

# Modelagem do crescimento da aprendizagem nos anos iniciais com dados longitudinais da pesquisa GERES<sup>I</sup>

Nigel Brooke<sup>II</sup>

Neimar da Silva Fernandes<sup>III</sup>

Isabela Pagani Heringer de Miranda<sup>III</sup>

Tufi Machado Soares<sup>III</sup>

## Resumo

Este artigo compara duas abordagens de valor agregado para dados oriundos do *survey* educacional de recorte longitudinal, chamado GERES – Estudo Longitudinal da Geração Escolar 2005, que acompanhou uma coorte de alunos de mais de 300 escolas públicas e privadas ao longo dos primeiros quatro anos do Ensino Fundamental. Ambas as abordagens utilizam modelos lineares hierárquicos, permitindo o agrupamento natural dos dados educacionais provenientes dos três níveis: aluno, turma e escola. Na primeira abordagem de valor agregado, constroem-se modelos cuja variável dependente é a proficiência do aluno em cada ano avaliado. Com um modelo distinto para cada ano é possível detectar fatores do aluno, da turma e da escola associados ao desempenho dos alunos. A segunda abordagem cria modelos para mostrar o efeito das covariáveis de aluno, turma e escola nas curvas de evolução da proficiência ao longo do período do estudo. Quando comparados os dois tipos de modelos de valor agregado, o primeiro foi o mais eficiente em diagnosticar os efeitos do ambiente e da prática pedagógica do professor, mas somente em determinados anos. Já o segundo tipo de modelo foi capaz de identificar curvas de evolução de proficiência de formatos distintos de acordo com determinadas características das escolas e dos alunos, mas foi menos sensível na identificação de variáveis associadas ao processo de formação de grupos e à prática pedagógica do professor. Os dois tipos de modelos de valor agregado oferecem indicações de processos de aprendizagem diferenciados para as disciplinas Língua Portuguesa e Matemática que mereceriam estudos adicionais.

**I-** Os autores agradecem ao Centro de Políticas Públicas e Avaliação da Educação, CAEd/UFJF e ao Programa Observatório da Educação – CAPES/INEP

**II-** Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brasil.  
Contato: n.brooke@terra.com.br

**III-** Universidade Federal de Juiz de Fora, Juiz de Fora, MG, Brasil.  
Contato: neimar@caed.ufjf.br;  
isabelaaphm@yahoo.com.br;  
tufi@caed.ufjf.br;

## Palavras-chave

Estudo longitudinal – Avaliação educacional – Valor agregado – GERES – Modelos lineares hierárquicos.

# **Modeling of the growth of learning in the early years with longitudinal data of GERES research<sup>1</sup>**

Nigel Brooke<sup>I</sup>

Neimar da Silva Fernandes<sup>III</sup>

Isabela Pagani Heringer de Miranda<sup>III</sup>

Tufi Machado Soares<sup>III</sup>

## **Abstract**

*This article compares two value-added approaches to data from the longitudinal education survey called Estudo Longitudinal da Geração Escolar 2005 (GERES – Longitudinal Study of 2005 School Generation), which followed a cohort of students from more than 300 public and private schools over the first four years of primary education. Both approaches use hierarchical linear models, allowing the natural grouping of educational data from three levels: student, class and school. The first value-added approach builds models whose dependent variable is the student proficiency in each year evaluated. With a separate model for each year, it is possible to detect factors of the student, class and school associated with student performance. The second approach creates models to show the effect of the covariates of student, class and school on progress curves of proficiency throughout the study period. When comparing the two types of value-added models, the first one was the most efficient in diagnosing the effects of the environment and the teacher's pedagogical practice, but only in certain grades. The second type of model was able to identify progress curves of proficiency of different formats according to certain characteristics of schools and students, but was less sensitive to identify variables associated with the group formation process and the teacher's pedagogical practice. The two types of value-added models offer indications of differentiated learning processes for the disciplines of Portuguese and Mathematics, which deserve further study.*

**I-** The authors thank Centro de Políticas Públicas e Avaliação da Educação (CAEd/UFJF – Center for Public Policy and Evaluation of Education, CAEd / UFJF) and Programa Observatório da Educação (Observatory of Education Program) – CAPES/INEP

**II-** Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil.

Contact: [n.brooke@terra.com.br](mailto:n.brooke@terra.com.br)

**III-** Universidade Federal de Juiz de Fora, Juiz de Fora, MG, Brazil.

Contacts: [neimar@caed.ufjf.br](mailto:neimar@caed.ufjf.br);

[isabelaaphm@yahoo.com.br](mailto:isabelaaphm@yahoo.com.br);

[tufi@caed.ufjf.br](mailto:tufi@caed.ufjf.br);

## **Keywords**

*Longitudinal study – Educational assessment – Value added – GERES – Hierarchical linear models.*

## Introdução

A partir de estudos que correlacionam o desempenho dos alunos com aspectos materiais e organizacionais da escola e com características técnicas e humanas da equipe escolar, esperam-se conclusões acerca das razões que levam algumas escolas a resultados melhores e, como corolário, diferentes opções para a melhoria na qualidade do ensino. É com este propósito em mente que, ao mesmo tempo em que se fazem testes para medir o desempenho aos alunos, aplicam-se questionários aos professores e gestores para caracterizar o ambiente em que a aprendizagem se desenvolve. Tanto a Prova Brasil quanto todos os sistemas estaduais de avaliação que seguem o modelo pioneiro do Sistema de Avaliação da Educação Básica (SAEB) implantaram esse mesmo padrão de avaliação transversal para a coleta simultânea de informações a respeito do desempenho dos alunos e das condições da escola (BONAMINO, 2002).

O problema desse modelo reside no fato da aprendizagem ser um processo cumulativo, construído ao longo da trajetória educacional do aluno, e fruto de diversas influências, entre as quais todos os professores do aluno desde seu primeiro ano escolar. Ou seja, enquanto os testes que medem o desempenho do aluno estão sondando um agregado de aprendizagem de muitos anos, as informações coletadas sobre as condições escolares são específicas do ano da coleta de dados. Essa falta de sintonia fragiliza as análises e dificulta a formulação de políticas de qualidade e equidade mais sólidas (FRANCO, 2001). Pela falta de conexão com as origens da aprendizagem, provocada pelas incertezas sobre o ponto de partida dos alunos e das contribuições específicas do ambiente de aprendizagem, os pesquisadores só se permitem falar de “fatores associados” e, raramente, comprometem-se a indicar causas e efeitos.

Diversos autores mostram as dificuldades da utilização de dados transversais para investigar a relação entre fatores escolares e desempenho acadêmico. Além do problema em

atribuir causalidade (RAUDENBUSH; FOTIU; CHEONG, 1998), a literatura enfatiza a ausência de medidas anteriores do desempenho dos alunos para calcular o ganho atribuível aos fatores escolares (GOLDSTEIN, 1995) e a impossibilidade de fazer inferências confiáveis sobre a eficácia das escolas com base em uma única medida do desempenho (GOLDSTEIN et al. 2000). Essas dificuldades levam vários autores a especificar certos requisitos mínimos para o estudo dos fatores escolares, entre os quais se destaca a necessidade da coleta repetida de dados em um desenho de pesquisa longitudinal. Com medidas de aprendizagem aplicadas em momentos ou anos distintos é possível controlar as variáveis relativas à influência dos antecedentes socioeconômicos e educacionais dos alunos, de modo a extrair conclusões sobre os processos internos das escolas e sobre a qualidade do ensino oferecido. Assim, pode-se trabalhar não com o nível absoluto de proficiência alcançado pelo aluno, mas com a medida de seu avanço ou da aprendizagem nova adquirida em cada período e, a partir dela, calcular a contribuição da escola ou do próprio professor.

