

UNIVERSIDADE FEDERAL DE MINAS GERAIS

Faculdade de Educação

Programa de Pós-Graduação em Educação: Conhecimento e Inclusão Social

**Qualidade e eqüidade na educação básica brasileira:
as evidências do SAEB 1995 – 2003**

Renato Júdice de Andrade

Belo Horizonte

Outubro de 2008

Renato Júdice de Andrade

Qualidade e eqüidade na educação básica brasileira:
as evidências do SAEB 1995 – 2003

Tese apresentada ao Programa de Pós-Graduação em Educação: Conhecimento e Inclusão Social da Faculdade de Educação da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Doutor em Educação.

Área de concentração: Políticas públicas e educação – formulação, implementação e avaliação

Orientador: Prof. Dr. José Francisco Soares

Belo Horizonte
Faculdade de Educação da UFMG
2008

Dedico este trabalho, com todo amor, ao meu filho Antônio, fonte de inspiração em todos os momentos, principalmente, quando seu sorriso, sincero e ingênuo, me roubava do computador, mas me proporcionava novo ânimo para continuar a caminhada.

AGRADECIMENTOS

Agradeço a todos que de alguma forma me ajudaram na realização deste trabalho, em especial a Deus, pois foi a quem recorri mais vezes pedindo ajuda e agradecendo pelas etapas vencidas;

Aos meus pais, pela torcida e pelos conselhos. Em um trabalho em que se discute, entre outras coisas, os fatores que impactam o desempenho dos alunos, não posso deixar de reconhecer o “efeito-pais” nesta conquista;

À Jeanne, esposa e companheira, que sempre esteve ao meu lado, nesta e em todas as outras empreitadas; às vezes, proporcionando um necessário apoio pedagógico, em muitas outras, dando um carinho essencial;

Ao meu orientador, professor José Francisco Soares, o Chico, que, como um verdadeiro amigo, sempre “me puxou para cima”. Obrigado pelas “portas abertas” em muitos projetos. Com ele aprendi muito do que sei hoje sobre avaliação, em especial, a principal delas: aprendi a gostar do tema;

Aos parentes, em especial, às minhas duas vovós Conceição, que, com certeza, rezaram muito para que eu chegasse até aqui;

Aos amigos, que compartilharam comigo dessa aprendizagem sobre avaliação, entre eles, o Daniel, a Cris, a Lenise, a Lili, a Luana, a Maria Alba, o Nilton, a Teresa Gonzaga, o Zé Luiz e tantos outros...

À Rosa, que, além de amiga e sócia na Quali-E, foi uma das grandes professoras a me ensinar que tão importante quanto a qualidade é a equidade;

Ao professor Sérgio Talim, pelas orientações e por ter sido o precursor dessa minha caminhada na avaliação;

Ao professor Nigel Brooke, não só pelo contínuo interesse que demonstrava pelo trabalho, mas também pelas dicas e pelo apoio, inclusive financeiro, no Projeto

Geres. A propósito, Projeto com o qual aprendi muito e ganhei muita experiência. E na pessoa do seu coordenador geral, Nigel, aproveito para agradecer a oportunidade de participação;

Aos amigos da Newton Paiva e do Colégio Arnaldo, alunos, funcionários e colegas de profissão, obrigado pelo incentivo;

Ao pessoal do GAME (Grupo de Avaliações e Medidas Educacionais), seja pelo apoio administrativo da Mary e do Glaydson, seja pelo bate-papo com os porteiros, seja pelo trabalho com os colegas, seja pelo cafezinho da Tereza;

À FAE (Faculdade de Educação da UFMG), casa que me acolheu desde o início da pós-graduação no CECIMIG até hoje, local onde aprendi muito, fiz amigos e, efetivamente, me tornei um profissional da educação. Faço, também, um agradecimento sincero ao pessoal da Secretaria da Pós;

Por fim, um agradecimento à Fundação Ford, à CAPES (através do seu programa de apoio denominado Observatório da Educação), ao Projeto GERES e ao GAME, entidades que, de uma forma ou de outra, financiaram este trabalho, permitindo a sua conclusão.

SUMÁRIO

Lista de Tabelas	
Lista de Gráficos	
Lista de Quadros e Figuras	
Resumo	
CAPÍTULO 1 – INTRODUÇÃO	16
1.1 APRESENTANDO O PROBLEMA DE PESQUISA	16
1.2 JUSTIFICATIVA	20
1.3 QUADRO TEÓRICO.....	22
1.4 O SAEB.....	31
1.5 METODOLOGIA DA PESQUISA	38
CAPÍTULO 2 – METODOLOGIA: DADOS, MEDIDAS E MODELOS	42
2.1 A UTILIZAÇÃO DA TEORIA DE RESPOSTA AO ÍTEM (TRI)	42
2.1.1 Modelo para o cálculo da proficiência.....	45
2.1.2 Modelo para o cálculo do nível socioeconômico.....	49
2.2 MEDIDA DA PROFICIÊNCIA.....	51
2.2.1 A escala de proficiência do SAEB	51
2.2.2 Interpretações pedagógicas diversas de uma mesma escala.....	57
2.2.3 Pontos de corte da escala e seus respectivos adjetivos – olhares diferenciados sobre o mesmo resultado.....	59
2.2.4 Descrição das proficiências do sistema educacional brasileiro no SAEB ...	65
2.3 MEDIDA DO NÍVEL SOCIOECONÔMICO (NSE)	69
2.3.1 O que é o NSE?.....	69
2.3.2 Construção da medida de NSE e seus indicadores.....	70
2.3.3 Descrição do NSE a partir dos dados do SAEB	77
2.3.4 Análise da correlação entre NSE e proficiência.....	79
2.4 MODELOS DE REGRESSÃO MULTINÍVEL	81
2.4.1 Necessidade de uso desses modelos em educação	81
2.4.2 O modelo escolhido.....	84
2.4.3 Como analisar os resultados do modelo.....	90
2.4.4 Os primeiros resultados do modelo	92
CAPÍTULO 3 – O EFEITO DA ESCOLA BÁSICA BRASILEIRA	96
3.1 INTRODUÇÃO AO PROBLEMA	96
3.2 O EFEITO DA ESCOLA	98
3.3 A HETEROGENEIDADE DAS ESCOLAS BRASILEIRAS.....	103
3.4 CONCLUSÕES: A ESCOLA PODE FAZER A DIFERENÇA	110
CAPÍTULO 4 - DIFERENÇAS ENTRE A ESCOLA PÚBLICA E A PRIVADA	115
4.1 CONTEXTUALIZANDO O PROBLEMA	115
4.2 O ALUNADO DAS ESCOLAS PÚBLICAS E PRIVADAS.....	117
4.3 A QUALIDADE DO PÚBLICO E DO PRIVADO.....	121
4.4 AS DIFERENÇAS EM RELAÇÃO À EQUIDADE SOCIOECONÔMICA	123
4.5 CONCLUSÕES: UMA LEITURA CRÍTICA A PARTIR DE UM PADRÃO DE QUALIDADE	125

CAPÍTULO 5 – O EFEITO DA COR/RAÇA NO DESEMPENHO DOS ALUNOS	127
5.1 O CONTEXTO DA QUESTÃO RACIAL	127
5.1.1 <i>Identificação da pertença racial</i>	128
5.1.2 <i>A classificação racial</i>	129
5.1.3 <i>Cor/raça dos alunos do sistema brasileiro de educação básica</i>	131
5.2 DIFERENÇAS MÉDIAS.....	133
5.3 AS DIFERENÇAS DEPOIS DO CONTROLE PELO NÍVEL SOCIOECONÔMICO.....	136
5.4 CONCLUSÕES: PRETOS E PARDOS SÃO NEGROS?	141
6 – REFERÊNCIAS	143
7 – ANEXOS.....	153
7.1 METODOLOGIA E SÍNTESE DO PLANO AMOSTRAL DE CADA UM DOS CICLOS PESQUISADOS DO SAEB	153
7.2 TRANSFORMAÇÃO LINEAR DE ESCALAS.....	159
7.3 MODELO DE REGRESSÃO MULTINÍVEL NO FORMATO PROC MIXED.....	160
7.4 SAÍDAS DO SPSS PARA O MODELO BASE.....	161
7.5 SAÍDAS DO HLM PARA O MODELO BASE – RACIAL.....	188

LISTA DE TABELAS

TABELA 2A: INTERPRETAÇÃO DOS NÍVEIS DA ESCALA SAEB DE MATEMÁTICA SEGUNDO O SAEB 2003.....	60
TABELA 2B: INTERPRETAÇÃO DOS NÍVEIS DA ESCALA SAEB DE LÍNGUA PORTUGUESA SEGUNDO O SAEB 2003.....	60
TABELA 3A: INTERPRETAÇÃO DOS NÍVEIS DA ESCALA SAEB DE MATEMÁTICA SEGUNDO O PROEB 2003.....	61
TABELA 3B: INTERPRETAÇÃO DOS NÍVEIS DA ESCALA SAEB DE LÍNGUA PORTUGUESA SEGUNDO O PROEB 2002.....	61
TABELA 4A: INTERPRETAÇÃO DOS NÍVEIS DA ESCALA SAEB DE MATEMÁTICA SEGUNDO O SAEMS 2003.....	63
TABELA 4B: INTERPRETAÇÃO DOS NÍVEIS DA ESCALA SAEB DE LÍNGUA PORTUGUESA SEGUNDO O SAEMS 2003.....	63
TABELA 5A: INTERPRETAÇÃO DOS NÍVEIS DA ESCALA SAEB DE MATEMÁTICA SEGUNDO O ATLAS DE MG.....	63
TABELA 5B: INTERPRETAÇÃO DOS NÍVEIS DA ESCALA SAEB DE LÍNGUA PORTUGUESA SEGUNDO O ATLAS DE MG.....	63
TABELA 6: PROFICIÊNCIA MÉDIA EM MATEMÁTICA POR SÉRIE E POR CICLO DO SAEB	65
TABELA 7: PROFICIÊNCIA MÉDIA EM LEITURA POR SÉRIE E POR CICLO DO SAEB.....	65
TABELA 8: PERCENTUAL DE ALUNOS POR NÍVEL DA ESCALA SAEB DE MATEMÁTICA, SEGUNDO O SAEB 2003.....	66
TABELA 9: PERCENTUAL DE ALUNOS POR NÍVEL DA ESCALA SAEB DE LEITURA, SEGUNDO O SAEB 2003.....	67
TABELA 10: DIFERENÇA DA PROFICIÊNCIA MÉDIA ENTRE SÉRIES, EM MATEMÁTICA, POR CICLO DO SAEB.....	68
TABELA 11: DIFERENÇA DA PROFICIÊNCIA MÉDIA ENTRE SÉRIES, EM LÍNGUA PORTUGUESA, POR CICLO DO SAEB.....	68
TABELA 12: DESCRIÇÃO DAS VARIÁVEIS UTILIZADAS NO CÁLCULO DO NSE.....	73
TABELA 13: MÉDIA DO NSE* POR SÉRIE E POR CICLO DO SAEB.....	77
TABELA 14: CORRELAÇÃO DE PEARSON ENTRE PROFICIÊNCIA E NSE.....	81
TABELA 15: DESCRIÇÃO DAS VARIÁVEIS INCLUÍDAS NO MODELO BASE.....	88
TABELA 16 – COEFICIENTE β_{00} AJUSTADO PARA O MODELO BASE – INTERCEPTO	92
TABELA 17 – COEFICIENTE β_{10} AJUSTADO PARA O MODELO BASE – VARIÁVEL ASSOCIADA: NSE DO ALUNO.....	93
TABELA 18 – COEFICIENTE β_{01} AJUSTADO PARA O MODELO BASE – VARIÁVEL ASSOCIADA: NSE MÉDIO DA ESCOLA.....	93
TABELA 19 – COEFICIENTE β_{50} AJUSTADO PARA O MODELO BASE – VARIÁVEL ASSOCIADA: ATRASO ESCOLAR DO ALUNO.....	94

TABELA 20 – COEFICIENTE β_{03} AJUSTADO PARA O MODELO BASE – VARIÁVEL ASSOCIADA: ATRASO ESCOLAR MÉDIO	94
TABELA 21 – COEFICIENTE β_{70} AJUSTADO PARA O MODELO BASE – VARIÁVEL ASSOCIADA: ATITUDE	95
TABELA 22: PERCENTUAL DE ESCOLAS, EM CADA BASE SAEB, COM U_{0j} (EFEITO DA ESCOLA) INFERIOR OU SUPERIOR A 20 PONTOS	101
TABELA 23: PERCENTUAL DE ALUNOS NO BRASIL COM DESEMPENHO ADEQUADO* À SUA SÉRIE EM MATEMÁTICA	114
TABELA 24: PERCENTUAL DE ALUNOS NO BRASIL COM DESEMPENHO ADEQUADO* À SUA SÉRIE EM LÍNGUA PORTUGUESA	114
TABELA 25: PERCENTUAL DA CLASSIFICAÇÃO DE COR/RAÇA SEGUNDO A AUTODECLARAÇÃO DOS ALUNOS NOS QUESTIONÁRIOS DO SAEB, POR DEPENDÊNCIA ADMINISTRATIVA*	119
TABELA 26: PROFICIÊNCIA MÉDIA EM MATEMÁTICA POR REDE	120
TABELA 27: PROFICIÊNCIA MÉDIA EM LÍNGUA PORTUGUESA POR REDE	121
TABELA 28 – COEFICIENTE β_{02} AJUSTADO PARA O MODELO (9) – VARIÁVEL ASSOCIADA: REDE	122
TABELA 29 – COEFICIENTE β_{12} AJUSTADO PARA O MODELO BASE – VARIÁVEL ASSOCIADA: REDE*NSE	124
TABELA 30: PERCENTUAL DA CLASSIFICAÇÃO DE COR/RAÇA SEGUNDO A AUTODECLARAÇÃO DOS ALUNOS NOS QUESTIONÁRIOS DO SAEB*	131
TABELA 31: DIFERENÇA ENTRE AS PROFICIÊNCIAS MÉDIAS EM MATEMÁTICA SEGUNDO A CLASSIFICAÇÃO DE COR/RAÇA DOS ALUNOS NOS QUESTIONÁRIOS DO SAEB	133
TABELA 32: DIFERENÇA ENTRE AS PROFICIÊNCIAS MÉDIAS EM LÍNGUA PORTUGUESA SEGUNDO A CLASSIFICAÇÃO DE COR/RAÇA DOS ALUNOS NOS QUESTIONÁRIOS DO SAEB	134
TABELA 33 – COEFICIENTE β_{30} AJUSTADO PARA O MODELO (10) – VARIÁVEL ASSOCIADA: <i>PARDO</i>	137
TABELA 34 – COEFICIENTE β_{40} AJUSTADO PARA O MODELO (10) – VARIÁVEL ASSOCIADA: <i>PRETO</i>	138

LISTA DE GRÁFICOS

GRÁFICO 1: CURVA CARACTERÍSTICA DE UM ITEM COM PARÂMETROS: $A = 1,5$; $B = 0,5$; $C = 0,2$	47
GRÁFICO 2: CURVA CARACTERÍSTICA DO ITEM M1D01N1	48
GRÁFICO 3: CURVA DE INFORMAÇÃO DO ITEM M1D01N1	48
GRÁFICO 4: CURVA CARACTERÍSTICA DO ITEM M1D02N3	48
GRÁFICO 5: CURVA DE INFORMAÇÃO DO ITEM M1D02N3	48
GRÁFICO 6: CURVA CARACTERÍSTICA DO ITEM 10.....	50
GRÁFICO 7: CURVA DE INFORMAÇÃO DO ITEM 10.....	50
GRÁFICO 8: PROFICIÊNCIA EM MATEMÁTICA, POR SÉRIE, NO SAEB 2003	56
GRÁFICO 9: CURVA CARACTERÍSTICA DO ITEM 9.....	75
GRÁFICO 10: CURVA DE INFORMAÇÃO DO ITEM 9.....	75
GRÁFICO 11: CURVA CARACTERÍSTICA DO ITEM 17.....	75
GRÁFICO 12: CURVA DE INFORMAÇÃO DO ITEM 17.....	75
GRÁFICO 13: CURVA CARACTERÍSTICA DO ITEM 12.....	76
GRÁFICO 14: CURVA DE INFORMAÇÃO DO ITEM 12.....	76
GRÁFICO 15: NÍVEL SOCIOECONÔMICO POR REGIÃO DO BRASIL.....	78
GRÁFICO 16: NÍVEL SOCIOECONÔMICO POR DEPENDÊNCIA ADMINISTRATIVA.....	79
GRÁFICO 17: PROFICIÊNCIA EM MATEMÁTICA NA 8 ^A SÉRIE DO ENSINO FUNDAMENTAL VERSUS NSE – NÍVEL DO ALUNO	80
GRÁFICO 18: PROFICIÊNCIA MÉDIA EM MATEMÁTICA NA 8 ^A SÉRIE DO ENSINO FUNDAMENTAL VERSUS NSE – NÍVEL DA ESCOLA	80
GRÁFICO 19: GRÁFICO HIPOTÉTICO RELACIONANDO DESEMPENHO EM MATEMÁTICA E NSE.....	91
GRÁFICO 20: EFEITO DA ESCOLA EM MATEMÁTICA, 8 ^a SÉRIE EF, PARA A BASE 2003 100	
GRÁFICO 21: INTERVALO DE CONFIANÇA PARA OS EFEITOS DAS ESCOLAS ESTIMADO PELO MODELO 3 SEGUNDO OS DADOS DO VESTIBULAR DE 2000.....	103
GRÁFICO 22: EVOLUÇÃO DA HETEROGENEIDADE DAS ESCOLAS, POR SÉRIE – REDE PARTICULAR, MATEMÁTICA.....	105
GRÁFICO 23: EVOLUÇÃO DA HETEROGENEIDADE DAS ESCOLAS, POR SÉRIE – REDE MUNICIPAL, MATEMÁTICA	106
GRÁFICO 24: EVOLUÇÃO DA HETEROGENEIDADE DAS ESCOLAS, POR SÉRIE – REDE ESTADUAL, MATEMÁTICA.....	106
GRÁFICO 25: NÚMERO DE MATRÍCULAS NO ENSINO MÉDIO, POR ANO E POR DEPENDÊNCIA ADMINISTRATIVA	107
GRÁFICO 26: EVOLUÇÃO DA HETEROGENEIDADE DAS ESCOLAS, POR REDE – 4 ^A SÉRIE EF, MATEMÁTICA	108
GRÁFICO 27: EVOLUÇÃO DA HETEROGENEIDADE DAS ESCOLAS, POR REDE – 8 ^A SÉRIE EF, MATEMÁTICA	109

GRÁFICO 28: EVOLUÇÃO DA HETEROGENEIDADE DAS ESCOLAS, POR REDE – 3 ^A SÉRIE EM, MATEMÁTICA.....	110
GRÁFICO 29: NSE MÉDIO DAS ESCOLAS PÚBLICAS E PRIVADAS NOS CINCO CICLOS DO SAEB.....	118
GRÁFICO 30: PERCENTUAL DE ALUNOS REGULARES NOS CINCO CICLOS DO SAEB – 4 ^ª SÉRIE EF	120
GRÁFICO 31: PERCENTUAL DE ALUNOS REGULARES NOS CINCO CICLOS DO SAEB – 8 ^ª SÉRIE EF	120
GRÁFICO 32: PERCENTUAL DE ALUNOS REGULARES NOS CINCO CICLOS DO SAEB – 3 ^ª SÉRIE EM	120
GRÁFICO 33: DISTRIBUIÇÃO ACUMULADA DAS PROFICIÊNCIAS – 4 ^ª SÉRIE EF - MATEMÁTICA.....	135
GRÁFICO 34: DISTRIBUIÇÃO ACUMULADA DAS PROFICIÊNCIAS – 4 ^ª SÉRIE EF - LEITURA	135
GRÁFICO 35: DISTRIBUIÇÃO ACUMULADA DAS PROFICIÊNCIAS – 8 ^ª SÉRIE EF - MATEMÁTICA.....	135
GRÁFICO 36: DISTRIBUIÇÃO ACUMULADA DAS PROFICIÊNCIAS – 8 ^ª SÉRIE EF - LEITURA	135
GRÁFICO 37: DISTRIBUIÇÃO ACUMULADA DAS PROFICIÊNCIAS – 3 ^ª SÉRIE EM - MATEMÁTICA.....	136
GRÁFICO 38: DISTRIBUIÇÃO ACUMULADA DAS PROFICIÊNCIAS – 3 ^ª SÉRIE EM - LEITURA	136
GRÁFICO 39: EFEITO DA ESCOLA POR EQÜIDADE RACIAL – 8 ^ª SÉRIE EF – SAEB 1995	139
GRÁFICO 40: EFEITO DA ESCOLA POR EQÜIDADE RACIAL – 3 ^ª SÉRIE EM – SAEB 1995	139
GRÁFICO 41: EFEITO DA ESCOLA POR EQÜIDADE RACIAL – 4 ^ª SÉRIE EF – SAEB 1997	139
GRÁFICO 42: EFEITO DA ESCOLA POR EQÜIDADE RACIAL – 8 ^ª SÉRIE EF – SAEB 1997	139
GRÁFICO 43: EFEITO DA ESCOLA POR EQÜIDADE RACIAL – 3 ^ª SÉRIE EM – SAEB 1997	139
GRÁFICO 44: EFEITO DA ESCOLA POR EQÜIDADE RACIAL – 4 ^ª SÉRIE EF – SAEB 1999	140
GRÁFICO 45: EFEITO DA ESCOLA POR EQÜIDADE RACIAL – 8 ^ª SÉRIE EF – SAEB 1999	140
GRÁFICO 46: EFEITO DA ESCOLA POR EQÜIDADE RACIAL – 3 ^ª SÉRIE EM – SAEB 1999	140
GRÁFICO 47: EFEITO DA ESCOLA POR EQÜIDADE RACIAL – 4 ^ª SÉRIE EF – SAEB 2001	140
GRÁFICO 48: EFEITO DA ESCOLA POR EQÜIDADE RACIAL – 8 ^ª SÉRIE EF – SAEB 2001	140
GRÁFICO 49: EFEITO DA ESCOLA POR EQÜIDADE RACIAL – 3 ^ª SÉRIE EM – SAEB 2001	140
GRÁFICO 50: EFEITO DA ESCOLA POR EQÜIDADE RACIAL – 4 ^ª SÉRIE EF – SAEB 2003	140

GRÁFICO 51: EFEITO DA ESCOLA POR EQÜIDADE RACIAL – 8ª SÉRIE EF – SAEB 2003	
.....	140
GRÁFICO 52: EFEITO DA ESCOLA POR EQÜIDADE RACIAL – 3ª SÉRIE EM – SAEB 2003	
.....	140

LISTA DE QUADROS E FIGURAS

QUADRO 1: OS OBJETIVOS GERAIS DO SAEB	33
FIGURA 1: MODELO CONCEITUAL EXPLICATIVO DA PROFICIÊNCIA	98
FIGURA 2: RELAÇÃO ENTRE DESEMPENHO EM MATEMÁTICA E NSE PARA ESCOLAS PÚBLICAS E CATÓLICAS NOS ESTADOS UNIDOS.....	116

RESUMO

Mesmo reconhecendo que a escola assume várias funções perante a sociedade, é natural optar por alguns indicadores específicos quando se deseja avaliar um sistema educacional, em especial, indicadores ligados ao ensino de competências cognitivas. Iniciativas como o Sistema Nacional de Avaliação da Educação Básica (SAEB) são exemplos de esforços para se medir o desempenho cognitivo dos alunos. Esse desempenho, medido pela proficiência do aluno em testes padronizados, é fruto de suas opções pessoais, de seus antecedentes sociodemográficos, da estrutura e valores de sua família e da sociedade em que vive e, finalmente, da escola em que estuda. Assim, ao se avaliar um sistema de ensino, deve-se, antes de tudo, ser capaz de identificar o impacto de cada um desses fatores sobre o desempenho dos alunos.

O objetivo deste trabalho é caracterizar a qualidade e a equidade do sistema de educação básica brasileiro, a partir das informações obtidas ao longo de cinco ciclos do SAEB. Propõe-se um cálculo integrado e simultâneo do nível socioeconômico de cada aluno que compõe essa base; uma medida do efeito da escola sobre o desempenho cognitivo de seus alunos; uma análise do impacto da dependência administrativa, se pública ou privada, e, finalmente, um estudo sobre a influência da cor/raça do aluno.

O estudo empírico usa os dados do SAEB, referentes aos ciclos de 1995, 1997, 1999, 2001 e 2003; da 4ª e da 8ª séries do ensino fundamental e da 3ª série do ensino médio, tanto em Matemática como em Língua Portuguesa. Chama-se a atenção para o fato de que analisar a presença de um fenômeno em uma base de dados é bem mais frágil do que examinar a continuidade desse fenômeno em 30 bases diferentes, sendo esse o principal diferencial deste trabalho. No entanto, como existem muitos determinantes, os efeitos supracitados só podem ser isolados e medidos com o auxílio de modelos estatísticos que controlam a influência dos outros fatores, e, para isso, neste trabalho serão usados os modelos hierárquicos de regressão.

Os resultados mostram o quão segmentado é o sistema de educação básica brasileiro, bem como a influência do nível socioeconômico no desempenho dos alunos; mostram a existência de escolas com efeito diferenciado e que, portanto, detêm informações sobre a gestão pedagógica de seus recursos humanos e físicos, que, se usados em um número maior de escolas, podem melhorar o sistema brasileiro de educação básica; mostram, também, que a rede pública e a rede privada atendem a um alunado completamente distinto, mas, mesmo depois do controle pelas variáveis

contextuais, a rede privada ainda se mostra melhor que a pública, tanto na qualidade quanto na equidade; e, por último, os resultados indicam a perversidade do sistema com os alunos que se autodeclararam pretos, e a inconsistência, do ponto de vista dos resultados educacionais, de se agrupar pretos e pardos em uma categoria única denominada negros.

Palavras-chave: SAEB; modelos hierárquicos; qualidade da educação; equidade; efeito da escola.

Capítulo 1 – Introdução

1.1 Apresentando o problema de pesquisa

São variadas as funções que a escola desempenha, principalmente nos dias de hoje, o que a faz ocupar papel de destaque na sociedade atual. A escola é um espaço de aquisição de competências e habilidades cognitivas, de formação de hábitos, de convívio social, entre outros. Nenhuma função é menos importante que outra, porém não há como negar a expectativa da sociedade em relação ao ensino e à aprendizagem. Nesse contexto, são relevantes os esforços de caracterização de sistemas de escolas quanto a sua eficácia para implementar um ensino de qualidade, propiciando aprendizagem a todos os seus alunos, bem como esforços para identificar escolas de sucesso e até mesmo para se determinar quais fatores impactam mais o desempenho cognitivo dos alunos.

Nas décadas de 50 e 60, importantes estudos empíricos, realizados nos Estados Unidos, Inglaterra e França, chegaram à conclusão de que a escola não faz diferença, ou, em outras palavras, os fatores externos impactam mais o desempenho e o acesso dos alunos à escola do que os fatores ligados diretamente às escolas. E, uma vez controlado o nível socioeconômico dos alunos, “as diferenças entre as escolas somente são responsáveis por uma pequena fração das diferenças no desempenho dos alunos” (COLEMAN, 1966, p. 131). Essas publicações tornaram-se famosas nos Estados Unidos, o estudo conhecido como “Relatório Coleman”; na Inglaterra, o “Relatório Plowden”; e na França, um estudo longitudinal realizado pelo órgão governamental de estudos demográficos, o INED. O interessado pode consultar, além de Coleman (1966), já citado anteriormente, Mosteller e Moynihan (1972), Nogueira (1990), Forquin (1995) e Bressoux (1994).

Tais estudos justificaram um certo “determinismo sociológico”, que começou a ser questionado com o trabalho de Rutter *et al.* (1979), cujas repercussões encontram-se em Sammons *et al.* (1995) e Lee (2000). A literatura brasileira sobre o assunto é mais recente: Soares (2004) e Soares, Alves e Mari (2003). Esses trabalhos evidenciam que existe “escola eficaz”, ou melhor, dentre os vários fatores que

influenciam o desempenho cognitivo dos alunos, pode-se encontrar também o “efeito da escola”. Como assinalaram Brooke e Soares (2008, p. 16), “a escola faz muita diferença para o aluno, o que Coleman mostrou é que, pelo menos nas dimensões incluídas no estudo, as escolas não são muito diferentes entre si”.

No Brasil, o sistema de educação básica tem melhorado, ainda que lentamente, especialmente quando se focam os indicadores de acesso e fluxo, mas o mesmo não pode ser dito sobre os indicadores de desempenho (KLEIN, 2006). Popularmente, tal situação seria descrita da seguinte forma: o Brasil está próximo de alcançar a escola para todos, mas ainda está longe de uma escola de qualidade para todos. E é nesse contexto de preocupação em se analisar o sistema educacional brasileiro e os fatores que afetam o desempenho dos alunos que surge esta pesquisa.

Mesmo reconhecendo a importância de se considerar vários indicadores, na linha proposta por Klein (2006), neste trabalho foi feito um recorte analítico. Propõe-se estudar a qualidade do nosso sistema educacional, tomando como base os dados de desempenho de seus alunos, aferidos pelo Sistema Nacional de Avaliação da Educação Básica (SAEB), executado pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), autarquia do Ministério da Educação (MEC).

O SAEB é a única fonte existente, que tem abrangência nacional além de permitir a comparabilidade dos resultados entre os diversos ciclos de aplicação. Daí a questão de pesquisa que se coloca: que evidências sobre a qualidade e a equidade do sistema brasileiro de educação básica podem ser extraídas dos dados do SAEB, entre os anos de 1995 e 2003?

Constata-se que há uma pequena exploração dos dados gerados pelos sistemas de avaliação no Brasil, e, quando ocorre, como no caso do SAEB, são estudos que, embora sejam importantes para o avanço do conhecimento da área, concentram-se na análise de um único fator, a partir da base de dados de um único ano. Barbosa; Fernandes (2001), Andrade; Franco; Carvalho (2003), Soares; Alves (2003) e Soares; Candian (2007) são exemplos desse tipo de publicação.

Ao explorar um número maior de fatores, ao longo de cinco ciclos do SAEB, este trabalho visa preencher essa lacuna. No entanto isso não significa que todos os problemas serão abordados, ao contrário, este campo de pesquisa ainda precisa ser muito explorado, principalmente, com estudos longitudinais, que são muito caros, mas que não sofrem com as limitações metodológicas das avaliações transversais.

O SAEB teve seu primeiro ciclo em 1990, porém foi somente a partir do terceiro ciclo, em 1995, que os avanços metodológicos permitiram a produção de dados

comparáveis. Por isso, apesar de reconhecer a importância histórica dos ciclos de 90 e 93, este trabalho utiliza somente as bases de dados correspondentes aos ciclos que ocorreram em 1995, 1997, 1999, 2001 e 2003. Na prática, isso significa um trabalho com 30 conjuntos de dados (5 x 2 x 3): os cinco ciclos; as duas disciplinas que permaneceram em todos esses ciclos¹, Língua Portuguesa (Leitura) e Matemática; e as três séries avaliadas, 4ª e 8ª séries do ensino fundamental e 3ª série do ensino médio. Chama-se a atenção para o fato de que analisar a presença de um fenômeno em uma base de dados é bem mais frágil do que examinar a continuidade desse fenômeno em 30 bases diferentes, sendo esse o principal diferencial deste trabalho.

Fazer inferências sobre a qualidade da educação, a partir dos resultados produzidos por alunos em testes padronizados, ainda não é um consenso na área. Muitos educadores sequer aceitam que o desempenho de um aluno ou que o impacto da escola sobre esse desempenho possa ser de fato medido. Este trabalho, no entanto, tem como premissas básicas a mensurabilidade do desempenho cognitivo e a necessidade de se basear as decisões políticas em resultados bem interpretados. Portanto as avaliações de sistemas, cada vez mais presentes em nosso dia-a-dia, tornam-se o contexto desta pesquisa, uma vez que são elas as geradoras desses resultados.

Esta investigação enquadra-se na linha denominada “estudos sobre o efeito da escola” e no Programa de Pós-Graduação em Educação: Conhecimento e Inclusão Social da Faculdade de Educação da UFMG na sublinha de pesquisa Políticas Públicas e Educação: Formulação, Implementação e Avaliação. Tem como objetivo geral analisar a qualidade e a equidade do sistema brasileiro de educação básica, a partir das evidências do SAEB. São objetivos específicos: a criação de um índice que caracterize o nível socioeconômico do alunado do sistema educacional brasileiro, correlacionando-o com o seu desempenho; o cálculo do efeito da escola, que caracteriza o nosso sistema educacional; o cálculo do impacto da dependência administrativa (público *versus* privado) sobre o desempenho; e o o cálculo do impacto da raça/cor sobre o desempenho dos alunos.

¹ As disciplinas de Língua Portuguesa e Matemática são as únicas que figuram em todos os ciclos. Nos dois primeiros ciclos, além dessas duas, foram aplicados também testes de Ciências. Em 1995, o conhecimento dos alunos foi avaliado apenas nessas duas primeiras disciplinas. Em 1997, fez-se uma nova tentativa de introduzir o teste de Ciências no ensino fundamental e no ensino médio, os estudantes responderam também questões de Física, Química e Biologia. 1999 foi o ciclo em que mais disciplinas foram avaliadas. Além de todas aquelas de 97, foram incluídas também História e Geografia. Em 2001 e 2003, apenas os testes de Língua Portuguesa e Matemática foram aplicados.

Para isso parte-se de um quadro teórico, que caracteriza o desempenho dos alunos como fruto de um complexo de fatores inter-relacionados, a fim de se determinar como e quanto cada um dos fatores impacta o desempenho por meio do ajuste de modelos matemáticos. Alguns desses fatores são internos à escola e outros externos. Ao longo do texto, os modelos adotados, bem como as escolhas dos fatores explicativos e de controle, serão trabalhados em cinco capítulos.

Ainda nesta introdução, discute-se a relação entre os conceitos medir e avaliar, mostrando que ambos estão relacionados, mas que a avaliação é algo bem mais abrangente. É necessário registrar que o foco deste estudo não é a avaliação, mas os resultados que são gerados em uma avaliação de larga escala. Por esse motivo, dedica-se uma reflexão breve sobre as avaliações sistêmicas como contexto da pesquisa. O SAEB, principal avaliação de larga escala brasileira, será descrito detalhadamente por se tratar da fonte de dados utilizada.

No segundo capítulo, faz-se a descrição das duas medidas centrais desta análise: a medida da proficiência dos alunos em Matemática e em Leitura e a medida do nível socioeconômico de cada aluno. A proficiência pode ser encarada como uma nota, mas não deve ser entendida como o percentual de acertos em um teste. É, na verdade, a medida de um traço latente do aluno no que se refere à sua competência em Leitura ou Matemática. O nível socioeconômico também é um traço latente do aluno, que se manifesta por meio de indicadores de escolaridade, ocupação e renda de sua família. Utilizando-se do avanço metodológico da Teoria de Resposta ao Item (TRI), pode-se locar o aluno em um *continuum* de hierarquia, seja de proficiência, seja social. Por isso também se faz uma breve discussão sobre a construção e interpretação de escalas.

No terceiro capítulo, estuda-se o efeito da escola no sistema educacional brasileiro, a partir da implementação de modelos hierárquicos de regressão linear. Esse efeito é estudado sob dois pontos de vista distintos: um primeiro, no qual o efeito da escola é calculado de forma individual, isto é, um valor para cada escola participante do estudo; e um segundo, no qual o efeito da escola é calculado como uma medida de grupo, isto é, que parcela da variação dos resultados dos alunos pode ser explicada pela variação existente entre as escolas. Tendo em vista as análises implementadas neste capítulo, propõe-se uma mudança de nome para a segunda forma de se calcular o efeito da escola, passando a denominá-lo 'medida da heterogeneidade das escolas'.

Nos capítulos 4 e 5, foram escolhidos dois fatores de grande importância para o contexto brasileiro, respectivamente, a dependência administrativa da escola na qual ele estuda (se pública ou privada) e a cor/raça do aluno, para se ajustar modelos hierárquicos apropriados e analisar os indicadores do SAEB para o sistema educacional brasileiro em relação a esses fatores. Nesses capítulos, é feita uma análise não só da qualidade do sistema como também da equidade do mesmo, termos que serão mais bem definidos mais adiante.

Pretende-se, neste trabalho, portanto, investigar as evidências que podem ser obtidas dos dados do SAEB, ao longo de cinco ciclos, possibilitando, assim, a caracterização do nosso sistema de educação básica. É mais do que uma análise do impacto de determinado fator sobre o desempenho dos alunos em um dado momento, mas um estudo sobre a evolução desse fator de 1995 a 2003 e uma reflexão sobre o impacto desses resultados na formulação de políticas públicas educacionais.

Como cada capítulo desta tese traz a investigação sobre um determinado fator, que envolveu uma discussão teórica e uma análise de dados, optou-se por apresentar as conclusões do trabalho, pertinentes a cada fator, no próprio capítulo, em vez de concentrá-las no final do texto. Embora os capítulos foquem fatores específicos, entende-se que o esforço de caracterização de um sistema educacional seja algo complexo, que deve envolver a análise de vários fatores e não ser fruto de análises isoladas. Assim, essa diversidade de fatores, necessária para que haja uma análise educacional adequada, foi considerada, por meio da implementação dos modelos multivariados, neste trabalho. Além disso, uma variedade de fatores, que ainda devem ser explorados, mostra as possibilidades de continuidade deste trabalho de pesquisa.

1.2 Justificativa

Este trabalho propõe uma análise dos dados do SAEB, nos ciclos de 1995, 1997, 1999, 2001 e 2003, inédita até então. Como citado anteriormente, outros trabalhos já foram publicados sobre os resultados produzidos em sistemas de avaliação, entre eles, o SAEB, no entanto tais trabalhos restringem-se à análise de dados relativos a um único ano.

Esta tese tem como principal justificativa a abrangência de sua análise, já que propõe o estudo tanto em Matemática quanto em Leitura, nas três séries avaliadas e ao longo de cinco ciclos distintos. Espera-se, portanto, que, ao fim deste trabalho, seja

possível tecer uma série de considerações a respeito do atual estágio da educação brasileira em relação ao desempenho cognitivo de seus alunos, além de fornecer uma gama de subsídios para que os gestores públicos da educação possam embasar suas decisões.

A pertinência do tema, em relação à legislação vigente, também contribui para a relevância da questão de pesquisa. A atual Lei de Diretrizes e Base (BRASIL, 1996), por exemplo, indica o estabelecimento de um sistema nacional de avaliação. Tratar de avaliação sistêmica, na atualidade brasileira, é sem dúvida relevante, seja pela magnitude e abrangência do tema, seja pela carência de amadurecimento da análise dos resultados produzidos, seja pela própria exploração dos dados gerados. E mais, deve-se ter em mente que as “avaliações em larga escala ultrapassam os limites das instituições educativas e, por meio dos veículos de comunicação de massa, envolvem parcelas muito amplas da sociedade” (FRANCO; BONAMINO, 2001, p. 27).

Ainda buscando um “apoio” legal para a pesquisa, encontra-se, na própria Constituição Federal, um argumento que justifica o reconhecimento da escola como um espaço de desenvolvimento cognitivo sem, no entanto, desprezar a importância do seu papel de propiciar a formação de hábitos e o convívio social. Em seu artigo 210, a Constituição Brasileira preconiza a existência de conteúdos mínimos a serem trabalhados: “Serão fixados conteúdos mínimos para o ensino fundamental, de maneira a assegurar formação básica comum e respeito aos valores culturais e artísticos, nacionais e religiosos” (BRASIL, 1999, p. 117), os quais foram, posteriormente, explicitados pela série ‘Parâmetros Curriculares Nacionais’ (PCN). Assim, infere-se que é lícito medir a qualidade de um sistema pelo desenvolvimento cognitivo de seus alunos, além do fato de que competências cognitivas, como o domínio da leitura, são essenciais para o aprendizado das competências sociais. Mesmo entre os críticos da avaliação educacional ou entre os que apontam uma série de dificuldades para a implementação de programas como o SAEB, há menção às vantagens de sua adoção. Maluf (1996) elenca como pontos positivos:

sistematização de informações para processos decisórios e planejamento; sensibilização de professores e diretores a respeito do processo ensino-aprendizagem e estímulo à introdução de mudanças em sua prática; prover os professores com informação sobre seus alunos; contribuição para instalação de uma ‘cultura avaliativa’; prestação de contas à sociedade (MALUF, 1996, p. 7).

Portanto, apesar das dificuldades para a implementação de um sistema de avaliação, seus resultados não podem deixar de ser estudados profunda e

sistematicamente. O Brasil está muito próximo de alcançar a meta 'escola para todos'. A próxima etapa seria a da 'boa escola para todos'. Almejar uma escola de qualidade, em que as desigualdades sociais não sejam acirradas, exige monitoramento permanente do desempenho dos alunos, de acordo com uma metodologia adequada. É nesse sentido que a implantação de um sistema de avaliação é necessária, tanto quanto o trabalho que deve ser feito com os seus resultados, que devem ser bem compreendidos e utilizados para que a formulação de políticas públicas educacionais contribua para a melhoria da educação como um todo.

1.3 Quadro teórico

A educação é um tema sempre contemporâneo e, talvez por isso, bastante complexo. A avaliação, pano de fundo deste trabalho de pesquisa, é apenas uma de suas muitas ramificações. Por isso, nesta seção, pretende-se desenvolver uma reflexão, partindo do geral para o específico, com o objetivo de se definir alguns conceitos que farão parte do arcabouço teórico. Inicia-se com uma discussão sobre o que é avaliar e as diferenças para o ato de medir; em seguida, abordam-se os diversos focos que uma avaliação pode assumir, em especial, a avaliação de sistemas, e conclui-se com os conceitos de qualidade e equidade, indicadores centrais que caracterizarão o desempenho dos alunos do sistema de educação básica brasileiro neste trabalho.

