

2

Origem dos dados: Estudo longitudinal da geração escolar 2005 – GERES

2.1

Definição e metodologia

Nesta seção, num primeiro momento, apresento os princípios que norteiam a pesquisa GERES. Nesse sentido, discuto as vantagens do uso de dados longitudinais e da medida de “valor agregado” para estudos educacionais. A seguir, detalho os principais procedimentos de pesquisa, esclarecendo sobre a construção e aplicação dos testes e sobre a metodologia da teoria da resposta ao Item utilizada para o cálculo da proficiência. Num segundo momento, apresento os princípios norteadores da matriz de referência de leitura que orienta a elaboração dos testes utilizados na pesquisa.

O GERES é uma “pesquisa longitudinal de painel”. Sua aplicação teve início em 2005 e acompanhou, por quatro anos, a evolução da aprendizagem de Leitura e de Matemática de alunos de escolas públicas e privadas das séries iniciais do Ensino Fundamental.

A pesquisa, financiada pela Fundação Ford e pelo PRONEX, foi aplicada em cinco grandes cidades brasileiras: Belo Horizonte (MG), Campinas (SP), Campo Grande (MS), Rio de Janeiro (RJ) e Salvador (BA).

A coordenação e execução do Projeto GERES contou com a associação de seis centros universitários: o Laboratório de Avaliação da Educação da PUC – Rio; o Grupo de Avaliação e Medidas Educacionais da UFMG; o Laboratório de Observação e Estudos Descritivos, da UNICAMP; o Centro de Estudos Interdisciplinares para o Setor Público, da UFBA, o Centro de Avaliação da Educação da UFJF e a Universidade Estadual do Mato Grosso do Sul (UEMS).

O objetivo central da pesquisa é investigar quais práticas educativas e quais condições escolares contribuem para a promoção da eficácia escolar e da equidade intraescolar. Para atingir tal propósito, o GERES pretende identificar, especificamente, características escolares que maximizam a aprendizagem dos alunos, minimizam o impacto da origem social sobre a aprendizagem, diminuem a probabilidade de repetência e absenteísmo e maximizam a autoestima e a motivação para estudar.

Os sistemas de avaliação produzidos sobre a educação brasileira nos últimos anos têm sido extremamente úteis, porque criam condições para pesquisas sobre a Eficácia Escolar. O SAEB, Sistema de Avaliação da Educação Básica, tem coletado medidas cognitivas e contextuais que concretizam um campo fértil para a pesquisa educacional e vêm fomentando o debate sobre características de uma boa escola. Porém, esses dados apresentam limites consideráveis, dependendo dos objetivos das pesquisas. A coleta de dados do SAEB obtém medidas do rendimento de certos grupos de alunos em determinado momento, ou seja, cada aluno é avaliado uma única vez. Esse tipo de avaliação mede o desempenho dos alunos, mas não é apropriado para medir o seu aprendizado e as condições específicas em que esse aprendizado ocorreu, uma vez que não permite medir o valor agregado pelo estabelecimento de ensino.

Nesse sentido, alguns pesquisadores (Lee, 2004; Franco, 2004a) passaram a apontar para a necessidade de se transitar do exercício da avaliação para sistemas de avaliação, por meio de pesquisas longitudinais.

Partindo dessa premissa, o GERES adota a modelagem longitudinal e, conseqüentemente, seu aparato metodológico. Essa característica da pesquisa atende ao propósito de se investigar as alterações no desempenho dos alunos acontecidas entre um momento e outro da escolarização, ou seja, o “valor agregado” pelo estabelecimento escolar e, ainda, de se identificar as características das escolas e salas de aulas promotoras desse progresso. Essa decisão metodológica é muito importante, uma vez que permite controlar os resultados dos estudantes ao final de um determinado período pelos seus próprios resultados em teste prévio e aferir o valor agregado nesse período determinado, sob condições específicas de ensino.

Ainda sobre a metodologia, a partir da segunda etapa de observação, o GERES passou a contar com estudos que incluem análises qualitativas, investindo na colaboração entre diferentes metodologias. A necessidade de estudos qualitativos está relacionada à compreensão de que variáveis e medidas utilizadas por pesquisas de survey contêm, tipicamente, um viés conservador, decorrente da lógica de observação e mensuração que lhe é própria. Estudos qualitativos colaboram para a diminuição desse viés, porque permitem melhor adequação dos instrumentos de medida e das análises quantitativas à realidade que se pretende conhecer (cf. Bonamino (2004). Desta forma, se, por um lado, a pesquisa

qualitativa pode se valer de evidências estatísticas para informar suas análises, por outro lado, sua forma específica de abordagem permite ultrapassar limites inerentes à metodologia estatística, contribuindo para um conhecimento mais complexo sobre os dados coletados nos estabelecimentos escolares.

