



COMPREENSÕES SOBRE AMOSTRA AO MANIPULAR DADOS NO SOFTWARE TINKERPLOTS: UM CASO DE UMA PROFESSORA POLIVALENTE

UNDERSTANDINGS ABOUT SAMPLING FROM MANIPULATION OF DATA IN TINKERPLOTS SOFTWARE: A CASE OF A POLYVALENT TEACHER

Maria Niedja Pereira Martins¹; Carlos Eduardo Ferreira Monteiro²; Tamires Nogueira Queiroz³

Universidade Federal de Pernambuco, UFPE, Brasil

Resumo

Este artigo discute os resultados de um estudo de caso que investigou a compreensão sobre amostra por uma professora polivalente ao utilizar o *software TinkerPlots*. Buscou-se identificar a compreensão da professora a respeito de três aspectos básicos do conceito de amostra: tamanho, tipo e representatividade. O estudo também analisou as influências das ferramentas do *TinkerPlots* na compreensão desses três aspectos. A professora passou por uma entrevista inicial sobre o conceito de amostra; uma sessão de apresentação do *software*; realizou duas atividades de amostragem no *TinkerPlots*; e ao final respondeu a uma entrevista que teve o objetivo de identificar possíveis mudanças na compreensão sobre o conceito de amostra. As etapas da pesquisa foram registradas em áudio e vídeo e as transcrições geraram protocolos. Observou-se que a professora apresentou mudanças na compreensão sobre o tamanho e a representatividade em algumas amostras. Ela pode identificar tamanhos e preocupar-se com vieses em amostras pequenas, a partir de simulações usando a ferramenta *Sampler*, e da manipulação de dados usando a ferramenta *Plot*. As análises sugeriram que as atividades de amostragem desenvolvidas com esse *software* facilitaram compreensões sobre amostra.

Palavras-chave: Educação Estatística; Amostra; *Software TinkerPlots*.

Agência de fomento: CAPES

Abstract

This paper discusses the results of a case study that investigated the understanding of sampling by a teacher while using the software *TinkerPlots*. The aim was identifying the teacher's understanding about the three basic aspects of sample concept: size, type and representativeness. The study also examined the influence of *TinkerPlots* tools in understanding these three aspects. The teacher answered an initial interview about her knowledge concerning the concept of sample; then the researcher presented the software; she developed two activities with the *TinkerPlots* about sampling; and at the end she answered an interview that aimed to identify possible changes in her understanding about sampling. The stages of the research were recorded in audio and video and their transcripts generated protocols. The analysis indicated that the teacher presented changes about her understanding of size and representativeness in some cases. She can identify sizes and worry about biases in small samples from simulations using the *Sampler* tool and data manipulation using the *Plot* tool. The data analysis suggested that activities related to sampling using the *TinkerPlots* facilitated the understanding about sampling.

Keywords: Statistics Education. Sampling. *TinkerPlots* Software.

¹ Mestranda do Programa de Pós-Graduação em Educação Matemática e Tecnológica da Universidade Federal de Pernambuco (UFPE), Recife, PE, Brasil. E-mail: martinsniedja@hotmail.com

² PhD in Mathematics Education pela University of Warwick, Inglaterra. Professor da Universidade Federal de Pernambuco (UFPE), Recife, PE, Brasil. E-mail: carlos.fmonteiro@gmail.com

³ Mestranda do Programa de Pós-Graduação em Educação Matemática e Tecnológica da Universidade Federal de Pernambuco (UFPE), Recife, PE, Brasil. E-mail: tamires_nq@hotmail.com



COMPREENSÕES SOBRE AMOSTRA AO MANIPULAR DADOS NO SOFTWARE TINKERPLOTS: UM CASO DE UMA PROFESSORA POLIVALENTE

Introdução

A partir da necessidade de compreender como as pessoas interpretam dados estatísticos comumente disseminados pelos meios de comunicação social, tem-se desenvolvido pesquisas enfocando tais ações humanas. A Educação Estatística, em particular, possui o papel de possibilitar aos indivíduos a compreensão dos fenômenos sociais por meio da compreensão de conceitos estatísticos.

Temos observado, no entanto, a necessidade de mais investigações no Brasil a respeito de diversos conceitos específicos da atividade envolvendo dados estatísticos; tais como: variabilidade, média, inferência, probabilidade, dentre outros. Mais especificamente, poucas pesquisas têm sido realizadas, até recentemente, sobre a compreensão da amostra no contexto da Educação Básica.

A amostra tem sido defendida por Bolfarine e Bulsab (2005) como um elemento presente em diversas atividades do dia-a-dia. Os procedimentos que envolvem a escolha de uma amostra são denominados de amostragem e se fazem presentes nas atividades cujo sujeito procura realizar uma generalização. Por isso, o trabalho com amostras na Estatística pode conduzir o sujeito a definir escolhas, refletir sobre os dados, prever tendências, estimar resultados e generalizar dados. Percebe-se assim que tais habilidades podem ser incorporadas socialmente pelos sujeitos, evidenciando a importância do trabalho com esse conceito desde os anos iniciais.

Pesquisas que investigam os processos de ensino e de aprendizagem da Estatística (GUIMARÃES; GITIRANA; ROAZZI, 2001; LIRA, 2010; ASSEKER, 2011) têm demonstrado que atividades estatísticas desenvolvidas com auxílio do computador diferem significativamente das atividades que envolvem apenas lápis e papel. Com a utilização de um *software* para análise de dados, por exemplo, é possível gerar dados, formular diferentes tipos de representação, manipular diferentes aspectos de uma mesma representação, dentre outras ações. Todavia, Padilha (2012) afirma para que a utilização pedagógica do computador seja proveitosa e eficiente, é preciso que os professores possuam uma formação técnica-pedagógica específica.

Este artigo apresenta aspectos de uma pesquisa de dissertação de mestrado que buscou analisar a compreensão de professores dos anos iniciais sobre amostra ao utilizar o *software TinkerPlots*. Os resultados deste estudo podem contribuir para ampliar a compreensão sobre aspectos da aprendizagem sobre amostra; para identificar elementos no *software* que podem auxiliar no desenvolvimento desse conceito por professores, e para incentivar a aproximação de docentes dos anos iniciais com ferramentas tecnológicas.



Amostra e Amostragem

Bolfarine e Bussab (2005) conceituam amostra como qualquer subconjunto de uma dada população que, por sua vez, pode ser compreendida como sendo um conjunto de elementos que possuem pelo menos uma característica em comum. A partir das características da população a ser estudada e das restrições de orçamento de uma pesquisa, bem como das características que se deseja obter a partir da amostra, podemos considerar diferentes técnicas de amostragem.

O conceito de Amostragem refere-se a um procedimento tipicamente inferencial que permite chegar a uma conclusão a partir do estudo de uma ou mais amostras (COUTO JUNIOR, 2009). As diferentes técnicas de amostragem podem ser agrupadas em dois grandes grupos: probabilísticas e não probabilísticas. As probabilísticas envolvem a possibilidade de cada elemento da população ter a mesma probabilidade de pertencer à amostra. As não probabilísticas, por sua vez, caracterizam a ausência de mesma probabilidade na escolha dos elementos a serem incluídos.

Os métodos de amostragem probabilísticos se ramificam em diferentes tipos. Os mais comuns são: a Amostragem Aleatória Simples, a Amostragem Estratificada Proporcional e a Amostragem Sistemática.

A Amostragem Aleatória Simples é uma técnica que permite a seleção dos elementos de uma população por meio de uma mesma probabilidade. Assim, cada um dos elementos que compõem uma determinada população tem igual probabilidade de serem escolhidos para formar uma amostra. A Amostragem Estratificada Proporcional só pode ser utilizada quando os dados aparecem de forma ordenada. Esse procedimento consiste na retirada de K elementos da população de modo que os elementos a serem escolhidos para compor a amostra serão determinados por um intervalo. A Amostragem Sistemática, por sua vez, pode ser utilizada quando a população permite ser fragmentada em subconjuntos.

Bolfarine e Bussab (2005) argumentam sobre a importância de se planejar, bem como extrair uma amostra a partir do procedimento de amostragem mais adequado à situação de pesquisa. Para eles, o propósito é fornecer informações que permitam descrever da maneira mais adequada possível os parâmetros do universo da população a ser investigada. Para determinações adequadas de uma amostra, no entanto, é preciso levar em consideração alguns aspectos relacionados. Notadamente três elementos parecem ser essenciais para iniciar uma pesquisa por amostragem, a saber, o tamanho, a representatividade e tipo de amostra. De um modo geral, esse processo por amostragem esbarra nas limitações do orçamento da pesquisa e no tamanho da amostra. Parece-nos, no entanto, que definir o que seria uma amostra grande ou pequena demais não é uma tarefa tão simples.

Na visão de Oliveira e Grácio (2005) para determinar o tamanho de uma amostra aleatória simples, por exemplo, deve-se considerar se os dados da população se apresentam de forma homogênea ou heterogênea, ou seja, se os



elementos da população apresentam um grau de variabilidade pequeno ou grande. Assim, para determinar adequadamente o tamanho de uma amostra é preciso relacioná-la a outros elementos.

