

## CONSTRUÇÃO DE UM MODELO DE REGRESSÃO HIERÁRQUICO PARA OS DADOS DO SIMAVE-2000

**Tufi Machado Soares \***

Departamento de Estatística  
Centro de Avaliação Educacional (CAEd)  
Universidade Federal de Juiz de Fora – UFJF  
Juiz de Fora – MG  
[tufi@estatistica.ufjf.br](mailto:tufi@estatistica.ufjf.br)

**Márcia Cristina Meneghin Mendonça**

Colégio Politécnico Pio XII  
Universidade Federal de Juiz de Fora – UFJF  
Juiz de Fora – MG

\* *Corresponding author*/autor para quem as correspondências devem ser encaminhadas

*Recebido em 02/2003; aceito em 11/2003*  
*Received February 2003; accepted November 2003*

### Resumo

Os sistemas de avaliação desenvolvidos na última década mantêm um mesmo objetivo prioritário: encontrar mecanismos para melhorar a qualidade do ensino oferecido à sociedade de forma eficaz e eficiente. Além de identificar resultados da aprendizagem dos alunos, avalia-se o conjunto do sistema educacional. Dentro desse contexto o Sistema Mineiro de Avaliação da Educação Pública (SIMAVE) foi instituído pela Secretaria de Estado de Educação no ano de 2000. O presente estudo empregou modelos hierárquicos (multinível) com o objetivo de identificar a relação entre o desempenho escolar, as características técnico-pedagógicas das escolas e o perfil sócio-econômico dos alunos das 4ª séries do Ensino Fundamental que fizeram os testes do SIMAVE-2000. Buscando investigar os fatores que atuam na diferenciação entre as turmas dentro das escolas foram considerados os modelos com os níveis de hierarquia: turma (nível 1) e escola (nível 2). Foi encontrada explicação para uma grande proporção da variação dos escores entre as turmas.

**Palavras-chave:** avaliação educacional; dados hierárquicos; modelo multinível.

### Abstract

The evaluation systems developed on the last decade keep a same main objective: to find ways to efficiently improve the quality of the education. Besides identifying results of the pupil's learning, the educational system is also evaluated. In this context, the Sistema Mineiro de Avaliação da Educação Pública (SIMAVE) was instituted in 2000 year by the Secretaria de Estado de Educação. The present study had used multilevel models with the objective to identify the relation between the educational performance, the technic-pedagogical characteristics of the schools and the social-economic profile of the 4th. grade pupils that had made SIMAVE-2000 tests. Searching to investigate the factors that act in the classrooms differentiation in schools had been considered the multilevel models with the hierarchy levels: classroom (level 1) and school (level 2). It was found explanation for a great proportion of the variation of the scores of the classroom.

**Keywords:** educational evaluation; hierarchical data; multilevel model.

## 1. Introdução

Os problemas educacionais não são recentes. Segundo Ribeiro (1981), em linhas gerais a evolução do processo educativo foi construído pelos choques entre a tradição cultural e as novas exigências educacionais da sociedade. É importante salientar que o país evoluiu historicamente de uma reforma educacional para outra. Mas é possível observar também que os problemas educacionais que foram surgindo em decorrência dos acontecimentos foram se agravando e crescendo exponencialmente. De acordo com pesquisas educacionais recentes, tais como as de Klein (1997), Castro (1998) e Relatório Nacional-PISA 2000 (2001), alguns dos mais graves problemas do sistema educacional brasileiro, que permaneceram até os dias atuais, podem ser os seguintes:

- a) Defasagem qualitativa, com predominância de ensino escolar desvinculado da prática;
- b) Discriminação acentuada favorecendo a manutenção do *status quo*;
- c) Grandes disparidades regionais;
- d) Insuficiência de vagas no ensino superior em determinadas áreas, principalmente no ensino público, agravado com a tendência de universalização, especialmente do ensino médio.

Por outro lado, os sistemas de avaliação da educação, conforme frisa Firme (1994), também evoluíram substancialmente, de tal forma que, a partir dos resultados obtidos, foi possível introduzir mudanças nos aspectos educacionais de interesse da sociedade. Atualmente o foco dessas avaliações é a escola. Para tanto, na composição dos sistemas de avaliação educacional mais recentes estão sendo avaliados não apenas o rendimento acadêmico dos alunos, mas também aspectos humanos, sociais, culturais, éticos, metodológicos e instrumentais. Nesse sentido, espera-se que a evolução dos instrumentos de avaliação educacional permita aos pais, professores, empresários, governos, e outros atores sociais, verificarem se as escolas estão enfrentando adequadamente os desafios das transformações econômicas e sociais.

Assim, atualmente, o principal interesse não é o de se comparar escolas apenas em relação ao rendimento de seus alunos. Pelo menos não deveria ser. Mas, entre outros, o de procurar identificar características – de natureza humana, social, cultural, ética, metodológica e instrumental – que podem estar influenciando o desempenho escolar dos alunos. Além disso, tendo em vista a necessidade de se buscar, para os sistemas educacionais, maior eficiência e eficácia, sob qualquer que seja o enfoque dado a esses atributos, é importante medir o impacto e as interações entre essas características para cada sistema educacional específico. É preciso notar que os dados obtidos são estruturados hierarquicamente, isto é, alunos, turmas, escolas, constituem uma seqüência natural de agrupamentos aninhados, de tal forma, que as variáveis representativas das características nos diversos níveis podem interagir com outras variáveis dentro do mesmo nível hierárquico e, também, com variáveis de outro nível. Tendo em vista, principalmente, essa natureza dos dados, é freqüente a utilização de modelos de regressão hierárquicos (*cf.* Lee, 2001; Goldestein, 1995), que permitem investigar a influência das características de cada nível da hierarquia no desempenho escolar dos alunos e na diferenciação entre as escolas, por exemplo, e, ainda, separar a variabilidade nos resultados associada às escolas da variabilidade dentro de cada escola – associada aos alunos ou turmas de alunos. Muito comuns no exterior, estudos que empregam modelos de regressão hierárquicos estão sendo cada vez mais utilizados no Brasil desde os anos finais da década de 90. Entre esses trabalhos, neste artigo, faz-se referências aos seguintes.

Fletcher (1998) analisou em sua pesquisa educacional os efeitos das características do ambiente escolar e do ambiente familiar no rendimento dos alunos. Para o autor, as médias dos rendimentos dos alunos por escola, sem o ajuste das diferenças na composição social do alunado, distorcem os resultados das análises. Os dados utilizados por Fletcher (1998) foram os dados do SAEB/95 colhidos através da prova de matemática e de questionário, aplicados aos alunos de 8ª série. Barbosa & Fernandes (2000), utilizaram em sua pesquisa dados do SAEB-97, colhidos através de testes e questionários com o objetivo de estabelecer uma relação entre as variáveis explicativas de dois níveis (alunos e escolas) e o rendimento escolar dos alunos da 8ª série. O objetivo do estudo de Soares *et al.* (2001), foi o de conhecer o efeito das escolas de nível médio no vestibular da UFMG nos anos de 1998, 1999 e 2000 e ao mesmo tempo apresentar uma forma alternativa de avaliar os efeitos dessas escolas.

*Grosso modo*, o que se pode observar nesses trabalhos mais recentes é que os níveis de hierarquia considerados são, quase sempre, aluno e escola. Os resultados dessas pesquisas indicam que existe sempre uma influência de agrupamento, presente no sistema escolar, no modelo de impacto das variáveis sobre o rendimento escolar, reforçando a idéia de que os alunos não estão distribuídos aleatoriamente pelas escolas.

Neste estudo, constrói-se um modelo de regressão hierárquico para a população composta pelos alunos da 4ª série que participaram da avaliação de matemática e português realizada no ano de 2000 pelo SIMAVE/PROEB, num total de 2350 escolas, 5463 turmas e 152634 alunos. Este estudo foi planejado para servir de apoio e complementar, com informações de natureza qualitativa e quantitativa, as análises do relatório do SIMAVE-2000 (*cf.* Relatório Técnico do PROEB/SIMAVE-2000 (2001)), visando obter um maior conhecimento do sistema educacional mineiro, fornecer subsídios para futuras medidas de governo destinadas a reverterem o quadro de baixo rendimento acadêmico dos alunos do sistema estadual de educação. Além disso, estudos posteriores, mais específicos, também podem ser construídos utilizando-se dos resultados aqui encontrados. A proposta do presente trabalho é o de identificar a relação entre o desempenho escolar, medido pelo escore obtido nas provas do SIMAVE/PROEB-2000, e as características técnico-pedagógicas das escolas e sócio-econômicas dos alunos das 4ª séries do Ensino Fundamental, conforme medidas nos questionários apresentados aos alunos e professores das escolas durante a avaliação.

