números seriam os mesmos?

Ressalta-se que as questões expostas anteriormente possuem um carácter norteador, pois o próprio processo de ensino requer adaptação da linguagem e explicitações de acordo com as necessidades advindas de cada turma, expondo questionamentos ou dúvidas.

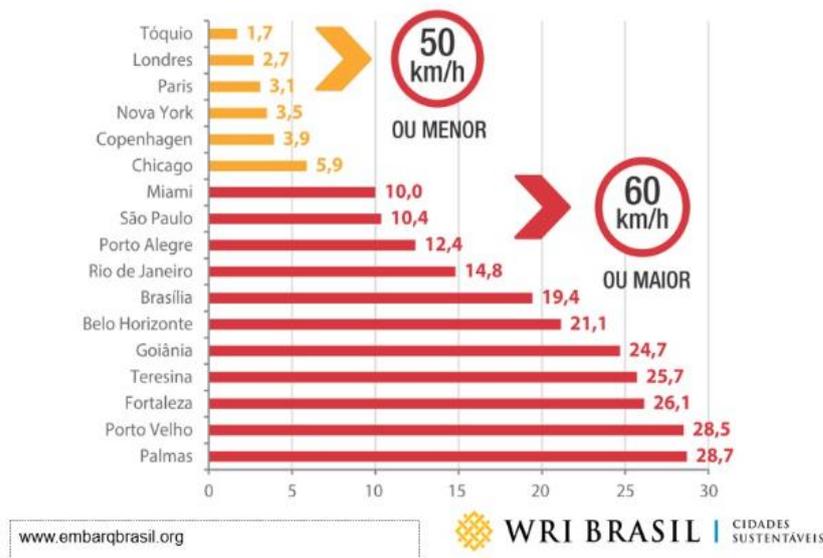
Ressalta-se que as perguntas de levantamento de novas hipóteses não foram trabalhadas nos testes (inicial e final), todavia, o processo interativo e tempo das aulas permitiram a proposição de mais questões, cujo encaminhamento para resolução aborda suposições, o que requer linguagem probabilística também.

Atividade 1.2

O tema mortes no trânsito e limites de velocidade foi o explorado na Atividade 1.2, a respeito da relação entre duas variáveis. O gráfico apresentado aos estudantes foi utilizado tal qual divulgação dos dados pela fonte pesquisadora. Foi importante apresentar os dados conforme veiculação, pois é desse modo que a população toma conhecimento das informações estatísticas, em ambientes não escolares.

Levantamento de hipótese: O que a velocidade dos veículos tem a ver com o número de mortes no trânsito?

Mortes no trânsito x limites de velocidade



Fonte: World Resources Institute (WRI) Cidades Sustentáveis Brasil.

Dados sobre relação entre limite de velocidade permitida em grandes cidades e número de mortes no trânsito (para cada 100 mil habitantes).

Interpretação dos dados reais: O que se percebe da velocidade dos veículos e do número de mortes no trânsito nas cidades mostradas?

Confronto entre hipótese e dados reais: Os dados reais se parecem com o que vocês pensaram antes?

Avaliação de conclusões: Pensando nos dados do gráfico, podemos dizer que diminuir a velocidade impede que mortes no trânsito aconteçam? Podemos afirmar também que a velocidade menor impede os acidentes?

Novas hipóteses: Por que será que o número de mortes maior está relacionado com a velocidade maior no trânsito?

Vê-se que na atividade anterior (Atividade 1.2) a habilidade de avaliação de conclusão contou com mais de uma questão norteadora. Foi interessante explorar mais essa habilidade, por se tratar de dados muito relacionados ao cotidiano dos estudantes, ou seja, a relação entre as variáveis é constantemente noticiada pelas diferentes mídias, o que poderia fazer com que as crenças dos mesmos os levassem a fazer generalizações não permitidas de acordo com os dados. Afinal, pode-se dizer que a menor velocidade está relacionada com a diminuição dos acidentes, mas não impede, além disso, o gráfico não trata do número de acidentes, mas do número de mortes, o que precisa ser compreendido pelo leitor dos dados.

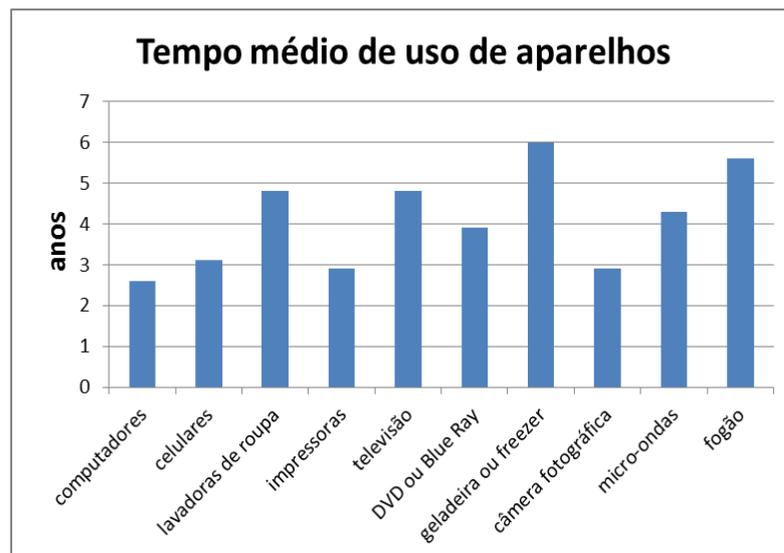
Segundo dia de intervenção

O segundo encontro de intervenção com os estudantes explorou apenas dados univariados. Porém, abordou-se encaminhamentos pedagógicos diferentes no dia, como atividade coletiva e em pequenos grupos.

Atividade 2.1

Os dados da Atividade 2.1 foram coletados no *site* do Instituto Brasileiro de Defesa do Consumidor (IDEC), para assim serem representados em um gráfico de barras.

Levantamento de hipótese: Qual é o tempo médio de uso de aparelhos eletrônicos que normalmente temos em casa?



Fonte: IDEC, 2013.

Interpretação de dados reais: O que se percebe dos dados no gráfico?

Confronto entre hipótese e dados: Os dados da pesquisa se parecem com o que vocês pensaram antes?

Avaliação de conclusões: Podemos afirmar que o baixo tempo de uso de computadores é porque eles dão mais defeitos? Podemos afirmar também que os produtos com tempo mais longo de uso são aqueles usados na cozinha?

Novas hipóteses: Que hipóteses vocês têm para o que a pesquisa mostrou, ou seja, por que vocês imaginam que isso aconteceu?

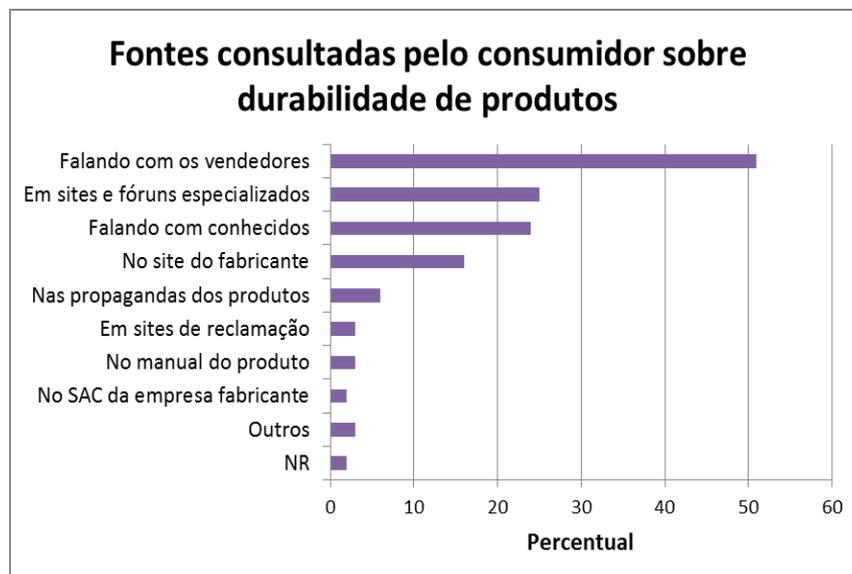
Uso de linguagem probabilística ao prever: Se essa pesquisa fosse repetida neste ano ou no próximo, os números seriam os mesmos?

Como na atividade descrita anteriormente, na Atividade 2.1 há mais de uma questão de avaliação de conclusão, nas quais a primeira requer uma negativa por parte do estudante, quando se considera os dados, e na segunda pode-se haver concordância, desde que se observe apenas os dois produtos com maior tempo de uso.

Atividade 2.2

Os dados utilizados para a Atividade 2.2 são da mesma fonte da atividade anterior, o IDEC. Foi importante mantê-las no mesmo encontro para uma discussão mais profunda acerca da fonte dos dados divulgados pelas mídias. Inclusive no gráfico, a seguir, informações sobre como a coleta foi realizada são postas, o que dá mais elemento para análise dos dados pelos estudantes.

Levantamento de hipótese: Que fontes (pessoas, sites, recursos e etc.) você imagina que o consumidor procura para saber a durabilidade de produtos?



Fonte: IDEC, 2013.

Foram entrevistados, por telefone, 806 homens e mulheres, de 18 a 69 anos, de diferentes classes sociais das seguintes cidades: Belo Horizonte (MG), Brasília (DF), Curitiba (PR), Goiânia (GO), Porto Alegre (RS), Recife (PE), Rio de Janeiro (RJ), Salvador (BA) e São Paulo (SP). O número de entrevistados em cada capital foi proporcional à população de

cada capital. O levantamento foi feito entre agosto e outubro de 2013. A margem de erro é de 3,5% para mais ou para menos.

Interpretação de dados reais: O que se percebe dos dados apresentados?

Confronto entre hipótese e dados reais: Os dados reais se parecem com o que vocês pensaram antes? Por quê?

Tomada de decisão a partir dos dados: Escolham duas dessas fontes do gráfico que vocês consultariam antes de comprar um produto. Explique.

Avaliação de conclusão: Pensando nos dados do gráfico, podemos dizer que as pessoas entrevistadas consultam fontes confiáveis? Por quê?

Novas hipóteses: Vocês imaginam porque tantas pessoas consultam vendedores para saber quanto dura o produto que elas querem comprar?

Percebe-se na descrição anterior que há uma nova habilidade envolvida com a Atividade 2.2, a tomada de decisão a partir dos dados. A literatura apresenta a tomada de decisão como ações posteriores a análise dos dados (MAKAR, 2013). Desse modo, aqui se propõe a mesma como habilidade constituinte do processo de conclusão a partir dos dados, que é a última fase da pesquisa estatística, conforme ciclo investigativo utilizado na tese.

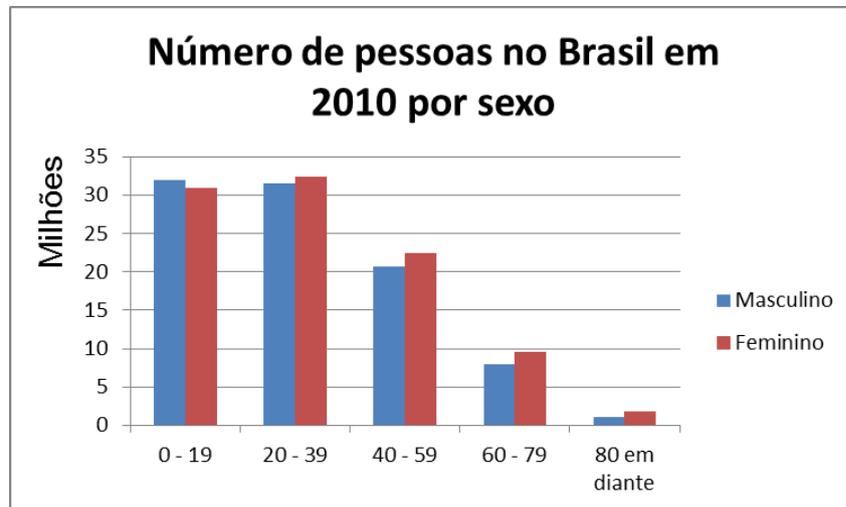
Terceiro dia de intervenção

O último encontro de intervenção com as turmas explorou apenas dados bivariados (correlação), nos diferentes formatos de interação já explicitados. O censo demográfico do Instituto Brasileiro de Geografia e Estatística (IBGE) foi utilizado em ambas as atividades.

Atividade 3.1

Na Atividade 3.1 os dados do censo demográfico apresentaram a relação entre as variáveis idade e sexo, na qual vê-se que nascem mais pessoas do sexo masculino, entretanto, em poucos anos o número de pessoas do sexo feminino se torna maior.

Levantamento de hipótese: Você acredita que nascem mais meninos ou mais meninas no Brasil? O número de pessoas do sexo feminino e do sexo masculino no Brasil tem a ver com a idade?



Fonte: IBGE, Censo Demográfico 2010.

Interpretação de dados reais: O que acontece com o número de meninos e de meninas em relação as idades?

Confronto entre hipótese e dados reais: Os dados reais se parecem com o que vocês pensaram antes?

Interpretação de dados reais (reanálise): O que acontece com a população masculina e feminina com o passar dos anos?

Novas hipóteses: Que hipóteses vocês têm para o que a pesquisa mostrou, ou seja, por que vocês imaginam que isso aconteceu?

Avaliação de conclusão: De acordo com o gráfico, podemos concluir que a população feminina é mais saudável que a população masculina e por isso supera a população de homens dos 20 anos em diante?

Uso de linguagem probabilística ao predizer: Quando essa pesquisa se repetir, no próximo censo, o que pode acontecer?

Para o levantamento de hipótese na descrição da atividade anterior, precisou-se investigar antes de tudo quais as concepções dos estudantes sobre o sexo das crianças nascidas, se havia ou não diferença nos números, para a partir de então ser lançada a questão que levaria a análise da relação com o passar do tempo.

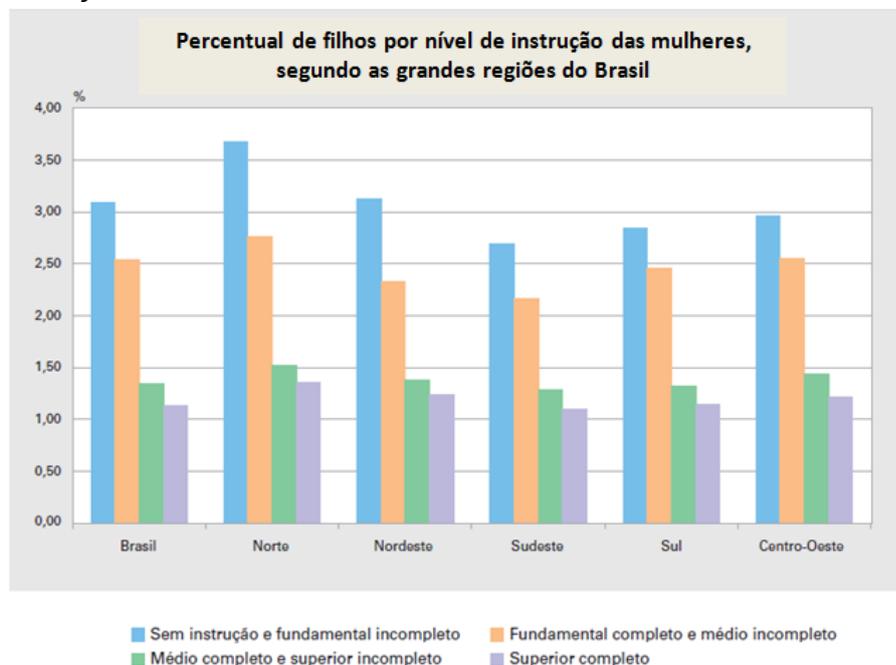
Como incentivo à reanálise dos dados, uma questão de interpretação dos dados após habilidade de confronto entre hipótese e dados foi posta. Entende-se

que retomar os dados pode contribuir para a elaboração de argumentações mais apropriadas por considerarem evidências.

Atividade 3.2

Outro tipo de relação entre variáveis proposta aos estudantes a partir do censo demográfico do IBGE diz respeito ao número de filhos de mulheres e o nível de instrução das mesmas.

Levantamento de hipótese: O que o número de filhos das mulheres tem a ver com o nível de instrução delas?



Fonte: IBGE, Censo Demográfico, 2010.

Interpretação de dados reais: O que se percebe do número de filhos das mulheres de acordo com a instrução (estudo) delas nas diferentes regiões do Brasil?

Confronto entre hipótese e dados reais: Os dados reais se parecem com o que vocês pensaram antes?

Interpretação dos dados reais (reanálise): O que acontece com a quantidade média de filhos que as mulheres têm em relação à instrução delas?

Novas hipóteses: Que hipóteses vocês têm para o fato das mulheres mais instruídas terem menor quantidade de filhos que aquelas com pouca instrução?

Avaliação de conclusão: A partir do gráfico, podemos afirmar que no Sudeste a média de filhos das mulheres é menor que nos outros lugares do Brasil? Explique.

Verifica-se na descrição da Atividade 3.2 que a questão norteadora da habilidade de avaliação de conclusão, diferente das demais, permitia a concordância, sem ressalvas, com o que foi proposto, uma vez que no Sudeste do país, de fato, as mulheres possuem um percentual de filhos menor que o restante do país, seja qual for o nível de instrução delas.

2.5.2.3 Atividades do Pós-teste

As atividades do pós-teste são as mesmas do pré-teste. Isso significa que os estudantes que responderam ao Teste A no pré-teste responderam ao Teste B no pós-teste. Conforme já explicitado, os dois tipos de testes estiveram presentes em todas as turmas investigadas. Metade dos estudantes de cada turma, ou aproximadamente isso, fizeram um dos testes e a outra parte fez o teste similar.

4 ANÁLISE DAS COMPREENSÕES DOS ESTUDANTES NO ESTUDO 1

O Estudo 1 teve como objetivo:

- Analisar como o tipo de distribuição (univariada ou bivariada) interfere nos conhecimentos demonstrados por estudantes do Ensino Fundamental (5º e 7º anos) ao levantarem hipóteses, analisarem dados e tirarem conclusões a partir da análise.

Para tal, 22 alunos do 5º ano e 25 alunos do 7º ano responderam individualmente um teste contendo 4 (quatro) atividades. As duas primeiras atividades partiram de questões gerais, que gerariam uma distribuição univariada dos dados e as duas últimas questões foram estruturadas de modo a fazer com que o sujeito levantasse hipóteses a partir da avaliação de existência ou não de relação entre duas variáveis, sendo uma direta e outra inversa. As quatro atividades propostas foram apresentadas numa mesma ordem para todos os estudantes.

A análise dos protocolos com respostas dos estudantes contou com a avaliação de dois juízes. No sentido de melhor compreender o raciocínio utilizado pelo estudante durante cada atividade e a relação que o mesmo estabeleceu entre as questões, o protocolo de cada atividade foi considerado como um todo, isso significa que para entender se o sujeito relacionou os dados com sua hipótese, as questões um, dois e três precisaram ser analisadas conjuntamente.

Verificou-se que as duas primeiras atividades contaram com um número maior de sujeitos (quarenta e sete) engajados na resolução das mesmas. Nas duas últimas atividades, diminuiu o número de estudantes concluindo-as (quarenta e três no total, dois sujeitos a menos de cada ano não realizaram todas as atividades no tempo aproximado de duas horas). Acredita-se que o efeito cansaço tenha sido responsável pela redução no número de estudantes concluindo todas as atividades.

4.1 ANÁLISE DA ATIVIDADE 1

A primeira questão da Atividade 1 propôs o levantamento de hipótese a partir da pergunta *“Em quais momentos do dia você acredita que as pessoas mais utilizam a internet dos celulares? Por quê?”*. Constatou-se que tanto os estudantes do 5º

ano quanto do 7º apresentaram hipóteses para a questão proposta (Figura 4.1). Apenas 27,3% dos estudantes do 5º ano e 28% do 7º não escolheram um momento, esses responderam para que usavam a internet dos celulares (Figura 4.2).

Q1: Em quais momentos do dia você acredita que as pessoas mais utilizam a internet dos celulares? Por quê?

De madrugada. Porque as vezes a pessoa não tem sono a noite e vai mexer no celular e acabam passando toda a madrugada da talvez até ficar de dia mexendo no celular.

[De madrugada. Porque as vezes a pessoa não tem sono a noite e vai mexer no celular e acabam passando toda a madrugada talvez até ficar de dia mexendo no celular].

Figura 4.1 – Estudante (S42, 7º ano) que levanta hipótese na Atividade 1.

Q1: Em quais momentos do dia você acredita que as pessoas mais utilizam a internet dos celulares? Por quê?

por que a pessoa pode fazer varias coisas como ver filme baixar musicas baixar jogos e outras coisas pra estudar ver video no youtube e só.

[Porque a pessoa pode fazer várias coisas como ver filme, baixar músicas, baixar jogos e outras coisas pra estudar, ver vídeo no *youtube* e só].

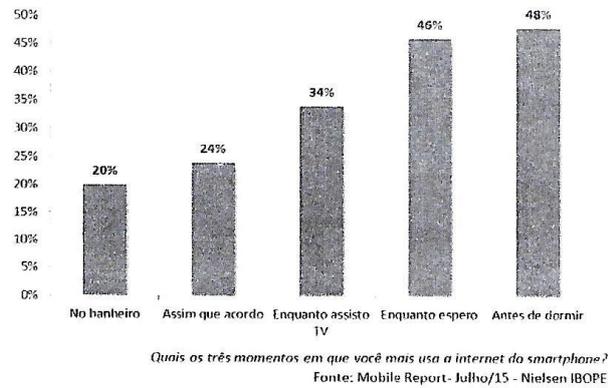
Figura 4.2 – Estudante (S7, 5º ano) que não cita momento algum, refere-se ao uso.

Vê-se na Figura 4.1, que o estudante escolhe um momento de maior uso da *internet* dos celulares, que foi a madrugada, diferentemente do estudante (Figura 4.2) que expõe motivações para o uso, não citando qualquer momento, que foi o solicitado na questão.

Após o devido registro pelos estudantes de suas hipóteses, dados reais representados em um gráfico de barras a respeito do tema foi apresentado para interpretação. Nesse momento questionou-se “*Como as pessoas têm utilizado a internet de seus celulares de acordo com o gráfico?*”. Observou-se que a maioria dos estudantes, tanto do 5º ano (81,8%), quanto do 7º (60%) mostrou evidências de que foram hábeis em interpretar o gráfico (Figura 4.3).

Figura: Maiores momentos de uso do celular com internet em 2015

Os maiores momentos de uso do smartphone com internet - Brasil - julho de 2015



Q2: Como as pessoas têm utilizado a internet de seus celulares de acordo com o gráfico? Explique.

Muitas das pessoas usava muito antes de dormir no banheiro nem tanto, 24% assim que acorda, 34% enquanto assiste TV, 46% enquanto espera, 48% antes de dormir.

[Muitas das pessoas usava muito antes de dormir, no banheiro nem tanto, 24% assim que acorda, 34% enquanto assiste TV, 46% enquanto espera, 48% antes de dormir].

Figura 4.3 – Estudante (S35, 7º ano) que interpreta os dados representados no gráfico.

O estudante no exemplo anterior (Figura 4.3) demonstra que compreendeu o solicitado na questão e chama a atenção em sua resposta para a variação nas porcentagens entre as categorias do gráfico.

Na questão que solicitou o confronto entre as hipóteses formuladas e os dados apresentados no gráfico - “Os dados do gráfico coincidem com o que você imaginou? Por quê?” - observou-se que a maioria dos estudantes conseguiu retomar a primeira questão para relacioná-las (81,8% para o 5º ano e 68% para o 7º ano). Assim, a habilidade de relacionar hipóteses criadas com dados apresentados posteriormente foi possível para ambos os grupos de estudantes. Os exemplos que seguem referem-se a respostas de estudantes do 5º (Figura 4.4) e do 7º ano (Figura 4.5).

Q3: Os dados do gráfico coincidem com o que você imaginou? Por quê?

Coincidem. Eu imaginei as minhas tias antes de irem dormir mexem sempre mexem

[Coincidem. Eu imaginei as minhas tias, antes de irem dormir mexem sempre mexem].

Figura 4.4 – Exemplo de resposta que confronta (Q3) hipótese e dados (S16, 5º ano).

Q3: Os dados do gráfico coincidem com o que você imaginou? Por quê?

Sim parece por que eu botei a noite e o maior uso de internet usado

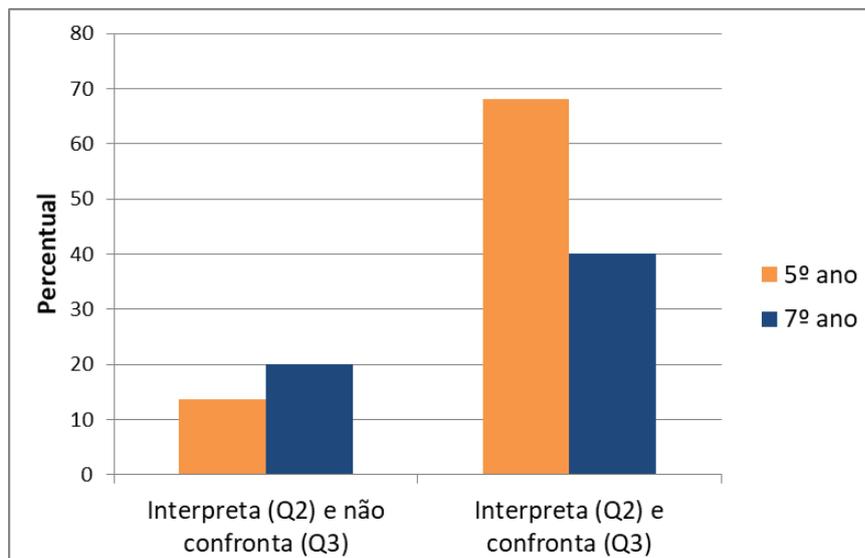
[Sim, parecem. Porque eu botei a noite e o maior uso de internet utilizado].

Figura 4.5 – Exemplo de resposta que confronta (Q3) hipótese e dados (S33, 7º ano).

A análise das compreensões dos estudantes a respeito da questão de confronto entre hipótese e dados requereu a retomada das questões de levantamento de hipótese (Q1) e de interpretação (Q2) para a verificação da argumentação do estudante. Assim, a Figura 4.4 e a Figura 4.5 são exemplos de respostas em que se constatou que havia coerência no que os estudantes afirmavam, ao relacionarem a hipótese levantada e o que perceberam dos dados.

Ao se relacionar a questão de confronto entre hipótese e dados com a habilidade de interpretar (Q2) constatou-se que no 5º ano, a maioria dos estudantes que interpretou também retomou suas hipóteses. No 7º ano o percentual foi de 40%. Quanto aos sujeitos que interpretam, mas não respondem adequadamente a terceira questão esse número foi menor, conforme Gráfico 4.1.

Gráfico 4.1 – Análise da relação entre habilidades de interpretar (Q2) e confrontar (Q3) na Atividade 1 para o 5º e 7º ano



Fonte: o autor.

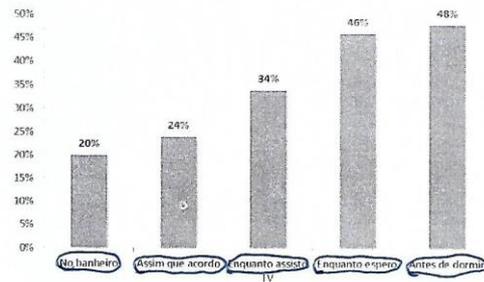
Ainda em se tratando da relação entre as questões de interpretação e confronto, verificou-se também que 13,6% dos estudantes do 5º ano e 28% do 7º ano não interpretaram os dados no momento solicitado, mas quando precisaram confrontar suas hipóteses com os dados voltaram ao gráfico, realizando uma reanálise (Figura 4.6), o que não é uma atitude comumente encontrada na escola. Por outro lado, houve estudantes (13,6% do 5º ano e 20% do 7º ano) que realizaram uma interpretação correta do gráfico, entretanto não conseguiram relacionar com suas hipóteses.

Q1: Em quais momentos do dia você acredita que as pessoas mais utilizam a internet dos celulares? Por quê?

No período da noite as pessoas utilizam mais a internet pois é o período que quase todo mundo está online.
É isso costume a atrapalhar o sono das pessoas e a vida de dia.

Figura: Maiores momentos de uso do celular com internet em 2015

Os maiores momentos de uso do smartphone com internet - Brasil - julho de 2015



Quais os três momentos em que você mais usa a internet do smartphone?
 Fonte: Mobile Report - Julho/15 - Nielsen IBOPE

Q2: Como as pessoas têm utilizado a internet de seus celulares de acordo com o gráfico? Explique.

No gráfico aparenta que as pessoas estão usando a internet do celular a todo momento, desde quando acorda até a hora de dormir.

Q3: Os dados do gráfico coincidem com o que você imaginou? Por quê?

Sim por que eu imaginei que as pessoas usavam mais a internet do celular quando estão no período da noite.

Q1: [No período da noite as pessoas utilizam a *internet*, pois é o período que quase todo mundo está *online*. E isso costuma a atrapalhar o sono das pessoas e a vida de dia].

Q2: [No gráfico aparenta que as pessoas estão usando a *internet* do celular a todo momento, desde quando acorda até a hora de dormir].

Q3: [Sim porque eu imaginei que as pessoas usavam mais a *internet* do celular quando estão no período da noite].

Figura 4.6 – Estudante (S28, 7º ano) que retoma hipótese levantada (Q3), mas não interpretou quando solicitado (Q2).