Há diferentes formas de determinar o tamanho da amostra, mas algo fundamental é verificar se a população é constituída de casos finitos ou infinitos. Ao contrário de uma população finita, a população denominada infinita é aquela na qual o número de observações de seus elementos é, a princípio, indeterminável. Pode-se dizer que populações infinitas pressupõem a seleção de uma amostra grande, levando-se em consideração a variável a ser estudada e a heterogeneidade da mesma e que, populações finitas permitem a seleção de pequenas amostras levando-se em consideração os mesmos aspectos. Para todos os casos, no entanto, a amostra (n) sempre assumirá um valor menor de elementos que a população (N). Assim, sempre teremos $n < N$.

Tais preocupações tornam-se importantes no trabalho com amostragem, pois nesse tipo de atividade há sempre a necessidade de extrair amostras que sejam representativas. Podemos dizer, inclusive, que a representatividade é a condição mais importante numa investigação quando se pretende generalizar dados. Conforme explicitam Bolfarine e Bussab (2005) uma amostra é dita representativa quando suas características se assemelham ao máximo à população investigada.

Visando buscar sempre amostras representativas, alguns estatísticos fazem uso de procedimentos de amostragem em que é possível prever as condições de seleção das unidades amostrais, como, por exemplo, na amostra aleatória simples. Pode-se dizer que esse procedimento de amostragem poderia ser o mais indicado para extrair amostras representativas em várias situações de pesquisa

Em vista dos diferentes elementos necessários à realização de uma pesquisa por amostragem, pode-se concluir que o trabalho com amostras envolve uma série de conceitos e ideias que se interrelacionam. Isso tem levado alguns pesquisadores a considerar que os conceitos de amostra e amostragem são estruturalmente complexos de serem desenvolvidos (BEN-ZVI et al, 2011; WATSON, 2004; PFANNKUCH, 2008).

De acordo com Ben-Zvi et al (2011) o conceito de amostra apresenta a ligação com elementos, tais como intervalo e distribuição, inferência, probabilidade, aleatoriedade e interpretação de gráficos. Apesar disso, é possível identificar pesquisas que indicam a possibilidade de crianças em idade escolar apresentarem noções iniciais sobre amostragem, como por exemplo: Watson, 2004; Gomes e Guimarães, 2013.

Em um estudo desenvolvido por Gomes e Guimarães (2013) com estudantes do 5° e do 9° ano, no Brasil, foi visto que alunos desde o 5° ano já conseguem compreender conceitos ligados a amostragem a partir de relações que eles estabelecem com suas vivências no cotidiano. As autoras consideram ainda que envolver os estudantes em diferentes contextos e situações foi importante para que os alunos apresentassem conflitos com suas justificativas e avançassem nas ideias sobre um procedimento por amostragem.

Na sua investigação, Pfannkuch (2008) também observou que ao envolver estudantes em contextos e histórias com discussões sobre amostras, os alunos conseguiram identificar elementos como variabilidade amostral e relacionar população e amostra. Para essa autora, noções sobre amostragem são importantes de serem construídas por estudantes desde os anos iniciais, pois tal conceito faz parte de um bloco de conteúdos presente na Estatística Inferencial que envolve o aprendizado sobre situações do mundo real.

Watson (2004), por sua vez, observou avanços no raciocínio apresentado por alunos australianos entre 8 e 15 anos sobre amostragem. Alguns, inclusive passando de um estágio elementar e inicial para um estágio complexo de compreensão sobre o conceito. Assim, como forma de resumir seus estudos sobre amostragem com alunos australianos, aquela autora realizou categorizações sobre o nível de raciocínio que estudantes demonstraram possuir ao longo de suas pesquisas. A seguir, na Figura 1 podemos visualizar três níveis distintos de concepções sobre amostragem e suas categorias relacionadas apontadas por Watson (2004):

Figura 1: Níveis de Raciocínio sobre Amostragem

Nível 1 – Compreendendo a Terminologia.	
(Categoria 1)	Amostras pequenas sem método de seleção
(Categoria 2)	Amostras pequenas com seleção aleatória primitiva
Nível 2 – Entendendo a Terminologia no Contexto.	
(Categoria 3)	Amostras pequenas com pré-seleção de resultados
(Categoria 4)	Amostras equivocadas
(Categoria 5)	Grandes amostras com aleatoriedade ou distribuição
Nível 3 – Questionamento crítico de alegações sem fazer justificações	
(Categoria 6)	Grandes amostras de viés sensíveis

Fonte: Baseado no estudo de Watson (2004).

Conforme a Figura 1, estudantes situados no primeiro nível de raciocínio conseguiam definir o conceito de amostra. Nesse primeiro nível, encontraram-se ainda duas categorias de compreensões: na primeira, estudantes concordavam com uma amostra pequena sem fazer relação com a variação dos dados e não conseguiam dizer como fazer uma seleção; na segunda, os alunos concordavam com a seleção de pequenas amostras, conseguiam fornecer exemplos e indicavam uma seleção com aspectos sutis de aleatoriedade.

No segundo nível de raciocínio o sujeito conseguia definir a amostra num contexto de pesquisa. Na categoria 3, temos alunos que concordavam com pequenas amostras e identificavam variáveis tendenciosas. Também se encontram aqueles que realizam combinações equivocadas para a escolha da amostra, sugerindo a seleção de uma amostra pequena com um método adequado ou uma amostra grande com um método inadequado na categoria 4.



Na categoria 5, o sujeito sempre oferece uma amostra grande associada a um método de seleção aleatório.

O último nível de raciocínio refere-se às habilidades críticas necessárias para questionar afirmações sobre amostras feitas sem base estatística adequada. Na categoria 6 há os sujeitos que podem expressar interesse em evitar viés de seleção e identificar amostras tendenciosas em notícias, por exemplo.

O estudo de Watson (2004) se configura como uma importante referência a respeito do desenvolvimento conceitual sobre amostragem por se tratar de um estudo longitudinal que apresenta raciocínios variados sobre amostragem de alunos em idade escolar.

O software *TinkerPlots*

O *TinkerPlots* é um *software* projetado para construir e manipular dados estatísticos com uma interface simples (KONOLD, 2006), algo não muito comum nos *softwares* para análise de dados estatísticos existentes no mercado.

A escolha por esse *software* para fazer parte de nosso estudo foi baseada nas análises de pesquisas anteriores que apontam para a facilidade de manipulação que professores brasileiros têm apresentado no seu uso. No estudo de Alves (2011) estudantes do 5º ano de uma escola rural puderam manipular esse *software* conseguindo criar uma série de representações e oferecer sentido aos dados.

O estudo de Eugênio (2013) investigou a compreensão do conceito de média usando o *TinkerPlots* por alunos do ensino fundamental e evidenciou que o *software* ofereceu contribuições para que os alunos expressassem respostas mais elaboradas sobre média.

De modo geral, no entanto, percebe-se que as pesquisas realizadas com esse *software* no Brasil ainda precisam avançar no que concerne às investigações de conceitos específicos (LIRA; MONTEIRO, 2011), uma vez que o *software* apresenta diferentes recursos que possibilitam abordar conceitos estatísticos, tais como, distribuição, variabilidade, probabilidade e amostragem.

Apesar do crescimento no número de investigações com o *software TinkerPlots*, não encontramos no Brasil pesquisas que investigassem propriamente a compreensão do conceito de amostra com ele. No âmbito internacional, no entanto, podemos recorrer aos estudos de Ben-Zvi et al. (2011), Prodromou (2011) e Kasak e Konold (2010) envolvendo o uso do *TinkerPlots*.

A pesquisa de Ben-Zvi et al. (2011) envolvendo crianças de 11 anos de idade em Israel, mostrou que ao utilizarem amostras crescentes com o *TinkerPlots*, as crianças apresentaram ideias importantes relacionadas à inferência realizada em amostras de diferentes tamanhos, como também indicaram ao longo da pesquisa um desenvolvimento dessas noções.

Prodromou (2011), por sua vez, desenvolveu um estudo semelhante ao de Ben-Zvi, com alunos australianos entre 13 e 14 anos de idade. De acordo com a autora, os participantes realizaram conexões com conceitos estatísticos

importantes durante o aumento das amostras ao utilizarem o *TinkerPlots*, tais como, distribuição, variação, incerteza e interpretação de dados.

As observações de Prodromou (2011) e Ben-Zvi et al. (2011) oferecem bons esclarecimentos sobre o processo de desenvolvimento inferencial com amostras a partir de um recurso tecnológico. Esses dois estudos sugerem que o trabalho com o *TinkerPlots* no contexto de amostras crescentes puderam auxiliar estudantes a desenvolver noções sobre inferência e relacionar conceitos presentes numa pesquisa por amostragem.