Na seção 2 apresenta-se a origem e a estruturação dos dados, discriminando-se as variáveis selecionadas para o estudo. A seção 3 traz uma introdução à formulação teórica para o modelo utilizado nas análises e na seção 4 apresenta-se os modelos construídos para o impacto, isoladamente, de cada variável sobre o rendimento escolar do aluno. Na seção 5 apresentam-se os modelos hierárquicos finais, o primeiro sem interações e o segundo considerando interações entre as variáveis, com a devida interpretação e análise dos resultados.

## **2. A Origem e a Estrutura dos Dados**

### **2.1 O PROEB/SIMAVE**

Em fevereiro de 2000, o governo do Estado de Minas Gerais, através da secretaria estadual de educação, instituiu o Sistema Mineiro de Avaliação da Educação Pública (**SIMAVE**). O SIMAVE se propõe a contribuir para uma nova cultura de avaliação educacional no estado, “compromissada com o sucesso escolar e com a educação de qualidade para todos” (*cf.* Relatório Técnico do PROEB/SIMAVE-2000 (2001)). Dentro dessa concepção está sendo implementado, a cada dois anos, o Programa de Avaliação da Rede Pública de

Educação Básica (**PROEB**), que integra o SIMAVE. Atualmente, o PROEB vem sendo operacionalizado pelo Centro de Políticas Públicas e Avaliação da Educação (**CAEd**) da Universidade Federal de Juiz de Fora (UFJF) – Minas Gerais. O PROEB é um programa de avaliação que tem por objetivo avaliar as escolas da rede estadual. O ciclo de avaliação se completa a cada dois anos. Em novembro de 2000, foram aplicados testes com o objetivo de avaliar as competências dos alunos em Língua Portuguesa e Matemática, em 2001 foram avaliadas as competências em Ciências Humanas e Naturais. Os testes são aplicados a todos os alunos da quarta e oitava séries do ensino fundamental e da terceira série do ensino médio da rede estadual. Além dos testes que avaliam as competências nessas disciplinas, o processo de avaliação inclui outros instrumentos importantes: um questionário aplicado aos alunos, com o objetivo de obter dados sobre o perfil sócio-econômico e trajetória escolar dos estudantes; um questionário aplicado aos professores e especialistas da escola, para se buscar informações sobre características desses profissionais e da escola. A metodologia empregada para avaliação pelo PROEB é, em geral, a mesma utilizada pelo Sistema (Nacional) de Avaliação da Educação Básica (**SAEB**) implementado pelo Instituto Nacional de Estudos e Pesquisa (**INEP**) do Ministério da Educação, e é adequada para avaliação em larga escala, com precisão muito maior na avaliação das proficiências médias de grupos de alunos do que, propriamente, na avaliação individual. Os itens (no caso, questões de múltiplas escolhas) desses testes foram construídos por especialistas tendo como base a proposta curricular de Minas Gerais e as matrizes de competências utilizadas pelo SAEB. Utiliza-se uma metodologia de construção de testes que é denominada de Blocos Incompletos Balanceados (BIB). A partir de um conjunto inicial de cerca de 1500 itens para cada disciplina e série avaliada, foram selecionados de acordo com suas características estatísticas, como as que medem a dificuldade e a capacidade de discriminação do item, um total de 169 itens. Esse total de itens é disposto em 13 blocos com 13 itens cada. Segundo uma combinação apropriada dos blocos, forma-se 26 cadernos de teste constituídos de 3 blocos cada. Cada aluno responde a um caderno, e, conseqüentemente, a 39 itens. A forma de construção dos cadernos faz com que haja blocos comuns entre eles produzindo, assim, o que se denomina de *equalização horizontal* dos escores estimados (a equiparação dos escores) pela calibração simultânea dos 169 itens – isto é, a estimação dos parâmetros dos modelos de comportamento impostos aos itens. O modelo que, geralmente, vem sendo utilizado na produção das proficiências é um modelo logístico de três parâmetros cuja construção é baseada na Teoria da Resposta ao Item (TRI) (cf. Hambleton *et al.*, 1991; Soares & Pereira, 2001; Lord, 1980). Esse tipo de modelo permite produzir “escores” de proficiências que sejam independentes dos testes aplicados e da população de examinandos (*ibidem*), estas são as principais razões pelas quais seu emprego vem se universalizando em avaliações educacionais. Assim, após análises estatísticas para verificar propriedades dos itens e dos testes, tais como, a *unidimensionalidade* (uma única proficiência sendo medida) do teste aplicado, ausência de comportamento diferencial do item entre as principais regiões do estado, ajuste dos dados empíricos ao modelo do item proposto – o que às vezes conduzem a eliminação de alguns itens, produz-se a chamada calibração dos itens (isto é, a estimação dos parâmetros dos modelos dos itens) e a estimação dos escores de **proficiência** (ou habilidade) de cada aluno. Para o presente estudo, as proficiências foram fornecidas pelo CAED\UFJF (maiores detalhes construtivos podem ser encontrados no Relatório Técnico do PROEB/SIMAVE-2000 (2001), pp. 2-20, PARTE 2).

As variáveis referentes à escola e ao aluno foram coletadas através dos questionários. Os dados da Avaliação fornecidos pelo CAED/SIMAVE/UFJF se encontravam em três arquivos: questionário do aluno, questionário do professor e arquivo com os escores de

proficiência em matemática e em português. Devido às restrições na aplicação dos questionários, não foi possível associar de forma satisfatória a informação de identificação do aluno à sua identificação nos arquivos dos escores de proficiências, exceto para uma amostra desses alunos. No entanto, a amostra referida pode não ser uma boa representante para o universo do Estado, perdendo qualidade no aspecto de aleatoriedade com que foi colhida. Então, neste trabalho, optou-se por agregar as informações ao nível de turma, pois, em grande parte, espera-se que a turma reflita propriedades individuais dos alunos, tendo em vista a semelhança entre os alunos de uma mesma turma, comum nos sistemas educacionais. Além disso, desconhecem-se estudos que tenham focado o rendimento médio da turma como uma variável dependente. Assim sendo, o modelo hierárquico construído neste estudo considera dois níveis: turma (1º nível) e escola (2º nível). As variáveis explicativas do modelo para o nível de escola foram obtidas através do questionário aplicado aos professores da escola. E, as variáveis explicativas para o nível de turma foram obtidas através do questionário aplicado aos alunos, agregados para cada turma.

Devido a existência de uma grande quantidade de variáveis (cerca de 180) nos bancos de dados, foi necessário selecionar um conjunto menor dessas variáveis que constituíssem uma *proxy* das principais características individuais dos alunos, agregadas ao nível da turma a qual ele está inserido, e das características das escolas.

## 2.2 As Variáveis Selecionadas

Uma variável que representa a condição sócio-econômica dos alunos, de impacto reconhecidamente importante no rendimento escolar, foi obtida através de um escore construído também, com auxílio da Teoria da Resposta ao Item, a partir de uma série de variáveis indicadoras da condição sócio-econômica, como por exemplo: “Número de quartos e banheiros na casa”, “número de geladeiras”, “número de televisores”, e assim por diante. Existem outros critérios para obtenção do índice sócio-econômico, por exemplo, através da variável “renda familiar” ou da variável “nível educacional” dos pais. Porém, geralmente, a maioria dos entrevistados – no caso presente alunos da 4ª série do ensino fundamental – desconhecem esses dados familiares. Por este motivo optou-se por utilizar os itens de conforto doméstico para obtenção do índice sócio-econômico do aluno, que por sua vez foi também fornecido pelo CAED (maiores detalhes construtivos podem ser encontrados no Relatório Técnico do PROEB/SIMAVE-2000 (2001), p. 7, PARTE 4). O anexo A traz algumas informações adicionais sobre o processo construtivo do índice sócio-econômico.

Após a construção do índice sócio-econômico do aluno, as 42 variáveis indicadoras da condição sócio-econômica foram substituídas por um único índice, e as demais variáveis provenientes do questionário do aluno e as variáveis do questionário do professor, disponíveis para a construção de um modelo que explique a proficiência de matemática, totalizaram 137 variáveis. Este número, ainda elevado, de variáveis inviabiliza a utilização adequada do modelo, além disso, como já discutido anteriormente um banco de dados desse tipo contém as mais diversas variáveis; entre as quais algumas fornecem pouca informação adicional além daquela fornecida por outras muito mais importantes na explicação da variável resposta, no caso, o escore obtido na prova de matemática aplicada através do SIMAVE. Uma maneira de tentar solucionar esse problema é utilizar a análise fatorial como uma ferramenta exploratória para redução do número de variáveis.

Um modelo de análise fatorial procura “explicar” as correlações entre um conjunto de variáveis observadas através de combinações lineares dessas variáveis com um conjunto

menor (desconhecido) de variáveis latentes (não observadas) denominadas de fatores. Admita que  $X_i$ ,  $i = 1, \dots, n$ , represente o conjunto total das variáveis observadas e que  $\theta_j$ ,  $j=1, \dots, d$ ,  $d < n$ , o conjunto de fatores correspondentes. O modelo de análise fatorial é então definido da seguinte forma (cf. Timm, 2002):

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \dots & \lambda_{1d} \\ \lambda_{21} & \lambda_{22} & \dots & \lambda_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{n1} & \lambda_{n2} & \dots & \lambda_{nd} \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_d \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_d \end{bmatrix} = A \theta + e$$

Os valores  $\lambda_{ij}$  são conhecidos como as cargas (do inglês “loadings”) associadas ao fator  $\theta_j$  e à variável  $X_i$ , sendo uma medida do grau de associação entre o fator e a variável.