Na quarta questão da atividade, questionou-se “É possível afirmar que as pessoas usam mais a internet do celular em momentos de trabalho?”. A habilidade envolvida refere-se à avaliação de conclusão, os dados precisariam ser considerados nas respostas dos estudantes. Constatou-se que apenas 14,9% do total de sujeitos (três do 5º ano e quatro do 7º ano) justificaram tomando os dados como evidências (Figura 4.7). Os demais justificaram por suas vivências ou não justificaram.

Q4: É possível afirmar que as pessoas usam mais a internet do celular em momentos de trabalho? Por quê?

~~nao~~ ~~Por~~ ~~nao~~ Por que no banheiro não é trabalho
Assim que acorda também não é trabalho
enquanto assiste também não enquanto espera
também não e antes de dormir também não
~~e~~ e no trabalho

[Não porque no banheiro não é trabalho, assim que acorda também não é trabalho, enquanto assiste também não, enquanto espera também não e antes de dormir também não é no trabalho].

Figura 4.7 – Exemplo de resposta adequada, que justifica pelos dados (S24, 7º ano).

A última questão, que buscou investigar o tipo de vocabulário utilizado pelos estudantes em relação à incerteza da resposta, partiu da pergunta: “*Você acredita que os momentos em que as pessoas mais usam a internet no celular podem ser diferentes em 2016? Por quê?*”. Essa habilidade mostrou-se difícil para ambas as turmas, o que confirma resultados de pesquisa anterior (HENRIQUES; OLIVEIRA, 2015). No 5º ano, apenas 9,5% dos estudantes utilizaram linguagem probabilística e no 7º ano, 12%. A Figura 4.8 exemplifica a resposta de um estudante à questão que investiga tal habilidade.

Q5: Você acredita que os momentos em que as pessoas mais usam a internet no celular podem ser diferente em 2016? Por quê?

sim Por que as pessoas pode ter mudado as
suas Para de mexer no celular

[Sim porque as pessoas podem ter mudado (...) parar de mexer no celular].

Figura 4.8 – Exemplo de resposta que utiliza linguagem probabilística (S26, 7º ano).

Percebe-se na Figura 4.8 que o estudante considera a incerteza, quando utiliza na escrita o termo relacionado a “pode ter”. Situações de incerteza requerem uso de linguagem probabilística, como ressaltado na literatura.

4.2 ANÁLISE DA ATIVIDADE 2

Na Atividade 2, as questões exploraram as mesmas habilidades por parte dos estudantes. Como levantamento de hipóteses (primeira questão) propôs-se: “*Para você a água consumida no Brasil é usada em maior quantidade onde (casas, indústrias, plantações, criação de animais)? Explique*”. Verificou-se que a grande maioria dos estudantes de ambas as turmas escolheu um dos setores de maior uso da água apresentados. Apenas 9,1% dos estudantes do 5º ano e 8% do 7º ano não citou um dos usos da água apresentados, registraram informações que não atendiam à questão, como onde se vende ou onde encontrar.

A segunda questão requereu análise de dados sobre uso da água no Brasil apresentados em uma tabela, com o seguinte questionamento: “*Como tem sido a distribuição do uso da água em diferentes setores do país, de acordo com a tabela? Explique*”. Constatou-se que a maioria dos estudantes (77,3% no 5º ano e 52% no 7º ano) conseguiu realizar a interpretação dos dados. (Figura 4.9).

Q2: Como tem sido a distribuição do uso de água em diferentes setores do país, de acordo com a tabela? Explique.

Irrigação de plantações é a mais que gasta no Brasil e a menos é a indústria porque usam menos água

[Irrigação de plantações é a mais que gasta no Brasil e a menos é a indústria porque usam menos água].

Figura 4.9 – Exemplo de resposta que interpreta os dados (S6, 5º ano).

Vê-se que o estudante na Figura 4.9 expõe em sua argumentação os valores extremos da tabela, como forma de explanar a variabilidade dos dados.

Para retomada da hipótese formulada, após conhecimento dos dados reais, foi apresentada a questão: “*Os dados da tabela coincidem com o que você imaginou? Por quê?*”. Observou-se que a maioria dos sujeitos de ambas as turmas (54,5% no 5º ano e 60% no 7º ano) também se mostrou hábil em rever respostas dadas nas hipóteses (Figura 4.10).

Q3: Os dados da tabela coincidem com o que você imaginou? Por quê?

Eu pensei que indústria tinha mais uso de água, mas não era, na tabela a indústria é menor.

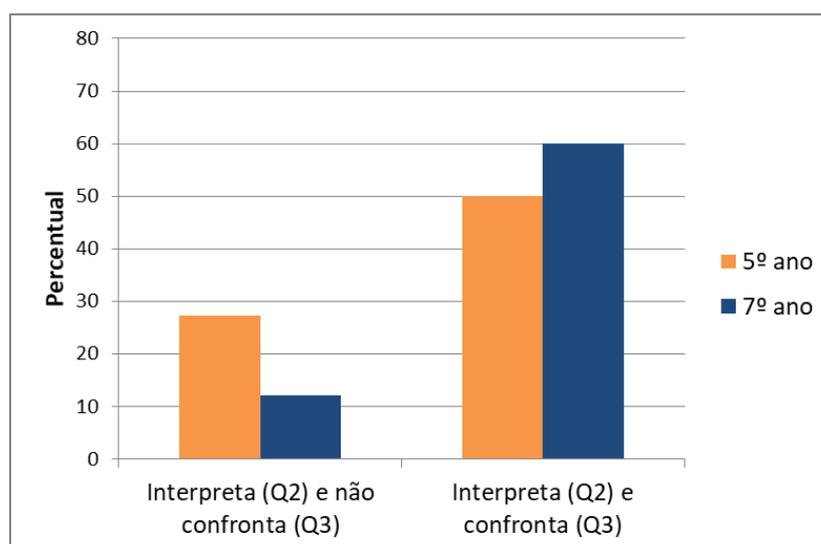
[Eu pensei que indústria tinha mais uso de água, mas não era, na tabela a indústria é menor].

Figura 4.10 – Exemplo de resposta que retoma hipótese após os dados (S8, 5º ano).

Para a análise das compreensões dos estudantes na realização de confrontos entre hipótese e dados, foi necessário retomar a primeira questão respondida pelo estudante como verificação. Assim, na Figura 4.10 evidenciou-se que o sujeito no levantamento de hipótese, de fato, tinha citado a indústria como o setor que gastaria mais água. No exemplo anterior fica claro que a hipótese não se confirmou, o que foi bem aceito pelo estudante.

Outra análise possível para a questão de confrontação entre hipótese e dados (Q3) foi a relação dessa habilidade com a interpretação de dados (Q2). Constatou-se que metade dos estudantes do 5º ano que interpretaram também retomaram a primeira questão em função dos dados, no 7º ano o percentual foi maior (Gráfico 4.2).

Gráfico 4.2 – Análise da relação entre habilidades de interpretar (Q2) e confrontar (Q3) na Atividade 2 para o 5º e 7º ano.



Fonte: o autor.

Na análise da relação entre questões de interpretação (Q2) e levantamento de hipótese (Q1), percebeu-se também que 4,5% dos estudantes do 5º ano e 20% dos estudantes do 7º ano relacionaram os dados com a hipótese levantada, mas não conseguiram interpretar os dados no momento solicitado, o que indica que houve uma reanálise dos dados, assim como na primeira atividade. O extrato de protocolo (Figura 4.11) exemplifica tal discussão.

Q1: Para você a água consumida no Brasil é usada em maior quantidade onde (casas, indústrias, plantações, criação de animais)? Explique.

Indústrias.
Pois indústria gasta muita água com a produção dos produtos.

Tabela: Consumo da água em diferentes setores

Uso da água	Percentual (%)
Irrigação de plantações	72
Indústria	7
Criação animal	11
Abastecimento urbano e rural	10

Fonte: Conjuntura de Recursos Hídricos do Brasil, ANA, 2013.

Q2: Como tem sido a distribuição do uso de água em diferentes setores do país, de acordo com a tabela? Explique.

O uso da água é muito grande.

Q3: Os dados da tabela coincidem com o que você imaginou? Por quê?

não.
Porque na folhinha eu botei que quem mais consome a água foi a indústria, mas o certo é a irrigação de

Q1: [Indústrias, pois gasta muita água com a produção de produtos].

Q2: [O uso da água é muito grande].

Q3: [Não porque na folhinha eu botei que quem mais consome a água foi a indústria, mas o certo é a irrigação de...].

Figura 4.11 – Estudante (S47, 7º ano) que retoma hipótese levantada (Q3), mas não interpretou quando solicitado (Q2).

A questão de conclusão a partir dos dados propôs: “De acordo com a tabela, é possível afirmar que economizar água em casa resolve o problema de falta de água? Por quê?”. A habilidade exigida relacionava-se com a avaliação de uma possível conclusão para os dados apresentados, os quais precisariam ser considerados, por parte dos estudantes, para que a resposta fosse apropriada. Tal

habilidade mostrou-se difícil para ambos os grupos de estudantes. Apenas 12,8% do total (3 estudantes de cada turma) justificaram suas respostas a partir dos dados (Figura 4.12). Os demais, no total, justificaram a partir de suas vivências (61,7%) ou não justificaram (25,5%).

Q4: De acordo com a tabela, é possível afirmar que economizar água em casa resolve o problema de falta de água? Por quê?

Não. Por que nós usamos pouca água

[Não. Porque nós usamos pouca água].

Figura 4.12 – Exemplo de resposta adequada (Q4), que justifica pelos dados (S16, 5º ano).

Percebe-se na Figura 4.12 que o estudante considera os dados reais representados para justificar sua resposta. Afinal, o “nós” expresso pelo sujeito refere-se à realidade próxima dele que é a de utilizar a água em casa, o que não consome tanto quanto outros setores.

A última questão da atividade, “*Os números da tabela sobre o uso da água podem ser muito diferentes em 2016?*”, verificou o vocabulário utilizado pelos estudantes diante da necessidade de realizarem predições a partir dos dados. Os resultados confirmaram as dificuldades já reveladas na questão correspondente da primeira atividade. Apesar do aumento do percentual de acerto, em relação à Questão 5 da Atividade 1 (A1Q5), apenas 25% dos sujeitos do 5º ano e 20,8% do 7º ano utilizaram linguagem probabilística em suas respostas, o que é necessário quando se prediz em estatística.

4.3 ANÁLISE DA ATIVIDADE 3

A Atividade 3 envolveu relação entre duas variáveis, a partir da questão: “*Para você o que o câncer de pulmão tem a ver com o hábito de fumar cigarros, cachimbos, charutos e outros (produtos com nicotina)? Explique*”. As respostas dos estudantes para o levantamento de hipótese, solicitado na primeira questão, são apresentadas na Tabela 4.1.

Tabela 4.1 – Percentual de tipo de resposta para A3Q1 por estudantes do 5º e 7º anos

Ano	Em branco	Refere-se a uma variável apenas	Relaciona as duas variáveis	
			Concorda com relação	Discorda da relação
5º	0,0	10,5	84,2	5,3
7º	4,3	0,0	95,7	0,0
Total	2,4	4,8	90,5	2,4

Para resolução da primeira questão, ambas as variáveis precisariam ser consideradas, independente de haver concordância ou discordância em relação à associação. Percebe-se na Tabela 4.1 que a grande maioria dos estudantes conseguiu explicitar em suas hipóteses a percepção das duas variáveis, relacionando-as de modo direto (84,2% no 5º ano e 95,7% no 7º ano), conforme exemplificação na Figura 4.13. Um pequeno número de sujeitos no 5º ano (10,5%) considerou apenas uma das variáveis em suas respostas (Figura 4.14).

Q1: Para você o que o câncer de pulmão tem a ver com o hábito de fumar cigarros, cachimbos, charutos e outros (produtos com nicotina)? Explique.

Por que câncer de pulmão como o hábito de fumar cigarros etc. eles causam câncer porque eles têm um produto com nicotina, é um produto que faz muito mal para quem fuma.

[Porque câncer de pulmão com o hábito de fumar cigarros etc. eles causam câncer porque eles têm um produto com nicotina, é um produto que faz muito mal para quem fuma].

Figura 4.13 – Exemplo de resposta (S43, 7º ano) que levanta hipótese na Atividade 3.

Q1: Para você o que o câncer de pulmão tem a ver com o hábito de fumar cigarros, cachimbos, charutos e outros (produtos com nicotina)? Explique.

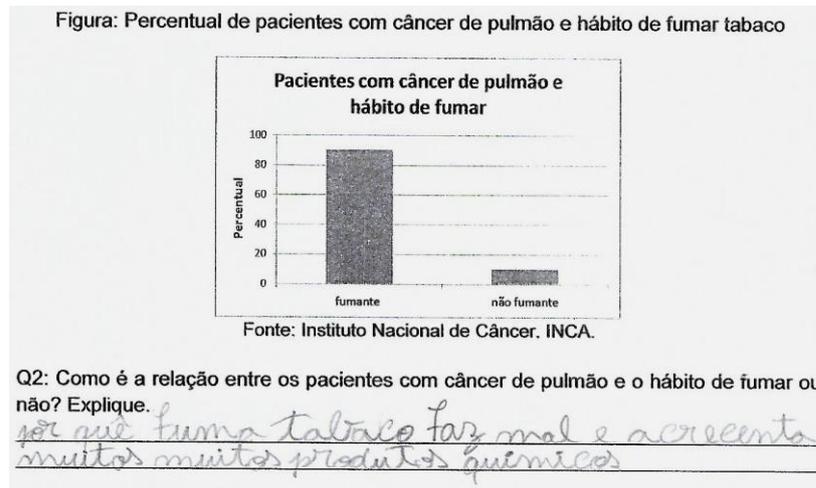
sim charuto porque ele é mais forte

[Sim, charuto porque ele é mais forte].

Figura 4.14 – Exemplo de resposta (S6, 5º ano) que não considera as duas variáveis na hipótese da Atividade 3.

A segunda questão, de interpretação dos dados apresentados, partiu da pergunta: “Como é a relação entre os pacientes com câncer de pulmão e o hábito de fumar ou não? Explique”. A mesma mostrou-se difícil para o 5º ano (45% ofereceram

respostas adequadas) e para o 7º (43,5% responderam corretamente). A maioria dos estudantes elaborou respostas sem indícios de análise do gráfico (Figura 4.15) ou deixou questão em branco. Respostas consideradas adequadas apresentaram um ou mais elementos dos dados representados (Figura 4.16).



[Porque fumar tabaco faz mal e acrescenta muitos produtos químicos].

Figura 4.15 – Exemplo de resposta (S07, 5º ano) que não faz interpretação dos dados.

Q2: Como é a relação entre os pacientes com câncer de pulmão e o hábito de fumar ou não? Explique.

que há mais gente com cancer que fuma do que não fumantes

[Que há mais gente com câncer que fuma do que não fumantes].

Figura 4.16 – Exemplo de resposta (S45, 7º ano) que mostra interpretação dos dados.

A terceira questão partiu da pergunta: “Os dados do gráfico coincidem com o que você imaginou? Por quê?”. Confronto entre hipótese levantada e dados analisados foi a habilidade explorada, a mesma gerou dificuldades maiores que a questão de interpretação. Nesse momento, respostas consideradas adequadas precisavam apresentar indicações de alguma análise do gráfico. Nenhum estudante do 7º ano conseguiu retomar a hipótese levantada na primeira questão e relacioná-la com os dados. Entretanto, a resposta de um sujeito sugere observação do ponto máximo no gráfico (Figura 4.17). No 5º ano, apenas 10,2% dos estudantes relacionaram hipóteses e dados (Figura 4.18). Os estudantes que não confrontaram

hipóteses com análises apresentaram respostas baseadas em crenças (Figura 4.19) ou não responderam.

Q3: Os dados do gráfico coincidem com o que você imaginou? Por quê?

Sim porque a maioria das Escolas pegam câncer de pulmão

[Sim porque a maioria dos fumantes pega câncer de pulmão].

Figura 4.17 – Exemplo de resposta (S35, 7º ano) que sugere percepção de ponto máximo no gráfico.

Q3: Os dados do gráfico coincidem com o que você imaginou? Por quê?

não por que fumante tá 80 por cento e não fumante tá 10 por cento

[Não porque fumante tá 80 por cento e não fumante tá 10 por cento].

Figura 4.18 – Exemplo de resposta (S10, 5º ano) que retoma hipótese após os dados.

Q3: Os dados do gráfico coincidem com o que você imaginou? Por quê?

sim porque a fumaça do cigarro corrói o pulmão

[Sim porque a fumaça do cigarro corrói o pulmão].

Figura 4.19 – Exemplo de resposta (S29, 7º ano) que não relaciona hipótese com dados.

Ainda na terceira questão, outro tipo de análise possível foi a observação da relação entre a habilidade de interpretar os dados (Q2) e a habilidade de confrontar hipótese criada (Q1 – Q3). Constatou-se que apenas estudantes que interpretaram (Q2) conseguiram retomar a primeira questão (de levantamento de hipótese) e relacioná-la com os dados. Esse dado faz com que essa atividade se distinga das atividades anteriores, quando se constatou que estudantes que não tinham interpretado no momento solicitado, ao precisarem responder a terceira questão conseguiram relacionar hipótese e análise de dados, sugerindo um processo de reanálise ou reinterpretção dos dados representados. Isso se justifica pelo pequeno número de estudantes respondendo adequadamente a terceira questão, como já discutido.

A última questão da atividade abordou avaliação por parte do estudante de conclusão dada. Questionou-se: “De acordo com o gráfico, pode-se dizer que o risco

de fumar e ter câncer de pulmão é pequeno? ". Verificou-se que apenas 14% do total de estudantes responderam de modo apropriado, no sentido de considerar os dados como evidências. Desse percentual, um sujeito pertencia ao 5º ano e cinco pertenciam ao 7º ano. Dentre as respostas tidas como adequadas, a Figura 4.20 é um exemplo. Já a Figura 4.21, mostra o oposto.

Q4: De acordo com o gráfico, pode-se dizer que o risco de fumar e ter câncer de pulmão é pequeno?

Não. Porque a maioria dos pacientes que fumam tem câncer de pulmão

[Não porque a maioria dos pacientes que fumam tem câncer de pulmão].

Figura 4.20 – Exemplo de resposta (S42, 7º ano) que avalia conclusão a partir dos dados.

Q4: De acordo com o gráfico, pode-se dizer que o risco de fumar e ter câncer de pulmão é pequeno?

mas não que todo mundo que fuma tem câncer

[Não porque todo mundo que fuma tem câncer].

Figura 4.21 – Exemplo de resposta (S4, 5º ano) que não avalia conclusão a partir dos dados.

Percebe-se que habilidades relacionadas a avaliação de conclusões baseadas em evidências são tarefas complexas, independentemente do tipo de distribuição (se de dados univariados ou bivariados) e do contexto tratado na atividade, já que houve variação nesse sentido.

4.4 ANÁLISE DA ATIVIDADE 4

A primeira questão da atividade partiu da pergunta: “*Você acredita que a quantidade de água no corpo humano tem relação com a idade das pessoas? Por quê?*”. Levantamento de hipótese foi a habilidade investigada na questão, que, assim como na atividade anterior, apresentou duas variáveis para que o sujeito

relacionasse as mesmas, concordando ou discordando. A Tabela 4.2 apresenta as respostas de ambas as turmas para a questão.

Tabela 4.2 – Percentual de tipo de resposta para A4Q1 por estudantes do 5º e 7º anos.

Ano	Em branco	Refere-se a uma variável apenas	Relaciona as duas variáveis	
			Concorda com relação	Discorda da relação
5º	5,0	20,0	30,0	45,0
7º	13,0	8,7	30,4	47,8
Total	9,3	14,0	30,2	46,5

Na Tabela 2, observa-se um percentual maior de respostas em branco, comparada com a mesma questão da atividade 3, o efeito cansaço pode ter contribuído para esse dado. A maioria dos estudantes do 5º ano (75%) e do 7º ano (78,2%) conseguiu criar hipótese que relacionasse as duas variáveis explicitadas na questão. Foi interessante perceber que, nesse caso, houve um número maior de sujeitos discordando da relação entre idade das pessoas e percentual de água no corpo (Figura 4.22). Acredita-se que o contexto da atividade foi de encontro às crenças desses estudantes.

Q1: Você acredita que a quantidade de água no corpo humano tem relação com a idade das pessoas? Por quê?

Eu acho que não, acho que a gente tem a mesma quantidade de água (bebê, adulto, adolescente, idosos) ...

[Eu acho que não, acho que a gente tem a mesma quantidade de água (bebê, adulto, adolescente, idosos) ...].

Figura 4.22 – Exemplo de resposta (S47, 7º ano) que relaciona variáveis, com discordância.

Foi possível perceber também que mais estudantes na Atividade 4 consideraram apenas uma das variáveis em suas respostas (20% no 5º ano e 8,7% no 7º ano), o que é inadequado. A Figura 4.23 exemplifica esse tipo de resposta.

Q1: Você acredita que a quantidade de água no corpo humano tem relação com a idade das pessoas? Por quê?

Sim. Por que precisamos beber bastante água

[Sim, porque precisamos beber bastante água].

Figura 4.23 – Exemplo de resposta (S16, 5º ano) que registra considerações apenas sobre a variável água.

A segunda questão “O que acontece com a quantidade de água no corpo com o passar dos anos? Explique” solicitou interpretação dos dados apresentados em um banco de dados. Como resultados, verificou-se que muitos estudantes deixaram a questão em branco (50% no 5º ano e 39,1% no 7º ano). Como respostas consideradas apropriadas, espera-se que os sujeitos percebessem a relação indireta entre as duas variáveis representadas. Nesse sentido, no 5º ano 20% dos estudantes indicaram em suas respostas essa análise, no 7º ano o número foi maior, 34,8% (Figura 4.24). Outros estudantes estabeleceram uma relação direta entre os dados (Figura 4.25), o que não era apropriado.

Tabela: Percentual de água no corpo humano com o aumento da idade

Idade (ser humano)	Percentual de água no corpo (%)
0 a menos de 2 anos	75 a 80
2 a menos de 5 anos	70 a 75
5 a menos de 10 anos	65 a 70
10 a menos de 15 anos	63 a 65
15 a menos de 20 anos	60 a 63
20 a menos de 40 anos	58 a 60
40 a menos de 60 anos	50 a 58
60 anos acima	menos de 50

Fonte: <http://slideplayer.com.br/slide/2879216/>

Q2: O que acontece com a quantidade de água no corpo com o passar dos anos? Explique.

Ela está diminuindo com o passar dos anos.

[Ela está diminuindo com o passar dos anos].

Figura 4.24 – Exemplo de resposta (S37, 7º ano) que faz uma interpretação adequada dos dados.

Q2: O que acontece com a quantidade de água no corpo com o passar dos anos?
Explique.

aumenta porque se a pessoa fica velha a quantidade de água aumenta.

[Aumenta porque se a pessoa fica velha aí a quantidade aumenta].

Figura 4.25 – Exemplo de resposta (S15, 5º ano) que não faz uma interpretação adequada dos dados.

A última questão da atividade, “Os dados da tabela confirmam o que você pensou? Por quê?”, buscou investigar a habilidade dos estudantes em rever suas hipóteses iniciais e compará-las com os dados. Constatou-se que a maioria dos estudantes de ambas as turmas não conseguiu realizar esse confronto, apenas 45% no 5º ano retomou hipótese em função dos dados e 39,1% no 7º ano (Figura 4.26). Contudo, esse resultado da terceira questão da Atividade 4 ainda é superior ao que se constatou na questão correspondente da Atividade 3, na qual pouquíssimos estudantes responderam de modo adequado.

Q3: Os dados da tabela confirmam o que você pensou? Por quê?

Não. Achei que tínhamos a mesma quantidade de água (bebê, idoso, adolescente, adulto)...

[Não. Achei que tínhamos a mesma quantidade de água (bebê, idoso, adolescente, adulto) ...].

Figura 4.26 – Exemplo de resposta (S47, 7º ano) que retoma hipótese após análise dos dados.

Como explicação para as diferenças encontradas nas duas últimas atividades, em relação à terceira questão (de confronto entre levantamento de hipótese e dados reais), pode-se pensar que a Atividade 3 contou com um contexto em que as crenças dos estudantes foram ao encontro dos dados, o que justifica a grande quantidade de estudantes concordando com a relação entre variáveis no levantamento de hipótese, 90,5% dos estudantes, no geral. Parece que essa familiaridade interferiu negativamente no momento em que o estudante foi solicitado a explicar a comparação que realizou entre hipótese e dados. Como os dados

representados apenas confirmaram o que muitos pensavam, acredita-se que não foram considerados necessários (pois já sabiam) para serem mobilizados em suas explanações. Já na Atividade 4, estudantes tanto do 5º ano, quanto do 7º ano, expressaram com clareza como pensaram diferente em suas hipóteses, o que se evidencia na Figura 4.27.

Q3: Os dados da tabela confirmam o que você pensou? Por quê?
 não pensei que quanto mais crescesse bebia mais, mas
 mais vai diminuindo

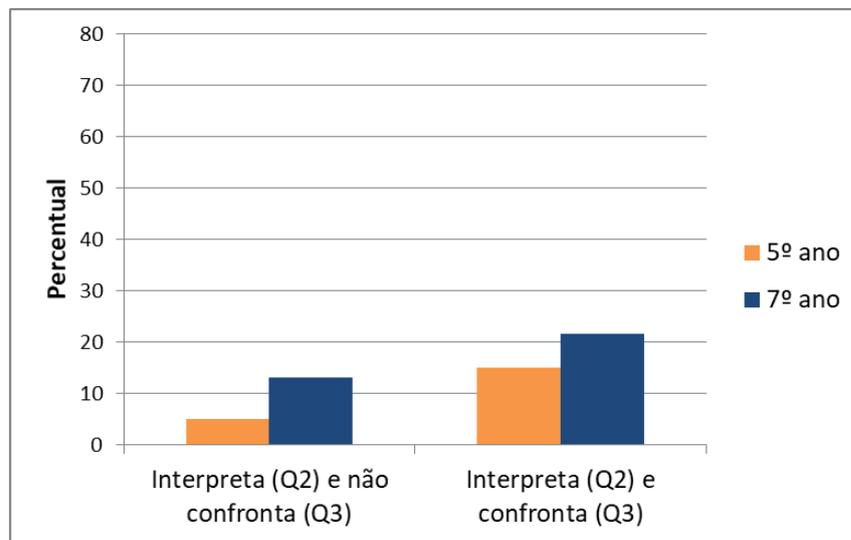
[Não. Pensei que quanto mais crescesse bebia mais, mas vai diminuindo].

Figura 4.27 – Exemplo de resposta (S29, 7º ano) que retoma hipótese e interpreta os dados.

Cumpre reforçar que a questão de levantamento de hipótese (Q1) respondida pelo estudante (Figura 4.27) foi verificada na análise das compreensões, uma vez que nem todos os estudantes que afirmavam ou negavam que os dados confirmaram suas hipóteses realmente perceberam a aproximação ou distância entre suas crenças e os dados expostos.

A terceira questão ainda contou com outro tipo de análise. Trata-se da verificação da relação entre a habilidade de interpretar os dados (Q2) e a habilidade de confrontar hipótese criada (Q1 – Q3). Como resultados, percebeu-se que o percentual de estudantes que interpretaram e retomaram a hipótese foi maior do que o percentual daqueles que apenas interpretaram (Gráfico 4.3), principalmente no 7º ano. Nas Atividades 1 e 2, o percentual de estudantes que interpretaram e também retomaram hipóteses foi maior do que aqueles que só interpretaram.

Gráfico 4.3 – Análise da relação entre habilidades de interpretar (Q2) e confrontar (Q3) na Atividade 4 para o 5º e 7º ano



Fonte: o autor.

Percebeu-se também na análise da relação entre interpretação e retomada de hipótese na Atividade 4, que 18,6% dos estudantes no geral (quatro sujeitos do 5º ano e quatro do 7º ano) retomaram hipótese, respondendo corretamente a terceira questão, mas na interpretação não obtiveram sucesso. Esses sujeitos erraram ao interpretar por não perceberem a relação indireta entre as variáveis. O extrato de protocolo, a seguir, exemplifica essa discussão (Figura 4.28).

Q1: Você acredita que a quantidade de água no corpo humano tem relação com a idade das pessoas? Por quê?

não sei que todo mundo tem uma quantidade de água

Tabela: Percentual de água no corpo humano com o aumento da idade

Idade (ser humano)	Percentual de água no corpo (%)
0 a menos de 2 anos	75 a 80
2 a menos de 5 anos	70 a 75
5 a menos de 10 anos	65 a 70
10 a menos de 15 anos	63 a 65
15 a menos de 20 anos	60 a 63
20 a menos de 40 anos	58 a 60
40 a menos de 60 anos	50 a 58
60 anos acima	menos de 50

Fonte: <http://slideplayer.com.br/slide/2879216/>

Q2: O que acontece com a quantidade de água no corpo com o passar dos anos? Explique.

almenta quantidade porque o corpo cresce

Q3: Os dados da tabela confirmam o que você pensou? Por quê?

nao porque eu pensei que cada um tem sua quantidade

Q1: [Não porque todo mundo tem uma quantidade de água].

Q2: [Aumenta a quantidade porque o corpo cresce].

Q3: [Não porque eu pensei que cada um tem sua quantidade].

Figura 4.28 – Extrato de protocolo (S11, 5^o ano), em que o estudante retoma hipótese, mas errou ao interpretar os dados.

Os resultados apontados no Estudo 1, apresentaram diferentes elementos já ressaltados na fundamentação teórica e devida revisão da literatura (seção 2), que precisam ser contemplados no Estudo 2, o qual tem o carácter intervencionista. Crenças, como destacado por Nickerson (1998) e por Nunes e Bryant (2011), interferem nas compreensões dos indivíduos diante da análise de informações. Acredita-se que as crenças demonstradas por estudantes em contextos mais familiares aos mesmos, no Estudo 1, desempenharam um papel importante nas suas explicitações. Uso de linguagem apropriada diante de situações de incerteza, como já ressaltado por Henriques e Oliveira (2015) não é uma habilidade facilmente compreendida, a mesma requer instrução adequada.