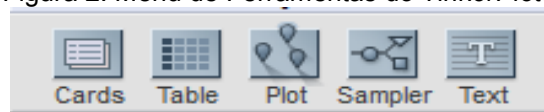
Outro estudo que também indica elementos para compreender as possibilidades de uso do *TinkerPlots* no desenvolvimento de noções sobre amostragem é a pesquisa de Kasak e Konold (2010) com estudantes do ensino secundário dos EUA. Tal pesquisa buscou investigar o raciocínio sobre probabilidade que os estudantes apresentavam a partir da utilização de um recurso computacional. Enquanto resultado, esses autores observaram que os participantes puderam “construir um modelo, executar um grande número de repetições em vários testes, e exibir como os dados estão reunidos” (KASAK; KONOLD, 2010, p.5). Uma das conclusões para esses resultados refere-se ao fato de que o processo realizado com os estudantes foi útil para que eles pudessem construir planos amostrais e desenvolvessem compreensões mais complexas sobre a aleatoriedade nas amostras utilizando uma base tecnológica.

Esses indícios nos fazem acreditar no potencial do *software TinkerPlots* para o trabalho com amostra, concebendo-o, inclusive, enquanto um recurso apropriado para o trabalho com professores que lecionam nos anos iniciais da escolarização.

3.1 Ferramentas do *software TinkerPlots*

A tela inicial do *TinkerPlots* é representada por uma área em branco. O menu de ferramentas é apresentado no idioma inglês, assim como todo o *software*, e traz cinco ferramentas básicas: *Cards*, *Table*, *Plot*, *Sampler* e *Text* direcionadas a construção e representação dos dados, conforme mostra a Figura 2:

Figura 2: Menu de Ferramentas do *TinkerPlots*.

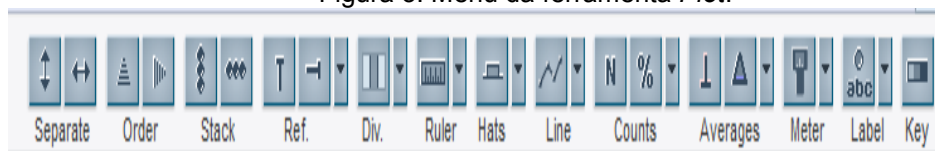


Fonte: Konold e Miller (2012).

A ferramenta *Cards* serve para registrar a criação de dados da pesquisa a ser realizada no *software*. O recurso *Table*, por sua vez, permite que os dados incluídos nos *Cards* sejam representados em uma tabela. Com a ferramenta *Plot* também é possível visualizar os casos incluídos nos *Cards* em uma janela. Essa ferramenta oferece a possibilidade de manipulação dos casos incluídos nos *Cards* permitindo que o usuário produza gráficos. Para manipular os casos visualizados no *Plot* o usuário deverá fazer uso do menu da ferramenta do *Plot* que é habilitado ao selecionarmos tal ferramenta.

As principais ferramentas disponíveis para a manipulação do *Plot* são: *Separate*, *Order*, *Stack*, cujos ícones aparecem no extremo esquerdo na configuração reproduzida na Figura 3. Essas ferramentas referem-se às funções de *separar*, *ordenar* e *empilhar* os casos incluídos nos *Cards*.

Figura 3: Menu da ferramenta *Plot*.



Fonte: Konold e Miller (2012).

Além dessas funções, o menu da ferramenta *Plot* também traz outras, como a *Ref.*, a qual aciona uma linha de referência e a constrói em cima dos dados presentes na janela *Plot*. A ferramenta *Hat*, por sua vez, aciona a construção de um *Hat Plot* ou de um *Box Plot*. A ferramenta *Counts* realiza contagens dos casos inseridos no *Plot* e apresenta duas opções que são a contagem numérica representada por n e a contagem percentual. Por fim, a ferramenta *Averages* pode apresentar a média e a mediana dos dados presentes na ferramenta *Plot*.

As ferramentas presentes no menu de manipulação do *Plot* podem ser utilizadas de forma simultânea e auxiliam na identificação de estatísticas, bem como facilitam o processo de análise de dados em algumas representações (ALVES, 2011).

Além das ferramentas *Cards*, *Table*, *Plot*, o menu inicial do *TinkerPlots* (Ver Figura 2) oferece um ícone para as ferramentas *Sampler* e *Text*. A ferramenta *Text* serve para incluir comentários numa caixa de texto sobre as tarefas desenvolvidas no *software*. A ferramenta *Sampler* é direcionada para o trabalho com amostras e, por isso, assume uma relevância particular neste estudo. Por esse motivo, e pela complexidade da ferramenta, explicaremos o seu funcionamento em uma seção específica.

A ferramenta *Sampler*

O *Sampler* é um simulador de probabilidade do *TinkerPlots 2.0* que permite a construção de um plano amostral aleatório simples, proporcional estratificado ou sistemático.

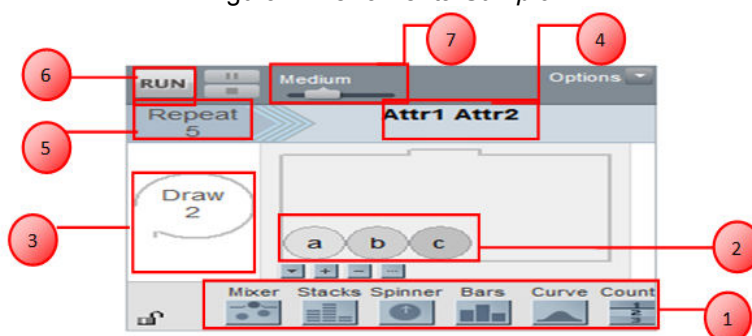
A base para a construção do *Sampler* foi realização da pesquisa teste com estudantes americanos e se apoiaram em ideias desenvolvidas por Liu e Thompson (2007). Em seus estudos, aqueles autores defendem que desenvolvimento conceitual em probabilidade, pode ser facilitado com auxílio da visualização de um processo repetível, probabilístico e com resultados claramente distribuídos.

Atividades que envolvem um contexto de amostragem estão sujeitas à manipulação de uma grande quantidade de dados. A utilização de uma base tecnológica para o desenvolvimento de atividades desse tipo estará, então, a manipulação das informações sem a necessidade de uma etapa extensa e dispendiosa para a sua coleta. Tal fato oferece possibilidades interessantes para

o trabalho com o conceito de amostra em sala de aula, uma vez que no contexto escolar o tempo demandado a esse tipo de atividade nem sempre apresenta relação com as aprendizagens objetivadas pelo professor.

Nesse sentido, o *software TinkerPlots* e, especificamente a ferramenta *Sampler* aliada aos recursos do *Plot*, torna-se uma ferramenta adequada para a realização de atividades sobre amostragem, pois possibilita a criação, exploração, análise e representação de vários dados. Os componentes básicos do *Sampler* incluem várias ferramentas e funções que são mostradas na Figura 4 e detalhadas no Quadro 1 a seguir:

Figura 4: Ferramenta *Sampler*.



Fonte: Konold e Miller (2012).

Quadro 1: Funções da Ferramenta *Sampler*.

1) Dispositivos: Objetos em que se pode preencher com diferentes elementos ou propriedades. Existem seis dispositivos: <i>Mixer</i> , <i>Stacks</i> , <i>Spinner</i> , <i>Bars</i> , <i>Curve</i> e <i>Count</i> .
2) Elementos do dispositivo: Apresenta os casos incluídos em um dispositivo. Eles indicam as propriedades de uma variável/atributo. Por exemplo, em um <i>Mixer</i> , os elementos são visualizados como círculos. Já em um <i>Spinner</i> , os elementos são apresentados como fatias, igualmente a um gráfico de setores.
3) Draw: Indica o número de variáveis/atributos que comporão uma amostra ou resultado.
4) Atributos: Apresenta e nomeia atributo(s) ou variável(is) que comporão a amostra.
5) Repeat Number: Indica o número de casos gerados a cada execução do <i>Sampler</i> .
6) Run: Inicia a extração de casos para amostra. Também acompanha botões que pausam ou interrompem a extração dos casos.
7) Cursor de Velocidade: Determina a velocidade com que os casos de uma amostra será extraída. Apresenta 9 velocidades diferentes.

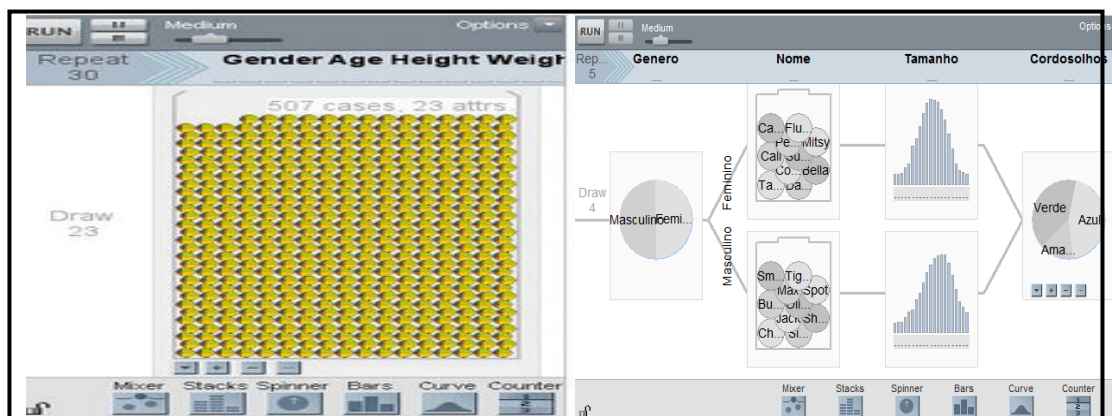
Fonte: Adaptado de Konold & Miller (2012).