Representa-se o vetor de dimensões latentes associadas por  $\theta$  e, por hipótese, admite-se que  $E(\theta) = 0$  e, ainda, que  $e \sim N(0, \Psi)$ , com  $\Psi$  diagonal. Dessa forma, sob essas hipóteses, é fácil mostrar que a covariância de  $X$  é dada por:

$$\Sigma = A \Theta A^t + \Psi,$$

onde  $\Theta$  é a matriz de covariância de  $\theta$ . Admitindo-se que os fatores são ortogonais, isto é,  $E(\theta_i \theta_j) = 0$  para  $i \neq j$  e, de variâncias iguais a 1, tal que  $\Theta = I$ , segue que  $\Sigma = A A^t + \Psi$ .

Nesse caso, mostra-se que a variância de cada variável  $X_i$  é dada por  $\sigma_i^2 = \sum_{j=1}^d \lambda_{ij}^2 + \psi_i^2$ ,

e a proporção dessa variância explicada pelos  $d$  fatores ( $h_i^2 = \sum_{j=1}^d \lambda_{ij}^2$ ) é chamada de *comunalidade* (do inglês “communality”).

Pode-se, *grosso modo*, utilizar esse tipo de modelo em dois tipos de análises: a exploratória e, ou, a confirmatória. Uma análise confirmatória é realizada quando se deseja testar estatisticamente uma hipótese específica quanto à estrutura do modelo fatorial. Por exemplo, uma análise de unidimensionalidade, isto é, a de que existe um único fator preponderante na explicação de um conjunto de variáveis. Como análise exploratória, a análise fatorial pode ser utilizada como uma ferramenta para se determinar, por exemplo, um número mínimo de fatores que retenha uma parcela “razoável” da variabilidade devida às variáveis originais. Essa forma de utilização da análise fatorial é, provavelmente, a mais comum, e está relacionada à redução do número de variáveis originais a um conjunto menor que será utilizado em análises subsequentes, como na construção de modelos de regressão. Conforme opção e interesse do pesquisador e da finalidade do estudo, pode-se ou utilizar os próprios fatores estimados (conforme um método apropriado) ou utilizar as variáveis mais fortemente associadas a cada fator, escolhidas, por exemplo, entre as que apresentam maiores valores de cargas associadas.

Tendo em vista que tanto as variáveis do questionário aluno quanto as variáveis do questionário do professor foram agregadas (ao nível de turma e ao nível de escola, respectivamente), representando médias, e ou percentuais, dos alunos nas turmas e dos professores nas escolas, sendo, portanto, variáveis reais, pôde-se proceder à análise fatorial se utilizando diretamente do modelo acima e dos métodos disponíveis no *software SPSS®*.

Assim, utilizou-se o método de extração por Componentes Principais – que é o procedimento mais utilizado na prática e é o método padrão no *software SPSS* (SPSS Inc., 1999) – com a matriz das componentes principais “rotacionada”, a partir do método *Varimax* de rotação. O método de extração por componentes principais é justificado, em linhas gerais, pelos seguintes argumentos. Sendo  $\Sigma$ , a matriz de covariância do conjunto de variáveis  $X$ , uma matriz quadrada simétrica, ela pode ser decomposta da seguinte forma:

$$\Sigma = \alpha_1 u_1 u_1^t + \alpha_2 u_2 u_2^t + \dots + \alpha_n u_n u_n^t,$$

onde  $\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_n \geq 0$  representa o conjunto de autovalores de  $\Sigma$  (lembrando que  $\Sigma$  é positivo semi-definida) e,  $U = [u_1 u_2 \dots u_n]$  o respectivo conjunto de autovetores associados, tais que  $UU^t = I$  (isto é, a base de autovetores associados é ortonormal). Esse resultado é conhecido como Teorema da Decomposição Espectral (cf. Strang, 1988, pp. 309-310). Assim sendo,  $\Sigma = L_n L_n^t$ , onde  $L_n = [\sqrt{\alpha_1} u_1 \dots \sqrt{\alpha_n} u_n]$ . Retendo-se as  $d$  primeiras componentes (principais) de  $L_n$ , isto é, tomando-se  $A = A_d \doteq [\sqrt{\alpha_1} u_1 \dots \sqrt{\alpha_d} u_d]$ , dentro do pressuposto de que as demais componentes podem ser desprezadas, o que é justificado pela propriedade seguinte.

Uma propriedade muito importante nessa representação é que a contribuição do  $j$ -ésimo fator sobre a variância total é  $\alpha_j$ . De fato, seja a  $\sigma_i^2$  variância da variável  $X_i$ , logo,

$$\sigma_i^2 = \sum_{j=1}^d \lambda_{ij}^2 + \psi_i^2 = \sum_{j=1}^d \alpha_j [u_j u_j^t]_{ii} + \psi_i^2, \text{ e, portanto, a contribuição do } j\text{-ésimo fator sobre a}$$

variância total  $\sum_{i=1}^n \sigma_i^2$  é dada por:

$$\sum_{i=1}^n \alpha_j [u_j u_j^t]_{ii} = \alpha_j \sum_{i=1}^n [u_j u_j^t]_{ii} = \alpha_j u_j^t u_j = \alpha_j.$$

Tem-se, portanto, a representação do modelo de análise fatorial através das  $d$  componentes principais de  $L_n$ , tal que  $\Sigma \cong A_d A_d^t + \Psi$ , onde  $\psi_i = \sigma_i^2 - \sum_{j=1}^d \alpha_{ij}$ , para  $i=1, \dots, n$ . O

método de extração por componentes principais consiste, então, em calcular os  $d$  maiores autovalores da matriz de covariância da amostra  $S$ ,  $[\tilde{\alpha}_1 \tilde{\alpha}_2 \dots \tilde{\alpha}_d]$  tomando-se como estimativa para  $A_d$  a matriz  $\tilde{A}_d \doteq [\sqrt{\tilde{\alpha}_1} \tilde{u}_1 \dots \sqrt{\tilde{\alpha}_d} \tilde{u}_d]$ , sendo  $[\tilde{u}_1 \tilde{u}_2 \dots \tilde{u}_d]$  a base de autovetores ortonormais associados.

A partir, ainda, da propriedade acima se pode definir, por exemplo, critérios para decisão quanto ao número de fatores e, conseqüentemente, um limite para a redução do número de variáveis. O *SPSS*, automaticamente, retém todos os fatores para os quais os autovalores correspondentes são superiores a 1. Esse critério, no entanto, é altamente subjetivo e, o critério mais utilizado consiste em extrair um número de fatores tal que o percentual da variância total explicada pelos fatores seja substancial, tipicamente de 50% a 70%, dependendo da aplicação. Neste trabalho, reteve-se um número de fatores tal que o percentual da variância explicada pelo modelo de análise fatorial correspondente a pelo menos cerca de 60% da variância total.

Em geral, o critério para manter as variáveis foi o de manter aquelas variáveis com cargas mais elevadas nos fatores extraídos. Além do padrão de correlações observadas, considerou-se também como critério de escolha de variáveis a importância apresentada por algumas variáveis em estudos correlatos. No caso do questionário dos alunos foram obtidos 10 fatores e no caso dos questionários dos professores foram obtidos 29 fatores, uma descrição mais detalhada desse procedimento e resultados pode ser encontrada em Mendonça (2002). A descrição dessas variáveis encontra-se na próxima seção.