Gal (2002) alertou que diferentes elementos, disposicionais e cognitivos, participam do desenvolvimento de um indivíduo letrado estatisticamente. Pensar, desse modo, em compreensões advindas de contextos diversos, em habilidades

estatísticas e matemáticas, além de outros conhecimentos é importante para a elaboração de processos de ensino que visem a formação de sujeitos críticos.

4.5 CONCLUSÕES DO ESTUDO 1

O Estudo 1 buscou investigar compreensões dos estudantes do 5º e 7º ano, a respeito de habilidades importantes em fases do ciclo investigativo. As habilidades exploradas relacionavam-se ao levantamento de hipótese em diferentes tipos de questões de pesquisa, à análise de dados univariados e bivariados, à confrontação entre hipótese elaborada e dados, à conclusão a partir dos dados e à linguagem probabilística em situações de predição baseadas nos dados.

Nas quatro atividades propostas, a maioria dos estudantes conseguiu levantar hipóteses para as questões de pesquisa propostas, seja nas questões de carácter geral (distribuição de dados univariados), seja nas questões de relação entre duas variáveis (distribuição de dados bivariados). Esse resultado foi surpreendente, pois até então não foram encontradas pesquisas mostrando que estudantes do 5º ano já são capazes de levantar hipóteses para questões de pesquisas estatísticas. Percebe-se também que situações de ensino precisam ser pensadas, uma vez que os estudantes do 7º ano não demonstraram maiores avanços, comparados com os estudantes do 5º ano.

Quanto à habilidade de analisar dados em diferentes distribuições, constatou-se que quando os dados foram apresentados em distribuições simples (univariadas) a interpretação não representou grandes dificuldades para a maioria dos estudantes. Diferente do que se percebeu quando a análise dos dados requeria a interpretação dos dados numa distribuição bivariada, ou seja, quando havia relação entre duas variáveis. Nesse caso, a maioria dos sujeitos (do 5º ano e do 7º ano) apresentou problemas de inadequação de respostas, principalmente quando a relação entre as variáveis foi indireta.

Em se tratando da habilidade de confrontar hipótese criada após conhecimento dos dados, o tipo de distribuição gerada para análise dos dados também parece ter influenciado no desempenho dos estudantes de ambos os anos. Nas duas primeiras atividades, as quais partiram de questões gerais, os estudantes, no geral, mostraram-se bastante hábeis. Já nas últimas atividades, de relação entre

variáveis, a questão mostrou-se difícil, principalmente quando o contexto da atividade foi mais familiar, o que sugere que nesse caso os dados não acrescentaram informação às crenças dos estudantes. Ainda a respeito do confronto entre hipótese e dados, foi possível perceber, nesse momento, estudantes retomando a análise dos dados posteriormente à interpretação, uma vez que quando a habilidade de interpretar foi solicitada não o fizeram.

A habilidade que requeria avaliação de conclusão a partir dos dados mostrou-se bastante difícil para o 5º e 7º ano, em todas as atividades em que a habilidade foi explorada.

A utilização de linguagem apropriada a situações de incerteza também se revelou uma habilidade de difícil compreensão por parte dos estudantes de ambas as turmas. Em se tratando de informações estatísticas, a linguagem probabilística é a mais adequada.

Os resultados alcançados indicaram encaminhamentos importantes para a efetivação do Estudo 2, o qual conta com intervenções em turmas de 5º e 7º ano. Acredita-se que trabalhar situações de ensino de modo sistemático pode favorecer o desenvolvimento de habilidades relacionadas à formulação de hipóteses e à interpretação de dados, assim como a relação entre essas habilidades, por meio de confrontações e reanálise de dados. Isso em diferentes tipos de distribuição de dados.

A exploração exaustiva dos dados, como propõe a AED, em situações de ensino adequadas pode permitir, inclusive, que conclusões sejam tomadas pelos estudantes de modo a considerar os dados e o uso de linguagem apropriada a situações envolvendo incerteza, além de ajudar aos estudantes na identificação de crenças que influenciam suas análises.

Em se tratando de crenças, cabe salientar que o Estudo 1 apresentou, ainda, novidades no que se refere as mesmas. Não foram encontrados estudos anteriores que confrontassem crenças, no sentido de se investigar o que se pensa antes e após o conhecimento dos dados, o que se tornou possível com o levantamento de hipótese e análise dos dados. As crenças podem direcionar leituras de mundo e de informações que chegam a quem as consomem, o que é bastante relevante para o ensino e para a pesquisa de carácter acadêmico.

5 ANÁLISE DAS APRENDIZAGENS DOS ESTUDANTES NO ESTUDO 2

O objetivo do Estudo 2 foi analisar as aprendizagens de estudantes do Ensino Fundamental (5º e 7º anos) a partir de um estudo experimental de intervenção, diante de distribuições univariadas e bivariadas, ao levantarem hipóteses, analisarem dados e tirarem conclusões a partir da análise.

O desenvolvimento desse estudo foi possível a partir do que se constatou no estudo anterior, o qual apresentou evidências de que os estudantes tanto do 5º ano quanto do 7º são capazes de compreender as fases do ciclo investigativo referentes ao levantamento de hipóteses antes e depois da apresentação de dados e a tirar conclusões possíveis a partir dos mesmos.

O estudo experimental contou com um pré-teste seguido de três encontros de intervenção e um pós-teste, na semana seguinte ao último encontro. Participaram do pré-teste três turmas de estudantes do 5º ano e três turmas de estudantes do 7º ano de escolas públicas de Recife (Tabela 5.1).

Tabela 5.1 – Frequência de estudantes por turma e ano de escolaridade no pré-teste

	Turmas A	Turmas B	Turmas C	Total
5º ano	23	15	15	53
7º ano	22	26	20	68

O número de estudantes constituiu um total de 121 sujeitos no pré-teste. Em termos percentuais, 43,8% pertenciam ao 5º ano e 56,2% pertenciam ao 7º.

Iniciamos comparando a média de acerto de cada turma por ano de escolaridade. As 9 (nove) questões propostas aos alunos no pré-teste (apresentadas nas atividades 1 e 2 foram corrigidas para essa análise, como certa ou errada, o que possibilita a comparação entre as médias de acerto.

Tabela 5.2 – Média de acerto por ano de escolaridade no pré-teste

Ano de escolaridade	Turma A	Turma B	Turma C
	Média	Média	Média
5º	4,13	3,13	4,13
7º	5,32	4,81	5,10

Para averiguar se as turmas apresentaram desempenho semelhante no pré-teste, foi realizada uma análise de variância (Anova) para cada ano de escolaridade. Não foi encontrada diferenças significativa para o desempenho dos alunos nem para o 5º ano [$F(2, 52) = 1,717$; $p = 0,190$], nem para o 7º ano [$F(2, 67) = 0,348$; $p = 0,707$]. Esses resultados evidenciam que não havia diferenças entre as turmas de cada ano, o que possibilita o modelo experimental proposto.

Como foi apresentado no método, foram utilizados 2 (dois) testes similares, os quais variaram apenas o contexto das informações. Dessa forma, em cada uma das turmas, metade dos estudantes resolveu o Teste A e a outra metade o Teste B. Como constatado na análise de variância, não houve diferença significativa no desempenho dos estudantes ao responderem o Teste A ou o Teste B para o 5º ano [$F(1, 52) = 1,975$; $p = 0,166$], nem para os estudantes do 7º ano [$F(1, 67) = 3,070$; $p = 0,084$]. Uma vez que não foram encontradas diferenças significativas entre os testes para nenhum dos anos investigados, podemos afirmar que ambos os testes podem ser considerados equivalentes. Desse modo, os testes utilizados no pré-teste mostraram-se importantes instrumentos também para serem utilizados no pós-teste como forma de investigar a possível aprendizagem.

Devido à equivalência entre os testes, as análises que seguem, as quais comparam o desempenho dos estudantes entre pré-teste e pós-teste não distinguem o tipo de teste resolvido pelos sujeitos.

Como já discutido, foram coletados dados em 6 (seis) turmas, sendo 3 (três) do 5º ano (Turmas 5A, 5B e 5C) e 3 (três) do 7º ano (Turmas 7A, 7B e 7C). Nas Turmas A e B (5A, 5B, 7A e 7B) foram realizados processos de intervenção de ensino, seguindo um mesmo planejamento e tempo de duração. Já nas turmas C (5C e 7C) houve apenas o pré-teste e o pós-teste, sendo denominado assim de grupos controle. Foi importante termos grupos controle, que não passaram por um processo de ensino planejado pela pesquisadora, uma vez que também desejamos saber se os estudantes poderiam aprender as habilidades investigadas, sem as atividades planejadas e direcionamentos dados a elas. Termos mais de uma turma de intervenção para cada ano, vivenciando as mesmas atividades, foi necessário, uma vez que teríamos mais elementos para enriquecer as análises partindo-se do contexto de cada turma, enquanto núcleo de relações interpessoais e com o conhecimento. Além do mais, imprevistos que inviabilizassem a continuidade da pesquisa em uma dada turma poderiam acontecer.

Após essa fase, os alunos participaram individualmente de um pós-teste. Os sujeitos que responderam ao Teste A no pré-teste responderam ao Teste B no pós-teste, e vice-versa. Ressalta-se que a aplicação dos testes, assim como os encontros de intervenção aconteceram no mesmo período escolar, ou seja, com datas próximas. O tempo aproximado entre pré-teste e pós-teste foi de um mês para todos os estudantes. Finalizando-se a pesquisa até a segunda semana de dezembro.

É importante esclarecer que durante a pesquisa, alguns estudantes não participaram de todas as fases. No 5º ano houve uma diminuição do número de sujeitos que realizou o pós-teste. Isso em uma das turmas de intervenção. Esse caso específico será discutido mais adiante. Assim, as análises, a seguir, envolvem apenas os sujeitos que participaram tanto do pré-teste como do pós-teste. Além disso, a partir deste momento, as análises quantitativas unem as turmas de intervenção de cada ano de escolaridade, uma vez que o processo de intervenção foi o mesmo, as turmas não apresentaram desempenho significativamente diferente no pré-teste e estavam inseridas em contextos sociais próximos. As médias de acerto para cada grupo de estudantes e fase estão apresentadas na Tabela 5.3.

Tabela 5.3 – Média de acerto por grupo de estudantes no pré-teste e pós-teste

Grupo de estudantes	Média Pré-teste	Média Pós-teste	N
G5I - 5º (Intervenção)	3,78	5,41	32
G5C - 5º (Controle)	4,13	4,07	15
G7I - 7º (Intervenção)	5,09	5,89	47
G7C - 7º (Controle)	5,10	4,75	20

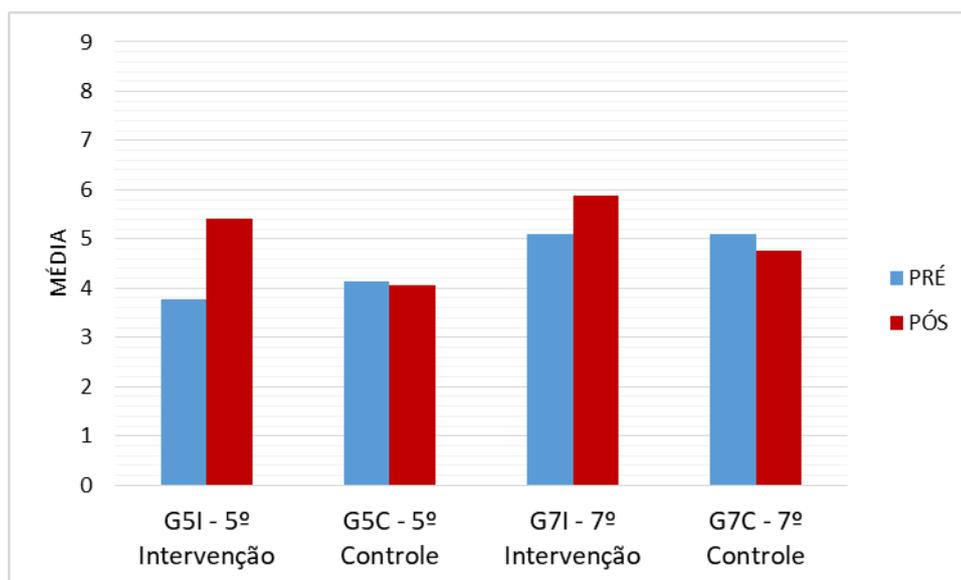
Do quantitativo de sujeitos (N) expresso na Tabela 5.3, apenas um estudante do 7º ano esteve ausente no pós-teste, o qual participou do processo de intervenção. Já no 5º ano esse número foi maior, 6 (seis) estudantes que participaram dos encontros de intervenção não compareceram ao pós-teste. Salienta-se que tais sujeitos pertenciam a uma mesma turma.

Trata-se da turma 5B, na qual houve uma situação atípica. O pós-teste, como já explicitado, aconteceu em dezembro e sendo o último mês do ano letivo o professor da turma dispensou todos os estudantes que estavam aprovados pela

média escolar, mesmo tendo concordado com a pesquisa e todas as suas fases, as quais contaram com datas previamente estabelecidas. É importante esclarecer que essa não é a normativa da rede de ensino que as escolas pesquisadas integram, tanto que o fato aconteceu em apenas uma das seis turmas estudadas. Dessa forma, o pós-teste nessa turma específica foi realizado com os estudantes com um baixo desempenho escolar. Tal fato interferiu na média de acerto do grupo de intervenção do 5º ano, uma vez que os estudantes que apresentavam melhor desempenho e que participaram ativamente durante os encontros de intervenção, infelizmente não realizaram o pós-teste. Apesar dessa interferência negativa, ainda pode-se observar o bom desempenho do 5º ano.

Para uma melhor visualização das médias de acertos por grupo de estudantes podemos observar o Gráfico 5.1.

Gráfico 5.1 – Média de acerto no pré-teste e pós-teste por grupo de estudantes



Fonte: o autor.

Buscando investigar se houve diferença significativa entre o pré-teste e o pós-teste, considerando o desempenho de cada um dos grupos de estudantes, foi realizado um T teste. Comparando-se pré-teste e pós-teste do grupo de intervenção do 5º ano (G5I) verificamos uma diferença bastante significativa ($t(31) = 3,908$; $p < 0,001$) na média de acerto dos estudantes após os encontros de intervenção, ou seja, houve aprendizagem efetiva de habilidades exploradas na investigação. Já o grupo controle do 5º ano (G5C) não diferiu significativamente, o que era esperado, de acordo com nossa hipótese de pesquisa ($t(14) = 0,135$; $p = 0,894$).

Em se tratando dos grupos de estudantes do 7º ano, o mesmo tipo de teste (T teste) revelou que o grupo de intervenção (G7I) também conseguiu avançar em seus conhecimentos de modo significativo ($t(46) = 2,718$; $p = 0,009$), o que não foi constatado no grupo controle do 7º ano ($t(19) = 0,863$; $p = 0,399$).

Verifica-se que os resultados do desempenho dos estudantes dos grupos controle entre pré-teste e pós-teste, no que diz respeito à média de acerto, foram similares quando comparamos 5º e 7º anos. Isso significa que as atividades desenvolvidas nos encontros de intervenção foram imprescindíveis para a elevação da compreensão dos estudantes envolvidos no estudo, ou seja, habilidades investigadas na pesquisa requerem intencionalidade de ensino e situações adequadas para que os estudantes consigam expressá-las em um teste específico. A ausência de diferença entre pré-teste e pós-teste nos grupos controle confirmam essa constatação.

Nesse sentido, a exploração de fases do ciclo investigativo de modo interativo beneficiou aos estudantes que participaram do estudo de intervenção. A defesa do uso do ciclo investigativo e interação entre suas fases ou etapas é amplamente difundida na literatura (WILD; PFANKUCH, 1999; BRESLOW, 1999; GAISE, 2007; SILVA, 2013) e o presente estudo confirma o que pesquisas anteriores vêm apontando.

Outra comparação realizada, considerando-se a média de acerto nos testes, foi a diferença entre os grupos de estudantes. Para isso, uma análise de variância (Anova) foi necessária. A análise de variância revelou diferença significativa entre os grupos de estudantes na média de acertos, tanto no pré-teste [$F(3,113) = 3,318$; $p = 0,023$], quanto no pós-teste [$F(3,113) = 4,728$; $p = 0,004$]. Contudo, cumpre esclarecer que a diferença significativa na média de acerto no pré-teste é explicada pela comparação entre grupos de anos diferentes, de acordo com o post hoc Tukey. Isso aconteceu apenas na comparação entre os grupos de intervenção (G5I e G7I), o que significa que o grupo de intervenção G5I era estatisticamente diferente do G7I ($p = 0,029$). Desse modo, os 7º anos que participaram do processo de intervenção possuíam um desempenho melhor que os estudantes dos 5º anos de intervenção durante o pré-teste. Entretanto, a diferença deixa de existir após a intervenção, como demonstrou a comparação entre as médias no pós-teste dos mesmos grupos ($p = 0,638$).

Quanto a comparação entre os dois grupos de estudantes que constituíram o controle das intervenções (G5C e G7C) verificamos que os mesmos não se comportaram de modo significativamente diferente nem no pré-teste ($p = 0,503$), nem no pós-teste ($p = 0,682$).

A análise das turmas que não vivenciaram o processo de intervenção planejado e a comparação entre as mesmas foi interessante para verificarmos que as habilidades investigadas no presente estudo requerem intervenções pedagógicas para que a aprendizagem aconteça entre os anos de escolaridade considerados.

A partir das análises de variância aqui destacadas evidenciou-se que os estudantes, tanto do 5º quanto do 7º ano são capazes de aprender as habilidades investigadas nesta tese.

Entretanto, vê-se que o grupo de intervenção do 5º ano se beneficiou mais com as atividades planejadas do que o grupo do 7º ano, uma vez que os sujeitos do G5I se equipararam em desempenho com os sujeitos do G7I. Ressalta-se que as turmas dos 7º anos possuíam uma resistência maior à realização de registros escritos, o que provavelmente corroborou para os resultados alcançados na presente pesquisa. Com as peculiaridades de cada ano de escolaridade, pode-se pensar que apesar das adversidades, os estudantes dos anos investigados podem compreender e aprender, com poucos encontros de intervenção, habilidades essenciais à pesquisa estatística, a partir de temas adequados.

A perspectiva de que estudantes podem aprender habilidades necessárias ao pensar de modo estatístico, desde que sejam utilizados dados reais e de interesse do estudante (CAMPOS; WODEWOTZKI; JACOBINI, 2013), assim como seja valorizada a exploração dos dados, conforme orientações da Análise Exploratória de Dados (BEN-ZVI, 2016) foi confirmada nos resultados alcançados e discutidos aqui.

Salienta-se, entretanto, que na literatura específica da área de Estatística na Educação Básica não foram estabelecidas todas as habilidades investigadas no presente estudo. A revisão da literatura revelou a ausência de trabalhos que explorassem de modo integrado levantamento de hipóteses, interpretação de dados reais, confronto entre hipóteses e dados, além de avaliação de conclusões e a linguagem nas predições a partir dos dados já analisados, tanto em distribuições univariadas quanto bivariadas. Aliás, excetuando a interpretação de dados e o uso de linguagem probabilística em situações de incerteza, as demais habilidades não

foram encontradas em estudos anteriores, o que trás para essa tese um carácter inovador.

5.1 APRENDIZAGEM DAS HABILIDADES INVESTIGADAS A PARTIR DE FASES DO CICLO INVESTIGATIVO

O ciclo investigativo é um processo necessário à pesquisa de carácter estatístico, afinal delinear etapas para a resolução de um problema é fundamental. Dos diferentes ciclos investigativos propostos na literatura, no presente estudo, tem-se aquele defendido por Guimarães e Gitirana (2013) como suporte para a eleição das fases envolvidas com o objeto de estudo da tese. A partir da seleção das fases, houve a elaboração das habilidades inerentes às mesmas. A Figura 5.1, a seguir, destaca as fases exploradas no estudo.

Seleção de fases do Ciclo Investigativo



Figura 5.1 – Fases do ciclo investigativo com adaptação (GUIMARÃES; GITIRANA, 2013, p. 97).

As três fases em destaque na Figura 5.1 possuem características diferenciadas que exigem uma abordagem específica quando se pensa no ensino de

habilidades relacionadas às mesmas. Logo, foram definidas as seguintes habilidades:

- Levantamento de hipótese: relaciona-se com a fase de levantamento de dados, como o próprio nome diz, salienta-se que a conjectura requer a exposição de uma questão de pesquisa;
- Interpretação de dados: relaciona-se com a fase de análise dos dados reais, os quais estão sendo apresentados em gráficos;
- Confronto entre hipótese e dados reais: requer a inter-relação entre fases distintas, a hipótese lançada e a análise dos dados reais são confrontadas;
- Avaliação de conclusão dada: Trata-se de um aspecto da fase de conclusão, na qual os dados já foram analisados e precisa-se tomar decisões a partir dos mesmos;
- Linguagem que demonstre incerteza ao se realizar previsões: relaciona-se com a fase de conclusão que é a realização de previsões, o que exige uso de linguagem adequada.

Após a retomada das habilidades envolvidas no estudo da tese, compreende-se, com os resultados apresentados e discutidos no início do capítulo, que os estudantes do 5º e do 7º ano são capazes de compreender e avançar em seus conhecimentos referentes ao levantamento de hipóteses, à interpretação de dados reais, ao confronto entre hipóteses e dados, à avaliação de conclusões dadas e ao uso de linguagem probabilística.

Uma análise das questões e habilidades específicas é o foco a partir de então, tendo-se o pré-teste e pós-teste apresentado ao estudante como apoio. Para a análise que segue o acerto ou erro da questão proposta cede lugar à justificativa dada pelo estudante. Assim, os dados apresentados, a seguir, relacionam-se com a argumentação expressada pelo estudante. É importante essa análise, pois nem sempre o acerto contou com uma justificativa elaborada de modo a elucidar como o estudante pensou. Nesses casos o olhar sobre o protocolo do estudante como um todo foi necessário, a fim de percebermos a compreensão ou não por parte do estudante do solicitado em cada questão. A visão global do protocolo de cada aluno foi fundamental, inclusive, para a análise da habilidade de confronto entre hipótese e dados reais mostrados no gráfico. Portanto, após o conhecimento dos dados

verificar se os mesmos estavam sendo mobilizados como evidência de argumentos colocados pelos estudantes foi uma perspectiva da análise.

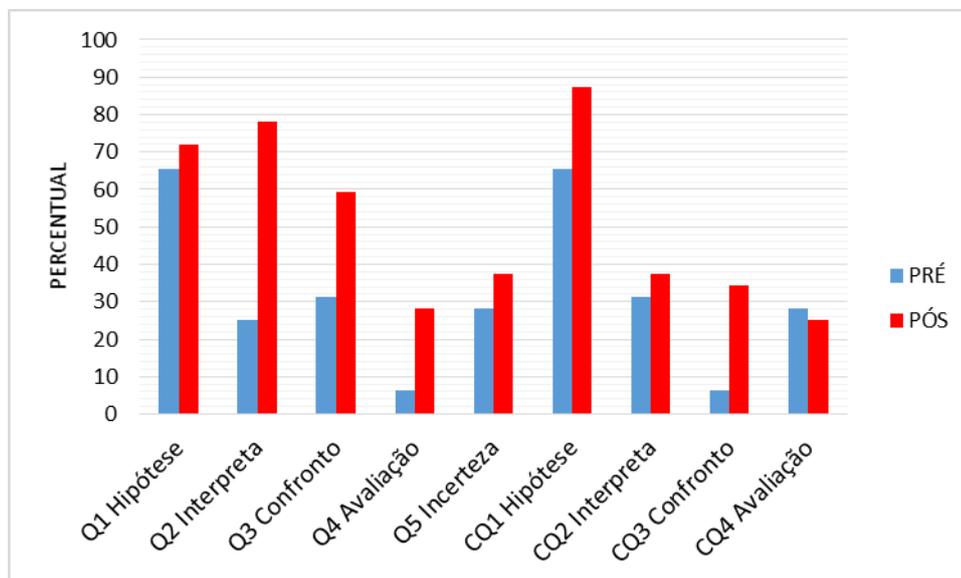
Nos tópicos seguintes, verificar a argumentação realizada pelos estudantes que participaram do processo de intervenção de ensino é o foco, no qual o esperado é a evolução da capacidade de se realizar registros de justificativas mais consistentes com os dados reais apresentados. Assim, buscou-se a comparação entre pré-teste e pós-teste.

5.1.1 Aprendizagem no 5º ano diante das justificativas

No capítulo que trata do método (seção 3), as habilidades investigadas, de acordo com fases do ciclo investigativo, são exploradas nas diferentes questões propostas nos testes, tanto nas Atividades 1 quanto nas Atividades 2. Nas questões apresentadas aos estudantes havia a solicitação, na maioria dos casos, de que se explicasse a resposta dada. O registro da explicitação por parte do sujeito foi um elemento importante como análise da compreensão dos estudantes, mas também é uma habilidade em si. Isso porque foi possível perceber que alguns estudantes sabiam o que a questão proposta requeria, contudo não sabiam expressar o entendimento dos mesmos a partir do registro escrito solicitado.

Desse modo, analisar a evolução da justificativa oferecida pelos estudantes diante das questões que eles precisavam responder é possível, observando-se o percentual de justificativas adequadas dadas em cada questão (que contou com habilidade específica). Isso comparando-se pré-teste e pós-teste. O Gráfico 5.2 apresenta os resultados dessa análise para o grupo de intervenção do 5º ano.

Gráfico 5.2 – Percentual de justificativa adequada em cada habilidade no pré-teste e pós-teste do G5I



Fonte: o autor.

O Gráfico 5.2 demonstra que houve uma melhora percentual nas justificativas ofertadas pelos estudantes do 5º ano que participaram do processo de intervenção. Entretanto, vê-se também que a situação ou contexto com dados univariados (Q1 a Q5) ou bivariados (CQ1 a CQ4) interferiu no uso de justificativas realizado pelos estudantes do 5º ano. Logo, no sentido de um maior esclarecimento dos dados apresentados, análises estatísticas (T teste) são discutidas adiante, assim como evidências com exemplos de respostas dos estudantes também são apresentadas. Para tanto, a discussão de cada habilidade investigada está em destaque.

5.1.1.1 Levantamento de hipótese

Em dados univariados, a habilidade de levantamento de hipótese (Q1), apesar de apresentar um percentual maior de acerto, não apresentou diferença significativa entre pré-teste e pós-teste ($t(31) = 0,528$; $p = 0,601$), no que diz respeito a argumentação utilizada pelos estudantes. Um exemplo de uma argumentação pouco apropriada está na Figura 5.2. Já na Figura 5.3 tem-se um exemplo de um estudante que conseguiu elaborar uma resposta apropriada para a questão colocada.

Q1: Em quais momentos do dia você acredita que as pessoas mais utilizam a internet dos celulares? Por quê?

Porque os celulares são pequenos e são fácil de mexer (usá-lo), São bons porque você pode ver tudo o que quiser, tira o estresse, e nos deixa ocupado. A tecnologia deles é muito avançada, e também porque contem jogos e vídeos, a gente sempre usa eles pra se divertir e nos ocupa, são muito precisos nos dias porque ajudam em praticamente tudo.

[Porque os celulares são pequenos e são fácil de mexer (usá-lo). São bons porque você pode ver tudo o que quiser. Tira o estresse e nos deixa ocupado. A tecnologia deles é muito avançada e também porque contem jogos e vídeos, a gente sempre usa eles pra se divertir e nos ocupa, são muito precisos nos dias porque ajudam em praticamente tudo.]

Figura 5.2 – Sujeito (S06) que não levanta hipótese (Q1) no pré-teste.

Q1: Para você a água consumida no Brasil é usada em maior quantidade onde (casas, indústrias, plantações, criação de animais)? Explique.

Em Casas por que em casa a pessoa toma banho lava lousa e outras coisa mais porque tem gente que gasta muita água por isso que eu acho que é a casa.

[Em casas, porque em casas a pessoa toma banho, lava lousa e outras coisas mais. Porque tem gente que gasta muita água, por isso que eu acho que é a casa.]

Figura 5.3 – Sujeito (S11) que levanta hipótese (Q1) no pós-teste.

No exemplo da Figura 5.2, o estudante na argumentação não expressou os momentos que ele acredita serem mais utilizados pelas pessoas quando utilizam a internet de seus celulares. O sujeito se deteve a explicar sobre a função desses celulares. Quanto ao levantamento de hipótese classificado como adequado (Figura 5.3), o estudante escolhe uma categoria (casas) e justifica o motivo de sua decisão.

Já nas situações de correlação ou dados bivariados, a argumentação utilizada pelos estudantes do 5º ano ao levantarem hipóteses (CQ1) apresentou avanços significativos entre pré-teste e pós-teste ($t(31) = 2,239$; $p = 0,032$). Nos exemplos que seguem há evidências que mostram estudantes que não conseguem elaborar

uma argumentação no levantamento de hipótese em situação de correlação (Figura 5.4) e estudantes que conseguem realizar essa tarefa (Figura 5.5).

Q1: Para você o que a hipertensão arterial (conhecida como pressão alta) tem a ver com o peso (massa corpórea) das pessoas? Explique.

*Tem a ver que comer muita coisa Salgada pode sim
Provocar pressão Alta.*

[Tem a ver que comer muita coisa salgada pode sim provocar pressão alta].