Apesar de não ser o foco da pesquisa, a avaliação é um tema correlacionado ao objeto do estudo. Constata-se que há resistências e críticas em relação à avaliação educacional e que não existe um consenso no meio acadêmico. Mesmo assim, a avaliação é tema de inúmeros projetos de pesquisa educacional. Perguntas como: por que avaliar? O quê avaliar? Quando? Para quê? Como?, certamente são formuladas nos projetos de pesquisa em educação. Afinal, na vida de um educador, boa parte do tempo é dedicada à avaliação de algum fator, entretanto, em que consiste o avaliar dentro do campo educacional? Seria a avaliação um acerto de contas, como certamente contestaria Vasco Moretto (2001)? Ou a avaliação em educação depende fortemente do contexto a que se propõe, e, portanto, estaria inserida numa dialética de duas lógicas, como sugere Perrenoud (1999)? Ou ainda, poder-se-ia concordar com Hadji (1994; 2001), quando ele afirma que avaliar consiste em se formular um juízo de valor sobre um objeto, por meio da comparação de uma realidade observável e de

uma realidade idealizada? Ou, então, seria possível até simplificar, nem por isso com grandes prejuízos, e afirmar, como Tyler, que “o processo de avaliação consiste essencialmente em determinar em que medida os objetivos educacionais estão sendo realmente alcançados pelo programa do currículo e ensino” (TYLER, 1981, p. 98)? Enfim, embora avaliar seja uma prática comum entre educadores e pesquisadores, sua definição não é consensual no meio acadêmico.

Esse é um tema polissêmico, o que não significa, entretanto, que não se deva tentar encontrar uma definição para ele. Neste trabalho, a avaliação será encarada como um processo, não só de monitoramento, mas, principalmente, de melhoria do desempenho cognitivo dos alunos. Para isso, entende-se que o sentido de se avaliar firma-se sobre três pilares: verificar alguma coisa – a medida; situá-la em relação a um nível, fazendo um julgamento de valor – a comparação; interpretar os resultados, seja no plano de gestão, seja no plano pedagógico – a explicação, isto é, avaliação é a leitura que se faz ao observar um mundo real, tendo em vista as concepções de um mundo ideal e as interpretações que são elaboradas com base nessa leitura.

De acordo com essa definição, avaliar implica em uma medida anterior, mas não se restringe a ela. A medida é especialmente necessária quando há um grande número de alunos, como no caso de um sistema educacional, porém, no decorrer deste trabalho, será evidenciado que os resultados de uma avaliação educacional não são consequência linear imediata dos insumos. A equação, nesse caso, é mais complexa e exige a observação de outros fatores, que não apenas as práticas internas da escola.

Segundo alguns autores, os dois conceitos, avaliar e medir, até se confundem, mas não são equivalentes. De acordo com Lindeman, pode-se afirmar que “a medida em educação é essencial para que o processo avaliativo seja levado a termo com precisão e eficiência” (LINDEMAN, 1976, p. 5). Franco (2004) faz uma distinção entre aferição e avaliação, afirmando que “aferição envolve mensuração de um construto; já avaliação é algo distinto, pois não há avaliação que deixe de envolver relações complexas entre diversos construtos” (FRANCO, 2004, p. 46). A medição não é, portanto, um fim em si mesmo, mas uma das partes das quais se constitui o processo de avaliação. Ou seja, mede-se para se obter informações a respeito de cada aluno em particular, ou de cada turma, ou de cada escola, ou mesmo para viabilizar a avaliação de um programa.

Outra visão é a que apresenta Silva (1992), na qual ela afirma que, em educação, a avaliação:

a) difere semanticamente de medida, porém a inclui como condição indispensável à sua objetividade e precisão; b) realiza-se em função de objetivos claramente definidos; c) é um processo amplo, voltado para diferentes aspectos da situação educacional: rendimento escolar de alunos, atuação de professores, supervisores e administradores, eficácia de programas, currículos, processos de ensino, etc.; d) fornece informações úteis para a tomada de decisões com relação a alunos e com relação a projetos e programas educacionais; e) é um processo contínuo, sistemático e integral (julga o aluno em todos os domínios do seu comportamento) (SILVA, 1992, p.15).

Dessa forma, medir e avaliar são conceitos diferentes, mas relacionados entre si. Avaliar é muito mais complexo. Medir, apesar de necessário, é apenas uma das etapas do processo. Como uma possível síntese, pode-se afirmar que medir está associado à coleta de dados, por outro lado, avaliar está associado não só à coleta, mas também às inferências, que podem ser feitas quando esses dados são analisados, bem como às decisões tomadas a partir dessas inferências.

Independentemente da importância que se dê à medida, o fato é que ela é essencial para os processos de avaliação em larga escala, no entanto deve ficar claro que, em qualquer processo de avaliação, é sempre o aluno que é avaliado, mesmo que a unidade de análise não o seja. O próprio SAEB exemplifica bem essa situação. São aplicados testes de Leitura e Matemática, juntamente com questionários contextuais, para os alunos, mas o foco da análise é o sistema educacional como um todo e não cada aluno individualmente. Essa distinção entre a unidade de coleta e a unidade de análise é uma primeira característica, que “pinta com cores fortes” as diferenças existentes entre a avaliação, que é feita em sala de aula, pelo professor, e as avaliações sistêmicas, feitas em larga escala.

Como em qualquer processo de medição, quando se fala em medida na área educacional, também é preciso abordar temas como a fidedignidade e a validade. A fidedignidade refere-se à precisão da medida. Um instrumento de medida é fidedigno, é preciso, quando apresenta uma coerência de resultados em várias mensurações sucessivas. A validade, por sua vez, refere-se ao fato de um teste estar medindo realmente aquilo que se pretendia medir. A validade não é um conceito absoluto, mas, ao contrário, é “um julgamento avaliativo integrado do grau em que evidências empíricas e raciocínios teóricos suportam a adequação e a apropriação de inferências e ações baseadas nos resultados de testes ou outros modos de avaliação” (MESSICK, 1993, p.13). Portanto o que é validado não é o teste ou o aparelho de avaliação, mas sim as inferências feitas a partir dos resultados desse teste ou de

outros indicadores. Ainda como afirma Messick (1993), a validade de um conjunto de evidências potencializa as conseqüências de se interpretar e utilizar os resultados desse conjunto, logo, a validade de um teste está associada ao grau em que ele é capaz de servir para determinados propósitos. E para que um teste seja válido, antes de qualquer coisa, ele deve ser confiável. Para clarear essa idéia, pode-se analisar o exemplo a seguir: um termômetro indica 1º C (um grau Celsius) a mais todas as vezes que é utilizado. Se o utilizarmos para os mesmos valores de temperatura, os resultados serão sempre iguais, porém as medidas estão erradas. Assim, pode-se afirmar que a fidedignidade é uma condição necessária, mas não suficiente para a validade.

Outros dois fatores relacionados à medida são, às vezes, tratados de forma superficial. O primeiro refere-se ao fato de que não se tem um acesso direto à competência do aluno, ou seja, a aprendizagem, a compreensão de determinado conteúdo, não é um fato diretamente observável, ela é formalizada como um traço latente do aluno. Ao fazer uma avaliação, são produzidos apenas indicadores. Assim, se o instrumento utilizado na avaliação for válido, podem ser feitas inferências sobre a aprendizagem do aluno. Moretto (2001, p. 28-29) exemplifica bem essa situação ao discutir que a performance é apenas “um indicador da possível competência do sujeito”. Portanto, segundo ele, o que se avalia em uma prova “é a performance do aluno, obtendo assim um indicador de sua competência”. Mas, assim como o time líder de um campeonato pode perder um jogo, em uma das várias avaliações durante o ano letivo, se o aluno obtiver “uma performance aquém do esperado, não significa, necessariamente, falta de competência” (MORETTO, 2001, p. 28-29). No entanto, o fato da nota aferida ser um indicador limitado não tira sua validade, muito menos sua utilidade prática.

Nas avaliações de larga escala, esse ‘traço latente’ é medido por um número que não é a quantidade de acertos em um teste. Para fazer a diferença, essa nota é normalmente denominada proficiência e pode ser definida como uma medida das competências já desenvolvidas pelo aluno. Deve-se observar que, neste caso, os dados realmente existentes são as respostas de cada aluno aos itens que lhe foram apresentados. São feitas inferências sobre essas respostas, por meio de um conjunto de modelos matemáticos, denominado Teoria de Resposta ao Item (TRI), para gerar a proficiência. Essa técnica permite que todos os alunos, apesar de responderem a diferentes itens, tenham suas respectivas notas colocadas numa mesma escala, denominada escala de proficiência. A construção e interpretação de escalas ainda é

um tema incipiente no cenário acadêmico brasileiro (SOARES; ANDRADE, 2005) e será mais detalhado no capítulo 2. No entanto, é necessário registrar que a dificuldade para se medir a competência do aluno, em um determinado assunto, é ainda maior no caso das avaliações de larga escala, se comparada à dificuldade encontrada pelos professores em sala de aula. Enquanto o professor convive cotidianamente com seus alunos, nas avaliações sistêmicas, normalmente, uma equipe externa à escola aplica um único instrumento, em um momento específico do ano letivo. Desse fato decorre a importância de se observar a qualidade da medida que é feita nessas avaliações.

A comparabilidade entre medidas distintas também é importante neste contexto de avaliações sistêmicas. Consegue-se tornar essa medida comparável entre diferentes testes ao colocar itens em comum entre eles (o mesmo raciocínio pode ser aplicado para tornar comparável a medida de duas séries diferentes ou até mesmo de avaliações aplicadas em dois anos diferentes). Ou seja, a implementação de uma medida na área educacional pressupõe, entre outras coisas, o domínio de metodologias estatísticas avançadas. No caso específico da TRI, pode-se consultar: Hambleton, Swaminathan e Rogers, 1991; Hambleton, 1993; Valle, 2000; e Baker, 2001.

O segundo fator relacionado às medidas e que, às vezes, é tratado também superficialmente é o fato de que os testes funcionam, em algum momento, como um instrumento de medição. Esse instrumento, como outro qualquer, apresenta uma imprecisão, que é inerente ao próprio ato de medir. Esse assunto deve ser tratado com cuidado, principalmente em avaliações de larga escala. A imprecisão na medida de proficiência média invalida o uso automático de qualquer lista de hierarquização de escolas, como detalham Goldstein (2001), Goldstein e Thomas (1996) e Goldstein e Spiegelhalter (1996).

Identificar e diferenciar os diversos tipos de avaliação educacional são tarefas das mais importantes, sobretudo para se evitar decisões precipitadas ou equivocadas a respeito de como avaliar e de como utilizar resultados de avaliação para tomar decisões que levem à melhoria do ensino e do desempenho dos alunos e escolas (OLIVEIRA; CHADWICK, 2006). É preciso repetir que, mesmo que o aluno seja sempre a unidade de coleta de dados, a unidade de análise pode variar. Por exemplo, um professor aplica testes diagnósticos para averiguar se seus alunos realmente aprenderam aquilo que ele se propôs a ensinar. Nesse caso, tanto a unidade de coleta quanto a de análise é o próprio aluno, por outro lado, um Estado avalia seus alunos para detectar se em cada escola os alunos aprenderam aquilo que os programas de

ensino da Secretaria determinam como relevante, nesse caso a unidade de coleta continua sendo o aluno, porém a unidade de análise é a escola. Nesta pesquisa, o foco se volta para uma avaliação na qual a unidade de coleta é o aluno e a unidade de análise é o sistema educacional brasileiro.

Tais avaliações sistêmicas, segundo Pestana, podem trazer “informações úteis sobre: o que estava sendo gerado no setor educacional, como, onde, quando e quem era responsável pelo produto obtido” (PESTANA, 1998, p. 16). A avaliação sistêmica é uma espécie de monitoramento, em que se coleta e organiza, periodicamente, uma série de informações, que sejam comparáveis sobre aspectos diversos do sistema educacional. E mais, os dados coletados em avaliações deste tipo podem ter profundo impacto nas decisões de políticas públicas para a educação, como sugere Vianna (2000). Em especial, o SAEB tem como objetivo explícito gerar e organizar informações sobre a qualidade, a equidade e a eficiência da educação nacional, permitindo o monitoramento das políticas brasileiras, fornecendo subsídios para a melhoria da qualidade, equidade e eficiência da educação no Brasil (BRASIL/MEC/INEP, 2001; BRASIL/MEC/INEP, 2002a; BRASIL/MEC/INEP, 2002b; PESTANA, 1998; FRANCO, 2001; FRANCO; BONAMINO, 2001).

Deve ficar claro, portanto, que uma avaliação sistêmica vai muito além de uma coleta de dados ou mesmo da relação professor-aluno. Avaliação de sistemas e avaliação de aprendizagem, feita em sala de aula, não são equivalentes. Primeiro, porque as do tipo sistêmica subsidiam decisões e políticas educacionais e, segundo, porque o número de alunos envolvidos é enorme, impedindo estratégias semelhantes às que são utilizadas em sala de aula para que o professor possa avaliar a aprendizagem de cada estudante. Na tentativa de encontrar uma definição, Ruben Klein e Nilma Fontanive poderiam ser lembrados, uma vez que eles afirmam que “a avaliação educacional é um sistema de informações que tem como objetivos fornecer diagnóstico e subsídios para a implementação ou manutenção de políticas educacionais” (KLEIN; FONTANIVE, 1995, p. 29), sem, no entanto, fornecer informações sobre alunos de forma individual. E mais, para que essa definição de sistemas de avaliação educacional seja completada, deve-se concordar que o principal objetivo da avaliação é justamente “a construção de uma dimensão pedagógica mais descritiva, que permita aos agentes do processo educacional, em diferentes níveis, não só entender como agir, tomando iniciativas para que a escola, efetivamente, possa atender aos seus alunos, respeitada a diversidade entre estes” (LOCATELLI, 2002, p.18). Ou seja, é um sistema de informação, que, mesmo sendo caracterizado

como “baixo impacto”², reflete na vida de toda uma comunidade, ou, como afirma Sobrinho (2002, p.14), a avaliação “ultrapassa os muros das salas de aula, se instaura também e decisivamente no centro do poder e espalha seus efeitos para toda a sociedade”.

Nesta pesquisa, assume-se a importância dos sistemas de avaliação para a melhoria da qualidade da educação, seja em relação aos sistemas particulares de rede de ensino ou sistemas públicos municipais, estaduais, nacionais, seja, até mesmo, em nível internacional. Arregui (2000) vai além e propõe que as avaliações de sistema podem ser o instrumento para garantir a qualidade e também a equidade no meio educacional. Esta postura baseia-se no pressuposto, defendido também nesta tese, de que não basta uma ‘boa escola’, é necessário uma ‘boa escola para todos’.

No entanto, a avaliação de um sistema educacional não se resume apenas aos indicadores de desempenho de seus alunos. Klein (2006) defende de maneira bastante clara que para se

analisar a qualidade da educação utilizam-se indicadores de atendimento (proporção da população em idade escolar matriculada), movimentação escolar (taxas de aprovação, reprovação e abandono durante o ano escolar) e fluxo escolar (taxas de promoção, repetência e evasão entre séries), taxas de acesso à escola, taxas de conclusão da 4ª série, do EF e do EM e indicadores de aprendizado baseados em escalas interpretadas de proficiência (habilidade) para algumas disciplinas, como, por exemplo, compreensão de leitura e matemática. Para julgar a qualidade é necessário definir padrões (ou valores mínimos) para esses indicadores. (KLEIN, 2006, p. 141)

Para exemplificar a importância de uma análise correta desses indicadores, Klein (2006) defende que para melhorar a qualidade da educação são necessárias, entre outras, políticas de correção de fluxo, objetivando-se baixar consideravelmente as taxas de repetência e evasão. No Brasil, os resultados disponíveis na década de 80, se analisados corretamente, indicavam problemas sérios na primeira taxa, porém as políticas públicas, implementadas com base nessas avaliações, atacavam apenas o segundo problema.

Ou seja, um sistema educacional é considerado de qualidade se (1) o acesso à escola está universalizado, isto é, uma considerável proporção de crianças e jovens, em idade escolar, está realmente matriculada na escola; (2) a conclusão dos seus

² Os termos *baixo impacto* e *alto impacto* referem-se à tradução de *low stakes* e *high stakes*, os quais, no cenário educacional, caracterizam os efeitos que uma determinada avaliação sistêmica pode ter sobre os indivíduos diretamente avaliados.

diversos níveis está universalizada, isto é, há um fluxo escolar estável, com baixas taxas de repetência, abandono e evasão; (3) níveis de desempenho, baseados em escalas interpretadas de proficiência, estejam sendo alcançados, isto é, os alunos estejam aprendendo os conteúdos apresentados; e, por fim, deveria se acrescentar aos critérios propostos por Klein (2006) mais um (4) referente à medida do aprendizado recente dos alunos, isto é, não basta ter um desempenho satisfatório, é preciso que o avanço na aprendizagem, ano a ano, também seja significativo.

No Brasil, vários estudos oficiais têm sido implementados a fim de medir a qualidade do nosso sistema educacional. Para verificar os dois primeiros pontos mencionados, há os levantamentos realizados pelo Instituto Brasileiro de Geografia e Estatística (IBGE), por meio do Censo Populacional e da Pesquisa Nacional por Amostra de Domicílios (PNAD), e pelo Ministério da Educação (MEC), por meio do Censo Escolar. Fletcher (2005), por exemplo, analisa a qualidade do sistema educacional brasileiro pelas suas taxas bruta e líquida de matrículas. Para verificar o terceiro item, no Brasil, dispõe-se do SAEB.

Em relação ao quarto critério proposto, é necessário registrar que a escolha do SAEB, como fonte de dados a ser analisada, se deu por conta da sua abrangência em todo o sistema educacional brasileiro e pela comparabilidade que se pode estabelecer. No entanto, é preciso reconhecer a fragilidade do SAEB como sistema de avaliação transversal. Essa é outra distinção, também sugerida por Franco (2004), e refere-se ao “aprendizado recente” e ao “conjunto de aprendizados”. O primeiro é a melhor medida para se analisar, por exemplo, o efeito da escola, mas é muito mais cara e mais difícil de ser obtida, pois demanda pelo menos um teste no início do processo e outro no término. Essa medida está, de fato, associada à aprendizagem do aluno e, conseqüentemente, muito mais associada ao valor agregado pela escola na qual ele estuda. Já o conjunto de aprendizados pode ser medido com um único teste ao final do processo; é o tipo mais comum nas avaliações sistêmicas brasileiras. No entanto, essa medida relaciona-se com o desempenho dos alunos e não com sua aprendizagem e, conseqüentemente, apresenta uma visão menos nítida quando se pretende obter o efeito da escola. Tal distinção precisa ser explicitada, pois a medida utilizada na análise dos dados desta pesquisa é o desempenho cognitivo dos alunos e não a sua aprendizagem. A medida do desempenho do aluno reflete toda a sua história naquela escola; em outra, na qual por ventura ele possa ter estudado; e até mesmo fora da escola. Para se ter uma medida mais próxima do que o aluno realmente aprendeu, durante um ano letivo em uma escola, seria necessário averiguar

de que ponto ele partiu, ou seja, o que ele já sabia no início do ano, ou então, utilizar outras formas de controle que assegurem, ainda que parcialmente, uma comparação entre grupos mais semelhantes de alunos e funcionem como uma *proxi* do desempenho inicial. Somente assim é prudente falar em ‘valor agregado’ pela escola ao conhecimento do aluno. Isso é conseguido de forma mais eficiente com uma metodologia mais complexa e muito mais cara, os estudos longitudinais de painel, nos quais se acompanha o mesmo aluno ao longo de um determinado tempo³. Na falta de estudos longitudinais, com abrangência nacional, elegeu-se o SAEB, mas todas as precauções estão sendo tomadas para não se atribuir resultados mais ambiciosos dos que esses dados podem fornecer.

Polissêmico, o conceito de qualidade da educação escolar ainda gera incompreensão, especialmente quanto ao seu valor no processo de democratização da escola (BARBOSA, 2005). Neste ponto, já se pode convergir para os conceitos de qualidade e equidade, tal qual serão utilizados neste texto. Os resultados que se deseja analisar estão expressos na escala interpretada de proficiências do SAEB, portanto a medida de qualidade do sistema de educação básica brasileiro está relacionada ao desempenho de seus alunos e fica, então, definida como sendo a parcela dessa proficiência do aluno, que foi produzida pela escola na qual ele estuda. Uma educação escolar de qualidade é aquela na qual se agrega valor ao desempenho esperado dos alunos.

Atualmente, a defesa não apenas da qualidade dos sistemas, mas também da busca por equidade é uma posição claramente defendida por pesquisadores em educação. Não basta que o sistema proporcione condições de alto desempenho aos seus alunos, é preciso que mudanças, implementadas para a melhoria do sistema, reflitam-se igualmente nos diversos ‘setores’, nos quais ele se subdivide, quais sejam, sociais, econômicos, raciais, de gênero etc. Equidade não é sinônimo de igualdade; a escola, na qual todos os alunos têm o mesmo resultado, é utopia. O real é composto de diferenças, devido à variação intrínseca do ser humano, mas o que é inaceitável é

³ Para os interessados em estudos longitudinais, sugere-se a consulta ao trabalho: ALVES, M. T. G. *Efeito das escolas e fatores associados ao desempenho de alunos de 5ª a 6ª série do ensino fundamental: um estudo longitudinal em sete escolas da rede pública no município de Belo Horizonte – MG*. Tese de doutorado. FaE/UFMG. Belo Horizonte, 2006. Ou, também, os documentos do Estudo Longitudinal da Geração Escolar de 2005, o Projeto GERES, uma iniciativa de pesquisadores em avaliação educacional do Brasil lotados na PUC-Rio, UEMS, UFBA, UFJF, UFMG e UNICAMP. O estudo se propõe a acompanhar a aprendizagem de mais de 20.000 alunos, em fase de alfabetização, em cinco cidades brasileiras, levando em conta os fatores escolares e sociofamiliares que incidem sobre o desempenho escolar, além de outras dimensões, como a auto-estima e a motivação, que podem afetar o desenvolvimento dos alunos. O sítio do projeto é www.geres.ufmg.br.

admitir que o sistema não precise ser justo. Dubet (2004) apresenta uma discussão pertinente sobre o que é uma escola justa. A medida de equidade relaciona-se com a capacidade que cada escola tem para acirrar ou amortecer as diferenças pré-existentes entre o seu grupo de alunos. A variação nos resultados de desempenho dos alunos sempre vai existir, portanto uma escola equitativa é aquela na qual o padrão da variação para a população se repete para os subgrupos, sejam eles sociais, raciais etc. Lee (2001b) coloca a equidade social como sendo relacionada à aprendizagem que não é dependente do contexto social dos alunos. Define-se, portanto, equidade como sendo uma medida da capacidade do sistema (ou de uma escola) de fazer com que o desempenho de seus alunos independa do seu próprio contexto ou de características pré-existentes.

1.4 O SAEB

Nas décadas de 60 e 70, observava-se uma preocupação com a ampliação do acesso ao ensino fundamental no Brasil. Paralelamente, as pesquisas na área educacional, no Brasil ou internacionalmente, justificavam o desempenho escolar, principalmente, com base em fatores externos à instituição (NOGUEIRA; CATANI, 1999). Na década de 80, as pesquisas internacionais passaram a considerar que a solução para os problemas da educação estariam ao alcance dos profissionais da área. Surgem estudos que enfatizam as características intra-escolares como um dos fatores explicativos do desempenho do aluno. Na América Latina, não foi diferente e várias reformas educacionais foram implementadas. O Brasil vivenciava um processo de redemocratização, que teve reflexos em todos os setores, inclusive, no educacional. Já naquela época, surgiam no país as primeiras discussões efetivas sobre a democratização e universalização do ensino. Nesse novo contexto, dois aspectos ganhavam destaque: o acesso à escola básica e a qualidade do ensino oferecido (BONAMINO; FRANCO, 1999).

Em relação ao primeiro aspecto, deve-se destacar o esforço e a eficiência dos governos para ampliar o número de escolas e de alunos que efetivamente freqüentavam as salas de aula, porém o segundo aspecto não gozava de ênfase semelhante. Na verdade, e esse é o ponto relevante desta reflexão, havia um sentimento, uma impressão geral, de que o ensino oferecido não possuía a qualidade esperada. No entanto, não havia, até então no país, como garantir qualidade, ou

melhor, não havia como medir essa qualidade do ensino oferecido naquela época. Ficou claro, em meio às discussões sobre a democratização do ensino, que, para se atingir a qualidade na educação, era necessário implementar um sistema que aferisse essa qualidade. E mais, “a persistência de altas taxas de repetência no país estimulou, já desde a entrada da década de 80, a realização de pesquisas e a formulação de políticas voltadas para o enfrentamento desse problema” (BONAMINO; FRANCO, 1999, p. 105-106).

Alguns estudos iniciaram-se nesta linha, sendo, então, lançadas as primeiras sementes das avaliações sistêmicas em solo brasileiro. Em trabalhos como o de Maluf (1996), tem-se acesso à história não só dos primeiros SAEB's como também do período anterior, referente aos estudos para implantação de tais sistemas.

No final da década de 80, ocorreram as primeiras ações voltadas para a implementação no Brasil de um sistema nacional de avaliação da educação básica, buscando verificar não apenas a cobertura do atendimento educacional oferecido à população, mas, principalmente, o desempenho dos alunos dentro do sistema (BONAMINO; FRANCO, 1999, p. 108).

Inicialmente, ele foi denominado Sistema Nacional de Avaliação do Ensino Público (SAEP), com um projeto piloto datado de 1988, realizado nos estados do Paraná e do Rio Grande do Norte. A falta de recursos, no entanto, só permitiu que ele fosse efetivamente implantado em 1990, ano da realização do primeiro ciclo do Sistema Nacional de Avaliação da Educação Básica, o SAEB.

Porém, como descreve Bonamino (2002), o seu desenvolvimento atravessou diferentes governos e conjunturas políticas.

É preciso ter em mente que o SAEB se desenvolveu ao longo de praticamente uma década, durante a qual diferentes conjunturas configuraram o contexto político dos anos 80 e 90 e concorreram para as mudanças verificadas no arranjo institucional e nas posições ocupadas pelos agentes responsáveis por este sistema de avaliação. O SAEB começou a ser desenhado durante o governo Sarney (1985-1990), por meio da primeira experiência com um sistema ampliado de avaliação educacional, ensaiada em 1988, atravessou os governos Collor (1990-1992) e Itamar Franco (1992-1994), para vir a consolidar-se durante a primeira gestão de Fernando Henrique Cardoso na presidência da República, a partir de 1995. (BONAMINO, 2002, p. 90)

Os objetivos do SAEB não ficaram fixos ao longo dos anos. Nos dois primeiros ciclos (1990 e 1993), a ênfase era em aspectos processuais. Nos ciclos que se

seguiram, a ênfase passou para a produção de resultados sobre o desempenho cognitivo dos alunos, visando ao monitoramento do sistema e o subsídio à formulação de políticas públicas educacionais (Bonamino; Franco, 1999). Essa ‘evolução’ nos objetivos do SAEB é sintetizada no Quadro 1

QUADRO 1: Os objetivos gerais do SAEB

Ciclo – Ano	Objetivos gerais
1º ciclo - 1990	Desenvolver e aprofundar a capacidade avaliativa das unidades gestoras do sistema educacional (MEC, secretarias estaduais e órgãos municipais); regionalizar a operacionalização do processo avaliativo, criando nexos e estímulos para o desenvolvimento de infra-estrutura de pesquisa e avaliação educacional; propor uma estratégia de articulação dos resultados das pesquisas e avaliações já realizadas ou em vias de implementação. (BRASIL/MEC/INEP, s.d., p. 3 <i>apud</i> BONAMINO; FRANCO, 1999).
2º ciclo – 1993	Fornecer elementos para apoiar a formulação, reformulação e monitoramento de políticas voltadas para a melhoria da qualidade da educação (Brasil/MEC/INEP, 1995 <i>apud</i> Bonamino; Franco, 1999). Promover o desenvolvimento e o aperfeiçoamento institucional, organizacional e operacional do Sistema Nacional de Avaliação da Educação Básica – SAEB; incrementar, descentralizar e desconcentrar a capacidade técnico-metodológica na área de avaliação educacional no Brasil. (BRASIL/PNUD, 1992 <i>apud</i> BONAMINO; FRANCO, 1999).
3º ciclo – 1995	Monitorar os resultados das políticas e estratégias educacionais, visando à melhoria da qualidade e do desempenho da Educação Básica do país. (BRASIL/MEC/INEP, 1998).
4º ciclo – 1997	Gerar e organizar informações sobre a qualidade, a equidade e a eficiência da educação nacional, de forma a permitir o monitoramento das políticas brasileiras. (PESTANA, 1998).
5º ciclo – 1999	Oferecer subsídios para a formulação, reformulação e monitoramento de políticas públicas, contribuindo, dessa maneira, para a universalização do acesso e a ampliação da qualidade, da equidade e da eficiência da educação brasileira. (BRASIL/MEC/INEP, 2000).

6º ciclo – 2001	Contribuir para a melhoria da qualidade da educação brasileira e para a universalização do acesso à escola, oferecendo subsídios concretos para a formulação, reformulação e o monitoramento das políticas públicas voltadas para a educação básica. (BRASIL/MEC/INEP, 2001).
7º ciclo - 2003	Acompanhar a qualidade da educação oferecida no Brasil, mediante a avaliação do conhecimento e das habilidades e competências adquiridas e desenvolvidas pelos alunos. (BRASIL/MEC/INEP, 2003a; 2003b; 2003c). Apoiar municípios, estados e a União na formulação de políticas que visam à melhoria da qualidade do ensino. (BRASIL/MEC/INEP, 2005).

Atualmente, os objetivos detalhados no sítio oficial do INEP são:

- *oferecer subsídios à formulação, reformulação e monitoramento de políticas públicas e programas de intervenção ajustados às necessidades diagnosticadas;*
- *identificar os problemas e as diferenças regionais do ensino;*
- *produzir informações sobre os fatores do contexto socioeconômico, cultural e escolar que influenciam o desempenho dos alunos;*
- *proporcionar aos agentes educacionais e à sociedade uma visão clara dos resultados dos processos de ensino e aprendizagem e das condições em que são desenvolvidas;*
- *desenvolver competência técnica e científica na área de avaliação educacional, ativando o intercâmbio entre instituições educacionais de ensino e pesquisa.*

As informações obtidas a partir dos levantamentos do SAEB também permitem acompanhar a evolução da qualidade da educação ao longo dos anos, sendo utilizadas principalmente pelo Ministério da Educação e Secretarias Estaduais e Municipais de Educação na definição de ações voltadas para a solução dos problemas identificados, assim como, no direcionamento dos seus recursos técnicos e financeiros às áreas prioritárias, visando ao desenvolvimento do Sistema Educacional Brasileiro e à redução das desigualdades existentes nele (BRASIL/MEC/INEP, 2005).

Franco (2001) afirma que os objetivos perseguidos pelo SAEB, ao longo dos anos, podem ser resumidos em três grupos: acompanhamento do sistema educacional; acompanhamento de como o sistema educacional vem se comportando em termos de equidade dos resultados educacionais; e apreensão dos fatores escolares, que podem explicar resultados escolares. Esses objetivos, de alguma forma, confundem-se com o foco deste trabalho de doutorado, além de serem fundamentais para as decisões do gestor de políticas públicas.

A metodologia utilizada pelo SAEB variou ao longo desses cinco ciclos. O ciclo de 1995, por exemplo, foi marcado por um grande número de mudanças, de avanços metodológicos, a saber: o ensino médio e a rede particular de ensino foram incluídos na amostra; técnicas estatísticas mais adequadas foram adotadas para a obtenção e análise dos dados; foram incorporados os questionários contextuais; e redefinidas as séries avaliadas. Porém vale destacar que, desde 1993, ele vem sendo aplicado de dois em dois anos, com algumas características importantes, tais como:

- Cada item dos testes é construído, a partir de orientações, na forma de descritores, que estão agrupadas em um documento denominado Matriz de Referência⁴ para o SAEB (BRASIL/MEC/INEP, 2001). A Matriz de Referência é, portanto, a primeira etapa a ser desenvolvida em uma avaliação sistêmica. Nesse caso, ela foi iniciada

com uma ampla consulta nacional sobre os conteúdos praticados nas escolas brasileiras de ensino fundamental e médio, incorporando a análise de professores, pesquisadores e especialistas sobre a produção científica em cada área que se tornou objeto de conhecimento escolar (BRASIL/MEC/INEP, 2001, p.11).

Essa inovação só foi incorporada ao SAEB em 1997 e reformulada em 2001, buscando-se incorporar a associação dos conteúdos às competências utilizadas no processo de construção do conhecimento. Mas é importante salientar que uma matriz de referência não define um currículo. Ela concentra o que os especialistas julgam como importante e o que é possível de ser medido em um teste deste tipo e, portanto, não pode ser confundida com procedimentos, estratégias de ensino ou orientações metodológicas. No caso do SAEB, é feita explicitamente, nos documentos oficiais e relatórios, uma opção teórica de natureza cognitivista, priorizando assim a medida do domínio de conteúdos na perspectiva das habilidades neles implícitas.

Na perspectiva cognitivista, o conhecimento não é 'cópia do real', isto é, o objeto não é compreendido em seu significado e sentido apenas porque é exposto ao sujeito. Nesta concepção está implícito o ato de raciocinar, coordenar as informações em consonância com as questões propostas e eventualmente produzir novas informações significativas e fazer inferências quando necessário (Brasil/MEC/INEP, 2001, p. 12).

⁴ Para mais estudos sobre Matriz de Referência, ler BRITTO, L. P. L., NÓBREGA, M. J. Matriz geradora de itens de avaliação: concepção e aplicações. *Revista de Estudos de Educação*, 2000.

- Em se tratando de uma avaliação desse tipo, o ideal é que um teste consiga produzir informações sobre grande parte do currículo vivenciado pelos alunos, porém, para medir uma determinada competência com maior precisão, é necessário um grande número de itens. Ou seja, um teste ideal exige grande número de itens e, portanto, não pode ser respondido por um só aluno, em uma única ocasião. Essa dificuldade é resolvida pelo SAEB com a adoção de uma técnica denominada Blocos Incompletos Balanceados (BIB)⁵, por meio da qual se constrói um certo número de cadernos de provas, contendo blocos de itens em comum. Com a adoção dessa técnica, o SAEB tem optado por aplicar, no total, 169 itens de cada disciplina, divididos em 13 blocos com 13 itens cada, agrupados, em seguida, de três em três blocos. Ou seja, o teste, como um todo, possui 26 cadernos diferentes, mas um aluno individualmente responde apenas um caderno com 39 itens.

- O instrumento cognitivo é composto por itens de Língua Portuguesa e Matemática, aplicado aos alunos na conclusão de etapas do seu percurso escolar, quais sejam, 4ª e 8ª séries do ensino fundamental e 3ª série do ensino médio. Esses itens são construídos por especialistas, tendo como base a lista de descritores da Matriz de Referência. As provas de Língua Portuguesa centram-se na leitura e compreensão de textos, abrangendo os seguintes tópicos: Procedimentos de leitura; Implicações do suporte, do gênero e/ou enunciador na compreensão do texto; Relação entre textos; Coerência e coesão no processamento de textos; Relações entre recursos expressivos e efeitos de sentido; e Variação lingüística. Já as provas de Matemática abordam os seguintes tópicos: Espaço e forma; Grandezas e medidas; Números e operações/ álgebra e funções; e Tratamento da informação.

- A coleta de dados não se resume apenas ao instrumento cognitivo. Questionários contextuais⁶ são aplicados com o objetivo de possibilitar uma análise de fatores – dos próprios alunos, professores, diretores e da escola – associados ao desempenho final. Em relação aos alunos, têm sido privilegiados os seguintes construtos: a caracterização sociodemográfica; o capital cultural; o capital social; a

⁵ Uma descrição detalhada desta técnica pode ser encontrada em BEKMAN, R. M. Aplicação dos Blocos Incompletos Balanceados na Teoria de Resposta ao Item. *Estudos em Avaliação Educacional*, São Paulo, n. 24, p. 119-135, jul./dez. 2001.

⁶ No artigo FRANCO, C. *et al.* O Referencial teórico na construção dos questionários contextuais do SAEB 2001. *Estudos em Avaliação Educacional*, São Paulo, n. 28, p. 39-74, jul./dez. 2003 pode-se encontrar mais informações sobre esse tipo de instrumento e o seu referencial teórico. Além dos construtos que são privilegiados e o modo pelo qual esses se relacionam com os itens dos questionários.

motivação e auto-estima; as práticas de estudo e a trajetória escolar. Em relação ao professor, são privilegiados os construtos: caracterização sociodemográfica dos professores; a formação; a experiência profissional; as condições de trabalho; o estilo pedagógico e a expectativa desses em relação aos alunos. Quanto à escola, os construtos são: a caracterização sociodemográfica do diretor; sua formação, experiência e liderança; as condições de trabalho do diretor e equipe; trabalho colaborativo; a organização do ensino e políticas de promoção; o clima acadêmico; o clima disciplinar; os recursos pedagógicos disponíveis; a situação das instalações e equipamentos e as atividades extracurriculares.

- Um avanço considerável na metodologia do SAEB foi implantado a partir de 1995, com a adoção da Teoria de Resposta ao Item (TRI). Além de gerar as proficiências dos alunos em cada teste, essa teoria permite os pesos corretos de cada item e a comparabilidade tanto entre séries quanto entre momentos diferentes da aplicação. Ao final, tem-se uma escala de proficiência que deve ser interpretada para descrever o desempenho dos alunos. Como há comparabilidade entre os testes de alunos oriundos de séries distintas, o SAEB assume a existência de uma competência única e que, portanto, permite locar, em uma única escala de proficiência, todos os alunos, independente da série⁷.

- Como o objetivo central do SAEB é informar sobre o sistema e não sobre cada escola individualmente, utiliza-se uma amostra representativa da população de alunos matriculados nas escolas brasileiras, é o chamado Plano Amostral do SAEB. Esse plano pode variar nos diversos ciclos, por isso, optou-se por descrever sucintamente as características do Plano Amostral do SAEB em 2001 (BRASIL/MEC/INEP, 2001). Inicialmente, selecionam-se as escolas, depois as turmas dentro das escolas e os alunos dessas turmas presentes no dia da avaliação. Esse Plano Amostral define cinco critérios de estratificação: o primeiro deles refere-se à série em que o aluno está matriculado (4^a e 8^a séries do ensino fundamental e 3^a série do ensino médio, totalizando três estratos nessa categoria); o segundo é a Unidade da Federação (todos os 26 estados mais o Distrito Federal, totalizando 27 estratos nessa categoria); o terceiro critério é a dependência administrativa das escolas (estadual, municipal e particular, totalizando mais três estratos); o quarto é a localização da

⁷ Mais esclarecimentos sobre o que é uma escala de proficiência, como ela é construída e como ela é interpretada, podem ser encontrados em VALLE, R. C. Construção e interpretação de escalas de conhecimento: Um estudo de caso. *Estudos em Avaliação Educacional*, São Paulo: Fundação Carlos Chagas, n. 23, jan./jul. p. 71-92, 2001.

escola, ou seja, se ela está localizada na capital ou no interior. Sendo do interior, em algumas situações, elas ainda se subdividiam em rural ou urbana; e o quinto critério de estratificação é o tamanho da escola (um grupo com até duas turmas da série pesquisada, denominadas de “pequenas”, e outro grupo com três ou mais turmas da série, denominadas de “grandes”). Combinados todos esses critérios, tem-se um Plano Amostral, em 2001, com 438 estratos possíveis. Porém vários desses estratos são vazios ou contêm um número muito baixo de alunos, o que gera a necessidade de se fazer agregações, reduzindo o número dos estratos de interesse.

No anexo 7.1, apresenta-se, sucintamente, a metodologia adotada em cada um dos cinco ciclos do SAEB, estudados nesta tese, extraída dos documentos publicados pelo INEP, no formato de CD-ROM, em novembro de 2006, por meio de sua Diretoria de Tratamento e Disseminação de Informações Educacionais.

1.5 Metodologia da pesquisa

Esta pesquisa deve ser caracterizada como um estudo quantitativo, realizado com uma base de dados secundários. A partir de uma análise detalhada dos resultados do SAEB, nos cinco ciclos entre 1995 e 2003, foram discutidas as inferências que podem ser produzidas sobre o sistema básico de ensino do Brasil e os fatores intra e extra-escolares que interferem no desempenho desses alunos. O estudo baseou-se nos relatórios oficiais do MEC, nas publicações acadêmicas pertinentes e, principalmente, na base de dados do INEP, tal como está registrada em local de acesso público, o Consórcio de Informações Sociais⁸ (CIS)⁹.

O interesse pelo problema de pesquisa surgiu da familiaridade do autor e de seu orientador com análises quantitativas. Como já citado no item 1.1, o que se constata é uma pequena exploração dos dados gerados pelos sistemas de avaliação

⁸ O Consórcio de Informações Sociais (CIS) resultou de um projeto do Núcleo de Apoio à Pesquisa sobre Democratização e Desenvolvimento da Universidade de São Paulo (NADD-USP) em parceria com a Associação Nacional de Pós-graduação e Pesquisa em Ciências Sociais (ANPOCS). O projeto está sendo desenvolvido sob a chancela de um convênio entre a Universidade de São Paulo e a ANPOCS e conta com o suporte financeiro da Fundação Ford e da CAPES. O sítio do Consórcio é www.cis.org.br.

⁹ Quando a pesquisa já estava sendo desenvolvida com a base de dados disponibilizada no CIS, o INEP publicou os Microdados do SAEB de 1995 a 2003 (Brasil/MEC/INEP, 2006). Optou-se por não refazer todos os procedimentos estatísticos da tese utilizando a nova fonte de dados, porém foi implementada uma minuciosa conferência de ambas as bases para se garantir que os dados trabalhados nesta tese correspondam ao Microdados publicados pelo INEP em 2006.

no Brasil. Este trabalho diferencia-se por explorar um número muito maior de bases de dados. Na prática, é um trabalho com 30 conjuntos de dados, que correspondem a cinco ciclos, duas disciplinas e três séries avaliadas. Para lidar com esses grandes números, que superam a casa dos 770.000 alunos, utilizou-se o software SPSS 13.0, que permitiu não só a manipulação da base de dados, como também a realização de análise estatística.