### 2.3 Descrição das Variáveis

As variáveis do questionário do aluno que foram mantidas após o critério de seleção conforme descrito na seção anterior, e que serão empregadas na construção do modelo hierárquico são apresentadas e descritas na tabela abaixo:

**Tabela 1** – Variáveis selecionadas do questionário do aluno.

FATOR	VARIÁVEIS
Fator 1	<ul style="list-style-type: none"> <li>• Escore sócio-econômico produzido no SIMAVE. ( <b>E-SOCIO*</b> )</li> <li>• Grau de escolaridade do pai. ( <b>E-PAI</b> )</li> <li>• Grau de escolaridade da mãe. ( <b>E-MAE</b> )</li> <li>• Se o aluno sabe usar computador. ( <b>S-COMP</b> )</li> <li>• Número de livros que tem em casa. ( <b>N-LIVROS</b> )</li> <li>• Tipo de escola que já estudou. ( <b>T-ESC</b> )</li> </ul>
Fator 2	<ul style="list-style-type: none"> <li>• Número de anos que deixou de freqüentar a escola. ( <b>N-ABAND</b> )</li> <li>• Defasagem escolar, diferença entre a idade do aluno e a idade prevista para um aluno da 4ª série. ( <b>DEFAS</b> )</li> </ul>
Fator 3	<ul style="list-style-type: none"> <li>• Se o aluno teve boas notas no português neste ano. ( <b>T-BP</b> )</li> </ul>
Fator 4	<ul style="list-style-type: none"> <li>• Se o aluno mora com seus irmãos ou irmãs. ( <b>M-IRM</b> )</li> </ul>
Fator 5	<ul style="list-style-type: none"> <li>• Frequência com que o aluno vai a cultos religiosos. ( <b>F-CULTOS</b> )</li> </ul>
Fator 6	<ul style="list-style-type: none"> <li>• Se é de raça parda/mulato. ( <b>R-PARDA</b> )</li> </ul>
Fator 7	<ul style="list-style-type: none"> <li>• Se é de raça negra. ( <b>R-NEGRO</b> )</li> </ul>
Fator 8	<ul style="list-style-type: none"> <li>• Se o aluno gosta de português. ( <b>G-PORT</b> )</li> </ul>
Fator 9	<ul style="list-style-type: none"> <li>• Frequência de faltas dos alunos. ( <b>F-FALTAS</b> )</li> </ul>
Fator 10	<ul style="list-style-type: none"> <li>• Se o aluno gosta de matemática. ( <b>G-MAT</b> )</li> <li>• Se o aluno é do sexo masculino. ( <b>S-MASC</b> )</li> </ul>

\* Variável de turma Correspondente.

Todas essas variáveis são agregadas para a turma, tal que as variáveis que serão utilizadas na construção do modelo de regressão representam ou valores médios da turma ou, ainda, percentuais de determinadas características dentro da turma. Por exemplo, **E-SOCIO** representa o escore sócio-econômico médio dos alunos da turma. Por outro lado, **S-MASC** representa o percentual de alunos do sexo masculino da turma.

As variáveis do questionário do professor selecionadas para construção do modelo estão descritas na Tabela 2. Note-se que nesse caso a variável “renda familiar” foi utilizada como medida da condição sócio-econômica do professor, pois se espera que, neste caso, as respostas sejam mais confiáveis que as respostas dadas pelos alunos da 4ª série sobre a renda familiar de seus pais.



**Tabela 2** – Variáveis selecionadas do questionário do professor.

<b>FATOR</b>	<b>VARIÁVEIS</b>
Fator 1	<ul style="list-style-type: none"> <li>• Renda familiar. ( <b>R-FAM</b> *)</li> <li>• Nível de escolaridade. ( <b>N-ESC</b> )</li> </ul>
Fator 2	<ul style="list-style-type: none"> <li>• Se o professor é solteiro. ( <b>SOLTEIRO</b> )</li> </ul>
Fator 3	<ul style="list-style-type: none"> <li>• Frequência de pesquisas realizadas pelos alunos. ( <b>F-PESQ</b> )</li> <li>• Frequência de lição de casa. ( <b>F-LICAO</b> )</li> </ul>
Fator 4	<ul style="list-style-type: none"> <li>• Se o professor usa livros didáticos durante as aulas. ( <b>U-LIVROD</b> )</li> </ul>
Fator 5	<ul style="list-style-type: none"> <li>• Frequência de roubo dentro da escola. ( <b>F-ROUBO</b> )</li> </ul>
Fator 6	<ul style="list-style-type: none"> <li>• Frequência de realização profissional. ( <b>F-REAL</b> )</li> </ul>
Fator 7	<ul style="list-style-type: none"> <li>• Se o professor é de raça branca. ( <b>R-BRANCA</b> )</li> </ul>
Fator 8	<ul style="list-style-type: none"> <li>• Se o professor tem casa própria. ( <b>T-CASAP</b> )</li> </ul>
Fator 9	<ul style="list-style-type: none"> <li>• Número de salas na casa do professor. ( <b>N-SALAS</b> )</li> </ul>
Fator 10	<ul style="list-style-type: none"> <li>• Concordância com: A melhoria da educação depende só do governo. ( <b>C-MEDG</b> )</li> </ul>
Fator 11	<ul style="list-style-type: none"> <li>• Frequência da insuficiência de pessoal administrativo na escola. ( <b>F-INS_ADM</b> )</li> </ul>
Fator 12	<ul style="list-style-type: none"> <li>• Frequência de insuficiência de professores. ( <b>F-INS_PROF</b> )</li> </ul>
Fator 13	<ul style="list-style-type: none"> <li>• Se o professor avalia a participação do aluno em sala de aula. ( <b>A-PART</b> )</li> </ul>
Fator 14	<ul style="list-style-type: none"> <li>• Se o professor fez aperfeiçoamento. ( <b>F-APERF</b> )</li> </ul>
Fator 15	<ul style="list-style-type: none"> <li>• Se o professor fez grupo de estudos nos últimos 2 anos. ( <b>F-GRUPE</b> )</li> </ul>
Fator 16	<ul style="list-style-type: none"> <li>• Se o professor tem água encanada em sua casa. ( <b>T-AGUAENC</b> )</li> </ul>
Fator 17	<ul style="list-style-type: none"> <li>• Se a instituição em que formou é privada. ( <b>I-FORM</b> )</li> </ul>
Fator 18	<ul style="list-style-type: none"> <li>• Se o professor trabalha em escola estadual. ( <b>T-ESCEST</b> )</li> </ul>
Fator 19	<ul style="list-style-type: none"> <li>• Número de horas trabalhadas em sala de aula. ( <b>N-HORAS</b> )</li> </ul>
Fator 20	<ul style="list-style-type: none"> <li>• Se o professor fez doutorado. ( <b>F-DOUT</b> )</li> </ul>
Fator 21	<ul style="list-style-type: none"> <li>• Número de conselhos de classe nesta escola este ano. ( <b>N-CONS</b> )</li> </ul>
Fator 22	<ul style="list-style-type: none"> <li>• Se o professor avalia com provas individuais. ( <b>A-PROVIND</b> )</li> </ul>
Fator 23	<ul style="list-style-type: none"> <li>• Se o projeto pedagógico foi elaborado pela secretaria de educação. ( <b>PROJPED</b> )</li> </ul>
Fator 24	<ul style="list-style-type: none"> <li>• Frequência da exposição oral da matéria. ( <b>F-EXPORAL</b> )</li> </ul>
Fator 25	<ul style="list-style-type: none"> <li>• Se o professor usa laboratório de ciências durante as aulas. ( <b>U-LABCI</b> )</li> </ul>
Fator 26	<ul style="list-style-type: none"> <li>• Se o professor usa computador durante as aulas. ( <b>U-COMP</b> )</li> </ul>
Fator 27	<ul style="list-style-type: none"> <li>• Número de livros que o professor tem em casa. ( <b>N-LIVROS</b> )</li> </ul>
Fator 28	<ul style="list-style-type: none"> <li>• Se o professor exerce outra profissão. ( <b>OUTPROF</b> )</li> </ul>
Fator 29	<ul style="list-style-type: none"> <li>• Se o projeto pedagógico este ano foi elaborado pelo diretor. ( <b>PROJPEDD</b> )</li> <li>• Escore sócio-econômico dos alunos médio por escola. ( <b>E-SOCIOM</b> )</li> </ul>

\* Variável de Escola Correspondente.

De forma análoga ao caso das variáveis dos alunos, as variáveis dos professores são agregadas para representarem características das escolas, tal que as variáveis que serão utilizadas na construção do modelo de regressão representam ou valores médios da escola ou, ainda, percentuais de determinadas características dentro da escola. Por exemplo, **F-EXPORAL** representa a frequência média de exposição oral da matéria dos professores da escola e, **A-PROVIND** representa o percentual de professores da escola que avaliam os alunos com provas individuais.

Além das 48 variáveis selecionadas através da análise fatorial, foi escolhida, ainda, o escore sócio-econômico dos alunos médio por escola ( **E-SOCIOM** ) como uma variável ao nível de escola. Essa variável tem sido freqüentemente utilizada em estudos correlatos, e tem se mostrado de grande importância na construção dos modelos (ver, por exemplo, Fletcher, 1995). Portanto, ao todo, foram escolhidas 49 variáveis no total para construção do modelo.

### 3. Modelos Hierárquicos

Os sistemas escolares são um exemplo típico de estrutura hierárquica, pois os alunos estão agrupados em turmas, as turmas agrupadas em escolas, as escolas em uma determinada localidade, e assim por diante.

O modelo multinível (*cf.* Goldstein, 1995) também chamado de modelo hierárquico (*cf.* Bryk & Raudenbush, 1992), leva em consideração a estrutura de agrupamento dos dados. Concretamente isso se reflete na especificação do modelo multinível, por exemplo, da seguinte forma: para o modelo de regressão clássico o intercepto e o coeficiente de inclinação são parâmetros fixos enquanto que para o modelo multinível o intercepto e o coeficiente de inclinação são considerados parâmetros aleatórios, dependentes da influência do nível hierárquico mais alto.