Figura 5.4 – Sujeito (S21) que não levanta hipótese (CQ1) no pré-teste.

Q1: Para você o que o câncer de pulmão tem a ver com o hábito de fumar cigarros, cachimbos, charutos e outros (produtos com nicotina)? Explique.

*Tem a ver quase com tudo, porque com
o tanto que a pessoa fuma a fumaça que li-
gava-se ao charuto enfim, elas entram
no pulmão e causa câncer.*

[Tem a ver quase com tudo, porque com o tanto que a pessoa fuma, a fumaça do cigarro ou do charuto enfim, elas entram no pulmão e causa câncer].

Figura 5.5 – Sujeito (S03) que levanta hipótese (CQ1) no pré-teste.

Na Figura 5.4, o estudante não cita em seu registro a relação ou ausência de relação entre as duas variáveis: hipertensão e massa corpórea. Apenas uma das variáveis foi considerada pelo estudante, a hipertensão. O levantamento de hipótese na situação de correlação exigia por parte do estudante a reflexão a respeito de que a resposta requeria a contemplação de ambas as variáveis, mesmo que o sujeito não concordasse que a relação existisse. Quanto à Figura 5.5, que trata da relação ou não entre as variáveis câncer de pulmão e hábito de fumar produtos com nicotina, o estudante cita as duas variáveis e concorda que há relação entre as mesmas. Cumpre discutir que o estudante estabelece uma relação causal entre as variáveis, o que é um erro comum, conforme Triola (2013) e Friel, Curcio e Bright (2001), entretanto não invalida o levantamento de hipótese pelo estudante no presente estudo.

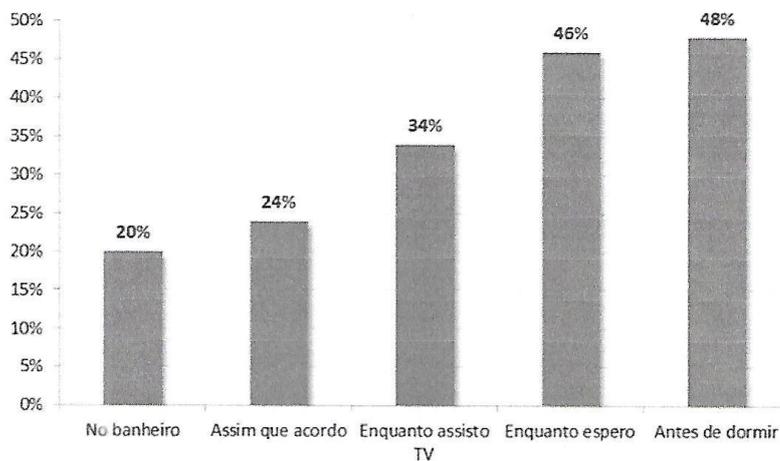
Os dados apresentados no Gráfico 5.2 e análises estatísticas possibilitadas pelo T teste revelaram que as situações de correlação contaram com um grande desenvolvimento na argumentação, após os encontros de intervenção, diferente das situações de dados univariados. Isso foi surpreendente, uma vez que pensar a respeito de duas variáveis, expondo conjecturas não é uma atividade comumente encontrada na sala de aula. Sabe-se que a tarefa de levantar hipótese a respeito de um fato mobilizará crenças ou ideias preconcebidas (NICKERSON, 1998; NUNES; BRYANT, 2011). Isso sugere que as crenças dos estudantes emergiram com maior facilidade diante das situações de correlação.

5.1.1.2 Interpretação dos dados reais

A justificativa oferecida pelos estudantes do 5º ano na interpretação, quando em dados univariados (Q2) melhorou de modo bastante significativo ($t(31) = 5,299$; $p < 0,001$). É importante ressaltar que a questão de interpretação foi elaborada de modo a se permitir uma visão global dos dados representados, o que exigiria uma discussão de carácter mais global na explanação realizada pelo aluno. Isso significa que destacar um ponto do gráfico (ponto máximo, por exemplo) não invalida a interpretação, até porque a exigência da identificação de pontos extremos nas atividades escolares é bastante frequente nos livros didáticos, quando se trabalha com interpretação de gráficos de barras.

Desse modo, a Figura 5.6 e a Figura 5.7 exemplificam a aprendizagem do mesmo sujeito, entre pré-teste e pós-teste.

Os maiores momentos de uso do smartphone com internet - Brasil - julho de 2015



Quais os três momentos em que você mais usa a internet do smartphone?

Fonte: Mobile Report - Julho/15 - Nielsen IBOPE

Q2: De acordo com o gráfico, quando as pessoas têm utilizado a internet de seus celulares? Explique.

As pessoas agora, pela minha opinião de 2015 para 2017 elas não mudaram em uso de Smartphone.

[As pessoas, agora pela minha opinião, de 2015 para 2017 elas não mudaram em uso de smartphone].

Figura 5.6 – Sujeito (S03) que não interpreta o gráfico (Q2) no pré-teste.



Fonte: Conjuntura de Recursos Hídricos do Brasil, ANA, 2013.

Q2: De acordo com o gráfico, como tem sido a distribuição do uso de água em diferentes setores do país? Explique.

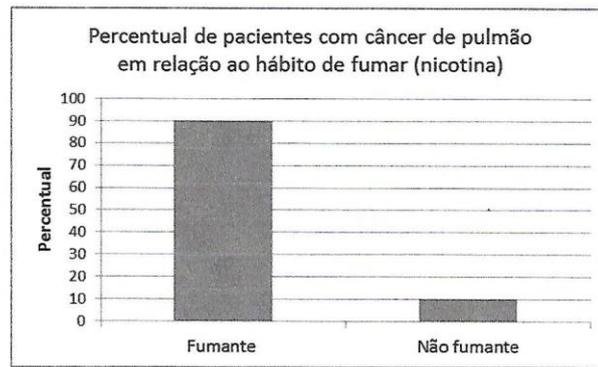
Ruim. Porque um tem mais do que o outro.

[Ruim. Porque um tem mais do que o outro].

Figura 5.7 – Sujeito (S03) que interpreta (Q2) e argumenta no pós-teste.

Vê-se que no pré-teste (Figura 5.6) o estudante expõe um registro que não atende ao solicitado. Questionou-se a respeito de quando as pessoas têm usado a internet dos celulares, mas o estudante ofereceu uma resposta fora de tal contexto, demonstrando não compreensão da pergunta ou do que precisava ser observado no gráfico. Já no pós-teste (Figura 5.7) o mesmo estudante ao opinar demonstra ter percebido a variabilidade dos dados representados. Entende-se que apesar do juízo de valor realizado pelo sujeito, a sua resposta está coerente com a formulação da questão.

A habilidade de interpretação de dados em correlação (CQ2) não apresentou diferença significativa ($t(31) = 0,701$; $p = 0,488$) entre pré-teste e pós-teste no que diz respeito à justificativa utilizada pelos estudantes do 5º ano. Apesar disso, foi possível verificar estudantes que não conseguiam interpretar gráficos com duas variáveis em correlação no pré-teste (Figura 5.8) e após os encontros de intervenção passaram a desempenhar essa tarefa, justificando de modo a demonstrar que perceberam os dados (Figura 5.9).



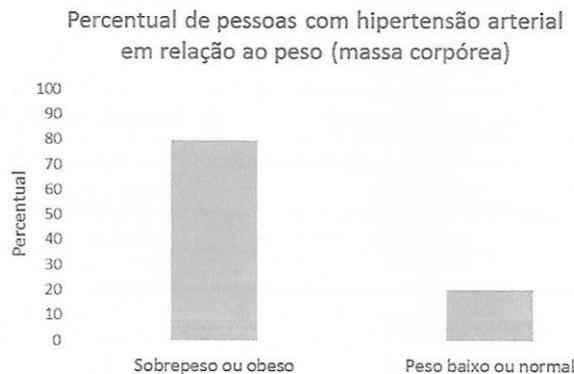
Fonte: Instituto Nacional de Câncer, INCA.

Q2: De acordo com o gráfico, existe relação entre os pacientes que tem câncer de pulmão e o hábito de fumar ou não fumar? Explique.

nao

[Não].

Figura 5.8 – Sujeito (S29) que não interpreta (CQ2) no pré-teste.



Fonte: Revista Brasileira de Medicina do Trabalho, 2009

Q2: De acordo com o gráfico, existe relação entre hipertensão e peso (massa corpórea)? Explique.

as pessoas mais gordinhas e mais facil de ter pressão alta e as pessoas no peso baixo ou normal e dificil ter pressão alta

[As pessoas mais gordinhas é mais fácil de ter pressão alta e as pessoas no peso baixo ou normal é mais difícil ter pressão alta].

Figura 5.9 – Sujeito (S29) que interpreta (CQ2) e argumenta no pós-teste.

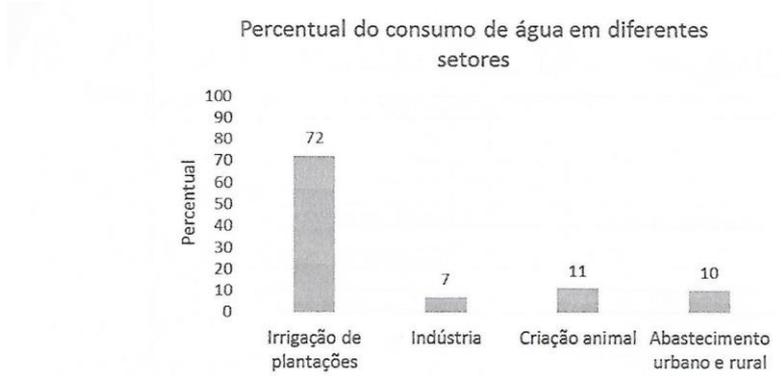
Percebe-se na Figura 5.8 que o estudante não consegue interpretar e elaborar uma explicação para o que os dados apresentam. No pós-teste, o mesmo estudante expõe as duas variáveis do gráfico em seu registro e estabelece uma argumentação em acordo com o que é informado nos dados representados. Interessante que o sujeito utiliza uma linguagem própria e não a utilizada no gráfico, o que reforça que os dados foram bem compreendidos, uma vez que os termos próximos à realidade do estudante estão relacionados ao processo de assimilação e reelaboração. Além disso, sua resposta demonstra uma linguagem de incerteza, uma vez que argumenta utilizando a expressão “é mais fácil”.

5.1.1.3 Confronto entre hipóteses e dados reais

A habilidade de confrontar hipótese levantada e dados reais, tanto na situação de dados univariados (Q3), quanto na situação de correlação ou dados bivariados (CQ3) representou grandes avanços para os estudantes do 5º ano. Quando os dados foram univariados (Q3) o T teste indicou diferença significativa ($t(31) = 2,509$; $p = 0,018$) e quando trataram-se de dados bivariados o índice foi mais expressivo ($t(31) = 3,483$; $p = 0,002$). Na análise da habilidade de confronto entre fases da pesquisa, considerar as três primeiras questões da atividade apresentada ao estudante foi imprescindível. Assim, a Figura 5.10 exemplifica o desempenho de um estudante que não conseguiu realizar o confronto requerido e na Figura 5.11 está o exemplo de um estudante que o faz.

Q1: Para você a água consumida no Brasil é usada em maior quantidade onde (casas, indústrias, plantações, criação de animais)? Explique.

EM PLANTAÇÕES, POR QUE SEM OS VEGETAIS QUASE NINGUÉM VIVE



Fonte: Conjuntura de Recursos Hídricos do Brasil, ANA, 2013.

Q2: De acordo com o gráfico, como tem sido a distribuição do uso de água em diferentes setores do país? Explique.

MUITO INJUSTO POR QUE ESTÁ MUITO POUCO NAS INDÚSTRIA, CRIAÇÃO DE ANIMAL, ABASTECIMENTO DE URBANO E RURAL

Q3: Os dados do gráfico apresentam informações parecidas com as que você imaginou na primeira questão? Por quê?

SIM, POR QUE FALA SOBRE DISTRIBUIÇÃO DE ÁGUA

[Q1: Em plantações porque sem os vegetais quase ninguém vive].

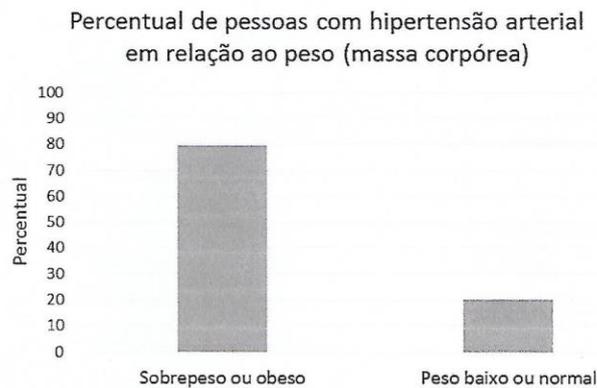
[Q2: Muito injusto, porque está muito pouco nas indústrias, criação de animal, abastecimento urbano e rural].

[Q3: Sim, porque fala sobre distribuição de água].

Figura 5.10 – Sujeito (S17) que não confronta a hipótese com os dados (Q3) no pré-teste.

Q1: Para você o que a hipertensão arterial (conhecida como pressão alta) tem a ver com o peso (massa corpórea) das pessoas? Explique.

Sim, porque a massa vai vai cada vez mais e então a pessoa se pressa a cada mais dando a pressão subi mais rapido.



Fonte: Revista Brasileira de Medicina do Trabalho, 2009

Q2: De acordo com o gráfico, existe relação entre hipertensão e peso (massa corpórea)?

Explique.

A mais pessoa sobrepeso ou obeso.

Q3: Os dados do gráfico apresentam informações parecidas com o que você imaginou na primeira questão? Por quê?

Sim, por que quem tem mais pressão alta e tem sobrepeso ou obeso.

[CQ1: Sim, porque a massa vai cada vez maior e quando a pessoa se preocupa, ajuda mais ainda a pressão subir mais rápido].

[CQ2: Há mais pessoas com sobrepeso ou obeso].

[CQ3: Sim, porque quem tem mais pressão alta é ou sobrepeso ou obeso].

Figura 5.11 – Sujeito (S11) que confronta a hipótese com os dados (CQ3) no pós-teste.

Percebe-se na Figura 5.10 que o estudante (S17) levanta hipótese, interpreta os dados do gráfico emitindo uma consideração, porém, ao necessitar confrontar os dados, na situação proposta, de dados univariados, o mesmo não relaciona os registros realizados por ele nas questões. Numa outra situação, de correlação (Figura 5.11) há o exemplo de um estudante que levantou hipótese e conseguiu confrontar os dados do gráfico com o seu registro, apesar da sua interpretação (CQ2) não abranger as duas variáveis representadas no gráfico.

No Estudo anterior da tese (seção 4) já se constatou que estudantes que não interpretaram os dados adequadamente no momento proposto (Q2 e CQ2),

realizaram uma reanálise quando solicitados a confrontarem hipótese e dados reais (Q3 e CQ3). Isso significa que na questão de interpretação alguns estudantes não conseguiram expressar uma análise que correspondesse aos dados mostrados. Entretanto, os mesmos sujeitos ao serem solicitados a confrontarem hipótese levantada e dados reais retomaram o gráfico em um processo de reanálise, conseguindo explicitar uma resposta adequada, que demonstrasse percepção dos dados. Assim, a habilidade de confronto entre hipótese e dados reais revela também uma reaproximação com o que o gráfico apresenta, o que antes não foi percebido ou explicitado coerentemente, possibilitando uma nova chance para a aquisição do conhecimento ali exposto.

5.1.1.4 Avaliação de conclusões

Na avaliação de conclusão dada em situação univariada (Q4) houve avanços significativos ($t(31) = 2,521$; $p = 0,017$) no uso de justificativas consideradas adequadas para a questão proposta. Nas figuras que seguem (Figura 5.12 e Figura 5.13) exemplifica-se o avanço na compreensão do mesmo estudante entre pré-teste e pós-teste.

Q4: A partir do gráfico, é possível afirmar que economizar água em casa resolve o problema de falta de água? Por quê?

Sim, porque um exemplo nós aqui gastamos água enquanto no sertão eles precisam de água

[Sim, porque um exemplo nós aqui gastamos água enquanto no sertão eles precisam de água].

Figura 5.12 – Sujeito (S21) não avalia conclusão a partir dos dados no pré-teste.

Q4: A partir do gráfico, é possível afirmar que as pessoas usam mais a internet do celular em momentos de trabalho? Por quê?

Não, porque não pode dormir no trabalho.

[Não, porque não pode dormir no trabalho].

Figura 5.13 – Sujeito (S21) avalia conclusão a partir dos dados no pós-teste.

Nos exemplos anteriores, a habilidade de avaliação de uma conclusão proposta (Q4), a princípio, mostrou-se baseada na vivência do estudante, ou seja, suas crenças foram mais importantes que as informações estatísticas do gráfico. Diferente do que aconteceu no pós-teste, o estudante já considera que as categorias do gráfico, sobre os momentos de maior uso da internet dos celulares, não estão relacionadas com ações esperadas em ambientes de trabalho.

Quanto à mesma habilidade, avaliação de conclusão dada, em situação de correlação (CQ4), os resultados mostraram-se bastante distintos. Os estudantes do 5º ano apresentaram grandes dificuldades de justificarem suas respostas de modo adequado, havendo inclusive uma diminuição percentual no pós-teste. Entretanto, a análise estatística possibilitada pelo T teste indicou que a variação entre pré-teste e pós-teste não apresentou uma diferença significativa ($t(31) = 0,273$; $p = 0,786$). Os exemplos que seguem (Figura 5.14 e Figura 5.15) referem-se ao desempenho de um mesmo sujeito e evidenciam a adequação da escrita argumentativa do estudante.

Q4: A partir do gráfico, é possível afirmar que o risco de fumar e ter câncer de pulmão é pequeno? Por quê?

não, porque 10% das pessoas que não fumam tem câncer de pulmão e 90% das pessoas que fumam tem.

[Não, porque 10% das pessoas que não fumam tem câncer de pulmão e 90% das pessoas que fumam tem].

Figura 5.14 – Sujeito (S08) que avalia conclusão (CQ4) e argumenta no pré-teste.

Q4: De acordo com o gráfico, pode-se dizer que o risco de ser hipertenso e estar acima do peso ideal é pequeno? Por quê?

não porque o risco é muito alto.

[Não, porque o risco é muito alto].

Figura 5.15 – Sujeito (S08) que avalia conclusão (CQ4) no pós-teste.

Verifica-se na Figura 5.14 que o estudante elabora uma justificativa contemplando os dados do gráfico, o que é importante, de acordo com o proposto na questão. Todavia, o mesmo estudante no pós-teste (Figura 5.15) expõe uma

resposta adequada para a questão dada, sem mobilizar dados do gráfico que reforcem sua explicitação, o que seria esperado, uma vez que o sujeito já mostrou no pré-teste a capacidade de justificar registrando dados do gráfico lido.

Nas questões relacionadas à habilidade de avaliação de conclusão, crenças emergem com grande facilidade, como é perceptível na Figura 5.12. De acordo com Nickerson (1998), quando as pessoas lidam com informações estatísticas tendem a dar maior peso àquilo que reforça crenças, descartando evidências dos dados que vão de encontro às ideias preconcebidas do sujeito. Nunes e Bryant (2011) salientam que esse processo baseado em crenças acontece tanto com crianças como com adultos. Nesse sentido, os resultados aqui alcançados sugerem que os estudantes podem estar selecionando registrar em suas justificativas aquilo que do gráfico confirma sua ideia original sobre um tema. Todavia, os resultados mostram também que esses estudantes são capazes de modificar condutas em relação aos dados e tomada de decisão com intervenções adequadas. Afinal, nas situações de dados univariados, principalmente, houve avanço significativo dos sujeitos.

5.1.1.5 Uso de linguagem probabilística a partir de predições

O uso de linguagem que demostre incerteza (Q5), quando os estudantes precisaram realizar predições, não diferiu entre o pré-teste e o pós-teste ($t(31) = 0,902$; $p = 0,374$). Cabe retomar, que o uso de uma linguagem probabilística, ou seja, com consideração da incerteza inerente às situações estatísticas precisaria contar com expressões que fossem além de falas do uso cotidiano, as quais podem ser percebidas como vícios de linguagem, como é o caso da expressão “eu acho”. Desse modo, a Figura 5.16 exemplifica a resposta do estudante que não utiliza uma linguagem que contemple a incerteza da previsão e a Figura 5.17 é um exemplo de resposta que justifica adequadamente o que foi requerido.

Q5: Depois de conhecer os dados do gráfico, você acredita que os momentos em que as pessoas mais usam a internet do celular podem ser diferentes em 2017? Por quê?

Eu acho que vai ser a mesma coisa, porque não vai mudar nada

[Eu acho que vai ser a mesma coisa, porque não vai mudar nada].

Figura 5.16 – Sujeito (S25) que não utiliza linguagem probabilística ao prever no pré-teste.

Q5: Depois de conhecer os dados do gráfico, você acredita que o modo que se usa a água no país pode ser muito diferente em 2017? Por quê?

Pode ser diferente por que em 2017 a gente pode economizar mais.

[Pode ser diferente, porque em 2017 a gente pode economizar mais].

Figura 5.17 – Sujeito (S29) que utiliza linguagem probabilística ao prever no pós-teste.

Na Figura 5.16 é possível verificar que o estudante afirma que nada vai mudar com o passar do tempo, apesar de usar o termo “eu acho”. Já na Figura 5.17 há um exemplo de resposta em que o estudante afirma “pode ser diferente”, linguagem considerada de incerteza. Cumpre salientar que o mesmo sujeito que oferece uma resposta adequada no pós-teste, deixou a questão em branco no pré-teste, o que é um salto qualitativo importante, quando se pensa na individualidade do sujeito.

A literatura já ressalta as dificuldades de estudantes dos anos finais do Ensino Fundamental quando são avaliados a respeito da utilização de linguagem apropriada às situações estatísticas, mesmo após intervenções de ensino (HENRIQUES; OLIVEIRA, 2015). Isso indica que a escola pode estar contribuindo para o processo determinístico de certezas, o qual invalida a dúvida por parte do aluno, ou seja, o estudante pode estar sendo estimulado no ensino a expressar apenas linguagens que demonstrem que ele detém o saber. Ideia reforçada quando se compara os resultados de pesquisas que analisaram o uso de linguagem probabilística por estudantes maiores com aqueles resultados alcançados por McPhee e Makar (2014), com crianças entre 5 e 6 anos de idade, as quais conseguiram fazer uso da linguagem de incerteza em predições ao trabalharem em duplas.

Nesse sentido, emerge da análise realizada das justificativas dos estudantes do 5º ano, que passaram pelo processo de intervenção, elementos de conhecimento e elementos disposicionais, tal qual proposto no modelo de Letramento Estatístico de Gal (2002). Afinal, constatou-se que houve evolução por parte sujeitos de muitas das habilidades investigadas e até mesmo naquelas habilidades que não contaram com avanços significativos, é possível verificar casos particulares de estudantes que

não haviam respondido no pré-teste e conseguiram no pós-teste oferecer respostas adequadas, como já se discutiu.

Gal (2002) estabelece que a argumentação diante de informações estatísticas é uma habilidade necessária de ser desenvolvida. Habilidade que está contemplada em seu modelo para o Letramento Estatístico (elementos do conhecimento e disposicionais).

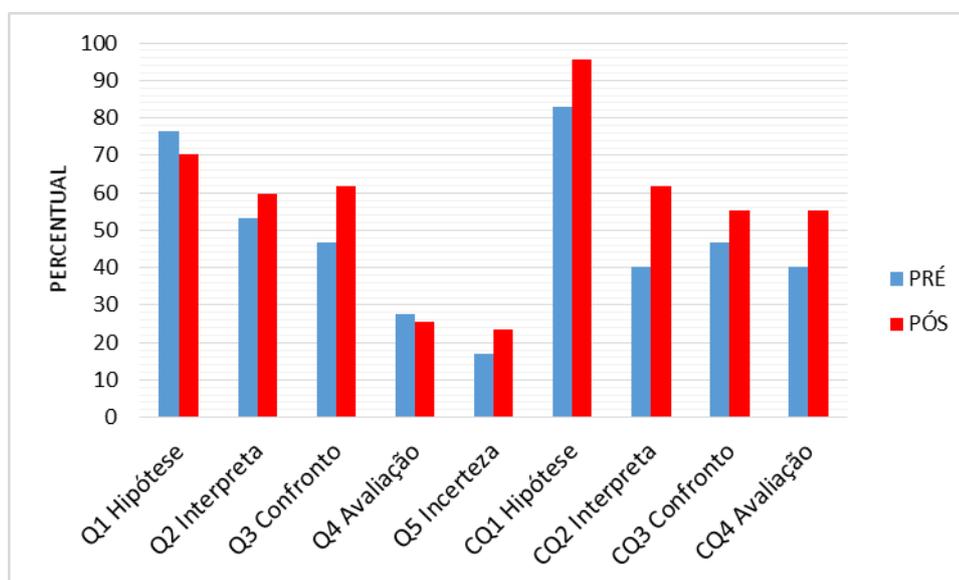
Compreende-se que quando o autor cita habilidades de letramento, a elaboração de justificativas e argumentos pertinentes aos dados reais veiculados integram-se a esse tipo de conhecimento. Conhecimento estatístico e conhecimento matemático são fundamentais na interpretação de gráficos, como os explorados no estudo e os estudantes conseguiram melhorar nesse sentido. Quanto ao conhecimento de contexto proposto por Iddo Gal, tal conhecimento é essencial ao levantamento de hipótese, o que foi significativo para os estudantes, em situações de dados bivariados. Dos elementos disposicionais (crenças, atitudes e postura crítica), o autor salienta que os mesmos atuam de modo integrado com os demais, emergindo quando o estudante se sente inclinado a ativar os conhecimentos cognitivos descritos em seu modelo e compartilha com outros seu entendimento dos dados. Verificamos que, de fato, os estudantes comunicaram crenças e evoluíram em seus conhecimentos cognitivos durante o processo investigativo que vivenciaram.

5.1.2 Aprendizagem no 7º ano diante das justificativas

Da mesma forma que analisamos as explicações das respostas dadas pelos alunos do 5º ano, passamos agora a analisar as explicações dos alunos do 7º ano. Como dito anteriormente, o registro da explicitação por parte do sujeito foi um elemento importante como análise da compreensão dos estudantes, mas também é uma habilidade em si, uma vez que alguns estudantes sabiam o que a questão proposta requeria, contudo não sabiam expressar o entendimento dos mesmos a partir do registro escrito solicitado.

Desse modo, no Gráfico 5.3 está apresentada a evolução das justificativas consideradas apropriadas, quando comparamos pré-teste e pós-teste.

Gráfico 5.3 - Percentual de justificativa adequada em cada habilidade no pré-teste e pós-teste do G7I



Fonte: o autor.

Verifica-se no Gráfico 5.3 que em todas as habilidades investigadas houve aumento percentual na argumentação oferecida pelos estudantes do 7º ano, após a intervenção, com exceção do levantamento de hipótese (Q1) e da avaliação de conclusão (Q4), ambas em situações de dados univariados. O tipo de situação, se de dados univariados ou dados bivariados (correlação), parece ter influenciado mais no comportamento desse grupo de sujeitos. A fim de maiores esclarecimentos acerca do desempenho desse grupo de estudantes, cumpre o estabelecimento de análises estatísticas mais consistentes, para isso o T teste foi utilizado.

A seguir, análises mais específicas de cada habilidade explorada no estudo são exploradas, com explicitação dos devidos resultados estatísticos de variância e exemplos de respostas que evidenciam as análises realizadas.

5.1.2.1 Levantamento de hipótese

Apesar dos percentuais mostrados no gráfico (Gráfico 5.3), com uma pequena diminuição do pré-teste para o pós-teste, a habilidade de levantamento de hipótese não revelou uma diferença significativa nas situações de dados univariados (Q1) ($t(46) = 0,724$; $p = 0,473$). Isso quanto ao uso de justificativas adequadas ao que a questão solicita. Importante retomar que crenças influenciam bastante os estudantes

na pesquisa estatística. Além disso, os mesmos são solicitados a registrarem uma resposta que está no campo da incerteza, ou seja, o que os estudantes poderiam responder seriam suposições a partir de seus conhecimentos prévios. O levantamento de hipótese solicitou uma justificativa, contudo, não se tinha conhecimentos dos dados reais para que uma resposta baseada em evidências de uma pesquisa fosse exposta. A Figura 5.18 e a Figura 5.19 exemplificam, respectivamente, a resposta de um estudante que não consegue elaborar uma justificativa apropriada ao requerido na questão e a resposta de um estudante que o faz.

Q1: Em quais momentos do dia você acredita que as pessoas mais utilizam a internet dos celulares? Por quê?

Todos os dias, porque precisamos dela para mantermos informados e cientes do que está acontecendo.

[Todos os dias, porque precisamos dela para (nos) mantermos informados e cientes do que está acontecendo].

Figura 5.18 – Sujeito (S83) que não levanta hipótese (Q1) no pré-teste.

Q1: Em quais momentos do dia você acredita que as pessoas mais utilizam a internet dos celulares? Por quê?

de manhã. Por que a pessoa de manhã quando se acorda sempre quer ver as novidades que aconteceram e sair informado em tudo.

[De manhã, porque a pessoa de manhã quando se acorda sempre quer ver as novidades que aconteceram e sair informado em tudo].

Figura 5.19 – Sujeito (S54) que levanta hipótese (Q1) no pré-teste.

Nos exemplos expostos anteriormente, percebe-se que o estudante na Figura 5.18 não cita qualquer momento do dia de provável maior uso da *internet* dos celulares. Já o estudante na Figura 5.19, para o mesmo contexto, seleciona um período do dia, que era a expectativa.