A Figura 4 reproduz a tela do *TinkerPlots* após o acionamento da ferramenta *Sampler*. Conforme apresentado no Quadro 1, a ferramenta 1 corresponde aos dispositivos, os quais têm a função de determinar a forma como cada atributo incluído no *Sampler* será selecionado. Cada um dos dispositivos apresenta uma visualização mais adequada a um ou outro tipo de variável (ou atributo).

O *TinkerPlots* é configurado para que o usuário manipule suas ferramentas para criar uma amostra desejada. Tais dispositivos podem ser utilizados simultaneamente para a construção de planos amostrais, ou utilizado

somente um dispositivo com vários atributos distintos, como mostra as duas reproduções da tela do *TinkerPlots* na Figura 5.

Figura 5: Reproduções de telas do *TinkerPlots* com uso do *Sampler*. *Mixer* contendo 23 atributos; *Sampler* com vários dispositivos contendo apenas 1 atributo cada um.



Fonte: Konold e Miller (2012).

A tela do lado esquerdo na Figura 5 mostra um *Mixer* composto por 507 casos e 23 atributos. A tela do lado direito na Figura 5 mostra um *Sampler* organizado com três tipos de dispositivos diferentes para formar quatro atributos. A partir das diferentes organizações do *Sampler*, o usuário poderá criar diferentes tipos de amostras. No total, três tipos de amostra poderão ser criados por meio do *Sampler*: Amostra Aleatória Simples; Amostra Proporcional Estratificada e Amostra Sistemática.

Método

Nossa investigação procurou responder a seguinte questão de pesquisa: “Como se dá a compreensão sobre amostragem por professores dos anos iniciais ao utilizarem o *software TinkerPlots*?” O objetivo geral deste trabalho foi investigar a compreensão de uma professora polivalente sobre amostra ao utilizar o *software TinkerPlots*. Especificamente, buscou-se identificar a compreensão da professora a respeito do tamanho, da representatividade e do tipo de amostra trabalhada; bem como analisar o efeito das ferramentas do *TinkerPlots* utilizadas pela professora para o entendimento desses três aspectos. Como forma de responder nossos objetivos, realizamos as escolhas metodológicas descritas a seguir.



Contexto do estudo e participantes

A pesquisa foi realizada em uma escola municipal da Região Metropolitana do Recife - RMR, que se encontra registrada em um levantamento realizado pelo GPEME – Grupo de Pesquisa em Educação Matemática e Estatística (CARVALHO; MONTEIRO, 2012). O trabalho realizado por esse grupo buscou identificar as condições de uso dos laboratórios de informática das escolas municipais e estaduais da RMR, uma vez que se tem visto pouca exploração de recursos tecnológicos nas práticas educativas em algumas escolas da RMR.

Nas visitas iniciais do grupo, uma escola destacou-se dentre as demais, apresentando um espaço pequeno para o laboratório, mas sendo utilizado pelos estudantes entre 3 e 4 dias semanais, com 13 computadores e Internet (CARVALHO; MONTEIRO, 2012).

Tal instituição é uma escola municipal rural da cidade de Paulista. A escola atende a 353 estudantes do Ensino Fundamental distribuídos entre o 1º e 9º anos e na modalidade EJA – Educação de Jovens e Adultos.

Para a realização deste estudo, selecionamos uma professora polivalente que trabalha atualmente no laboratório de informática dessa instituição de ensino e por questões éticas será chamada de Ana, neste artigo. As aulas de Ana foram observadas durante o levantamento realizado pelo GPEME e definidas por esse grupo como sendo ações pedagógicas que mereciam destaque em relação às demais práticas desenvolvidas em outras escolas pesquisadas. Por essa conotação, apresentamos interesse em investigar tal caso.

Procedimentos de coleta de dados

Como forma de alcançar nossos objetivos, a abordagem metodológica foi composta por dados provenientes de diferentes instrumentos de coleta, tal qual a entrevista semi-estruturada e a videografia.

Na primeira sessão da pesquisa, realizou-se uma entrevista para identificação de conhecimentos da professora sobre amostra, e, em seguida, realizamos uma atividade de familiarização com o objetivo de levar Ana a reconhecer o funcionamento de alguns ícones dispostos no *TinkerPlots*. No segundo encontro, submetemos duas atividades sobre amostragem à professora. Finalmente, depois de um intervalo de uma semana em relação ao segundo encontro, realizamos uma terceira sessão de pesquisa na qual foram apresentadas as mesmas questões submetidas no primeiro encontro. Esses procedimentos buscou identificar as mudanças apresentadas pela participante em relação ao conceito de amostra. As questões que compuseram as entrevistas foram baseadas no estudo de Watson (2004), conforme mostra o Quadro 2:



Quadro 2: Questões das entrevistas realizadas na primeira e terceira sessão de pesquisa

1. a) Você já ou viu a palavra amostra antes? Onde? O que ela significa? b) um repórter da TV disse: "Em um estudo realizado sobre o peso de crianças do 5º ano alguns pesquisadores entrevistaram uma amostra de crianças do 5º ano no Estado." O que a palavra amostra quer dizer nessa frase?
2. a) Por que você acha que os pesquisadores usaram uma amostra de crianças do 5º ano, ao invés de estudar todas as crianças do 5º ano série do Estado? b) Você acha que eles usaram uma amostra de cerca de 10 alunos? Por que sim e Porque não? Quantas crianças eles deveriam escolher para essa amostra? Por que? c) Como eles poderiam escolher as crianças para a amostra deles? Por que?
3. Os pesquisadores foram a duas escolas: uma escola no centro da cidade e uma escola no interior. Cada escola teve cerca de metade meninas e metade meninos. Os pesquisadores escolheram uma amostra aleatória de cada escola: 50 crianças da escola da cidade e 20 crianças da escola do interior. Uma dessas amostras foi incomum: teve mais de 80% de meninos. Esses 80% é mais provável de ter surgido na amostra grande de 50 da escola da cidade ou na amostra pequena de 20 da escola do interior ou as ambas poderiam ter sido igualmente incomum? Por favor, explique sua resposta.
4. Cerca de 6 em 10 estudantes escolares do ensino médio nos Estados Unidos dizem que eles poderiam conseguir uma arma se eles quisessem uma, um terço deles dentro de uma hora, mostra uma pesquisa. A sondagem que envolvia 2.508 estudantes do primeiro e último ano do ensino médio em Chicago também encontrou que 15% deles carregaram um revólver nos últimos 30 dias, com 4% tendo levado à escola. a) Você faria alguma crítica ao que é dito nesse artigo? b) Se você fosse um professor do ensino médio, essa pesquisa faria você recusar uma oferta de emprego em algum lugar nos Estados Unidos, tal como Colorado ou Arizona? Por que sim ou Por que não?
5. Cerca de 96% dos ouvintes que ligaram para a estação de rádio jovem triple J disseram que o uso da maconha deveria ser descriminalizado na Austrália. A enquete feita com os ouvintes pelo, que encerrou ontem, mostrou que 9.924 - em 10.000 ligações - são a favor da descriminalização, disse a estação. Apenas 389 acreditavam que a posse de drogas poderia continuar como uma ofensa criminal. Muitas pessoas que ligaram enfatizaram que não fumavam maconha, mas acreditavam na descriminalização de seu uso, apontou a Triple J. a) Qual foi o tamanho da amostra nesse artigo? b) Essa amostra que foi relatada aqui é uma forma confiável para descobrir o apoio do público para a descriminalização da maconha? Por que sim ou por que não?

Fonte: Adaptado do estudo de Watson (2004).

Nas sessões direcionadas a realização das entrevistas, utilizamos apenas a gravação de áudio. A videografia serviu para registrar o áudio, a imagem da entrevistada e suas manipulações feitas no computador no momento da Atividade de Familiarização e nas Atividades de Exploração sobre Amostra no *Software*.

As atividades de familiarização e exploração foram criadas a partir dos arquivos de dados incluídos do *software TinkerPlots*. Na atividade de familiarização utilizamos um banco de dados com informações sobre características de gatos. Essa etapa ocorreu após a realização da primeira entrevista e consistiu em apresentar como se utilizava as principais ferramentas do *TinkerPlots*.

Em seguida, nas atividades de exploração, utilizamos dois bancos de dados distintos presentes nos arquivos do *TinkerPlots*. O primeiro refere-se a um banco de dados sobre peixes normais e peixes geneticamente modificados. A intenção dessa atividade era fazer com que a professora aumentasse gradativamente o tamanho das amostras aleatórias de peixes e definisse qual grupo de peixes apresentava um comprimento maior em relação ao outro. Para tanto, conduzimos a professora a extrair amostras do simulador *Sampler* e oferecer níveis de confiança informais sobre suas inferências em cada amostra.