As análises que consideram em seus modelos a estrutura de agrupamento dos dados têm várias vantagens: (i) se baseiam em modelos mais flexíveis e estruturados que utilizam melhor a informação presente na amostra, fornecem, ainda, uma equação para cada escola, por exemplo, o que permite análises individuais para cada grupo (ii) o uso da informação do agrupamento dos dados possibilita formular e testar hipóteses relativas a efeitos entre os níveis; (iii) permite a partição da variabilidade da variável resposta nos diversos níveis.

Os modelos de regressão multinível têm por objetivo descrever, através de um modelo estatístico, a relação entre variáveis explicativas e independentes, representadas genericamente por  $x$ , e uma variável dependente  $y$  (ou, mais de uma no caso de modelos multivariados). Neste trabalho só são considerados modelos com dois níveis hierárquicos. Assim, considere-se que, genericamente, uma amostra aleatória de dados tenha sido coletada a partir de uma estrutura em dois níveis, estando as unidades do 1º nível (turma) agrupadas segundo as unidades do 2º nível (escola). Cada unidade de turma é representada pelo índice  $i$  e, o índice  $j$  representa cada unidade de escola. Suponha que  $x$  represente uma variável de turma e  $w$  uma variável de escola. O modelo multinível então, terá a seguinte expressão geral:

$$y_{ij} = \beta_{0j} + \beta_{1j} x_{ij} + e_{ij} \quad (1)$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01} w_j + u_{0j} \quad (2)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} w_j + u_{1j} \quad (3)$$

Substituindo (2) e (3) em (1) obtém-se:

$$y_{ij} = \gamma_{00} + \gamma_{10} x_{ij} + \gamma_{01} w_j + \gamma_{11} w_j x_{ij} + u_{1j} x_{ij} + u_{0j} + e_{ij} \quad (4)$$

No modelo acima, os coeficientes apresentam a seguinte interpretação:

$y_{ij}$  representa a proficiência média da  $i$ -ésima turma da  $j$ -ésima escola;  $\beta_{0j}$  é o intercepto geral do modelo, sendo definido como variável aleatória;  $\beta_{1j}$  é o coeficiente de inclinação associado à variável  $x$ , representa o impacto da variável explicativa no rendimento médio da turma, e também é definido como variável aleatória;  $\gamma_{00}$ ,  $\gamma_{01}$ ,  $\gamma_{10}$ ,  $\gamma_{11}$  são parâmetros fixos a

serem estimados;  $u_{0j}$  é o denominado de efeito individual da escola, que é a componente de erro aleatório do nível 2 associada ao intercepto, pressupõe-se ter distribuição normal com média zero e variância  $\sigma_{u0}^2$ ;  $u_{1j}$  é a componente de erro aleatório do nível de escola associada ao coeficiente de inclinação, pressupõe-se ter distribuição normal com média zero e variância  $\sigma_{u1}^2$ ;  $e_{ij}$  é a componente de erro aleatório associado à turma, representa o resíduo da medida do rendimento da turma não explicado pelo modelo, pressupõe-se ter distribuição normal com média zero e variância  $\sigma_e^2$ ;  $\sigma_{u0}^2$ ,  $\sigma_{u1}^2$  e  $\sigma_e^2$  são denominados de componentes de variância do modelo.

Por hipótese, admite-se que o erro  $e$ , de nível de turma, seja independente dos erros de nível de escola. Note-se, ainda que  $\beta_{0l}w_j$  representa o impacto da variável explicativa de nível de escola no rendimento médio da turma e,  $\beta_{1l}w_jx_{ij}$  representa o termo de interação entre as duas variáveis explicativas (de turma e escola).

Na equação de regressão anterior poderão ser incluídas outras variáveis explicativas de nível de turma e também de nível de escola. A estrutura para o modelo resultante é análoga àquela apresentada através das equações de (1) a (4). A extensão do modelo multinível para outras variáveis permite obter o impacto das novas variáveis no rendimento escolar bem como obter outros termos de interação, alcançando uma maior diminuição da variabilidade total e, conseqüente aumento da capacidade de explicação da variável dependente pelo modelo resultante. Além disso, é possível analisar como as diversas variáveis se interagem e como seus impactos sobre a variável dependente se comportam na presença de outras variáveis.

Segundo Goldstein (1995), um indicador do grau de agrupamento da população em estudo é o coeficiente de “intra-correlação”. Para o modelo multinível especificado o que se mede com o coeficiente de intra-correlação é a proporção da variância total do resultado dos alunos que é devida às características das escolas. A forma mais simples para se obter o coeficiente de intra-correlação é a seguinte: primeiramente constrói-se um modelo multinível sem variáveis explicativas, também chamado de modelo nulo, que tem apenas três termos:  $\beta_0$ ,  $u_{0j}$  e  $e_{ij}$ , de tal forma que a variância total para a variável dependente nesse modelo é dada apenas por  $\sigma_{u0}^2 + \sigma_e^2$ . O coeficiente de intra-correlação, então, é calculado pela fórmula:

$$\rho = \frac{\sigma_{u0}^2}{\sigma_e^2 + \sigma_{u0}^2} \quad (5)$$

Esse coeficiente toma valores no intervalo [0,1] e, quanto maior o seu valor, maior a proporção da variância que é devida ao segundo nível. Seu cálculo é usado para justificar o emprego de um modelo multinível ao invés de um modelo de regressão clássico.

Neste trabalho, a estimação dos coeficientes fixos é realizada através do método de mínimos quadrados generalizados (cf. Bryk & Raudenbush, 1992) e, a estimação das componentes de variância é realizada através dos métodos de máxima verossimilhança plena e máxima verossimilhança restrita (*ibidem*). Para tanto, foi utilizado o *software HLM5*<sup>®</sup> (Raudenbush *et al.*, 2000). A medida de ajuste do modelo utilizada foi a chamada estatística de *deviance*, definida por:

$$D = -2 \text{ LOG } (L),$$

onde  $L$  é o valor da função de verossimilhança (maximizada segundo os valores dos parâmetros do modelo) nos valores observados da variável dependente e das variáveis explicativas (cf. Bryk & Raudenbush, 1992).

#### 4. Modelos de Impacto Direto das Variáveis

As pesquisas educacionais mais recentes procuram estruturar seus dados hierarquicamente para que seja possível investigar a influência das características de cada nível da hierarquia no desempenho escolar dos alunos e na diferenciação das escolas. No entanto, não existem referências na literatura brasileira sobre pesquisas que tenham utilizado as características da turma como nível básico de hierarquia. Nesse sentido, uma investigação que considera o nível de turma poderia ajudar a esclarecer se existe variabilidade no desempenho médio das turmas e, nesse caso, quais as características próprias desse nível de hierarquia estariam influenciando o desempenho delas mesmas e contribuindo para a diferenciação também das escolas.

Considerando o objetivo exposto e a característica dos dados educacionais do SIMAVE-2000, optou-se por utilizar um modelo de análise multinível com os níveis: turma (nível 1) e escola (nível 2). O modelo aqui adotado difere da maioria daqueles estudos que utilizaram modelos multinível, já que apresenta como unidade básica de análise a turma e não o aluno. Nesse sentido a contribuição do presente estudo é poder investigar: (i) a existência de variabilidade de desempenho médio da proficiência entre as turmas; (ii) o quanto dessa variabilidade, ainda assim, pode ser atribuída a características das escolas; (iii) se existe aleatoriedade de distribuição das turmas pelas escolas; (iv) quais os fatores, ou variáveis, têm impacto no rendimento médio das turmas; (v) qual o valor desse impacto no rendimento médio das turmas; (vi) e o comportamento, enquanto modelo de nível turma e escola, das variáveis cujo impacto no rendimento escolar foi comprovado por outros trabalhos, nos quais o nível básico considerado é o aluno.

O primeiro procedimento de análise do modelo multinível foi o de estimar o valor das variâncias dos erros aleatórios em cada nível de hierarquia através da construção do modelo nulo. Nesse caso, o coeficiente de intra-correlação encontrado foi de 0.182, no caso da proficiência de matemática – no caso da proficiência de português o valor foi de 0.191, indicando que cerca de 18% da variabilidade é devida a reais diferenças entre escolas e o restante da variabilidade, cerca de 82%, é devida a reais diferenças entre as turmas. Esse valor é inferior àquele observado nos trabalhos de Flether (1995) e Barbosa & Fernandez (1999) (0.312 e 0.37, respectivamente) principalmente pelo fato de os dados do SAEB incluírem escolas particulares o que, naturalmente, introduz maior variabilidade entre as escolas, mas tendo em vista, também, que esses trabalhos consideram como nível básico o aluno, e ao se agregar a informação por turma, boa parte da variabilidade devida à escola está sendo absorvida pela turma. De fato as variâncias estimadas para os erros de 1º e 2º nível, no modelo nulo, foram, respectivamente, 813,98 e 181,8 (ambas significativas pelo teste qui-quadrado correspondente para  $p < 0.001$ ). Esses resultados são suficientemente expressivos para se justificar o emprego do modelo hierárquico na análise. O valor da estatística de *deviance* para o modelo nulo foi de 39477 e, esse valor pode servir de comparação com os resultados alcançados nos demais modelos.