Quanto ao levantamento de hipótese nas situações de correlação, a argumentação dos estudantes do 7º ano também não apresentou diferença

significativa (CQ1) ($t(46) = 1,953$; $p = 0,057$). Contudo, ressalta-se que os mesmos já contavam no pré-teste com um percentual alto de justificativas adequadas (83%) e o número cresceu no pós-teste (95,7%) demonstrando que quase todos os alunos conseguiram realizar um levantamento de hipótese com justificativa adequada. A Figura 5.20 explicita a resposta oferecida por um sujeito que não conseguiu levantar hipótese argumentando adequadamente e a Figura 5.21 um exemplo de resposta bem formulada para a habilidade investigada.

Q1: Para você o que o câncer de pulmão tem a ver com o hábito de fumar cigarros, cachimbos, charutos e outros (produtos com nicotina)? Explique.

porque o cigarro contém infecção no pulmão e doenças no coração e os churrascos as vezes faz mal por causa da fumaça

[Porque o cigarro contém infecção no pulmão e doenças no coração e os churrascos as vezes faz mal por causa da fumaça].

Figura 5.20 – Sujeito (S68) que não levanta hipótese (CQ1) no pós-teste.

Q1: Para você o que a hipertensão arterial (conhecida como pressão alta) tem a ver com o peso (massa corpórea) das pessoas? Explique.

nada, porque minha mãe é magrinha e tem pressão alta e o vizinho é gordo também e tem

[Nada, porque minha mãe é magrinha e tem pressão alta e o vizinho é gordo também e tem].

Figura 5.21 – Sujeito (S88) que levanta hipótese (CQ1) no pós-teste.

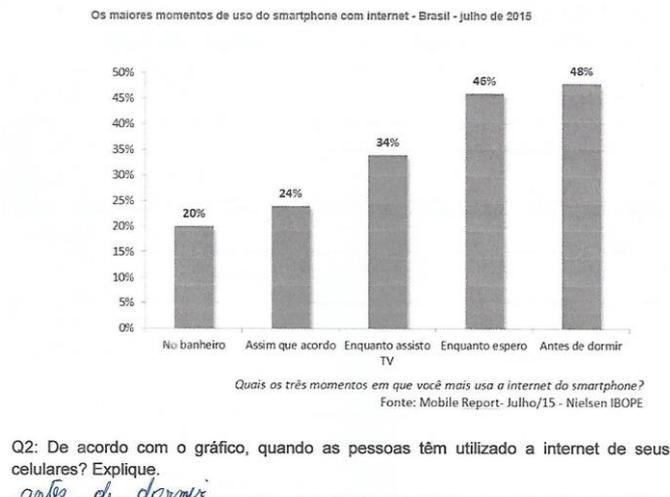
Na Figura 5.20, o estudante demonstra crer que o cigarro traz malefícios à saúde, mas não consegue elaborar uma argumentação que expresse com clareza suas ideias. Por outro lado, na Figura 5.21, o estudante evidencia que percebeu as duas variáveis presentes na questão, apesar de não concordar que elas estejam relacionadas, seja direta, seja indiretamente. Assim, o estudante no último exemplo, guiado por experiências de vida, nega a relação entre massa corpórea e hipertensão arterial. Em se tratando de contexto, salienta-se que essa atividade (correlação entre massa corpórea e hipertensão) causou maior desacordo entre os estudantes, do que

a atividade de correlação entre câncer de pulmão e hábito de fumar produtos com nicotina, uma vez que a última relação é bastante enfatizada no cotidiano.

Emergem, nesse sentido, conhecimentos de contexto, como descritos por Gal (2002). English (2014) em estudo longitudinal com crianças dos anos iniciais já apontou a interferência do contexto na habilidade dos estudantes de lidarem com variabilidade e predição. Além disso, podemos citar o papel das crenças influenciando as hipóteses elaboradas pelos sujeitos (NICKERSON, 1998; NUNES; BRYANT, 2011).

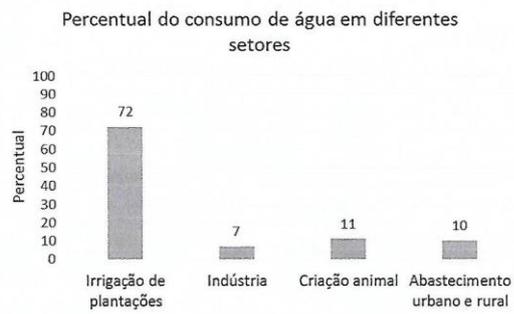
5.1.2.2 Interpretação dos dados reais

As justificativas oferecidas pelos estudantes do 7^o ano ao interpretarem dados em situação de dados univariados (Q2) não apresentou uma melhora significativa ($t(46) = 0,621$; $p = 0,537$), como seria o desejado. Bem diferente do que se constatou com os estudantes do 5^o ano a respeito desta mesma análise. Acredita-se que a resistência dos estudantes maiores ao registro por meio da escrita (o solicitado em todos os testes do estudo) interferiu nos resultados alcançados. Essa discussão será retomada mais adiante. Como exemplificação de resposta considerada inadequada está a Figura 5.22, o mesmo estudante, já no pós-teste (Figura 5.23) consegue avançar nas justificativas.



[Antes de dormir].

Figura 5.22 – Sujeito (S80) que não justifica ao interpretar (Q2) no pré-teste.



Fonte: Conjuntura de Recursos Hídricos do Brasil, ANA, 2013.

Q2: De acordo com o gráfico, como tem sido a distribuição do uso de água em diferentes setores do país? Explique.

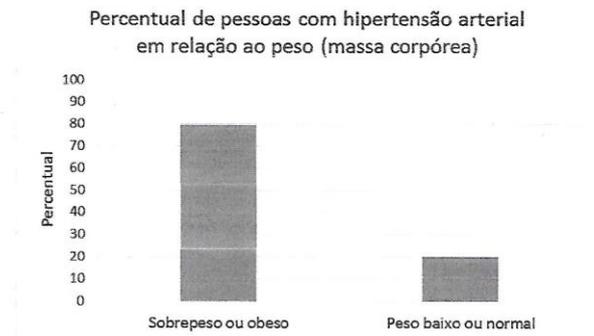
A água é gastada mais com a irrigação e pouco com outras coisas como indústria e etc.

[A água é “gastada” mais com a irrigação e pouco com outras coisas como indústria e etc.].

Figura 5.23 – Sujeito (S80) que justifica ao interpretar (Q2) no pós-teste.

Nos exemplos anteriores, vê-se que o estudante durante o pré-teste (Figura 5.22) oferece uma resposta que expõe o ponto máximo do gráfico, diante de uma questão que envolvia uma visão mais global dos dados representados, além disso, não há uma explicação que ajude na compreensão das ideias expressas pelo sujeito. No pós-teste, o mesmo estudante (Figura 5.23) já consegue desenvolver uma resposta que vai além do ponto máximo, relacionando categorias do gráfico em sua escrita.

Quando os dados foram bivariados (CQ2), houve grande avanço dos estudantes do 7º ano. A comparação entre pré-teste e pós-teste nas justificativas dos estudantes revelou crescimento significativo ($t(46) = 2,483$; $p = 0,017$). A Figura 5.24 e a Figura 5.25 evidenciam o desenvolvimento do mesmo estudante entre pré-teste e pós-teste.



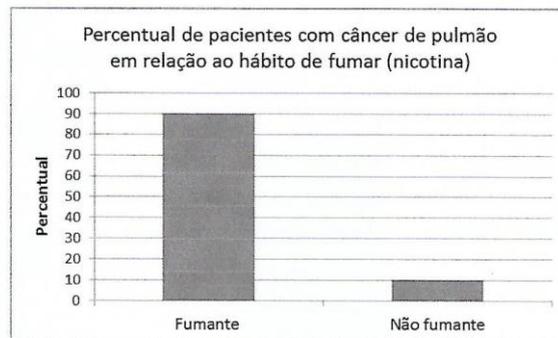
Fonte: Revista Brasileira de Medicina do Trabalho, 2009

Q2: De acordo com o gráfico, existe relação entre hipertensão e peso (massa corpórea)? Explique.

Sim, porque é mais comum os mais velhos ter hipertensão.

[Sim, porque é mais comum os mais velhos ter hipertensão].

Figura 5.24 – Sujeito (S67) que não interpreta no pré-teste.



Fonte: Instituto Nacional de Câncer, INCA.

Q2: De acordo com o gráfico, existe relação entre os pacientes que tem câncer de pulmão e o hábito de fumar ou não fumar? Explique.

Pode ser que a maioria que fuma tem câncer de pulmão.

[A maioria que fuma tem câncer de pulmão].

Figura 5.25 – Sujeito (S67) que interpreta e justifica no pós-teste.

Percebe-se na Figura 5.24, que apesar do estudante ser solicitado a interpretar os dados de um gráfico que representa a relação entre duas variáveis, o mesmo oferece uma resposta que pode ser considerada imaginativa, mas que na verdade faz parte das crenças do sujeito, pois a afirmação elaborada não condiz com as informações representadas. Cavalcanti (2011) em estudo diagnóstico com

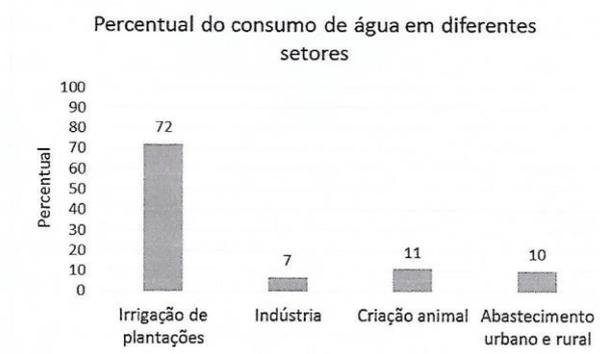
estudantes do Ensino Fundamental já havia apontado que os sujeitos ao serem solicitados a tomarem decisões baseados nos dados tendem a justificarem de acordo com o que acreditam ser “certo” ou “justo”, quando desconsideram os dados representados. Na Figura 5.25, o mesmo estudante expõe a relação entre as duas variáveis e utiliza uma linguagem que está em acordo com o que os dados mostram, ao afirmar que a maioria fuma e tem câncer de pulmão. É importante o uso da expressão “a maioria”, pois revela compreensão dos dados, não estendendo o que acontece em uma categoria mais frequente para todos os sujeitos pesquisados.

5.1.2.3 Confronto entre hipóteses e dados reais

O confronto entre a hipótese levantada e os dados reais (Q3) nas situações de dados univariados não revelou diferença significativa entre pré-teste e pós-teste ($t(46) = 1,854$; $p = 0,070$), mesmo com o aumento percentual verificado. A Figura 5.26 é uma evidência de sujeito que consegue realizar o confronto entre hipótese e dados no pós-teste.

Q1: Para você a água consumida no Brasil é usada em maior quantidade onde (casas, indústrias, plantações, criação de animais)? Explique.

casas, porque a gente consome muita
água diariamente.



Fonte: Conjuntura de Recursos Hídricos do Brasil, ANA, 2013.

Q2: De acordo com o gráfico, como tem sido a distribuição do uso de água em diferentes setores do país? Explique.

Tem sido diferente nas plantações tem sido mais e nas indústrias menos.

Q3: Os dados do gráfico apresentam informações parecidas com as que você imaginou na primeira questão? Por quê?

Não, porque eu achava o consumo maior era nas casas, mas são nas plantações.

[Q1: Casas, porque a gente consome muita água diariamente].

[Q2: Tem sido diferente, nas plantações tem sido mais e nas indústrias menos].

[Q3: Não, porque eu achava o consumo maior era nas casas, mas são nas plantações].

Figura 5.26 – Sujeito (S56) que realiza confronto entre hipótese e dados (Q3) no pós-teste.

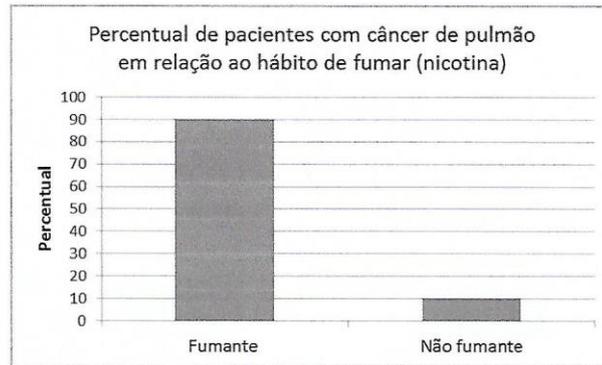
Percebe-se na Figura 5.26 que o estudante se utiliza de crenças pessoais para fazer sua escolha e justificá-la no levantamento da hipótese. Contudo, ao tomar conhecimento que os dados representados no gráfico vão de encontro ao que ele acreditava, o estudante reconhece o desacordo, externando na sua argumentação o que ele constatou.

Em se tratando de dados bivariados, as justificativas oferecidas pelos estudantes do 7º ano diante do confronto entre hipótese e dados reais (CQ3) também não resultou em diferença significativa ($t(46) = 0,942$; $p = 0,351$) na comparação entre os dois testes realizados pelos estudantes. A Figura 5.27 é um

exemplo de resposta que demonstra a ausência de inter-relação na questão de confronto (CQ3) entre a hipótese que o estudante expôs e o que os dados mostram.

Q1: Para você o que o câncer de pulmão tem a ver com o hábito de fumar cigarros, cachimbos, charutos e outros (produtos com nicotina)? Explique.

Porque os cigarros, contém vários tipos de coisas, que fazem mal a saúde, coisas que fazem muito mal a saúde.



Fonte: Instituto Nacional de Câncer, INCA.

Q2: De acordo com o gráfico, existe relação entre os pacientes que tem câncer de pulmão e o hábito de fumar ou não fumar? Explique.

Sim o fumante ele corre o risco de pegar o câncer de pulmão por causa dos produtos que são colocados no cigarro.

Q3: Os dados do gráfico apresentam informações parecidas com o que você imaginou na primeira questão? Por quê?

Sim porque quase toda a população fuma o cigarro.

[CQ1: Porque os cigarros contêm vários tipos de coisas que fazem mal à saúde, coisas que fazem muito mal à saúde].

[CQ2: Sim, o fumante ele corre o risco de pegar o câncer de pulmão por causa dos produtos que são colocados no cigarro].

[CQ3: Sim, porque quase toda a população fuma o cigarro].

Figura 5.27 – Sujeito (S69) que não faz confronto entre hipótese e dados (CQ3) no pós-teste.

Verifica-se na Figura 5.27 que o estudante levanta uma hipótese que de certa forma relaciona o hábito de fumar produtos com nicotina a prejuízos à saúde,

embora haja dificuldades em se justificar a resposta de modo claro. Ao ser solicitado que respondesse e explicasse sobre os dados representados no gráfico, o estudante não consegue extrair do gráfico informações que lhe auxiliem na argumentação e no confronto (CQ3) não se constata um processo de reanálise dos dados ou retomada para discussão do que foi registrado na questão de hipótese. Tal atividade foi proposta no pós-teste e torna evidente o papel desempenhado pelas crenças, uma vez que mesmo após repetidos momentos de análise dos dados, como solicitado nas questões seguintes à exposição do gráfico, o estudante seleciona dos dados o que reforça suas vivências (NICKERSON, 1998), não o considerando integralmente. Além disso, a temática que explora a relação entre as variáveis câncer de pulmão e hábito de fumar produtos com nicotina é uma correlação bastante enfatizada no cotidiano com uso de propagandas advertindo sobre os males da nicotina para a saúde, fazendo com que muitos estudantes generalizem, afirmando que todos que fumam têm câncer, o que não é uma conclusão apropriada de acordo com os dados reais do gráfico. Outro efeito dessa propaganda é a relação de causalidade estabelecida entre as variáveis, de acordo com Daniels (2017), o que não é recomendável em se tratando de pesquisa estatística. Daniels (2017), assim como Triola (2013) e Friel, Curcio e Bright (2001) alertam para o erro de se estabelecer uma relação causa/efeito quando na interpretação de correlação.

5.1.2.4 Avaliação de conclusões

A justificativa oferecida pelos estudantes do 7º ano quando avaliaram uma conclusão dada, em situações de dados univariados (Q4) não revelou diferença significativa entre os testes ($t(46) = 0,330$; $p = 0,743$). Vê-se que a habilidade se mostrou de difícil compreensão para os estudantes de ambos os anos de escolaridade investigados. Apesar disso, foi possível encontrar exemplos de estudantes que não conseguiam realizar a tarefa solicitada (Figura 5.28) e que no pós-teste alcançaram um nível de argumentação adequado ao que foi exigido na questão (Figura 5.29).

Q4: A partir do gráfico, é possível afirmar que economizar água em casa resolve o problema de falta de água? Por quê?

Sim, porque os moradores não colaboram.

[Sim, porque os moradores não colaboram].

Figura 5.28 – Sujeito (S67) que não avalia conclusão (Q4) no pré-teste.

Q4: A partir do gráfico, é possível afirmar que as pessoas usam mais a internet do celular em momentos de trabalho? Por quê?

Não, porque não tem nem dizendo que no gráfico que usamos em local de trabalho.

[Não, porque não tem nem dizendo no gráfico que usamos em local de trabalho].

Figura 5.29 – Sujeito (S67) que avalia conclusão (Q4) no pós-teste.

Percebe-se na Figura 5.28 que o estudante não se baseia nos dados para justificar sua resposta, o que é solicitado. O que faz é produzir uma resposta fundamentada no discurso amplamente divulgado de que as pessoas consomem muita água em casa. Já no pós-teste (Figura 5.29), o mesmo estudante passa a utilizar os dados do gráfico como argumento para a sua resposta.

Quando os dados foram bivariados, a avaliação de conclusão (CQ4) pelos estudantes também não alcançou uma melhora significativa ($t(46) = 1,478$; $p = 0,146$), apesar do aumento percentual entre os testes. A seguir, são evidenciados exemplos de respostas de um estudante que demonstra superação de dificuldades entre o pré-teste (Figura 5.30) e pós-teste (Figura 5.31).

Q4: De acordo com o gráfico, pode-se dizer que o risco de ser hipertenso e estar acima do peso ideal é pequeno? Por quê?

Sim, por conta da alimentação

[Sim, por conta da alimentação].

Figura 5.30 – Sujeito (S100) que não avalia conclusão (CQ4) no pré-teste.

Q4: A partir do gráfico, é possível afirmar que o risco de fumar e ter câncer de pulmão é pequeno? Por quê?

Não. Porque quando não é fumante os riscos são menores e quando é fumante os riscos são maiores.

[Não, porque quando não é fumante os riscos são menores e quando é fumante os riscos são maiores].

Figura 5.31 – Sujeito (S100) que avalia conclusão (CQ4) no pós-teste.

O estudante no primeiro momento (Figura 5.30) oferece uma resposta equivocada, uma vez que o risco de ser hipertenso e estar acima do “peso” ideal é grande, de acordo com o gráfico. Além disso, o gráfico não traz informações sobre alimentação, o que torna a explicação dada imaginativa. Já no momento seguinte, no pós-teste (Figura 5.31), o mesmo estudante contempla informações do gráfico em sua argumentação.

A revisão da literatura não apresenta dados que possam ser comparados com os resultados aqui alcançados, os quais tratam-se da análise do desempenho dos estudantes diante de dados reais não produzidos pelos estudantes. Diferente dos estudos realizados por Makar (2013) e Allmond e Makar (2014), os quais partiram de situações investigativas de coleta de dados pelos próprios estudantes. Tais pesquisas constataram que houve melhoras importantes nas compreensões dos estudantes ao tirarem conclusões tomando-se os dados como evidências. Ressalta-se que refletir sobre os dados reais em situação de leitura e não apenas de produção é imprescindível ao mundo de constantes tomada de decisões justificadas por pesquisas estatísticas de fato ou assim veiculadas. Caseiro, Ponte e Monteiro (2014) ao investigarem professor e estudantes do Ensino Fundamental expõem situação de avaliação de conclusão para uma afirmação dada, mas não apresentam resultados do desempenho dos estudantes, o foco esteve na intervenção do professor elaborando questões para seus alunos.

5.1.2.5 Uso de linguagem probabilística a partir de predições

As justificativas dadas pelos estudantes do 7º ano, quando solicitados a preverem acontecimentos futuros (Q5), após análise dos dados do gráfico, não resultou avanços significativos ($t(46) = 0,903$; $p = 0,371$) na linguagem utilizada pelos sujeitos. Cumpre retomar que demonstrar incerteza diante de informações estatísticas é o desejável, porém, o aumento percentual entre pré-teste e pós-teste não revelou significância estatística. Os exemplos que seguem mostram o desempenho do mesmo estudante no pré-teste (Figura 5.32) e pós-teste (Figura 5.33).

Q5: Depois de conhecer os dados do gráfico, você acredita que os momentos em que as pessoas mais usam a internet do celular podem ser diferentes em 2017? Por quê?

Sim. Porque as coisas sempre evoluem.

[Sim, porque as coisas sempre evoluem].

Figura 5.32 – Sujeito (S85) que não utiliza linguagem probabilística no pré-teste.

Q5: Depois de conhecer os dados do gráfico, você acredita que o modo que se usa a água no país pode ser muito diferente em 2017? Por quê?

Sim a tecnologia pode mudar muitas coisas.

[Sim, a tecnologia pode mudar muitas coisas].

Figura 5.33 – Sujeito (S85) que utiliza linguagem probabilística no pós-teste.

Percebe-se que o estudante na Figura 5.32 afirma de modo determinístico que as coisas “sempre” evoluem, nesse sentido, não há espaço na argumentação do estudante para a dúvida ou possibilidade de um resultado diferente do imaginado. No pós-teste (Figura 5.33), o estudante já apresenta outra postura e utiliza o termo “pode mudar”, o que caracteriza uma visão probabilística da predição em estatística.

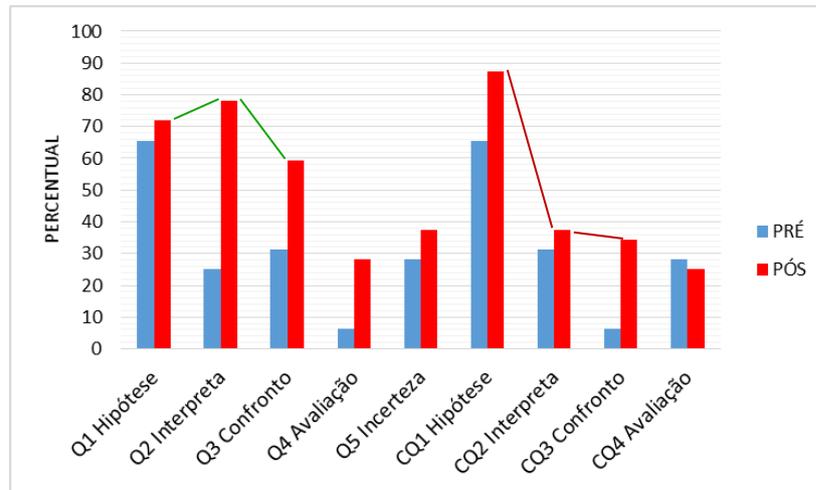
Com os resultados descritos e discutidos até aqui verifica-se que os estudantes do 5º ano, pertencentes ao grupo de intervenção (G5I), conseguiram crescer mais em suas aprendizagens, do que os estudantes participantes da intervenção no 7º ano (G7I). Isso porque as justificativas adequadas dos estudantes do 5º ano partiram de percentuais menores no pré-teste, em praticamente todas as

questões investigadas, excetuando na questão que tratou da habilidade de uso de linguagem probabilística (Q5) ao predizer, na qual o 5º ano contou com 28,1% e o 7º ano com 17%. Após o período de intervenção, ao realizarem o pós-teste, constatou-se que os estudantes do 5º ano conseguiram avanços significativos em 5 (cinco) das 9 (nove) questões: no levantamento de hipótese em dados bivariados (CQ1); na interpretação de dados reais univariados (Q2); no confronto entre hipóteses e dados reais em dados univariados (Q3) e em dados bivariados (CQ3); e na avaliação de conclusão a partir de dados univariados (Q4). Já os estudantes do 7º ano apresentaram desempenho significativamente diferente na justificativa oferecida no pós-teste apenas na questão de interpretação de dados reais bivariados (CQ2).

Não é o objetivo nessa discussão comparar estudantes do 5º e do 7º ano, os quais estão inseridos em contextos específicos. A intenção é demonstrar que ambos os grupos aprenderam. Entretanto, não partiram de um mesmo saber, nem avançaram nos mesmos aspectos investigados. Os grupos tiveram caminhos próprios, os quais são importantes de serem ressaltados e compreendidos para futuras pesquisas e intervenções.

A respeito dos caminhos próprios citados, vê-se que a evolução no pós-teste da compreensão de habilidades investigadas aconteceu de modo singular em cada grupo de intervenção. No 5º ano, as habilidades envolvidas com o processo de confronto entre hipótese e dados reais (Q1, Q2 e Q3) univariados parecem ter se desenvolvido de maneira a se ter uma menor variação no percentual de justificativas adequadas no pós-teste. Já nos dados bivariados, tais habilidades não compartilharam dessa proximidade. A seguir, os dados já apresentados anteriormente são resgatados, com destaques no pós-teste do que se está discutindo aqui (Gráfico 5.4).

Gráfico 5.4 – Justificativas dos estudantes do 5º ano em questões dos testes, com destaque no pós-teste.

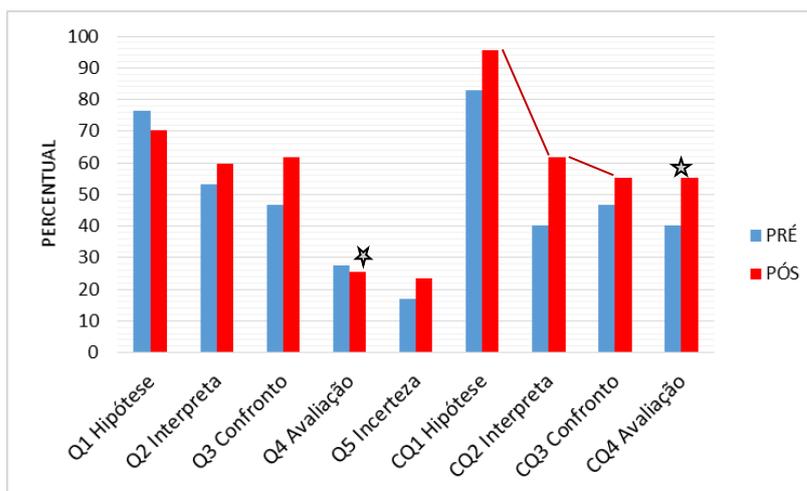


— Dados univariados.
— Dados bivariados.

Fonte: o autor.

Quanto ao 7º ano, também parece haver uma proximidade maior no percentual de justificativas adequadas das questões inter-relacionadas com o confronto entre hipótese e dados reais. Entretanto, isso representa um avanço apenas nas situações de dados bivariados (CQ1, CQ2 e CQ3), quando os estudantes conseguem melhoras significativas na interpretação. Outra habilidade que mostrou desempenho bastante diferente quando se compara o tipo de situação de dados é a avaliação de conclusão. As situações de dados bivariados acarretaram justificativas mais apropriadas do que as situações de dados univariados. O Gráfico 5.5, que apresenta dados já discutidos, evidencia o que se discute e destaca as singularidades do pós-teste para o 7º ano.

Gráfico 5.5 – Justificativas dos estudantes do 7º ano em questões dos testes, com destaque no pós-teste.



— Dados bivariados.

☆ Avaliação de conclusão nas diferentes situações.

Fonte: o autor.

As análises desenvolvidas até o presente momento, quanto à justificativa dos estudantes nas diferentes habilidades investigadas no estudo e entre os testes solucionados, revelaram que os estudantes de cada grupo de intervenção se comportaram de maneira singular. Constatação reforçada quando se observa o número de questões sem justificativa, ou seja, deixadas em branco, mesmo após o processo de intervenção. Não foram muitas as questões não justificadas no pós-teste, contudo, vê-se que o 7º ano tende a oferecer respostas mais breves, comprometendo a explicitação de ideias. A questão que contou com o maior número de falta de justificativa no pós-teste foi a de interpretação de dados univariados (Q2), na qual 12,8% dos estudantes do 7º ano não responderam ou não ofereceram explicações. No 5º ano o percentual de estudantes na mesma situação foi de 6,2%.

Salienta-se que todas as questões do teste foram elaboradas de modo que reflexões, opiniões ou argumentos fossem possíveis e bem aceitos. Para isso, questões discursivas foram criadas e espaços pertinentes foram disponibilizados para registro. Entretanto, percebe-se que a produção escrita com justificação adequada nas situações investigadas não apresentou desenvolvimento com a escolaridade, considerando-se os anos pesquisados. Acredita-se que as atividades cotidianas de ensino e de avaliação da aprendizagem escolar, mais voltadas para a escolha de alternativas corretas, sem ou com pouca elaboração de respostas baseadas em produção textual influenciam no hábito de escrita dos estudantes.

Sabe-se que essa é a realidade em que os estudantes, principalmente os do 7º ano, estão inseridos.

Se por um lado o uso de atividades de produção escrita com justificativas tem sido pouco desenvolvido nas turmas investigadas, por diversas razões de cunho pedagógico, como tempo de correção de atividades, entende-se que tal investimento é essencial. Isso porque a formação de cidadãos críticos em um mundo que requer letramento perpassa a argumentação diante de informações que chegam a cada sujeito por diferentes meios midiáticos. Nesse sentido, o Letramento Estatístico faz toda a diferença para a constituição de cidadãos capazes de interpretar, avaliar, discutir e desconfiar de informação produzidas estatisticamente, como está exposto nas habilidades elencadas por Gal (2002, p.2-3). Acrescentaríamos ainda a habilidade de levantar hipóteses, visto que a mesma faz uma intersecção entre as crenças ou opiniões e as análises que o sujeito realiza dos dados que passa a conhecer.