A proficiência e o nível socioeconômico (NSE) dos alunos são as duas medidas centrais desta pesquisa, como será explicado no próximo capítulo. A proficiência de cada aluno já foi disponibilizada, previamente, pelo INEP, mas o cálculo do NSE para cada aluno, em uma escala única, teve que ser implementado pelo próprio autor, utilizando-se a TRI. Tal tarefa se constitui como uma contribuição significativa para a área e, por isso, também, será detalhada no capítulo 2.

Uma vez que essas duas medidas centrais estavam disponíveis, iniciou-se a análise dos resultados tendo a proficiência como variável dependente e o NSE como uma das variáveis de controle. Os cálculos da qualidade e da eqüidade para o sistema foram construídos a partir da comparação dos estratos de interesse. Essas medidas têm como principal característica o fato de estar em um nível coletivo¹⁰. O desempenho analisado para cada estrato não foi apenas a 'média bruta', mas valores de proficiência corrigidos por fatores como o NSE de cada aluno e o NSE médio da escola na qual ele estuda. Para se entender a realidade apresentada pelo SAEB sobre a educação nacional, é preciso considerar as diferenças socioeconômicas entre os alunos, já que as desigualdades sociais no país são muito grandes.

A metodologia da pesquisa consistiu das seguintes etapas, repetidas para cada um dos três capítulos subseqüentes, com o seguinte padrão: (1) definição do fator a ser estudado; (2) leitura da literatura relacionada, especificamente, ao fator escolhido; (3) adaptação do modelo hierárquico base, que será implementado para análise do impacto desse fator no desempenho dos alunos; (4) implementação desse modelo nas 28 bases¹¹ disponíveis; (5) análise dos resultados obtidos.

¹⁰ Em hipótese alguma serão identificadas as escolas, construindo-se listas de hierarquização. Mesmo porque a metodologia do plano amostral do SAEB não permite tal ação.

¹¹ Como serão mais bem explicados no próximo capítulo, os questionários contextuais da 4ª série de 1995 eram muito precários e não possibilitaram a produção de uma medida de NSE confiável, por isso, tanto os alunos que realizaram o teste de Matemática quanto os alunos que fizeram o teste de Leitura, na 4ª série em 1995, foram retirados. Ou seja, a pesquisa foi realizada, na verdade, com um universo de 28 bases de dados do SAEB.

É necessário registrar que os modelos utilizados para análise do impacto dos fatores no desempenho dos alunos referem-se aos cálculos de regressão linear com modelos hierárquicos multiníveis (FERRÃO, 2003). A área educacional exige modelos de regressão hierárquicos tendo em vista a própria estrutura intrínseca dos dados. Isto é, os alunos estão organizados em turmas, que, por sua vez, estão agrupadas em escolas, que estão reunidas em regionais de ensino etc., uma estrutura que é, naturalmente, hierárquica. A escolha de um modelo hierárquico adequado é um passo crucial para a análise dos dados, principalmente, “nas ciências humanas, [na qual] as possibilidades de fazer experimentos controlados são muito mais restritas. Daí a necessidade de usar a estatística para, por assim dizer, simular os experimentos controlados” (SOARES; ALVES, 2003, p. 154).

Foram escolhidos três fatores para análise, não só pela sua representatividade na literatura sobre o tema, mas também pela expectativa de impacto no desempenho dos alunos: o efeito da escola sobre o desempenho dos alunos, a cor/raça dos alunos e a dependência administrativa da escola na qual ele estuda. Toda metodologia de uma pesquisa é um conjunto de decisões. E, uma vez que se priorizam três fatores para análise, outros ‘n’ fatores, não menos importantes, estão sendo deixados de lado. As etapas de trabalho para cada fator são complexas e, por isso, vários temas ainda se encontram ‘em aberto’ para outras iniciativas de pesquisa, tais como questões ligadas ao aluno, por exemplo, o gênero, a defasagem idade-série, o início da escolarização em ambientes formais de educação, o seu capital cultural, a participação e envolvimento dos pais nos estudos dos filhos; questões ligadas ao professor, por exemplo, o salário, a sua formação, o seu tempo de trabalho na profissão, a sua carga de trabalho; e questões ligadas à escola, por exemplo, infraestrutura, rotina de trabalho, seleção de alunos; além de várias outras questões possíveis de serem analisadas com essa metodologia, utilizando-se das mesmas bases de dados¹².

Ao final, espera-se que este trabalho possa ter respondido, de forma satisfatória, a pergunta de pesquisa proposta: “quais as evidências sobre a qualidade e a equidade do sistema brasileiro de educação básica podem ser extraídas dos dados do SAEB entre os anos de 1995 e 2003?”; e que se torne uma referência para

¹² Vale o registro de que novas bases de dados estão sendo produzidas a cada momento. O ciclo do SAEB de 2005 é um exemplo disso. No entanto, optou-se por manter a temporalidade do trabalho de 1995 a 2003 do que recomeçar tudo do zero quando esses dados foram disponibilizados para uso público. Além deles, hoje em dia, dispõe-se também de dados da Prova Brasil, os quais não foram analisados neste trabalho, mas certamente serão alvo de muitos outros trabalhos de pesquisa daqui pra frente.

os estudos sobre o SAEB, em especial, os estudos quantitativos que se valem dos modelos hierárquicos.

Capítulo 2 – Metodologia: dados, medidas e modelos

2.1 A utilização da Teoria de Resposta ao Item (TRI)

O capítulo anterior deixa clara a posição adotada, na qual se defende que a educação deveria ser olhada pelos seus resultados. Vários são os dados produzidos e que fornecem informações para se analisar o sistema educacional brasileiro. O Censo Escolar, do INEP, e a Pesquisa Nacional por Amostra de Domicílios (PNAD), do Instituto Brasileiro de Geografia e Estatística (IBGE), e as avaliações externas, como, por exemplo, o próprio SAEB, podem ser eleitos como as melhores fontes de resultados disponíveis sobre a educação brasileira.

Klein (2006, p. 140) cita que um “sistema educacional é de qualidade quando seus alunos aprendem e passam de ano. Além disso, tem que atender a todas as suas crianças e jovens”. Ou seja, Klein analisa a qualidade de um sistema educacional por meio de indicadores de atendimento (proporção da população em idade escolar matriculada), movimentação (taxas de aprovação, reprovação e abandono durante o ano letivo) e fluxo escolar (taxas de promoção, repetência e evasão entre séries), taxas de acesso e de conclusão e indicadores de desempenho. Sem dúvida, esses indicadores são importantes e devem ser levados em consideração quando são formuladas as políticas públicas. No entanto, o foco deste trabalho é a análise da qualidade e da equidade do sistema educacional brasileiro a partir dos resultados fornecidos pelos indicadores de desempenho, baseados em escalas interpretadas de proficiência.

Quando se fala em grandes números de alunos, isto é, quando se trata de avaliações de larga escala, as dificuldades técnicas para a produção de resultados são muito maiores que aquelas encontradas por um professor em sala de aula, onde ele pode basear seus resultados em um grande conjunto de observações sobre seus alunos. Este capítulo trata das duas medidas que servirão de base para todas as análises que serão implementadas sobre o sistema educacional brasileiro: a medida da proficiência em Matemática e Leitura e a medida do nível socioeconômico de cada aluno, bem como da teoria estatística utilizada para sua produção.

Ambas as medidas são expressões de traços latentes do aluno e, conseqüentemente, não podem ser acessadas diretamente. Também não ficam muito bem definidas se olhadas como o percentual de acertos em um teste, no caso da proficiência, ou a soma dos escores e um questionário contextual, no caso do NSE. Tais dificuldades exigiam um avanço metodológico para serem resolvidas. E esse avanço começa a surgir, no contexto educacional, por volta da década de 80, com a incorporação de uma nova teoria, que “oferece recursos mais flexíveis e eficazes na confecção, análise e apresentação dos resultados de prova que quaisquer outros recursos equivalentes, derivados a partir da Teoria Clássica” (FLETCHER, 1994, p. 23). A chamada Teoria da Resposta ao Item (TRI) surge na década de 50, mas somente na década de 80 passa a ser aplicada nas avaliações educacionais. E no Brasil é utilizada, pela primeira vez, em 1995, na análise dos dados do SAEB daquele ano. Segundo Soares e Pereira (2001), ela surge com os trabalhos de Lord (1952), o primeiro a desenvolver o modelo unidimensional de 2 parâmetros, baseado na distribuição normal acumulada, e Rasch (1960), que, independentemente do trabalho de Lord, propôs o modelo unidimensional de 1 parâmetro, expresso também como modelo de ogiva normal. O próprio Lord, algum tempo depois, percebe a necessidade de aprimoramento e propõe o modelo de 3 parâmetros. Birnbaum (1968) propõe modelos que substituem a ogiva normal pela função logística, a qual é mais conveniente matematicamente. Hoje em dia, esses modelos logísticos para itens dicotômicos são os mais utilizados. Samejima (1969), assim como outros autores, introduzem modelos para mais de duas categorias de respostas, os quais possibilitam a obtenção de mais informação do que apenas o ‘certo/errado’ dos modelos anteriores. Recentemente, Bock e Zimowski (1997) propuseram modelos logísticos de 1, 2 e 3 parâmetros, só que para duas ou mais populações de respondentes.

“A TRI é um conjunto de modelos matemáticos que procuram representar a probabilidade de um indivíduo dar uma certa resposta a um item como função dos parâmetros do item e da habilidade (ou habilidades) do respondente” (ANDRADE; TAVARES; VALLE, 2000, p. 7). Talvez a principal característica dessa nova teoria seja o fato, ou melhor, a promessa de proporcionar medidas invariantes do desempenho cognitivo. Isso significa que os resultados dos instrumentos de medida passam a independem dos itens que o compõem, bem como da amostra investigada. Uma conseqüência importante desta nova teoria é que o item passa a ser a nova unidade de análise, ao contrário do teste como um todo, que era a unidade de análise da Teoria Clássica dos Testes (TCT). Para os interessados em mais detalhes sobre essa

metodologia, sugere-se: Andrade, Tavares e Valle (2000); Hambleton, Swaminathan e Rogers (1991); Hambleton (1993); Baker (2001); e Valle (2000).

Retomando a questão de se tentar medir um traço latente, um *não-observável*, Hambleton (1993) afirma que a relação entre a quantidade observável (escore no item, performance, desempenho) e a quantidade não-observável (traço latente, habilidade do examinando, proficiência, NSE) é descrita por uma função matemática. A Teoria da Resposta ao Item fornece modelos (as funções matemáticas), que podem descrever essa tal relação por meio de uma equação monotônica crescente, chamada de Curva Característica do Item (CCI). Cada item tem a sua própria Curva Característica (BAKER, 2001). Observe-se, portanto, que dependendo dos objetivos ou dos dados disponíveis, modelos diversos podem ser escolhidos. “O que esta metodologia sugere são algumas formas de representar a relação entre a probabilidade de um aluno responder corretamente a um item e seus traços latentes ou habilidades na área de conhecimento avaliada” (VALLE, 2000, p. 7). Quanto maior a proficiência, maiores serão as chances desse aluno acertar um determinado item. No caso de testes de desempenho cognitivo, a proficiência fica definida, então, como uma medida da competência desenvolvida pelo respondente. Um avanço significativo se comparado ao percentual de acertos apenas.

O avanço recente da TRI justifica-se, entre outras coisas, pela dependência de desenvolvimento computacional que viabilizasse sua implementação. Os softwares comerciais, mais comumente utilizados para estimação dos parâmetros da TRI, são: BILOG (MISLEVY; BOCK, 1990), específico para itens dicotômicos com respostas provenientes de um único grupo. O BILOG-MG (ZIMOWSKI et al, 1996), também para itens dicotômicos, porém apropriado para a análise de mais de um grupo de respondentes. E o MULTILOG (THISSEN, 1991), o qual permite a análise de itens com respostas graduadas ou ordinais.

A TRI está baseada em duas suposições: a unidimensionalidade do teste, ou seja, o conjunto de itens apresentado deve medir apenas um único traço latente. Na verdade é suficiente admitir que haja uma competência dominante, responsável pelas respostas dadas ao conjunto de itens. E a independência local, na qual se assume que, para uma dada competência, as respostas aos diferentes itens da prova são independentes.

Um ponto não trivial da TRI, que exige, entre outras coisas, um avanço computacional, é a estimação dos parâmetros envolvidos nos modelos. Essa dificuldade acentua-se ainda mais quando se necessita estimar tanto os parâmetros

dos itens quanto as habilidades. Métodos diferentes podem ser utilizados para estimar tais parâmetros. Hoje em dia, para se estimar os parâmetros dos itens, os métodos mais utilizados são a estimação por Máxima Verossimilhança Marginal e a estimação Bayesiana. Já para a estimação das habilidades, tem-se a Máxima Verossimilhança, o Bayesiano EAP e o Bayesiano MAP. E para a estimação conjunta dos parâmetros dos itens e das habilidades, utiliza-se a Máxima Verossimilhança Conjunta. Para mais detalhes sobre métodos de estimação, consultar Andrade; Tavares; Valle (2000).

Talvez, a principal vantagem de utilização da TRI seja o fato de que é possível apresentar os resultados de populações diferentes e que foram submetidas a testes parcialmente diferentes em uma mesma escala. Tal vantagem é essencial para a concretização desta tese, seja na comparabilidade das proficiências entre os diversos ciclos do SAEB, seja na comparabilidade do NSE, calculado para cada aluno de bases distintas. Esse processo é denominado equalização.

Equalizar significa equiparar, tornar comparável, o que no caso da TRI significa colocar os parâmetros de itens vindos de provas distintas ou habilidades de respondentes de diferentes grupos, na mesma métrica, isto é, numa escala comum, tornando os itens e/ou as habilidades comparáveis.

Existem dois tipos de equalização: a equalização via população e a equalização via itens comuns. Isto significa que há duas maneiras de colocar parâmetros, tanto de itens quanto de habilidades, numa mesma métrica: na primeira usamos o fato de que se um único grupo de respondentes é submetido a provas distintas, basta que todos os itens sejam calibrados conjuntamente para termos a garantia de que todos estarão na mesma métrica. Já na equalização via itens comuns, a garantia de que as populações envolvidas terão seus parâmetros em uma única escala será dada pelos itens comuns entre as populações, que servirão de ligação entre elas. (ANDRADE; TAVARES; VALLE, 2000, p. 80-81)

2.1.1 Modelo para o cálculo da proficiência

O cálculo da proficiência dos alunos em Matemática e Leitura no SAEB já é realizado pelo próprio INEP, não sendo necessário implementá-lo neste trabalho de doutorado. O modelo utilizado é o modelo logístico unidimensional de três parâmetros.

A equação matemática, que relaciona a probabilidade de um aluno, com habilidade θ , de responder corretamente ao i-ésimo item, é dada por:

$$P_i(U_{ij} = 1 | \theta_j) = c_i + (1 - c_i) \frac{1}{1 + \exp[-Da_i(\theta_j - b_i)]} \quad (1)$$

com $i = 1, 2, \dots, l$, e $j = 1, 2, \dots, n$, onde:

U_{ij} é uma variável dicotômica que assume os valores 1 quando o indivíduo j responde corretamente ao item i , ou 0 quando o indivíduo j não responde corretamente ao item i ;

θ_j representa a competência (traço latente) do j -ésimo indivíduo;

$P(U_{ij} = 1 | \theta_j)$ é a probabilidade de um indivíduo j com habilidade θ_j responder corretamente ao item i ;

a_i é o parâmetro de discriminação (ou de inclinação) do item i , com valor proporcional à inclinação da Curva Característica do Item (CCI) no ponto b_i ;

b_i é o parâmetro de dificuldade (ou de posição) do item i , medido na mesma escala da competência;

c_i é o parâmetro do item que representa a probabilidade de indivíduos com baixa competência responderem corretamente ao item i (muitas vezes referido como a probabilidade de acerto casual);

D é um fator de escala, constante e igual a 1. Utiliza-se o valor 1,7 quando deseja-se que a função logística forneça resultados semelhantes ao da função ogiva normal.

O Gráfico 1 apresenta uma CCI (teórica), na qual a influência dos três parâmetros na medida da proficiência pode ser visualizada graficamente. Essa curva foi gerada a partir da equação 1.

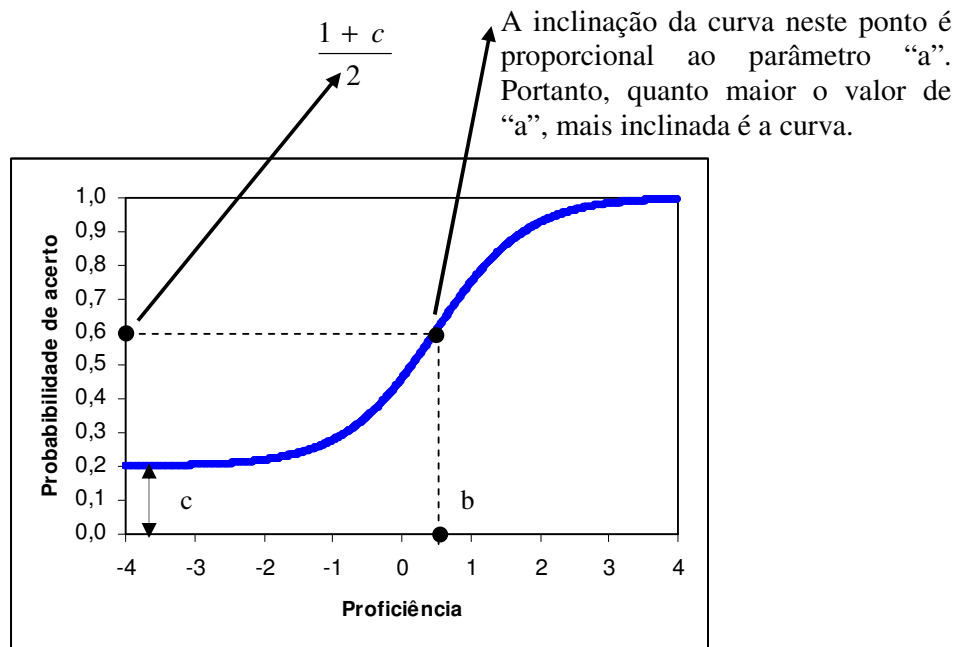


GRÁFICO 1: Curva Característica de um item com parâmetros: $a = 1,5$; $b = 0,5$; $c = 0,2$.

A escala de proficiência, utilizada nesses gráficos de CCI, não é absoluta. Aqui, por exemplo, ela está variando de -4 até $+4$, os quais são valores genéricos e normalmente disponibilizados pelos softwares. Deve-se observar que o valor de “ b ” está localizado na escala do traço latente (proficiência) e representa a competência necessária para uma probabilidade de acerto igual a $(1+c)/2$, ou seja, quanto menor o valor de “ b ” mais fácil será o item, e vice-versa.

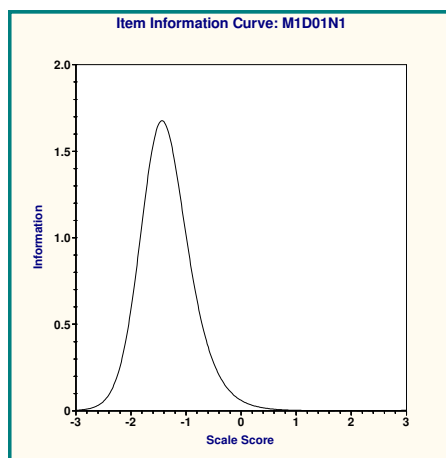
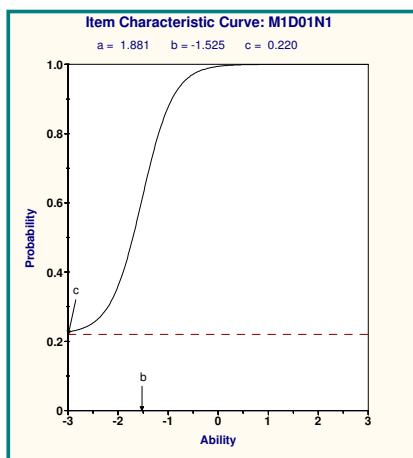
Já o parâmetro “ c ”, é uma medida de acerto ao acaso, portanto vale constatar que mesmo os alunos de menor proficiência ainda têm uma chance de acertar este item “no chute”.

O parâmetro “ a ” está relacionado com a inclinação da CCI. Assim, não se espera itens com “ a ” negativo, pois indicariam que a probabilidade de se responder corretamente àquele item diminui com o aumento da competência. Quanto maior o valor de “ a ” maior será o poder de discriminação daquele item, isto é, o “ a ” relaciona-se com a capacidade do item para separar os alunos que sabem aquilo que está sendo perguntado daqueles que não sabem.

Os gráficos a seguir foram gerados no software BILOG-MG e referem-se a dois itens utilizados em um teste de Matemática. Eles servem para ilustrar a influência desses parâmetros na CCI do item. Para cada CCI, associa-se, também, uma Curva de Informação do Item, a qual “permite analisar quanto um item (ou teste) contém de

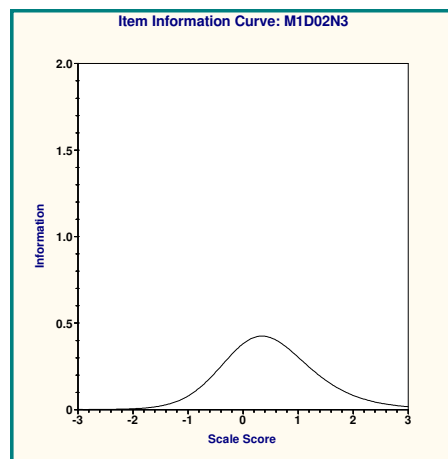
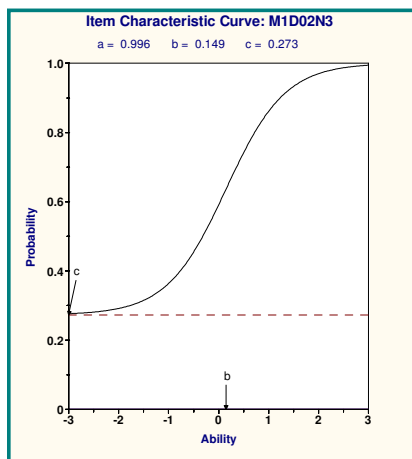
informação para a medida de habilidade” (ANDRADE; TAVARES; VALLE, 2000, p. 12).

GRÁFICO 2: Curva Característica do Item M1D01N1 GRÁFICO 3: Curva de Informação do Item M1D01N1



FONTE: Elaboração própria com os dados do SAEB – INEP/MEC.

GRÁFICO 4: Curva Característica do Item M1D02N3 GRÁFICO 5: Curva de Informação do Item M1D02N3



FONTE: Elaboração própria com os dados do SAEB – INEP/MEC.

Comparando-se esses gráficos, percebe-se que quanto maior o valor de “b” mais a CCI ‘demora a subir’, isto é, somente os alunos com maior competência têm maiores chances de acertar este item, justificando porque o item M1D02N3 é mais difícil que o item M1D01N1. Comparando-se o parâmetro “c” dos dois itens, percebe-

se a sua influência na chance de os alunos de baixa competência acertarem o item. Quanto maior o valor de “a”, mais inclinada é a curva, portanto o item M1D01N1 discrimina melhor os alunos na faixa de proficiência próxima de “b”. Por fim, os gráficos das curvas de informação nos permitem analisar não só o quanto um item contém de informação, mas também para qual faixa de proficiência ele é apropriado. Esse dado é particularmente relevante quando se está montando um teste em uma avaliação sistêmica, uma vez que a mesma deve conter itens em todas as faixas da competência avaliada.

2.1.2 Modelo para o cálculo do nível socioeconômico

Para se obter uma medida do nível socioeconômico (NSE) de cada aluno, não é apropriada a utilização de modelos dicotômicos da TRI, uma vez que os itens utilizados não podem ter as suas respostas categorizadas em certo ou errado. Em geral, são itens com respostas graduadas, como, por exemplo, quando se pergunta a escolaridade da mãe. Opta-se, então, pela adoção de um modelo para itens não dicotômicos, os quais não consideram somente se o indivíduo respondeu à alternativa correta ou não, mas também se leva em consideração qual foi a resposta dada por ele. Em especial, no caso do NSE, como as respostas dos itens podem ser ordenadas entre si, o modelo mais apropriado é o modelo de resposta gradual de Samejima (1969).

Neste caso, os escores das categorias de cada item são arranjados, previamente, do menor para o maior, e a probabilidade de um indivíduo j escolher uma particular categoria ou outra mais alta do item i pode ser dada por uma extensão do modelo logístico de dois parâmetros:

$$P_{i,k}(\theta_j) = \frac{1}{1 + \exp[-Da_i(\theta_j - b_{i,k})]} - \frac{1}{1 + \exp[-Da_i(\theta_j - b_{i,k+1})]} \quad (2)$$

com $i = 1, 2, \dots, l$, $j = 1, 2, \dots, n$, e $k = 0, 1, \dots, m_i$, onde:

$b_{i,k}$ é o parâmetro de dificuldade da k -ésima categoria do item i ;

e os demais parâmetros no modelo são análogos aos já definidos no modelo para o cálculo da proficiência.

Ao contrário do modelo para itens dicotômicos, nesse caso tem-se uma ordenação entre o nível de dificuldade das categorias de um dado item, de acordo com a classificação de seus escores. Os Gráficos 6 e 7, gerados pelo software MULTILOG, ilustram essa situação, na medida em que apresentam a CCI e a Curva de Informação para um dos itens do questionário contextual do SAEB, que foi utilizado no cálculo do NSE. No caso, trata-se do item 10 que questiona a quantidade de automóveis existentes no lugar onde o aluno mora e que possui três alternativas: 1) Não tem; 2) Tem um; 3) Tem dois ou mais.

GRÁFICO 6: Curva Característica do Item 10

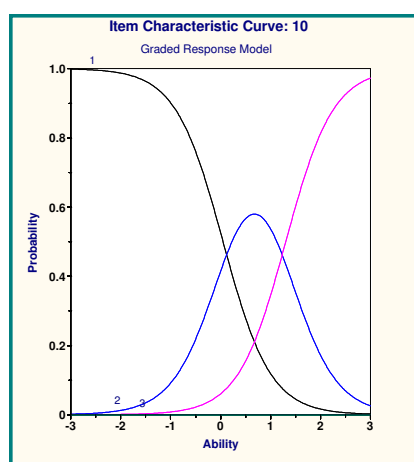
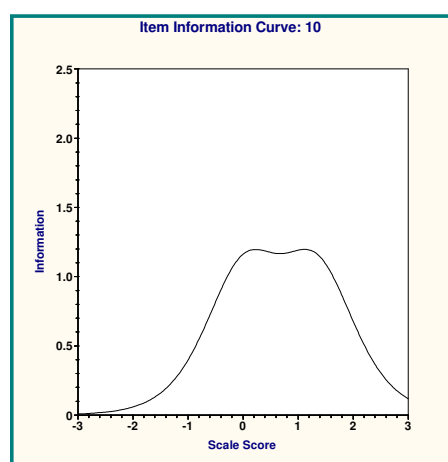


GRÁFICO 7: Curva de Informação do Item 10



FONTE: Elaboração própria com os dados do SAEB – INEP/MEC.

No eixo das abscissas, está colocado o NSE dos alunos e no eixo das ordenadas, a probabilidade de uma pessoa, com um dado NSE, escolher cada uma das respostas. Pode-se notar que, com exceção dos extremos da escala, não existem regiões nas quais as alternativas 1, 2 ou 3 são escolhidas apenas pelos alunos de uma determinada faixa de NSE. Além disso, à medida que aumenta o NSE, aumenta a probabilidade de se escolher a opção 3 e diminui a probabilidade de se escolher as outras duas opções. Ou seja, enquanto alunos classificados como sendo de menor NSE respondem que não têm automóvel em casa, alunos classificados com NSE mais elevado respondem que, em sua casa, existe pelo menos um automóvel. Essa CCI, expressa no Gráfico 6, é fruto da relação matemática expressa em (2), com os dados selecionados para este trabalho. O Gráfico 7 é um pouco mais abstrato e fornece a região da escala na qual aquele item fornece melhores resultados. Observa-se, por exemplo, que este não é um bom item para se estimar o NSE de alunos mais pobres, pois, em geral, todos optarão pela alternativa 1.

Tanto a medida de proficiência quanto a de NSE serão mais bem explicadas nos tópicos a seguir, entretanto faz-se necessário registrar a importância da TRI para a produção dessas medidas, um aparato metodológico sem o qual a abrangência deste trabalho não seria possível.

2.2 Medida da proficiência

2.2.1 A escala de proficiência do SAEB

Refletir sobre a lógica de estruturação dessas avaliações de larga escala, contrapondo-as às avaliações de aprendizagem, aplicadas pelo próprio professor em sala de aula, pode ser um caminho para se entender a necessidade da utilização de uma única escala interpretada de proficiência em nível nacional, no caso a escala do SAEB. Essa lógica, grosso modo, poderia ser resumida em nove etapas: (1) definição da população a ser avaliada ou de um plano amostral; (2) elaboração ou definição de uma matriz de referência; (3) produção de itens; (4) pré-teste desses itens; (5) montagem dos testes; (6) aplicação; (7) correção dos testes e produção de estatísticas; (8) alocação de alunos e itens a uma escala interpretada de proficiências; (9) e, finalmente, a divulgação dos resultados para a sociedade. A descrição de todas essas etapas metodológicas para o caso específico do SAEB pode ser encontrada em Brasil/MEC/INEP (2006; 2005; 2002a; 2001). Particularmente, neste capítulo, nos interessa a etapa (8), alocação de alunos e itens a uma escala interpretada de proficiências.

Inicialmente, apenas a Teoria Clássica dos Testes (TCT) era utilizada para análise dos resultados de avaliações. Isso, de certa forma, restringia as possibilidades de um trabalho pedagógico de qualidade com os resultados de avaliações em larga escala. O teste era a unidade de análise e o resultado era função do número de itens acertados pelo examinando. As comparações ficavam fragilizadas na medida em que se exigia que os diversos indivíduos fossem submetidos à mesma prova, ou, pelo menos, ao que se denomina de formas paralelas de testes. E mais, ter dois alunos com notas idênticas não significava, necessariamente, similaridade na competência de ambos. Isso porque um deles poderia ter acertado uma questão mais difícil, conseqüentemente, de peso maior, enquanto o outro acertou duas ou três questões mais fáceis. Ou mesmo porque, utilizando a Matemática para exemplificar, um acertou

as questões de Geometria, enquanto o outro acertou as questões de Álgebra. Resumindo, limitações na tecnologia disponível implicavam em limitações na qualidade da análise dos resultados.

No entanto, os professores continuam aplicando testes e contando o número de acertos dos alunos, sem prejuízos para a avaliação do processo de aprendizagem. Por que é possível dentro das salas de aula, mas não nas avaliações sistêmicas? A explicação reside no fato de que, no caso da sala de aula, o próprio professor é o autor dos testes, o corretor, o analista dos resultados, o divulgador e, principalmente, porque esse não é o único instrumento utilizado por ele. O professor convive com seus alunos, que são poucos se comparados com os de um sistema educacional; observa-os todos os dias; corrige seus deveres de casa; tira suas dúvidas; aplica trabalhos em grupo; ou seja, o professor tem várias maneiras diferentes para identificar a competência de cada aluno, e o resultado não precisa ser interpretado, porque ele participou de todas as etapas e conhece seus alunos. Ou melhor, a interpretação está subjacente a todo o processo. Já nas avaliações em larga escala acontece o oposto. Em geral, são grupos (ou pessoas) diferenciados, que se responsabilizam por determinadas etapas; o resultado é fruto de um único contato com os alunos; e o principal, deve ser divulgado para setores da sociedade, que, evidentemente, não conhecem cada um dos alunos avaliados.

Dessa necessidade de apresentação e divulgação dos resultados, de forma compreensível para os vários setores da sociedade, nascem as interpretações das escalas de proficiência. “Uma escala é uma maneira de ordenar medidas de acordo com valores arbitrados” (BRASIL/MEC/INEP, [1997?], p. 6). As escalas de temperatura são uma boa analogia para se entender as escalas de proficiência, já que ambas são, na verdade, arbitrárias. Dizer que a proficiência de um aluno, em um determinado aspecto, é 200 significa que ele já desenvolveu mais competências, neste aspecto, que um aluno cuja proficiência é 150; assim como uma temperatura de 20° C implica em algo mais quente que uma temperatura de 10° C, porém qual roupa vestir em um dia em que a temperatura ambiente está próxima dos 20° C? Responder a essa pergunta implica em se dominar uma interpretação da escala de temperatura. Da mesma forma, se se perguntasse quais tarefas um aluno com proficiência 200 consegue realizar em relação àquela determinada competência, só seria possível responder se existisse uma interpretação pedagógica da escala de proficiência.

O papel dos itens, em uma avaliação educacional, é análogo ao de um oftalmologista quando, em uma consulta, tenta determinar o grau de deficiência visual

de um paciente. O médico apresenta lentes diferentes, uma após a outra, perguntando sempre se a visão do paciente melhorou ou piorou. Esse processo é repetido tantas vezes quanto necessário para se convergir para um ponto específico de uma “escala visual”, se é que se pode chamar assim o grau do paciente. No caso das avaliações, é como se cada item perguntasse ao aluno se ele já domina ou não aquela habilidade à qual o item se refere. E o conjunto de itens dará a informação que se deseja com o teste, ou seja, a medida da competência dos alunos. Deve-se chamar a atenção para o fato de que é o conjunto de itens que permite inferências sobre a competência do aluno e não um item isoladamente, assim, como um oftalmologista necessita utilizar mais de uma lente para inferir sobre o problema visual do paciente.

Vários itens, com índices de dificuldades diferentes e que abordam habilidades específicas, são apresentados, e os acertos de cada aluno vão definindo em que ponto da escala ele está, isto é, qual é a proficiência dele. Fica claro porque não se deve apresentar um único item que meça uma determinada habilidade, a informação seria pouco precisa, pouco confiável. Mas também não adianta apresentar inúmeros itens de uma mesma habilidade, pois estaria sendo coletada sempre a mesma informação.

Se o oftalmologista conhece o histórico do paciente, menos tentativas para determinar o seu grau ele terá que fazer. Isso também vale para a escala de proficiência. Se já se tem alguma idéia *a priori* da distribuição de proficiências de uma população de alunos, menos itens serão necessários para se fazer uma medida com a mesma precisão, pois já se vai procurar o aluno muito próximo de onde ele deve estar realmente.

Para se construir escalas de conhecimento interpretáveis, ou como estão sendo denominadas neste texto, escalas de proficiência, é necessário que todos os parâmetros envolvidos sejam comparáveis. Isto é, tanto os parâmetros estimados para os itens do teste quanto os parâmetros das habilidades dos respondentes devem estar em uma mesma métrica (ANDRADE; TAVARES; VALLE, 2000). Existe, conseqüentemente, uma correspondência entre a proficiência do aluno e o item que é respondido. Ambos estão colocados na mesma escala de posicionamento. Itens mais fáceis ficarão alocados na parte inferior dessa escala e os mais difíceis, na parte superior. Os alunos que só conseguem acertar os itens mais fáceis têm baixa proficiência e também ficam posicionados na parte inferior da escala, e os de maior proficiência, na parte superior. Esse fato, fundamental ao processo, é que permite fazer uma ligação entre os valores medidos de proficiência e os conteúdos e

habilidades envolvidos na avaliação, ou em palavras já utilizadas, se interpretar pedagogicamente a escala de proficiência.

Diferentemente do que acontece nos testes corrigidos pelo número de acertos, nos quais a menor nota possível é zero e a maior é igual ao número de itens no teste, utilizando-se a TRI, esses valores podem variar, teoricamente, de $-\infty$ até $+\infty$ (de menos infinito a mais infinito). Na prática, todos os softwares que utilizam a TRI geram valores de proficiência numa escala com média igual a zero e desvio padrão igual a um, escala que é comumente denominada (0;1). Como a habilidade de um indivíduo é invariante à escala de medida, normalmente, utilizam-se outros valores para se divulgar tais resultados. Isso porque a escala (0;1) apresenta um grande número de proficiências negativas, proficiências com casas decimais e compreende o zero, valores que facilmente podem implicar em interpretações equivocadas.

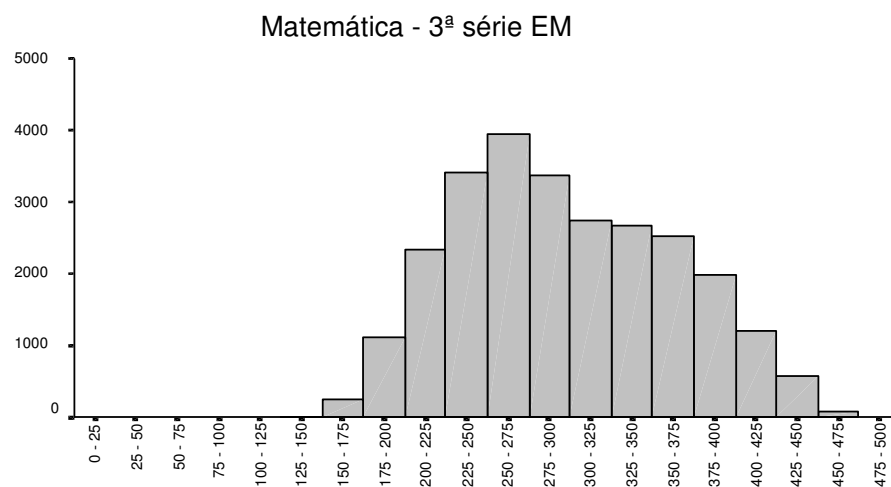
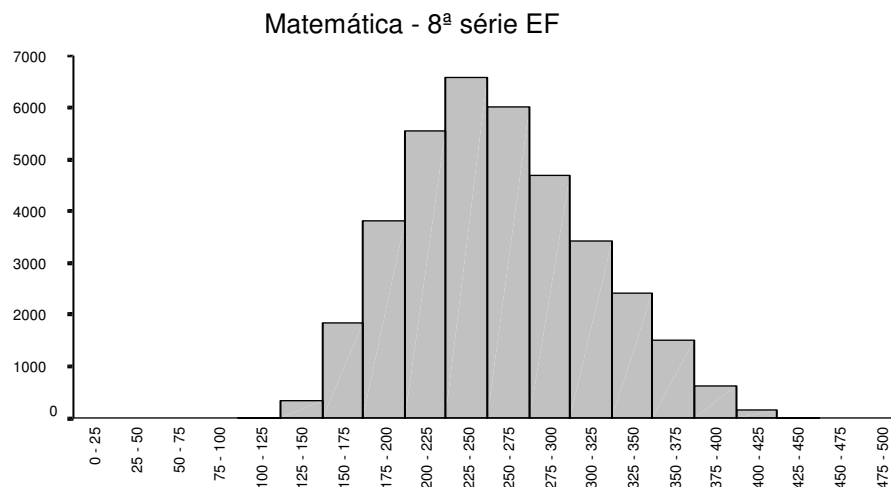
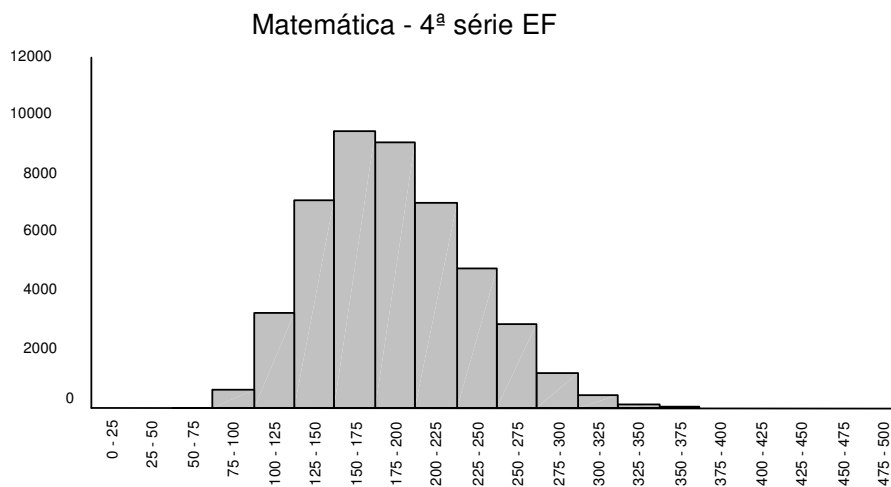
O INEP estabeleceu que a escala do SAEB seria tal que os alunos da 8ª série do ensino fundamental, que participaram da avaliação do SAEB em 1997, teriam uma média de 250 pontos e um desvio-padrão de 50¹³. Com isso eliminam-se os valores negativos e o zero, mesmo para as séries iniciais do fundamental e evita-se também que a escala contenha apenas valores entre 0 e 100, os quais poderiam levar a uma confusão com percentual de acertos no teste. A mudança nos valores de uma escala é relativamente simples, exigindo apenas uma transformação linear. Para os interessados, essa transformação linear está descrita nos anexos.

No entanto, não se deve confundir a simples tarefa de transladar os valores de uma escala com a difícil tarefa de colocar uma escala em uma mesma métrica de outra escala previamente construída. Esse processo, como já informado anteriormente, é denominado equalização (ANDRADE; TAVARES; VALLE, 2000; KLEIN, 2003) e consiste em encontrar os valores que seriam assumidos pelas novas proficiências caso estivessem sujeitas às regras de produção da primeira escala. Ou como define Andrade (2001), é o processo de obtenção de valores dos parâmetros dos itens e das habilidades numa mesma escala. Andrade (2001) mostra ainda que existem diversas técnicas para se realizar uma equalização, dependendo do número de grupos envolvidos, dos tipos de provas aplicados e do conjunto de itens que foram selecionados para as provas.

¹³ Na verdade, esse valor é ligeiramente diferente para cada uma das disciplinas avaliadas. Em Matemática, ele vale 249,9644 para a média e 55,89233 para o desvio padrão. E em Língua Portuguesa, os valores são 249,985 e 55,093, respectivamente, para a média e o desvio padrão.