Numa análise preliminar é investigado o efeito de cada uma das 49 variáveis descritas na seção 3 sobre a proficiência de matemática (e português) isoladamente. Basicamente, os resultados para ambas as proficiências (matemática e português) foram os mesmos quando são considerados os resultados dos testes de significância (isto é, se o teste t indica correlação significativa entre a proficiência e cada uma das variáveis). Os resultados encontrados, no caso da proficiência de matemática, estão apresentados nas tabelas seguintes:

**Tabela 3** – Modelos de impacto direto de cada variável (variáveis de turma).

VARIÁVEL	$\gamma_{00}$	$\gamma_{10}$	$\sigma_e^2$	$\sigma_{u0}^2$	$\sigma_{u1}^2$	Deviance
E-SOCIO	149.07	0.50	691.90	161.02	0.01	38489
E-MAE	143.20	30.12	649.40	54.01	***	38125
E-PAI	149.43	26.62	698.20	231.24	22.46	38410
S-COMP	174.37	58.68	777.40	67.05	***	38826
N-LIVROS	171.89	0.73	762.20	90.50	***	38852
T-BP	116.90	90.17	724.10	170.40	***	38972
N-ABAND	209.10	-80.60	604.80	149.90	126.80	38216
DEFAS	217.24	-31.02	536.40	180.17	213.72	37882
T-ESC	184.76	58.70	799.40	113.17	***	39139
M-IRM	168.85	33.40	778.80	2720.00	3345.00	39333
F-CULTOS	105.73	34.90	610.48	202.18	***	38485
R-PARDA	194.00	5.30	809.80	186.20	1177.00	39334
R-NEGRO	209.98	-120.00	705.40	131.40	***	38748
G-PORT	220.05	-28.70	810.70	169.20	***	39310
F-FALTAS	***	***	***	***	***	***
G-MAT	190.30	12.11	832.60	200.16	***	39511
S-MASC	245.47	-94.20	649.40	222.50	***	38756

\*\*\* – valores não significativos

**Tabela 4** – Modelos de impacto direto de cada variável (variáveis de escola).

VARIÁVEL	$\gamma_{00}$	$\gamma_{01}$	$\sigma_e^2$	$\sigma_{u0}^2$	Deviance
R-FAM	126.92	18.89	808.08	142.43	39238
N-ESC	146.08	17.85	809.26	136.45	39224
SOLTEIRO	202.91	-22.10	807.45	169.28	39355
F-PESQ	166.75	13.72	813.60	170.81	39348
F-LICAO	173.42	6.45	814.53	174.48	39364
U-LIVROD	***	***	***	***	***
F-ROUBO	213.91	-10.40	817.06	154.07	39314
F-REAL	***	***	***	***	***
R-BRANCA	175.40	35.64	809.88	136.18	39225
T-CASAP	***	***	***	***	***
N-SALAS	***	***	***	***	***
C-MEDG	***	***	***	***	***
F-INS_ADM	***	***	***	***	***
F-INS_PROF	205.88	-6.35	815.26	175.01	39368
F-APERF	191.47	15.40	813.68	171.96	39352

VARIÁVEL	$\gamma_{00}$	$\gamma_{01}$	$\sigma_e^2$	$\sigma_{u0}^2$	Deviance
F-GRUPE	186.28	14.33	812.72	177.62	39362
T-AGUAENC	137.02	60.27	812.99	178.61	39364
I-FORM	***	***	***	***	***
T-ESCEST	***	***	***	***	***
N-HORAS	207.22	-0.42	813.49	180.62	39381
F-DOUT	***	***	***	***	***
N-CONS	***	***	***	***	***
A-PROVIND	177.85	20.58	814.05	178.11	39361
U-LABCI	195.51	18.92	812.85	179.39	39368
U-COMP	***	***	***	***	***
N-LIVROS	***	***	***	***	***
OUTPROF	***	***	***	***	***
PROJEDD	***	***	***	***	***
PROJEP	***	***	***	***	***
E-SOCIOM	155.60	0.425	798.85	50.26	38871
A-PART	141.06	56.95	814.52	176.90	39220

\*\*\* – valores não significativos

*Grosso modo*, os resultados são os esperados. Observa-se que as variáveis associadas à defasagem e ao abandono escolar são as que apresentam modelos mais explicativos (menor *deviance*). De fato, além de apresentarem um impacto negativo sobre a proficiência individual do aluno, o que já é esperado, parece que essas variáveis ganham maior importância, ainda, quando explicam a proficiência média da turma. Provavelmente, esse fato deve, em parte, estar associado à forma seletiva como as turmas são dispostas dentro da escola. Chama a atenção, o impacto negativo produzido pela maior incidência na turma de alunos de raça negra sobre a proficiência média. Esse fato poderia estar associado ao fato de esses alunos advirem de camadas da população mais pobre, mas na análise do modelo com várias variáveis os efeitos dos escores sócio-econômicos e da variável raça-negra são analisados conjuntamente, e como se verá, persiste o efeito negativo sobre a proficiência dessa variável. Maior incidência de alunos do sexo masculino em uma turma produz, também, um efeito negativo sobre a proficiência média da turma. Esse fenômeno necessita ser melhor avaliado, mas pode estar associado em maior intensidade às questões de comportamento dentro da turma do que, propriamente, às características individuais.

Com impacto positivo destaca-se o efeito da escolaridade da mãe sobre a proficiência, e surpreende a influência positiva da frequência a cultos religiosos.

A análise do impacto individual das variáveis é apenas um passo inicial na construção do modelo. Ela fornecerá um ponto de partida e uma ordem para inclusão de variáveis no modelo de regressão. Na próxima seção apresentam-se dois modelos hierárquicos de regressão, um com efeitos de interação entre variáveis e outro sem efeitos de interação, que permitem analisar a influência das diversas variáveis simultaneamente.

## 5. A Aplicação do Modelo Multinível Aos Dados do SIMAVE-2000

O processo mais utilizado na construção de um modelo hierárquico é do tipo “Bottom-up”, isto é, parte-se do modelo nulo e vai-se incluindo as variáveis segundo uma heurística específica. Normalmente, recomenda-se (cf. Hox, 2001) verificar as interações entre as variáveis após a última variável ter sido incluída no modelo.

No presente trabalho, o primeiro passo foi o de verificar a significância dos coeficientes (parâmetros fixos e aleatórios) para cada modelo individual das 49 variáveis, conforme apresentado na seção anterior. Numa etapa seguinte, após serem escolhidas as variáveis cujos coeficientes sejam significativos, o segundo critério adotado para priorizar a inclusão de variáveis a partir do modelo nulo foi o critério de menor valor para a *deviance*. A estatística *deviance* é uma medida do ajuste do modelo aos dados. Utiliza-se a *deviance* para comparar um modelo mais simples com um modelo mais geral. Normalmente, os modelos com a *deviance* mais baixa são melhores, no entanto, para testar se a diferença entre as *deviances* de dois modelos é significativa, ou não, emprega-se o teste de significância  $\chi^2$  com o número de graus de liberdade igual à diferença de parâmetros entre os dois modelos. Outro critério empregado para inclusão ou não de variáveis no modelo foi o critério AIC (Akaike, 1974, *apud* Hox, 2001). Este critério é utilizado para comparar modelos diferentes (normalmente aninhados) e é calculado a partir do valor da *deviance* adicionado a um fator que penaliza o número de parâmetros estimados. Segundo a sugestão de Hox (2001), este critério pode ser empregado para se decidir entre dois modelos hierárquicos. O critério AIC é dado por:

$$AIC = d + 2q \quad (6)$$

sendo que  $d$  é a estatística *deviance* e,  $q$  é o número de parâmetros estimados no modelo.

A vantagem do AIC em relação ao teste do  $\chi^2$  para testar a diferença entre as *deviances*, é que no critério AIC não é necessário estipular um nível de significância, o que é sempre um critério subjetivo. Nesse trabalho, ambos os métodos foram empregados e não houve conflito entre os resultados.

Todo essa seqüência de procedimentos descritos acima teve como objetivo a construção de um melhor modelo multinível. No entanto, essa abordagem não é a única possível, mas pretende-se que com o seu emprego possa se alcançar um modelo suficientemente informativo para fins de análise e decisões daí decorrentes. O processo é, portanto, iterativo, e em cada passo uma variável (do nível de turma ou escola) é introduzida ou não no modelo de acordo com os critérios descritos acima. Eventualmente, uma variável anteriormente introduzida pode ser excluída a partir do momento que outra “mais explicativa” seja introduzida. As possíveis interações entre as variáveis são testadas ao final do processo a partir daquelas variáveis que estão no modelo e outras que não se encontram no modelo.