Foi com base nesse raciocínio que foi realizada a pesquisa GERES – Estudo Longitudinal da Geração Escolar 2005, cuja metodologia e resultados preliminares já foram relatados em outras publicações (FRANCO; ALVES; BROOKE, 2008; BROOKE; BONAMINO, 2011). Desenhada para produzir informação a respeito das trajetórias de aprendizagem em Língua Portuguesa e Matemática de uma coorte de alunos desde o princípio da então 1ª série até o final da 4ª série, a pesquisa acompanhou mais de 20.000 alunos de escolas públicas e privadas em cinco cidades do Nordeste, Sudeste e Centro-Oeste. Sendo a primeira investigação longitudinal a ser completada com sucesso no Brasil, a pesquisa GERES permite refletir a respeito dos fatores intra e extraescolares que incidem no processo de aprendizagem com base não somente no nível de desempenho alcançado pelos alunos ao final dos anos iniciais, mas também nos ganhos de

aprendizagem diferenciados ano a ano ao longo de um período de quatro anos.

## **A pesquisa GERES**

A pesquisa GERES adotou um desenho longitudinal de painel: o mesmo conjunto de escolas e estudantes foi observado ao longo de quatro anos. A amostra foi composta por alunos de 303 escolas dos municípios de Belo Horizonte (MG), Rio de Janeiro (RJ), Salvador (BA), Campo Grande (MS) e Campinas (SP) que, no início de 2005, estavam matriculados na 1ª série do ensino fundamental (ou seu equivalente, quando a organização do ensino era em ciclos, ou na 2ª série se os alunos não tiveram nenhuma experiência prévia de alfabetização). A primeira aplicação (onda) de medidas de aprendizagem foi realizada em março desse mesmo ano, para gerar uma linha de base para os alunos, já a segunda, ocorreu em novembro. As seguintes ondas de aplicação de instrumentos aconteceram ao final dos anos de 2006, 2007 e 2008, viabilizando o acompanhamento da amostra ao longo de quatro anos letivos. O estudo buscou fazer a distinção entre “valor agregado pela escola” e o efeito da “seleção” devido à eventual reprovação de alunos de desempenho menor. Em termos práticos, isso significou que aqueles que não foram aprovados continuaram sendo observados pela pesquisa desde que permanecessem retidos na mesma escola ou se transferissem para outra escola da amostra.

Cada uma das cidades foi considerada como um estrato e, dentro de cada cidade, foi selecionada uma amostra probabilística complexa de escolas, turmas e alunos a partir do cadastro do Censo Escolar de 2003, excluídas as escolas que não possuíam 10 alunos ou mais matriculados na 1ª série do ensino fundamental. Testes adequados aos anos iniciais do ensino fundamental foram elaborados por especialistas de três das seis universidades participantes da pesquisa (Universidade Federal de Minas Gerais – UFMG, Pontifícia Universidade Católica do Rio de Janeiro – PUC-Rio e Universidade Federal

de Juiz de Fora), com base em matrizes de habilidades de Leitura e Matemática. As outras universidades participantes foram Universidade Estadual de Campinas – UNICAMP, Universidade Estadual de Mato Grosso do Sul e Universidade Federal da Bahia – UFBA.

As questões foram previamente testadas em escolas públicas e privadas das cidades de Juiz de Fora e Rio de Janeiro. Em cada onda, todos os alunos presentes foram submetidos a testes de Leitura e de Matemática, elaborados em duas versões: uma mais fácil e outra mais difícil. As diferentes versões possuíam, no entanto, itens comuns, de modo a viabilizar escores equalizados a partir da Teoria de Resposta ao Item (TRI). Além dos testes, foram aplicados questionários contextuais aos diretores, professores e famílias, para levantar informações a respeito da escola e sua organização, da prática pedagógica dos professores e o nível socioeconômico dos alunos. Essas informações fazem parte da análise dos fatores relevantes na explicação das diferenças nos ganhos de aprendizagem entre alunos e escolas.

Em consonância com as matrizes de Leitura e Matemática foram desenvolvidas duas escalas para descrever o desempenho dos alunos em termos das habilidades em fase de consolidação e as habilidades já consolidadas. Essas escalas foram indispensáveis para a comunicação dos resultados às escolas e para a interpretação pedagógica do desempenho das turmas. Os cálculos das proficiências em todas as ondas foram feitos pela mesma equipe, a partir de uma mesma metodologia, o que garante a confiabilidade dos resultados cognitivos obtidos.

## **Modelagem do crescimento da aprendizagem na pesquisa GERES**

O presente trabalho tem como objetivo a observação de associações entre diversos fatores contextuais em diferentes níveis – alunos, turmas e escolas – e ganhos de proficiência dos alunos GERES em Matemática e Língua Portuguesa mensurados a partir de testes aplicados aos mesmos alunos em ondas sucessivas de testagem,

segundo o modelo longitudinal já descrito. Isso possibilitou controlar a proficiência final pela inicial, além da realização de análises os diversos padrões de evolução do aprendizado ao longo do tempo de acordo com variáveis, tais como a rede de ensino (público e privado), grupo de cor e nível socioeconômico.

O presente artigo inicia-se com uma descrição dos modelos lineares hierárquicos utilizados nas análises aqui realizadas, acompanhada de uma breve explicação sobre a lógica que justifica a sua utilização. A seguir, discorre a respeito da análise dos modelos agregados onda a onda e das covariáveis neles utilizados. Essas covariáveis são, então, introduzidas nos modelos onda a onda, tanto no caso da disciplina de Matemática quanto de Língua Portuguesa. Após uma análise desses

resultados, são introduzidos os modelos de evolução. Também se apresentam os detalhes da construção da base para a análise multinível, seguidos dos resultados obtidos com os modelos nulos. Finalmente, são apresentadas as conclusões do trabalho e uma discussão acerca dos principais pontos levantados no estudo.

O projeto GERES mede a proficiência dos alunos, a nossa variável dependente, em Língua Portuguesa e Matemática através de escalas próprias. No caso da Língua Portuguesa, a amplitude da escala, em termos dos resultados alcançados pelos alunos, vai do mínimo de 48 pontos na primeira onda ao máximo de 223 na quinta onda. A escala de Matemática, por sua vez, vai do mínimo de 20 pontos ao máximo de 397. A média e os desvios padrão para cada onda também são apresentados na Tabela 1:

**Tabela 1** – Média e desvio padrão das proficiências em Língua Portuguesa e Matemática por onda na pesquisa GERES

Disciplina	Onda	Média	Desvio Padrão	Mínimo	Máximo
Língua Portuguesa	1 <sup>a</sup>	108,42	26,28	47,80	159,90
	2 <sup>a</sup>	127,16	22,91	68,63	183,93
	3 <sup>a</sup>	143,25	27,54	85,68	214,52
	4 <sup>a</sup>	157,23	27,06	87,78	214,24
	5 <sup>a</sup>	169,05	25,37	89,45	222,96
Matemática	1 <sup>a</sup>	109,14	28,82	19,52	173,05
	2 <sup>a</sup>	138,22	32,22	54,73	217,11
	3 <sup>a</sup>	161,01	57,39	39,10	290,35
	4 <sup>a</sup>	202,06	63,24	45,49	356,03
	5 <sup>a</sup>	246,84	66,87	57,20	396,55

Fonte: dados da pesquisa.

Foram usadas como variáveis de alunos as seguintes categorias: *gênero; cor/raça; defasagem idade/série e a condição socioeconômica* NSE, medida a partir de um questionário de posse de bens análogo aos aplicados pelo IBGE. Para maiores detalhes da construção dessa medida ver Brooke e Bonamino (2011).