Concretizando o que é uma escala de proficiência, por meio de um exemplo com os dados do SAEB, o Gráfico 8 apresenta três histogramas, propositalmente apresentados em um mesmo intervalo da escala e um sobre o outro. Com isso, pode-se perceber como estão distribuídas as proficiências dos alunos brasileiros em Matemática no SAEB de 2003. A tendência média é de crescimento de uma série para outra, apesar de as diferenças entre a 8ª série do ensino fundamental e a 3ª série do ensino médio serem pequenas. Repetindo, isso só é possível porque existem itens em comum entre os testes das três séries avaliadas, o que torna a proficiência produzida comparável via um processo de equalização dos resultados. Observe-se também que, no caso específico do SAEB, níveis muito baixos da escala (menores que 75) não fazem sentido, assim como níveis muito altos (acima de 500), isto é, não existem alunos em nenhuma das três séries avaliadas com esses valores de proficiência.

Gráfico 8: Proficiência em Matemática, por série, no SAEB 2003



FONTE: Elaboração própria com os dados do SAEB 2003 – INEP/MEC

2.2.2 Interpretações pedagógicas diversas de uma mesma escala

Uma vez vencida a tarefa de se mostrar a necessidade da utilização de escalas de proficiência para a apresentação e divulgação de resultados de avaliações de larga escala, surge uma nova pergunta: já que as escalas de proficiência nada mais são do que escalas de posicionamento, sua efetiva utilização pelos gestores públicos da educação e, principalmente, pela comunidade escolar só acontecerá mediante uma interpretação pedagógica da escala. Pergunta-se então, como acontece esse processo de interpretação de uma escala desse tipo?

Para que se responda a essa pergunta, propõe-se recorrer novamente à analogia com a escala de temperatura. A escala Celsius já faz parte do dia a dia do povo brasileiro, porém o mesmo não pode ser dito da escala Fahrenheit¹⁴. De tal forma que, se fosse apresentada uma temperatura de 40° F (quarenta graus Fahrenheit) aos brasileiros, a maioria não saberia dizer o que significa essa medida, qual a sua interpretação física, apenas poderia afirmar que é algo mais quente que 10 ou 20° F. No entanto, se a temperatura apresentada fosse de 40° C, o mesmo grupo saberia o seu significado físico, podendo, por exemplo, tirar conclusões sobre que tipo de roupa vestir.

Para as escalas de proficiência, o raciocínio é o mesmo. Os números indicam apenas uma posição. É preciso fazer uma interpretação pedagógica dessa escala para que aqueles números passem a significar algo. Olhando-se a história do SAEB percebe-se que, nos ciclos de 95 e 97, essa interpretação era feita escolhendo-se pontos arbitrários da escala, identificando itens que pudessem ser ancorados àqueles pontos e, finalmente, inferindo o que os alunos, alocados próximos àqueles pontos, eram capazes de fazer (BRASIL/MEC/INEP, [1997?]; 1998). Essa inferência é feita por especialistas de cada disciplina, analisando-se um painel de itens alocados em diversos pontos da escala, porém a descrição do intervalo da escala próximo àquele ponto “não é uma formulação teórica em termos do que os alunos devem saber, mas sim o que eles efetivamente demonstraram saber e foram capazes de fazer nos testes

¹⁴ Escala de temperatura utilizada em alguns países de língua inglesa, tais como Estados Unidos e Inglaterra.

aplicados, uma vez que os itens-âncora¹⁵ foram obtidos empiricamente” (BRASIL/MEC/INEP, 1998, p. 25).

A forma de interpretação da escala de proficiência, descrita no parágrafo anterior, apresenta um problema, pois com ela

não é possível atribuir significado pedagógico às medidas de desempenho que se situam entre dois pontos interpretados da escala. Sob esta perspectiva, portanto, havia lacunas na interpretação das escalas de desempenho, dificultando a análise e utilização dos resultados pelos agentes educacionais (BRASIL/MEC/INEP, 2000, p. 10).

Para vencer esse obstáculo, a partir do SAEB 99, a forma de interpretação da escala passou a levar em consideração não um ponto arbitrário, mas um intervalo de proficiências, denominados níveis de desempenho. Essa técnica permite que a escala como um todo, e não somente alguns pontos, ganhe sua própria interpretação. No caso da escala do SAEB, Klein (2006) descreve a maneira como a escala foi dividida para ser interpretada:

A escala do SAEB é interpretada no nível 250 e nos níveis obtidos somando-se e subtraindo um múltiplo de 25 (meio desvio padrão). Assim, na disciplina Matemática, os níveis interpretados são 125, 150, [...], 400 e 425. A interpretação é acumulativa. Os alunos em um nível sabem e são capazes de fazer o que foi descrito nos níveis anteriores. Em cada nível, adicionam-se as habilidades que a população começa a dominar naquele nível. (KLEIN, 2006, p. 152-153)

Valle (2001) apresenta, em um texto simples, considerações gerais sobre construção e interpretação das escalas de desempenho cognitivo. Mas, independente da forma pela qual se opta para fazer a interpretação da escala, o fato é que a história das avaliações mostra que esse processo não é simples. Retomando a analogia com

¹⁵ Itens-âncora são itens que caracterizam os níveis das escalas no sentido de que a grande maioria dos alunos, situados em cada um dos níveis, acerta o item, enquanto poucos alunos, situados no nível imediatamente inferior, também o acertam. Inicialmente, eram utilizados critérios mais rigorosos para a definição de um item-âncora, como citam Andrade; Tavares; Valle (2000, p. 110): “para um item ser âncora em um determinado nível âncora da escala, ele precisa ser respondido corretamente por uma grande proporção de indivíduos (pelo menos 65%) com este nível de habilidade e por uma proporção menor de indivíduos (no máximo 50%) com o nível de habilidade imediatamente anterior. Além disso, a diferença entre a proporção de indivíduos com esses níveis de habilidade que acertam a esse item deve ser de pelo menos 30%.” Hoje em dia, percebendo-se a dificuldade de localizar itens aos níveis da escala, são adotados critérios mais flexíveis, como, por exemplo, exigir, apenas, que o item seja acertado por, no mínimo, 65% dos indivíduos, para que ele seja classificado como item âncora do primeiro nível em que isso acontecer.

a temperatura, Medeiros (1999), em um texto sobre a história da termometria, afirma que, assim como na proficiência, o processo de medição e interpretação da escala de temperatura fora lento e cheio de problemas, portanto ainda não se tem um consenso sobre o assunto. Esta falta de consenso gera, entre outras coisas, interpretações divergentes para um mesmo resultado, fato que será detalhado no próximo item.

Mas não se deve reduzir uma escala de proficiência a apenas um conjunto de itens ordenados e interpretados em níveis. Ela é, também, um espelho do ensino ao qual foi submetida a população que ‘calibrou’ a escala. No caso deste trabalho, quando se afirma que todas as proficiências dos cinco ciclos utilizados estão numa mesma escala, significa dizer, na verdade, que todos os demais ciclos foram colocados na mesma métrica do SAEB de 1997 e, portanto, estão sujeitos aos padrões de ensino vivenciados naquela época.

2.2.3 Pontos de corte da escala e seus respectivos adjetivos – olhares diferenciados sobre o mesmo resultado

Uma vez que os resultados de uma avaliação foram produzidos, tem-se uma escala de proficiências gerada. De acordo com o que foi apresentado no item anterior, é fortemente desejável que essa escala seja única e de abrangência nacional. Todas as avaliações sistêmicas deveriam se esforçar por divulgar seus resultados na escala do SAEB, assim, os resultados produzidos seriam comparáveis.

No entanto, apenas a divulgação de resultados equalizados na escala SAEB não é uma garantia de entendimento e utilização dos mesmos. Essa escala, por já ter sido interpretada, fornece uma lista das competências já desenvolvidas em cada faixa de proficiência, mas essa é uma informação complexa e que só pode ser analisada por quem tenha conhecimentos pedagógicos da disciplina em questão. Então, é natural imaginar que alguns esforços tenham sido feitos para categorizar esta escala, isto é, criar adjetivos para os seus diversos intervalos, tornando-a acessível a qualquer pessoa, mesmo que leiga na disciplina avaliada. Essa categorização implica em responder três perguntas: (1) em quantos intervalos dividir a escala?; (2) quais serão os pontos de corte de cada intervalo?; (3) qual o nome será dado para cada

intervalo¹⁶?. Como inexistem estudos para se determinar as respostas a essas três questões, várias são as classificações que têm surgido, como se pode ver a partir dos exemplos nas tabelas a seguir.

A Tabela 2A foi proposta pelo Brasil/MEC/INEP (2003a; 2003b; 2003c) para análise dos resultados de Matemática do SAEB 2003, e a Tabela 2B para os resultados de Língua Portuguesa.

TABELA 2A: Interpretação dos níveis da escala SAEB de Matemática segundo o SAEB 2003

Série escolar	Muito crítico	Crítico	Intermediário	Adequado	Avançado
4 ^a EF	Abaixo de 125	125 – 175	175 – 250	250 – 350	Acima de 350
8 ^a EF	Abaixo de 175	175 – 250	250 – 350	350 – 400	Acima de 400
3 ^a EM	Abaixo de 200	200 – 300	300 – 375	Acima de 375	-

FONTE: SAEB – INEP/MEC

TABELA 2B: Interpretação dos níveis da escala SAEB de Língua Portuguesa segundo o SAEB 2003

Série escolar	Muito crítico	Crítico	Intermediário	Adequado	Avançado
4 ^a EF	Abaixo de 125	125 – 175	175 – 250	250 – 350	Acima de 350
8 ^a EF	Abaixo de 150	150 – 200	200 – 300	300 – 375	Acima de 375
3 ^a EM	Abaixo de 175	175 – 250	250 – 350	Acima de 350	-

FONTE: SAEB – INEP/MEC

Deve-se salientar que as escalas de Matemática e Língua Portuguesa não são comparáveis, o que significa que os resultados em uma disciplina devem ser olhados tendo em vista o que seria esperado para ela, e não em comparação com os resultados da outra disciplina.

As tabelas 3A e 3B foram propostas pelo CAEd¹⁷ (MINAS GERAIS/SEE-MG, 2002; MINAS GERIAS/SEE-MG, 2003), respectivamente, para se interpretar os resultados do Programa de Avaliação da Rede Pública da Educação Básica de Minas

¹⁶ Um intervalo pode conter vários níveis de uma escala. No entanto, a literatura deste tema costuma utilizar a palavra “nível” como sinônimo de “intervalo”. Reconhece-se que isso pode provocar alguma confusão, mas por força do hábito, a partir de agora se opta, também, pelo uso da palavra “nível”.

¹⁷ CAED é o Centro de Políticas Públicas e Avaliação da Educação da Universidade Federal de Juiz de Fora.

Gerias (PROEB) em Matemática, no ano de 2003, e em Língua Portuguesa, no ano de 2002.

TABELA 3A: Interpretação dos níveis da escala SAEB de Matemática segundo o PROEB 2003

Série escolar	Crítico	Intermediário para crítico	Intermediário para suficiente	Suficiente
4 ^a EF	Abaixo de 150	150 – 175	175 – 200	Acima de 200
8 ^a EF	Abaixo de 200	200 – 250	250 – 300	Acima de 300
3 ^a EM	Abaixo de 275	275 – 325	325 – 375	Acima de 375

FORNTE: PROEB – SEE/MG

TABELA 3B: Interpretação dos níveis da escala SAEB de Língua Portuguesa segundo o PROEB 2002

Série escolar	Crítico	Intermediário	Adequado
4 ^a EF	Abaixo de 175	175 – 200	Acima de 200
8 ^a EF	Abaixo de 225	225 – 275	Acima de 275
3 ^a EM	Abaixo de 250	250 – 300	Acima de 300

FORNTE: PROEB – SEE/MG

Antes de comparar as tabelas 2A e 2B com as tabelas 3A e 3B, vale uma comparação entre as duas últimas tabelas. Observe-se que o mesmo grupo de especialistas, no intervalo de apenas um ano, optou por mudar a quantidade de níveis e os adjetivos. À mudança nos pontos de corte não cabe análise, pois são disciplinas diferentes, e, como já foi mencionado, tais escalas não são comparáveis. Nome também é política, e isso pode ser constatado na opção de se trocar o adjetivo do nível mais alto de “adequado” para “suficiente”.

Outra questão que surge, ao se analisar essas tabelas, é a diferença no tamanho dos intervalos entre séries. Olhando o nível “adequado” da Tabela 3B, tem-se uma diferença de 75 pontos entre a 4^a e a 8^a séries do ensino fundamental, mas apenas 25 pontos separam a 8^a série da 3^a série do ensino médio. Analisando-se a coluna do “crítico”, tem-se a mesma proporção dos intervalos. Terá sido essa uma decisão baseada no aspecto pedagógico ou nos resultados empíricos? E sendo a segunda opção a resposta para essa pergunta, não se estaria subestimando os níveis de excelência para o ensino médio só porque, atualmente, a qualidade dos resultados para esse seguimento ainda é baixa? No caso de Matemática, na tabela 3A, o tamanho dos intervalos na coluna do “suficiente” é mais coerente com o número de

anos de estudo existente entre a 4ª EF e 8ª EF e entre a 8ª EF e a 3ª EM (há um aumento de 100 pontos entre os quatro anos que se passam da 4ª para a 8ª série do EF; e 75 pontos de aumento entre os três anos que se passam entre a 8ª EF e a 3ª EM). Já na coluna do “crítico”, a variação na proficiência está incoerente com o tempo de estudo (aumentam-se 50 pontos nos quatro anos entre a 4ª e a 8ª; e 75 pontos nos três anos entre a 8ª e a 3ª EM).

Comparando-se as propostas do SAEB com as do PROEB, observam-se mais diferenças que semelhanças. Entre elas, vale salientar a diferença no número de níveis, com a adoção, por parte do INEP, de um nível denominado “avançado”, para além do que seria adequado para aquela série. Provavelmente, a justificativa para a utilização desse nível seja o incentivo para aquelas escolas que querem um algo a mais além do básico e que possuem um número maior de alunos no extremo direito da distribuição de proficiências.

Outra diferença relevante é a divergência entre os pontos de corte para níveis com o mesmo nome. Como exemplo, tem-se o caso da 4ª série EF, em Matemática, quando alunos, cuja proficiência fosse entre 125 e 175, seriam classificados como “crítico” pelo SAEB. No entanto, no mesmo ano de 2003, o PROEB só classificou como “crítico” os alunos cuja proficiência foi inferior a 150. E aí surge a pergunta óbvia: afinal, esses alunos, cuja proficiência em Matemática, no ano de 2003, se encontrava entre 150 e 175, estão em um estado crítico de aprendizagem ou não? Deve-se implementar políticas públicas para melhorar o quadro de aprendizagem desses alunos ou o nível deles não é tão preocupante assim? Independente da resposta, o fato é que a classificação adotada em 2003 pelo SAEB é bem mais ‘exigente’.

A diversidade nas interpretações é vista, também, em outros exemplos. As Tabelas 4A e 4B referem-se aos níveis do Sistema de Avaliação da Educação de Mato Grosso do Sul (SAEMS), no ano de 2003, (MATO GROSSO DO SUL/SED-MS, 2003) e comprovam, se comparadas às Tabelas 3A e 3B, que um mesmo grupo de especialistas pode sugerir, em um mesmo ano, interpretações diversas para estados diferentes (no caso Minas Gerais, com o PROEB, e Mato Grosso do Sul, com o SAEMS). E as tabelas 5A e 5B finalizam essa seção com mais uma interpretação proposta. Novamente, no estado de Minas Gerais, só que agora dentro do contexto de produção do Atlas da Educação de Minas Gerais (MINAS GERAIS/SEE-MG, 2005).

TABELA 4A: Interpretação dos níveis da escala SAEB de Matemática segundo o SAEMS 2003

Série escolar	Muito crítico	Crítico	Intermediário	Adequado
4ª EF	Abaixo de 125	125 – 175	175 – 250	Acima de 250
8ª EF	Abaixo de 150	150 – 250	250 – 300	Acima de 300
3ª EM	Abaixo de 200	200 – 300	300 – 375	Acima de 375

FONTES: SAEMS – SEE/MS

TABELA 4B: Interpretação dos níveis da escala SAEB de Língua Portuguesa segundo o SAEMS 2003

Série escolar	Muito crítico	Crítico	Intermediário	Adequado
4ª EF	Abaixo de 125	125 – 150	150 – 200	Acima de 200
8ª EF	Abaixo de 150	150 – 175	175 – 250	Acima de 250
3ª EM	Abaixo de 175	175 – 200	200 – 300	Acima de 300

FONTES: SAEMS – SEE/MS

TABELA 5A: Interpretação dos níveis da escala SAEB de Matemática segundo o Atlas de MG¹⁸

Série escolar	Básico	Recomendado
4ª EF	175 – 200	Acima de 200
8ª EF	250 – 300	Acima de 300
3ª EM	325 – 375	Acima de 375

FONTES: Atlas da Educação de MG – SEE/MG

TABELA 5B: Interpretação dos níveis da escala SAEB de Língua Portuguesa segundo o Atlas de MG

Série escolar	Básico	Recomendado
4ª EF	175 – 200	Acima de 200
8ª EF	225 – 275	Acima de 275
3ª EM	250 – 300	Acima de 300

FONTES: Atlas da Educação de MG – SEE/MG

Dentre todas as classificações apresentadas, a proposta do Atlas de Minas Gerais (MINAS GERAIS/SEE-MG, 2005) é a que opta pelo menor número de níveis,

¹⁸ No Atlas da Educação de MG, os autores propõem apenas duas linhas de corte, denominadas “básico” e “recomendado”, conseqüentemente, aqueles alunos que tiverem proficiência inferior à primeira linha de corte estariam em uma categoria, não representada aqui, mas denominada “abaixo do básico”.

provavelmente por considerar que uma classificação mais condensada representa um número mais adequado de índices de qualidade e favorece a sua utilização pelos gestores públicos. Tendo como justificativa a ausência de estudos que embasassem a mudança dos pontos de corte, este documento optou por escolher os mesmos limites propostos pelo CAED, nos PROEB's 2002 e 2003, respeitando assim a competência acumulada desse grupo de especialistas. Apenas os adjetivos foram trocados segundo a máxima de que nome também é política.

Mais recentemente, a Secretaria de Estado de Educação de São Paulo encomendou um documento que subsidiasse as discussões preparatórias para a escolha de metas para as escolas estaduais paulistas. Seria produzido um indicador que permitisse a verificação do cumprimento de tais metas, visando à melhoria do ensino oferecido por essas escolas. Esse documento (MENEZES FILHO; SOARES; TERRA, 2008) não só define a métrica a ser usada para medir o sucesso das escolas, como também considera uma interpretação pedagógica dessas metas, proporcionando, assim, mais uma forma de apresentação de níveis para os desempenhos de alunos na escala do SAEB.

É fortemente desejável que as políticas públicas educacionais sejam formuladas com base em resultados produzidos, porém a discussão anterior gera um sentimento de insegurança, uma vez que não existe ainda um padrão de interpretação para a escala do SAEB, nossa principal referência no que se refere ao conjunto de aprendizados dos alunos. Conseqüentemente, a sensação da qualidade de um sistema irá variar dependendo da interpretação adotada.

De certa forma, percebe-se que a classificação adotada no SAEB 2003 é um tanto quanto rigorosa, colocando os resultados em níveis bastante pessimistas. As demais classificações não são idênticas, apesar de apresentarem uma tendência de proximidade. A proposta das metas de desempenho para as escolas de São Paulo é, talvez, a mais consistente, tendo em vista a forma de escolha dos níveis, mas, mesmo assim, não se vai aqui adotar um posicionamento sobre qual seria a melhor classificação. O esforço foi feito para explicitar que a convivência de diversas classificações, para uma mesma escala, não é benéfica.

Uma vez que a escala de proficiência é única, é necessário que sua interpretação também seja única. Ou é aceitável, até mesmo do ponto de vista ético, que os alunos, por exemplo, do Nordeste, saiam sem saber ler da 4ª série do ensino fundamental enquanto os do Sudeste devem dominar até algumas habilidades de leitura mais complexas? Sendo assim, todos os esforços deveriam ser feitos para se

decidir por uma classificação dos níveis que fosse padrão e que surgisse embasada no pedagógico e não apenas em escolhas empíricas ou conveniências políticas imediatas.

Defende-se, portanto, neste capítulo, não só a adoção em nível nacional da escala de proficiência do SAEB, como também uma adoção única para a divisão e classificação dessa escala em níveis interpretados. Muitos investimentos têm sido empregados na produção de resultados de avaliação. Talvez fosse o momento de concentrar parte dessa verba para melhorar o processo de divulgação desses resultados, principalmente, no que se refere aos aspectos pedagógicos dessa divulgação.

2.2.4 Descrição das proficiências do sistema educacional brasileiro no SAEB

Esta sessão não tem a pretensão de cobrir todas as possibilidades de análise da proficiência dos alunos nos cinco ciclos do SAEB em estudo, mas de fornecer um panorama do sistema educacional brasileiro e de sua evolução nesses últimos anos.

As Tabelas 6 e 7 sintetizam, respectivamente, a medida da proficiência média em Matemática e em Leitura de todos os alunos do sistema educacional brasileiro ao longo desses cinco ciclos.

TABELA 6: Proficiência média em Matemática por série e por ciclo do SAEB

Série escolar	SAEB 1995	SAEB 1997	SAEB 1999	SAEB 2001	SAEB 2003
4 ^a EF	-	184	176	171	174
8 ^a EF	245	244	243	239	242
3 ^a EM	276	292	275	272	278

FONTE: Elaboração própria com os dados do SAEB – INEP/MEC

TABELA 7: Proficiência média em Leitura por série e por ciclo do SAEB

Série escolar	SAEB 1995	SAEB 1997	SAEB 1999	SAEB 2001	SAEB 2003
4 ^a EF	-	180	165	159	166
8 ^a EF	247	245	231	232	230
3 ^a EM	282	282	261	258	266

FONTE: Elaboração própria com os dados do SAEB – INEP/MEC

Retomar a classificação dos níveis, proposta pelo próprio INEP em 2003 e descrita nas Tabelas 2A e 2B para se analisar a média das proficiências, seria um equívoco. Os níveis de interpretação da escala são para alunos e não para média; nenhuma combinação série/disciplina apresentaria proficiência média no nível “adequado”. Exigir resultados médios muito elevados para um sistema ou mesmo para escolas seria uma injustiça. Trata-se de uma distribuição de proficiências, então, o mais sensato seria exigir que essa curva de distribuição se deslocasse para a direita, ou, em outras palavras, que cada vez mais tivéssemos alunos nos níveis mais altos da escala. Portanto, uma maneira mais adequada de se trabalhar com os níveis e os resultados de proficiência seria observando o percentual de aluno por nível da escala SAEB. Esse percentual, em cada um dos cinco ciclos, pode ser visto nas Tabelas 8 e 9, e aí sim, uma análise mais correta pode ser feita a respeito da qualidade do sistema educacional brasileiro.

TABELA 8: Percentual de alunos por nível da escala SAEB de Matemática, segundo o SAEB 2003

Série – Ciclo	Muito crítico	Crítico	Intermediário	Adequado	Avançado
4 ^a EF – 1995	2,9	40,8	49,9	6,4	0,1
8 ^a EF – 1995	6,1	50,0	41,9	1,9	0,1
3 ^a EM – 1995	4,3	65,8	24,8	5,1	-
4 ^a EF – 1997	4,8	41,3	46,8	7,0	0,1
8 ^a EF – 1997	5,9	51,1	40,8	2,1	0,1
3 ^a EM – 1997	2,4	55,8	30,9	10,8	-
4 ^a EF – 1999	7,1	45,2	43,4	4,2	0,0
8 ^a EF – 1999	5,3	53,1	40,0	1,5	0,1
3 ^a EM – 1999	4,9	65,5	24,1	5,5	-
4 ^a EF – 2001	13,6	44,4	37,2	4,7	0,0
8 ^a EF – 2001	7,1	54,4	36,5	1,9	0,1
3 ^a EM - 2001	5,3	65,4	24,1	5,1	-
4 ^a EF - 2003	11,6	43,0	40,4	5,0	0,0
8 ^a EF - 2003	7,8	51,5	38,3	2,3	0,2
3 ^a EM - 2003	7,1	61,2	24,5	7,3	-

FONTE: Elaboração própria com os dados do SAEB – INEP/MEC

TABELA 9: Percentual de alunos por nível da escala SAEB de Leitura, segundo o SAEB 2003

Série - Ciclo	Muito crítico	Crítico	Intermediário	Adequado	Avançado
4 ^a EF - 1995	10,2	32,8	49,2	7,8	-
8 ^a EF - 1995	4,3	13,8	65,9	15,9	0,2
3 ^a EM - 1995	3,6	24,7	60,8	10,9	-
4 ^a EF - 1997	8,2	41,0	44,2	6,6	0,0
8 ^a EF - 1997	2,4	15,4	69,0	12,9	0,3
3 ^a EM - 1997	2,8	25,0	60,8	11,4	-
4 ^a EF - 1999	17,1	46,0	33,4	3,5	0,0
8 ^a EF - 1999	2,5	23,0	67,8	6,5	0,1
3 ^a EM - 1999	2,0	43,5	48,8	5,7	-
4 ^a EF - 2001	24,4	40,0	32,4	3,3	-
8 ^a EF - 2001	4,5	21,1	66,1	8,2	0,1
3 ^a EM - 2001	5,7	39,7	50,1	4,4	-
4 ^a EF - 2003	19,4	39,1	37,7	3,8	-
8 ^a EF - 2003	4,8	22,3	64,7	8,1	0,1
3 ^a EM - 2003	3,9	35,6	54,7	5,9	-

FONTE: Elaboração própria com os dados do SAEB – INEP/MEC.

Inicialmente, as Tabelas 8 e 9 mostram que o desempenho cognitivo dos alunos do sistema brasileiro de educação básica é muito ruim. Isso pode ser constatado no grande povoamento dos níveis mais críticos em detrimento de uma pequena representação de alunos nos níveis mais altos. Os dados mostram, também, uma tendência de aumento do percentual de alunos nos níveis classificados como 'muito crítico' e 'crítico' e uma pequena queda no percentual do nível considerado como 'adequado', para todas as séries, ao longo dos cinco ciclos, ou seja, o desempenho é ruim e está piorando!

Uma grande vantagem de um resultado comparável entre séries é a possibilidade de se analisar o aumento na proficiência ao longo dos anos de estudo. Essa informação poderia ser traduzida, ainda que precariamente, como a quantidade de pontos que os alunos aumentam por ano em sua proficiência. Diz-se 'traduzida precariamente', pois o ideal, para se chegar a esse resultado, seria uma medida longitudinal, isto é, uma medida do aprendizado recente dos alunos. Isso é feito aplicando-se dois testes para um mesmo grupo de alunos, sendo um no início e outro no final do período que se deseja avaliar. A diferença entre essas duas proficiências está, certamente, muito mais associada ao valor agregado pelo aluno, ao longo de um

período, do que à simples diferença entre a proficiência média de duas séries. Para se fazer esse cálculo com os dados do SAEB, tendo em vista que a medida não é feita com o mesmo aluno (quando ele está, por exemplo, na 4ª série e depois quando ele chega na 8ª série), é necessária uma suposição de que a população de alunos, em uma determinada série, não varia muito suas características de um ano para outro. Aceita essa suposição, os dados do SAEB nos permitem construir as seguintes tabelas:

TABELA 10: Diferença da proficiência média entre séries, em Matemática, por ciclo do SAEB

Série escolar	SAEB 1995	SAEB 1997	SAEB 1999	SAEB 2001	SAEB 2003	Média por ano
8ª EF - 4ª EF	-	60	67	68	68	16,4
3ª EM - 8ª EF	31	48	32	33	36	12,0

FONTE: Elaboração própria com os dados do SAEB – INEP/MEC.

TABELA 11: Diferença da proficiência média entre séries, em Língua Portuguesa, por ciclo do SAEB

Série escolar	SAEB 1995	SAEB 1997	SAEB 1999	SAEB 2001	SAEB 2003	Média por ano
8ª EF - 4ª EF	-	65	66	73	64	16,8
3ª EM - 8ª EF	35	37	30	26	36	10,9

FONTE: Elaboração própria com os dados do SAEB – INEP/MEC.

As Tabelas 10 e 11 permitem visualizar um fato: os alunos brasileiros, no ensino médio, avançam menos do que avançam os alunos do ensino fundamental. Isso fica claro no momento em que se calcula, por exemplo, em Matemática, um deslocamento de mais de 16 pontos por ano entre a 4ª e a 8ª séries do ensino fundamental e de apenas 12 pontos por ano entre a 8ª série do ensino fundamental e a 3ª série do ensino médio. E a diferença é ainda maior em relação à Língua Portuguesa. Isso pode estar relacionado a uma maior dificuldade no desenvolvimento de habilidades no ensino médio (não se tem notícia de nenhum estudo que analise a velocidade de aprendizagem ao longo das séries e, portanto, nada garante que essa velocidade seja constante); ou relaciona-se a um fenômeno que parece ser natural: quanto maior o grau de especialidade de um ser humano na realização de uma determinada tarefa mais sutil é a mudança na evolução da realização dessa tarefa (três meses de natação fazem uma diferença enorme para quem está aprendendo a nadar, ao passo que não se nota quase nenhuma mudança em um nadador profissional); ou, infelizmente, pode estar relacionado a uma baixa qualidade desse

nível de ensino no país. Outros estudos precisam ser realizados para explicar esse fato.

2.3 Medida do nível socioeconômico (NSE)

2.3.1 O que é o NSE?

A estratificação social só existe porque existem desigualdades sociais, mesmo em uma sociedade que reconhece a igualdade de todos os cidadãos perante a lei. No Brasil, em especial, essas desigualdades se mostram de forma ainda mais acentuada do que em outros países desenvolvidos. A medida do NSE é, conseqüentemente, uma tentativa de se agregar estatisticamente esses estratos sociais, produzindo, assim, um resultado que possa contextualizar a medida de desempenho cognitivo dos alunos do sistema de educação básica brasileiro.

A pergunta que dá título a este tópico é bastante audaciosa, uma vez que não existe um consenso na literatura sobre a conceituação de nível socioeconômico (NSE), nem sobre a melhor forma de medi-lo em pesquisas empíricas, mas é preciso reconhecer que

em qualquer sociedade, diferenças entre os indivíduos são consistentemente observadas quanto ao lugar que eles ocupam na hierarquia social. Reconhece-se que tais diferenças se associam às oportunidades educacionais, às trajetórias ocupacionais, ao prestígio social, ao acesso aos bens e serviços, ao comportamento político e social etc. O estudo dessas diferenças, seja como um fenômeno a ser explicado ou em associação a outros fenômenos sociais, constitui uma área de grande importância nas pesquisas sociais. Mas para que as diferenças na hierarquia social possam ser apreendidas na pesquisa empírica é necessário um esforço no sentido de definir e operacionalizar medidas destas diferenças (ALVES; SOARES, 2008, p. 4).

O esforço feito por Alves e Soares (2008) para definir e operacionalizar uma medida válida e fidedigna de nível socioeconômico, a partir de dados obtidos sobre indivíduos participantes de *surveys*, em especial os alunos do pólo Belo Horizonte, participantes do Estudo Longitudinal da Geração Escolar de 2005 (Projeto GERES), é digno de nota. Não só pelo compromisso dos autores em gerar uma medida que fosse

teoricamente fundamentada à luz da literatura sociológica, mas também pelo ineditismo na forma de se operacionalizar a medida de ocupação e renda¹⁹.

Ganzeboom e Treiman (1996), por sua vez, defendem que o NSE se manifesta, primordialmente, nos indicadores de ocupação dos pais. Outros autores defendem a escolha de outras variáveis. O fato é que “na prática da pesquisa empírica, a definição de uma medida de NSE depende não só das opções teóricas do pesquisador, mas também da disponibilidade de dados adequados para esse tipo de análise” (ALVES; SOARES, 2008, p. 9-10). E é nesse contexto que se registra a não existência, nos questionários do SAEB, de variáveis sobre ocupação dos pais e local de moradia do aluno. A opção adotada, então, foi semelhante ao cálculo do Critério de Classificação Econômica Brasil, proposto pela Associação Nacional de Empresas de Pesquisa (ANEP), em que informações sobre a escolaridade dos pais foram associadas a uma medida indireta de renda, obtida a partir da inclusão de itens sobre o consumo de bens e serviços, prática que “busca reconstruir as hierarquias sociais através da destinação que as famílias ou os indivíduos dão à riqueza obtida devido a sua origem social” (ALVES; SOARES, 2008, p. 11).

O esforço implementado neste trabalho, que merece destaque pela novidade, está em se produzir uma medida desse construto, o NSE, para cada aluno das bases estudadas do SAEB, numa mesma métrica e, portanto, comparável como a proficiência. Está se falando em mais de 770.000 alunos, que tiveram esse traço latente manifestando-se por meio de indicadores de escolaridade e renda de sua família.

2.3.2 Construção da medida de NSE e seus indicadores

Assim como na produção da proficiência dos alunos, a TRI também é utilizada para se determinar o NSE de cada aluno, criando-se, portanto, uma escala de

¹⁹ Ocupação e renda foram dois dos três indicadores utilizados por Alves e Soares (2008) para calcular o nível socioeconômico dos indivíduos envolvidos. O terceiro foi a escolaridade dos pais. Em relação à ocupação, primeiramente as respostas foram codificadas de acordo com a Classificação Brasileira das Ocupações (CBO2002), em seguida, para o tratamento dessa informação, foi utilizado o *International Socio-economic Index* (ISEI) e uma hierarquização das ocupações de acordo com os estratos ocupacionais propostos por Pastore e Valle Silva (2002). Em relação à renda, foi feita uma medida indireta por meio do local de moradia do aluno, associando o Código de Endereçamento Postal (CEP) e a renda média do setor censitário associado, cujos dados são fornecidos pelo IBGE. O fato é que esse trabalho fornece ferramentas metodológicas para as pesquisas empíricas que utilizam, de alguma forma, o NSE como uma das variáveis de estudo.

posicionamento socioeconômico (SOARES; ANDRADE, 2005). Para isso, foi feito um esforço juntando-se todas as bases dos questionários contextuais dos cinco ciclos em questão. Como já foi explicado na metodologia do trabalho, são ao todo trinta²⁰ bases do SAEB (3 séries x 2 disciplinas x 5 ciclos). Essa associação de bases contextuais não é simples, pois algumas questões nem sempre aparecem em todos os questionários de todos os ciclos, além disso, às vezes, uma mesma questão é formulada de forma ligeiramente diferente de um ano para o outro, exigindo uma recodificação nos valores das respostas dadas pelos alunos, para que se possa considerá-la como um item comum entre as diversas bases; além dos devidos cuidados para que uma simples recodificação no item não altere a essência da questão. O modelo da TRI adequado para este cálculo do NSE, como explicado no item 2.1.2 deste capítulo, é o de Samejima (1969). De posse desse arquivo único, calcula-se, utilizando-se o software MULTILOG (THISSEN, 1991), um valor de NSE para cada um dos alunos, tendo em vista as respostas dadas por eles às seguintes questões:

- Quantos dos seguintes itens há no lugar onde você mora? TELEVISÃO A CORES
- Quantos dos seguintes itens há no lugar onde você mora? APARELHO DE VÍDEO CASSETE
- Quantos dos seguintes itens há no lugar onde você mora? RÁDIO
- Quantos dos seguintes itens há no lugar onde você mora? GELADEIRA
- Quantos dos seguintes itens há no lugar onde você mora? MÁQUINA DE LAVAR
- Quantos dos seguintes itens há no lugar onde você mora? ASPIRADOR DE PÓ
- Quantos dos seguintes itens há no lugar onde você mora? TELEFONE RESIDENCIAL
- Quantos dos seguintes itens há no lugar onde você mora? TELEFONE CELULAR
- Quantos dos seguintes itens há no lugar onde você mora? AUTOMÓVEL
- Quantos dos seguintes itens há no lugar onde você mora? BANHEIRO

²⁰ Na verdade, são 28 bases, pois os dados de 4ª série do ensino fundamental de 1995, tanto em Matemática como em Leitura, não puderam ser aproveitados devido ao formato do questionário contextual, o qual inviabilizou o cálculo do NSE dos alunos.

- Você possui computador em casa?
- No lugar onde você reside existe eletricidade?
- Em sua casa trabalha alguma empregada doméstica, todos os dias úteis?
Quantas?
- Em sua casa compra-se jornal?
- Número de pessoas por quarto?²¹
- Qual o nível de instrução do seu pai?
- Qual o nível de instrução da sua mãe?
- Você teve que abandonar a escola durante algum tempo?

A Tabela 12 mostra, detalhadamente, todas as variáveis que foram utilizadas para o cálculo do NSE, como as respostas dos alunos foram categorizadas e os ciclos do SAEB em que cada uma aparece.

²¹ Esta variável foi calculada a partir das respostas dadas às questões: 1) Quantas pessoas moram com você em sua casa?; 2) Quantos quartos há em sua casa?.

TABELA 12: Descrição das variáveis utilizadas no cálculo do NSE

NÚMERO	VARIÁVEL	ITENS	CATEGORIAS	CICLOS*				
				1995	1997	1999	2001	2003
1	tv cores	TV em cores	1 = Não tem 2 = Sim, uma 3 = Sim, duas 4 = Sim, três ou mais			X	X	X
2	video	Videocassete	1 = Não 2 = Sim			X	X	X
3	rádio	Rádio	1 = Não tem 2 = Sim, um 3 = Sim, dois 4 = Sim, três ou mais			X	X	X
4	geladeir	Geladeira	1 = Não tem 2 = Sim			X	X	X
5	maqlav	Máquina de lavar roupa	1 = Não tem 2 = Sim			X	X	X
6	aspirapo	Aspirador de pó	1 = Não tem 2 = Sim			X	X	X
7	telefone	Telefone residencial	1 = Não tem 2 = Tem um 3 = Tem dois ou mais			X		
8	celular	Telefone celular	1 = Não tem 2 = Tem um 3 = Tem dois ou mais			X		
9	comput	Computador	1 = Não tem 2 = Sim		X	X	X	X
10	carro	Automóvel	1 = Não tem 2 = Sim, um 3 = Sim, dois ou mais			X	X	X
11	banheiro	Banheiro	1 = Não tem 2 = Sim, um 3 = Sim, dois 4 = Sim, três ou mais			X	X	X

12	luz	Existe eletricidade	1 = Não 2 = Sim			X	X	X
13	empreg	Empregada doméstica	1 = Não tem 2 = Diarista, uma ou duas vezes por semana 3 = Uma ou mais, todos os dias úteis			X	X	X
14	jornal	Compra-se jornal diariamente	1 = Não 2 = Sim	X	X		X	X
15	pesquart	Número de pessoas por quarto	1 = Mais de três pessoas por quarto 2 = Entre 2 e 3, inclusive, pessoas por quarto 3 = Entre 1 e 2, inclusive, pessoas por quarto 4 = Uma ou menos pessoas por quarto			X	X	X
16	escolpai	Escolaridade do pai	1 = Nunca estudou 2 = 1ª a 4ª série do EF 3 = 5ª a 8ª série do EF 4 = Ensino Médio 5 = Superior	8ª EF e 3ª EM	X	X	X	X
17	escolmae	Escolaridade da mãe	1 = Nunca estudou 2 = 1ª a 4ª série do EF 3 = 5ª a 8ª série do EF 4 = Ensino Médio 5 = Superior	8ª EF e 3ª EM	X	4ª EF	X	X
18	abandono	Abandonou a escola	1 = Sim, por dois anos ou mais 2 = Sim, por até um ano 3 = Não	4ª EF e 3ª EM	X		X	X

FONTE: Elaboração própria com os dados do SAEB – INEP/MEC.

* Onde existe a série em vez de um “x” é porque, por algum motivo, esse dado não estava disponível na base inteira daquele ano.

O item 9, que pergunta se o aluno possui computador em casa, é sempre um item que fornece muita informação quando se deseja medir o NSE, fato que pode ser constatado nos gráficos abaixo. No entanto, deve-se observar também que esse item não discrimina bem os alunos dos dois extremos da escala de nível socioeconômico, certamente, porque no extremo inferior ninguém possui computador enquanto no extremo superior quase todos possuem.

GRÁFICO 9: Curva Característica do Item 9

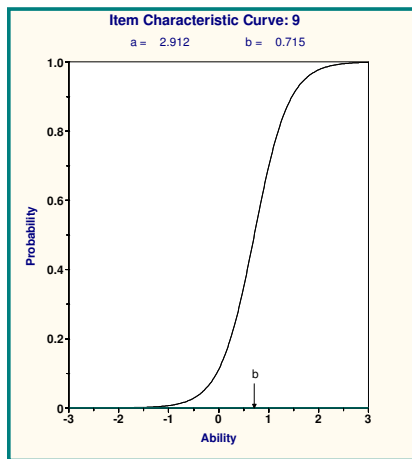
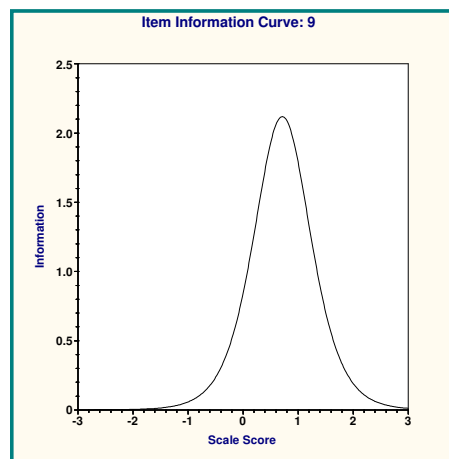


GRÁFICO 10: Curva de Informação do Item 9



FONTE: Elaboração própria com os dados do SAEB – INEP/MEC.

Já o item 17, 'qual a escolaridade da sua mãe', é o melhor exemplo de um item com cinco opções de respostas graduadas e que fornece boa informação em toda a escala de NSE.