Produziu-se neste estudo, dois modelos finais, nos quais se basearão as análises. O primeiro deles não contém termos de interação entre variáveis e, o segundo, foi obtido após se testar as interações possíveis. Os dois modelos são representados pela equação abaixo, no caso do modelo sem interações os valores dos coeficientes (de interações) são  $\gamma_{11} = \gamma_{91} = 0$ . As estimativas para seus coeficientes (assim como o erro-padrão de cada estimativa) são apresentadas na Tabela 5.

$$y_{ij} = \gamma_{00} + \gamma_{01} E\_SOCIOM + \gamma_{02} F\_LIC + \gamma_{03} F\_INS\_PROF + (\gamma_{10} + \gamma_{11} E\_SOCIOM) DEFAS + \gamma_{20} N\_ABAND + \gamma_{30} E\_MAE + \gamma_{40} E\_SOCIO + \gamma_{50} F\_CULTOS + \gamma_{60} S\_MASC + \gamma_{70} T\_ESC + \gamma_{80} G\_PORT + (\gamma_{90} + \gamma_{91} E\_SOCIOD) R\_NEGRO + u_{1j} DEFAS + u_{2j} N\_ABAND + u_{0j} + e_{ij}.$$

**Tabela 5** – Modelos finais (proficiências de matemática).

VARIÁVEL	PARÂMETRO	MODELO SEM INTERAÇÃO		MODELO COM INTERAÇÃO	
		COEFICIENTE	ERRO PADRÃO	COEFICIENTE	ERRO PADRÃO
<b>INTERCEPTO</b>	$(\beta_{0j})$				
Intercepto	$(\gamma_{00})$	152,02	11,80	141,39	11,43
A-PART	$(\gamma_{01})$	28,70	11,10	27,39	10,73
F-LIC	$(\gamma_{02})$	2,89	1,01	2,68	0,99
F-INS_PROF	$(\gamma_{03})$	-2,08	1,07	-2,00	1,05
<b>DEFAS</b>	$(\beta_{1j})$				
Intercepto	$(\gamma_{10})$	-8,40	0,88	7,27	1,78
E_SOCIOM	$(\gamma_{11})$	-	-	-0,18	0,02
<b>N-ABAND</b>	$(\beta_{2j})$				
Intercepto	$(\gamma_{20})$	-32,94	2,23	-31,84	2,22
<b>E-MÃE</b>	$(\beta_{3j})$				
Intercepto	$(\gamma_{30})$	9,53	0,97	8,12	0,96
<b>E-SOCIO</b>	$(\beta_{4j})$				
Intercepto	$(\gamma_{40})$	0,11	0,02	0,28	0,02
<b>F-CULTOS</b>	$(\beta_{5j})$				
Intercepto	$(\gamma_{50})$	12,03	0,97	11,73	0,96
<b>S-MASC</b>	$(\beta_{6j})$				
Intercepto	$(\gamma_{60})$	-36,45	3,16	-34,87	3,12
<b>T-ESC</b>	$(\beta_{7j})$				
Intercepto	$(\gamma_{70})$	8,07	2,76	***	***
<b>G-PORT</b>	$(\beta_{8j})$				
Intercepto	$(\gamma_{80})$	-17,51	2,48	-17,61	2,69
<b>R-NEGRO</b>	$(\beta_{9j})$				
Intercepto	$(\gamma_{90})$	-37,39	3,80	-29,53	4,25
E-SOCIOD	$(\gamma_{91})$	-	-	-23,05	6,05
	$\sigma_e^2$	345,12	-	334,62	-
	$\sigma_{u0}^2$ (intercepto)	42,90	-	43,38	-
	$\sigma_{u1}^2$ (DEFAS)	108,67	-	108,77	-
	$\sigma_{u2}^2$ (N-ABAND)	644,20	-	748,22	-
<b>DEVIANCE</b>		36122	-	36006	-

\*\*\* – valores não significativos

Analisando o modelo sem interações destacam-se, com impacto positivo sobre as proficiências em matemática (ambos os modelos finais foram testados na explicação das proficiências em português e os resultados das análises são os rigorosamente os mesmos), as variáveis do nível de turma E-MÃE (escolaridade da mãe), E-SOCIO (escore sócio-econômico)



e F-CULTOS (frequência a cultos religiosos), sendo que no caso das duas primeiras esse resultado já era esperado e, no caso desta última, embora presente em outros estudos similares, não deixa de suscitar conjecturas e explicações sociológicas das mais variadas. De fato, parece que essas variáveis afetam a explicação da proficiência a partir das características individuais dos alunos. Destacam-se, ainda com impacto positivo, as variáveis do nível de escola A\_PART (avaliação pelos professores da escola da participação escolar do aluno) e F\_LINC (frequência com que os professores da escola passam lições de casa), sendo variáveis interessantes para atuação e controle pelos órgãos públicos.

Como visto na seção anterior as variáveis DEFAS (média de defasagem escolar da turma) e N-ABAND (número médio de anos de abandono da escola na turma) são variáveis com alto impacto negativo sobre a proficiência média da turma. Esse impacto origina-se, certamente, das características individuais dos alunos, e que se refletem na turma devido à seletividade com essas turmas são construídas. Mas, também, pode refletir uma característica ruim de parte do sistema de ensino que, através da seletividade, despreza os alunos que possuem essas características. Estudos mais aprofundados devem ser construídos para corroborar ou não essa conjectura.

Com impacto negativo, destacam-se, ainda, as variáveis S\_MASC (percentual de alunos do sexo masculino na turma) e R-NEGRO (percentual de alunos de raça negra na turma). A primeira, provavelmente, reflete muito mais uma característica de turmas com maior incidência de alunos do sexo masculino do que, propriamente, características individuais e, também, pode ser objeto de estudos posteriores com vistas à intervenção por parte de gestores educacionais. A segunda variável não deixa de ser inquietante, tendo em vista, que o modelo final apresenta, também, outras variáveis, que medem a característica sócio-econômica da turma e, que estão diretamente associadas a essa variável. Explicações sociológicas podem ser formuladas, mas nesse nível de estudo seriam meras especulações e, portanto, vamos omiti-las. No entanto, a análise do modelo com interações traz interessantes informações adicionais sobre o impacto dessa variável sobre a proficiência.

Nas análises das interações entre as variáveis de turma e escola, encontraram-se interações entre a variável DEFAS e E\_SOCIOM (escore sócio-econômico médio da escola), indicando que E\_SOCIOM potencializa o efeito negativo da defasagem escolar, isto é, quanto maior for o escore sócio-econômico médio da escola maior é o impacto negativo sobre a proficiência média da turma produzido pela média da defasagem escolar da turma. Esse efeito foi confirmado numa análise paralela em que foram consideradas apenas as variáveis DEFAS e E\_SOCIOM e suas interações. Uma questão natural a ser colocada, e estudada em trabalhos posteriores, é se esse fenômeno estaria associado ao fato de escolas com maior nível sócio-econômico médio serem mais seletivas e, ou, serem mais elitistas.

A variável E\_SOCIOD é uma variável artificial de escola criada a partir da variável E\_SOCIOM. Ela é uma variável dicotômica que assume um valor 1 para cerca de 35% das escolas com valores de E\_SOCIOM mais elevados e 0 para as demais. O efeito da interação dessa variável com a variável R\_NEGRO, da mesma forma que no caso anterior, é o de potencializar o efeito negativo da maior incidência de alunos de raça negra sobre a proficiência média da turma. Conjecturas similares ao caso anterior podem ser colocadas e devem ser objeto de estudos posteriores.

A análise dos resíduos dos modelos finais confirma as suposições de normalidade dos resíduos, a homocedasticidade, e a desconexão entre os valores previstos pelo modelo e os resíduos do modelo.

## 6. Conclusões

Estudos, como o aqui apresentado, buscam obter conhecimento para subsidiar as soluções para problemas que são constantes preocupações de todos aqueles que estão envolvidos com sistema educacional. A utilização da turma como unidade de 1º nível mostrou que existe uma porcentagem da variação total do desempenho que é devido a diferenças entre as turmas e outra porcentagem da variação total que é devido a diferenças entre as escolas. Isto é, as características das escolas afetam o rendimento médio das turmas.

Alguns resultados já eram esperados, como o impacto positivo sobre a proficiência média de uma turma de alunos de variáveis que medem o nível de escore sócio-econômico, os anos de estudo da mãe, e o impacto negativo de variáveis que medem a defasagem escolar e anos de abandono da escola. Surpreende, em parte, o efeito positivo da variável que mede a frequência do aluno a cultos religiosos e, o efeito negativo da variável que mede incidência de alunos de raça negra em uma turma, principalmente, no caso dessa última, quando se considera um modelo que já leva consideração as condições sócio-econômicas dos alunos. Estes fenômenos devem ser estudados com mais cuidado em futuros estudos procurando, ainda, verificar possíveis causas para o que foi observado.