O segundo banco de dados foi introduzido na sequência e apresentava um contexto de números inteiros reais. O banco continha 500 números inteiros distribuídos de 0 a 100 inseridos no simulador *Sampler*. O desafio consistia em estimar informalmente o aglomerado em que se situava a população a partir de uma amostra menor possível. Para tanto, oferecemos a situação fictícia que determinava à professora o pagamento de R\$1,00 a cada retirada de uma amostra com $n = 5$. A professora deveria inferir o aglomerado da amostra à população, gastando o menos possível.

Análises de dados

Para a análise dos dados, procuramos transcrever as falas da professora em todas as etapas da pesquisa e construímos protocolos. Nas sessões referentes à familiarização e intervenção resgatamos as representações, as ferramentas selecionadas e as manipulações realizadas no *software* a fim de construirmos um roteiro que visualizasse esses dois processos nos protocolos. Assim, a partir de uma leitura inicial dos protocolos destacamos momentos significativos que evidenciam nossa intencionalidade no estudo.

Resultados e Discussão.

Nas entrevistas com a professora, pudemos identificar que tipo de conhecimentos sobre amostra Ana mantinha antes e após a realização das atividades no *TinkerPlots*. Como forma de sintetizar os resultados, elaboramos o seguinte quadro que apresenta as principais mudanças evidenciadas pela participante:

Quadro 3: Resultados da Entrevista 1 e 2.

Aspectos analisados	Entrevista 1	Entrevista 2
Amostra	<ul style="list-style-type: none">• Amostra é uma prova;• O termo tem sentidos diferentes a depender do contexto;• Confunde amostra e população.	<ul style="list-style-type: none">• Amostra é um modelo; uma pequena porção;• Indica a utilidade de uma amostra.
Tamanho	<ul style="list-style-type: none">• Discordou com uma amostra de $n = 10$;• Não questionou a heterogeneidade dos dados;• Não reconheceu o tamanho de uma amostra num contexto.	<ul style="list-style-type: none">• Concordou com uma amostra de $n = 10$;• Não questionou a heterogeneidade dos dados;• Reconheceu o tamanho de uma amostra num contexto.
Representatividade	<ul style="list-style-type: none">• Não identifica amostras tendenciosas em um contexto;• Não demonstra preocupação com vieses ao indicar seleção de amostra;• Utiliza o termo “média” como sinônimo de representatividade.	<ul style="list-style-type: none">• Não identifica amostras tendenciosas em um contexto;• Demonstra preocupação com vieses ao indicar seleção de amostra.
Amostragem	<ul style="list-style-type: none">• Destaca critérios incompatíveis com a finalidade da pesquisa para a seleção de amostra;• Escolhe elementos ao acaso;• Faz referência ao termo sorteio como sinônimo de escolha ao acaso.	<ul style="list-style-type: none">• Destaca critérios compatíveis com a finalidade de uma pesquisa para a seleção de amostra;• Escolhe elementos ao acaso.

Fonte: Os autores.



Na primeira entrevista, anterior à intervenção, a professora definiu a amostra como uma prova e demonstrou compreender esse conceito em uma situação social, identificando que o termo toma significados distintos a depender do contexto. Apesar do conhecimento social sobre a amostra, ao que parece a professora expressou uma dificuldade em identificar características específicas que diferenciam uma amostra de uma população numa situação de pesquisa, confundindo suas denominações e, por consequência, apresentando uma concepção equivocada sobre o motivo de se trabalhar com o conceito.

Ana discordou da seleção de uma amostra com 10 unidades advinda de uma população hipoteticamente grande, mas não conseguiu realizar uma seleção adequada, inserindo variáveis tendenciosas e um método por conveniência. Acreditamos que isso se deu em função da dificuldade que a professora apresentou para diferenciar amostra e população. Além disso, ela se valeu de expressões informais ou mesmo equivocadas para tratar de conceitos relacionados à amostragem, tais como a palavra “média” para indicar uma possível amostra representativa e “sorteio” como uma expressão simples para indicar a escolha ao acaso de uma unidade amostral.

Esses dados indicam que Ana estava, na primeira entrevista, em nível elementar de compreensão sobre amostragem. Baseados no estudo de Watson (2004), cujos níveis de raciocínio sobre amostragem são apresentados na Figura 1, podemos considerar que a professora pareceu estar no Nível 2, uma vez que, ora ela apresentava considerar o conceito de amostra por meio de um contexto de pesquisa social oferecendo grandes amostras, ora a professora apresentava dificuldades para indicar um método de seleção adequado.

Os resultados após a intervenção evidenciam algumas mudanças em relação à primeira etapa da coleta. Na entrevista final, Ana considerou a amostra enquanto um modelo, bem como a ressaltou do ponto de vista de uma pequena porção. Além disso, foi possível perceber uma compreensão sobre a utilidade da amostra. Ainda na entrevista final, destaca-se que Ana demonstrou reconhecer o motivo de se utilizar uma amostra quando verbalizou que um levantamento seria dispendioso se estudássemos todos os casos de uma população.

Em relação ao tamanho da amostra, a professora concordou com uma amostra de $n = 10$, desconsiderando aspectos como o tamanho da população e a heterogeneidade dos dados. De modo similar, a entrevistada ainda insistiu em um método de seleção inadequado, no entanto, apresentou características importantes que nos fizeram considerar um avanço na sua compreensão sobre amostras. A educadora ofereceu critérios que se relacionavam com a intencionalidade da pesquisa para a seleção de uma amostra. Acreditamos que essas escolhas feitas pela entrevista vinculam-se ao início de sua reflexão sobre vieses parciais em amostras e que pode estar relacionado ao envolvimento com uma atividade de amostragem na intervenção.

Mesmo apresentando inicial interesse em evitar vieses em amostras nessa segunda etapa da sondagem, tal construto não foi suficiente para fazer com que a professora conseguisse identificar amostras tendenciosas em notícias, sendo conivente com os resultados expressos nas situações expostas em

notícias. Apenas no que se refere ao caráter de generalização dos dados - na quarta questão referente à localidade em que a pesquisa sobre porte de armas poderia ser representativa - houve uma análise adequada da professora na entrevista final.

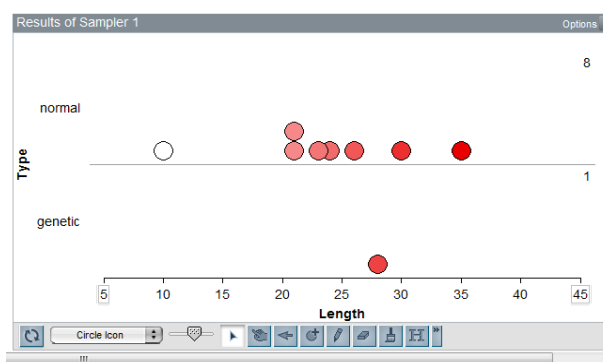
A fim de explicitar como tais mudanças puderam se constituir, trazemos na próxima sessão as principais aquisições de Ana.

Compreensões sobre tamanho, representatividade e tipo de amostra na Atividade 1

Nessa primeira atividade a professora foi conduzida a utilizar o *Sampler* para extrair amostras aleatórias de uma população de peixes a fim de identificar qual grupo de peixes apresentava um comprimento maior. A professora extraiu amostras com $n = 9$, $n = 18$, $n = 45$, $n = 95$ e $n = 145$ até chegar à população total. Neste artigo, discutem-se as considerações realizadas pela professora a partir de algumas dessas amostras.

De modo diferente da primeira entrevista, na entrevista final a professora Ana demonstrou considerar que uma pequena amostra poderia trazer conclusões satisfatórias sobre a população quando selecionou uma amostra de $n = 9$ no *TinkerPlots*. Sobre isso, extratos da entrevista com a professora a seguir ilustram suas compreensões quando das manipulações no *software*:

Figura 6: Amostra com $n = 9$ na Atividade 1



Fonte: Os autores.

Pesquisadora: Então, com essa amostra que a gente pegou, dá pra saber se os geneticamente modificados são maiores?

Ana: Tem! Dá!

P: Então, o que você diria ao piscicultor?

A: Eu diria que... os peixes geneticamente... não... que os normais... É. Deu dois... (Referindo-se a dois peixes normais que ultrapassam o único peixe geneticamente modificado da amostra).

P: Dá pra dizer algo?

A: Dá pra dizer que não precisa gastar com genética não! (risos) Deixa os peixes normais mesmo.

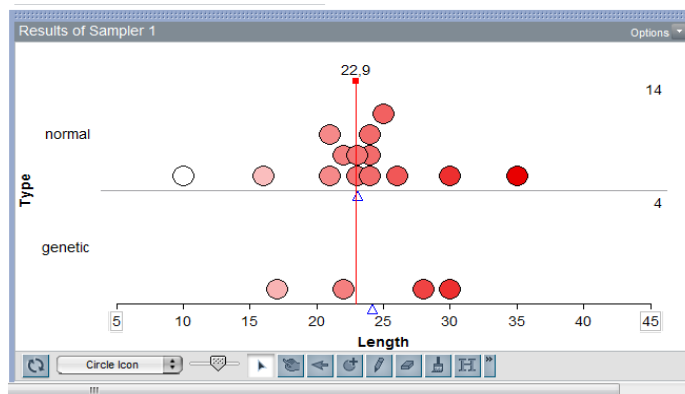
Na primeira amostra retirada a professora não conseguiu estabelecer uma interpretação adequada da representação e por isso, generalizou precipitadamente os resultados da amostra à população. Na verdade,

o gráfico mostrado na Figura 6 não poderia favorecer uma interpretação confiável, pois, o tamanho da amostra era pequeno em relação à população total de peixes. Ao fazer isso, no entanto, a professora ofereceu a resposta de que os peixes normais tendiam a crescer mais que os peixes geneticamente modificados. Acreditamos que a visualização de mais casos de peixes normais que de peixes geneticamente modificados no *Plot* influenciou a professora a oferecer essa conclusão, por esses não apresentarem uma quantidade equivalente entre si.