Para a coleta dos dados de natureza contextual, o GERES, além de utilizar informações relevantes já disponíveis na escola, fez uso de questionários aplicados aos pais dos alunos, aos alunos, aos professores e aos diretores. O questionário dos professores, particularmente importante para a minha tese, procura identificar diferentes perfis dos profissionais e delinear sua prática em sala de aula, buscando características sobre exigências acadêmicas, modos de organização da sala de aula, uso do tempo, estilos e recursos pedagógicos e práticas de avaliação.

Para coletar os dados de caráter cognitivo, o GERES utilizou testes de Leitura e de Matemática, focalizando habilidades básicas tipicamente demandadas pela escola a alunos das séries iniciais, tendo como orientação uma matriz de referência.

Profissionais da área de Língua Portuguesa e de Matemática planejaram um conjunto de itens para serem pré-testados. Os testes de proficiência em leitura foram elaborados a partir de uma média de 140 itens de múltipla escolha. A pré-testagem foi realizada com alunos de escolas das redes pública e particular do Rio de Janeiro e de Juiz de Fora. Para a análise dos itens do pré-teste foram utilizadas as estatísticas clássicas, representadas pelo percentual de acerto (parâmetro da dificuldade) e a Teoria da Resposta ao Item – TRI, detalhada adiante, com o objetivo de selecionar os itens que apresentassem maior qualidade técnica e pedagógica. Dentre os itens mais adequados foram escolhidos os que compuseram as versões dos testes aplicadas aos alunos.

Para cada aplicação houve duas versões de teste mescladas com itens de diferentes graus de dificuldade. Um dos testes foi composto de itens considerados mais fáceis e de itens intermediários e o outro foi composto com os itens intermediários do anterior acrescido de itens considerados mais difíceis. A cada aplicação foram retirados os itens anteriores mais fáceis e acrescentados itens mais difíceis, sendo mantidos, sempre, itens em comum entre os diferentes testes. Essa metodologia, que mantém itens em comum entre os testes, permite o ajuste de uma escala única que possibilita comparar os níveis de proficiência de alunos que participaram de diferentes versões de teste (mais fácil e mais difícil).

Para o cálculo da proficiência e a análise dos resultados dos testes, o GERES adota o modelo logístico da Teoria de Resposta ao Item Paramétrica⁹. Essa escolha metodológica deve-se ao propósito de utilização de um modelo que relaciona a habilidade de um indivíduo com a probabilidade que ele tem de acertar a resposta ao item. Tal modelo associa três parâmetros que consideram a discriminação do item, o grau de dificuldade e a probabilidade de acerto ao acaso.

A partir das informações fornecidas pelos itens e para determinar o quanto, em média, as habilidades dos alunos estão desenvolvidas, são consideradas três etapas principais do processo: o início do desenvolvimento da habilidade, o pleno desenvolvimento da habilidade e a consolidação da habilidade, conforme pode ser observado a seguir:

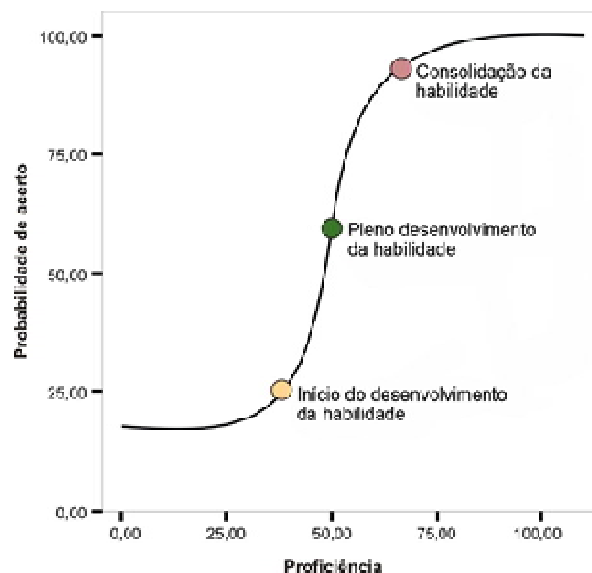


Figura 1: Curva característica de um item com os principais pontos do desenvolvimento de habilidades

A utilização do parâmetro da dificuldade do item nos possibilita inferir que, provavelmente, alunos com proficiência muito abaixo de desse parâmetro ainda não desenvolveram a habilidade específica demandada pelo item; que alunos com proficiência próxima ao ponto que marca a dificuldade do item estão em pleno desenvolvimento da referida habilidade; e que os alunos com proficiência muito acima de tal parâmetro consolidaram a habilidade exigida pelo item.