GRÁFICO 11: Curva Característica do Item 17

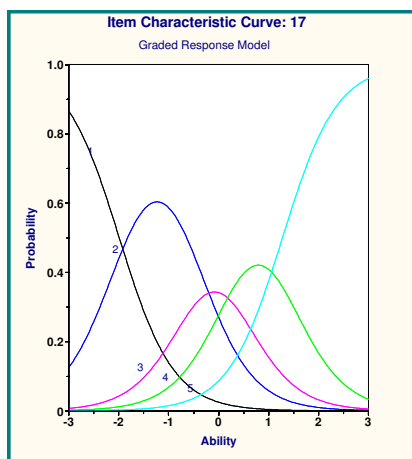
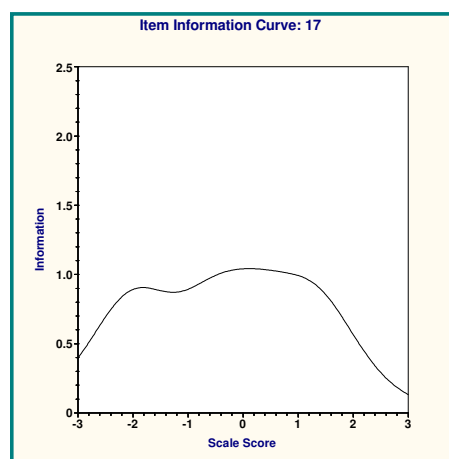


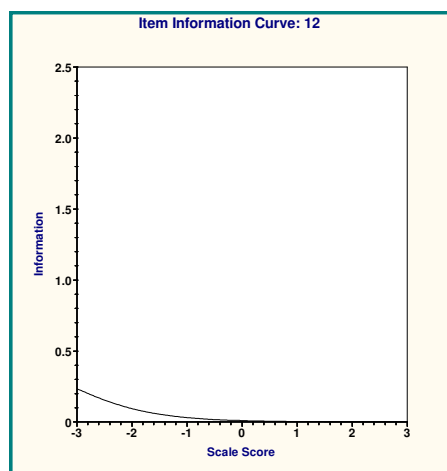
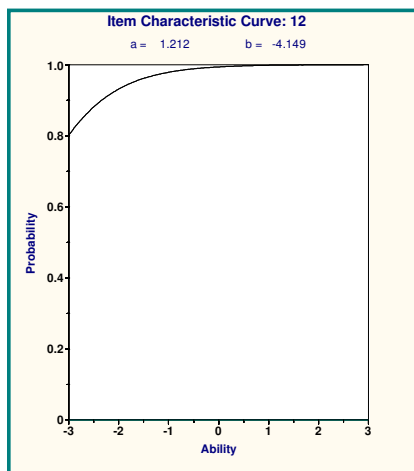
GRÁFICO 12: Curva de Informação do Item 17



FONTE: Elaboração própria com os dados do SAEB – INEP/MEC.

Outras variáveis de potencial interesse constam dos questionários contextuais, no entanto, elas foram retiradas do cálculo por não produzirem informação relevante, como se pode ver nos Gráficos 13 e 14, referentes ao item ‘cozinha’ (Quantos dos seguintes itens há no lugar onde você mora? COZINHA). Além dessa, foram retiradas também, pelo mesmo motivo, as variáveis: ‘sala’, ‘água encanada’, ‘calçamento’, ‘se o aluno exerce atividade remunerada’, ‘seu pai está trabalhando’ e ‘sua mãe está trabalhando’.

GRÁFICO 13: Curva Característica do Item 12 GRÁFICO 14: Curva de Informação do Item 12



FONTE: Elaboração própria com os dados do SAEB – INEP/MEC.

O item ‘sua casa possui eletricidade’ foi mantido, pois, apesar de fornecer pouca informação, seu pico concentra-se no extremo esquerdo da escala, região carente de uma medida de melhor qualidade. Já as variáveis ‘jornal’ e ‘abandonou a escola’ foram mantidas, mesmo não fornecendo boa informação, porque eram os dois únicos itens restantes da base de 4^a série de 1995.

No entanto, após a análise final da medida de NSE, verificou-se que, para essa base, o NSE era extremamente instável, com uma qualidade duvidosa, e optou-se por excluir a 4^a série desse ciclo. Isso só corrobora a importância dos itens sobre escolaridade para o cálculo deste construto, uma vez que essa pergunta não figurou no questionário contextual daquele ano. Por isso todas as análises dos capítulos subsequentes incluem 28 bases, já que as bases de Matemática e Língua Portuguesa da 4^a série de 1995 estão sem a variável explicativa.

Além de todos os alunos da 4^a série de 1995, também foram retirados do arquivo único, para o cálculo do NSE, aqueles alunos que apresentam dado ausente

em todas as questões. Com isso, o arquivo base para o cálculo do NSE contém 772.964 casos.

Além da equalização via população, implementada pela TRI, ao rodar todas as bases conjuntamente, outra vantagem de utilizar essa teoria para gerar a medida de NSE é o fato de que ela permite o cálculo apesar de alguns dados ausentes que existam em algumas variáveis. Como se pode observar na Tabela 11, nenhuma variável está presente em todas as bases, mas os cuidados para se garantir itens comuns entre algumas delas e rodar toda base conjuntamente propiciam a tranqüilidade de que a medida de NSE produzida estará equalizada, ou seja, estará toda em uma mesma métrica podendo ser comparada.

2.3.3 Descrição do NSE a partir dos dados do SAEB

Como explicado no item 2.2.1 deste capítulo, ao se utilizar a TRI, deve haver a preocupação em se apresentar os resultados numa escala, cujos valores sejam mais simples para se compreender. No caso da medida de NSE, produzida aqui, obteve-se, inicialmente, valores que variavam entre -2 e +2. Então, decidiu-se por transformar essa escala de tal forma que o menor NSE encontrado no grupo fosse 0 e o maior fosse 10. A transformação linear adotada foi:

$$NSE_novo = 2,064 \times NSE_original + 4,976$$

Uma síntese dos resultados obtidos com esse cálculo do NSE está na tabela abaixo, a qual mostra o aumento do NSE com o aumento da escolaridade em todos os ciclos analisados. Isso por si só já é um reflexo da realidade brasileira, na qual os menos favorecidos socioeconomicamente abandonam a escola mais cedo para trabalhar.

TABELA 13: Média do NSE* por série e por ciclo do SAEB

Série escolar	SAEB 1995	SAEB 1997	SAEB 1999	SAEB 2001	SAEB 2003
4ª EF	-	4,67	4,21	4,71	4,70
8ª EF	4,44	4,81	4,64	4,98	4,96
3ª EM	4,87	5,05	4,84	5,05	5,14
Total	4,60	4,79	4,50	4,89	4,89

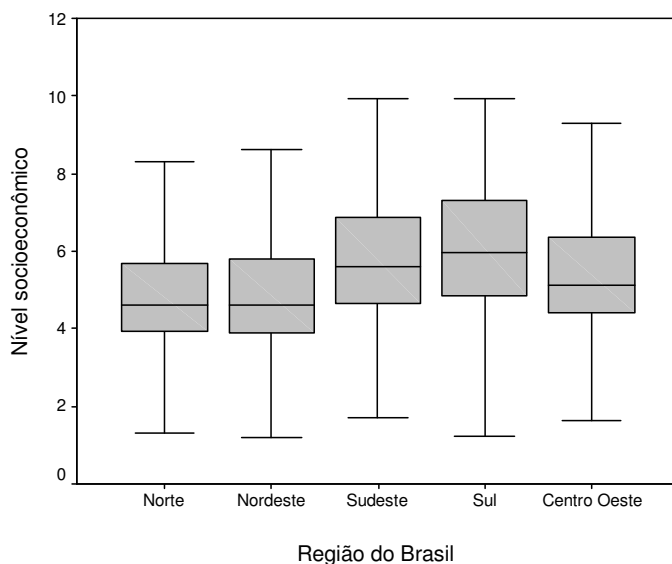
FONTE: Elaboração própria com os dados do SAEB – INEP/MEC

* Numa escala que vai de 0 a 10.

Em países desenvolvidos, como a Finlândia, por exemplo, as escolas são muito semelhantes entre si. No Brasil, menos semelhantes, ainda que a semelhança ocorra em alto grau, como será visto no próximo capítulo. No entanto, a dependência entre o desempenho cognitivo dos alunos e o seu nível socioeconômico (NSE) é marcante em países como o Brasil. Esse fato poderia ser inferido, ainda que de forma precária, da Tabela 13, uma vez que se observa uma relação entre a quantidade de anos que um aluno consegue permanecer na escola e o seu NSE.

Já o gráfico de caixas abaixo mostra um retrato importante do Brasil. Observa-se que as regiões Norte e Nordeste, além de concentrarem os índices de menor proficiência, possuem também os menores índices de NSE. Isso era de se esperar, considerando a forte relação entre esses dois construtos. Em contrapartida, as regiões Sul e Sudeste concentram os maiores valores de NSE, mas também as maiores variações.

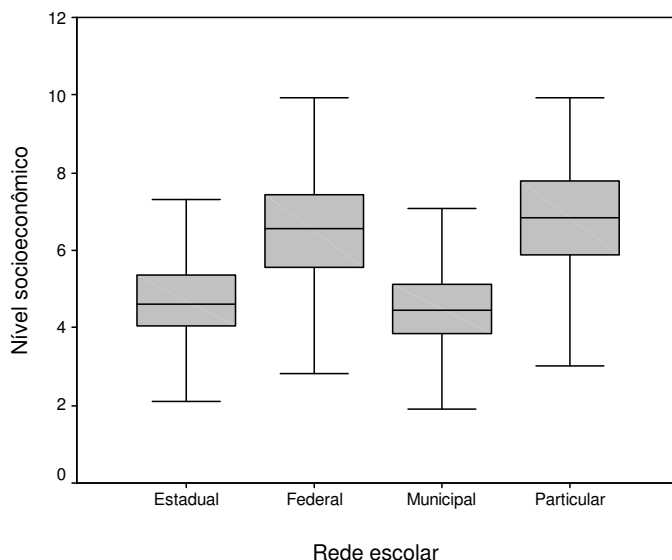
GRÁFICO 15: Nível socioeconômico por região do Brasil



FONTE: Elaboração própria com os dados do SAEB 2003 – INEP/MEC. Amostra aleatória de 10%.

E o Gráfico 16 mostra como esses alunos se distribuem nas quatro redes de escolas: Estadual, Federal, Municipal e Particular. Como era esperado, as redes Federal e Particular atraem alunos de maior NSE se comparado aos alunos que estão nas redes Estadual e Municipal. Essa distinção do público atendido por cada rede precisa ser considerada no momento em que se deseja, por exemplo, calcular o efeito de cada uma das redes sobre o desempenho cognitivo dos seus alunos.

GRÁFICO 16: Nível socioeconômico por dependência administrativa



FONTE: Elaboração própria com os dados do SAEB 2003 – INEP/MEC. Amostra aleatória de 10%

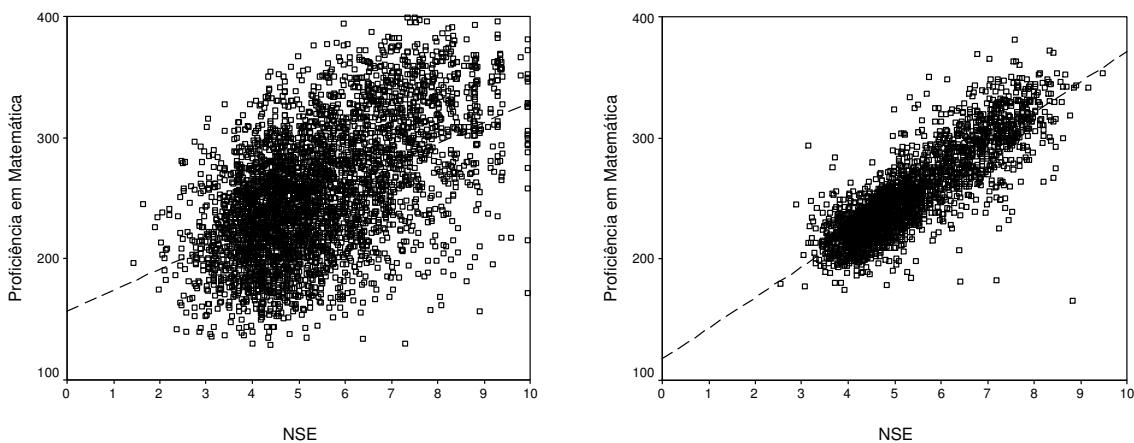
Nos próximos capítulos, ambas as medidas aqui descritas, a proficiência e o nível socioeconômico, serão utilizadas em modelos matemáticos para tentar descrever o sistema educacional brasileiro a partir dos dados do SAEB.

2.3.4 Análise da correlação entre NSE e proficiência

Para este trabalho, ainda mais pertinente do que descrever o NSE, de acordo com os dados do SAEB, é analisar a possibilidade de correlação entre esse construto e a proficiência dos alunos.

Os gráficos a seguir mostram uma forte associação existente entre o NSE e o desempenho cognitivo do aluno, justificando, conseqüentemente, a necessidade do seu cálculo para análise do desempenho. O Gráfico 17 mostra a relação entre essas duas variáveis no nível do aluno, enquanto o Gráfico 18 está no nível da escola.

GRÁFICO 17: Proficiência em Matemática na 8ª série do ensino fundamental *versus* NSE – Nível do aluno
 GRÁFICO 18: Proficiência média em Matemática na 8ª série do ensino fundamental *versus* NSE – Nível da escola



FONTE: Elaboração própria com os dados do SAEB 2003 – INEP/MEC.

Como era de se esperar, há uma maior variação no nível do aluno do que no nível da escola quando o dado se encontra agregado. Porém fica nítida a necessidade de se considerar a influência do NSE em medidas de efeito da escola no Brasil. Esta influência está presente no nível do aluno, o que pode ser visto a partir da reta auxiliar de regressão linear, que foi traçada no Gráfico 17, porém ela é ainda maior no nível da escola, como mostra a maior inclinação da reta de regressão linear. Desconsiderar o efeito do nível socioeconômico do aluno e, principalmente, da escola na qual ele estuda implica em se obter uma medida mascarada pela segmentação socioeconômica existente no país.

Essa dependência entre NSE e desempenho é impactante, mesmo que não determinante, como mostram Soares e Andrade (2006), em um estudo feito com escolas de Belo Horizonte. Por isso, no Brasil, grande parte da explicação para a heterogeneidade das escolas é a própria heterogeneidade socioeconômica dos alunos. A consequência desse fato é que as medidas de interesse devem ser, então, controladas pelo NSE dos alunos e pelo NSE médio da escola que ele frequenta.

O efeito visual mostrado nos gráficos anteriores é importante, mas o diferencial deste trabalho encontra-se, principalmente, na análise da continuidade de um fenômeno, ou seja, o estudo de um fenômeno a partir de várias bases do SAEB. A Tabela 14 sintetiza, para as 28 bases estudadas, essa correlação entre desempenho cognitivo dos alunos e o seu NSE.

TABELA 14: Correlação de Pearson entre Proficiência e NSE

Ano	Matemática			Língua Portuguesa		
	4ª EF	8ª EF	3ª EM	4ª EF	8ª EF	3ª EM
1995	-	0,414	0,426	-	0,334	0,363
1997	0,397	0,505	0,521	0,354	0,395	0,428
1999	0,439	0,471	0,506	0,417	0,369	0,446
2001	0,484	0,535	0,557	0,440	0,455	0,474
2003	0,517	0,544	0,564	0,461	0,447	0,488

FONTE: Elaboração própria com os dados do SAEB – INEP/MEC.

* Todos os valores apresentados nesta tabela são significativos no nível 0,01.

Os dados numéricos corroboram o que estava representado no gráfico, mostrando a correlação existente entre proficiência e NSE. Essa correlação parece aumentar entre o primeiro e o último ciclo analisado e é nitidamente maior em Matemática do que em Leitura.

Essa correlação, sempre positiva, entre os dois construtos, é que justifica a necessidade, principalmente no Brasil, de se controlar as análises de desempenho pelo nível socioeconômico. Qualquer iniciativa diferenciada estaria, no mínimo, mascarando uma realidade na qual o social é um dos principais fatores explicativos dos resultados dos alunos.

Chama-se uma especial atenção para o item 2.4, a seguir, em que será apresentada uma discussão sobre os modelos apropriados para análise desse tipo de dado. Independente do modelo a ser escolhido, percebe-se uma coerência na literatura nacional, na qual os autores são unânimes em controlar o desempenho pelo NSE, como se pode ver em Soares, César e Mambrini (2001), Barbosa e Fernandes (2001), Fukuda (2003), Jesus (2004) e Andrade e Laros (2007).

2.4 Modelos de regressão multinível

2.4.1 Necessidade de uso desses modelos em educação

Se este trabalho fosse na área da agricultura, outras metodologias de análise poderiam, facilmente, ser aplicadas. Por exemplo, para se testar o efeito de um determinado agrotóxico, poder-se-ia optar por manter uma cultura de controle, na qual não se utiliza o produto, e outra cultura, com especificações semelhantes, mas em que

o agrotóxico seria aplicado. E, ao final, comparar-se-ia o resultado obtido nos dois plantios. Ainda que o processo de pesquisa também não seja tão simples assim na agricultura, na área educacional tais metodologias são especialmente complicadas de serem implementadas. Para sanar tais dificuldades, a estatística desempenha um papel fundamental nas Ciências Sociais.

Uma das técnicas estatísticas mais utilizadas para se analisar a relação entre variáveis é o modelo de regressão, mais popularmente o modelo de regressão linear. Implementar um modelo desse tipo é, na verdade, buscar uma relação matemática entre dois conjuntos de dados, as variáveis. Essa relação matemática pode ser descrita pela equação a seguir:

$$Y_i = \beta_0 + \beta_1 * x_i + r_i \quad (3)$$

Nesse modelo, o “Y” denota o conjunto de dados que se pretende explicar, ou seja, a variável dependente; o “x” é a variável independente, portanto a parte $\beta_0 + \beta_1 * x_i$ é a componente explicativa do modelo ou preditor linear. Como a equação anterior apresenta apenas uma variável explicativa, pode-se denominá-la de modelo de regressão linear simples, mas, normalmente, pretende-se estudar o efeito de várias variáveis e, neste caso, o modelo seria de regressão linear múltipla. Na prática, os dados não “caem” exatamente sobre a reta de regressão e, por isso, é necessário acrescentar ao modelo um termo que represente esse erro. O “r” da equação mede justamente a diferença entre o “Y” observado e “Y” teórico, fornecido pela equação da reta, e sua estimativa é conhecida como resíduo.

Pode-se afirmar, então, que o principal objetivo, ao se utilizar essa técnica estatística, é o de ajustar o modelo aos dados. Isso significa estimar os parâmetros desconhecidos β_0 , β_1 e demais betas, caso a regressão seja múltipla. Vários modelos podem se ajustar ao conjunto de dados, o melhor ajuste será aquele que torne os resíduos tão próximos de zero quanto possível.

Registre-se aqui que uma relação entre duas variáveis, explicitada por um modelo de regressão, ainda que muito forte, não implica na existência de causalidade. A relação causa-efeito, se necessária, precisa ser buscada por outros meios.

No entanto, esta técnica de regressão linear múltipla não considera o fato, corriqueiro em situações escolares, que os diferentes alunos de uma turma ou escola se influenciam mutuamente e, portanto, suas proficiências não são independentes. Ou

seja, o dado educacional é naturalmente mais complexo e exige assim instrumentos de modelagem, com um nível comparável de complexidade (GOLDSTEIN, 2001, p. 86). A solução encontrada e desenvolvida, concomitantemente, por Goldstein, no Instituto de Educação da Universidade de Londres, e Bryk e Randenbush, na Universidade de Harvard, é hoje denominada modelos hierárquicos ou multiníveis, que se tornaram padrão em pesquisa educacional empírica (LEE, 2001b).

Esses modelos incorporam

naturalmente, e de uma forma parcimoniosa, a estrutura hierárquica ou de agrupamento da população em estudo, tratando o intercepto e os coeficientes de inclinação como variáveis aleatórias. Desta forma, o modelo permite a variabilidade das estimativas entre os grupos (FERRÃO, 2003, p. 31).

Esta técnica estatística permite captar a complexidade da relação entre os fatores de cada um dos níveis e como esses níveis se influenciam mutuamente. Além disso, os dados de alunos são utilizados nos modelos de análise, mas o interesse analítico é a organização escolar (SOARES *et al*, 2004)²², ou, como argumenta Lee (2001a), existem algumas questões de pesquisa que são intrinsecamente multiníveis, e, portanto, utilizar técnicas analíticas de um único nível implica em uma série de limitações. A modelagem multinível é que proporciona uma abordagem capaz de lidar adequadamente com questões multiníveis. Para detalhes técnicos sobre os modelos multinível, sugere-se Ferrão (2003), Raundenbush e Bryk (2002) e Natis (2001).

Nas análises multinível deve-se especificar a que nível a variável pertence e quais os efeitos diretos, bem como quais os efeitos de interação entre os níveis que são esperados. Em relação à medição das variáveis, estas podem ser feitas diretamente no próprio nível da variável e, também, por agregação ou desagregação” (Andrade; Laros, 2007, p. 34).

Dependendo dessas escolhas, vários tipos de modelos podem ser implementados. Os modelos multinível mais utilizados são aqueles em que apenas o intercepto é aleatório e a variância da variável dependente é decomposta entre os níveis, denominados de componentes de variância; e aqueles em que se testa o efeito randômico dos coeficientes de inclinação, denominados coeficientes aleatórios (Ferrão, 2003).

²² O que é muito relevante, tendo em vista que “esses modelos produziram uma solução para o sério problema da unidade de análise, cujo equacionamento limitou, durante anos, a análise de dados provenientes de organizações” (Soares *et al*, 2004. p. 21).

A equação abaixo ilustra, de forma introdutória, o que vem a ser um modelo de regressão multinível. No próximo item, será detalhado o modelo base para as análises desta tese.

$$\text{Nível 1: } Y_{ij} = \beta_{0j} + \beta_{1j} * X_{ij} + r_{ij} \quad (4)$$

$$\text{Nível 2: } \beta_{0j} = \gamma_{00} + \gamma_{01} * Z_{ij} + u_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

onde:

- Y_{ij} é a variável dependente. No caso desta pesquisa, a proficiência em Matemática ou em Leitura, do i -ésimo aluno da escola j ;
- β_{0j} é o intercepto, ou seja, a proficiência média dos alunos quando X_{ij} é zero;
- β_{1j} é o coeficiente de inclinação, ou seja, é a mudança esperada na proficiência quando X_{ij} aumenta uma unidade;
- X_{ij} é a primeira variável explicativa do nível do aluno, isto é, do nível 1;
- r_{ij} é o resíduo associado ao i -ésimo aluno da escola j , também denominado de resíduo do nível 1. É a parte da proficiência não explicada pelo modelo;
- γ_{00} é a média dos interceptos de j escolas;
- Z_{ij} é a primeira variável explicativa do nível da escola, isto é, do nível 2;
- u_{0j} é o efeito associado a j -ésima escola, também denominada de variância residual do nível 2;
- γ_{10} é a média dos coeficientes de regressão de j escolas para a primeira variável do nível 1.

2.4.2 O modelo escolhido

A escolha do modelo hierárquico de análise, ou melhor, das variáveis, tanto do nível do aluno como do nível da escola, que constarão desse modelo, deve refletir os objetivos da questão de pesquisa e, assim, constitui-se em uma etapa fundamental de todo o processo de modelagem dos dados.

Análises feitas previamente sobre a influência do nível socioeconômico no desempenho escolar dos alunos justificam que o estudo de qualquer realidade educacional, em especial a brasileira, deve, necessariamente, considerar a condição socioeconômica dos alunos (SOARES; ANDRADE, 2005). Além disso, suas características sociodemográficas, tais como o sexo, a cor/raça e sua história escolar, aqui representada pelo atraso escolar e pela sua atitude frente aos estudos de Matemática e Língua Portuguesa, sintetizam importantes experiências vividas e que impactam, ainda que de forma distinta, o seu desempenho. Sendo assim, tais variáveis também devem ser incluídas como controle no modelo de análise.

Mas não são apenas as variáveis individuais que afetam a proficiência dos alunos, ao contrário, o efeito coletivo dessas características tem um impacto ainda maior.

Este é um ponto importante e freqüentemente pouco apreciado. O impacto dos fatores sociais do coletivo freqüentado pelo aluno, por exemplo, sua escola, no desempenho do aluno é maior do que no nível individual. Ou seja, o aluno que convive com colegas de alta condição social ou cultural é particularmente privilegiado. Seguindo as recomendações de Willms (2000), incluímos entre as características das escolas utilizadas como controle a média do nível socioeconômico e a média do atraso escolar dos alunos da escola. Willms justifica o uso desses fatores citando estudos realizados em vários países, que mostram que escolas e mesmo salas de aula com alunos de posição social mais alta e/ou de maior nível cognitivo tendem a desfrutar de várias vantagens associadas ao contexto criado por esses alunos. Em média, essas escolas têm maior apoio dos pais, menor número de problemas disciplinares e um clima que valoriza a obtenção de melhores resultados. Além disso, com maior facilidade atrai e mantém professores talentosos e motivados. Todas essas condições caracterizam o que na literatura é chamado de "efeito dos pares", observado quando alunos privilegiados social e culturalmente freqüentam a mesma escola.[...] A não consideração de todos estes fatores de controle é fonte geradora de muitas imprecisões principalmente na grande imprensa, onde muitos fatores apontados como decisivos para a melhoria do sistema educacional são apenas expressões destes fatores antecedentes (SOARES *et al*, 2004, p. 21).

Concordando com essa visão, este trabalho utiliza, como variáveis de controle, no nível da escola, a média do nível socioeconômico, a média do atraso escolar dos alunos e a dependência administrativa (rede) da escola.

No entanto, é preciso esclarecer que não há um consenso sobre a escolha das variáveis de controle. Em algumas situações, argumenta-se que o controle pelo NSE da escola pode obscurecer o bom trabalho realizado pelo corpo docente das escolas, que atraem os alunos de maior NSE. Neste trabalho, entretanto, tais escolhas se

justificam já que produzem um teste bastante rigoroso para a importância de qualquer fator escolar. Adota-se, portanto, a lógica do efeito tipo B, proposta por Willms e Raudenbush (1989).²³

Na literatura brasileira, ainda não há um consenso sobre o uso de dois níveis nos modelos de análise. Gremaud, Felício e Biondi (2006) propõem um modelo com apenas um nível, o da escola. Neste trabalho, optou-se por manter dois níveis – o do aluno e o da escola. Evita-se, assim, incorrer em um vício de agregação, que afetaria a estimativa do efeito das escolas, o que poderia ocorrer caso fosse utilizado um modelo de um nível com os resultados dos alunos agregados.

O modelo apresentado na equação (5), em dois níveis, é utilizado para o cálculo do efeito da escola em Matemática e Leitura. Nele, destaca-se a opção por retirar da proficiência do aluno o efeito das variáveis demográficas, variáveis sociológicas e variáveis que sintetizem a trajetória escolar desse aluno, completamente fora do controle da escola.

$$\begin{aligned} \text{Nível 1:} \quad Y_{ij} = & \beta_{0j} + \beta_{1j} NSE_{ij} + \beta_{2j} SEXO_{ij} + \beta_{3j} PARDO_{ij} + \beta_{4j} PRETO_{ij} + \\ & \beta_{5j} ATRASO_{ij} + \beta_{6j} DISCIPLINA_{ij} + \beta_{7j} ATITUDE_{ij} + r_{ij} \end{aligned} \quad (5)$$

$$\begin{aligned} \text{Nível 2:} \quad \beta_{0j} = & \gamma_{00} + \gamma_{01} MNSE_j + \gamma_{02} REDE_j + \gamma_{03} MATRASO_j + u_{0j} \\ \beta_{1j} = & \gamma_{10} \\ \beta_{2j} = & \gamma_{20} \\ \beta_{3j} = & \gamma_{30} \\ \beta_{4j} = & \gamma_{40} \\ \beta_{5j} = & \gamma_{50} \\ \beta_{6j} = & \gamma_{60} \end{aligned}$$

²³ “Willms e Raudenbush (1989) definem dois tipos de estimativa do efeito das escolas: efeito tipo A e efeito tipo B. Para estimar o efeito Tipo A, são usados modelos estatísticos que incluem fatores de controle associados ao aluno, como uma medida do nível socioeconômico e de seu desempenho prévio. O efeito tipo A corresponde à estimativa do desempenho escolar, esperado para um aluno com características “médias”, matriculado na escola X comparando com o desempenho médio de todos os alunos. Por exemplo, se a média geral dos alunos é 7,1 e o efeito da escola X é 0,8, então podemos estimar uma nota de 7,9 para este aluno. Esse efeito tem interesse para os pais que querem ter uma expectativa do resultado da escola sobre os seus filhos. A estimativa do efeito Tipo B requer modelos que agreguem, além dos fatores de controle usados para o cálculo do efeito Tipo A, fatores contextuais, que fogem ao controle da escola, oriundos, por exemplo, da composição do corpo discente. O efeito Tipo B estima qual o desempenho da escola X em comparação com outras escolas com contextos semelhantes. Esse efeito interessa mais aos atores envolvidos nas decisões sobre políticas e práticas educacionais (diretores, professores e gestores públicos), que podem obter uma medida comparativa das escolas, equalizadas por suas respectivas clientela” (ALVES, 2006).

$$\beta_{7j} = \gamma_{70}$$

$$\text{Onde } r_{ij} \sim N(0, \sigma^2) \text{ e } \begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{pmatrix} \right]$$

onde:

- Y_{ij} é a proficiência, em Matemática ou em Leitura, do i -ésimo aluno da escola²⁴ j ;
- β_{0j} é a proficiência média dos alunos da escola j ;
- β_{1j} é o efeito do NSE do aluno na sua proficiência;
- NSE_{ij} é o NSE do aluno centrado na média do grupo, isto é, quão distante da média da escola o NSE do aluno está;
- $SEXO_{ij}$, $PARDO_{ij}$, $PRETO_{ij}$, $DISCIPLINA_{ij}$ e $REDE_{ij}$ são variáveis indicadoras e estão mais bem descritas na tabela a seguir;
- $ATRASSO_{ij}$ é uma medida discreta do número de anos que o aluno está atrasado em relação à série adequada para a sua idade;
- $ATITUDE_{ij}$ é uma medida da postura do aluno frente ao ensino, variável que foi centrada na média do grupo, isto é, quão distante da média da escola a $ATITUDE$ do aluno está;
- $MATRASSO_j$ é a média do atraso escolar dos alunos que compõem a escola j ;
- γ_{00} é a proficiência média dos alunos da população em questão;
- γ_{01} é o efeito do NSE da escola na proficiência do aluno;
- $MNSE_j$ é o NSE médio da escola j , centrado na média total, isto é, quão distante da média geral o NSE da escola está;
- u_{0j} é o efeito associado a j -ésima escola, com variância σ_{u0}^2 entre escolas, isto é, $\sigma_{u0}^2 = \text{Var}(u_{0j})$;
- r_{ij} é o resíduo associado ao i -ésimo aluno da escola j , com variância σ_r^2 intra-escolas, isto é, $\sigma_r^2 = \text{Var}(r_{ij})$.

²⁴ O SAEB é um estudo, no qual os alunos não formam uma amostra da escola. Tem-se, na verdade, uma amostra de turmas da escola. Esse ponto é crucial para a definição dos modelos a serem considerados, pois o mais correto seria falar em “efeito turma”, e não em “efeito-escola”. Propõe-se, então, a idéia de pseudo-escola para cada turma amostrada. Apesar disso, por abuso de linguagem, continuaremos a nos referir ao “efeito das escolas”.

A inclusão da variável disciplina é uma necessidade quando são utilizados os dados do SAEB. Trata-se de uma variável indicadora, que permite o cálculo de um efeito único por escola, com um número maior de alunos. Quando pertinente, a variável *disciplina* é excluída do modelo e são trabalhadas, separadamente, as bases com as proficiências em Matemática e em Leitura. Essa variável não está presente no modelo como uma variável de controle do aluno. Sua inserção apenas possibilita o trabalho com uma base única, na qual a variável dependente passa a ser o desempenho cognitivo do aluno. A metodologia amostral do SAEB também sugere uma decisão nesse sentido, afinal, o que são amostradas são turmas de uma escola e, dentro dessa turma, aproximadamente metade dos alunos responde ao teste de Matemática e a outra metade, ao teste de Leitura. Ou seja, a inserção da variável *disciplina* no modelo proporciona um ganho, pois permite um número maior de observações por escola (ALBERNAZ; FRANCO; FERREIRA, 2002).

A tabela a seguir apresenta uma descrição mais detalhada das variáveis incluídas no modelo.

TABELA 15: Descrição das variáveis incluídas no modelo base

Variável	Descrição	Forma de Medida
NSE	Nível socioeconômico*	Escala contínua variando de 0 a 10.
ATRASO	Atraso escolar	Variável discreta: 0 – Regular 1 – 1 ano atrasado 2 – 2 anos atrasado 3 – 3 anos atrasado 4 – 4 anos atrasado 5 – 5 anos atrasado ou mais
SEXO	Sexo do aluno	Variável indicadora: 0 – Masculino 1 – Feminino
PARDO	Autodeclaração da cor “parda”	Variável indicadora: 0 – Não pardo 1 – Pardo
PRETO	Autodeclaração cor “preta”	Variável indicadora: 0 – Não preto 1 – Preto
REDE	Rede de ensino	Variável indicadora: 0 – Pública 1 – Particular

DISCIPLINA	Disciplina	Variável indicadora: 0 – Língua Portuguesa 1 – Matemática
ATITUDE	Atitude ²⁵	Escala contínua obtida via TRI
MNSE	Média do nível socioeconômico	Nível socioeconômico médio dos alunos de uma mesma escola
MATRASSO	Média do atraso escola	Atraso escolar médio dos alunos de uma mesma escola

* A variável NSE foi centralizada na média da escola, quando medida no nível do aluno, e na média geral, quando medida no nível da escola.

Algumas considerações sobre o programa computacional utilizado são necessárias. O software, normalmente utilizado para o trabalho com esse tipo de modelo de regressão multinível, é o HLM. Neste caso, porém, como iria se trabalhar com um número enorme de bases de dados, optou-se por um software que realizasse, ao mesmo tempo, as duas funções, preparação do banco de dados e capacidade para rodar os modelos hierárquicos. O software encontrado foi o SPSS 13.0, que é amplamente utilizado no tratamento com bancos de dados, com análises estatísticas e, nesta versão, possui também a função *Linear Mixed Models*. Raudenbush e Bryk (2002) especificam seu modelo, escrevendo, separadamente, uma equação para cada nível como está descrito, por exemplo, na equação (5), e o seu software (HLM) também trabalha com o modelo, sendo escrito em dois níveis. Já o SPSS, por utilizar a tecnologia *PROC MIXED*, exige que o modelo seja representado em apenas um nível, isto é, em apenas uma única equação. Tal decisão exigiu que todo modelo de regressão multinível implementado nesta pesquisa fosse representado em uma única equação, tal como se pode ver nos anexos a representação do modelo (5). Nos anexos também se encontra a *syntax* utilizada no SPSS para rodar o modelo base. No entanto, por questões didáticas, optou-se por continuar apresentando os modelos em dois níveis sempre que necessário. Singer (1998) apresenta uma discussão bastante detalhada sobre como utilizar *PROC MIXED* para estimar os parâmetros dos modelos multinível mais comuns.

²⁵ Para o cálculo da variável “Atitude”, foi utilizada a mesma metodologia do cálculo do nível socioeconômico, ou seja, a Teoria de Resposta ao Item (TRI). Como o objetivo aqui também era sintetizar as informações das 28 bases do SAEB, o construto acabou sendo composto por três variáveis, obtidas nos questionários contextuais. São elas: 1) Você gosta de estudar Matemática?; 2) Você gosta de estudar Língua Portuguesa?; e 3) Você faz lição de casa? Sendo que as alternativas de resposta para essas três questões foram, sempre, “sim” ou “não”.

2.4.3 Como analisar os resultados do modelo

No item 1.3, definiram-se os conceitos de qualidade e eqüidade no contexto deste trabalho. A qualidade está associada a um alto desempenho cognitivo, e a eqüidade associa-se à capacidade de proporcionar aumento no desempenho, independente das características de *background* do grupo de alunos. Registrou-se também que uma escola, nesse caso um sistema educacional, seria considerada tanto melhor quanto maiores fossem sua qualidade e sua eqüidade. Agora, é o momento de associar essas definições à maneira como elas serão obtidas matematicamente com a ajuda dos modelos de regressão hierárquica.

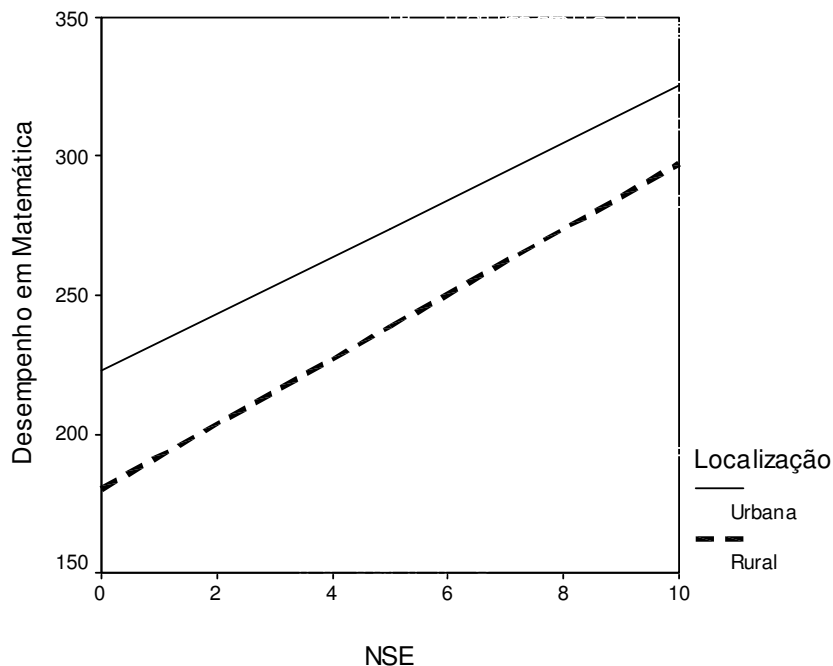
Como foi explicado anteriormente, os modelos multinível nada mais são que modelos de regressão linear. E como qualquer modelo de regressão linear, seu produto final é a equação de uma reta, que sintetiza um conjunto de pontos, ou seja, uma reta que exprime a relação existente entre duas ou mais variáveis. As retas são funções lineares, que podem ser representadas matematicamente por uma equação do tipo:

$$Y = a * X + b \quad (6)$$

Onde “Y” é a variável dependente e “X” é uma variável explicativa ou independente. O parâmetro “a” é inclinação da reta e informa sobre a variação ocorrida em “Y” para cada aumento de uma unidade ocorrido em “X”. O parâmetro “b” é o intercepto ou valor de y para o qual o x é igual a zero. No caso desta pesquisa, em que a variável dependente é a proficiência dos alunos e as variáveis independentes são os fatores escolhidos para compor o modelo, tem-se que a qualidade do sistema educacional fica definida pelo coeficiente do intercepto, ou seja, por “b”, e a eqüidade fica definida pelo inverso do valor da inclinação, ou seja, quanto menos inclinada for a reta maior é a eqüidade.

O Gráfico 19 foi construído hipoteticamente com o objetivo de explicitar esses conceitos. Nele estão representadas duas retas de regressão linear, que sintetizam a relação existente entre o desempenho dos alunos em Matemática e o seu NSE, para escolas localizadas na zona urbana e na zona rural.

GRÁFICO 19: Gráfico hipotético relacionando desempenho em Matemática e NSE



Como se pode constatar, as escolas da zona urbana, nessa situação hipotética, são consideravelmente melhores que as da zona rural. Primeiro, porque seu intercepto é maior, ou seja, sua qualidade é maior. Observe-se que a reta das escolas urbanas corta o eixo y em torno dos 225, enquanto a das escolas rurais corta em torno dos 180 pontos. E também porque sua inclinação é menor, ou seja, sua equidade é maior. Observe-se que a separação entre as duas vai diminuindo com o aumento do NSE, indicativo de menor inclinação da reta de cima.

Apesar de muito elucidativa, a análise gráfica, em diversos casos, torna-se ineficiente, pois a diferença na inclinação não é marcante. Soma-se a isso o fato de que, neste trabalho, está sendo analisado um número muito grande de bases de dados. Esses fatores justificam o porquê da preferência que, às vezes, será dada à análise de tais resultados na forma de tabelas, registrando-se os parâmetros ajustados do modelo para o intercepto e para a inclinação, em vez de se construir gráficos para cada uma das questões de análise.

2.4.4 Os primeiros resultados do modelo

Apesar de refletir o conhecimento real do aluno, a proficiência bruta não deve ser utilizada, isoladamente, como indicador de qualidade dos sistemas de ensino. “Para uma comparação mais justa e pedagogicamente mais útil, é preciso retirar estatisticamente das proficiências o efeito do nível socioeconômico tanto individual como da escola freqüentada” (SOARES *et al*, 2004, p. 14), além do efeito de outros fatores, como, por exemplo, o atraso escolar. Os modelos de regressão multinível são a solução encontrada para este problema, e os resultados apresentados nesta seção refletem bem o efeito de cada um dos fatores escolhidos para compor o modelo sobre a proficiência dos alunos²⁶.

As tabelas a seguir mostram alguns dos valores para os coeficientes resultantes do ajuste do modelo base (5). Os demais resultados serão apresentados, mais detalhadamente, em capítulos subseqüentes.

TABELA 16 – Coeficiente γ_0 ajustado para o modelo base – Intercepto

Ano	4ª EF	8ª EF	3ª EM
1995	-	265,36	300,79
1997	181,23	258,06	299,79
1999	170,56	246,18	282,22
2001	169,22	250,04	282,37
2003	174,42	243,3	284,35

Essa é uma tabela que reflete a média da proficiência em Língua Portuguesa, para alunos (sexo masculino apenas) da rede pública, que se autodeclararam brancos, que estejam regulares no que se refere à idade-série, cujo nível socioeconômico seja igual ao da média da população e que estudem em escolas com média de atraso igual a zero e com nível socioeconômico médio igual à média da população. Para analisar

²⁶ Os resultados completos do modelo base, obtidos como saída do software SPSS, encontram-se nos anexos.

resultados de outros grupos, é necessário observar o valor dos coeficientes estimados para os mesmos.