A professora também fez uso da ferramenta média quando a amostra permanecia com poucos casos. Por estar sendo utilizada em poucos dados, essa ferramenta não auxiliou a professora a confirmar sua resposta. Assim, os dados iniciais mostraram que a professora apresentou uma dificuldade em interpretar corretamente os dados das amostras e utilizar estatísticas que pudesse contribuir para o processo de interpretação. Percebeu-se que quando envolvemos a professora num contexto de interpretação de gráficos, sua dificuldade inicial de interpretação pareceu influenciar na identificação de vieses na amostra. Esses resultados podem ter explicação na relação que a amostragem pode apresentar com a interpretação de gráficos, conforme sugerem Ben-Zvi et al (2011) e Pfannkuch (2008).

Somente ao aumentarmos a amostra para $n = 18$, Ana pôde nos oferecer uma conclusão mais precisa sobre os dados ao utilizar a ferramenta média:

Figura 7: Ferramenta *Average* (média) acionada numa amostra com $n = 18$.



Fonte: Os autores.

Ana: Eu diria que se ele investisse no genético ele teria mais chances de aumentar os peixes de forma mais exata. Já nos peixes normais, o tamanho iria variar. Não vai ter sempre aquele peixe do mesmo tamanho.

Pesquisadora: E se eu quisesse quantificar a certeza que você tem? Numa escala de 0 a 10. Qual a certeza que você tem que essa resposta é correta?

A: 10.

P: 10? Então, olhando para cá, você acha que as conclusões que você tirou olhando para 18 peixes vai ser a mesma que você pode tirar para todo o resto dos peixes?

A: Eu acho!



A partir desse fragmento da entrevista, percebe-se que Ana identificou a natureza heterogênea da variável peixes normais. Apesar de esse ser um dado importante para realizar considerações sobre a possibilidade de a amostra ser ou não representativa, a professora parece se guiar apenas pela tendência do gráfico para oferecer uma resposta final. Assim, vimos que Ana pôde perceber a heterogeneidade da amostra, mas não conseguiu estabelecer nenhuma relação entre esta e a representatividade da amostra ou mesmo com o tamanho necessário para que uma amostra com dados heterogêneos fosse representativa.

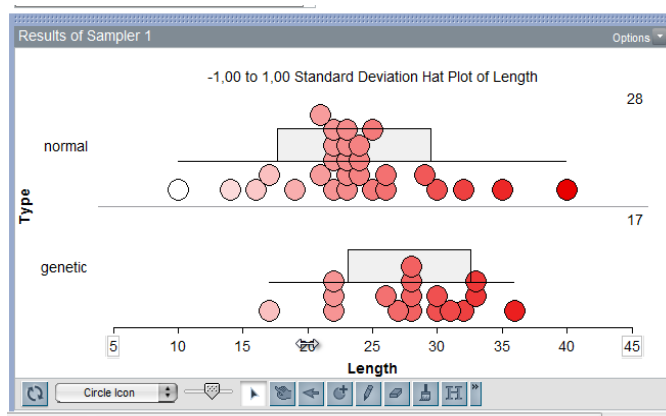
Mesmo apresentando uma compreensão sobre a necessidade de aumentar mais a amostra para obter uma representatividade melhor, ou uma inferência final mais confiável, Ana não ofereceu níveis de confiança adequados em relação ao tamanho da amostra ao longo da atividade. Um indício importante para esta conclusão foram os intervalos atribuídos pela professora os quais eram relativamente altos.

Ana pôde oferecer níveis de confiança nas amostras com $n = 18$, $n = 45$ e $n = 95$, mas apenas a partir da amostra com $n = 45$ ela apresentou uma interpretação correta sobre gráfico da Atividade 1. Essa resposta se manteve nas demais amostras apresentadas, sendo o intervalo de confiança oferecido pela professora, bem como a possibilidade de visualizar no crescimento da amostra importantes elementos para refletir sobre o impacto do aumento da amostra na representatividade dos dados.

Percebemos que ao aumentarmos o tamanho da amostra, a professora foi estabelecendo respostas mais seguras em relação à população. A professora baseou-se principalmente na tendência dos gráficos para retirar conclusões sobre cada amostra e comparou com os resultados das amostras menores.

Mesmo apresentando intervalos de confiança relativamente altos, identificou-se que o trabalho com amostras crescentes no *TinkerPlots* permitiu à professora compreender que amostras menores tendem a representar de forma menos confiável os dados de uma população. Uma fala que expressa essa afirmação, pode ser visualizada a seguir:

Figura 8: Ferramenta *Hat* acionada em amostra com $n = 45$.



Fonte: Os autores.

Pesquisadora: Qual seria o aglomerado que demonstra um tamanho maior em centímetros?

Ana: Esse (Geneticamente modificados).

P: Então, eu posso dizer que os genéticos estão aglomerados num lugar que indica um tamanho maior na escala do que os normais?

A: É. Os normais tem muito, mas estão atrás. Entendi.

P.: Então, o que você percebe? Qual é o conjunto que está maior?

A.: Os genéticos.

P.: Com essa quantidade de peixes que você pegou, você poderia dizer que essa ideia que você concluiu se aplicaria aos 625?

A.: Sim.

P: Quanto de confiança você tem nessa sua resposta? De 0 a 10. É 10 também? Tu acha que é isso?

A: Não. Mais ou menos 9.

P: Você diz que tem 9, numa escala de 0 a 10, de confiança que esse resultado se aplica ali. Se ao invés de colocar mais nove, eu colocasse 50?

A: Eu ia ter mais confiança ainda.

P: Por quê?

A: Porque se é mais, a gente pode ter mais certeza, né?

A professora argumentou a inferência sobre os dados tenderia a ser mais confiável ao selecionarmos mais casos para a amostra. Parece-nos, que o contexto de uso do *TinkerPlots* associado a amostras crescentes foi importante para que a professora apresentasse uma noção mais elaborada sobre a relação entre representatividade e tamanho das amostras aleatórias, assim como nos estudos de Ben-zvi et al (2011) e Prodromou (2011). Sabemos, porém, que essa compreensão expressa uma ideia parcial sobre a representatividade em amostras, uma vez que faltaram evidências de compreensões a respeito do processo de seleção da amostra e da variabilidade dos dados nessa atividade.

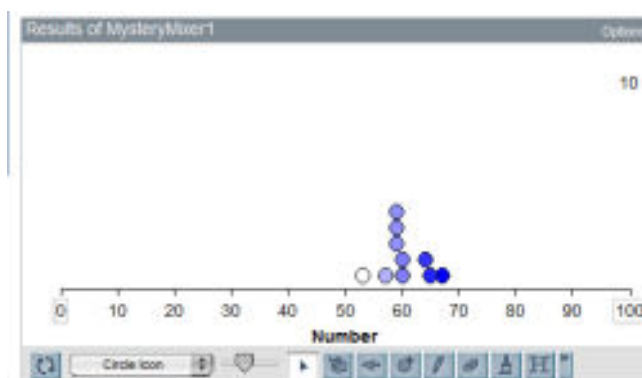
Ainda nesse mesmo fragmento, notamos que a ferramenta *Hat* pareceu facilitar a identificação da tendência dos dados no *Plot*, o que possivelmente facilitou a interpretação dos resultados da amostra pela professora.

Com isso, foi visto que a professora conseguiu obter clareza sobre o objetivo da atividade e sobre suas ações, uma vez que Ana demonstrou refletir sobre o aumento da amostra e a mudança na sua representatividade. Observamos ainda que a noção de representatividade esteve implícita na tarefa, na medida em que Ana demonstrava preocupação em obter um nível de confiança bom para dar força a sua generalização.

Compreensões sobre tamanho, representatividade e tipo de amostra na Atividade 2

Na segunda atividade, Ana deveria extrair amostras aleatórias do *Sampler* a fim de identificar um intervalo para a concentração dos dados extraídos do simulador. Ela selecionou amostras do simulador sem grandes dificuldades. A princípio, Ana extraiu uma amostra de $n = 5$, gastando apenas 1 real e achou conveniente retirar mais 5 casos da população, oferecendo na sequência a seguinte conclusão:

Figura 9: Amostra com $n = 10$ do banco de dados "MysteryMixer1."



Fonte: Os autores.

Ana: Menina, acho que eu não vou gastar mais real, não.

Pesquisadora: Por quê?

A: Acho que é o 60. De 10, 6 foram para o 60.

P: Então, numa amostra de 10, 6, aliás, 5 casos estão no 60.