As variáveis do professor foram construídas utilizando-se as respostas dadas aos questionários aplicados aos docentes, utilizando-se uma metodologia descrita em Brooke e Bonamino (2011). Essas variáveis serão apresentadas a seguir:

a) *Percepção do professor sobre os obstáculos que impedem o melhoramento da*

*escola* – refere-se ao quanto os professores percebem os obstáculos que venham a impedir a melhora da escola, tais como depredações das dependências das escolas, bem como o consumo e tráfico de drogas, furto de equipamentos didáticos e/ou pedagógicos, violência e intimidação aos professores, funcionários e alunos.

b) *Frequência de uso de certos recursos pedagógicos disponíveis na escola* – corresponde a alguns dos recursos pedagógicos usados como mapas geográficos e materiais concretos de matemática (material dourado, tangram, etc).

c) *Percepção do professor sobre a frequência de interrupção da aula* – sinaliza o nível das interrupções da aula. Caso seja em

maior nível, há um maior grau de interrupção das aulas por indisciplina e imp pontualidade dos alunos, barulho e anúncios de direção.

d) *Frequência com que o professor realiza determinadas práticas nas suas aulas de Língua Portuguesa* – está associada às práticas positivas de leitura em aulas de Língua Portuguesa adotadas pelos professores, tais como leitura silenciosa ou em voz alta, por professores e/ou alunos, de histórias ou do livro didático. Também incluímos a produção de textos, bem como cópias de textos e ditados.

e) *Frequência com que o professor realiza determinadas práticas nas suas aulas de Matemática* – mede o uso pelos professores de práticas positivas de diferentes níveis de complexidade, desde a aplicação direta de conceitos até a representação de situações em linguagem matemática. Inclui a prática da matemática ligada ao cotidiano, bem como situações de problemas ligados a jogos e quebra-cabeça, o compartilhamento com os colegas das estratégias para solução de problemas e a prática de exercícios visando à memorização e ao aumento da velocidade de cálculo.

f) *Frequência de leitura do professor* – abrange a prática dos professores de leitura de textos, livros e também revistas especializadas da área de educação e literatura em geral.

g) *Frequência de participação em atividades culturais* – refere-se à participação dos alunos em atividades culturais, como visitas a livrarias e museus, bem como a frequência a espetáculos de teatro, música e dança.

h) *Frequência com que o professor realiza procedimentos pedagógicos promotores de motivação intrínseca* – está associada ao uso de procedimentos pedagógicos que promovem a motivação intrínseca dos alunos. Esses procedimentos procuram valorizar a autoestima mediante a adoção de atividades práticas ou experiências que despertem a curiosidade, estimulem a socialização e tragam novidades.

i) *Frequência com que o professor realiza procedimentos pedagógicos promotores de motivação extrínseca com regulação introjada*

– é a frequência com que os docentes aplicam procedimentos pedagógicos que promovam a motivação extrínseca, com regulação interna, como a valorização do desempenho do aluno perante a turma, escola e família.

j) *Frequência com que o professor realiza procedimentos pedagógicos promotores de motivação extrínseca com regulação externa* – refere-se à frequência com que os professores aplicam procedimentos pedagógicos que promovam a motivação extrínseca, com regulação externa, tais como argumentos de ascensão social ou sucesso financeiro, bem como prêmios por trabalhos de aula ou de casa e resultados de provas, a fim de estimular o estudo.

Na presente pesquisa é possível, em cada onda, localizar o aluno em sua turma. Utilizando os questionários aplicados aos professores, é possível associar o professor à sua turma em cada onda. Com essa estrutura também é possível calcular o percentual de meninos na turma, a média do NSE da turma, a média da defasagem idade/série da turma, dentre outras variáveis. Também a cada turma são atribuídas as características de seus professores. Essas variáveis compõem o nível de turma.

As variáveis no nível da escola são obtidas pelo cálculo de médias e de percentuais de características específicas de seus alunos. Dessa forma, obtêm-se variáveis tais como o “Percentual de Meninos na Escola”, a “Defasagem Média da Escola”, entre outras. Também variáveis do nível do professor podem ser alocadas no nível de escola pelo cálculo de médias e percentuais como, por exemplo, a “Prática de Leitura na Escola”.

As variáveis utilizadas podem ser classificadas em quatro categorias: 1) dicotômicas; 2) percentuais; 3) padronizadas; e 4) não padronizadas. As variáveis dicotômicas assumem o valor 1 quando o elemento pertence à categoria indicada, por exemplo, o gênero masculino. Assume o valor 1 quando se trata de um aluno e 0 quando de uma aluna. Nos modelos, o valor dos coeficientes corresponde à perda/ganho em pontos das escalas GERES de Língua Portuguesa e Matemática caso o aluno pertença àquela categoria.

Variáveis em percentuais assumem valores entre 0 e 1. Quando igual a 1, representa 100% do total. Nos modelos, o coeficiente indica a perda/ganho caso assumira valor 1. Caso assumira valor 0,5 (50%) na unidade analisada a perda/ganho seria a metade desse coeficiente, e assim por diante.

Variáveis padronizadas são as contínuas que podem, a princípio, assumir qualquer valor, mas sempre possuem média igual a 0 e desvio padrão igual a 1. Exemplos são a variável NSE e todas as variáveis oriundas do questionário do professor. Isso significa que nas tabelas o número de pontos corresponde à perda/ganho a cada desvio padrão.

Variáveis não padronizadas são as contínuas utilizadas nos modelos de forma direta, ou seja, sem nenhuma transformação, como a defasagem idade/série e a proficiência prévia. A interpretação de seus coeficientes no modelo se dá de forma direta.

No Quadro 1 são apresentadas todas as variáveis preditoras estudadas para permitir visualizar sua forma de mensuração bem como o nível de análise em que a variável se encontra. Ou seja, a variável média do NSE no nível de turma refere-se à média do NSE da turma.

**Quadro1 –** Variáveis incluídas nos modelos, por nível e tipo de medida

Nível	Variáveis	Tipo de Medidas
Aluno	Proficiência prévia	Não Padronizada
	Gênero masculino	Dicotômica
	NSE	Padronizada
	Cor preta*	Dicotômica
	Defasagem idade/série	Não padronizada
Turma	Média das proficiências prévias	Não padronizada
	Média do NSE	Padronizada
	Percentual de gênero masculino	Percentual
	Percentual de cor/raça "preta"	Percentual
	Média da defasagem idade/série	Não Padronizada
	Frequência de uso de recursos pedagógicos disponíveis	Padronizada
	Percepção do professor sobre os obstáculos que impedem o melhoramento da escola	Padronizada
	Percepção do professor sobre a frequência de interrupção da aula	Padronizada
	Frequência com que o professor realiza determinadas práticas nas suas aulas de Língua Portuguesa	Padronizada
	Frequência com que o professor realiza determinadas práticas nas suas aulas de Matemática	Padronizada
	Frequência de leitura do professor	Padronizada
	Frequência de participação em atividades culturais	Padronizada
	Frequência com que o professor realiza procedimentos pedagógicos promotores de motivação intrínseca	Padronizada
	Frequência com que o professor realiza procedimentos pedagógicos promotores de motivação extrínseca com regulação introjetada	Padronizada
Frequência com que o professor realiza procedimentos pedagógicos promotores de motivação extrínseca com regulação externa	Padronizada	
Escola	Média das proficiências prévias	Não padronizada
	Média do NSE	Padronizada
	Percentual do gênero masculino	Percentual
	Percentual de cor/raça "preta"	Percentual
	Média da defasagem idade/série	Não padronizada
	Rede privada	Dicotômica
	Média da frequência de uso de recursos pedagógicos disponíveis	Padronizada
	Média da percepção do professor sobre a frequência de interrupção da aula	Padronizada
	Média da frequência com que o professor realiza determinadas práticas nas suas aulas de Língua Portuguesa	Padronizada
	Média da frequência com que o professor realiza determinadas práticas nas suas aulas de Matemática	Padronizada
	Média da frequência de leitura do professor	Padronizada
	Média da frequência de participação em atividades culturais	Padronizada
	Média da frequência com que o professor realiza procedimentos pedagógicos promotores de motivação intrínseca	Padronizada
	Média da frequência com que o professor realiza procedimentos pedagógicos promotores de motivação extrínseca com regulação introjetada	Padronizada
Média da frequência com que o professor realiza procedimentos pedagógicos promotores de motivação extrínseca com regulação externa	Padronizada	

Fonte: elaboração dos autores

\* Resposta à opção c da seguinte pergunta: Como você se considera? a) Branco(a); b) Pardo(a); c) Preto(a); d) Amarelo(a); e) Indígena.

## Descrição dos modelos de valor agregado

Com a finalidade de identificar os fatores que influenciam a aprendizagem ano a ano e os padrões de crescimento ao longo dos anos, dois tipos de modelos são construídos, assim denominados: modelo de valor agregado (*Value-added model*) onda a onda e modelo de evolução.