Outros estudos, aprofundando o que aqui foi iniciado poderia ser o da análise mais acurada do efeito diferenciado entre as escolas para as variáveis que medem a defasagem, os anos de abandono escolar e a incidência de alunos de raça negra, principalmente, buscando encontrar grupos de escolas para as quais esse efeito é minimizado. Ou, em outras palavras, buscar escolas que promovam maior “equidade” nos resultados das proficiências médias das turmas, estudando suas práticas pedagógicas e organização interna. Se essas escolas, além de promoverem maior “equidade” também promoverem maiores níveis médios de proficiência para seus alunos independentes de suas características sócio-econômicas, poderão ser consideradas como eficientes e servirem de modelos dentro de um sistema educacional. Como o estudo realizado neste trabalho teve como enfoque a 4ª série do ensino básico, seria interessante que estudos desse tipo também fossem realizados para a 8ª série e a 3ª série do 2º grau. De tal forma que seja possível verificar se os resultados aqui encontrados se mantêm para essas séries ou se há especificidades que os diferenciam.

## Agradecimentos

Os autores agradecem ao CAEd e à Secretaria de Educação do Estado de Minas Gerais pela cessão dos dados e apoio ao trabalho. Agradecem, também, aos comentários e sugestões dos três revisores anônimos.

## Referências Bibliográficas

- (1) Andrade, D.F.; Tavares, H.R. & Valle, R.C. (2001). *Teoria da Resposta ao Item: Conceitos e Aplicações*. ABE – Associação Brasileira de Estatística, 2000, São Paulo.
- (2) Barbosa, M.E. & Fernandes, C. (2000). Modelo multinível: uma aplicação a dados de avaliação educacional. Pontifícia Universidade Católica do Rio de Janeiro. Disponível em: <<http://www.puc-rio.br>>. Acesso em: dez 2000.
- (3) Barbosa, M.E. (1999). Modelo Multinível para Dados Discretos Longitudinais. Tese de Doutorado, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro.

- (4) Bryk, S.A. & Raudenbush, W. (1992). *Hierarchical Linear Models*. Sage Publications, Inc., Newbury Park, Califórnia.
- (5) Castro, M.H. (1998). Avaliação do Sistema Educacional Brasileiro, Tendências e Perspectivas. Avaliação de Políticas Públicas Educacionais. *Ensaio*, **20**, 303-364.
- (6) Fletcher, P.R. (1998). À Procura do Ensino Eficaz. Ministério da Educação e Cultura, Departamento da Avaliação da Educação Básica, Rio de Janeiro.
- (7) Firme, T.P. (1994). Avaliação: Tendências e Tendenciosidades. Avaliação e Políticas Públicas em Educação. *Ensaio*, **2**, 5-12.
- (8) Goldstein, H. (1995). *Multilevel Statistical Models*. 2ª edição, John Wiley & Sons, New York.
- (9) Hambleton, R.K.; Swaininathan, H. & Rogers, H. (1991). *Fundamentals of Item Response Theory*. Sage Publications, Inc., Newbury Park, Califórnia.
- (10) Hox, J. (2001). Multilevel Analysis of regression and Structural equation models. **In:** *Hierarchical models of Survey Data* [edited by J. Hox], 54<sup>th</sup> Summer Institute, Michigan.
- (11) Klein, R. (1997). Indicadores educacionais: disparidades regionais e sócio-econômicas no Brasil. **In:** *Avaliação e determinação de padrões na educação Latino-Americana. Realidades e desafios* [organizado por H. Bomeny], Fundação Getulio Vargas, Rio de Janeiro.
- (12) Lee, V.L. (2001). What are Multilevel Questions, and How might we explore them with quantitative methods? *Estudos em Avaliação Educacional*, **24**, 31-68.
- (13) Lord, F.M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale: Lawrence Erlbaum, New York.
- (14) Mendonça, M.C. (2002). Em Busca Pela Qualidade na Educação: Modelo Multinível Aplicado aos Dados do SIMAVE-2000. Dissertação de Mestrado, Faculdade de Educação, Universidade Federal de Juiz de Fora, Juiz de Fora.
- (15) Raudenbush, S.; Bryk, A.; Cheong, Y.F. & Congdon, R. (2000). *HLM5 – Hierarchical Linear and Nonlinear Modeling*. Scientific Software International, Inc., Lincolnwood, Illinois.
- (16) Relatório Nacional – PISA 2000 (2001). OCDE Brasília. <<http://www.pisa.ocde.org/NatReports/PISA2000/Brazilnatrep.pdf>>. Acesso em: jan 2001.
- (17) Relatório Técnico do PROEB/SIMAVE-2000 (2001). Minas Gerais: Avaliação da educação. Secretaria Estadual de Educação, Governo do Estado de Minas Gerais, Juiz de Fora.
- (18) Ribeiro, M. (1981). *História da Educação Brasileira: A Organização Escolar*. 3ª edição, Editora Moraes, São Paulo.
- (19) Soares, J.F.; Alves, M.T.G. & Oliveira, R.M. (2001). O efeito de 248 escolas de nível médio no vestibular da UFMG nos anos de 1998, 1999 e 2000. *Estudos em Avaliação Educacional*, **24**, 69-117.
- (20) Soares, T.M. & Pereira, D.R.M. (2001). Estudo de Adequação para Modelos da TRI Aplicado ao Caso do Ensino Fundamental da Micro-Região de Juiz de Fora em 1999. *Educação em Foco*, **6**, 91-108.
- (21) Wilson, D.T.; Wood, R. & Gibbons, R. (1998). *Testfact: Test Scoring, and Item Factor Analysis*. Scientific Software International, Inc. Lincolnwood, Illinois.

**Anexo A: Informações Adicionais Sobre a Construção do Índice Sócio-Econômico.**

As variáveis originais destinadas à produção do escore sócio-econômico foram *dicotomizadas* (isto é, transformadas em variáveis binárias). O primeiro passo na construção do índice foi o da análise da correlação bisserial (*cf.* Lord, 1980) para cada uma dessas variáveis. A correlação bisserial é uma medida clássica da capacidade de discriminação do item, representando o grau de correlação que cada item binário tem com o escore bruto calculado com todos os itens. Foi utilizado o *software testfact*<sup>®</sup>, e, excluídas as variáveis que apresentaram correlação bisserial inferior a 0,3. Todas as variáveis restantes apresentaram coeficiente estimado para a correlação bisserial superior a 0,50. Essas variáveis foram as seguintes:

**Tabela 6** – Variáveis indicadoras da condição sócio-econômica

Variável	DESCRIÇÃO
V1	Onde você mora existe água encanada?
V2	Onde você mora existe eletricidade?
V3	A rua em que você mora tem calçamento ou asfalto?
V4	A rua em que você mora tem coleta de lixo?
V5	A casa em que você mora tem cozinha?
V6	A casa em que você mora tem banheiro?
V7	A casa em que você mora tem sala?
V8	A casa em que você mora tem quarto?
V9	A casa em que você mora tem televisão a cores?
V10	A casa em que você mora tem geladeira?
V11	A casa em que você mora tem máquina de lavar roupa?
V12	A casa em que você mora tem aparelho de som?
V13	A casa em que você mora tem aparelho de videocassete?
V14	A casa em que você mora tem freezer?
V15	A casa em que você mora tem telefone?
V16	A casa em que você mora tem telefone celular?
V17	A casa em que você mora tem computador?
V18	A casa em que você mora tem automóvel?

Em seguida, uma análise da dimensionalidade associada foi produzida utilizando-se o método de análise fatorial *Full Bayesian Factorial Analysis* (*cf.* Wilson, Wood & Gibbons, 1998) em amostra de 1000 alunos. Através do critério  $G^2$  (*ibidem*), baseado num teste de razões de verossimilhanças, constatou-se a presença de dois fatores dominantes, e em função desse resultado as variáveis V1, V2, V4, V5 e V8 foram excluídas, com o que se alcançou um único fator dominante para as demais.

O índice final foi construído (usando-se o *software* BILOG-MG®) estimando-se para cada variável restante os parâmetros do modelo logístico (de dois parâmetros) da teoria da resposta ao item (*cf.* Hambleton, Swaminathan & Rogers, 1991; Andrade, Tavares & Valle, 2001; Lord, 1980) pelo método de verossimilhança marginal máxima e, obtendo-se o escore sócio-econômico através do método de verossimilhança máxima (*cf.* Andrade, Tavares & Valle, 2001). O modelo logístico de dois parâmetros é definido através da seguinte equação:

$$P_i(Y = 1; \theta, a_i, b_i) = \frac{e^{D a_i (\theta - b_i)}}{1 + e^{D a_i (\theta - b_i)}},$$

onde  $P_i(Y = 1; \theta, a_i, b_i)$  representa a probabilidade do indivíduo possuir o bem correspondente ao *i*-ésimo item (variável da Tabela 6),  $\theta$  representa o escore sócio-econômico do indivíduo, a ser estimado,  $a_i$  representa o chamado parâmetro de discriminação do item e,  $b_i$  representa o parâmetro de dificuldade do item. É fácil verificar que a probabilidade do indivíduo possuir o bem, correspondente ao item *i*, quando o seu escore sócio-econômico atinge o valor igual ao parâmetro  $b_i$  é igual a 0.5.