As duas próximas tabelas corroboram a discussão feita no item 2.3 sobre a relação entre NSE e desempenho cognitivo. De forma especial, percebe-se que o efeito do NSE é sempre positivo, ou seja, contribui para aumentar a proficiência do aluno. Esta relação é nítida no nível do aluno, mas muito mais forte quando se olha a influência da característica socioeconômica da escola que ele frequenta, “talvez porque escolas com alunos de maior nível socioeconômico formem ambientes mais adequados ao aprendizado” (SOARES *et al*, 2004, p. 23).

TABELA 17 – Coeficiente γ_{10} ajustado para o modelo base –
Variável associada: NSE do aluno

Ano	4ª EF	8ª EF	3ª EM
1995	-	2,80**	1,40*
1997	1,47**	1,22**	1,26*
1999	0,59*	0,07	0,09
2001	1,61**	0,68**	0,08
2003	1,85**	2,18**	1,08**

(*) = $p < 0,05$ e (**) = $p < 0,001$

TABELA 18 – Coeficiente γ_{01} ajustado para o modelo base –
Variável associada: NSE médio da escola

Ano	4ª EF	8ª EF	3ª EM
1995	-	17,80**	20,38**
1997	15,29**	13,33**	12,80**
1999	8,53**	9,13**	10,97**
2001	14,54**	11,96**	15,07**
2003	14,16**	12,93**	14,95**

(*) = $p < 0,05$ e (**) = $p < 0,001$

Os coeficientes ajustados para a variável *atraso*, sintetizados nas duas tabelas a seguir, mostram o efeito negativo de cada ano de defasagem sobre o desempenho cognitivo do aluno. Ao contrário do que pode indicar o senso comum, a retenção de alunos, além dos seus prejuízos financeiros, não implica em aumento no desempenho. Resultado amplamente divulgado por Klein (2006), no qual se defende a busca da

qualidade de sistemas educacionais, não só a partir do acesso, mas também se levando em consideração o fluxo escolar e o desempenho dos alunos.

TABELA 19 – Coeficiente γ_{50} ajustado para o modelo base –
Variável associada: Atraso escolar do aluno

Ano	4ª EF	8ª EF	3ª EM
1995	-	-6,07**	-4,43**
1997	-2,65**	-5,12**	-5,19**
1999	-3,37**	-5,63**	-5,49**
2001	-3,20**	-6,74**	-6,88**
2003	-5,14**	-8,01**	-8,18**

(*) = $p < 0,05$ e (**) = $p < 0,001$

TABELA 20 – Coeficiente γ_{03} ajustado para o modelo base –
Variável associada: Atraso escolar médio

Ano	4ª EF	8ª EF	3ª EM
1995	-	-1,76*	-2,77*
1997	0,57	-0,70	-4,97**
1999	0,43	-1,62**	-4,42**
2001	0,57	-1,24**	-2,94**
2003	1,65**	-0,62	-4,05**

(*) = $p < 0,05$ e (**) = $p < 0,001$

Os resultados da Tabela 20 chamam a atenção para um fato que, talvez, precise ser mais bem estudado. Na 8ª série do ensino fundamental e na 3ª série do ensino médio, cada ano de atraso escolar médio contribui negativamente para o desempenho final do aluno. No entanto, na 4ª série do ensino fundamental, esse fato, estranhamente, não se repete. Ainda que os valores sejam muito próximos de zero e, na maioria das vezes, não significativos, eles são positivos. É como se o efeito da repetência agisse ao contrário nas séries iniciais do fundamental. O fato é que essa questão demanda outras investigações e permanecerá em aberto.

Outro fator que merece mais atenção e, principalmente, mais estudos é a questão da atitude do aluno em relação ao ensino. Como explicado no item 2.4.2, o construto “atitude” foi medido a partir de três questões do questionário do SAEB, as quais perguntavam se o aluno gosta de estudar Matemática, se ele gosta de estudar Língua Portuguesa e se ele faz a lição de casa. A tabela a seguir mostra que esse efeito é sempre significativo e, nos cinco ciclos, apresenta um efeito médio, que

acrescenta mais de 16 pontos à nota do aluno, para cada aumento de uma unidade na medida da atitude. Especificamente na 3ª série do ensino médio, esse efeito médio chega a quase 19 pontos, ou seja, é necessário preocupar-se com a motivação dos alunos frente aos estudos e com o fato de eles fazerem ou não o “dever de casa”.

TABELA 21 – Coeficiente γ_0 ajustado para o modelo base –
Variável associada: Atitude

Ano	4ª EF	8ª EF	3ª EM
1995	-	12,59**	10,96**
1997	21,76**	14,21**	13,64**
1999	13,32**	18,74**	20,20**
2001	23,38**	15,96**	19,96**
2003	6,45**	16,29**	19,73**

(*) = $p < 0,05$ e (**) = $p < 0,001$

Capítulo 3 – O efeito da escola básica brasileira

3.1 Introdução ao problema

Hoje, a sociedade brasileira espera que a escola de educação básica garanta a seus alunos o aprendizado das competências necessárias para uma inserção crítica e produtiva na sociedade. Tais competências, que são de natureza cognitiva, mas também cultural e ética, incluem a formação de hábitos, o convívio com as diferenças e a solução de conflitos pacificamente.

Neste texto, o conceito de “efeito da escola” terá, entretanto, uma definição mais restrita, referindo-se somente à influência da escola no desempenho cognitivo de seus alunos. No Brasil, a medida desse efeito só se tornou possível com a consolidação do SAEB, como afirmam Brooke e Soares (2008):

A consolidação do SAEB – Sistema de Avaliação da Educação Básica a partir de 1995 e a conseqüente publicação regular de dados comparáveis sobre o desempenho dos alunos da educação básica provocaram uma mudança profunda na forma de analisar os sistemas de ensino no Brasil. Após o SAEB, os sistemas educacionais puderam ser analisados não só em relação a sua capacidade de atendimento às crianças em idade escolar, mas também em relação ao aprendizado de seus alunos. Pela primeira vez, tornou-se possível avaliar aquelas características das escolas que mais se associavam à aprendizagem. Mediante o cruzamento de informações sobre as condições de funcionamento das escolas, as características socioeconômicas dos alunos e o nível de aprendizagem alcançado, a pesquisa brasileira começou a quantificar o efeito-escola e a desvendar as especificidades da escola eficaz (BROOKE; SOARES, 2008, p. 9).

Ou seja, o conceito de “efeito da escola” é utilizado para medir a capacidade de as escolas, por meio de seu projeto pedagógico e de suas políticas internas, influenciarem o desempenho cognitivo de seus alunos.

Os estudos para a caracterização do efeito das escolas e dos fatores e estruturas promotoras de desempenho surgiram como contraponto aos resultados pessimistas obtidos pelo Relatório Coleman (COLEMAN *et al.*, 1966). Hoje se sabe que a escola pode fazer diferença no aprendizado do aluno, ainda que sozinha não

possa nem ser responsabilizada por todas as dificuldades dos alunos, nem por todos os seus sucessos. Sobre isso, ver, por exemplo, a ampla revisão da literatura, elaborada por Teddlie e Reynolds (2000), que aborda desde aspectos históricos e empíricos da pesquisa sobre o efeito das escolas até a metodologia e características científicas dessa linha de pesquisa.

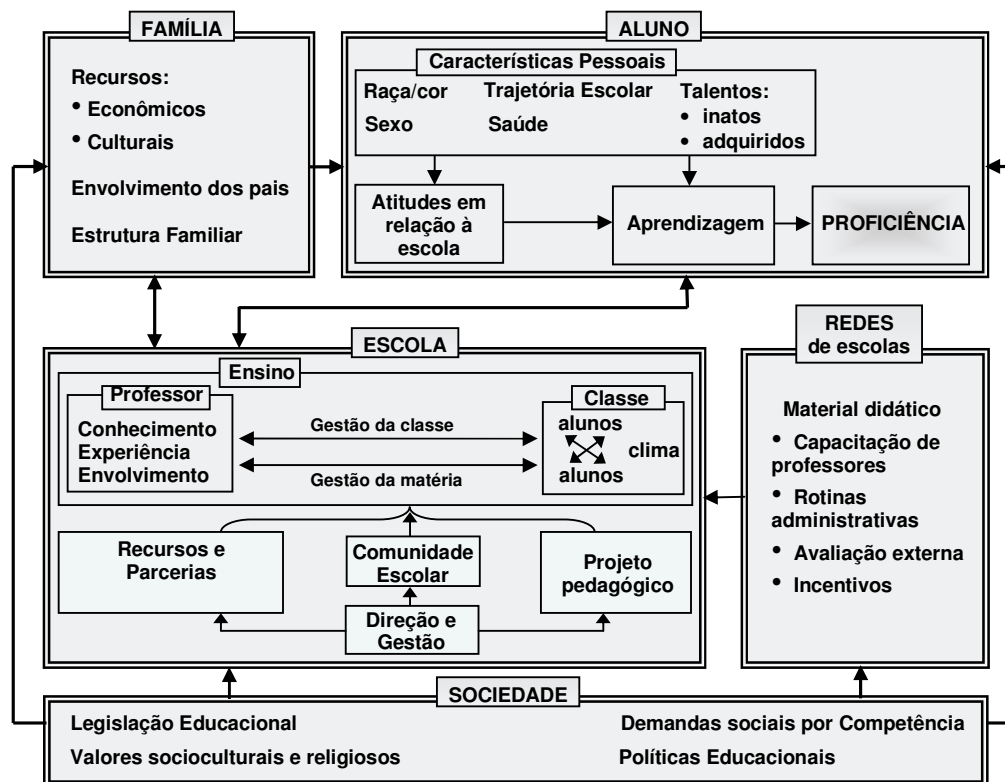
Essa literatura tem concluído que:

embora parte importante da explicação dos baixos níveis de desempenho dos alunos esteja em fatores extra-escolares, há uma enorme variação entre resultados de escolas de um mesmo sistema que atendem alunos muito similares em termos socioeconômicos. Ou seja, a unidade escolar freqüentada pelo aluno pode fazer diferença significativa na sua vida escolar (BROOKE; SOARES, 2008, p. 9).

Soares (2008) propõe um modelo conceitual, que relaciona cinco estruturas associadas ao desempenho cognitivo dos alunos: o aluno, sua família, a escola e a rede, ou sistema a que está associado, e, finalmente, a sociedade em geral. O autor registra que hoje está estabelecido, além da dúvida razoável, que tanto os fatores intra-escolares quanto os extra-escolares estão ligados ao desempenho dos alunos. Mesmo restringindo-se aos fatores ligados à escola, o

modelo mostra que são tantos os fatores escolares associados ao desempenho dos alunos que nenhum deles é capaz de garantir, isoladamente, bons resultados escolares. A ênfase dada a fatores específicos em alguns momentos históricos deve ser atribuída mais à fé dos que os advogam, e não a evidências científicas (SOARES, 2004, p. 5).

FIGURA 1: Modelo conceitual explicativo da proficiência



FONTE: Soares, 2008.

Como dito, o objetivo deste texto é caracterizar, por meio de medidas, o efeito das escolas, o que exige a explicitação dos modelos estatísticos e dos dados. Isso será feito na próxima seção.

3.2 O Efeito da escola

Hoje, na literatura, há duas definições do conceito de efeito da escola. A primeira refere-se a uma medida específica para cada escola, que registra o efeito daquela unidade escolar sobre o aprendizado de seus alunos, consiste na parcela da proficiência do aluno que pode ser atribuída às práticas da escola. A segunda é uma medida do efeito global das escolas, de um determinado grupo de referência, com base numa partição da variância dos resultados dos alunos, isto é, se esse efeito é grande, há diferenças entre estudar em uma ou outra escola (BROOKE; SOARES, 2008). Essas duas definições estão presentes no modelo base proposto nesta tese. A

primeira trata-se do termo u_{0j} e a segunda é captada no quociente entre a variância do u_{0j} e a soma das variâncias do u_{0j} e r_{ij} .

Inicialmente, colocaremos foco na medida do efeito da escola individual, que foi introduzida por Willms (1992). Essa merece destaque, entre outras razões, por apresentar, criticamente, a forma de cálculo do efeito da escola por meio dos modelos estatísticos escolhidos para análise. Willms (1992) deixa claro que a interpretação dos efeitos depende da compreensão dos modelos subjacentes, ou seja, depende das escolhas feitas sobre quais variáveis compõem o modelo. Diferentes agentes podem se interessar por formas de controle diferenciadas. Neste trabalho, como já foi citado, optou-se por um controle mais rigoroso do efeito da escola. Esta opção metodológica se aproxima do 'efeito tipo B,' proposto por Willms e Raudenbush (1989). O estudo deste efeito é considerado mais adequado para a formulação e análise de políticas públicas educacionais. Ao contrário do 'efeito tipo A', mais interessante aos pais dos alunos, interessados no efeito da escola, sem distinção da origem deste efeito. Neste caso, os modelos de análise não incluem variáveis de controle das características internas da escola.

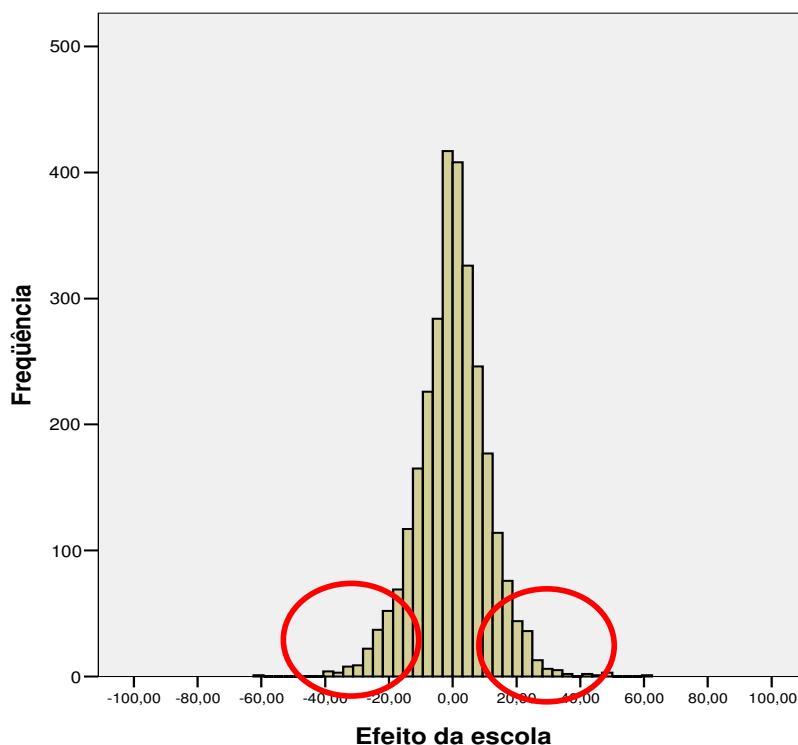
A medida individual do efeito escola, u_{0j} , é a diferença entre a performance média de uma escola específica e um nível padrão, controlando-se a performance por características de entrada (perfil do grupo de alunos da escola, características do sistema de ensino, infra-estrutura etc.), portanto fica claro que a medida depende das escolas que compõem a amostra. Obtendo-se um valor alto e positivo para uma determinada escola, conclui-se que essa escola possui práticas internas que levam seus alunos a resultados melhores do que era esperado, comparando-se com escolas com características de entrada semelhantes e incluídas na amostra. Obtendo-se valores negativos para esse efeito da escola, não significa que os alunos dessa escola específica estão 'desaprendendo'. A medida apenas indica um desempenho abaixo da média das escolas incluídas no estudo.

O trabalho de Soares, Alves e Oliveira (2001) ilustra bem esse tipo de medida, pois representa um esforço para se medir o efeito das escolas de ensino médio do estado de Minas Gerais, utilizando-se as bases de dados então disponíveis, no caso, as bases do vestibular da UFMG, nos anos de 1998, 1999 e 2000. Os autores ajustaram três modelos diferentes com esses dados, e, no contexto deste artigo, o principal resultado encontrado foi justamente que o efeito das escolas muda, isto é, depende do modelo escolhido. O primeiro modelo gera uma medida de efeito da escola, que não deve ser utilizada, pois não leva em consideração as características

de background do alunado. O segundo modelo já considera tais características, sendo, portanto, de maior interesse aos pais dos alunos em processo de escolha da escola ideal para seu filho, uma vez que ele identifica os estabelecimentos onde os alunos têm maior desempenho por conta do efeito dessas escolas. E, por fim, o terceiro modelo é o mais rigoroso dos três, uma vez que retira do efeito da escola a parte do desempenho que é atribuída às suas características contextuais. Esse último modelo é similar ao que foi adotado nesta pesquisa e interessa mais aos gestores educacionais, “pois discrimina escolas que estão apresentando melhores resultados pelos seus próprios méritos” (SOARES; ALVES; OLIVEIRA, 2001, p. 89).

Ajustando-se o modelo base (5) aos dados do SAEB, observa-se que o comportamento da medida do efeito da escola, u_{oj} , é muito semelhante em todos os cinco ciclos estudados. O histograma a seguir ilustra, para a base de Matemática 2003, 8ª série do ensino fundamental, que a maioria das instituições apresenta um efeito da escola próximo de zero. Isto significa que, na prática, as escolas têm efeito muito próximo entre si.

GRÁFICO 20: Efeito da escola em Matemática, 8ª série EF, para a base 2003



FONTE: Elaboração própria com dados do SAEB 2003 – INEP/MEC.

No entanto, essa constatação geral não pode impedir que se perceba que também existe um número considerável de escolas que estão impactando, positiva ou negativamente, o desempenho de seus alunos em mais de 20 pontos de proficiência, valor equivalente a aproximadamente um ano de escolaridade. A tabela a seguir explicita tal fato com maior clareza, uma vez que sintetiza o percentual de escolas, em cada base, que apresenta o u_{0j} maior ou igual a 20 pontos, tanto positiva quanto negativamente. Entretanto, vale lembrar que esse é um efeito relativo e, portanto, se uma escola for retirada da amostra o valor muda.

TABELA 22: Percentual de escolas, em cada base SAEB, com u_{0j} (efeito da escola) inferior ou superior a 20 pontos

Ano	$u_{0j} \leq -20,0$			$u_{0j} \geq 20,0$		
	4ª EF	8ª EF	3ª EM	4ª EF	8ª EF	3ª EM
1995	-	7,3	8,5	-	5,7	8,5
1997	3,0	3,6	7,5	4,6	3,9	8,2
1999	2,5	3,0	7,3	3,6	3,6	8,1
2001	3,1	3,3	6,1	3,1	2,7	6,9
2003	3,1	3,7	7,5	3,4	3,2	7,8

FONTE: Elaboração própria com dados do SAEB – INEP/MEC

Em especial, é digna de nota a existência de escolas com efeito positivo e alto, as quais precisam ser mais estudadas, por metodologias qualitativas, para se compreender melhor os fatores que impactam a aprendizagem na realidade brasileira. Nessa linha, vale o registro de uma iniciativa marcante: o “Aprova Brasil – O direito de aprender” (BRASIL/MEC/UNICEF, 2006), que sintetiza boas práticas existentes em 33 escolas públicas do país, que tiveram um alto valor para o efeito da escola²⁷, de acordo com os resultados da Prova Brasil²⁸.

²⁷ Tendo em vista o contexto deste trabalho, seria desejável citar a metodologia utilizada no “Aprova Brasil – O direito de aprender”, para se calcular o efeito da escola. No entanto, nos documentos, aos quais se teve acesso, em especial o Brasil/MEC (2006), a única menção a esse índice é de que se trata de um “indicador do impacto que a escola tem na vida e no aprendizado da criança. Pesquisado e testado em estudos em diversas realidades, o Índice é determinado a partir de critérios científicos. As escolas visitadas, pesquisadas e analisadas no estudo ‘Aprova Brasil - o direito de aprender’ não são exatamente as melhores escolas, mas aquelas com o mais alto ‘efeito escola’” (BRASIL/MEC, 2006, p. 10).

²⁸ Cada uma dessas 33 escolas foi visitada e centenas de pessoas foram ouvidas, entre elas, diretores, professores, funcionários, pais e alunos. As crianças que freqüentam essas escolas são filhas de famílias de baixa renda, vivem em municípios pobres ou de difícil acesso, ou em comunidades especialmente vulneráveis. As condições desses meninos e meninas poderiam conspirar para que eles não

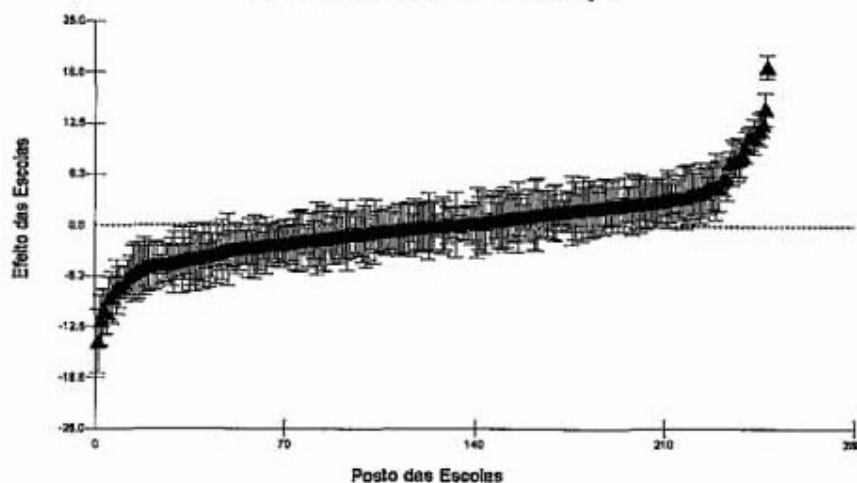
No entanto, a imprecisão da medida do efeito da escola é muito grande, tanto que torna inútil qualquer tentativa de hierarquização de escolas segundo o seu efeito. “As estimativas desses trabalhos contêm tanta incerteza estatística que é impossível fazer comparações válidas entre a maioria das escolas” (GOLDSTEIN, 2001, p. 91).

Infelizmente, essa postura é extremamente “apreciada” pela imprensa quando há qualquer divulgação de resultados de uma avaliação educacional, provocando, na maioria das vezes, conclusões e inferências sem fundamentos. As listas das escolas, com o resultado do Exame Nacional do Ensino Médio (ENEM), é um exemplo de como “não” se deveria divulgar os resultados. A hierarquização das escolas causa um tumulto grande, sem, no entanto, contribuir para uma completa compreensão desses resultados e da medida que foi feita. E mais, as escolas mudam freqüentemente sua posição no tal *ranking*, fruto muito mais da incerteza na medida do que dos esforços pedagógicos da comunidade escolar.

O próximo gráfico exemplifica essa inviabilidade já que hierarquiza as escolas segundo o valor do seu efeito, ao mesmo tempo em que associa o intervalo de confiança para cada medida. O gráfico foi extraído do trabalho de Soares, Alves e Oliveira (2001) e tem como referencial teórico as discussões propostas inicialmente por Goldstein e Thomas (1996) e detalhadas, em seguida, por Goldstein e Spiegelhalter (1996). Observa-se que não é possível distinguir a diferença entre duas escolas consecutivas, no máximo, podem ser identificados três grandes grupos: um primeiro, mais à esquerda, com as escolas com efeitos mais negativos; um segundo, o mais populoso dos três, ao centro, contendo as escolas, cujo efeito é nulo dentro da precisão da medida; e um terceiro, mais à direita, com escolas, cujo efeito é positivo. Ou seja, nas palavras de um desses autores, ao comentar a incerteza dessas medidas, “podemos descrever relações estruturais num alto nível de complexidade, mas não podemos medir precisamente a contribuição de cada escola individualmente” (GOLDSTEIN, 2001, p. 91).

tivessem bons rendimentos em seus estudos, mas as relações humanas, a criatividade, a participação, o respeito às condições e saberes de cada um, as práticas pedagógicas e a interação com a comunidade fazem a diferença. As crianças e os adolescentes aprendem e, assim, a escola demonstra toda sua força de transformar para melhor a vida desses meninos e meninas (Brasil/MEC, 2006, p. 4).

GRÁFICO 21: Intervalo de confiança para os efeitos das escolas estimado pelo modelo 3 segundo os dados do vestibular de 2000



FONTE: Soares; Alves; Oliveira, 2001, p. 90

3.3 A heterogeneidade das escolas brasileiras

A segunda definição do conceito efeito da escola é o “efeito da escola como medida da proporção da variância total, que pode ser atribuída às escolas em relação à variância total dos escores de seus alunos em testes individuais” (TEDDLIE; REYNOLDS, 2000, p. 68), ou seja, essa é uma medida de grupo de escolas e não individual como o primeiro tipo. Dessa forma, não se trata de atribuir uma parcela do desempenho do aluno à escola na qual ele estuda, mas atribuir a responsabilidade de uma parcela da variação dos desempenhos de todos os alunos às características internas de todas as escolas do grupo considerado.

Antes de usar essa definição de efeito da escola, é adequado questionar o nome usualmente dado a essa medida. A palavra efeito sugere que quanto maior o valor maior será o seu impacto, não é, entretanto, o caso dessa medida. Um aumento no percentual (ou na proporção da variância) indica, apenas, que há uma maior distinção no grupo de escolas. Isto é, há diferenças entre as escolas, cujas explicações residem nas características internas das escolas, seu alunado, seu projeto pedagógico, seus professores etc. Daí a opção pelo termo “medida da heterogeneidade” das escolas em detrimento do termo “efeito da escola”. Na verdade, um aumento desse percentual não implica em um aumento do efeito provocado por cada escola, que é uma medida individual. Ele corresponde apenas a uma maior

variabilidade dentro do grupo, que agora apresentaria escolas com alto efeito e outras com baixo, ou até mesmo com efeito negativo, por isso propõe-se uma distinção, na qual apenas o primeiro tipo de efeito continue sendo denominado efeito da escola, já que se trata de uma medida individual; enquanto o segundo tipo de efeito passaria a ser denominado medida da heterogeneidade de escolas, por se tratar de uma medida de grupo.

O objetivo específico da introdução desse conceito seria medir quão diferentes são as escolas brasileiras em relação ao desempenho em Matemática e Língua Portuguesa (Leitura) de seus alunos, utilizando-se os dados de cinco ciclos do Sistema Nacional de Avaliação da Educação Básica (SAEB). Ou seja, não se trata de considerar o efeito de cada escola em particular, mas o coeficiente de correlação intra-escolar como estimador da heterogeneidade das escolas, que é dado por:

$$\rho = \frac{\sigma_{u_0}^2}{\sigma_r^2 + \sigma_{u_0}^2} \quad (7)$$

onde:

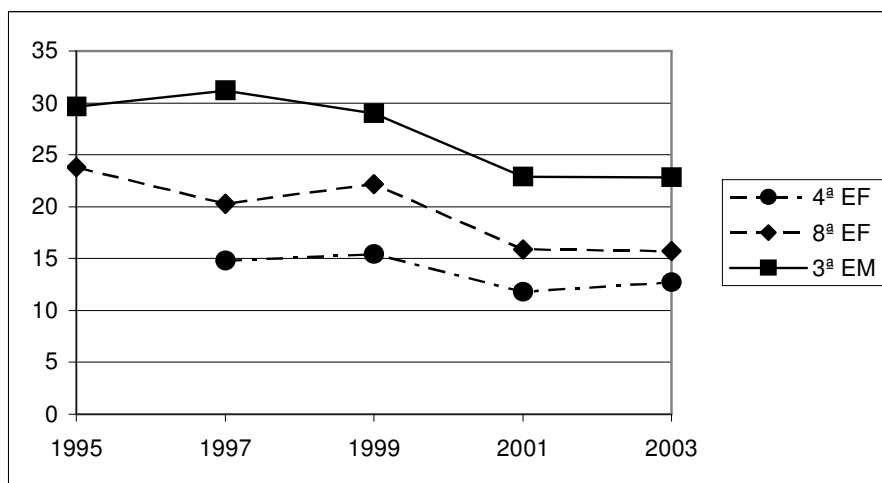
- ρ é a medida da heterogeneidade das escolas. Um coeficiente que varia entre 0 e 1, sendo que, quanto mais próximo de 1 mais a variabilidade na proficiência dos alunos deve-se à diferença entre as escolas; e
- os coeficientes σ foram definidos na equação (5), no detalhamento do Modelo Base escolhido.

Teddlie e Reynolds (2000) fazem uma revisão bibliográfica considerando a magnitude da heterogeneidade das escolas em estudos internacionais. Esses mostram valores entre 11 e 12% para as escolas da Holanda, de 12% para a Alemanha e de 10% para os EUA. Utilizando-se de dados do SAEB, este trabalho obteve uma medida da heterogeneidade para as escolas brasileiras de 19%, quando se leva em consideração apenas o desempenho em Matemática, e 15%, quando se considera como variável dependente a proficiência em Leitura. Como registrado anteriormente, as escolas brasileiras também são muito semelhantes. Registre-se, no entanto, que os valores dos países citados referem-se a um cálculo sem o controle prévio pelo NSE, enquanto os resultados para o Brasil foram obtidos após esse

controle. Já a diferença na medida para as duas disciplinas, provavelmente, reside no fato de que a Matemática sofre mais a influência dos fatores escolares, enquanto o estudo da língua materna é mais sensível às variáveis familiares (BARBOSA, 2005).

Os gráficos a seguir permitem que seja analisada não só a estabilidade da medida da heterogeneidade de 1995 a 2003, como também a influência de fatores como a dependência administrativa e a série dos alunos.

GRÁFICO 22: Evolução da Heterogeneidade das escolas, por série – Rede Particular, Matemática



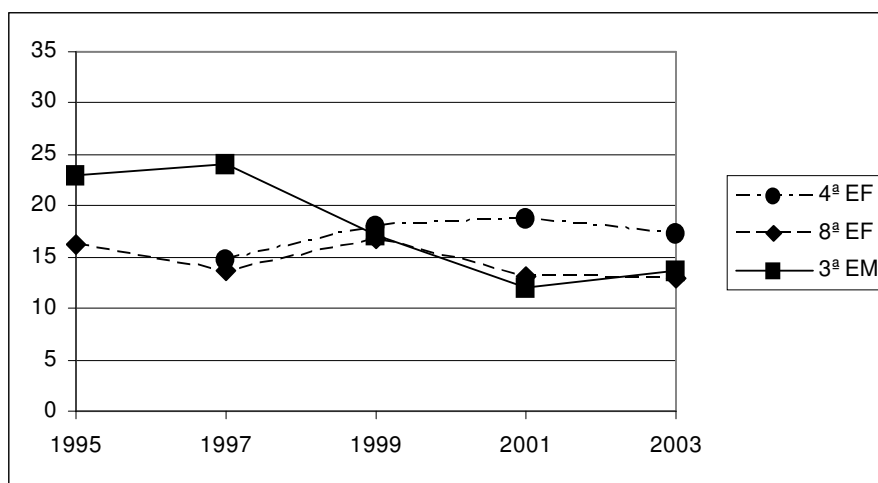
FONTE: Elaboração própria com dados do SAEB – INEP/MEC.

O Gráfico 22, considerando-se todas as escolas particulares da amostra, mostra que a semelhança entre elas é maior no ensino fundamental do que no ensino médio. Há um aumento na heterogeneidade entre escolas com o aumento das séries de estudo, em todos os ciclos considerados. Este gráfico sintetiza os dados de Matemática, mas o mesmo fenômeno é observado em relação aos dados de Leitura. Uma possível explicação para esse maior valor no ensino médio é a forte competição para o vestibular das instituições públicas de ensino superior. Isso faz com que algumas escolas foquem seus trabalhos no desenvolvimento cognitivo de seus alunos, enquanto outras estabelecem prioridades distintas em seus projetos pedagógicos. Uma segunda hipótese, sabendo-se que as escolas diferem bastante entre si, pode estar relacionada com o fato de que as competências desenvolvidas nas séries iniciais, como a alfabetização, por exemplo, sofrem grande influência familiar, enquanto as competências previstas para o ensino médio são de cunho muito mais escolar. Portanto, o efeito da escola, seja ele positivo ou negativo, estará muito mais em evidência nas últimas séries do que nas primeiras.

Apesar de se considerar uma tendência de queda na heterogeneidade das escolas em relação aos alunos da 3ª série do ensino médio, não se pode negar que, para a rede particular do Brasil, esses valores ainda são muito elevados.

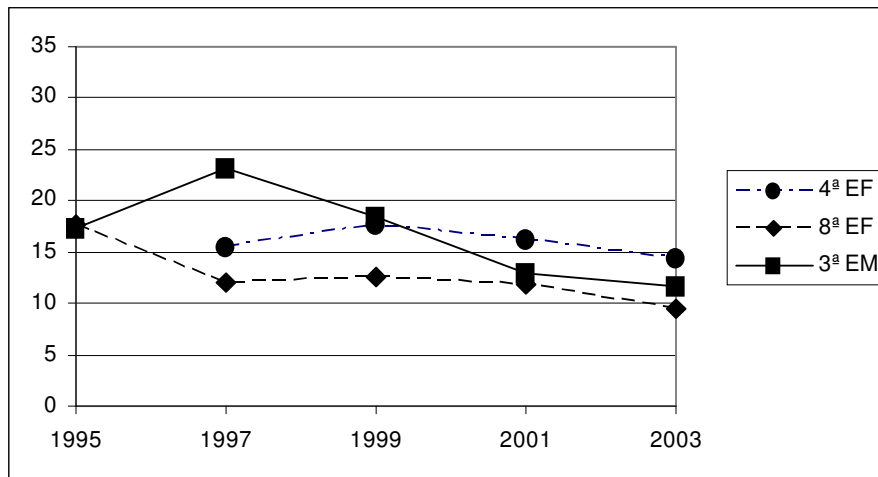
Comparando-se o Gráfico 22 com os Gráficos 23 e 24, a seguir, constata-se que, na rede pública, a supremacia da heterogeneidade das escolas para a 3ª série do ensino médio permanece nos ciclos do SAEB dos anos 90. Mas os ciclos de 2001 e 2003 mostram que há maior variação entre as escolas das redes municipal e estadual, que atendem à 4ª série do ensino fundamental.

GRÁFICO 23: Evolução da Heterogeneidade das escolas, por série – Rede Municipal, Matemática



FONTE: Elaboração própria com dados do SAEB – INEP/MEC.

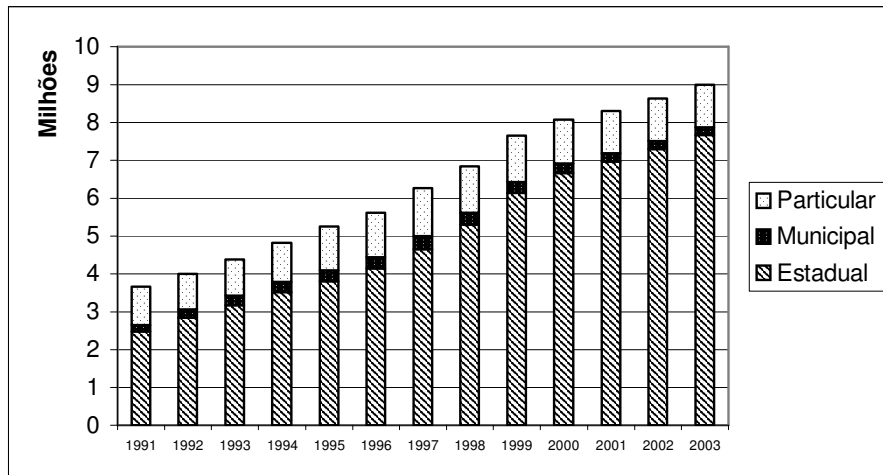
GRÁFICO 24: Evolução da Heterogeneidade das escolas, por série – Rede Estadual, Matemática



FONTE: Elaboração própria com dados do SAEB – INEP/MEC.

Essa inversão de valores na última década pode ser creditada, como mostram os Gráficos 23 e 24, muito mais a uma queda na heterogeneidade das escolas de 3ª série do ensino médio do que a um aumento na heterogeneidade das escolas de 4ª série do ensino fundamental. Essa queda se concentrou, principalmente, entre os anos de 1997 e 2001, porém foi muito mais acentuada na rede pública do que na privada. Isso pode ser explicado pelo grande aumento de matrículas no ensino médio da rede pública, fato não observado na rede particular, que permaneceu com um número de matrículas, praticamente, inalterado. Quanto maior o número de alunos, maior a homogeneidade entre as escolas. Em 1995, segundo os dados do Censo Escolar, somando-se os alunos matriculados no ensino médio na rede municipal e na rede estadual, havia um quantitativo superior a 4.400.000 matrículas. Em 2001, esse número estava 71% maior, superando 7.500.000 matrículas. Durante o mesmo período, as matrículas na rede particular decresceram de pouco mais de 1.170.000 para um pouco mais de 1.120.000 alunos matriculados no ensino médio. Esses dados podem ser mais bem visualizados no gráfico a seguir.

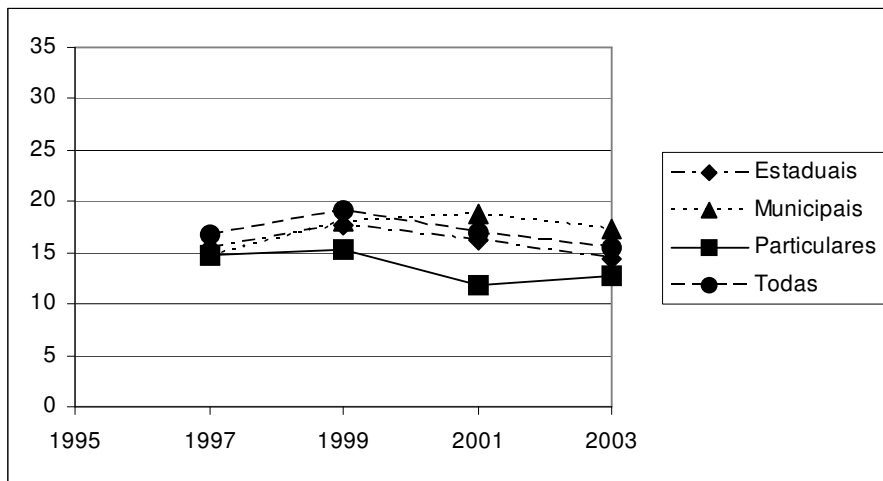
GRÁFICO 25: Número de matrículas no ensino médio, por ano e por dependência administrativa



FONTE: Elaboração própria com os dados do Censo Escolar – INEP/MEC.

Os gráficos 26, 27 e 28 apresentam a evolução da heterogeneidade das escolas para cada um dos níveis de ensino por dependência administrativa. Eles mostram que existem diferenças entre as redes, porém essas diferenças são mais acentuadas quando se olha a 3ª série do ensino médio do que a 4ª série do ensino fundamental. Esse fato corrobora os resultados mostrados na Tabela 22, em que se percebe um aumento do percentual de escolas com u_{0j} significativo, com o aumento da escolaridade.

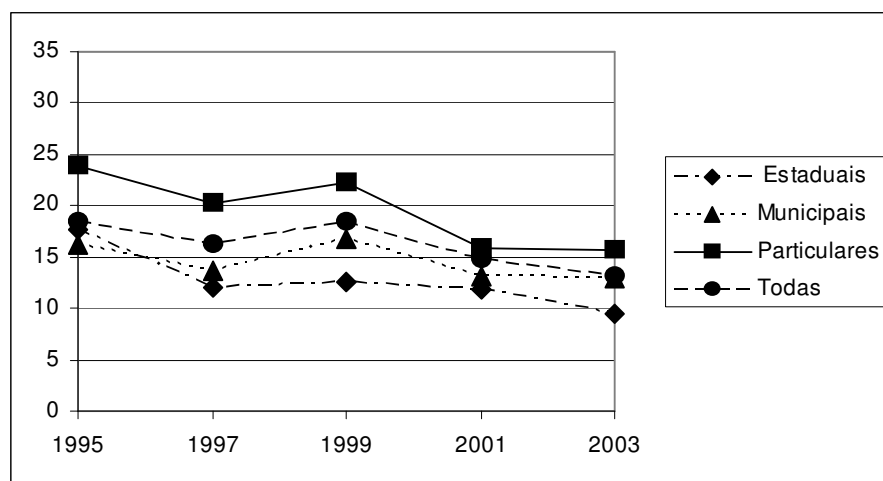
GRÁFICO 26: Evolução da Heterogeneidade das escolas, por rede – 4ª série EF, Matemática



FONTE: Elaboração própria com dados do SAEB – INEP/MEC.

O gráfico anterior confirma que, atualmente, em relação à 4ª série, as escolas da rede particular brasileira são mais homogêneas, e a rede municipal é a mais heterogênea de todas. Isso, certamente, é fruto dos investimentos na educação feitos por algumas prefeituras nos últimos anos, principalmente em algumas capitais, em detrimento de outras prefeituras, que mantêm a duras penas os níveis de ensino pelos quais são responsáveis.

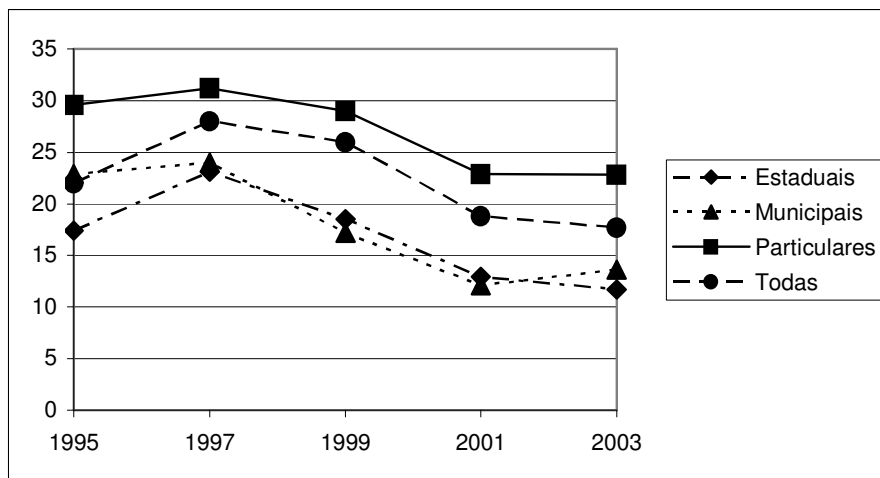
GRÁFICO 27: Evolução da Heterogeneidade das escolas, por rede – 8ª série EF, Matemática



FONTE: Elaboração própria com dados do SAEB – INEP/MEC.