Modelos de três níveis para valor agregado têm a intenção de aferir quanta proficiência se agregou em ondas sucessivas de testagem. Para isso, os modelos onda a onda obrigatoriamente possuem a proficiência prévia do aluno, que é a proficiência aferida na entrada do ciclo considerado. Esses modelos contemplam variáveis com a intenção de explicar o ganho de proficiência do aluno.

Na elaboração desses modelos procura-se identificar como as práticas pedagógicas de cada professor, bem como o clima escolar e políticas de alocação e formação de turmas na escola, entre outras variáveis do contexto escolar e as características pessoais dos alunos, influenciam a proficiência discente. Assim, é possível explicar as diferenças de proficiência entre alunos, turmas e escolas.

O segundo tipo de modelo mostra o efeito produzido pelas variáveis na curva de evolução da proficiência. A estrutura do modelo é de três níveis, na qual o primeiro nível corresponde ao desempenho em cada tempo; o segundo, às características dos alunos; e o terceiro, às características da escola. Assim, dadas as covariáveis, é possível determinar em cada onda, para cada aluno, em cada escola, qual era a proficiência esperada.

## Modelos hierárquicos

Os modelos de regressão linear múltipla tradicionais, utilizados para a explicação de uma variável dependente com base num conjunto de variáveis independentes, baseiam-se em quatro pressupostos básicos para as características dos dados: linearidade, normalidade, homocedasticidade e independência entre os

elementos amostrais. Em geral, os três primeiros pressupostos são razoavelmente verificados nos dados educacionais ou contornados a partir da utilização de grandes amostras. Por outro lado, a independência dos elementos amostrais não é razoável em dados de pesquisas educacionais, uma vez que a população de alunos está organizada em turmas e estas em escolas. Logo, a estrutura dos dados na população é naturalmente hierárquica.

Neste caso, torna-se pouco razoável admitir a independência para as observações individuais como, por exemplo, os alunos, já que estaria sendo desprezado o efeito de agregação: alunos de uma mesma turma tendem a apresentar características mais semelhantes do que alunos de turmas diferentes, mesmo que difiram entre si quanto a vários aspectos individuais.

Nos modelos hierárquicos de três níveis é levada em consideração a estrutura de agrupamento dos dados, admitindo que cada turma e escola, por exemplo, tenham um modelo de regressão particular. Nesses modelos a influência que cada variável exerce sobre a proficiência do aluno pode depender da agregação das unidades amostrais, além de também eventualmente vir a depender de variáveis encontradas em níveis de agregação superiores como, por exemplo, as variáveis de escola.

Respeitando a estrutura hierárquica presente no modelo, a expressão matemática utilizada contém os índices  $i, j$  e  $k$  que são os indexadores do 1° ao 3° nível respectivamente, exemplo,  $w$  e  $z$  que representam as variáveis do 1° ao 3° nível analogamente. O modelo hierárquico então terá a seguinte expressão geral, segundo Bryk & Raudenbush (1992):

### Fórmula 1 – Equações do modelo hierárquico

$$Y_{ijk} = \pi_{0jk} + \sum_{f=1}^F \pi_{fijk} X_{fijk} + e_{ijk}$$

$$\pi_{fjk} = \beta_{f0k} + \sum_{s=1}^S \beta_{fsk} W_{sjk} + u_{fjk}, f = 0, \dots, F \quad (1)$$

$$\beta_{fsk} = \gamma_{fso} + \sum_{t=1}^T \gamma_{fst} Z_{tk} + r_{fsk}, f = 0, \dots, F \text{ e } s = 0, \dots, S$$

A variável dependente  $Y_{ijk}$  representa a proficiência em Língua Portuguesa ou em Matemática do  $i$ -ésimo aluno pertencente a  $j$ -ésima turma da  $k$ -ésima escola nas ondas 2 até 5 dependendo do modelo considerado. O termo  $X_{fijk}$  representa a covariável  $f$  de primeiro nível (aluno), tais como o gênero masculino e a defasagem idade/série, do  $i$ -ésimo aluno. A  $s$ -ésima covariável de segundo nível (turma), como por exemplo, percentual de gênero masculino na turma e defasagem média da turma, é representada pelo termo  $W_{sjk}$ . O termo  $Z_{tk}$  representa as variáveis do terceiro nível, como percentual de gênero masculino na escola e defasagem média da escola.

Os coeficientes  $\pi_{fjk}$ ,  $\beta_{fsk}$  e  $\gamma_{fst}$  representam os efeitos das covariáveis  $X_{fijk}$ ,  $W_{sjk}$  e  $Z_{tk}$  respectivamente, sobre a proficiência do aluno em Língua Portuguesa ou Matemática em cada uma das ondas. Já os termos  $U_{fjk}$  e  $r_{f_{sk}}$  medem a incerteza associada aos coeficientes  $\pi_{fjk}$  e  $\beta_{fsk}$  em seus respectivos níveis de agrupamento, turmas (2° nível) ou escolas (3° nível). Finalmente, o termo  $\epsilon_{ijk}$  representa a incerteza de todos os efeitos associados na proficiência de cada aluno.

### Construção dos modelos multiníveis

O processo básico mais utilizado na construção de um modelo hierárquico é o *bottom-up*, isto é, parte-se de um modelo nulo no qual somente se ajustam constantes relativas a cada nível representado, utilizado como ponto de partida para a inclusão das demais variáveis sempre mantendo suas constantes iniciais e incluindo-se as variáveis segundo uma heurística definida pelo especialista que, neste estudo, se baseia na verificação da significância dos coeficientes (parâmetros fixos e aleatórios) para cada modelo.

De modo geral, a construção dos modelos aqui apresentados seguiu esses passos. Inicialmente, analisou-se o modelo nulo com o objetivo de avaliar a proporção da variância devida a cada nível hierárquico. Em seguida, introduziram-se as principais variáveis do

nível de aluno para produzir um modelo que convencionalmente é chamado de básico ou de referência. Em uma segunda etapa, sempre seguindo a mesma heurística, foram introduzidas as variáveis de nível de professor, turma e, finalmente, variáveis de nível de escola. Dessa forma, pode-se analisar a evolução da explicação alcançada após a introdução de cada variável, sempre testando inicialmente os efeitos aditivos, seguidos pelas interações entre as variáveis de diferentes níveis.

### Interpretação dos modelos multiníveis

Para interpretar os modelos multiníveis é fundamental conhecer os elementos estimados por esses modelos, dentre os quais se destacam o intercepto, os coeficientes fixos e os coeficientes randômicos.

O intercepto é um termo constante que, apesar de não possuir interpretação direta na maior parte dos modelos aqui construídos, é muito importante para o ajuste da equação de cada modelo, portanto, jamais deve ser suprimido. Os coeficientes fixos medem o efeito de cada variável e não mudam em cada unidade de análise.

Caso o coeficiente apresente efeito randômico significativo, ele passa a variar nos diferentes níveis de agregação como, por exemplo, diferentes turmas podem apresentar diferentes coeficientes caso este coeficiente apresente efeito randômico no nível de turma, o mesmo pode acontecer no nível de escola; a média deste coeficiente é apresentada na tabela juntamente com os desvios padrões nos níveis nos quais ele apresente significância.

O termo  $e$  apresentado ao final da tabela representa a incerteza média esperada, quanto menor for essa incerteza em relação ao desvio padrão original, maior o poder explicativo do modelo.

### Modelos onda a onda

Foram construídos cinco modelos segundo essa linha, iniciando com o modelo da onda 2 predita pela onda 1, até a onda 5 predita pela

onda 4. Um modelo adicional também construído foi o da onda 5 predita pela onda 1, que permitiu a modelagem do valor agregado durante o ciclo total do GERES.

### Resultado dos modelos nulos dos modelos onda a onda

A decomposição da variância nos três níveis – aluno, turma e escola – é representada na tabela 2.