A: É. o outro ta encostadinho no 60! Eu tenho crédito ou não?

P: Tem! Você pode gastar quanto quiser, agora você é uma pessoa econômica, seria melhor gastar o menos possível.

A: Ah, é? E, se eu quisesse parar?

P: Aí, você teria de ter uma confiança boa de que sua resposta estaria correta.

A: E, se eu perder?

P: Aí, todo o trabalho que você teve não vai valer de nada.

A: Então eu prefiro comprar mais... são 500! [seleciona mais 5 casos]

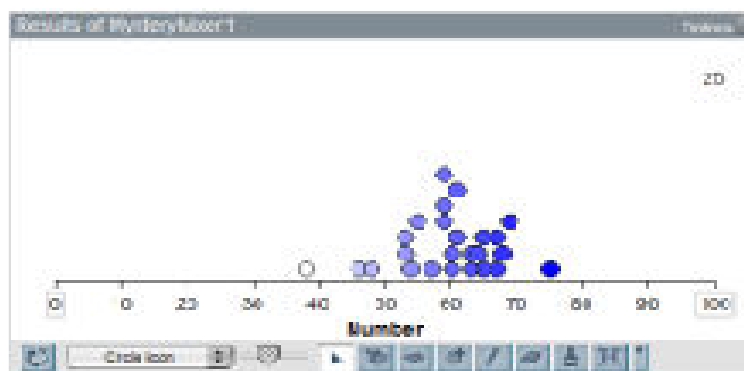
Nota-se na conversa estabelecida com a professora que ela denotou preocupação em estabelecer uma confiança razoável para sua amostra. Ela percebeu que tal conclusão ainda é incerta em função da população total de 500 casos. As conclusões retiradas pela professora nesse banco de dados sugerem

que a participante considerou que o custo para a pesquisa era significativo na determinação do tamanho da amostra. Acreditamos assim, que o contexto envolvendo o custo hipotético com a pesquisa possa ter contribuído para que a professora pudesse equilibrar o tamanho da amostra e a representatividade da mesma.

Diferentemente da Atividade 2, notamos que a professora não demonstrou dificuldades em interpretar os dados desse gráfico. É possível visualizar nos comentários da professora que a mesma conseguiu observar a concentração dos dados nessa amostra rapidamente. Acreditamos que essa facilidade deveu-se ao fato da possibilidade de visualização dos casos no *Plot*, mas, principalmente pelos dados se apresentarem de forma mais homogênea e com apenas uma variável.

Ao aumentar o tamanho da amostra ($n = 15$), Ana indicou com mais precisão que os dados estariam concentrados no número 60. Quando problematizamos o valor da população em relação ao tamanho da sua amostra, Ana nos ofereceu o entendimento de que os processos aleatórios não garantiam a representatividade das amostras menores. Com isso, Ana sentiu a necessidade de extrair mais 10 casos do simulador, totalizando 25 casos para a amostra.

Figura 10: Amostra com $n = 25$ do banco de dados "MysteryMixer1 ."



Fonte: Os autores.

Ana: Acho que vai ficar entre 60-70.

Pesquisador: Você gastou 4,00 reais. Tu acha que poderia dizer com certeza?

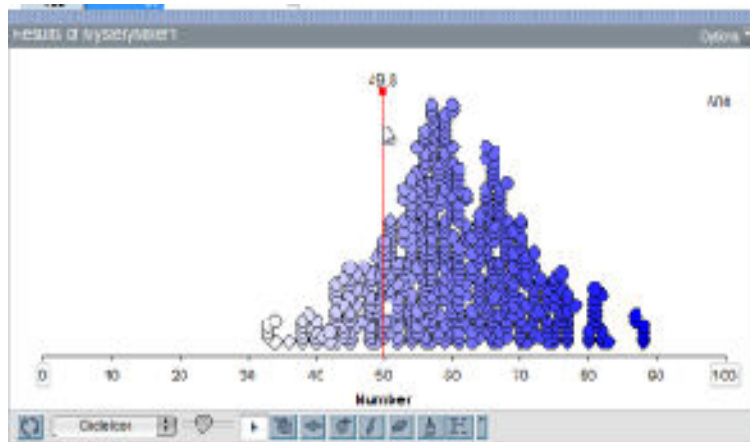
A: humrun. Pode escolher um intervalo, não é? Então eu acho que vai ficar por aqui entre 60-70.

P: Porque tu acha isso?

A: Porque as primeiras amostras já caíram aqui. [...] Vamos ver se esse intervalo vai permanecer. [Escolhe mais 5 casos] Olha aí. Se eu puder escolher um intervalo, eu ficaria com esse entre 60-70 intervalo de 10.

Com uma amostra de $n = 25$, Ana conseguiu oferecer um intervalo como resposta à questão. A mesma só achou conveniente o uso de outras ferramentas do *TinkerPlots*, quando surgiu a preocupação em oferecer um intervalo mais preciso sobre o aglomerado dos dados ao inserirmos toda a população no *Plot*, como mostra a Figura 11 a seguir:

Figura 11: População do Banco de dados MysteryMixer 1.

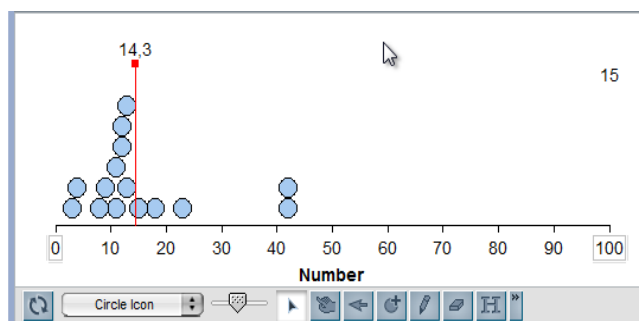


Fonte: Os autores.

Mesmo só se valendo da ferramenta Ref. no final da Atividade, Ana ofereceu um intervalo próximo a que os dados da população mostravam. Pode-se visualizar na Figura 11 que boa parte dos dados se concentra aproximadamente entre os níveis 50 a 70 e o intervalo oferecido pela professora foi de 60 a 70. Notamos com isso, que boa parte da atividade foi realizada sem a necessidade do uso de outras ferramentas de manipulação dos casos. Apenas o uso da ferramenta Plot, na Atividade 2, bem como a Atividade 1, pareceu ser relevante para a interpretação das amostras crescentes.

Os dados dispostos no banco de dados MysteryMixers 2, por sua vez, estavam aglomerados no início da escala, conforme mostra a Figura 12. Ao ser convidada a extrair uma amostra desse banco de dados, a professora demonstrou preocupar-se com o tamanho adequado da amostra para que a mesma fosse representativa. Justo por isso, a mesma selecionou uma amostra com $n = 10$ e logo na sequência, uma amostra com $n = 15$, oferecendo as seguintes ideias sobre a representação:

Figura 12: Ferramenta Ref. acionada na Amostra com $n = 15$ do banco de dados "MysteryMixer2 ."



Fonte: Os autores.

Ana: Eu acho que vai ficar entre 10-15.

Pesquisadora: E esses que apareceram aqui (Referindo-se aos casos que ocupam outros locais da escala). Você não acha que os demais plots não irão correr mais para cá, não?

A: Pode acontecer.

P: Quanto tu acha que deveria ter no tamanho da amostra para ter certeza?

A: 500.

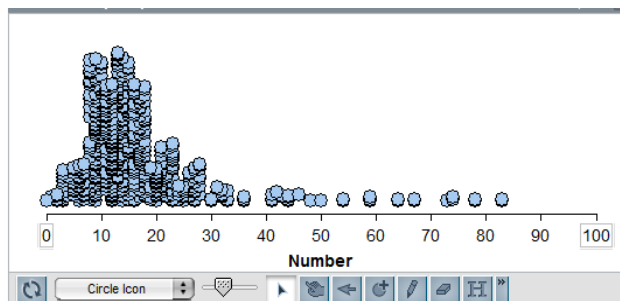
P: 500 é o tamanho total, não pode pegar tudo porque se não você vai ficar pobre demais.

A: Ai já deixava de ser uma pesquisa econômica, né? Eu acho que esse aí já é o suficiente.

Notamos que Ana conseguiu oferecer uma resposta a partir de uma amostra menor do que a extraída do banco de dados anterior. Isso pode ter relação com uma autonomia maior que a professora apresentou em relação à primeira atividade, bem como pelo concentrado do banco de dados “MysteryMixer 2” estar com uma tendência voltada para um dos extremos da escala. Observou-se que a ferramenta *Ref.* pôde auxiliar a professora a estipular um intervalo na escala para oferecer estimativas à população baseadas nas análises das amostras.

Após oferecer uma inferência final no segundo banco de dados, Ana observou o aglomerado da população e pôde confirmar suas hipóteses com base na observação do gráfico:

Figura 13: População com N = 500 do banco de dados “MysteryMixer2.”



Fonte: Os autores.

Ana: De 10-20, né? Então é aquela ideia mesmo, quanto menos você pegar, menos certeza você vai ter. No outro eu peguei mais, não foi?