Quando o foco está na 8ª série do ensino fundamental, esse cenário começa a mudar, e a rede particular começa a agregar escolas mais distintas quanto ao impacto produzido por elas sobre o desempenho cognitivo de seus alunos. Observa-se também que os valores da heterogeneidade vêm sofrendo constante queda, e os maiores percentuais atingidos em 2003 assemelham-se aos menores de 1995.

GRÁFICO 28: Evolução da Heterogeneidade das escolas, por rede – 3ª série EM, Matemática



FONTE: Elaboração própria com dados do SAEB – INEP/MEC.

Na 3ª série do ensino médio, as diferenças entre o público e o privado são as maiores dentre as observadas. Nesse nível de ensino, a rede estadual e a municipal assumem valores parecidos para a heterogeneidade. O destaque fica por conta da rede particular, que atingiu 30% em 1995, e, em 2003, ainda ostenta quase 25%. Suspeita-se que a rede particular seja mais heterogênea, não só por atender a um alunado mais diversificado, mas também por não estar sujeita a padrões pré-estabelecidos pelas Secretarias de Educação, como é o caso da rede pública.

3.4 Conclusões: a escola pode fazer a diferença

A relevância das discussões sobre o efeito da escola já está mais do que consolidada. No entanto, é preciso lembrar que o cálculo desse efeito é apenas o primeiro passo para a busca de fatores que expliquem os resultados alcançados, ou seja, há um caminho a ser percorrido para se conectar a eficácia e o melhoramento escolar (HOPKINS; STRINGFIELD, 2000 *apud* BROOKE; SOARES, 2008).

Os impactos desses resultados sobre as políticas públicas educacionais e sobre as futuras pesquisas do país, nessa linha de efeito da escola, relacionam-se com o fato de que somos mais parecidos com a situação internacional do que é usual admitir. Nossa medida de heterogeneidade apresenta uma média de 15%, quando a disciplina é a Língua Portuguesa, aproximando-se dos 10% que caracterizam os países desenvolvidos. Por outro lado, ainda que existam semelhanças, há escolas de

destaque. Tais expoentes também podem funcionar como referencial para a formulação de políticas públicas, visando à melhoria da educação, a partir da melhoria da escola.

No entanto, a variação existente, quando são analisados os subconjuntos de escolas, dentro do sistema educacional como um todo, é muito grande. Para os subconjuntos analisados, tem-se desde 9,5% até 31,2%. Isso é um indicador de que existem diferenças significativas entre redes ou entre séries, em relação à heterogeneidade das escolas, portanto a média dessa medida não caracteriza bem a diversidade do sistema educacional brasileiro.

Um registro fundamental para esse tipo de análise quantitativa da heterogeneidade das escolas, que, às vezes, é negligenciado em trabalhos da área, refere-se ao fato de que essa medida só faz sentido quando o modelo multinível escolhido tem apenas o intercepto aleatório (RAUDENBUSH; WILLMS, 1995). No caso do modelo (5), adotado neste trabalho, observa-se que o intercepto β_{0j} possui sua parte randômica, mas os demais fatores, tais como β_{1j} , β_{2j} etc. estão fixos, possibilitando assim o cálculo da partição da variância feito no item 3.3. Portanto, a medida de efeito da escola, u_{0j} , é mais geral, pois ela estará presente como resultado de qualquer modelo implementado, ao contrário da medida de heterogeneidade, que depende das restrições dos modelos, já citadas anteriormente (RAUDENBUSH; WILLMS, 1995).

Em termos de efeito da escola propriamente dito, e não de heterogeneidade, percebe-se que no Brasil a escola também faz diferença. Pode-se afirmar, portanto, que o estabelecimento de metas de desempenho deveria ser tema freqüente nas discussões acadêmicas e governamentais, de tal forma que as práticas internas às escolas pudessem ser constantemente fomentadas, visando a um aumento do resultado final.

Neste trabalho, propôs-se a diferenciação conceitual entre heterogeneidade das escolas e efeito das escolas, com base nas evidências empíricas apresentadas. Apesar disso, quando é possível calcular as duas medidas, é preciso observar que ambas as definições baseiam-se nos mesmos princípios básicos, isto é, utilizam os mesmos elementos de um modelo hierárquico, que fora ajustado a um determinado conjunto de dados. A diferença encontra-se no fato de que a heterogeneidade sintetiza tais elementos em um único número, enquanto na outra definição esses elementos são mantidos separados para cada escola. As estimativas dessas variâncias, entre escolas $\sigma_{u0}^2 = \text{Var}(u_{0j})$, quanto intra-escola $\sigma_r^2 = \text{Var}(r_{ij})$, são fornecidas

rotineiramente pelos softwares utilizados para se estimar os modelos hierárquicos, no caso, o HLM é o mais comum, mas, nesta pesquisa, foi utilizado o SPSS 13.0. No entanto, pode-se ilustrar o fato de que ambas advêm dos mesmos elementos de um modelo, calculando-se a medida da heterogeneidade diretamente a partir dos valores das variâncias dos u_{0j} e dos e_{ij} . Esta estimativa mostra que, na realidade, os dois conceitos de efeito da escola são similares e consistem em dar ênfase a aspectos diferentes de parâmetros do modelo de análise. Essa situação é comprovada pela convergência dos resultados apresentados na Tabela 22 e nos gráficos 26, 27 e 28.

A princípio, parece que existe uma relação entre a diferença das duas formas de se calcular a heterogeneidade e o número de alunos. Na 4ª série do ensino fundamental da rede pública, o número de alunos e de escolas é bem maior que na 3ª série do ensino médio da mesma rede, e, na medida em que o número de alunos diminui, também diminui a diferença. Mas isto é apenas uma constatação e não uma explicação do ocorrido. Essa questão ainda está em aberto e precisa ser mais bem investigada.

A partir do modelo base escolhido, pôde-se propor análises a respeito do impacto do efeito da escola sobre a qualidade do sistema educacional brasileiro. No entanto, outras análises, não implementadas aqui, podem ser feitas a partir de configurações diversas do modelo base. Pode-se citar, por exemplo, a análise do efeito da escola sobre a equidade socioeconômica. Neste caso, o nível 1 do modelo base seria mantido, enquanto o nível 2 seria ligeiramente modificado. Pode ser observado abaixo que a medida de interesse, nesse caso, seria o u_{1j} , uma vez que o modelo passaria a ter não só o intercepto aleatório, mas também inclinações distintas para cada escola, no que se refere ao impacto do nível socioeconômico dos alunos sobre o seu desempenho cognitivo. Outras configurações também poderiam ser propostas e, conseqüentemente, novos estudos realizados.

$$\begin{aligned} \text{Nível 1:} \quad Y_{ij} = & \beta_{0j} + \beta_{1j} NSE_{ij} + \beta_{2j} SEXO_{ij} + \beta_{3j} PARDO_{ij} + \beta_{4j} PRETO_{ij} + \\ & \beta_{5j} ATRASO_{ij} + \beta_{6j} DISCIPLINA_{ij} + \beta_{7j} ATITUDE_{ij} + r_{ij} \quad (8) \end{aligned}$$

$$\begin{aligned} \text{Nível 2:} \quad \beta_{0j} = & \gamma_{00} + \gamma_{01} MNSE_j + \gamma_{02} REDE_j + \gamma_{03} MATRASO_j + u_{0j} \\ \beta_{1j} = & \gamma_{10} + u_{1j} \\ \beta_{2j} = & \gamma_{20} \\ \beta_{3j} = & \gamma_{30} \\ \beta_{4j} = & \gamma_{40} \end{aligned}$$

$$\beta_{5j} = \gamma_{50}$$

$$\beta_{6j} = \gamma_{60}$$

$$\beta_{7j} = \gamma_{70}$$

E o último ponto, mas, talvez o mais importante de todos, refere-se ao fato preocupante, que distingue o sistema educacional brasileiro dos sistemas dos países desenvolvidos ora citados. Como visto anteriormente, ambos apresentam uma medida de heterogeneidade baixa, porém, uma coisa é apresentar semelhança entre as escolas, mas com alto desempenho dos seus alunos; outra coisa, completamente diferente, é apresentar semelhança entre as escolas, mas com baixo desempenho. Infelizmente, o Brasil encontra-se nessa última situação, que poderia ser caracterizada como uma “homogeneidade perversa”. Resultados de estudos internacionais, como o PISA, comprovam esse baixo desempenho. Utilizando-se, também, de dados do próprio SAEB, o movimento “Todos pela Educação²⁹” mostra que o percentual de alunos brasileiros, com desempenho cognitivo adequado à série que estão cursando, é muito pequeno, principalmente em Matemática. As tabelas a seguir retratam essa realidade para os cinco ciclos do SAEB estudados nesta pesquisa, ou seja, a despeito de nossa medida de heterogeneidade ser comparável à medida dos países desenvolvidos, seria um equívoco comparar as nossas escolas com as deles.

²⁹ O projeto “Todos pela educação” é uma iniciativa da sociedade civil, da iniciativa privada, de organizações sociais, educadores e gestores públicos de Educação, que tem como objetivo garantir Educação Básica de qualidade para todos os brasileiros até 2022, bicentenário da Independência do país. A atuação do movimento inclui o monitoramento da Educação, por meio da divulgação de pesquisas, dados e informações relacionadas ao tema; a maior e melhor inserção da Educação na mídia, o fomento ao debate e à mobilização, e o estímulo à formação de agendas locais de acompanhamento, cobrança e apoio. Para alcançar a educação que o Brasil precisa, foram definidas cinco metas específicas, simples, compreensíveis e focadas em resultados mensuráveis, que devem ser atingidas até 7 de setembro de 2022: (1) Toda criança e jovem de 4 a 17 anos na escola. (2) Toda criança plenamente alfabetizada até os 8 anos. (3) Todo aluno com aprendizado adequado à sua série. (4) Todo jovem com o Ensino Médio concluído até os 19 anos. (5) Investimento em Educação ampliado e bem gerido. As Metas, comunicadas constantemente, servirão como direcionamento para que todos os brasileiros acompanhem e cobrem melhorias na Educação. Mais informações sobre o movimento podem ser consultadas no sítio <<http://www.todospelaeducacao.com.br>>.

TABELA 23: Percentual de alunos no Brasil com desempenho adequado* à sua série em Matemática

Série escolar	SAEB 1995	SAEB 1997	SAEB 1999	SAEB 2001	SAEB 2003
4ª EF	20,3	22,3	15,1	16,1	16,8
8ª EF	17,1	17,1	13,2	13,4	14,6
3ª EM	10,9	17,9	11,9	11,6	12,4

FONTE: Todos pela Educação

* Na 4ª EF foram considerados os alunos com desempenho superior a 225 pontos; Na 8ª EF foram considerados os alunos com desempenho superior a 300 pontos; e na 3ª EM foram considerados os alunos com desempenho superior a 350 pontos. Todos esses valores foram retirados da Meta 3 do compromisso “Todos pela educação”.

TABELA 24: Percentual de alunos no Brasil com desempenho adequado* à sua série em Língua Portuguesa

Série escolar	SAEB 1995	SAEB 1997	SAEB 1999	SAEB 2001	SAEB 2003
4ª EF	42,3	36,7	26,0	25,6	28,1
8ª EF	38,2	32,4	18,6	21,8	20,0
3ª EM	45,3	39,7	27,6	25,8	26,9

FONTE: Todos pela Educação

* Na 4ª EF foram considerados os alunos com desempenho superior a 200 pontos; Na 8ª EF foram considerados os alunos com desempenho superior a 275 pontos; e na 3ª EM foram considerados os alunos com desempenho superior a 300 pontos. Todos esses valores foram retirados da Meta 3 do compromisso “Todos pela educação”.

Capítulo 4 - Diferenças entre a escola pública e a privada

4.1 Contextualizando o problema

A Constituição Federal de 1988 estabeleceu que as escolas de ensino básico brasileiras podem ser públicas ou privadas. Nas públicas, organizadas e mantidas pelos governos municipais, estaduais ou federal, o ensino básico é gratuito. As privadas podem ser escolas comunitárias, filantrópicas e confessionais ou simplesmente empresas com fins lucrativos. Analisar as diferenças existentes, em termos de resultados educacionais, entre a rede pública de ensino e a rede particular é uma questão de pesquisa relevante e atual, embora não seja “de hoje que este problema vem atravessando nossa história republicana” (CURY, 1992, p. 34).

O senso comum da sociedade brasileira aponta que as melhores escolas estão na rede privada. Akkari (2001) justifica o surgimento dessa idéia com argumentos de três naturezas: primeiro, o fato de essa rede ter se tornado, ao longo dos tempos, a única opção usada pelos filhos da elite social; segundo, pelas disparidades, em termos de desempenho cognitivo, as quais implicariam, até mesmo, em dizer que o sistema educacional brasileiro se consolida em redes de várias velocidades, e, por último, as

despesas por aluno podem também dar um indício da diferença de qualidade entre as redes pública e particular. As mensalidades pagas pelas famílias para cada criança freqüentando o ensino particular são, em média, amplamente superiores às despesas públicas anuais por aluno do ensino público (AKKARI, 2001, p. 171).

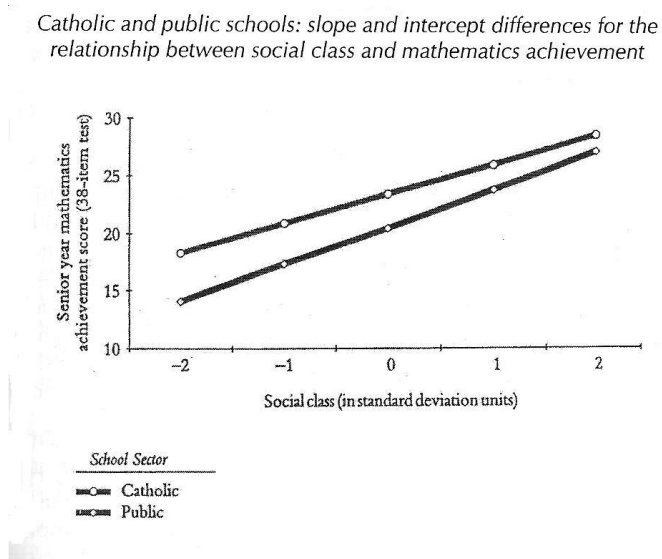
Se isso for verdade, é necessário que se aprenda com elas, ou, utilizando-se das palavras de Cury (1992), fazer uma leitura crítica da escola pública a partir do padrão de qualidade da privada.

Por outro lado, segundo os dados do Censo Escolar de 2006 (BRASIL/MEC/INEP, 2007), 44,8% dos alunos matriculados na educação básica estão na rede municipal e 41,7% estão na rede estadual, perfazendo um total de 86,5% das

matrículas de educação básica no ensino público³⁰. Diante disso, só haverá educação de qualidade no Brasil quando o ensino público for de qualidade, produzindo excelência do aprendizado dos alunos e também mediando os determinantes sociais de baixo desempenho.

Estudos internacionais, que são referência na linha de pesquisa sobre efeito da escola, indicam que nos Estados Unidos as escolas católicas (que são a maioria das escolas particulares dos EUA) são melhores que as públicas (OPDENAKKER; DAMME, 2006; LEE, 2001b; BRYK; LEE; HOLLAND, 1993; COLEMAN; HOFFER, 1987). Melhores, não só por terem maior qualidade, mas, ao contrário do que se poderia esperar, melhores também por terem maior equidade. Registre-se aqui a preocupação, nesta tese, não só com a qualidade dos sistemas de ensino, mas também com sua equidade. A Figura 2 mostra o que Lee (2001b) chama de “Distribuição social do conhecimento” e ilustra, com clareza, como as escolas católicas são melhores que as públicas no contexto norte-americano.

FIGURA 2: Relação entre desempenho em matemática e NSE para escolas públicas e católicas nos Estados Unidos



FONTE: Lee, 2001b. p. 51

³⁰ Em todas as análises desta tese, a menos que seja registrado o contrário, ensino público refere-se às redes municipal e estadual. A rede federal de ensino, apesar de pública, tem uma baixa representatividade em âmbito nacional, principalmente em alguns segmentos de ensino, como as séries iniciais do ensino fundamental. Além disso, são escolas que, normalmente, têm um rigoroso processo de seleção de alunos na entrada, distinguindo-as das demais escolas públicas. Um estudo detalhado sobre suas estratégias de funcionamento e o impacto disso no efeito da escola, sem dúvida, é importante, mas foge aos propósitos desta tese.

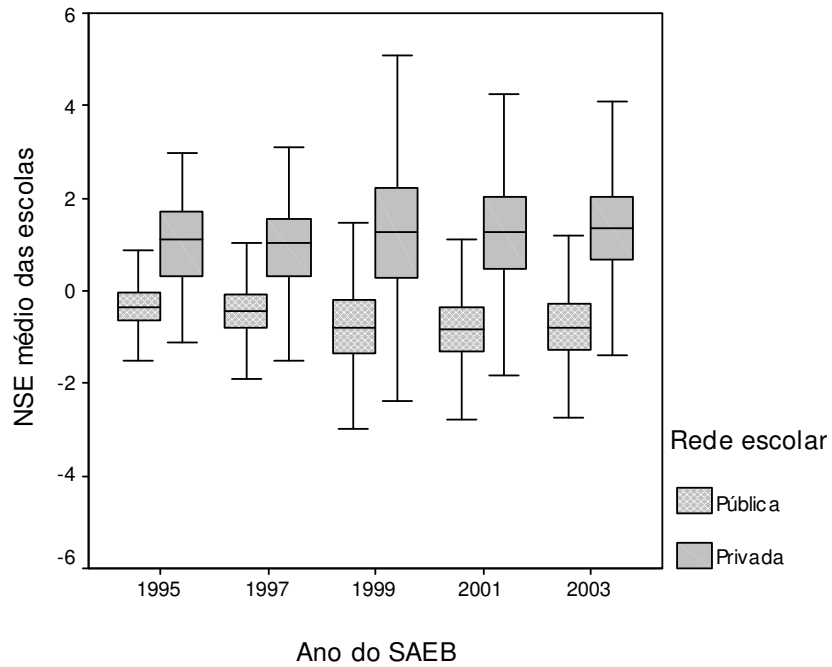
O intercepto das escolas católicas é maior, mesmo depois de se controlar o desempenho por características de *background* dos alunos, denotando a sua maior qualidade. E, enquanto a inclinação da reta, que sintetiza as escolas públicas é igual a 3,23, a inclinação das escolas católicas é menor, igual a 2,50, implicando também em maior eqüidade. Aqui, deve-se recordar o que foi apresentado no item 2.4.3 desta tese em relação à interpretação de resultados, ou seja, o intercepto está associado à qualidade e o inverso da inclinação à eqüidade. Quanto maior o intercepto, maior será a qualidade. E quanto menor a inclinação dessa reta de regressão, maior será a eqüidade.

Baseando-se neste referencial teórico, procede-se ao estudo comparativo de ambas as redes para o contexto brasileiro, o qual está organizado da seguinte forma: primeiro faz-se uma descrição do alunado das escolas públicas e das escolas privadas, explicitando quão diferentes eles são. Em seguida, mostra-se que as diferenças persistem, mesmo depois de se controlar pelas características sócio-demográficas e de trajetória escolar desses alunos, diferenças relacionadas tanto à qualidade quanto à eqüidade.

4.2 O alunado das escolas públicas e privadas

O controle por variáveis de caracterização sociodemográfica do aluno é essencial. Aqui, as diferenças do alunado atendido, em cada uma das redes, são marcantes. O exemplo mais expressivo é o fato da rede particular atender, em média, a alunos mais favorecidos socioeconomicamente, ainda que exista uma maior variabilidade, no que se refere ao NSE dos seus alunos, explicada pelas políticas de bolsas de estudos adotadas por muitas escolas. O gráfico a seguir compara o nível socioeconômico médio das escolas de cada uma das redes e, a partir dele, percebe-se que o maior NSE das escolas públicas chega a atingir, no máximo, a mediana das escolas particulares.

GRÁFICO 29: NSE médio das escolas públicas e privadas nos cinco ciclos do SAEB



FONTE: Elaboração própria com os dados do SAEB – INEP/MEC.

* Amostra aleatória de 10%

Outra fonte importante de informação, como se verá em detalhes no próximo capítulo, é em relação à raça, isto é, como se caracterizam as duas redes em termos de distribuição racial. A próxima tabela mostra que, sistematicamente, na rede privada, o percentual de alunos que se autodeclara branco é maior do que esse percentual na rede pública. Ao passo que o percentual de alunos que se autodeclara pardo ou preto é sempre maior na rede pública. A clientela atendida pelas duas redes é, também, distinta do ponto de vista racial.

TABELA 25: Percentual da classificação de cor/raça segundo a autodeclaração dos alunos nos questionários do SAEB, por dependência administrativa*

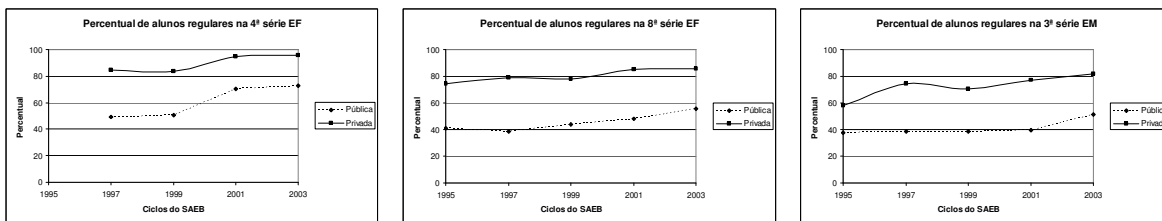
Série - Ciclo	Branco		Pardo		Preto	
	Pública	Privada	Pública	Privada	Pública	Privada
4 ^a EF – 1995	43,4	52,5	40,8	36,0	11,8	6,3
8 ^a EF – 1995	50,7	68,8	36,5	22,6	8,4	3,8
3 ^a EM – 1995	53,5	67,7	36,0	22,9	6,4	4,2
4 ^a EF – 1997	43,3	59,8	40,1	29,2	9,3	3,5
8 ^a EF – 1997	51,9	68,0	33,6	21,1	6,4	1,9
3 ^a EM – 1997	57,3	69,8	30,5	20,0	5,1	2,2
4 ^a EF – 1999	42,0	55,2	38,0	31,2	12,3	4,4
8 ^a EF – 1999	49,7	64,6	35,5	22,8	7,5	3,5
3 ^a EM – 1999	54,1	69,9	32,2	21,0	6,7	2,5
4 ^a EF – 2001	43,2	57,3	36,9	30,0	13,2	4,7
8 ^a EF – 2001	44,2	64,1	38,7	23,0	8,8	3,0
3 ^a EM – 2001	50,8	67,9	34,9	21,9	7,0	3,0
4 ^a EF – 2003	36,6	53,8	45,3	34,9	11,9	3,9
8 ^a EF – 2003	40,3	61,2	43,4	26,8	9,0	3,8
3 ^a EM – 2003	44,6	62,7	39,4	28,0	8,9	3,6

FONTE: Elaboração própria com os dados do SAEB – INEP/MEC.

* Nesta síntese, foram considerados apenas os dados válidos. A soma dos percentuais não está totalizando o 100%, pois não se apresentou nesta tabela os dados referentes a “amarelo” e “indígena”.

Do ponto de vista da trajetória escolar, o alunado atendido pelas redes pública e privada no Brasil também é muito distinto. Analisando-se a variável “atraso”, que mede a defasagem idade-série em número de anos, observa-se, para todas as combinações série-ciclo do SAEB, que o percentual de alunos regulares é sempre maior na rede particular. Mesmo constatando-se um aumento desses percentuais, a partir de 1999, para a 4^a e 8^a séries do ensino fundamental, e, a partir de 2001, para a 3^a série do ensino médio, a rede particular mantém, sempre, uma diferença igual ou superior a 20 pontos percentuais.

Gráfico 30: Percentual de alunos regulares nos cinco ciclos do SAEB – 4ª série EF
 GRÁFICO 31: Percentual de alunos regulares nos cinco ciclos do SAEB – 8ª série EF
 GRÁFICO 32: Percentual de alunos regulares nos cinco ciclos do SAEB – 3ª série EM



FONTE: Elaboração própria com os dados do SAEB – INEP/MEC.

E como era esperado, tendo em vista esta situação contextual descrita anteriormente, a proficiência média dos alunos da rede particular é maior que a proficiência média dos alunos da rede pública. As tabelas abaixo mostram este resultado para Matemática e para Língua Portuguesa, nas três séries avaliadas dos cinco ciclos do SAEB.

TABELA 26: Proficiência média em Matemática por rede

Série - Ciclo	Pública	Privada
4ª EF – 1995	186,1	222,2
8ª EF – 1995	245,7	292,8
3ª EM – 1995	272,5	306,5
4ª EF – 1997	185,0	231,0
8ª EF – 1997	240,8	300,6
3ª EM – 1997	271,1	337,6
4ª EF – 1999	176,4	217,6
8ª EF – 1999	239,3	293,7
3ª EM – 1999	267,9	329,8
4ª EF – 2001	170,8	221,8
8ª EF – 2001	235,4	301,1
3ª EM – 2001	264,7	338,6
4ª EF – 2003	171,5	223,8
8ª EF – 2003	236,8	304,2
3ª EM – 2003	266,0	340,2

FONTE: Elaboração própria com os dados do SAEB – INEP/MEC.

TABELA 27: Proficiência média em Língua Portuguesa por rede

Série - Ciclo	Pública	Privada
4 ^a EF – 1995	184,0	217,2
8 ^a EF – 1995	251,2	283,0
3 ^a EM – 1995	281,4	307,3
4 ^a EF – 1997	180,9	224,6
8 ^a EF – 1997	243,4	286,5
3 ^a EM – 1997	271,6	317,9
4 ^a EF – 1999	165,9	208,9
8 ^a EF – 1999	227,3	270,6
3 ^a EM – 1999	256,8	305,8
4 ^a EF – 2001	159,9	209,2
8 ^a EF – 2001	228,7	282,4
3 ^a EM – 2001	253,2	309,6
4 ^a EF – 2003	163,9	214,8
8 ^a EF – 2003	225,6	278,3
3 ^a EM – 2003	257,1	314,0

FONTE: Elaboração própria com os dados do SAEB – INEP/MEC.

4.3 A qualidade do público e do privado

Tendo em vista tantas distinções entre o alunado atendido em cada uma das redes, seria um erro comparar a qualidade de ambas pelas tabelas anteriores. É fato que as tabelas demonstram um retrato da realidade, mas deve-se ajustar um modelo que controle o efeito das variáveis contextuais, para que se obtenha uma medida mais fidedigna da qualidade dessas redes.

Para enfrentar essa questão da pesquisa, a variante adotada do modelo base é a seguinte:

$$\text{Nível 1: } Y_{ij} = \beta_{0j} + \beta_{1j} NSE_{ij} + \beta_{2j} SEXO_{ij} + \beta_{3j} PARDO_{ij} + \beta_{4j} PRETO_{ij} + \beta_{5j} ATRASO_{ij} + \beta_{6j} DISCIPLINA_{ij} + \beta_{7j} ATITUDE_{ij} + r_{ij} \quad (9)$$

$$\text{Nível 2: } \beta_{0j} = \gamma_{00} + \gamma_{01} MNSE_j + \gamma_{02} REDE_j + \gamma_{03} MATRASO_j + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} MNSE_j + \gamma_{12} REDE_j + \gamma_{13} MATRASO_j + u_{1j}$$

Em relação ao modelo base, a mudança aqui implementada foi o acréscimo no coeficiente β_{1j} das variáveis de controle do nível da escola, quais sejam: nível socioeconômico médio dos alunos da escola (*MNSE*), dependência administrativa da escola (*REDE*) e média dos anos de atraso dos alunos da escola (*MATRASO*). Isso se justifica, porque existem outras diferenças entre o público e o privado, além do tipo de aluno. Por exemplo, as diferenças na possibilidade de gestão ou as diferenças de seleção/manutenção de alunos para a escola. Além disso, tais controles permitem observar o impacto dessas variáveis, não só na qualidade de cada rede de ensino, mas também na equidade socioeconômica.

Tendo ajustado um modelo hierárquico aos dados dos cinco ciclos do SAEB, interessa analisar agora a medida da qualidade do sistema educacional brasileiro para cada um dos setores de ensino. Já é um ponto tácito que a proficiência bruta é importante, pois revela o nível atual dos alunos, mas não fornece informação sobre a qualidade do ensino. Tal informação, no que se refere à diferença entre as duas redes, é fornecida pelo coeficiente γ_{02} do modelo (9) e registrado na Tabela 28, o coeficiente da variável indicadora *REDE*. Essa variável assume o valor “0” quando a escola pertence à rede pública de ensino e o valor “1” se é da rede particular. Sendo assim, se o coeficiente γ_{02} for negativo, isso implicará em uma melhor qualidade da escola pública, mas se ele for positivo, significa quantos pontos são adicionados, em média, à nota dos alunos pelo fato de se estudar na rede particular. Ou seja, como se pode ver na tabela a seguir, a rede particular possui maior qualidade.

TABELA 28 – Coeficiente γ_{02} ajustado para o modelo (9) –
Variável associada: Rede

Ano	4ª EF	8ª EF	3ª EM
1995	-	2,44	5,46*
1997	20,53**	17,37**	19,05**
1999	18,97**	14,45**	16,93**
2001	17,05**	16,68**	14,45**
2003	16,15**	18,16**	17,36**

(*) = $p < 0,05$ e (**) = $p < 0,001$

Todos os valores para o coeficiente γ_{02} são marcadamente positivos. Ainda que para o ciclo de 1995 os valores sejam estranhamente baixos, para as demais combinações série/ciclo, os valores são sempre maiores que 14 pontos, com uma

média de 16,15 pontos. Retomando-se a discussão feita no item 2.2, percebe-se que esses 16,15 pontos indicam uma diferença substancial. Significam aproximadamente um ano de escolaridade a mais pelo simples fato de se estudar na rede particular, já que o efeito dos outros fatores foi controlado. E na 4ª série do ensino fundamental, essa diferença entre as duas redes é ainda maior, alcançando os 18 pontos em média.

Portanto, mesmo depois de se controlar o resultado pelas características de contexto dos alunos e das escolas, a rede particular apresenta melhores resultados em relação à qualidade do ensino ofertado.

4.4 As diferenças em relação à equidade socioeconômica

O tempo todo, nesta tese, tem-se deixado claro que a qualidade dos sistemas educativos está sendo investigada com base no desempenho cognitivo dos seus alunos, que é aferido por testes individuais de Matemática e Leitura. Esse é um recorte feito, afinal, sabe-se que as escolas desenvolvem muitas outras competências além das cognitivas. Esse recorte, às vezes, é utilizado como desculpa pela rede pública para justificar a má qualidade constatada. O argumento, nesses casos, é de que a rede particular preocupa-se muito mais com o cognitivo, principalmente, por conta da preparação de seus alunos para os exames vestibulares e pela autonomia que essas escolas têm para moldar seu projeto pedagógico. Em contrapartida, na rede pública, é uma minoria de alunos que busca os exames vestibulares mais concorridos, e o projeto pedagógico, ainda que construído pela comunidade escolar, tem seus moldes “amarrados” pelas propostas das secretarias de educação. Nesse contexto, o discurso de defesa, apresentado pela rede pública, gira em torno do argumento que não se tem uma qualidade como seria desejável, mas, em compensação, forma-se o aluno como cidadão e promove-se a educação para todos, ou seja, é uma rede rica em projetos que maximizem a equidade nas escolas.

No entanto, Castro (1997, p. 425) afirma que “não há qualquer evidência empírica de que as escolas públicas formem melhores cidadãos” e os resultados do ajuste do modelo (9) não corroboram esse discurso e explicitam que, se existe uma rede destacando-se em relação à equidade socioeconômica, é a rede particular de ensino. O coeficiente γ_{12} do modelo (9) fornece a medida da influência da rede no coeficiente do indicador socioeconômico. Valores negativos desse coeficiente indicam que a rede particular possui maior equidade socioeconômica, enquanto valores

positivos indicam a pública como mais equitativa. Isso porque a equidade está associada a um coeficiente de inclinação de uma reta. Tanto maior será a equidade socioeconômica quanto menor for a inclinação associada à variável *NSE*, ou seja, quanto mais próximo de zero for o valor de β_1 . Como os coeficientes associados ao *NSE* do aluno são sempre positivos, valores negativos de γ_{12} diminuem β_1 e aumentam a equidade.

A Tabela 29 mostra que a maior parte dos coeficientes ajustados para as diversas combinações série/ciclo é negativa. Na média, o efeito da rede no *NSE* do aluno é -0,40, sendo -0,16 para a 4ª série do ensino fundamental, -0,57 para a 8ª série do ensino fundamental e -0,47 para a 3ª série do ensino médio.

TABELA 29 – Coeficiente γ_{12} ajustado para o modelo base –
Variável associada: Rede*NSE

Ano	4ª EF	8ª EF	3ª EM
1995	-	-0,27	0,90
1997	0,97	1,51	-1,17
1999	0,03	-0,83	-1,84*
2001	-0,16	-0,71	0,67
2003	-1,08	-2,05*	-1,06

(*) = $p < 0,05$ e (**) = $p < 0,001$

Mais uma vez se justifica o propósito desta pesquisa, pois o estudo desse fenômeno, em uma das bases escolhidas aleatoriamente, poderia levar o pesquisador a uma determinada conclusão, que não se repete nos demais ciclos do SAEB ou nas demais séries avaliadas.

Para o sistema brasileiro de educação básica, conclui-se, então, que está refutado o argumento de que a rede particular é boa porque atrai bons alunos. Controladas as características de entrada dos alunos, a rede particular ainda se mostra muito melhor que a pública, seja na produção da qualidade, seja na capacidade de produzir a equidade socioeconômica desse desempenho.

4.5 Conclusões: uma leitura crítica a partir de um padrão de qualidade

Ainda que este capítulo tenha como título e como foco as diferenças entre o ensino público e o privado, não se pretende concluí-lo elegendo uma das duas redes como modelo de excelência. Não se deseja a inserção deste trabalho no cerne de uma batalha ideológica, com a aclamação de uma das redes e o aniquilamento da outra. Defender a educação para todos não significa, necessariamente, defender uma educação pública para todos. O ensino privado representa uma economia de recursos e de “energia administrativa” do poder público e, principalmente, aumenta a variedade de alternativa do sistema.

Sendo assim, a reflexão que precisa ser feita refere-se às práticas que levam ao sucesso escolar, afinal “o que interessa é encontrar as condições que fazem com que funcione bem o ensino, seja público ou privado” (CASTRO, 1997, p. 424). Ou, em outras palavras, mais duras do mesmo autor: “a real opção não é entre público e privado, mas entre regra boa e a regra podre” (CASTRO, 1997, p. 431).

Este trabalho não nos permite identificar exatamente quais fatores provocam as diferenças de resultados entre ambas as redes, mas apenas constatar que elas existem, mesmo depois do controle por uma série de variáveis. No entanto, nesta busca de uma regra que seja favorável à qualidade do ensino, defende-se a existência de uma rotina no dia a dia das escolas; uma busca por profissionais mais qualificados; e uma gestão mais autônoma das escolas no sistema, ainda que avaliada pelo Estado, mas em uma perspectiva de que, em economias mais competitivas, o mercado se auto-regula. É uma busca do “caminho para a modernidade” (ZIBAS, 1997, p. 64). Afinal,

a discussão da dinamização do sistema escolar não se localiza mais na questão do caráter privado ou estatal dos estabelecimentos, mas nos estilos de gestão que caracterizam um ou outro setor. A realidade dos países em desenvolvimento demonstra que o serviço estatal é o único que chega aos setores pobres (...) Mas essa mesma experiência mostra que a forma com a qual o Estado enfrenta o desafio de oferecer este serviço é ineficiente e excludente. Inversamente, o setor privado possui uma dinâmica de gestão que lhe permite ser eficiente, criativo e flexível, mas está dirigido somente aos setores sociais mais favorecidos (TEDESCO 1991, apud ZIBAS, 1997, p. 74).

Em resumo, se a hipótese de que existe maior qualidade na rede particular for comprovada nesta tese, é urgente aprender com o seu padrão de qualidade (CURY, 1992) para que se tenha uma formulação de políticas públicas educacionais bem fundamentadas, que visem igualar esse quadro e oferecer um ensino adequado, também, à grande maioria da população brasileira.

Para concluir, é necessário registrar que tais sugestões e os resultados apresentados neste capítulo não gozam de unanimidade na literatura. Alguns autores, apesar de reconhecerem as desigualdades educativas estruturais, insistem na tese de que a solução para se equiparar os resultados de ambas as redes passa unicamente por melhorar a estrutura e o apoio político ao setor público, ou, em outras palavras, “a questão-chave é a insuficiência dos recursos para a educação pública” (AKKARI, 2001, p. 185). Já outros autores nem sequer reconhecem a melhor posição da rede privada afirmando

que, em toda a América Latina, a iniciativa privada não tem demonstrado eficiência e dinamismo junto às camadas populares. Quanto ao seu êxito na educação das classes médias, é preciso lembrar que, quando dirigida a esses setores, a educação estatal também tem, em geral, obtido bons resultados (ZIBAS, 1997, p. 74).

Espera-se, no entanto, que a utilização de modelos multiníveis, controlados pelo NSE do aluno e da escola, e por outras características de *background*, seja suficiente para contrapor tais argumentos.

Capítulo 5 – O efeito da cor/raça no desempenho dos alunos

5.1 O contexto da questão racial

Este capítulo pretende contribuir para a discussão da dimensão educacional da questão racial no Brasil, sem, no entanto, pretender fornecer uma análise geral da questão, pois isso demandaria abordagens que ultrapassam, em muito, os limites desta tese. De forma específica, este capítulo pretende apresentar os resultados das diferenças de desempenho cognitivo entre grupos de alunos segmentados por sua cor/raça. Trabalhos como o de Osorio e Soares (2005) mostraram, ainda que utilizando metodologia e dados diferentes dos que serão utilizados neste trabalho, que o sistema educacional brasileiro, além de não conseguir compensar as diferenças iniciais, acrescenta outras distorções, provocando grandes disparidades entre brancos e pretos, e, diferenças mais tênues entre brancos e pardos. Adotamos a mesma postura desses autores quando afirmam:

[...] sejam quais forem os motivos da defasagem – origem familiar, baixa auto-estima, discriminações presentes, ou qualquer outra razão – o sistema de ensino tem a obrigação de compensar as desvantagens e trazer os alunos negros ao nível dos brancos. Isso não é uma questão de caridade, mas de direitos humanos e justiça social. (OSORIO; SOARES, 2005, p. 35)

Naturalmente, para analisar a influência da cor/raça no desempenho dos alunos, é preciso definir conceitualmente este fator. Há dificuldades não triviais em relação à raça. Alguns autores negam qualquer possibilidade de definição de raça entre os seres humanos, outros, como Osorio (2003), ressaltam que a existência de raças

expressa o fato de que há diferenças biológicas entre grandes grupos de indivíduos que são sensíveis e classificáveis, mas não autoriza o racismo, que é um conjunto de construções culturais sobre

essas diferenças que lhes atribui um sentido que não é 'natural' (OSORIO, 2003, p. 11).

Reconhece-se, portanto, que o conceito de raça nos seres humanos não é apenas um conceito biológico, mas um construto social, ou, como proposto por Guimarães (1999), o conceito a ser considerado é o de “raça social”. E, como vivemos “em uma sociedade em que as características físicas se mostram mais relevantes que a origem social” (SOARES, R. M., 2006, p. 27), a cor funciona como um dos mais importantes elementos para a tal construção social das relações raciais. Diante disso, nesta tese, quando se falar em raça ou cor, não estará sendo feita distinção conceitual entre os termos.

O fato é que há muitos outros aspectos e nuances na definição do conceito de raça. Aos interessados, reitera-se a sugestão dos textos de Soares, R. M. (2006) e Osorio (2003); e acrescentam-se os trabalhos de Pena (2000), que enfatiza a natureza triíbrida da população brasileira, a partir da miscigenação de africanos, europeus e ameríndios, tornando o Brasil uma “democracia racial”; de Kamel (2006), o qual propõe que nós brasileiros não somos racistas, uma vez que não existem barreiras institucionalizadas para a progressão do negro em nossa sociedade; e de autores como Hasenbalg (1979), que dizem exatamente o contrário, ou seja, que não somos uma democracia racial e que somos, sim, um país com preconceito de cor, portanto, um país com barreiras para progressão dos negros na sociedade.

5.1.1 Identificação da pertença racial

Pode-se entender como método de identificação a maneira utilizada para definir a pertença dos indivíduos aos determinados grupos raciais. Tão mais fácil será essa identificação quanto mais próximo do estereótipo do grupo estiver o indivíduo. E tão mais difícil será, quanto mais próximo estiver da fronteira entre dois grupos. Osorio (2003) detalha três métodos de classificação:

1) A identificação biológica - esse método utiliza a análise de DNA para inferir a qual grupo racial teria pertencido os ancestrais de uma pessoa. Mesmo considerando os avanços da biologia, em especial, da genética, esse método não é utilizado nos levantamentos de informação social. Primeiro, pelo alto custo associado; segundo, pois “a diversidade humana, a despeito da uniformidade do genoma, é tão grande que, se fossem definidas raças com base em similitudes genéticas, provavelmente haveria

tantas raças quanto existem línguas no mundo” (Cavalli-Sforza *apud* Osorio, 2003, p. 10); terceiro, e mais importante argumento, é o problema de que tal método não proporciona “correspondência entre as ‘raças’, que a genética pode oferecer, e o que as sociedades entendem por raça” (Osorio, 2003, p. 11).

2) A auto-atribuição - ainda que este método possa conter o viés da variação social da cor, existem recomendações internacionais para que ele seja preterido em função dos demais.

3) Heteroatribuição - é o método no qual a informação sobre a cor/raça é fornecida por uma outra pessoa. Esses dois últimos métodos são os preferidos nos levantamentos de informações nacionais.

Mas, ainda que Osorio finalize a discussão sobre os métodos afirmando que, “aparentemente, a auto-atribuição parece engendrar uma distribuição de cor mais acurada do que a heteroatribuição, embora os resultados desta não desautorizem seu uso” (Osorio, 2003, p. 17), precisa-se ter consciência, como ele mesmo afirma, que, “no fundo, a opção pela auto-atribuição ou pela heteroatribuição de pertença racial é uma escolha entre subjetividades: a do próprio sujeito da classificação ou a do observador externo” (Osorio, 2003, p. 13).