No primeiro nível (aluno), a diferença na variabilidade observada entre Língua Portuguesa e Matemática é maior no modelo da onda 2. Ambas vão aumentando até que praticamente se igualam no modelo da onda 5 predita pela onda 1. No segundo nível (nível de turma e professor) ambas as variabilidades

aumentam: a diferença entre elas é pequena na onda 2 e mais expressiva e em favor de Língua Portuguesa no modelo da onda 5 predita pela onda 1. O terceiro nível apresenta um comportamento diferente: as variabilidades diminuem, a de Língua Portuguesa inicialmente é maior, mas logo se torna menor. Os resultados permitem afirmar que as diferenças entre os alunos são sempre as mais importantes fontes de variabilidade, sendo responsáveis por entre 46% e 59% de toda a variação nos resultados. Mesmo assim, tanto as características das turmas e do professor quanto outros aspectos do contexto escolar também são muito relevantes na explicação da variância nos resultados dos alunos. Com o passar do tempo, a importância da escola diminui enquanto aumenta a relevância das diferenças entre os alunos.

**Tabela 2** – Partição da variabilidade explicada por diferentes níveis ao longo das ondas (%)

Onda	2		3		4		5		5*	
Proficiência predita	LP	MT								
Aluno	45,8	51,3	52,6	51,3	53,5	50,5	58,6	58,6	57,9	58,1
Turma	22,6	22,8	23,5	23,8	28,0	28,2	27,8	25,1	27,5	24,6
Escola	31,5	25,9	23,9	24,9	18,6	21,3	13,9	16,3	14,6	17,3
Total	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

\* Considerando os dados do modelo da onda 5, predita pela onda 1.

Fonte: elaboração dos autores

### Resultado dos modelos onda a onda

Os modelos de valor agregado onda a onda estão apresentados na tabela 3.

### Associação das variáveis do aluno com o desempenho

No nível do aluno, tanto em Língua Portuguesa quanto em Matemática, as variáveis que mais se associam a ganhos maiores de proficiência onda a onda são a proficiência prévia e o NSE. Quanto maior a proficiência prévia do aluno e maior seu NSE, maior seu ganho de pontos nas escalas GERES de uma onda para outra. Por exemplo, os alunos tendem

a apresentar uma vantagem de 1,1 pontos na escala de proficiência de Língua Portuguesa no modelo para a 2ª onda, e 1,11 pontos em Matemática. No modelo da 5ª onda predita pela 1ª, os ganhos associados a cada desvio padrão do NSE são de 1,85 pontos em Língua Portuguesa e 4,84 pontos em Matemática. Esses resultados confirmam estudos anteriores que mostram também uma associação estreita entre o NSE da família e o ponto de partida inicial do aluno.

Por outro lado, a relação significativa mais negativa entre gênero masculino e o avanço em Língua Portuguesa nos cinco modelos mostra que os meninos estão em desvantagem na aprendizagem da leitura quando comparados

**Tabela 3** – Coeficientes do modelo de valor agregado onda a onda<sup>a</sup>

Nível	Coeficientes	Ondas										
		2		3		4		5		5**		
		LP	MT	LP	MT	LP	MT	LP	MT	LP	MT	
Escola (variável, intercepto, intercepto)	NSE médio da escola									3,66	15,45	
	Percentual de pretos na escola					-22,6	-36,4					
	Média de prática de leitura na escola	0,64				0,82				1,55	3,21	
	Média de interrupção da aula na escola						-2,89		-2,45	-1,19	-5,92	
	Média da frequência de uso de recursos pedagógicos disponíveis na escola			0,66								
	Média de prática de aula de LP na Escola			0,77							1,33	
Turma** (variável, intercepto)	Média de proficiência prévia	0,15	0,28	0,08	0,15	0,30	0,41	0,34	0,32	0,32		
	Média do NSE	2,63	2,79	1,41	6,76		1,93					
	Percentual de gênero masculino		-11,9		-11,9						-19,95	
	Percentual de pretos			-7,35	-12,7							
	Defasagem média da turma				-12,7							
	Média de interrupção da aula na turma							-1,17			-1,11	
	Frequência de uso de recursos pedagógicos disponíveis na Turma								1,62			
Aluno	Intercepto	Intercepto	55,34	48,28	29,18	-1,6	25,79	31,56	9,44	33,59	96,20	149,26
		Rand de Turma	4,92	8,79	4,21	26,72	4,38	27,89	4,36	11,56	6,81	18,79
		Rand de escola	1,59	3,32	2,18	5,31	1,95	6,04	1,55	6,32	3,52	9,24
	Proficiência prévia	Intercepto	0,51	0,60	0,82	1,04	0,66	0,67	0,62	0,72	0,36	0,89
		Gênero masculino	Intercepto	-1,98		-3,09		-2,37		-2,53		-4,73
	NSE	Intercepto	1,10	1,11	1,68	3,80	0,78	2,04	1,07	2,82	1,85	4,84
		Rand de turma	2,24			0,20						
	Cor preta	Intercepto	-1,00	-2,65	-2,21	-4,19	-1,27	-2,90	-2,7	-5,34	-4,24	
	Defasagem idade/série	Intercepto			-2,93		-5,72	-11,7	-3,31	-6,22	-18,3	-47,34
	Σ		11,9	18,7	14,5	30,0	15,6	34,7	14,5	36,5	17,7	44,57

Fonte: dados da pesquisa.

\* Todos os coeficientes apresentados apresentaram significância ao nível de  $p \leq 0,05$ ;

\*\* Tendo como proficiência prévia a proficiência na onda 1;

\*\*\* Nenhum coeficiente do nível de turma apresentou efeito aleatório significativo.

com as meninas. A desvantagem em Língua Portuguesa dos alunos aumenta da 2ª para a 5ª onda, de -1,98 para -4,73 pontos, em relação à 1ª onda, mostrando que o déficit é progressivo e possivelmente cumulativo. Por outro lado, os resultados mostram que a aprendizagem de Matemática não é influenciada de forma

significativa pelo gênero dos alunos em nenhum dos modelos.

Os alunos que se autodenominaram pretos e aqueles defasados possuem coeficientes negativos de regressão associados a essas variáveis, nos modelos de ambas as disciplinas. Isso significa que tanto em Matemática quanto

em Língua Portuguesa, o grupo de alunos de cor preta está em desvantagem em relação aos outros grupos e, onda a onda, agrega menos pontos na escala de proficiência. O efeito negativo da experiência da defasagem, significativa em Língua Portuguesa, no modelo para as ondas 3, 4, 5 e especialmente na onda 5, predita pela onda 1, reflete o efeito cumulativo dos atrasos dos alunos. O mesmo ocorre em Matemática no modelo para as ondas 4, 5 e no modelo da onda 5 predita pela onda 1. Mesmo nos anos iniciais do ensino fundamental os atrasos já se mostram um impedimento à aprendizagem.

Apesar do número de pontos de desvantagem dos alunos autotransclassificados como pretos ser maior em Matemática que em Língua Portuguesa, não se deve concluir que o efeito da cor seja maior nessa disciplina. Deve-se lembrar que as duas escalas são diferentes e que a escala de Matemática tem o dobro da abrangência da escala de Língua Portuguesa, em termos de pontos. No entanto, o efeito da defasagem idade/série em Matemática no modelo da onda 5 predita pela onda 1, de -47,34 pontos extrapola a diferença esperada e pode indicar a existência de um impacto diferencial por disciplina. Haverá necessidade de estudar esse fenômeno com outros dados para confirmar a possibilidade de processos diferenciados de aprendizagem.

### **Associação das variáveis de turma com o desempenho dos alunos**

Alunos pertencentes a turmas com uma maior média das proficiências prévias tendem a apresentar uma maior proficiência em Língua Portuguesa, em todas as ondas, bem como no modelo da onda 5 predita pela onda 1. Em Matemática, o mesmo se passa nas cinco ondas, embora o efeito não seja significativo no modelo da onda 5 predita pela onda 1.

Nos modelos construídos para as ondas 2 e 3 foi significativa a variável NSE médio da turma. Tal fato indica que os alunos pertencentes a turmas com um maior NSE médio tendem a apresentar uma maior proficiência em Língua

Portuguesa, mostrando a influência dos pares sobre a evolução da aprendizagem individual. No caso de Matemática, essa variável foi significativa nos modelos das ondas 2, 3 e 4.