5.2 ANÁLISE DO PROCESSO DE INTERVENÇÃO PEDAGÓGICA

O processo de intervenção surgiu do Estudo 1, que teve como objetivo analisar as compreensões demonstradas por estudantes do Ensino Fundamental (5º e 7º anos), diante de distribuições univariadas e bivariadas, ao levantarem hipóteses, analisarem dados e tirarem conclusões a partir da análise. Além disso, o resultado da análise do pré-teste, que constituiu a primeira fase do Estudo 2, foi outro elemento importante de ser considerado para ajustes nas atividades planejadas para os encontros de intervenção de ensino.

5.2.1 Encaminhamento pedagógico das habilidades do estudo nos encontros de intervenção

Os encontros de interação pedagógica com os estudantes aconteceram em duas perspectivas de mediação: a partir de discussões coletivas, com ênfase na oralidade; e também pela proposição de atividades em duplas ou trios, nas quais a oralidade existia, contudo, incentivou-se o registro em papel do que se discutia.

Foram três encontros de intervenção em quatro turmas (duas do 5º ano e duas do 7º ano), com tempo de duração aproximado de duas horas. Duas atividades foram realizadas em cada encontro, as quais se distinguiram pelo contexto do conjunto de dados explorados, tipo de situação de dados reais (univariados ou bivariados) e/ou tipo de mediação proposta, conforme já explicitado no método (seção 3).

As diferentes habilidades investigadas no presente estudo foram contempladas em todos os encontros com os estudantes que participaram da intervenção. O processo dialógico privilegiado nas aulas permitiu, ainda, que fossem abordadas questões referentes a outras habilidades não incluídas nos testes realizados pelos estudantes, como o levantamento de novas hipóteses após conhecimento dos dados. O tema de cada conjunto de dados foi determinante no sentido de se buscar questionamentos após a leitura dos dados e confrontos entre hipótese e os mesmos. Esclarece-se que uma análise exploratória dos dados representados foi o foco, a fim de fazer com que os estudantes refletissem e fossem levados a tomarem decisões baseados em evidências das análises. Entende-se que as crenças têm um papel importante nas conjecturas e até no ato de confiabilidade dos dados, contudo, o pensamento crítico estatístico requer conhecimento dos dados, assim como busca de respostas para novos questionamentos que surjam, o que caracteriza o ciclo investigativo.

Nesse sentido, cada atividade desenvolvida com os estudantes contou com questões abordando a totalidade ou quase totalidade das habilidades investigadas no estudo. As especificidades de cada conjunto de dados constituíram-se no ponto de partida para reflexões, novos questionamentos e até discordâncias.

Verificou-se que a maior ou menor interação entre os estudantes e a atividade proposta durante as intervenções esteve relacionada à própria relação estabelecida entre os sujeitos enquanto turma, independentemente do ano ao qual pertenciam. Assim, para as duas turmas do 5º ano, uma (Turma 5A) mostrou-se mais participativa, ou seja, integrada às atividades, principalmente no que diz respeito às atividades escritas. O mesmo foi constatado no 7º ano, com a Turma 7A. Vê-se que utilizar mais de um meio de discussão (oral e escrito) durante os encontros foi importante para o processo de aprendizagem dos estudantes.

Se as discussões orais beneficiaram aqueles alunos com maiores dificuldades de se expressarem de modo escrito, o que aconteceu nos momentos coletivos, a

mediação em pequenos grupos (dupla ou trio) foram essenciais para aqueles outros estudantes que não falavam abertamente para a turma suas compreensões. Isso porque nos pequenos grupos havia a solicitação de registro escrito, além disso, muitos alunos se sentiram mais confiantes para falar suas impressões, dúvidas e discordâncias na presença da pesquisadora quando a mesma se aproximava.

Outra característica que distinguiu as atividades coletivas das atividades em pequenos grupos foi o tempo que os estudantes tiveram para a proposição reflexiva de uma resposta quando uma questão foi lançada à turma ou ao pequeno grupo. Nos momentos coletivos, quando diante do silêncio que surgia perante algo que nenhum estudante se sentia confiante para falar a mediadora acabava antecipando respostas. Nos grupos menores o processo reflexivo por parte dos estudantes para a geração de respostas contou com mais tempo, uma vez que tinham que discutir e escrever.

Assim, do total de 6 (seis) atividades, duas em cada dia, quatro delas envolveram momentos coletivos, nos quais perguntas e gráficos eram apresentados em *slides* com uso de *data show*. As outras 2 (duas) atividades constituíram-se momentos em grupos pequenos, com solicitação de discussão entre integrantes e produção escrita das respostas a que os mesmos chegaram. Como foram momentos de intervenção bem distintos, que suscitaram uma postura diferenciada da pesquisadora e dos estudantes também, aqui se opta por discutir as atividades desenvolvidas em dois blocos.

5.2.1.1 Atividades coletivas

Os três encontros de intervenção contaram com momentos de atividades coletivas. No primeiro dia as duas atividades foram dirigidas coletivamente, nos demais encontros a atividade coletiva iniciou cada aula. As respostas para questionamentos que contemplassem as habilidades exploradas no estudo, além de outras perguntas que aprofundassem a análise dos dados, foram registradas no quadro e discutidas. Para as atividades coletivas, precisou-se estabelecer acordos com as turmas. Aqueles que quisessem falar levantariam a mão e os outros ficariam em silêncio para podermos ouvir o que todos diziam. Assim, todos que tivessem algo a opinar, acrescentar ou discordar teriam sua vez.

Para o levantamento de hipótese, alguns conjuntos de dados demandaram que a pesquisadora explorasse antecipadamente o tema tratado, antes de lançar a questão que levaria a elaboração de hipóteses. Isso no sentido de se possibilitar o levantamento de hipóteses, uma vez que conhecer o tema abordado é imprescindível para a habilidade. Assim, para os dados dos resíduos de agrotóxicos em amostras de alimentos (Atividade 1.1) foi necessário interpelar os estudantes a respeito do que sabiam sobre agrotóxico e esclarecer caso não soubessem. Além disso, fazer com que os estudantes pensassem sobre os alimentos apresentados no gráfico demandou que tais alimentos fossem propostos para reflexão antes da apresentação dos dados. A seguir, tem-se um trecho do diálogo com uma das turmas (Turma 5A), que exemplifica esse procedimento. Os dados divulgados são apresentados logo depois (Gráfico 5.6).

Pesquisadora: A palavra agrotóxico, vocês sabem o que ela significa?

Turma: Sim.

Pesquisadora: Diz o que é [para o aluno que levantou a mão primeiro].

Aluno: É veneno, é matar...

Pesquisadora: Alguém mais quer falar? [Silêncio]. Exatamente isso. Tóxico é matar, mas vocês sabem o que vai matar? [Silêncio].

Pesquisadora: É para matar as pragas que prejudicam plantações, certo gente? Que alimentos vocês imaginam que têm maiores quantidades de agrotóxicos?

Turma: Maçã, manga, alface couve-flor, pera, brócolis, uva, laranja, milho, cenoura, morango, caju e kiwi [O que citavam era registrado no quadro].

Pesquisadora: Desses que vocês escolheram, qual deve ter mais agrotóxico? Levanta a mão para eu contar. Cada um só pode levantar a mão uma vez, tá bom?

Turma:

Maçã	2	Laranja	
Manga		Milho	10
Alface	4	Cenoura	
Couve-flor	1	Morango	
Pera		Caju	
Brócolis		Kiwi	
Uva	1		

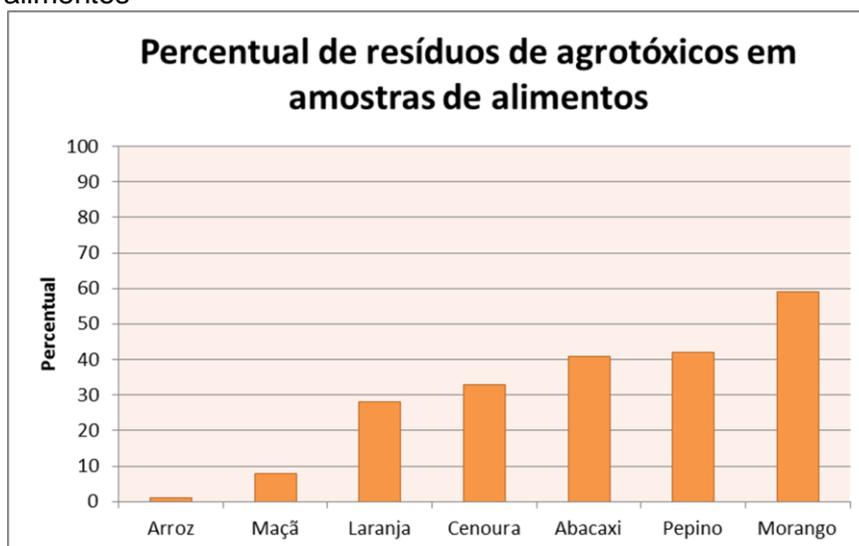
Pesquisadora: Quatro pessoas acham que é a alface... alguém quer explicar por quê? [Uma aluna levanta a mão].

Aluna: Por causa das folhas, elas atraem as pragas.

[...]

Pesquisadora: alguns desses que vocês falaram foram pesquisados realmente. [São apresentados no quadro apenas os nomes: abacaxi, arroz, cenoura, laranja, maçã, morango, pepino].

Gráfico 5.6 – Atividade 1.1 – dados sobre percentual de resíduos de agrotóxicos em amostras de alimentos



Fonte: ANVISA, Resultados de 2012.

No relato de diálogo exposto anteriormente, vê-se que algumas explicitações foram realizadas pela pesquisadora acerca da temática, uma vez que era fundamental eles entenderem o termo agrotóxico para poderem levantar hipóteses de modo coerente com o solicitado. Para que os estudantes pudessem confrontar suas hipóteses com os dados a serem apresentados, os alimentos divulgados no gráfico precisaram ser mostrados antes. Assim, estimulou-se a emergência de crenças, algumas delas já contempladas no gráfico, depois foram apresentados outros alimentos não citados livremente, mas que eram importantes de serem pensados pelos estudantes antes do conhecimento evidenciado pela pesquisa real.

Explicações a respeito do termo amostra, presente no gráfico sobre agrotóxicos em alimentos, foram necessárias a partir da apresentação dos dados. O relato que segue demonstra o diálogo realizado em uma das turmas (Turma 5B) a respeito desse conceito estatístico, uma vez que o mesmo estava diretamente

relacionado à interpretação dos dados reais. Apesar do conceito não ser o foco do estudo da tese.

Pesquisadora: A ANVISA fez uma pesquisa para saber sobre os agrotóxicos nesses alimentos aqui. A ANVISA é do governo e cuida da nossa saúde também. Olha, pesquisaram amostras em 2012. Vocês sabem o que é amostra?

Parte da turma: Não!

Aluna: Amostra é tipo um veneno.

Pesquisadora: O veneno está aqui no agrotóxico, certo? Como eles fizeram essa pesquisa? Será que pesquisaram todos os morangos? [Silêncio]. Eles pesquisaram alguns morangos, algumas cenouras, alguns pepinos... eles selecionaram alguns do Brasil para fazer a pesquisa, não podiam fazer com todos, tá?

Aluno: O morango tem 60% né professora?

Percebe-se no relato anterior que o termo estatístico “amostra” não fazia parte do repertório linguístico da turma e poucos se sentiram confiantes o suficiente para expor uma opinião sobre o significado da palavra nesse momento. Assim, a intervenção da pesquisadora com esclarecimentos foi fundamental.

Outro momento coletivo que demandou uma exploração da temática antes do levantamento de hipótese foi a Atividade 3.1, que tratou da relação entre duas variáveis: número de pessoas por sexo no Brasil e idade. Para que os estudantes levantassem hipótese para a questão “*O número de pessoas do sexo feminino e do sexo masculino no Brasil tem a ver com a idade?*”, optou-se por investigar o que eles pensavam sobre o número de crianças nascidas no Brasil de acordo com o sexo, se o número era igual ao nascer ou não. Acredita-se que, desse modo, a atenção dos estudantes na interpretação do gráfico, posteriormente, estaria mais direcionada para uma visão global dos dados, ou seja, estaria sendo estimulada a percepção de que há mudanças com o aumento da idade e não apenas a visualização de pontos extremos do gráfico. Logo, para tal fim questionou-se antecipadamente: “*Você acredita que nascem mais meninos ou mais meninas no Brasil?*”. A seguir, tem-se um trecho que evidencia a mediação realizada pela pesquisadora nesse momento da atividade (Turma 5B). Os dados apresentados após hipóteses levantadas são apresentados posteriormente (Gráfico 5.7).

Pesquisadora: Você acredita que nascem mais meninos ou mais meninas no Brasil?

Parte da turma: Meninos! [Gritam os meninos].

Aluna 1: Menina, professora!

Pesquisadora: Alguém acha que pode ser igual?

Parte da turma: Sim!

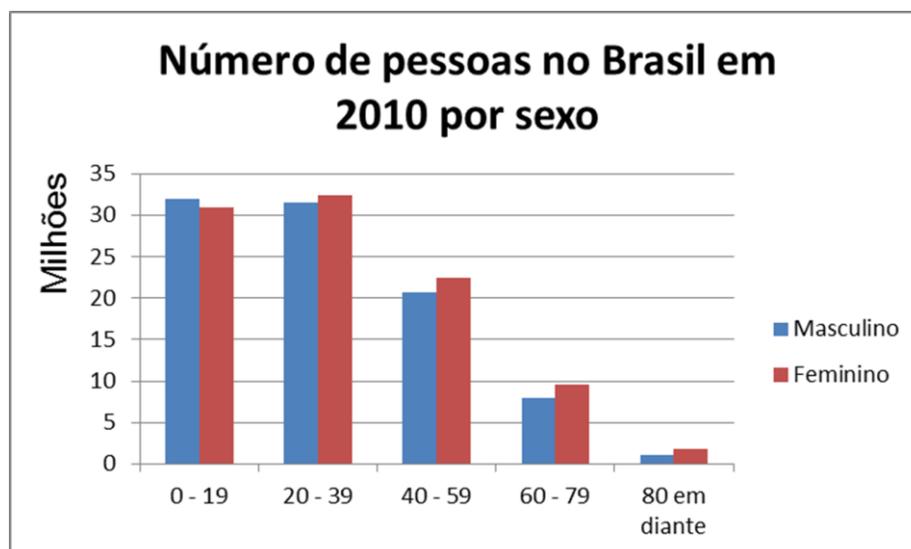
Parte da turma: Não!

Aluna 2: Não professora, eu já vi numa pesquisa que está nascendo mais menina...

Pesquisadora: Então vamos contar. Levanta a mão quem acha que é a mesma quantidade no Brasil... Duas pessoas. Quem acha que é mais menino? Um, dois, três... e meninas? [contagem] 6! Por que vocês acham isso?

[Silêncio]

Gráfico 5.7 – Atividade 3.1 – dados sobre número de pessoas no Brasil em 2010 por sexo e idade



Fonte: IBGE, Censo Demográfico 2010.

Percebe-se no trecho relatado antes a dificuldade dos estudantes expressarem uma justificativa para suas escolhas. A disputa entre meninos e meninas, potencializada pelo tema do gráfico, mostrou-se a partir dos gritos de parte da turma.

As demais atividades coletivas tiveram apenas a questão de levantamento de hipótese como introdução à atividade, foi o caso da Atividade 1.2, a respeito da relação entre as variáveis morte no trânsito e velocidade dos veículos, além da Atividade 2.1, que tratou do tempo médio do uso de aparelhos eletrônicos.

Percebe-se que algumas temáticas mobilizaram mais as discussões nas turmas do que outras, o que era esperado, pois de acordo com as vivências dos estudantes e suas crenças os mesmos poderiam se mostrar mais dispostos a expressarem opiniões em contextos conhecidos. Foi interessante observar que na interpretação dos dados, as explicações dos estudantes e justificativas para o que viam apoiavam-se em situações cotidianas. Um exemplo que explicita o exposto pode ser o trecho de diálogo que segue (Turma 5A), o qual surgiu após apresentação dos dados sobre a relação entre as variáveis morte no trânsito e velocidade dos veículos (Gráfico 5.8).

Pesquisadora: Olhem os nomes dessas cidades aqui. [Apontando para o gráfico]. Vamos ler? Os nomes estão pequenos...

Turma: Tóquio, Londres, Paris, Nova York, Copenhagen, Chicago, Miami, São Paulo, Porto Alegre, Rio de Janeiro, Brasília, Belo Horizonte, Goiânia, Teresina, Fortaleza, Porto Velho e Palmas.

Pesquisadora: De São Paulo pra baixo são capitais do Brasil e pra cima são grandes cidades também, só que de outros países.

Aluno 3: Ô tia Erica, no carnaval eu estava viajando com o meu pai e tinha uma placa que dizia o número da velocidade e meu pai estava acima. Ele estava pisando muito, muito mesmo e o carro começou a balançar e eu fiquei com medo.

Pesquisadora: Vocês deviam estar em uma rodovia, não é? Tem maior velocidade.

Aluno 3: Estava na PE 60.

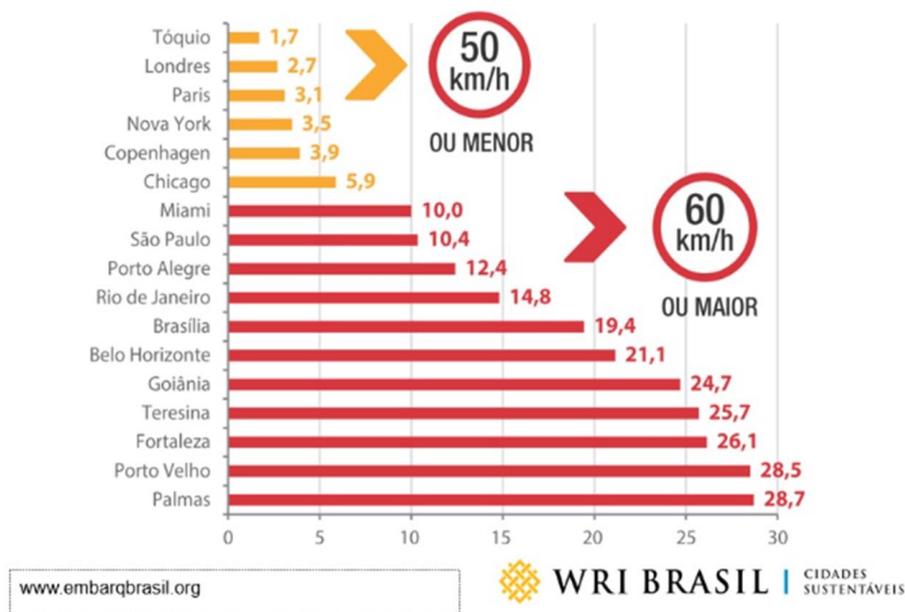
Aluna 4: Tem pessoas também que atravessam a rua no lugar errado...

Pesquisadora: É verdade, mas a responsabilidade maior é sempre do motorista. Voltando para os dados do gráfico, dá para perceber o quê?

Aluno 3: Velocidade maior tem mais mortes, tia.

Gráfico 5.8 – Atividade 1.2 – dados sobre número de mortes no trânsito para cada 100 mil habitantes e limite de velocidade

Mortes no trânsito x limites de velocidade



Fonte: World Resources Institute (WRI) Cidades Sustentáveis Brasil.

Dados sobre relação entre limite de velocidade permitida em grandes cidades e número de mortes no trânsito (para cada 100 mil habitantes).

Percebe-se no relato anterior que os estudantes entenderam o que os dados apresentavam e isso fez com que suas vivências viessem à tona. O que se revelou como forma de explicação para o acontecimento. Quando interpelados sobre o que dava para perceber dos dados, a relação entre as variáveis foi estabelecida.

Verifica-se que a pesquisadora dialogou a respeito dos conhecimentos que a turma apresentou de sua vivência acerca da temática, mas precisou direcionar as discussões para o que estava sendo apresentado no gráfico. Isto é, as discussões orais provavelmente não chegariam a expressão argumentativa com base nos dados, caso a pesquisadora não propusesse a volta ao gráfico. Tal fato sugere que os estudantes estão refletindo sobre os dados, mas precisam de assistência para exteriorizar suas compreensões de modo a considerar as relações entre as informações estatísticas.

O confronto entre hipótese e dados reais foi uma habilidade que demandou maior intervenção da pesquisadora nas atividades coletivas, devido à dificuldade de elaboração de justificativas por parte dos estudantes, pois precisavam analisar

diferenças e semelhanças entre as hipóteses levantadas na turma e as interpretações que realizaram dos dados reais. Assim, ao se questionar “*Os dados reais se parecem com o que vocês pensaram antes?*”, as turmas muitas vezes se limitaram a responder sim ou não, sem oferecer explicações, o que precisou ser indicado pela pesquisadora. Um exemplo de interferência maior na mediação pode ser o trecho, a seguir, da Turma 7A.

Pesquisadora: O que o gráfico está mostrando parece com o que vocês disseram? Está aqui ao lado o que vocês escolheram...

Parte da turma: Sim!

Pesquisadora: O que se parece?

Aluno 5: Erramos! [Grito].

Pesquisadora: Vocês não tinham como saber, gente.

Aluna 6: Falamos morango.

Pesquisadora: Só teve um voto para morango, não foi? Laranja teve dois... morango no gráfico tem mais agrotóxico do que laranja e os outros.

Vê-se a preocupação da turma com o erro, o que precisou ser explicitado que eles não tinham como saber aqueles resultados. Apesar da ausência de justificativas adequadas por parte da turma na atividade anterior, entende-se que os estudantes foram capazes de perceber suas hipóteses em relação aos dados. O estudante que afirmou terem errado percebeu que os dados não correspondem às escolhas da turma como um todo e a estudante que citou o morango (a amostra com maior concentração de agrotóxico) o fez porque o alimento foi um dos expressados livremente, mesmo com a baixa votação e no gráfico é o ponto máximo.

Importante esclarecer que a maior interferência, no sentido de busca de explicitações quando os estudantes foram solicitados a confrontarem hipóteses e dados reais, foi mais forte no primeiro encontro de intervenção. Nos demais encontros estimulou-se o processo de reanálise dos dados, uma vez que um vínculo entre pesquisadora e estudantes já estava se formando, fazendo com que os mesmos se sentissem mais à vontade para se expressarem diante de todos. Uma questão colocada após as dificuldades na justificativa da questão do confronto foi

verificada na Atividade 3.1: “*O que acontece com a população masculina e feminina com o passar dos anos?*”. O trecho de diálogo, a seguir, evidencia o que se discutiu. O relato corresponde ao verificado na Turma 7B.

Pesquisadora: Os dados reais se parecem com o que vocês pensaram antes? [Silêncio]. Vocês disseram que não tinha nada a ver o número de homens e mulheres com a idade não foi?

Aluno 8: Foi.

Pesquisadora: E agora?

Parte da turma: Está diferente.

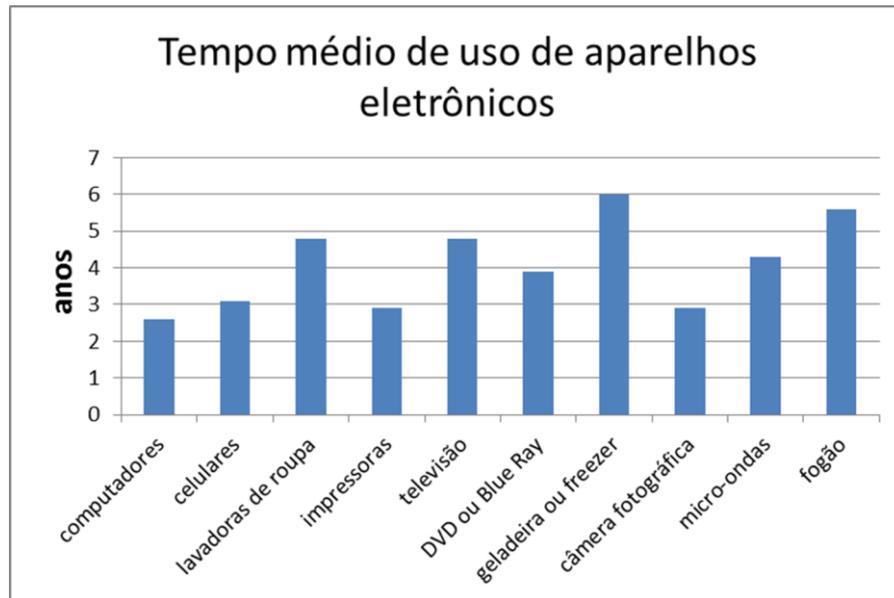
Pesquisadora: O que acontece com a população masculina e feminina com o passar dos anos?

Aluna 9: a expectativa de vida dos homens é menor...

Pesquisadora: Isso mesmo.

No relato anterior, está claro que por mais que os alunos percebam as diferenças entre suas respostas ao levantarem hipóteses e o que os dados estão mostrando, expressar explicações é uma habilidade que requer maiores investimentos, buscando-se outras questões que ajudem a turma na construção de argumentos coerentes com os dados reais.

As questões propostas para os estudantes, coletivamente, que visavam explorar a avaliação de conclusão também foram solicitadas em todas as atividades. Buscou-se ressaltar no diálogo a necessidade de se utilizar os dados dos gráficos como evidências. Quando os estudantes tomavam a iniciativa de responder às questões de conclusão, a pesquisadora questionava se alguém pensava diferente e pedia explicações apontando para o gráfico exposto no quadro (Gráfico 5.9). Um exemplo de mediação realizada pela pesquisadora no momento de avaliação de uma ou mais conclusões dadas, está descrito a seguir e aconteceu na Turma 5B. As questões surgiram da Atividade 2.1 e são: “*Podemos afirmar que o baixo tempo de uso de computadores é porque eles dão mais defeitos?*”; e “*Podemos afirmar também que os produtos com tempo mais longo de uso são aqueles usados na cozinha?*”.

Gráfico 5.9 – Atividade 2.1 – dados sobre o tempo médio de uso de aparelhos eletrônicos

Fonte: IDEC, 2013.

Pesquisadora: Podemos afirmar que o baixo tempo de uso de computadores é porque eles dão mais defeitos?

Parte da turma: Sim!

Pesquisadora: vamos ver o gráfico aqui. Está falando aqui de defeitos?

Parte da turma: Não!

Pesquisadora: As pessoas podem comprar outro produto porque é mais moderno, não é?

Aluna 10: Porque tem uma câmera melhor.

Pesquisadora: Isso! O computador poderia ser porque tem uma potência maior. Então aqui a gente não sabe se é por causa de defeito. A gente só sabe tempo de uso, tempo médio. Então tem conclusões que a gente não pode tirar olhando para os dados. Tá certo?

Parte da turma: Certo!

Pesquisadora: Podemos afirmar também que os produtos com tempo mais longo de uso são aqueles usados na cozinha?

Aluno 11: Sim, sim, sim, professora!

Pesquisadora: Por quê?

Aluno 11: Porque o da cozinha não tem como dar baque, tipo geladeira, fogão. Agora celular, computador, tabletes se a pessoa tiver distraída pode derrubar.

Pesquisadora: Ah, então geladeira não quebra, não é?

Aluno 11: Quebra! Mas pode ajeitar.

Pesquisadora: Então olhando para o gráfico podemos afirmar que os produtos com tempo mais longo são os da cozinha?

Aluna 12: Os dois maiores sim, professora.

Na Atividade 2.1 foi possível explorar com os estudantes duas questões de avaliação de conclusão: uma que não poderia ser tirada a partir dos dados e outra questão que poderia, desde que o indivíduo ressaltasse em sua resposta que na cozinha estão localizados comumente apenas os dois produtos com maior durabilidade. No trecho evidenciado anteriormente o estudante expõe para a turma essa compreensão, após os encaminhamentos dados pela mediadora, que busca sempre direcionar o olhar dos estudantes para os dados do gráfico.

Todas as atividades coletivas trabalharam questões que requeriam novas hipóteses para o que os dados mostraram, ou seja, solicitavam reflexões sobre o que poderia ter contribuído para os resultados informados nos gráficos. Apesar desse tipo de questão não tenha sido explorado nos testes propostos aos estudantes, até pelo efeito do cansaço na realização dos mesmos, acredita-se que o processo dialógico das aulas favorecia o levantamento de novas hipóteses, o que é característico de um ciclo investigativo. Além disso, nas novas hipóteses estimulou-se o uso de uma linguagem voltada para a incerteza, uma vez que se tratavam de suposições. Na primeira atividade desenvolvida com os estudantes (Atividade 1.1), há um exemplo desse tipo de questão: “*Por que será que na amostra de morango houve tanto agrotóxico comparada com a de arroz?*”. O trecho de diálogo, a seguir, evidencia a mediação realizada nesse momento com a Turma 5A.

Pesquisadora: Que hipóteses vocês têm para o que aconteceu na pesquisa? Explicando melhor, o que eu quero saber é por que vocês imaginam que na amostra de morango houve tanto agrotóxico comparada com a de arroz.

Aluno 13: Eu acho que é porque ele é muito doce.

Pesquisadora: Ah, então você acha que por isso o morango tem mais agrotóxico. Vocês concordam com o colega de vocês? [Silêncio]. Gente, não estamos vendo aqui o que é verdade e o que não é ou o que é certo e o que não é. Para dizermos algo assim teríamos que fazer outra pesquisa. Estamos apenas levantando hipóteses, o que a gente imagina que pode ter acontecido, tá bom?