Podemos observar que Ana estipulou um intervalo próximo ao aglomerado da população, que se apresentava em torno do número 10 e prosseguia para o número 20. Ela também considerou que o tamanho da amostra era pequeno para estabelecer conclusões mais precisas sobre a população após a visualização de todos os casos. Acreditamos que confrontar os resultados da população com as estimativas baseadas em amostras, pôde contribuir para que a professora refletisse sobre as decisões tomadas na escolha da amostra.



Considerações finais

No decorrer da Atividade 1, percebemos que a professora utilizou apenas três ferramentas para generalizar dados de uma amostra à uma população. Foram escolhidas as ferramentas *Average* que mostra a média e a ferramenta *Hot* que auxiliou a professora a identificar de maneira mais clara a tendência dos dados da amostra. Notamos, no entanto, que a possibilidade de visualização dos dados através da ferramenta *Plot* foi um elemento importante e que facilitou a observação das diferentes distribuições da amostra, tanto na primeira quanto na segunda atividade.

Observamos que em determinados momentos, a dificuldade de interpretação dos dados pela professora pôde interferir na identificação de amostras representativas na primeira atividade, mas que com o aumento das amostras, a professora pôde realizar melhores análises da tendência dos dados.

Na Atividade 2, utilizou-se poucas ferramentas do *TinkerPlots* se comparada a Atividade 1. Também nesta tarefa, não identificamos dificuldades em interpretar os gráficos, e tal fato pode estar relacionado com a situação problema só exigir a visualização de uma variável no gráfico.

De modo geral, os resultados desse estudo evidenciaram que uma compreensão inicial sobre representatividade e sua relação com o tamanho da amostra foi possível de ser construída em uma atividade de amostragem.

Sobre o tamanho das amostras, notamos que a professora pôde identificar aspectos de variação nos grupos dispostos na amostra nas duas atividades. Esse achado nos fez pensar que é possível explorar nos sujeitos a relação entre heterogeneidade/homogeneidade e representatividade em amostras em outros estudos. Observamos também que a professora conseguiu associar o custo hipotético de uma pesquisa por amostragem com o tamanho da amostra e sua representatividade.

Os dados sugerem que o contexto de amostras crescentes no *TinkerPlots* pôde facilitar a identificação de amostras representativas a partir da análise e comparação das amostras menores. Também concluímos que comparar os dados da população com as estimativas retiradas ao longo do processo de interpretação pode trazer boas reflexões sobre as decisões tomadas no procedimento de amostragem.

As análises dos protocolos evidenciaram que a professora mencionou por algumas vezes, que os dados eram escolhidos aleatoriamente. Contudo, os achados não oferecem evidências sobre como a professora concebia um fenômeno aleatório e por isso, não podemos concluir que a mesma construiu uma compreensão apropriada sobre a amostragem aleatória simples. Em vista disso, sugerimos que em estudos posteriores explorem-se atividades que envolvam princípios básicos de aleatoriedade com professores. Acreditamos que a partir da identificação do pensamento sobre aleatoriedade, se poderão obter entendimentos mais aprofundados sobre a compreensão relativa à amostragem aleatória.



Referências

ALVES, Iane Maria Pereira. **A interpretação de gráficos em um ambiente computacional por alunos de uma escola rural do município de Caruaru-PE.** 2011. 156f. Dissertação (Mestrado em Educação Matemática e Tecnológica) – Centro de Educação, UFPE, Recife, PE, 2011.

ASSEKER, Andreika. **O uso do TinkerPlots para exploração de dados por professores de escolas rurais.** 2011. 156f. Dissertação (Mestrado). Universidade Federal de Pernambuco, Recife, 2011.

BEN-ZVI, Dani. et al. Children's emergent inferential reasoning about samples in an inquiry based environment. In: CONGRESS OF THE EUROPEAN SOCIETY FOR RESEARCH IN MATHEMATICS EDUCATION, 7., 2011. Poland, **Anais eletrônicos...** Poland: Rzeszow, 2011.

BOLFARINE, Heleno; BUSSAB, Wilton. **Elementos de amostragem.** São Paulo: Blucher, 2005.

CARVALHO, Liliane; MONTEIRO, Carlos Eduardo. Reflexões sobre implementação e uso de laboratório de informática na escola pública. **Roteiro (UNOESC)**, v.37, p.343-360, 2012.

COUTO JUNIOR, Euro Barros. **Abordagem Não-Paramétrica para Cálculo do tamanho da Amostra com base em questionários ou escalas de avaliação na área da saúde.** 2009. 138f. Tese (Doutorado), Universidade de São Paulo, 2009.

EUGÊNIO, Robson. **Explorações sobre a média no TinkerPlots 2.0 por estudantes do ensino fundamental.** 2013. 231f. Dissertação (Mestrado em Educação Matemática e Tecnológica) – Centro de Educação, Universidade Federal de Pernambuco. Recife, 2013.

GUIMARÃES, Gilda; GITIRANA, Verônica; ROAZZI, Antônio. Interpretando e construindo gráficos. In: REUNIÃO ANUAL DA ASSOCIAÇÃO NACIONAL DE PÓS-GRADUAÇÃO E PESQUISA EM EDUCAÇÃO, 24, 2001, Caxambu. **Anais...**, Caxambu: ANPED, 2001. p. 01-19 Disponível em: <http://www.anped.org.br/reunioes/24/tp1.htm#qt19>. Acesso em 22 de Jul. 2012.

GOMES, Tamara. **O todo é a soma das partes, mas uma parte representa o todo?:** compreensão de estudantes do 5º e 9º ano sobre amostragem. 2013. 109f. Dissertação (Mestrado). Universidade Federal de Pernambuco. Recife, 2013.

KAZAK, Sibel; KONOLD, Cliff. Development of ideas in data and chance through the use of tools provided by computer-based technology. In: INTERNATIONAL ASSOCIATION OF STATISTICAL EDUCATION (IASE). ICOTS, 8., 2010.



Slovenia, **Anais eletrônicos...** Slovenia: International Statistical Institute, 2010.

Disponível em: <

http://www.stat.auckland.ac.nz/~iase/publications/icots8/ICOTS8_8D2_KAZAK.pdf

>. Acesso em: 9 jun. 2012.

KONOLD, Cliff. **Handling complexity in the design of educational software tools.** 2006. Disponível em:

<http://www.start.auckland.ac.nz/~iase/publications/17/2D-KONO.pdf>> Acesso em: 27 set. 2009.

KONOLD, Cliff; MILLER, Craig. **TinkerPlots: dynamic data exploration.**

TinkerPlots Help Version 2.0. Key Curriculum Press. 2012. Disponível em <

http://www.srri.umass.edu/sites/srri/files/tinkerplots2help/index.html?device_option_s.htm> Acesso em 31 ago. 2012.

LIRA, Olga. **Uso de Ferramentas do Software TinkerPlots para interpretação de dados.** 2010. 195f. Dissertação (Mestrado). Universidade Federal de Pernambuco. Recife, 2010.

LIRA, Olga; MONTEIRO, Carlos Eduardo. Interpretação de dados a partir da utilização de ferramentas do Software TinkerPlots. **Bolema.** Boletim de Educação Matemática (UNESP. Rio Claro. Impresso), 2011.

LIU, Yan; THOMPSON, Patrick. Teachers' understandings of probability. **Cognition and Instruction**, London, v. 2, n. 25, p.113-160, fev. 2007.

OLIVEIRA, Ely de; GRÁCIO, Maria. **Análise a respeito do tamanho de amostras aleatórias simples:** uma aplicação na área de ciência da informação. *DataGramaZero*. Rio de Janeiro, v.6, n.3, jun 2005. Disponível em: < http://www.dgzero.org/ago05/Art_01.htm>. Acesso em: 13 Nov. 2012.

PADILHA, Maria Auxiliadora. **A Formação de professores e as tecnologias da informação e comunicação:** Uma relação Possível e necessária. In: III Congresso Ibero-Americano de política e Administração da Educação, 2012, Zaragoza. Desafios para uma melhor formação e profissionalização dos educadores, 2012.

PFANNKUCH, Maxine. Building sampling concepts for statistical inference: A case study. In: INTERNATIONAL CONGRESS ON MATHEMATICS EDUCATION - ICME, 11., 2008, Mexico. **Anais eletrônicos...** Mexico: Proceedings, 2008. Disponível em: <<http://tsg.icme11.org/document/get/476>>. Acesso em: 12 jun. 2012.

PRODROMOU, Theodosia 2011. Students' emerging inferential reasoning about samples and sampling. In: ANNUAL CONFERENCE OF THE MATHEMATICS. EDUCATION RESEARCH GROUP OF AUSTRALIA, 34., 2011, Alice Springs,



Australia. **Anais eletrônicos...** Australia: Proceedings, 2011. Disponível em:
<http://www.merga.net.au/documents/RP_PRODROMOU_MERGA34-AAMT.pdf>.
Acesso em: 10 Set. 2013.

WATSON, J. M. Developing reasoning about samples. In: BEN-ZVI, D.; GARFIELD, J. (Org.). **The challenge of developing statistical literacy, reasoning and thinking**. Dordrecht, The Netherlands: Kluwer Academic Publishers, 2004.

Enviado em: 01/02/2013 Aceito em: 02/11/2013