Na onda 3, a variável percentual de pretos foi significativa tanto para Língua Portuguesa quanto para Matemática. A defasagem média na turma é significativa em Matemática na onda 3, enquanto em Língua Portuguesa essa variável não tem significância em nenhuma onda.

Diferentemente da mesma variável em nível de aluno, a variável percentual de gênero masculino na turma foi significativa em Matemática nas ondas 2 e 3, e no modelo da onda 5 predita pela onda 1. Já em Língua Portuguesa essa variável não é significativa em nenhuma onda. Com esses resultados, observa-se a inversão das relações observadas em nível de aluno.

A frequência de uso de recursos pedagógicos disponíveis na turma é significativa no modelo da onda 5 em Matemática. Já em Língua Portuguesa essa variável não é significativa em nenhuma onda.

Na onda 5 e no modelo da onda 5 predita pela onda 1, a variável média de interrupção da aula foi significativa, indicando que alunos pertencentes a turmas com maiores valores de média de interrupção da aula tendem a apresentar uma menor proficiência em Língua Portuguesa. Já em Matemática essa variável não é significativa em nenhuma onda. Alunos que estão alocados em turmas com maiores níveis de NSE tendem a apresentar um aumento de 2,63 pontos em Língua Portuguesa para cada desvio padrão a mais do que a média, e 2,79 em Matemática nos modelos da onda 2. Ou seja, pelo menos nos primeiros anos, o efeito do NSE da turma é maior do que o NSE do indivíduo. O efeito do NSE da turma é especialmente marcado na 3ª onda, quando sobe para 6,76 pontos em Matemática.

### **Associação das variáveis da escola com o desempenho dos alunos**

Tanto em Língua Portuguesa quanto em Matemática, alunos pertencentes a escolas

com maiores níveis de percentual de pretos e média de interrupção da aula na escola tendem a apresentar uma menor proficiência. No caso da variável percentual de pretos, isso acontece na onda 4 e, no caso da variável média de interrupção da aula na escola, isso acontece nas ondas 4, 5 e no modelo da onda 5 predita pela onda 1, em Matemática. O mesmo ocorre com o modelo da onda 5 predita pela onda 1 em Língua Portuguesa. Novamente, em ambos os modelos, os alunos pertencentes a escolas com maiores níveis de média da prática de leitura na escola, NSE e média de prática de aula na escola tendem a apresentar uma maior proficiência. No caso da primeira variável, isso acontece em Língua Portuguesa na onda 2 e no modelo da onda 5 predita pela onda 1, e, em Matemática é significativa apenas no modelo da onda 5 predita pela onda 1. Na segunda variável, isso acontece no modelo da onda 5 predita pela onda 1, em ambos os modelos e na terceira variável é observado em Língua Portuguesa na onda 3 e no modelo da onda 5 predita pela onda 1. Por fim, em Matemática, tal efeito é significativo apenas no modelo da onda 5 predita pela onda 1.

Algumas variáveis apresentam significância apenas no modelo de Língua Portuguesa mostrando, assim, que alunos pertencentes a escolas com um maior percentual de frequência de uso de recursos pedagógicos disponíveis na escola, maiores níveis de média de prática de aula de Língua Portuguesa na escola e uma maior média de prática de leitura na escola tendem a apresentar uma maior proficiência em Língua Portuguesa. No caso da primeira variável isso acontece na onda 3; na segunda variável, nas ondas 3 e 5; e na terceira variável, nas ondas 2, 4 e no modelo da onda 5 predita pela onda 1. Esses resultados indicam que as variáveis empregadas para captar diferenças entre os métodos de ensino dos professores demonstram efetivamente que diferentes abordagens para o ensino de português produzem resultados diferentes, com vantagens para aqueles alunos cujos professores usam recursos pedagógicos, dão

ênfase ao ensino de Língua Portuguesa e que fazem mais uso da leitura na sala de aula.

## Modelos de evolução

Os modelos aqui empregados podem ser representados pela seguinte equação.

### Fórmula 2 – Equação do modelo hierárquico longitudinal

$$Y_{tij} = \pi_{0ij} + \pi_{1ij}t + \pi_{2ij}t^2 + e_{tij} \quad (2)$$

Esse modelo permite analisar a evolução da proficiência do aluno segundo as ondas do GERES. Sua formulação é análoga ao modelo anteriormente apresentado, porém seus níveis agora considerados são o tempo (t, 1° nível), aluno (i, 2° nível) e escola (j, 3° nível). Na equação apresentada é ilustrado apenas o primeiro nível, o nível do tempo, a especificação dos demais níveis é semelhante àquela presente na fórmula 1.

Esta abordagem tem a vantagem de mostrar não somente o efeito produzido por uma covariável de aluno ou escola no patamar de proficiência (por meio do coeficiente  $\pi_{0ij}$ ), mas também o seu efeito na curva de evolução da proficiência (por meio dos coeficientes  $\pi_{1ij}$  e  $\pi_{2ij}$ ).

Nos modelos de evolução, o tempo, representado por t foi codificado segundo a sequência de inteiros de 0 a 4 (representando as ondas de 1 a 5).

### Resultado dos modelos nulos de evolução

Pela decomposição da variância nos três níveis (tempo, aluno e escola) temos a tabela 4:

**Tabela 4** – Decomposição da variância por nível e disciplina (%)

Proficiência	LP	MT
Tempo	67,0	74,1
Aluno	12,9	8,9
Escola	20,1	17,0
Total	100,0	100,0

Fonte: dados da pesquisa.

A variabilidade no primeiro nível (tempo) é maior que nos demais níveis e corresponde a 67% na proficiência de Língua Portuguesa e a cerca de 74% na de Matemática. No segundo nível (aluno), a variabilidade na proficiência decai de 12,9% em Língua Portuguesa para 8,9% em Matemática. O mesmo ocorre no terceiro nível (escola), porém, neste nível, a variabilidade é ainda menor, correspondendo a cerca de 20,1% de Língua Portuguesa e 17,0% em Matemática. Em resumo, a variável tempo explica mais da variação na proficiência dos

alunos que as variáveis em nível de escola e de aluno, e seu efeito em Matemática é maior que em Língua Portuguesa. A razão pela aparente dependência maior da aprendizagem de Matemática ao fator tempo será discutida posteriormente.

### Resultado dos modelos de evolução

O modelo de evolução para a proficiência em Língua Portuguesa e Matemática nas cinco ondas está apresentado na Tabela 5:

**Tabela 5** – Modelo de valor agregado para a proficiência em Língua Portuguesa e Matemática nas 5 ondas e covariáveis\*

Tempo	Variáveis		Língua Portuguesa		Matemática	
	Aluno	Escola	Coefficiente	Efeito Randômico	Coefficiente	Efeito Randômico
Intercepto				14,77		11,42
	Intercepto			9,39		7,16
		Intercepto	110,19		112,24	
		Média do NSE da escola	8,04		10,05	
		Média dos obstáculos na escola**	-1,76		-2,04	
		Rede privada	3,6			
	Cor preta	Intercepto	-2,75		-4,45	
	NSE			1,87		
		Intercepto	3,33		3,30	
	Gênero masculino	Intercepto	-3,16			
t				4,21		14,54
	Intercepto			8,15		8,12
		Intercepto	20,71		18,45	
		Rede privada			11,87	
		Média de interrupção na escola***			-1,15	
	NSE	Intercepto	0,15		2,30	
	Cor preta	Intercepto	-1,89		-3,89	
	Gênero masculino	Intercepto	-1,75			
	t <sup>2</sup>			0,63		2,36
	Intercepto			1,13		1,62
	Intercepto	-1,26		3,58		
	Rede privada			-1,46		
Gênero masculino			0,18			
	Intercepto	0,17				
Cor preta	Intercepto	0,2				
e				10,41		21,31

Fonte: dados da pesquisa.