Aluno 14: Porque o morango é maior que o arroz...

Pesquisadora: ah, então você pensa que pode ser porque ele tem mais espaço para pegar mais agrotóxico. Pode ser isso, gente?

Parte da turma: Pode!

Aluno 15: Ele pode atrair mais os insetos.

O diálogo descrito antes ressalta a necessidade de se realizar uma nova pesquisa para poder o motivo dos resultados alcançados. Desse modo, buscou-se estabelecer a incerteza inerente à situação apresentada. Outro ponto importante do diálogo que deve ser colocado em evidência é o uso de termos apropriados a incerteza, ou seja, uso de linguagem probabilística quando se supõe algo. No relato o termo utilizado pela pesquisadora foi o “pode ser”.

Nas predições também se explorou coletivamente o uso de linguagem probabilística. Entretanto, nem todas as atividades mostraram-se adequadas para a solicitação desse tipo de questão, o que só reforçou a necessidade de se explorar uma linguagem de incerteza nos momentos de levantamento de novas hipóteses, como as questões descritas anteriormente. A Atividade 3.1, utilizada no último encontro, por tratar de dados relacionados ao censo demográfico brasileiro, o qual parte de uma pesquisa que se repete a cada dez anos, possibilitou a seguinte questão: “*Quando essa pesquisa se repetir, no próximo censo, o que pode acontecer?*”. O trecho do diálogo realizado em sala, a partir da questão demonstra como os estudantes (Turma 7A) interagiram em tal momento.

Pesquisadora: Quando essa pesquisa se repetir, no próximo censo, o que pode acontecer? Vocês acham que vai ficar do mesmo jeito ou vai mudar esses números, o que pensam?

Aluna 16: Sei não.

Pesquisadora: Realmente não tem como saber, mas a gente pode prever. Podemos pensar: talvez isso aconteça. Percebam a linguagem, estamos tentando prever, não temos certeza. Então como podemos falar?

Aluna 17: Pode ser.

Pesquisadora: Isso. Tem outra forma além dessa? [Registro no quadro das palavras ditadas].

Aluno 16: Talvez.

Pesquisadora: Exato. Tem outras ainda? [Silêncio]. Provavelmente é uma palavra que também podemos usar, tá gente?

Percebe-se no relato anterior as dificuldades dos estudantes da turma em fazerem previsões. Nesse sentido, foi preciso a pesquisadora estabelecer o foco da elaboração da resposta não na previsão, já que a mesma era um meio para se chegar à habilidade investigada. A direção então foi para a linguagem utilizada quando se faz predições, que no caso por se tratar de incertezas, deve-se buscar linguagens em acordo. Assim, as palavras citadas pelos estudantes foram registradas no quadro, como um reforço visual do que se discutiu oralmente.

As atividades desenvolvidas de modo coletivo não apresentaram distinções comportamentais entre estudantes do 5º ano e do 7º ano. Verificou-se que os sujeitos que mais se prontificavam para responder às questões lançadas eram os mesmos na maioria das vezes, mesmo que outras opiniões fossem solicitadas e ainda que as atividades contassem com a atenção das turmas. Desse modo, as atividades realizadas em pequenos grupos possibilitaram que as mesmas habilidades do estudo da tese fossem vivenciadas em uma outra dinâmica, integrando melhor os estudantes com maiores dificuldades de se expressarem oralmente para a turma.

A seguir, análises a respeito dos momentos de interação dos estudantes em grupos de dois ou três participantes com a pesquisadora como mediadora são apresentadas e discutidas.

5.2.1.2 Atividades em pequenos grupos

As atividades realizadas em grupos menores foram apenas duas, uma no segundo encontro e outra no terceiro encontro. Foi solicitado que os estudantes que trabalharam juntos em uma atividade mantivessem a formação para a atividade do encontro posterior, o que foi atendido pela maior parte dos sujeitos.

Diferente das atividades coletivas, uma distinção clara entre os estudantes do 5º ano e do 7º nas atividades em pequenos grupos foi a constatação de questões em branco, ou seja, sem respostas, uma vez que houve a exigência de registro escrito de respostas. Isso aconteceu em ambas as turmas do 7º ano de intervenção,

o que não se verificou nas turmas de 5º ano. O que se observou reforça a ideia já discutida anteriormente, de que a maior escolaridade não contribuiu para que os estudantes do 7º ano avançassem em suas elaborações argumentativas. Isso pode estar relacionado com o contrato pedagógico estabelecido em tais turmas, pouco voltado para a produção textual. Além disso, outro fator que pode explicar o desinteresse de alguns estudantes do 7º ano pelas atividades de intervenção é o recurso utilizado por muitos professores de associar tarefas a serem cumpridas com as notas escolares. Os estudantes do 7º ano compreendiam que o trabalho realizado pela pesquisadora com eles não estava relacionado diretamente com as suas avaliações bimestrais.

Quanto às habilidades exploradas nos momentos de discussão em pequenos grupos, salienta-se que a habilidade de levantamento de hipótese foi exposta no quadro branco, com auxílio de *slides* em *data show*, quando os estudantes já estavam reunidos em duplas ou trios. Foi solicitado que conversassem sobre a questão e registrassem a hipótese que pensaram. A atividade 3.2, de relação entre dados bivariados requereu que a pesquisadora fosse de grupo em grupo ressaltando as duas variáveis da questão de levantamento de hipótese, pois houve solicitações dos estudantes. O diálogo, a seguir, evidencia um desses momentos com uma dupla (Turma 5A), após ter sido mostrada a questão de levantamento de hipótese no quadro: “*O que o número de filhos das mulheres tem a ver com o nível de instrução delas?*”.

Aluno 18: Professora, tô entendendo não.

Pesquisadora: Não? Deixa eu explicar. O número de filhos que a mulher tem, tem a ver com a instrução dela, ou seja, uma mulher que, por exemplo, estuda muito. Você acha que ela vai ter mais filhos ou menos filhos que aquelas que estudaram pouco? Ou você acha que uma coisa não tem nada a ver com a outra?

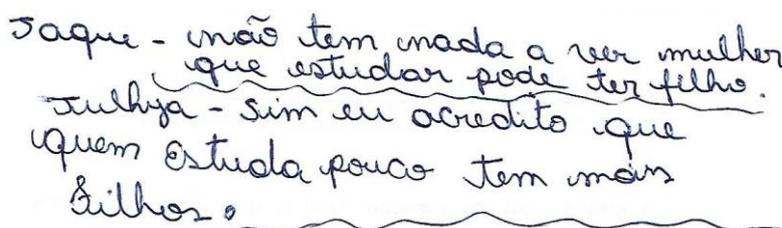
Aluno 18: Muito filho quem estuda pouco!

Pesquisadora: Então você acha que quem estuda pouco tem muito filho? Vocês agora precisam conversar para descobrirem se concordam. Aí explica também, certo?

Percebe-se no relato anterior que a questão de levantamento de hipótese na situação dada mostrou-se mais completa, sendo necessário maior interferência da mediadora com explicitações. Acredita-se que o termo instrução não é muito comum

para o público do Ensino Fundamental. Além disso, a palavra estava inserida em um contexto pouco explorado na sala de aula, que é a avaliação da relação entre duas variáveis. Entretanto, a assistência do adulto mediador tornou a situação mais próxima do grupo.

Na mesma atividade, a respeito da relação bivariada, pensar em uma hipótese na atividade proposta requeria uma negociação entre os participantes dos grupos. Apenas uma dupla decidiu registrar a resposta de cada integrante, uma vez que não concordaram e a escrita foi o modo encontrado de salientar o desacordo. A Figura 5.34, a seguir, demonstra o que se discutiu e trata-se da resposta da dupla a questão de dados bivariados: “O que o número de filhos das mulheres tem a ver com o nível de instrução delas?”. Logo depois são apresentados os dados referentes à questão (Gráfico 5.10).



Jaque - não tem nada a ver mulher
que estudar pode ter filho.

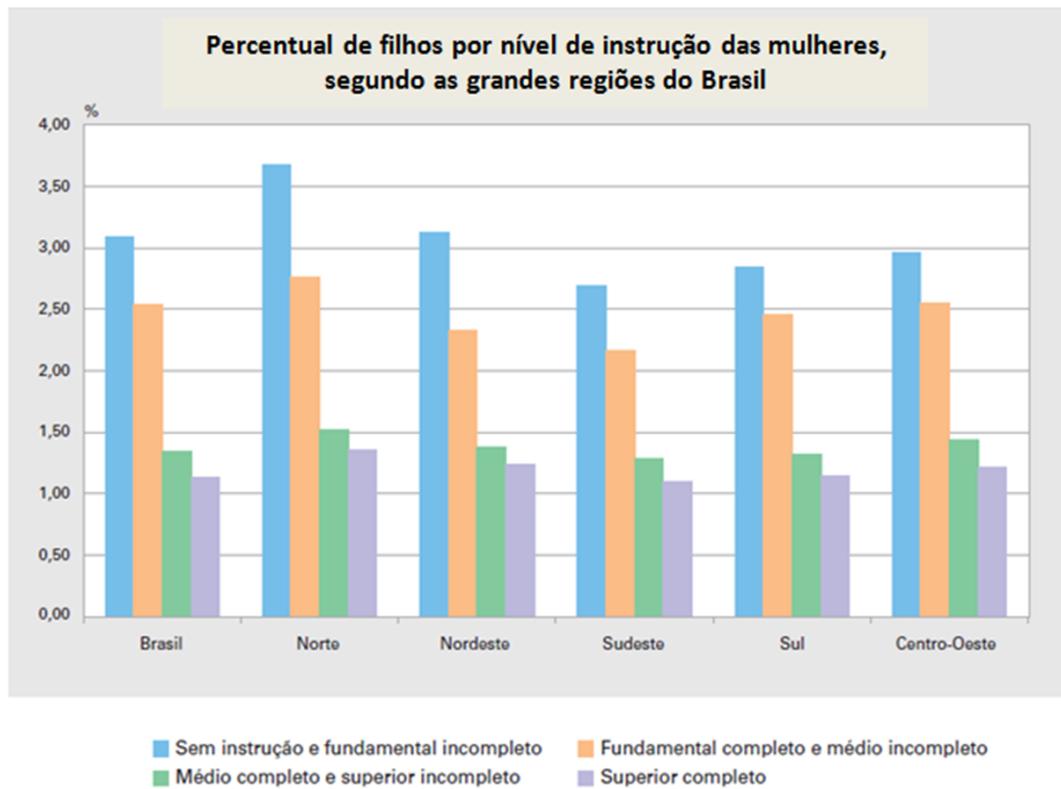
Julhya - sim eu acredito que
quem estuda pouco tem mais
filhos.

[Jaque – Não tem nada a ver, mulher que estuda pode ter filho].

[Julhya – Sim, eu acredito que quem estuda pouco tem mais filhos].

Figura 5.34 – Grupo (Turma 7A) que discorda no levantamento de hipótese, Atividade 3.2.

Gráfico 5.10 – Atividade 3.2 – dados sobre percentual de filhos das mulheres por nível de instrução nas regiões do Brasil



Fonte: IBGE, Censo Demográfico, 2010.

Os demais grupos (duplas ou trios) registraram a hipótese levantada como resposta única.

Um exemplo está na Figura 5.35, quando o grupo foi questionado: “*Que fontes (pessoas, sites, recursos e etc.) você imagina que o consumidor procura para saber a durabilidade de produtos?*”. A questão referiu-se a dados univariados da Atividade 2.2 (Gráfico 5.11), expostos após exemplo.

1º) R= Sites: porque os sites pode indicar mais sobre o produto

[1º) R= Sites, porque os sites podem indicar mais sobre o produto].

Figura 5.35 – Grupo (Turma 7B) que concorda no levantamento de hipótese, Atividade 2.2.

Gráfico 5.11 – Atividade 2.2 - dados sobre as fontes consultadas pelo consumidor ao pesquisar a durabilidade de produtos



Fonte: IDEC, 2013.

Foram entrevistados, por telefone, 806 homens e mulheres, de 18 a 69 anos, de diferentes classes sociais das seguintes cidades: Belo Horizonte (MG), Brasília (DF), Curitiba (PR), Goiânia (GO), Porto Alegre (RS), Recife (PE), Rio de Janeiro (RJ), Salvador (BA) e São Paulo (SP). O número de entrevistados em cada capital foi proporcional à população de cada capital. O levantamento foi feito entre agosto e outubro de 2013. A margem de erro é de 3,5% para mais ou para menos.

Quanto à interpretação de dados reais, não foi solicitado um registro, mas houve o incentivo por parte da pesquisadora em cada grupo, para que conversassem sobre o que os participantes percebiam dos dados apresentados no quadro. Isso antes de registrarem o confronto entre hipótese levantada e dados reais.

A questão de confronto em ambas as atividades foi: “Os *dados reais* se parecem com o que vocês pensaram antes? Por quê?”. A Figura 5.36 exemplifica os registros realizados por um grupo de estudantes diante da habilidade de confronto.

Que fontes (pessoas, sites, recursos e etc.) você imagina que o consumidor procura para saber a durabilidade de produtos?

os vendedores da loja. (os anos de garantia)

Os dados reais se parecem com o que vocês pensaram antes? Por quê?

Sim porque mais da metade das pessoas que responderam no gráfico responderam falando com o vendedor

[Ao vendedor da loja (anos de garantia)].

[Sim, porque mais da metade das pessoas que responderam no gráfico responderam falando com o vendedor]

Figura 5.36 – Grupo (Turma 7A) que realiza confronto entre hipótese e dados, Atividade 2.2.

Vê-se na Figura 5.36 que o grupo que trabalhou junto na atividade pensou em uma hipótese para a questão proposta e chegou a perceber que os dados reais apresentados coincidiam com o que imaginaram na Atividade 2.2, explorada no segundo encontro de intervenção.

A atividade 2.2, possibilitou ainda que outras discussões fossem abordadas, como a tomada de decisão a partir dos dados. A disposição dos estudantes em pequenos grupos viabilizou um debate entre os integrantes para que escolhessem duas das fontes do gráfico como consulta, caso fossem comprar um produto. Uma atividade coletiva não teria sido tão eficaz para a geração desse debate, pois como já discutido, nem todos os sujeitos se expressavam para a turma toda. Escolher duas das fontes mostradas no gráfico era importante para que os estudantes avaliassem melhor a confiabilidade ou não das fontes pesquisadas pelos entrevistados, que era a questão seguinte, de avaliação de conclusão: “*Pensando nos dados do gráfico, podemos dizer que as pessoas entrevistadas consultam fontes confiáveis? Por quê?*”. Assim, como exemplo de tomada de decisão por um grupo tem-se a Figura 5.37, a seguir.

Escolham duas dessas fontes do gráfico que vocês consultariam antes de comprar um produto. Explique.

Nós consultaríamos em sites de reclamação ou falando com conhecidos. Porque os sites de reclamação retiramos as dúvidas e as pessoas falam opiniões

[Nós consultaríamos em sites de reclamação ou falando com conhecidos. Porque os sites de reclamação nós tiramos as dúvidas e as pessoas falam opiniões]

Figura 5.37 – Grupo (Turma 5A) que faz tomada de decisão, Atividade 2.2.

Cabe discutir nesse momento que a discussão em torno da confiabilidade dos dados também está relacionada com o conhecimento dos estudantes sobre a importância da divulgação da fonte. Acredita-se que quando se comunica os resultados de informações estatísticas, ter a fonte expressa e chamar a atenção para a mesma em situações de ensino é uma atitude que corrobora para uma visão crítica dos dados e para o debate sobre confiabilidade dos mesmos. Assim, quando estudantes demonstraram não aceitar os dados representados, por ir de encontro a suas crenças, uma ação estabelecida pela pesquisadora foi apresentar informações adicionais da fonte responsável pelos dados. Isso para discussão acerca de interesses particulares ou não que poderiam estar direcionando uma dada pesquisa.

A Atividade 3.2, do último dia de intervenção, o qual tratou apenas de dados bivariados, também explorou a avaliação de conclusão com a questão: “*A partir do gráfico, podemos afirmar que no Sudeste a média de filhos das mulheres é menor que nos outros lugares do Brasil? Explique.*”. A seguir, evidências são dadas na Figura 5.38 que houve a preocupação em se retomar os dados na argumentação oferecida pelos estudantes.

A partir do gráfico, podemos afirmar que no sudeste a média de filhos das mulheres é menor que nos outros lugares do Brasil? Explique.

Sim, porque as barras no gráfico estão menores.

[Sim, porque as barras no gráfico estão menores]

Figura 5.38 – Grupo que avalia conclusão (Turma 7A), Atividade 3.2.

Diferente das questões propostas aos estudantes nos testes que eles realizaram, antes e após intervenções, estimulou-se bastante no processo de ensino a reinterpretção dos dados reais, ou seja, a reanálise dos dados para as questões posteriores ao confronto entre hipótese e dados reais. Isso foi importante para direcionar as justificativas dos estudantes com foco nos dados. Portanto, antes dos grupos elaborarem novas hipóteses, solicitou-se na Atividade 3.2: “*O que acontece com a quantidade média de filhos que as mulheres têm em relação à instrução delas?*”. A Figura 5.39 evidencia como um grupo respondeu ao solicitado.

O que acontece com a quantidade média de filhos que as mulheres têm em relação à instrução delas?

as mulheres mais instruída tem menos filhos e menos instruída tem mais

[As mulheres mais instruídas tem menos filhos e menos instruídas tem mais]

Figura 5.39 – Grupo (Turma 5B) que reanalisa os dados, Atividade 3.2.

Percebe-se na Figura 5.39 que o grupo de estudantes conseguiu realizar uma análise coerente dos dados bivariados apresentados. A partir de então, solicitou-se uma nova hipótese para os dados mostrados, como última questão proposta aos estudantes em ambas as atividades realizadas em pequenos grupos. A seguir, está explicitado um exemplo (Figura 5.40) de resposta para a pergunta: “*Que hipóteses vocês têm para o fato das mulheres mais instruídas terem menor quantidade de filhos que aquelas com pouca instrução?*”.

Que hipóteses vocês têm para o fato das mulheres mais instruídas terem menor quantidade de filhos que aquelas com pouca instrução?

Provavelmente as mulheres que estudam menos não tem consciência e a que estudam tem consciência.

[Provavelmente as mulheres que estudam menos não tem consciência e a que estudar tem consciência].

Figura 5.40 – Grupo (Turma 5A) que utiliza lança nova hipótese, Atividade 3.2.

Na Figura 5.40, vê-se que o estímulo ao uso de linguagem que demonstre incerteza também durante o levantamento de novas hipóteses foi eficaz, uma vez que o termo “provavelmente” foi explorado nos encontros de intervenção como linguagem probabilística e o mesmo não apareceu durante o pré-teste com os estudantes.

Após descrição dos resultados alcançados com as diferentes turmas que participaram do processo de intervenção pedagógica, a partir dos encaminhamentos didáticos realizados, cumpre analisa-los a luz das contribuições tanto do estudo de

Vygotsky, delineado enquanto aporte para o método interventivo, quanto a partir das inter-relações do mesmo com a fundamentação teórica e com os demais estudos anteriores à tese.

Assim, o tópico seguinte trata principalmente das contribuições da teoria de Vygotsky, com a Zona de Desenvolvimento Proximal (ZDP), à organização pensada para o processo didático do Estudo 2 e ao direcionamento interativo estabelecido entre pesquisadora, estudante e seus pares. Além das discussões com a ZDP de Vygotsky (1978), a seguir, há o resgate da teoria que fundamenta toda a tese, a qual trata-se do Letramento Estatístico proposto por Gal (2002) e entrelaçamentos com demais estudos.

5.2.2 Processo de intervenção pedagógica e Zona de Desenvolvimento Proximal (ZDP) de Vygotsky

Os encontros de intervenção com os estudantes do 5º ano e do 7º ano do Ensino Fundamental marcaram a Zona de Desenvolvimento Proximal (ZDP), tal qual proposta por Lev Vygotsky (1978). De acordo com o discutido na seção referente à fundamentação teórica (seção 2), a ZDP é exatamente a distância entre o que o sujeito pode realizar de modo independente, nível de desenvolvimento real, e o que ele é capaz de fazer com a assistência de um professor ou de seus pares.

Nesse sentido, as atividades propostas a partir dos testes (inicial e final) não poderiam definir a ZDP dos estudantes envolvidos no processo de intervenção. Os processos internos de desenvolvimento dos sujeitos, despertados pela aprendizagem organizada durante as aulas de intervenção, só são capazes de operar quando há interação do indivíduo com as pessoas de seu ambiente e em cooperação com os pares (VYGOTSKY, 1978, p. 90).

Saxe, Gearhart, Note e Paduano (1994) ressaltam que nas atividades de sala de aula e outras práticas cotidianas, as crianças se ocupam com problemas recorrentes e frequentemente há apoios sociais para a apropriação do conhecimento construído em um contexto para alcançar os objetivos emergentes em outro.

Para tanto, as atividades elaboradas e vivenciadas nos três encontros de intervenção contaram com momentos em que a pesquisadora desempenhou um papel mais ativo diante das habilidades com intensão didática, principalmente quando se constatou que os estudantes, enquanto grupo, não conseguiam

expressar uma argumentação ou opinião que atendesse ao solicitado. Em casos desse tipo, retomada dos dados foram feitas, com o encaminhamento de reanálises dos gráficos. Na argumentação, buscou-se sempre utilizar uma linguagem adequada ao que se requeria, pois em situações de predição o uso da incerteza é essencial, assim como no levantamento de hipótese e novas hipóteses, por se tratarem de suposições. Nas falas dos estudantes, chamar a atenção para o tipo de situação, se era uma suposição ou uma evidência dos dados, foi uma estratégia que visou a aprendizagem dos sujeitos.

Nos grupos menores, a cooperação entre duplas ou trios de estudantes foi o modo privilegiado de aquisição da aprendizagem por parte dos sujeitos. Nessas situações, a pesquisadora buscou muito mais incentivar o diálogo nos diferentes grupos do que responder a perguntas colocadas por eles. Foi interessante observar que nos grupos menores a expressão da discordância com os dados apresentados nos gráficos apareceu, o que não se constatou nas atividades coletivas, isso ocorreu pouco e com estudantes do 7º ano. Os dados que não foram aceitos como verdadeiros referiam-se à relação bivariada entre número de filhos e nível de instrução de mulheres. As atividades de correlação mostraram-se um campo fértil para a aceitabilidade ou não dos dados, como verificado também nas atividades de pré-teste e pós-teste. Nos encontros de intervenção, com os momentos de grupos menores, o diálogo aconteceu entre pares de modo a que crenças existentes viessem à tona.

Se a crítica diante da confiabilidade ou não dos dados veiculados pelos meios midiáticos é desejável, por outro lado, também é sabido que crenças do que esperam encontrar direcionam o olhar daquele que faz a leitura dos dados. Nickerson (1998) já alertou que ideias preconcebidas tendem a se sobressair quando se está perante informações, numa seleção não consciente dos dados apresentados. Assim, a pesquisadora atuou nos grupos também de modo a levá-los a observar a fonte responsável pela pesquisa. Se poderia haver algum interesse particular subjacente por parte daqueles que produziram os dados.

Compreende-se que a opinião e comunicação de impressões a respeito de informações estatísticas é uma habilidade necessária ao Letramento Estatístico da população para Iddo Gal (2002). Acredita-se que o estudante conseguir expressar oralmente ou de modo escrito seus argumentos e justificativas para decisões

tomadas é um passo importante para que o mesmo aprenda a distinguir uma postura crítica da defesa de crenças arraigadas.

As habilidades investigadas no presente estudo, de modo geral, mostraram-se na Zona de Desenvolvimento Proximal das turmas, uma vez que os estudantes nos encontros de intervenção conseguiram realizar as atividades solicitadas com a assistência e cooperação de todos os envolvidos no processo pedagógico.

O pós-teste marca um outro nível de desenvolvimento real dos estudantes das turmas pesquisadas, os quais aprenderam muitas das habilidades estudadas a partir de fases do ciclo investigativo, principalmente no que diz respeito a aprendizagem da argumentação no 5º ano. Porém, tratou-se de um momento em que os estudantes precisavam atuar de modo independente, não se caracterizando mais por ZDP.

Tem-se, portanto, que os estudantes que passaram pelo processo de intervenção pedagógica do Estudo 2 adquiriram novas aprendizagens, o que para Vygotsky (1978, p. 90) marca o início do processo de desenvolvimento desses sujeitos. Isso significa que a aprendizagem adequadamente organizada é o que viabiliza o desenvolvimento mental.

Na viabilização das aprendizagens dos estudantes, constatadas no pós-teste e nas respostas oferecidas pelos sujeitos durante o processo de intervenção, como descritas anteriormente, a Análise Exploratória de Dados (AED), introduzida por Tukey, foi bastante ressaltada. Para Tukey (1980), a visualização da representação dos dados é um aspecto que não deve ser desprezado na AED, assim como a extração dos dados de qualquer informação que possa levar a deduções. O autor ainda complementa que a AED é uma atitude, uma flexibilidade. Isso torna a AED apropriada para os estudantes do Ensino Fundamental, como revelou os resultados alcançados no presente estudo.

Os estudantes que participaram dos encontros de intervenção foram estimulados a analisarem e reanalisarem dados reais representados em gráficos de barras, tanto em situações de distribuição univariada quanto em situações bivariadas. Suposições a respeito dos dados, com o levantamento de novas hipóteses também foram exploradas, o que contribuiu para a elaboração de argumentações mais próximas dos dados evidenciados.

Nas diferentes turmas que passaram pelo processo de intervenção pedagógica percebeu-se aprendizagens distintas. Como proposto por English

(2014), as temáticas das atividades interferiram nas compreensões demonstradas pelos estudantes. Isso foi bastante evidente nas turmas de 7º ano. Contudo, pode-se afirmar que tanto a temática quanto o tipo de distribuição de dados influenciaram, pois averiguou-se que nas situações de correlação que iam de encontro às crenças dos estudantes, como na relação entre número de filhos de mulheres e nível de instrução as reações dos estudantes aos dados foram de negação, no sentido de se discordar dos dados. Importante esclarecer que não se pode concluir que a discordância prejudicou o desempenho dos estudantes do 7º ano, uma vez que o pós-teste revelou que em tais situações a maioria dos estudantes conseguiu oferecer argumentos apropriados, o que não se confirmou quando os dados foram univariados.

Vê-se que a organização das atividades, com os devidos encaminhamentos estava dentro da ZDP dos estudantes, já desde o 5º ano de escolaridade. Afinal, os estudantes do 5º ano conseguiram avanços significativos em todas as habilidades investigadas, isso sem a distinção entre os tipos de dados, exceto quando avaliados na predição com uso de linguagem probabilística, cuja dificuldade confirma resultados de pesquisas anteriores (HENRIQUES; OLIVEIRA, 2015).

Acredita-se que as relações estabelecidas no ensino da Matemática, nas quais normalmente ocorre o ensino de conceitos estatísticos na Educação Básica, influenciaram os estudantes na ausência de elaboração de linguagens apropriadas ao uso da incerteza. Constatou-se na análise das aulas que estudantes desde o 5º ano já são capazes de utilizar termos próprios das situações probabilísticas, como “provavelmente” e “chance”, após a devida exploração dos mesmos nas atividades propostas nos encontros de intervenção. Entretanto, as respostas às atividades normalmente utilizadas no ambiente escolarizado reivindicam certezas, principalmente no que diz respeito ao ensino da Matemática, enquanto disciplina. Desse modo, como já expôs Oliveira (2016), ao discutir a teoria Histórico-Cultural de Vygotsky, a cultura torna-se parte da natureza humana e molda o funcionamento psicológico do ser.

Entende-se, a partir de então, que repensar a prática educativa como organização da aprendizagem para o desenvolvimento já desde os primeiros anos de escolaridade, visando mudanças na cultura em relação ao saber estatístico, tende a favorecer a constituição de cidadãos. Fazer dos dados estatísticos um aliado na formação de adultos letrados está a cargo da escola.

Hedegaard (2002) expôs que o papel do professor é dirigir a ação de uma maneira apropriada ao nível atual de desenvolvimento do sujeito, ao contexto cultural e social e ao que se entende por fundamental em um dado assunto. Nesse sentido, durante os encontros de intervenção discussões orais, escritas, com assistência da pesquisadora e colaboração entre pares foram abordadas, o que se espera de um trabalho dentro da Zona de Desenvolvimento Proximal e o que possibilitou que muitos estudantes chegassem a um nível de desenvolvimento real no pós-teste, diferente daquele encontrado antes da intervenção. A maior interação nas aulas dadas e disponibilidade para a exposição de ideias, principalmente na escrita, entre os estudantes do 5º ano parece ter influenciado positivamente as turmas de tal ano.

A partir de todo o exposto, pode-se destacar as atitudes tomadas durante o processo de intervenção didática que se acredita viabilizaram as aprendizagens explicitadas nas análises anteriores. Assim, seguem as atitudes implementadas na elaboração das aulas ministradas e/ou durante as mesmas.

- Uso de variabilidade de contextos;
- Uso de dados estatísticos reais apresentados em gráficos de barras;
- Interdisciplinaridade intrínseca aos temas explorados;
- Levantamento de conhecimentos prévios sobre as temáticas antes da formulação das hipóteses;
- Abordagem de conceitos estatísticos específicos antes, durante ou depois do conhecimento dos dados;
- Encorajamento ao uso de vocabulário estatístico;
- Discussão com elaboração de justificativas voltadas para os dados observados;
- Confronto entre crenças estabelecidas e dados apresentados;
- Incentivo à percepção da fonte divulgadora dos dados;
- Incentivo à inter-relação entre as habilidades exploradas no estudo a partir da reanálise de dados;
- Inserção de outras habilidades como tomada de decisão e novas hipóteses (pós dados);
- Identificação das crenças como momentos de incerteza, o que exige linguagem probabilística.