⁹Algumas referências do Projeto GERES para a TRI são: Baker F., Kim (2004); Thissen D., Wainer H. (2001) *apud* Tufi M. Soares (2009).

A TRI permite que, após a aplicação de cada teste, as proficiências de vários anos sejam reestimadas e equalizadas, de modo que seja possível obter uma curva de crescimento da proficiência ao longo do tempo para cada aluno observado.

2.2 Matriz de referência e concepção de leitura

Todo trabalho desenvolvido no âmbito pedagógico assume uma concepção sobre como se adquire conhecimento. Os testes utilizados nas pesquisas para aferir o que os alunos aprenderam não estão de fora desse pressuposto. Neste sentido, esse tópico situa o leitor quanto à matriz de referência que informa a concepção de “aprender a ler” nos testes GERES. Por sinal, o próprio instrumento, que é construído para orientar a elaboração dos itens do teste, é denominado “matriz de referência”. A matriz de referência é uma etapa muito importante do processo de pesquisa porque é ela que define e delimita o que se pretende investigar, com o objetivo de gerar instrumentos adequados e confiáveis para medidas educacionais.

Uma matriz é composta de um conjunto de descritores¹⁰ que são elaborados de acordo com os conteúdos programáticos relevantes e as habilidades esperadas para determinada faixa etária/ano de escolaridade. Os descritores devem oportunizar também a elaboração de itens que captem desde habilidades consideradas fáceis como habilidades consideradas difíceis para alunos do ano escolar investigado. Em outras palavras, devem oportunizar a medida de conhecimentos que os alunos construíram fora da escola ou em anos escolares anteriores e habilidades que exigem conhecimentos mais complexos do que os normalmente esperados para determinado grupo. Esse cuidado possibilita aferir níveis de conhecimentos para todos os alunos testados.

A matriz de referência para a elaboração dos itens dos testes de leitura do Projeto GERES foi construída pelos especialistas Antônio Augusto Gomes Batista, Carla Viana Coscarelli e Magda Becker Soares (2004) e se orienta “por uma concepção ampla de letramento, de acordo com a qual o termo designa, em

¹⁰ Os descritores previstos para os testes do 2º ano e o teste do 3º ano encontram-se no anexo I. Para informações detalhadas sobre a matriz ver referências e/ou anexo I.

sua heterogeneidade e variação, o domínio de habilidades de uso da língua escrita e seu uso efetivo em práticas sociais, para a consecução de diferentes objetivos de natureza individual e social.” Os autores afirmam que, em função das particularidades da pesquisa, realizada em grande escala e com população em processo de aprendizagem da língua escrita, as capacidades de leitura foram privilegiadas em relação às de escrita e o conjunto das capacidades foi organizado em torno de quatro dimensões para avaliação dos estudantes. Tais dimensões têm em vista a realização de tarefas ligadas ao domínio (i) da tecnologia da escrita, especificamente do sistema de escrita e de seus suportes e de procedimentos de leitura responsáveis: (ii) pela recuperação de informações, (iii) pelo estabelecimento de relações entre as informações e (iv) pela avaliação e posicionamento em relação a informações.

Ressalto que, de acordo com a distinção proposta por Magda Soares (2004), as capacidades relacionadas ao domínio da tecnologia de escrita designam apropriadamente a alfabetização, ou seja, o aprendizado do sistema de escrita alfabético-ortográfico e o desenvolvimento de capacidades cognitivas e motoras envolvidas no uso e na manipulação de instrumentos e equipamento de escrita. Já as capacidades relacionadas à recuperação de informações, ao estabelecimento de relações entre informações, à avaliação e ao posicionamento designam as habilidades de uso do sistema de escrita e de seus instrumentos e equipamentos em práticas sociais, isto é, na compreensão de textos em variadas situações sociais.

Ainda segundo os autores, o conjunto das capacidades mencionadas compõe quatro escalas para a medição do letramento que se referem a diferentes esferas sociais de produção e recepção de textos e aos gêneros e tipos de textos associados a essas esferas. Os textos estão relacionados à esfera escolar, literária, jornalística, da publicidade, da vida doméstica, da vida pública e do espaço urbano.

Foram determinadas quatro grandes funções da leitura referentes às quais os estudantes são avaliados em suas capacidades de uso autônomo: informar-se, organizar-se, realizar comportamentos solicitados, aprender, avaliar informações.