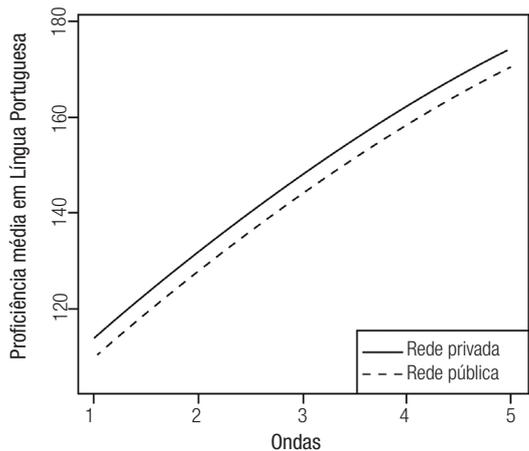
\* Todos os coeficientes são significantes ao nível de  $p \leq 0,05$ ;

\*\* Se refere à Média da Percepção do Professor sobre os Obstáculos que Impedem o Melhoramento da Escola;

\*\*\*Se refere à Média da Percepção do Professor sobre a Frequência de Interrupção da Aula na Escola.

Nos modelos de evolução construídos, identificaram-se quatro principais padrões de evolução, os quais, embora sejam ascendentes, diferem-se quanto à sua respectiva forma. A seguir são mostradas quatro figuras exemplificando os padrões identificados.

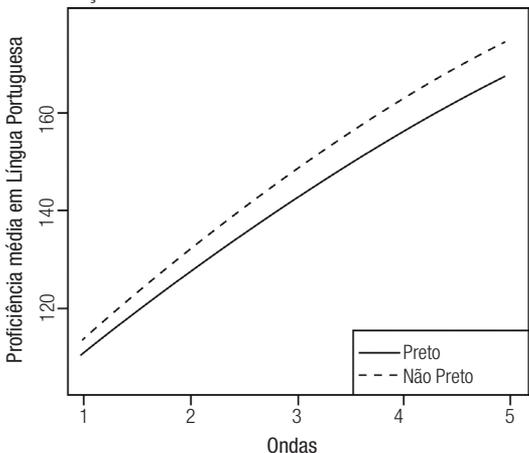
**Figura 1** – Evolução da proficiência em Língua Portuguesa segundo a rede da escola



Fonte: dados da pesquisa.

Na Figura 1, apresenta-se um exemplo do padrão 1 – Curvas paralelas com taxas de crescimento decrescentes, e os outros casos semelhantes serão apresentados e discutidos a seguir.

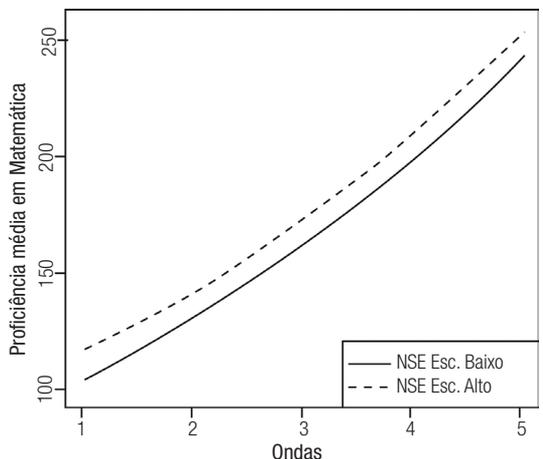
**Figura 2** – Evolução da proficiência em Língua Portuguesa pela cor/raça



Fonte: dados da pesquisa.

Na Figura 2, apresenta-se um exemplo do padrão 2 – Curvas com taxa de crescimento decrescente que se distanciam entre si com o passar do tempo, divergentes.

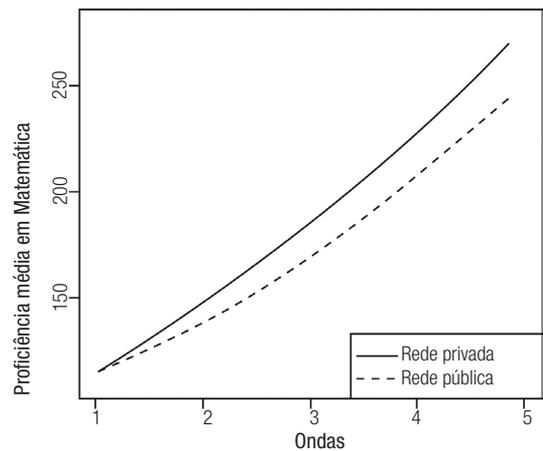
**Figura 3** – Evolução da proficiência em Matemática pelo NSE da escola



Fonte: dados da pesquisa.

Na Figura 3, apresenta-se um exemplo do padrão 3 – Curvas paralelas com taxa de crescimento crescentes.

**Figura 4** – Evolução da proficiência em Matemática segundo a rede da Escola



Fonte: dados da pesquisa.

Na Figura 4, apresenta-se um exemplo do padrão 4 – Curvas com taxa de crescimento

crecente que se distanciam com o passar do tempo, divergentes.

Tanto em Língua Portuguesa quanto em Matemática a evolução da proficiência evidenciada pelos modelos ocorre de forma ascendente, porém, em Língua Portuguesa, as taxas de crescimento decaem, enquanto que, em Matemática, elas aumentam.

A influência exercida pelas diferentes variáveis estudadas e incluídas nos modelos sempre se encaixou em um dos padrões acima especificados, percebendo-se variações apenas quanto à magnitude de seu efeito. Os padrões 1 (Figura 1) e 2 (Figura 2) foram encontrados em Língua Portuguesa, e os padrões 3 (Figura 3) e 4 (Figura 4), em Matemática.

A influência das variáveis NSE do Aluno, NSE médio da escola, rede privada e obstáculos na escola em Língua Portuguesa apresentam o primeiro padrão de crescimento, evidenciando que esses grupos se diferem apenas no patamar de crescimento. Já a influência das variáveis gênero masculino e cor preta faz com que esses grupos se difiram no patamar e na taxa de crescimento, correspondendo ao segundo padrão de crescimento, padrão no qual os alunos do gênero masculino e os que se denominam pretos, além de apresentarem inicialmente uma proficiência menor, tendem a ampliar essa diferença ao longo das ondas.

Em Matemática, a influência das variáveis NSE médio da escola e obstáculos na escola se encaixaram no terceiro padrão de crescimento, evidenciando que os valores assumidos por tais variáveis apenas afetam o patamar de crescimento, não provocando diferenças nas taxas de crescimento, que se mantêm constantes ao longo de todas as ondas. Já a influência das variáveis NSE, cor preta, rede privada e média da interrupção da aula na escola apresentam o quarto padrão de crescimento, ou seja, as diferenças entre os grupos que apresentam diferentes valores dessas variáveis tendem a se ampliar apesar do aumento das taxas de crescimento ao longo das ondas ocorrer em todos os casos.

## **Discussão dos resultados**

Na análise descritiva dos resultados dos modelos de ganho agregado onda a onda identificaram-se semelhanças na evolução da aprendizagem de Língua Portuguesa e Matemática. Ambas as disciplinas são igualmente influenciadas pela cor, pelo NSE e pela proficiência prévia dos alunos. Nesses casos, os resultados coincidentes das duas disciplinas parecem oferecer certa confirmação da correção dos métodos e dos efeitos observados. Ao mesmo tempo, chama atenção a indicação de diferenças entre a Língua Portuguesa e a Matemática na forma em que as duas matérias são influenciadas por outras variáveis em estudo. No caso do gênero do aluno, por exemplo, observa-se o fenômeno da desvantagem do gênero masculino na aprendizagem da leitura. Descrita por outros pesquisadores, essa desvantagem pode significar diferenças cognitivas entre os gêneros. Menos previsível é a aparente diferença entre Língua Portuguesa e Matemática no que diz respeito ao efeito da experiência da defasagem idade/série, sugerindo que o atraso na Matemática seja mais difícil de superar que o na Língua Portuguesa.

Diferenças nos efeitos das variáveis de turma e de escola nas duas disciplinas também merecem atenção. Em nível de turma, existe uma associação negativa entre o gênero masculino e a aprendizagem de Matemática em três dos modelos estudados, sugerindo um efeito dos pares na aprendizagem desta matéria. Em nível da escola, a média de interrupção da aula mostra uma associação negativa com a aprendizagem de Matemática em três dos modelos estudados, enquanto a aprendizagem de Língua Portuguesa está associada negativamente a essa variável somente na 5ª onda predita pela 1ª. Ou seja, as interrupções fazem mal em ambos os casos, mas têm efeitos negativos maiores e mais significativos no caso da aprendizagem de Matemática.