As atitudes pontuadas anteriores foram aquelas que se fizeram presentes sempre que o momento didático requereu, seja quando houve a apresentação do termo amostra para os estudantes, por exemplo, cujo entendimento conceitual é o que permitiria uma resposta adequada para um questionamento proposto, seja quando diante de uma discussão oral no grande grupo intervenções como tratamento do erro fizeram-se necessárias.

Vê-se então que o trabalho de análise exploratória de dados estatísticos reais nos anos iniciais do Ensino Fundamental foi bastante positivo, podendo a formação da população para lidar com informações estatísticas alcançar o Letramento Estatístico desejado, a partir das aprendizagens das habilidades investigadas no presente estudo. Isso porque poucos encontros resultaram em grandes avanços, o que se pode esperar que um trabalho instrucional contínuo e inserido na cultura escolar possa fazer mais. Isso no sentido de formar adultos conhecedores da sua condição no mundo e torná-los capazes de atitudes e decisões que promovam ações de inclusão e de justiça social.

6 CONCLUSÕES

A pesquisa aqui apresentada configurou-se a partir da problemática de investigação que surgiu da seguinte questão: De que modo o levantamento de hipóteses, análise de dados e habilidades relacionadas a conclusões por estudantes do Ensino Fundamental corroboram para a aprendizagem de justificativas baseadas em evidências?

Levantamento de hipótese, análise de dados e conclusões a partir dos dados constituem-se fases do ciclo investigativo (GUIMARÃES; GITIRANA, 2013; SILVA, 2013). A exploração do ciclo investigativo da pesquisa é essencial à Educação Estatística (BEN-ZVI, 2016; CAMPOS; WODEWOTZKI; JACOBINI, 2013; GUIMARÃES; GITIRANA, 2013). Ao se pensar nas informações de carácter estatístico, que chegam até o leitor a partir de diferentes veículos midiáticos, conhecer e refletir sobre os dados reais divulgados é um exercício de cidadania e requer uma postura crítica, principalmente porque o leitor, no caso, é um consumidor. Gal (2002, p. 2-3) defende que interpretar, avaliar criticamente informações estatísticas, assim como discutir, comunicar reações diante das mesmas informações são ações necessárias ao Letramento Estatístico. Desse modo, o modelo de letramento proposto pelo autor foi o fundamento para a tese. Por se tratar de uma pesquisa voltada para o Ensino Fundamental, a Análise Exploratória dos Dados (AED), foi privilegiada, uma vez que a mesma requer atitude e flexibilidade perante os dados, com a intensa reanálise dos dados, a fim de melhor compreendê-los (TUKEY, 1980).

Assim, estruturou-se o objetivo geral da tese, que foi analisar aprendizagens apresentadas por estudantes do Ensino Fundamental relacionadas ao levantamento de hipóteses, à análise de dados e às conclusões a partir dos dados, enquanto fases do ciclo investigativo. Como objetivos específicos, elencou-se:

- Analisar os conhecimentos demonstrados por estudantes do Ensino Fundamental (5^o e 7^o anos), diante de distribuições univariadas e bivariadas, quando levantam hipóteses, interpretam dados reais, confrontam hipóteses e dados, avaliam conclusões e fazem predições;
- Analisar as aprendizagens de estudantes do Ensino Fundamental (5^o e 7^o anos) em estudo experimental de intervenção, diante de distribuições

univariadas e bivariadas, quando levantam hipóteses, interpretam dados reais, confrontam hipóteses e dados, avaliam conclusões e fazem previsões.

Para a viabilização dos objetivos específicos postos, dois estudos foram realizados. O primeiro (Estudo 1) foi pensado de modo a atender o objetivo específico correspondente aos conhecimentos de estudantes dos anos de interesse (5º e 7º anos). O segundo estudo (Estudo 2) atendeu as expectativas do objetivo relacionado às aprendizagens dos estudantes de tais anos. Todos os estudos exploraram as habilidades de:

- Levantamento de hipótese: relaciona-se com a fase de levantamento de dados, como o próprio nome diz, salienta-se que a conjectura requer a exposição de uma questão de pesquisa;
- Interpretação de dados: relaciona-se com a fase de análise dos dados reais, os quais estão sendo apresentados em gráficos;
- Confronto entre hipótese e dados reais: requer a inter-relação entre fases distintas, a hipótese lançada e a análise dos dados reais são confrontadas;
- Avaliação de conclusão dada: Trata-se de um aspecto da fase de conclusão, na qual os dados já foram analisados e precisa-se tomar decisões a partir dos mesmos;
- Linguagem que demonstre incerteza ao se realizar previsões: relaciona-se com a fase de conclusão que é a realização de previsões, o que exige uso de linguagem adequada.

No Estudo 1, o número de 22 estudantes de uma turma do 5º ano e 25 estudantes de uma turma do 7º ano do Ensino Fundamental de escolas públicas da região Metropolitana de Recife responderam a um teste contendo 4 (quatro) atividades. As duas primeiras atividades envolveram distribuições univariadas e exploraram todas as habilidades citadas anteriormente. Distinguiram-se pelo contexto dos dados reais e tipo de representação, se gráfico (Atividade 1) ou tabela (Atividade 2). As duas outras atividades envolveram distribuições bivariadas (correlação). O contexto foi diferente também. Outra mudança foi que a Atividade 3 utilizou dados em uma relação direta no gráfico e a Atividade 4 utilizou dados em uma relação indireta em um quadro.

Verificou-se que os estudantes desde o 5º ano já são capazes de levantar hipóteses para uma questão de pesquisa proposta. Quanto a interpretação dos dados ao analisar uma representação com informações estatísticas, percebeu-se que o tipo de distribuição foi determinante para o desempenho de ambos os grupos de estudantes. Tanto o 5º ano, quanto o 7º sentiram grandes dificuldades nas situações de dados bivariados, principalmente na Atividade 4, de relação indireta entre duas variáveis. Na habilidade de confronto entre hipótese e dados o desempenho dos estudantes também foi baixo quando se investigou a distribuição de dados bivariados. Entretanto, foi interessante observar que ao confrontarem dados em relação direta, os estudantes apresentaram maiores dificuldades em perceberem as evidências do gráfico. Acredita-se que o contexto (câncer de pulmão e hábito de fumar produtos com nicotina) bastante explorado nas mídias fez com que os dados fossem desprezados. Constatou-se que eles buscavam apoio em crenças e não nos dados para a resolução de questões e exposições de justificativas após conhecimentos dos dados reais, o que vai ao encontro do que a literatura vem apontando (NUNES; BRYANT, 2011; NICKERSON, 1998). As habilidades de avaliação de conclusão e linguagem de incerteza em predições mostraram-se de difícil compreensão para os dois anos de escolaridade pesquisados, em todas as atividades que exploraram tais habilidades.

A partir dos resultados do Estudo 1, o Estudo 2 foi elaborado. De carácter intervencionista, no Estudo 2 utilizou-se o método experimental, com criação de pré-teste, atividades para uma aprendizagem organizada e pós-teste. Análises realizadas no primeiro estudo foram essenciais para as adaptações nos testes inicial e final do estudo posterior, assim como para a abordagem pensada no encaminhamento das aulas nos encontros de intervenção. Nos testes optou-se pela exploração de dados em gráficos de barras apenas e pela relação direta quando em dados bivariados (correlação). O Estudo 1 também revelou que alguns estudantes realizavam reanálises dos dados propostos após questões de interpretação, atitude que foi incentivada nas intervenções.

O Estudo 2 aconteceu com 6 (seis) turmas, 3 (três) do 5º ano e 3 (três) do 7º ano, em escolas da Rede Municipal de Recife que estavam inseridas em um contexto social próximo. Assim, o número de 121 sujeitos participou do pré-teste, 53 deles pertenciam ao 5º ano e 68 pertenciam ao 7º ano. Em ambos os anos uma das turmas constituiu-se o grupo controle do estudo, as quais não passaram por

qualquer intervenção pedagógica com a pesquisadora, apenas realizaram pré-teste e pós-teste. As demais turmas passaram pelo mesmo processo de intervenção. Importante retomar que nem todos os estudantes presentes no pré-teste participaram do pós-teste, havendo uma redução no número de estudantes, para 114 sujeitos, o que foi levado em consideração na comparação entre os testes inicial e final.

Constatou-se na análise de variância (Anova) do pré-teste que não houve diferença significativa entre as turmas, o que foi um dado positivo para a continuidade do estudo. Foram elaborados dois tipos de teste para ser aplicado em todas as turmas e em ambos os momentos (avaliação inicial e final), ou seja, metade da turma fez o Teste A e a outra metade o Teste B, os quais foram similares. Isso significa que as mesmas habilidades foram exploradas em duas atividades: Atividade 1, distribuição univariada; Atividade 2, distribuição bivariada (correlação). A análise de variância também revelou ausência de diferença significativa entre os tipos de testes, o que é necessário para efeitos e comparação.

Como as turmas de intervenção passaram pelo mesmo processo de ensino e as mesmas não se mostraram estatisticamente diferentes no pré-teste, houve uma junção das duas turmas do 5º ano e das duas turmas do 7º ano nas análises comparativas entre pré-teste e pós-teste.

Todos os estudantes foram solicitados a responderem a 9 (nove) questões em cada teste. Então a média de acerto do número de questões foi o parâmetro. Percebeu-se que tanto o 5º ano, quanto o 7º ano avançou de modo significativo, no que diz respeito às turmas de intervenção. Como esperado, as turmas do grupo controle não se mostraram estatisticamente diferentes entre pré-teste e pós-teste.

A análise das justificativas utilizadas pelos estudantes nas questões de ambos os testes (inicial e final) revelou que o 5º ano contou com melhoras expressivas do ponto de vista estatístico nos registros escritos dados ao interpretar dados univariados, ao confrontar hipóteses e dados reais nas distribuições univariadas e bivariadas (correlação), ao avaliar conclusões a partir dos dados univariados e também ao levantar hipóteses em se tratando de dados bivariados. Assim, entre os estudantes do 5º ano vê-se que as habilidades contaram com aprendizagem, isso em um tipo de distribuição ou em outro. A exceção foi a habilidade de uso de linguagem probabilística ao prever, que só foi abordada em dados univariados e não apresentou avanços significativos.

Já o 7º ano contou com avanço significativo nas justificativas apenas na habilidade de interpretar quando os dados foram bivariados. O que explica o desempenho dos estudantes do 7º ano é que tais sujeitos partiram de percentuais maiores do que os estudantes do 5º ano no pré-teste, exceto na habilidade de utilizar linguagem probabilística ao predizer. No uso de linguagem probabilística os estudantes do 7º ano obtiveram percentuais inferiores, ou seja, no pré-teste os estudantes do 7º apresentaram maiores dificuldades em demonstrar incerteza nas situações propostas. Além disso, a produção escrita, essencial no teste, contou com maior resistência entre os estudantes do 7º ano, o que interferiu na expressão argumentativa desses sujeitos. Por outro lado, constatou-se que as habilidades que partiram de dados bivariados (correlação) mobilizaram mais aos estudantes do 7º ano do que quando as mesmas habilidades partiram de dados univariados.

Quanto às aulas ministradas aos estudantes de ambos os anos, as mesmas aconteceram em 3 (três) encontros, de aproximadamente duas horas cada. Em cada encontro explorou-se duas atividades: no primeiro dia uma atividade com dados univariados e outra com dados bivariados, ambas encaminhadas coletivamente; no segundo dia abordou-se apenas dados univariados, de modo coletivo e em pequenos grupos; e por fim, no terceiro dia, explorou-se dados bivariados coletivamente e em pequenos grupos. A partir dos encaminhamentos dados, optou-se por analisar as aulas observando-se a mediação da pesquisadora com os estudantes e a assistência realizada nos grupos de dois ou três integrantes. Os estudos de Vygotsky (1978) a respeito da Zona de Desenvolvimento Proximal (ZDP) foram essenciais para a organização planejada das atividades de intervenção e para a posterior análise das mesmas. Assim, como totalizaram-se 6 (seis) atividades trabalhadas nos encontros, 4 (quatro) foram abordadas de modo coletivo e 2 (duas) foram realizadas em pequenos grupos.

As atividades de carácter coletivo priorizaram a oralidade, com exposição de argumentos, opiniões e compreensões a partir de questões-chave, pensadas de modo que as habilidades investigadas na tese emergissem. Nas atividades coletivas a pesquisadora desempenhou um papel mais atuante junto a cada turma, no sentido de esclarecer conceitos, promover reanálises dos dados para resolução de questões, incentivar o uso de linguagem probabilística diante de situações de incerteza e buscar nas falas dos estudantes, que se propuseram a expor ideias, o que era evidência nos dados reais e o que era suposição. A discussão da ZDP de

Vygotsky (1978) esclarece que a aprendizagem com adultos e em pares desperta processos internos de desenvolvimento, assim a cooperação no ambiente de ensino é essencial.

Como outro recurso de cooperação, buscou-se a realização de atividades também em pequenos grupos (duplas ou trios), nas quais a interação entre os integrantes foi priorizada, com discussão e negociação para o registro escrito de ideias. A pesquisadora em tais momentos atuou de modo a incentivar o diálogo entre pares e a produção escrita de respostas a questões dadas. As atividades de dados bivariados mostraram-se um campo fértil para discussão de crenças e confiabilidade a respeito dos dados reais, principalmente nas turmas de 7º anos. Os 7º anos, ainda, mostraram-se mais resistentes, mesmo trabalhando em grupos, a registrarem ideias de modo escrito, o que interfere na expressão argumentativa dos sujeitos, como verificado no pós-teste, fase que já não caracteriza a Zona de Desenvolvimento Proximal dos estudantes, a qual acontece em cooperação. Isso indica que a disponibilidade para discutir ideias com outros e registro das mesmas nos encontros de intervenção ajudaram os estudantes, especialmente do 5º ano, a chegarem ao nível de desenvolvimento real no pós-teste, no que diz respeito ao uso de justificativas adequadas diante de habilidades investigadas na tese.

Diante do exposto, é importante ressaltar o papel das crenças no presente estudo. As crenças são necessárias e permeiam todo o processo de investigação. O que se verifica é que o uso de informações estatísticas fidedignas e relevantes para o público ao qual se destinam no ensino, além dos encaminhamentos didáticos necessários, possibilitaram confronto e mudanças de crenças entre pré-teste e pós-teste. Isso se confirma pelo uso de justificativas adequadas pelos estudantes que participaram do processo de intervenção pedagógica.

Portanto, configura-se a tese desenvolvida no doutorado, a qual estabelece que o ensino voltado para o desenvolvimento da capacidade de formular hipóteses e analisar dados estatísticos em diferentes distribuições (uni e bivariadas) favoreceu a tomada de decisões baseadas em evidências, pois permitiu pensar fases da pesquisa estatística de modo inter-relacionado. Isso porque as justificativas oferecidas pelos estudantes diante de habilidades investigadas na tese contaram com aprendizagens importantes, que foram viabilizadas principalmente pela organização estratégica planejada nas aulas de exposição de ideias, retomada dos

dados reais em diferentes momentos e discussão das crenças arraigadas em contraposição ou confluência com os dados.

Por fim, acredita-se que as habilidades investigadas no presente estudo devem ser exploradas no ambiente escolarizado já desde os primeiros anos e de modo inter-relacionado. O desempenho dos estudantes do 5º ano é um indicador importante de que a argumentação com foco na evidência dos dados reais contou com um campo fértil em tal ano. Se em apenas três encontros, de duas horas cada, os estudantes foram capazes de avançarem de modo significativo em suas compreensões de diferentes habilidades e diferentes distribuições de dados, pode-se esperar que o ensino sistemático das habilidades relacionadas a fases do ciclo investigativo da pesquisa alavanque o desenvolvimento do Letramento Estatístico da população. Letramento imprescindível à formação de cidadãos críticos diante da enormidade de informações estatísticas ou assim veiculadas pelas mídias, as quais nem sempre têm a intenção de informar ou esclarecer, o que requer conhecimentos cognitivos e disposicionais (GAL, 2002) dos leitores perante dados largamente divulgados.

REFERÊNCIAS

ALLMOND, S.; MAKAR, K. From hat plots to box plots in tinkerplots: supporting students to write conclusions which account for variability in data. **Proceedings of the Ninth International Conference on Teaching Statistics** (ICOTS9, July, 2014), Flagstaff, Arizona, USA. Voorburg, The Netherlands: International Statistical Institute. 2014.

BATANERO, C.; ESTEPA; GODINO. 1991 BATANERO, C.; ESTEPA, A.; GODINO, J. D. Análisis Exploratorio de Datos: sus posibilidades em la enseñanza secundaria. **Suma**, n. 9, p. 25-31, 1991. Disponível em: <http://170.210.81.106/faea/academica/departamentos/estadistica/catedras/Matematica/unidad_2_analisis_exploratorio_SUMA_91.pdf> Acesso: 16 de dez. de 2009.

BEN-ZVI, Dani. El desafío de desarrollar el razonamiento estadístico de los estudiantes (DSSR). Conferência proferida em **XX Jornadas Nacionales de Educación Matemática**, Campus Curauma-Facultad de Ciencias, PUCV, dez., 2016. Disponível em: <http://s3.amazonaws.com/static.ima.ucv.cl/> Acesso em 30 de maio de 2017.

BRASIL. Ministério da Educação. **Base Nacional Comum Curricular**. Brasília: MEC, 2017. Terceira versão da Educação Infantil e Ensino Fundamental. Disponível em: <http://basenacionalcomum.mec.gov.br/images/BNCC_publicacao.pdf> Acesso em 09 de maio de 2017.

BRESLOW, Norman Edward. Discussion: Statistical Thinking in Practice. In: WILD, C. J.; PFANNKUCH, M. Statistical Thinking in Empirical Enquiry. **International Statistical Review**, 67(3), 223-265, 1999. Disponível em: <<http://www.stat.auckland.ac.nz/~iase/publications/isr/99.Wild.Pfannkuch.pdf>>. Acesso em: 09 de maio de 2017.

CAMPOS, C. R; WODEWOTZKI, M. L. L.; JACOBINI, O. R. **Educação Estatística: teoria e prática em ambientes de modelagem matemática**. 2 ed. Belo Horizonte: Autêntica, 2013. (Coleção Tendências em Educação Matemática).

CARVALHO, Liliane M. T. L de. **O papel dos artefatos na construção de significados matemáticos por estudantes do Ensino Fundamental II**. Tese (Doutorado). Universidade Federal do Ceará, Faculdade de Educação, Programa de Pós-Graduação em Educação Brasileira. Fortaleza, 2008.

CASEIRO, A.; PONTE, J. P.; MONTEIRO, C. Investigações no ensino de conceitos e representações estatísticas no 1º ciclo. **Encontro de Investigação em Educação Matemática**, p.239-253, 2014.

CAVALCANTI, Erica M. S. **Para variar: Compreensões de estudantes dos anos iniciais diante de aspectos da variabilidade**. Dissertação (Mestrado) - Universidade Federal de Pernambuco, CE, Programa de Pós-Graduação em Educação Matemática e Tecnológica, 2011.

CAVALCANTI, E. M. S.; GUIMARÃES, G. L. Statistics in primary and middle school: hypothesis formulation, data analysis and conclusions. Proceedings of the 42nd Conference of the International Group for the Psychology of Mathematics Education. Umeå, Suécia, 2018.

CERVO, A. L.; BERVIAN, P. A. **Metodologia científica**. 5 ed. São Paulo: Pearson Prentice Hall, 2002.

COUTINHO, C.; ALMOULOU, S.; SILVA, M. J. O desenvolvimento do letramento estatístico a partir do uso do Geogebra: um estudo com professores de matemática. **Revista Eletrônica de Educação Matemática Revemat**. Florianópolis, v. 07, n. 2, 2012. p. 246-265.

CURI, E.; NASCIMENTO, J. de C. O ensino de gráficos e tabelas nos anos iniciais do ensino fundamental: Resultados de pesquisa nas várias instâncias curriculares **Em Teia Revista de Educação Matemática e Tecnológica Iberoamericana**, v.7, n.1, 2016. Disponível em: <https://periodicos.ufpe.br/revistas/emteia/article/view/3886/pdf> Acesso em: 21 de maio de 2017.

DANIELS, John E. Studies have shown that ...' how to read a statistical study. **Teaching Statistics Trust**, v. 39, n.3, 2017 p. 75-78. Disponível: <https://onlinelibrary.wiley.com/doi/epdf/10.1111/test.12130> Acesso em: 16 de fev. de 2019.

DÍAZ-LEVICOY, D.; BATANERO, C.; ARTEAGA, P.; LÓPEZ-MARTÍN, M. Análisis de los gráficos estadísticos presentados en libros de texto de educación primaria chilena. **Educ. Matem. Pesq.**, São Paulo, v.17, n.4, 2015. p.715-739.

ENGLISH, Lyn D. Establishing statistical foundations early: data modeling with young learners. **Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9)**, Flagstaff, Arizona, USA. Voorburg, The Netherlands: International Statistical Institute, 2014.

EVANGELISTA, B.; GUIMARÃES, G. L. Tables in textbooks for elementary school grades. **II International Conference on Mathematics Textbook Research and development -ICMT 2**. Rio de Janeiro, maio de 2017.

FERNANDES, L. A.; GOMES, J. M. M. Relatórios de pesquisa nas ciências sociais: características e modalidades de investigação. **ConTexto**, Porto Alegre, v. 3, n. 4, 1º semestre, 2003.

FITZALLEN, N.; WATSON, J. Extending the curriculum with tinkerplots: opportunities for early development of informal inference. **Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9)**, Flagstaff, Arizona, USA. Voorburg, The Netherlands: International Statistical Institute, 2014.

FRANKLIN, C.; KADER, G.; MEWBORN, D.; MORENO, J.; PECK, R.; PERRY, K.; SCHEAFFER, R. **Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report: A Pre-K–12 Curriculum Framework**. Alexandria VA: American Statistical Association, 2007.

FREIRE, Paulo R. N. Educação e mudança. 15 ed. Rio de Janeiro: Paz e terra, 1979.

FRIEL, S.; CURCIO, F.; BRIGHT, G. Making sense of graphs: Critical factors influencing comprehension and instructional implications. **Journal for Research in Mathematics Education**, v. 32, n. 2, 2001. p.124-158. Disponível em: <http://www07.homepage.villanova.edu/michael.posner/Friel-Curcio-Bright-2001-JRME.pdf> Acesso em: 18 de maio de 2016.

GAL, Iddo. Adults Statistical Literacy: meanings, components, responsibilities. **International Statistical Review**, v.70, n.1, p. 1-25, 2002.

GARFIELD, J.; BEN-ZVI, D. Learning to reason about data. In: GARFIELD, J.; BEN-ZVI, D. (Org). **Developing Students' Statistical Reasoning: Connecting Research and Teaching Practice**. Springer, 2008. p. 1-31.

GITIRANA, Verônica. A pesquisa como eixo estruturador da Educação Estatística. In: BRASIL. **Pacto Nacional pela Alfabetização na Idade Certa**. Caderno 7 (Educação Estatística). Brasília: Ministério da Educação, 2014, p.7-20.

GUIMARÃES, G. L.; GITIRANA, V. Estatística no Ensino Fundamental: a pesquisa como eixo estruturador. In: BORBA, R. E.; MONTEIRO, C. E. (Org). **Processos de ensino e aprendizagem em Educação Matemática**. UFPE, 2013, p. 93-132.

GUIMARÃES, G. L.; OLIVEIRA, I. Construção e interpretação de gráficos e tabelas. In: BRASIL. **Pacto Nacional pela Alfabetização na Idade Certa**. Caderno 7 (Educação Estatística). Brasília: Ministério da Educação, 2014, p.21-38.

HEDEGAARD, Mariane. A Zona de Desenvolvimento Proximal como base para o ensino. In: DANIELS, Harry (Org). **Uma introdução à Vygotsky**. (Tradução Marcos Bagno). São Paulo: Loyola, 2002.

HENRIQUES, Ana; OLIVEIRA, Hélia. Students' informal inference when exploring a statistical investigation. **Ninth Congress of European Research in Mathematics Education**. Praga, República Tcheca, 2015.

JACOBBE, T.; FOTI, S.; WHITAKER, D. Middle school (ages 10-13) students' understanding of statistics. **Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9)**, Flagstaff, Arizona, USA. Voorburg, The Netherlands: International Statistical Institute. 2014.

KONOLD, Clifford. Handling complexity in the design of educational software tools. **Proceedings of the Seventh International Conference on Teaching Statistics (ICOTS7)**, Salvador, Bahia, Brasil, 2006. Disponível em: https://iase-web.org/documents/papers/icots7/7F2_KONO.pdf Acesso em: 21 de maio de 2017.

LAKATOS, E. M.; MARCONI, M. **Metodologia do trabalho científico: procedimentos básicos, pesquisa bibliográfica, projeto e relatório, publicações e trabalhos científicos**. 5 ed. São Paulo: Atlas, 1996.

MAKAR, Katie. Predict! Teaching statistics using informal statistical inference. **Australian Mathematics Teacher**, v.69, Adelaide, SA, Austrália. 2013.

MCPHEE, D.; MAKAR, K. Exposing young children to activities that develop emergent inferential practices in statistics. **Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9)**, Flagstaff, Arizona, USA. Voorburg, The Netherlands: International Statistical Institute. 2014.

MENDONÇA, C.; GITIRANA, V.; LIRA, A. K. Abordagem da estatística nos anos finais do ensino fundamental: uma análise comparativa de livros didáticos de matemática e ciências. In: **XI Encontro Nacional de Educação Matemática**, 2013, Curitiba. Anais do XI ENEM (Educação matemática: retrospectivas e perspectivas). Curitiba, 2013. p. 1-15.

MONTEIRO, Carlos Eduardo F. Letramento estatístico: conceituações e implicações para a educação estatística. **IX Encontro Paraibano de Educação Matemática**. 2016.

NICKERSON, Raymond S. Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. **Review of General Psychology**, 2, 1998. p. 175-220. Disponível em: <http://pages.ucsd.edu/~mckenzie/nickersonConfirmationBias.pdf> / Acesso em: 09 de maio de 2017.

NUNES, Terezinha. **A pesquisa sobre o ensino e a aprendizagem da matemática inicial**. Escola de Altos Estudos CAPES-UFRGS, 2015. (Curso ministrado).

NUNES, T.; BRYANT, P. 2011. Understanding risk and uncertainty: The importance of correlations. **EM TEIA, Revista de Educação Matemática e Tecnológica Iberoamericana**, v. 2, n. 2, 2011.

OLIVEIRA, Marta Kohl de. Vygotsky e o processo de formação de conceitos. In: LA TAILLE, Y.; OLIVEIRA, M. K.; DANTAS, H. **Piaget, Vygotsky e Wallon: teorias psicogenéticas em discussão**. 27 ed. São Paulo: Summus, 2016.

PRESTES, Z.; TUNES, E.; NASCIMENTO, R. Lev Semionovitch Vigotski: Um estudo da vida e da obra do criador da Psicologia Histórico-Cultural. In: LONGAREZI, A. M.; PUENTES, R. V. (org.). **Ensino desenvolvimental: vida, pensamento e obra dos principais representantes russos**. Uberlândia: EDUFU, 2013.

SAXE, G. B.; GEARHART, M.; NOTE, M.; PADUANO, P. A interação de crianças e o desenvolvimento das compreensões lógico-matemáticas: uma nova estrutura para a pesquisa e a prática educacional. In: DANIELS, Harry (org.). **Vygotsky em foco: pressupostos e desdobramentos**. (Tradução: Mônica Martins e Elisabeth Cestari). Campinas, SP: Papyrus, 1994.

SILVA, Edilza Maria C. **Como são propostas pesquisas em livros didáticos de ciências e matemática dos anos iniciais do ensino fundamental**. Dissertação (Mestrado) – Universidade Federal de Pernambuco, Centro de Educação. Programa de Pós-Graduação em Educação Matemática e Tecnológica, 2013.

SILVA, Rita de Cássia B. da. **É a moeda que diz, não é a gente que quer não: conhecimentos probabilísticos de crianças em situações de jogos**. Dissertação (Mestrado) – Universidade Federal de Pernambuco, Centro de Educação. Programa de Pós-Graduação em Educação Matemática e Tecnológica, 2016.

SOARES, M. A.; NEHRING, C. M. Proporcionalidade: uma análise de livros didáticos do ensino fundamental. **Anais do VI Congresso Internacional de Ensino da Matemática**. Canoas-RS. ULBRA, outubro de 2013.

TRIOLA, Mario F. **Introdução à estatística**: atualização da tecnologia. (Tradução e revisão técnica: Ana Farias e Vera Flores). 11 ed. Rio de Janeiro: LTC. 2013.

TUKEY, John W. We need both exploratory and confirmatory. **The American Statistician**, v.34, n.1, fev.1980. p. 23-25.

VERGARA, Sylvia C. **Projetos e relatórios de pesquisa em Administração**. 5 ed. São Paulo: Atlas, 2004.

VYGOTSKY, Lev S. Mind in Society: The development of higher psychological processes. Tradução (Cap. 6): Martin Lopez-Morillas. Editor: COLE, Michael e colaboradores. Cambridge: Harvard University, 1978.

WATSON, Jane. Foundations for improving statistical literacy. **Statistical Journal of the IAOS**, 27, jan, 2011. p. 197-204.

WILD, C. J.; PFANNKUCH, M. Statistical Thinking in Empirical Enquiry. **International Statistical Review**, 67(3), 223-265, 1999. Disponível em: <<http://www.stat.auckland.ac.nz/~iase/publications/isr/99.Wild.Pfannkuch.pdf>>. Acesso em: 09 de maio de 2017.