Os resultados da pesquisa Geres parecem oferecer indícios de diferenças entre as disciplinas na maneira que são aprendidas e

nos fatores que influenciam essa aprendizagem. O primeiro indício dessas diferenças foi o aumento significativo no desvio padrão da proficiência em Matemática em comparação com o de Língua Portuguesa. Entre as cinco ondas o desvio padrão em Língua Portuguesa se manteve bastante estável, variando entre 26,28 na 1ª onda e 25,37 na 5ª. Em Matemática, por outro lado, o desvio padrão aumentou 132% ao longo da pesquisa, crescendo de um patamar parecido com o de Língua Português na 1ª onda, de 28,82 para 66,87 na 5ª onda. Esse crescimento significa, no mínimo, uma aprendizagem progressivamente mais difícil para uma proporção significativa de alunos.

Os resultados dos modelos de evolução da aprendizagem mostram, em primeiro lugar, que a aquisição da Matemática começa mais devagar e depois acelera, enquanto a aprendizagem de Língua Portuguesa tem trajetória que sinaliza um processo de desaceleração a partir da 3ª onda. Independente das variáveis em estudo, a curva da Matemática é uma curva em aceleração enquanto a curva de Língua Portuguesa tem o formato diferente por ser uma curva em desaceleração. A diferença entre esses padrões pode confirmar uma importante e conhecida diferença entre as disciplinas no que diz respeito ao lugar da aprendizagem inicial. No caso da Língua Portuguesa, trata-se de uma disciplina ou área de conhecimento que já está em desenvolvimento antes da criança ser submetida ao processo formal de alfabetização, pelo uso da língua na convivência diária com a família. A Matemática, por outro lado, é um campo de conhecimento que não necessariamente recebe estímulos antes da criança entrar para a escola, o que torna essa uma disciplina mais nitidamente escolar. À luz dessa diferença, espera-se um desenvolvimento mais rápido da Matemática após um período inicial de familiarização.

Em segundo lugar, a variável que informa a rede da escola se relaciona de forma diferente com a aprendizagem de Matemática e Língua Portuguesa. No caso da Matemática, abre-se uma diferença no nível e ritmo de

aprendizagem entre as redes pública e privada, que só aumenta ao longo dos primeiros anos. Em Língua Portuguesa, por outro lado, a diferença que existe entre as redes desde o princípio do 2º ano se mantém ao longo do período sem alteração. Essa diferença poderia estar relacionada a diferenças no nível socioeconômico ou cor dos alunos, mas ao estudar essas relações, encontram-se curvas cujos padrões são parecidos. Por esse motivo, parece razoável supor que a diferença na aprendizagem de Matemática, entre escolas privadas e públicas, diz respeito a diferenças nos processos de ensino.

A descoberta das origens das diferenças fica além do escopo deste trabalho. No entanto, a análise dos modelos de valor agregado e de evolução não só confirma estas diferenças como começa a apontar componentes desta.

## **Conclusão**

As pesquisas educacionais feitas anteriormente só dispunham de um único resultado, não permitindo cálculos de ganhos agregados. As inferências eram feitas contextualizando a proficiência pelo NSE do aluno e de suas unidades de agrupamento como turmas e escolas. Com esses trabalhos era possível elencar fatores associados ao desempenho, ainda que houvesse um controle estatístico mais rigoroso, mas não era possível incorporar a informação acerca da ordem dos eventos, de forma a mapear causas e efeitos.

Utilizando-se um desenho de pesquisa longitudinal também é possível remover, em parte, o *background* dos alunos e acompanhar os efeitos das variáveis, identificando se seus efeitos atuam não somente nos patamares do desempenho, como também em seu ritmo de crescimento.

Nessas mesmas pesquisas podíamos mensurar as diferenças e a influência dos diversos valores das variáveis no desempenho do aluno, como por exemplo, entre as redes de ensino pública e privada, ou entre os gêneros; porém, não ficava claro se essas diferenças se

ampliavam ou reduziam ao longo do tempo. Desta forma não era possível contextualizar o papel da escola na tentativa de neutralizar, ou até mesmo de acentuar, tais processos. Observou-se o aumento nas diferenças de proficiências de grupos de alunos pretos *versus* não pretos, escola pública *versus* escola privada em Língua Portuguesa, gênero masculino *versus* feminino em Matemática, ao longo dos anos escolares. Por outro lado, outras diferenças para outros grupos, caracterizados por outras variáveis, como o NSE, se mantêm ao longo das etapas escolares. Assim sendo, esses resultados

trazem um alerta para os sistemas educacionais, pois a escola está longe de produzir a equidade desejada pela sociedade brasileira.

A pesquisa GERES apresenta um passo à frente e pode-se afirmar que constitui um estágio intermediário na pesquisa educacional brasileira. Seus resultados vão além das pesquisas transversais observacionais que constituíam quase a totalidade de pesquisas educacionais em grandes amostras no Brasil e prepara terreno para as pesquisas de natureza longitudinal e experimental, embasando as primeiras hipóteses a serem testadas nestas.

## Referências

- BONAMINO, Alicia C. **Tempos de avaliação educacional: o SAEB, seus agentes, referências e tendências.** Rio de Janeiro: Quartet, 2002.
- BROOKE, Nigel; BONAMINO, Alicia. (Org.). **Geres 2005: razões e resultados de uma pesquisa longitudinal sobre a eficácia escolar.** Rio de Janeiro, WalPrint, 2011.
- BRYK, Anthony; RAUDENBUSH, Stephen. **Hierarchical linear models: applications and data analysis methods.** Newbury Park: Sage, 1992.
- FRANCO, Creso; ALVES, Fátima; BROOKE, Nigel. Estudo longitudinal sobre qualidade e equidade no ensino fundamental brasileiro (GERES 2005). **Revista Ensaio**, Rio de Janeiro, n. 61, out/dez 2008.
- \_\_\_\_\_. O SAEB — Sistema de Avaliação da Educação Básica: potencialidades, problemas e desafios. **Revista Brasileira de Educação**, ANPEd, n. 17, p. 127-133, maio/jun/jul/ago 2001.
- GOLDSTEIN, Harvey. Hierarchical Data Modeling in the Social Sciences. **Journal of Educational and Behavioral Statistics**. v. 20, n. 2, p. 201-204, 1995.
- \_\_\_\_\_. et al. The use of value added information in judging school performance. **Perspectives on Education Policy**. London: Institute of Education, 2000. 40 p.
- RAUDENBUSH, Stephen W.; FOTIU, Randall P.; & CHEONG, Yuk Fai. Inequality of access to educational resources: a national report card for eighth grade math. **Revista Educational Evaluation and Policy Analysis**, v. 20, n. 4, p. 253-268, 1998.

*Recebido em: 27.02.2013.*

*Aprovado em: 13.08.2013.*

**Nigel Brooke** possui graduação em Psicologia Experimental - *University of Reading* (1969), mestrado em Psicologia Ocupacional - *Birkbeck College, University of London* (1973) e doutorado em Estudos do Desenvolvimento - *Institute of Development Studies* (1979). Foi assessor e depois representante da Fundação Ford no Brasil entre 1994 e 2003. Atualmente, é Professor Convidado da FAE/UFMG e pesquisador do Grupo de Avaliação e Medidas Educacionais - GAME/UFMG.

**Neimar da Silva Fernandes** é auxiliar de pesquisa do centro de Políticas Públicas e Avaliação da Educação— CAEd — da Universidade Federal de Juiz de Fora (MG).

**Isabela Pagani Heringer de Miranda** é aluna da Universidade Federal de Juiz de Fora (MG).

**Tufi Machado Soares** é professor Associado do Departamento de Estatística e do Programa de Doutorado e Mestrado em Educação, Coordenador de Pesquisa do Caed — Centro de Políticas Públicas e Avaliação da Educação, Universidade Federal de Juiz de Fora (MG). Possui doutorado em Teoria Matemática de Controle e Estatística pela PUC - Rio e pós-doutorado em estatística pela Universidade Federal do Rio de Janeiro.