O questionário contextual, preenchido pelo aluno quando realiza o teste do SAEB, é utilizado como fonte para a coleta da pertença racial. Assim sendo, analisou-se dados, que foram obtidos pela forma de auto-atribuição induzida. Aqui vale o registro de que o questionário do SAEB foi o primeiro instrumento de larga aplicação, no Brasil, que introduziu o quesito cor/raça.

5.1.2 A classificação racial

Se há uma relativa facilidade de convergência na escolha de métodos, o mesmo não se pode afirmar para a questão da classificação racial propriamente dita. Principalmente no Brasil, essa classificação mostra-se extremamente dependente de outras variáveis sociais. Carvalho (2005) afirma que:

no âmbito da escola, a classificação racial das crianças não estava relacionada apenas a suas características físicas e a seu status socioeconômico, como já foi verificado para o conjunto da sociedade brasileira, mas também a seu desempenho escolar, com uma associação, por parte das educadoras, entre pertencimento à raça negra e problemas de desempenho. (CARVALHO, 2005, p. 77)

Ou como cita Soares, R. M. (2006, p. 38) “mais do que informação objetiva acerca de suas origens étnicas, as respostas das crianças tenderiam a apontar o desejo delas de pertencer a grupos étnicos detentores de maior prestígio e privilégios sociais”.

Mas essa dificuldade não é característica apenas das crianças, os adultos também demonstram a não-trivialidade do tema. Osorio (2003) sintetiza uma série de estudos que correlacionam a auto-atribuição com a heteroatribuição ou mesmo a auto-atribuição livre com a induzida. Os resultados desses estudos mostram que, apesar da variedade de nomes que aparecem para se definir a pertença racial, a maioria da população, sempre, se concentra nas categorias branco, pardo ou preto, fato que se repete desde 1872, quando do primeiro Censo oficial brasileiro, o que justifica, assim, as categorias utilizadas atualmente em nosso sistema classificatório pelo Instituto Brasileiro de Geografia e Estatística (IBGE)³¹, que constitui-se como um “importante material de pesquisa, pois são os resultados de contradições e representações cultivadas e desenvolvidas durante séculos” (SOARES, R. M., 2006, p. 31). Isso contraria algumas críticas infundadas de que tais categorias teriam sido inventadas em laboratórios por burocratas do IBGE.

Em relação à forma de coleta desse dado, é necessário registrar que o IBGE classifica a raça ou cor dos entrevistados em branca, parda, preta, amarela ou indígena. Já os questionários contextuais do SAEB, utilizados nesta pesquisa, coletaram este dado da seguinte forma: a pergunta sempre era “Como você se considera?”. As categorias oferecidas, entretanto, sofreram variações ao longo dos cinco ciclos analisados. No ano de 1995, elas eram: branco(a), pardo(a), preto(a) ou amarelo(a). De 1997 em diante, foi acrescentada a categoria indígena. Além disso, nos ciclos de 1997, 1999 e 2001, a categoria parda era apresentada como pardo/mulato.

Para as análises aqui implementadas, não se fez distinção entre pardo e pardo/mulato. Assim, as categorias usadas neste trabalho não são o fruto de uma construção teórica dos autores, mas aquelas apresentadas nos questionários do SAEB. Importante ressaltar, entretanto, que o tratamento em separado das categorias

³¹ Além de “branco”, “pardo” e “preto”, o IBGE passou a adotar, também, as categorias “amarelo”, em função da grande imigração de asiáticos e, desde 1991, a “indígena” que transformou o sistema que era apenas de cor em sistema racial. Mas, ainda assim, deve-se ter em mente que “a cor desempenha vários papéis na classificação. O principal deles é funcionar como forma de aproximar a pertença racial dos indivíduos, pois é a marca mais importante considerada nas situações concretas para a definição dessa pertença” (OSORIO, 2003, p. 23).

pardo e preto não é consensual, sendo inúmeros os pesquisadores que defendem o uso das categorias branco e não-branco, ou branco e negro, em uma associação entre pardos e pretos (HASENBALG, 1979). No entanto, seja pela forma como os dados foram coletados, seja pelos resultados obtidos para a análise do desempenho educacional, argumenta-se, no próximo item, que a análise deve ser feita de forma separada.

5.1.3 Cor/raça dos alunos do sistema brasileiro de educação básica

A Tabela 30 mostra a classificação dos alunos nas categorias de cor/raça para as diferentes séries e ciclos do SAEB

TABELA 30: Percentual da classificação de cor/raça segundo a autodeclaração dos alunos nos questionários do SAEB*

Série - Ciclo	Branco	Pardo	Preto	Amarelo	Indígena
4 ^a EF - 1995	44,4	40,4	11,1	4,1	NA
8 ^a EF - 1995	53,4	34,6	7,7	4,3	NA
3 ^a EM - 1995	57,4	32,4	5,8	4,4	NA
4 ^a EF - 1997	45,3	38,7	8,5	4,5	2,9
8 ^a EF - 1997	54,4	31,7	5,7	6,2	2,0
3 ^a EM - 1997	60,6	27,7	4,3	5,5	1,8
4 ^a EF - 1999	43,5	37,2	11,4	4,7	3,2
8 ^a EF - 1999	51,7	33,8	7,0	5,4	2,2
3 ^a EM - 1999	57,2	29,9	5,9	5,2	1,8
4 ^a EF - 2001	44,7	36,1	12,3	3,8	3,1
8 ^a EF - 2001	46,6	36,8	8,1	6,2	2,3
3 ^a EM - 2001	53,6	32,8	6,4	5,7	1,5
4 ^a EF - 2003	38,6	44,1	11,1	2,6	3,7
8 ^a EF - 2003	42,8	41,4	8,4	3,8	3,6
3 ^a EM - 2003	47,8	37,4	8,0	4,6	2,2

FONTE: Elaboração própria com os dados do SAEB – INEP/MEC.

* Nesta síntese, foram considerados apenas os dados válidos.

Ainda que a metodologia do SAEB não implique no acompanhamento longitudinal dos mesmos alunos, da 4^a série do ensino fundamental até a 3^a série do ensino médio, a análise da frequência de classificação racial, em seus cinco ciclos,

corroborar os resultados mostrados por Osorio e Soares (2005) no documentário estatístico sobre a geração 80. Observa-se que acontece um “embranquecimento” sistemático da população, com o aumento da escolaridade, ou, em palavras mais explícitas: os pardos e pretos não conseguem permanecer no sistema educacional brasileiro na mesma proporção dos anos iniciais do ensino fundamental. Enquanto existem em torno de 44% de brancos nas 4^{as} séries do ensino fundamental, esse percentual se eleva para 57% nas 3^{as} séries do ensino médio. Ao passo que os pardos e pretos, somados, equivalem a cerca de 49% nas 4^{as} séries, mas reduzem-se a apenas 37% no final do ensino médio.

Os dados de 2003 revelam um fenômeno intrigante. A população de brancos diminuiu, refletindo, principalmente, em um aumento de pardos, em todas as séries, e de pretos, na 3^a série do ensino médio. Mantendo-se a suposição de que a distribuição racial brasileira permaneceu constante no período analisado, surgem duas possíveis explicações para o fenômeno. Considerando que a pertença racial é uma construção mais social que biológica, a hipótese mais razoável é que esses alunos passaram a se enxergar como não-brancos, um reflexo das primeiras ações afirmativas³² implantadas no país. Ou então, essa mudança na autoclassificação racial se deu como um reflexo indireto das discussões e implantação de políticas de cotas raciais no ensino superior. Independente da explicação, o fato merece registro, pois, pela primeira vez na história desse país, parte da população, que se autodeclarava branca, agora está optando por outras categorias até então muito discriminadas.

Registra-se um pequeno número de alunos que se autodeclararam amarelos ou indígenas. Além disso, normalmente, esses se encontram concentrados em algumas unidades da federação. Tais fatos, somados ao objetivo principal, que é descrever o sistema como um todo, justificam a opção por excluí-los das análises deste capítulo.

³² “As ações afirmativas podem ser definidas como um conjunto de políticas públicas e privadas de caráter compulsório, facultativo ou voluntário, concebidas com vistas ao combate à discriminação racial, de gênero, por deficiência física e de origem nacional, bem como para corrigir ou mitigar os efeitos presentes da discriminação praticada no passado, tendo por objetivo a concretização do ideal de efetiva igualdade de acesso a bens fundamentais, como a educação” (GOMES *apud* SOARES, R. M., 2006, p. 103).

5.2 Diferenças médias

As tabelas 31 e 32 sintetizam as diferenças de proficiência entre os grupos raciais, para cada série e para todos os ciclos do SAEB considerados, tanto em Matemática quanto em Leitura. O valor apresentado significa, portanto, a diferença entre a proficiência média daqueles alunos que se autodeclararam, no questionário do SAEB, como pardos e a proficiência média daqueles que se autodeclararam como brancos; a diferença entre pretos e brancos; e a diferença entre pardos e pretos. Embora essas sejam as diferenças reais, não se pode atribuir características explicativas às categorias raciais, já que, como dito anteriormente, são muitas as causas para o desempenho dos alunos.

Observa-se que, sistematicamente, as proficiências médias dos brancos são sempre maiores que as dos pardos, as quais, por sua vez, são sempre maiores que as dos pretos, independente da disciplina avaliada.

TABELA 31: Diferença entre as proficiências médias em Matemática segundo a classificação de cor/raça dos alunos nos questionários do SAEB

Série - Ciclo	Pardo - Branco	Preto – Branco	Pardo - Preto
4 ^a EF – 1995	-6,3	-15,3	-9,0
8 ^a EF – 1995	-15,0	-23,0	-8,0
3 ^a EM – 1995	-19,5	-21,4	-1,9
4 ^a EF – 1997	-8,2	-25,8	-17,6
8 ^a EF – 1997	-18,7	-31,1	-12,4
3 ^a EM – 1997	-29,1	-51,9	-22,8
4 ^a EF – 1999	-7,6	-23,1	-15,5
8 ^a EF – 1999	-17,1	-24,0	-6,9
3 ^a EM – 1999	-21,2	-35,8	-14,6
4 ^a EF – 2001	-10,3	-29,3	-19,0
8 ^a EF – 2001	-19,6	-31,5	-11,9
3 ^a EM – 2001	-23,1	-35,9	-12,8
4 ^a EF – 2003	-11,3	-30,5	-19,2
8 ^a EF – 2003	-20,3	-34,5	-14,2
3 ^a EM – 2003	-28,8	-41,1	-12,3

FONTE: Elaboração própria com os dados do SAEB – INEP/MEC.

TABELA 32: Diferença entre as proficiências médias em Língua Portuguesa segundo a classificação de cor/raça dos alunos nos questionários do SAEB

Série - Ciclo	Pardo - Branco	Preto – Branco	Pardo - Preto
4 ^a EF –1995	-4,7	-17,8	-13,1
8 ^a EF –1995	-14,0	-25,2	-11,2
3 ^a EM – 1995	-16,6	-22,7	-6,1
4 ^a EF – 1997	-6,9	-21,8	-14,9
8 ^a EF – 1997	-14,2	-25,3	-11,1
3 ^a EM – 1997	-15,3	-28,5	-13,2
4 ^a EF – 1999	-7,3	-25,9	-18,6
8 ^a EF – 1999	-13,9	-21,0	-7,1
3 ^a EM – 1999	-18,9	-28,8	-9,9
4 ^a EF – 2001	-8,7	-29,4	-20,7
8 ^a EF – 2001	-15,4	-28,8	-13,4
3 ^a EM – 2001	-19,4	-31,2	-11,8
4 ^a EF – 2003	-9,9	-30,5	-20,6
8 ^a EF – 2003	-15,0	-29,3	-14,3
3 ^a EM – 2003	-19,8	-34,2	-14,4

FONTE: Elaboração própria com os dados do SAEB – INEP/MEC.

Quando se compara o resultado médio dos dois grupos com o dos brancos, são observados, sistematicamente, dois movimentos: um primeiro, em que fica evidente que a diferença entre pretos e brancos é sempre maior que a diferença entre pardos e brancos. E um segundo, que mostra um aumento da diferença entre esses grupos, com o aumento da escolarização, fato que não se concretiza na íntegra apenas para Leitura, nos três últimos ciclos, considerando-se a diferença entre pretos e brancos.

5.3 Diferenças na distribuição

Além do desempenho médio, é muito importante que se analise a distribuição desse desempenho na população, afinal a variação é intrínseca ao ser humano. Nascemos diferentes e temos o livre-arbítrio para tomar decisões diferentes, portanto é natural que os resultados finais sejam diferentes. O que não é aceitável, principalmente em tempos de uma educação inclusiva, é que a variação apresente

algum viés, seja socioeconômico, racial, de gênero etc. Por isso é importante que se pergunte se existe equidade racial no sistema educacional brasileiro. Equidade é um conceito de grupo e não individual, que não está associado à igualdade de resultados. Almejar a equidade em relação aos resultados de desempenho na educação brasileira é, antes de tudo, desejar que a variação nos subgrupos apenas reflita a variação total da população.

Os gráficos abaixo mostram essa variação para Matemática e para Leitura nas três séries do SAEB 2003, comparando-se a distribuição acumulada das proficiências na população total com a distribuição acumulada na população que se autodeclara como branco, pardo e preto. Nesse contexto, Soares (2006) reflete sobre a possibilidade de se utilizar a diferença entre essas curvas para se criar um indicador de desigualdade educacional.

Gráfico 33: Distribuição acumulada das proficiências – 4ª série EF - Matemática

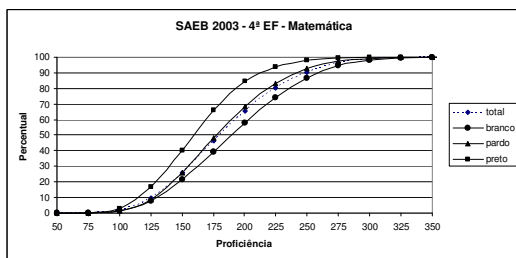


Gráfico 34: Distribuição acumulada das proficiências – 4ª série EF - Leitura

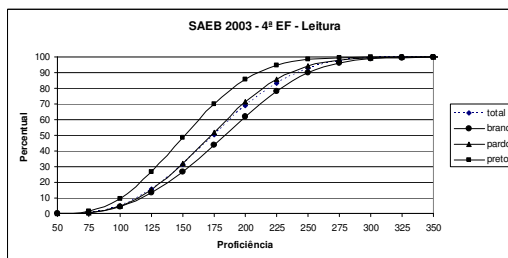


Gráfico 35: Distribuição acumulada das proficiências – 8ª série EF - Matemática

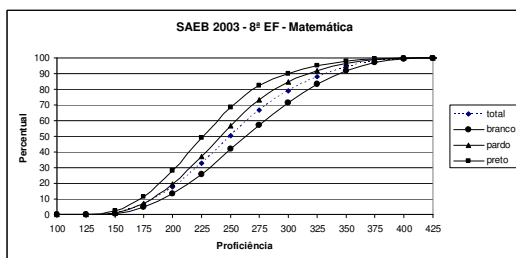


Gráfico 36: Distribuição acumulada das proficiências – 8ª série EF - Leitura

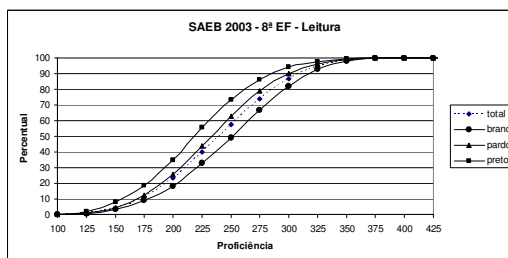


Gráfico 37: Distribuição acumulada das proficiências – 3ª série EM - Matemática

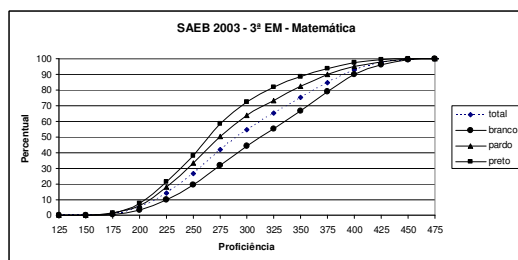
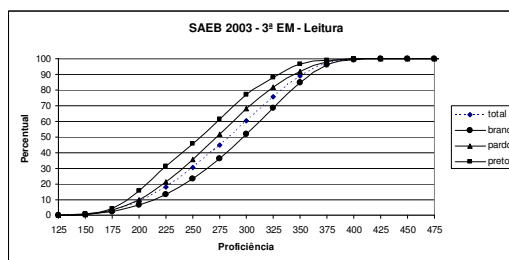


Gráfico 38: Distribuição acumulada das proficiências – 3ª série EM - Leitura



FONTE: Elaboração própria com os dados do SAEB 2003 – INEP/MEC.

Infelizmente, o que se pode constatar é que há uma diferença sistemática na distribuição das proficiências. Os pretos apresentam, sempre, a pior distribuição dos resultados, seguidos pelos pardos, pela curva da população total, e, por fim, apresentando a melhor distribuição, tem-se os alunos que se autodeclararam como brancos. Os mesmos resultados são obtidos nas demais bases, o que denota uma falta de equidade no sistema educacional brasileiro.

5.3 As diferenças depois do controle pelo nível socioeconômico

O modelo base, apresentado no item 2.4 (equação 5), foi ligeiramente modificado para permitir que o coeficiente, associado a cada variável, *PARDO*, β_3 , e *PRETO*, β_4 , variasse³³ entre as escolas. O modelo resultante permite a medida não só da diferença entre alunos de diferentes categorias de cor/raça como também pode medir a capacidade de as escolas mediar as diferenças observadas anteriormente e remanescentes, mesmo depois do controle pelo nível socioeconômico.

Sendo assim, a variante do modelo base, adotada neste caso é dada por:

³³ Uma decisão alternativa foi tomada pelo grupo de pesquisadores do GAME/UFMG (Grupo de Avaliação e Medidas Educacionais da Universidade Federal de Minas Gerais) em uma prestação de serviços de consultoria para a Prefeitura Municipal de Belo Horizonte. Naquele caso, tais pesquisadores optaram por permitir a variação apenas do coeficiente associado à variável “preto”. Isso só reflete o fato já abordado de que não existe uma solução única, ou melhor, não existe um modelo único para análise das variáveis que impactam o desempenho. Mas continua valendo a regra de que os resultados obtidos dependem do modelo escolhido. Foram produzidos relatórios técnicos deste trabalho para a Secretaria Municipal de Educação de Belo Horizonte, mas não há documentos públicos do mesmo.

$$\text{Nível 1: } Y_{ij} = \beta_{0j} + \beta_{1j} NSE_{ij} + \beta_{2j} SEXO_{ij} + \beta_{3j} PARDO_{ij} + \beta_{4j} PRETO_{ij} + \beta_{5j} ATRASO_{ij} + \beta_{6j} DISCIPLINA_{ij} + \beta_{7j} ATITUDE_{ij} + r_{ij} \quad (10)$$

$$\begin{aligned} \text{Nível 2: } \beta_{0j} &= \gamma_{00} + \gamma_{01} MNSE_j + \gamma_{02} REDE_j + \gamma_{03} MATRASO_j + u_{0j} \\ \beta_{3i} &= \gamma_{30} + u_{3j} \\ \beta_{4i} &= \gamma_{40} + u_{4j} \end{aligned}$$

Neste caso, vale destacar uma explicação mais substantiva do modelo em relação à questão racial. Os coeficientes γ_{30} e γ_{40} são os indicadores do comportamento do sistema educacional brasileiro em relação, respectivamente, às variáveis pardo e preto. O desejável é que ambos os coeficientes fossem zero, ou seja, o desempenho dos alunos não seria, tecnicamente, afetado pelo fato de ele ser pardo ou preto.

No entanto, no item 2.4.4 desta tese, foram apresentados alguns resultados para o ajuste do Modelo Base aos dados do SAEB. Naqueles resultados, entre outras coisas, percebe-se o quão perverso é o efeito do nível socioeconômico ou do atraso escolar sobre o desempenho dos alunos. Não menos perverso é o efeito sobre os alunos que se autotransferem como preto no sistema educacional brasileiro. As duas próximas tabelas foram geradas a partir do ajuste do modelo (10) às 28 bases do SAEB e representam o efeito das variáveis *PARDO* e *PRETO*, isto é, o valor dos coeficientes γ_{30} e γ_{40} .

TABELA 33 – Coeficiente γ_{30} ajustado para o modelo (10) – Variável associada: *PARDO*

Ano	4ª EF	8ª EF	3ª EM
1995	-	-1,21*	-1,61
1997	1,06*	-2,01**	-0,68
1999	0,02	-2,52**	-2,20*
2001	0,06*	-1,52**	-1,48**
2003	0,45	-0,58	-1,80**

(*) = $p < 0,05$ e (**) = $p < 0,001$

TABELA 34 – Coeficiente γ_{40} ajustado para o modelo (10) –
Variável associada: *PRETO*

Ano	4ª EF	8ª EF	3ª EM
1995	-	-7,78**	-4,22*
1997	-7,49**	-8,10**	-5,31*
1999	-9,32**	-5,32**	-6,96**
2001	-9,43**	-8,24**	-5,77**
2003	-10,15**	-8,90**	-6,44**

(*) = $p < 0,05$ e (**) = $p < 0,001$

Fica claro que ser preto no sistema educacional brasileiro implica em “já começar o jogo perdendo”. Ou, em uma linguagem mais técnica, implica que, em média, os equivalentes socioeconomicamente mais pretos têm quase dez pontos a menos de proficiência que um aluno branco nas mesmas condições. A diferença entre pardos e brancos já não é tão marcante assim. Calculando-se o efeito médio de ser pardo e ser preto, no sistema de educação básica brasileiro, percebe-se uma grande distinção entre esses dois grupos raciais, pois se obtém, respectivamente, -0,61 pontos e -8,34 pontos. Isso significa que enquanto os alunos pretos têm uma desvantagem de mais de oito pontos em relação aos alunos brancos, para os alunos pardos, essa desvantagem é pouco maior que meio ponto.

Apesar de os resultados, apresentados nas Tabelas 31 e 32, indicarem uma distância sistemática entre brancos e pardos, observa-se que essa separação praticamente some quando a análise é feita controlando-se o desempenho por variáveis sociológicas. A diferença entre esses dois grupos é socioeconômica e não racial, além disso, percebe-se que não é em todas as bases que o efeito da variável *PARDO* é significativo, enquanto que para a variável *PRETO*, o efeito é significativo no nível 0,001, em quase todas as bases. Portanto, independente de questões ideológicas, o fato é que os números não mostram uma associação entre pardos e pretos, a diferença marcante está entre brancos e pretos. Se se deseja buscar uma associação, no campo educacional, ela é muito mais forte entre pardos e brancos do que entre pardos e pretos.

O resultado anterior reflete uma realidade do sistema como um todo. Isso, entretanto, ainda não é motivo para se afirmar que não existe equidade racial no Brasil. Quando se analisa o mesmo resultado, agregado no nível da escola, percebe-se, nas diversas bases pesquisadas, que algumas delas conseguem, sim, atingir patamares desejáveis de equidade. Retomando o modelo multinível (10), apresentado

anteriormente, a equidade racial seria alcançada para aquelas escolas, cujo coeficiente de inclinação das variáveis “pardo” ou “preto” fosse nulo, ou seja, escolas cujo β_{3j} ou β_{4j} fossem iguais a zero, portanto dois outros coeficientes importantes do modelo são o u_{3j} e o u_{4j} . Eles estão associados a cada escola que compõe o grupo para o qual se ajustou esse modelo e sintetizam o esforço que cada escola deve fazer para anular a realidade do sistema, expressa pelos coeficientes γ_{30} e γ_{40} .

Os gráficos abaixo, que correlacionam o efeito da escola (u_{0j}) e a sua medida de equidade racial, em relação à variável “preto” (u_{4j}), categorizados pela dependência administrativa, mostram que escolas equitativas existem, mas, infelizmente, o sistema como um todo aponta que a solução para a equidade racial é encontrada à custa da diminuição do efeito da escola. Em relação à distinção entre público e privado, não se percebe, nesses gráficos, maior eficiência de uma ou de outra rede para se alcançar a equidade racial.

Gráfico 39: Efeito da escola por equidade racial – 8ª série EF – SAEB 1995
 Gráfico 40: Efeito da escola por equidade racial – 3ª série EM – SAEB 1995

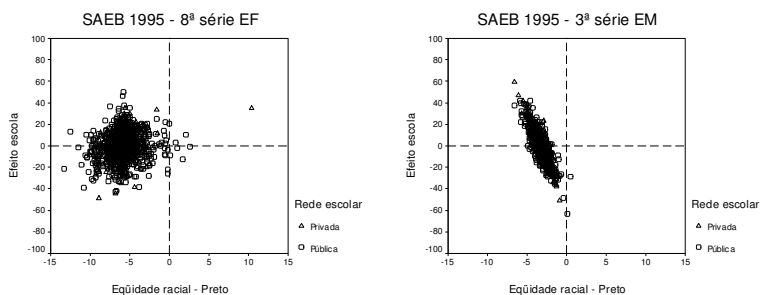


Gráfico 41: Efeito da escola por equidade racial – 4ª série EF – SAEB 1997
 Gráfico 42: Efeito da escola por equidade racial – 8ª série EF – SAEB 1997
 Gráfico 43: Efeito da escola por equidade racial – 3ª série EM – SAEB 1997

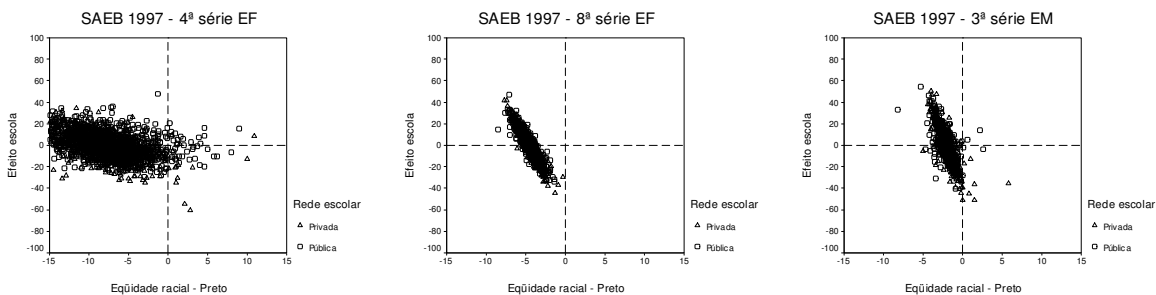


Gráfico 44: Efeito da escola por eqüidade racial – 4ª série EF – SAEB 1999
 Gráfico 45: Efeito da escola por eqüidade racial – 8ª série EF – SAEB 1999
 Gráfico 46: Efeito da escola por eqüidade racial – 3ª série EM – SAEB 1999

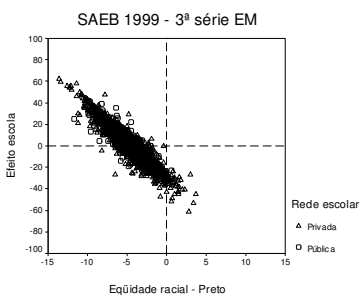
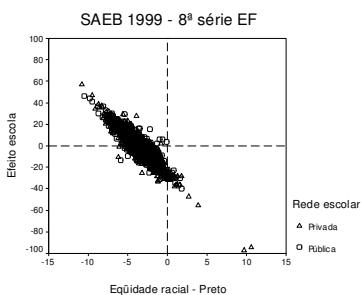
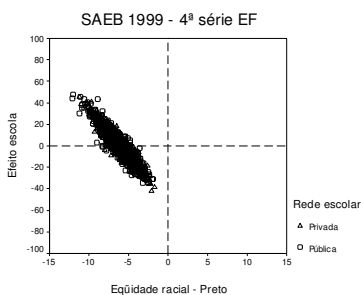


Gráfico 47: Efeito da escola por eqüidade racial – 4ª série EF – SAEB 2001
 Gráfico 48: Efeito da escola por eqüidade racial – 8ª série EF – SAEB 2001
 Gráfico 49: Efeito da escola por eqüidade racial – 3ª série EM – SAEB 2001

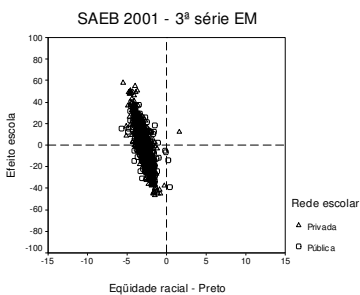
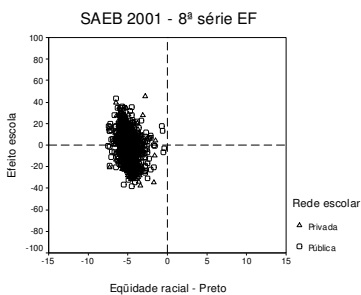
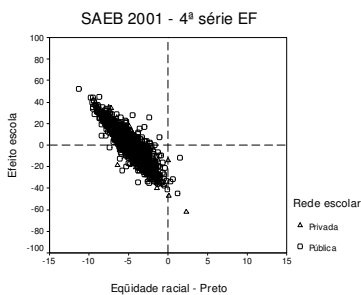
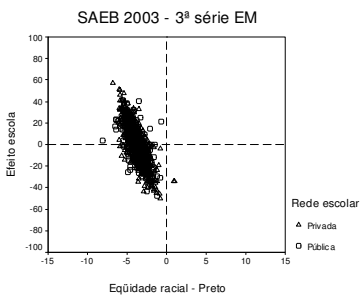
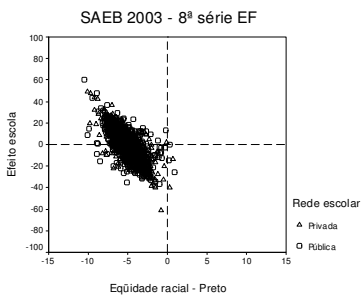
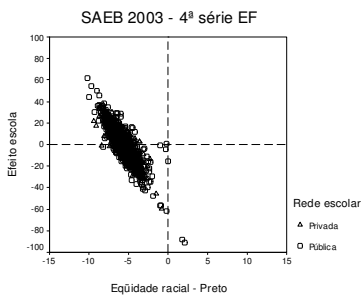


Gráfico 50: Efeito da escola por eqüidade racial – 4ª série EF – SAEB 2003
 Gráfico 51: Efeito da escola por eqüidade racial – 8ª série EF – SAEB 2003
 Gráfico 52: Efeito da escola por eqüidade racial – 3ª série EM – SAEB 2003



FONTE: Elaboração própria com os dados do SAEB – INEP/MEC.

No caso da variável “pardo”, as coisas são mais “suaves”, ou seja, apesar de permanecer a relação inversa entre efeito da escola e eqüidade racial, encontra-se um número muito maior de escolas com o coeficiente β_{3j} igual a zero, as quais já

conseguem promover uma equidade entre brancos e pardos, o que era esperado, tendo em vista que a diferença entre eles era muito menor.

5.4 Conclusões: Pretos e Pardos são Negros?

Como forma de conclusão deste capítulo, esse subtítulo propõe uma reflexão sobre os impactos, do ponto de vista dos resultados educacionais, da forma como são agregadas as categorias raciais, em especial, a junção de pretos e pardos em um grupo maior denominado de negros. Será essa a melhor forma de se agregar tais categoriais raciais? Ou cabe até mesmo perguntar: será que os resultados educacionais indicam que tal agregação deve ser feita? Existe, então, equidade racial no Brasil? Com a análise dos resultados dos cinco ciclos do SAEB, pôde-se construir uma resposta para essas questões.

Fica claro que as diferenças são grandes e reais, não só entre brancos e pardos, mas, principalmente, entre brancos e pretos. Ações devem ser pensadas e executadas para se diminuir essa separação, caso contrário, não se percebe uma tendência de alteração da curva de desempenho que separa esses grupos raciais. Entretanto, fica explícito que parte da diferença racial, na acepção que se dá ao termo negro, é socioeconômica; outra parte, entretanto, ultrapassa essa explicação. Este trabalho não pretende procurar explicações, mas apóia a tese de que temos uma questão racial a ser resolvida no Brasil, ainda que os termos de sua definição sejam diferentes do usual. Do ponto de vista dos resultados cognitivos em Matemática e Leitura, aferidos pelo SAEB, há uma distinção nítida entre preto e branco, e, também, entre preto e pardo.

Em momento algum está se contrariando a defesa a uma coexistência harmoniosa de diferenças. A busca pela equidade na educação passa por almejar a justiça entre os diversos subgrupos sociais. O fragmento a seguir retrata alguns dos ideais do Movimento Negro no Brasil, expressos por uma agência de notícias, que nasceu para levar a agenda social latino-americana e caribenha à mídia internacional, a Adital, e com os quais não há divergências ideológicas acentuadas.

O Movimento Negro é um movimento que está higienizando as relações entre negros e brancos na América Latina, entre negros e indígenas, entre brancos e indígenas, ou seja, essa higienização para a nova base epistemológica dentro de nossa sociedade, para compreender como construir uma sociedade juntos, como pode se

organizar a coexistência das diferenças, você não pode simplesmente, deixar as diferenças evoluírem de qualquer jeito, porque ela tem um acúmulo histórico que impede que ela se reproduza de maneira sã, ou seja, o racismo já está ali e você tem que lidar com ele, você tem que higienizar a base relacional da sociedade. (Adital, 2005).

Mas, se por um lado não há divergência ideológica em relação aos princípios de justiça anteriormente apresentados, por outro, é preciso registrar que, pelo menos do ponto de vista educacional, não se indica uma aglutinação de pardos e pretos em uma única categoria, denominada negros.

Para finalizar, propõe-se, como resposta à pergunta se existe eqüidade racial no Brasil, um 'sim com restrições'. Os resultados são claros e mostram que várias são as escolas que conseguem diminuir as diferenças pré-existentes entre brancos e pretos, no que se refere ao desempenho cognitivo em Matemática e Leitura. Essas escolas precisam ser identificadas, conhecidas, estudadas e suas práticas difundidas como um bom exemplo a se seguir. Registra-se aqui o trabalho de Soares, R. M. (2006), que acompanhou três escolas de educação básica da rede pública de Minas Gerais, as quais foram escolhidas, justamente por apresentarem destaque, positivo e negativo, nos índices de eqüidade racial. Metodologias qualitativas foram utilizadas para se identificar quais as práticas e/ou políticas internas à escola garantiriam

atenuar as desigualdades raciais, externas à escola, [tornando-se uma escola] cujos alunos negros detêm uma auto-imagem mais positiva, mais afirmativa, referendada numa dignidade apta para assumir a própria etnicidade, seja incorporando, com orgulho, o termo "negro", seja detendo mais maturidade e consciência crítica para entender a sua própria história étnico-racial, caminho, certamente, propício para a construção de aprendizagens libertárias e inter-relacionais (SOARES, R. M., 2006, p. 138).

No entanto, as restrições foram colocadas na avaliação da eqüidade racial brasileira como um alerta, já que não é desejável que se continue buscando a eqüidade à custa da diminuição do efeito da escola. É necessário que ambos caminhem juntos, aumentando a qualidade e a eqüidade do sistema de ensino brasileiro.

6 – Referências

ADITAL. **Movimento negro e educação**. Adital - Notícias da América Latina e Caribe. 2005. Disponível em: <<http://www.adital.com.br/site/noticia.asp?lang=PT&cod=15518>>. Acesso em: 15 Abr 2008.

AKKARI, A. J.. Desigualdades educativas estruturais no Brasil: entre estado, privatização e descentralização. **Educação & Sociedade**. Campinas, v. 22, n. 74, 2001. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0101-73302001000100010&lng=pt&nrm=iso>. Acesso em: 26 Nov 2007.

ALBERNAZ, A.; FRANCO, C.; FERREIRA, F. H. G. Qualidade e Equidade no Ensino Fundamental Brasileiro. **Pesquisa e Planejamento Econômico**. Rio de Janeiro: IPEA, v. 32, n. 3, p. 453-476, 2002.

ALVES, M. T. G. **Efeito-escola e fatores associados ao progresso acadêmico dos alunos entre o início da 5ª série e o fim da 6ª série do ensino fundamental: um estudo longitudinal em escolas públicas no município de Belo Horizonte – MG**. 2006. Tese de doutorado. Universidade Federal de Minas Gerais, Belo Horizonte, 2006.

ALVES, M. T. G.; SOARES, J. F. **Medidas de nível socioeconômico em pesquisas sociais: uma aplicação aos dados GERES**. In: I Seminário GERES, 2008, Belo Horizonte.

ANDRADE, D. F.; TAVARES, H. R.; VALLE, R. C. **Teoria da Resposta ao Item: Conceitos e aplicações**. São Paulo: Associação Brasileira de Estatística, 2000.

ANDRADE, D. F. Comparando desempenhos de grupos de alunos por intermédio da teoria de resposta ao item. **Estudos em Avaliação Educacional**. São Paulo: Fundação Carlos Chagas, n.23, p. 31-69, jan./jun. 2001.

ANDRADE, J. M. de; LAROS, J. A.. Fatores associados ao desempenho escolar: estudo multinível com dados do SAEB/2001. **Psic.: Teor. e Pesq.**, Brasília, v. 23, n. 1, 2007. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-37722007000100005&lng=pt&nrm=iso>. Acesso em: 14 Nov 2007.

ANDRADE, M.; FRANCO, C.; CARVALHO, J. B. P. Gênero e desempenho em Matemática ao final do ensino médio: quais as relações. **Estudos em Avaliação Educacional**, v. 27, n. 27, p. 77-95, 2003.

ARREGUI, P. **Sistemas de determinación y evaluación de metas de logros de aprendizaje escolar como instrumentos para mejorar la calidad, la equidad y la**

responsibilización en los procesos educativos en América Latina. Seminario sobre Perspectivas de la Educación en América Latina y el Caribe, Chile, 23 al 25 de agosto de 2000. Disponible en <<http://www.unesco.cl/promedl7/prospectivas/arregui.pdf>>

BAKER, F. B. **The basics of Item Response Theory.** USA: ERIC Clearinghouse on Assessment and Evaluation, 2001. p. 172

BARBOSA, M. E. F.; FERNANDES, C. A escola brasileira faz diferença? Uma investigação dos efeitos da escola na proficiência em matemática dos alunos da 4ª série. In: FRANCO, C. (org.) **Avaliação, ciclos e promoção na educação.** Porto Alegre: Artmed Editora, 2001.

BARBOSA, M. L. de O. A qualidade da escola e as desigualdades raciais no Brasil. In: SOARES, S. *et al.* (org.) **Os mecanismos de discriminação racial nas escolas brasileiras.** Rio de Janeiro: IPEA, 2005.

BIRNBAUM, A. Some latent trait models and their use in inferring an examinee's ability. In: LORD, F. M.; NOVICK, M. R. **Statistical theories of mental test scores.** Reading, MA: Addison-Wesley, 1968.

BOCK, R. D.; ZIMOWSKI, M. F. Multiple group IRT. In: **Handbook of modern Item Response Theory.** VAN DER LINDER, W. J.; HAMBLETTON, R. K. New York: Springer-Verlag, 1997.

BONAMINO, A. **Tempos de avaliação educacional: o SAEB, seus agentes, referências e tendências.** Rio de Janeiro, Quartet, 2002. 192 p.

BONAMINO, A.; FRANCO, C. Avaliação e política educacional: o processo de institucionalização do SAEB. **Cadernos de Pesquisa**, nº 108, p. 101-132, nov./1999.

BRASIL. **Lei nº 9.394, de 20 de dezembro de 1996.** Lei de Diretrizes e Bases da Educação Nacional: publicada no Diário Oficial da União, em 23 de dezembro de 1996, Seção 1, p. 27.839.

_____. **Constituição da República Federativa do Brasil: promulgada em 05 de outubro de 1988 e publicada no Diário Oficial da União, nº 191-A, 1988.** Obra coletiva da Editora Saraiva com a colaboração de Antônio Luiz de Toledo Pinto e Márcia Cristina Vaz dos Santos Windt. 22. ed. atual. e ampl. São Paulo: Saraiva, 1999. 273 p.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **SAEB 97: Primeiros resultados.** Brasília: INEP/MEC, [1997?].

_____. **SAEB/95 – Relatório Final.** Brasília: INEP/MEC, 1998.

_____. **Relatório SAEB-99.** Brasília: INEP/MEC. Disponível em <<http://www.inep.gov.br>>. Acesso em nov. 2000.

_____. **SAEB 2001: Novas perspectivas.** Brasília: INEP/MEC, 2001.

_____. **SAEB 2001: Relatório Nacional.** Brasília: INEP/MEC, 2002a.

_____. **SAEB 2001: Relatório Saeb 2001: Língua Portuguesa.** Brasília: INEP/MEC, 2002b.

_____. **Qualidade da Educação: uma nova leitura do desempenho dos estudantes da 4ª série do Ensino Fundamental.** Brasília: INEP/MEC, 2003a.

_____. **Qualidade da Educação: uma nova leitura do desempenho dos estudantes da 8ª série do Ensino Fundamental.** Brasília: INEP/MEC, 2003b.

_____. **Qualidade da Educação: uma nova leitura do desempenho dos estudantes da 3ª série do Ensino Médio.** Brasília: INEP/MEC, 2003c.

_____. **SAEB.** Brasília: INEP/MEC. Disponível em <<http://www.inep.gov.br/basica/saeb>>. Acesso em 25 jun. 2005.

_____. **Microdados do Sistema Nacional de Avaliação da Educação Básica 1995, 1997, 1999, 2001 e 2003.** Brasília: Diretoria de Tratamento e Disseminação de Informações Educacionais/INEP/MEC, 2006. 5 CD-Roms.

_____. **Censo Escolar – Sinopses Estatísticas da Educação Básica 1991 - 2006.** Brasília: INEP/MEC. Disponível em <<http://www.inep.gov.br/basica/censo/Escolar/Sinopse/sinopse.asp>> Acesso em: mar. 2007.

BRASIL. Ministério da Educação. Unicef. **Aprova Brasil – O direito de aprender: Boas práticas em escolas públicas avaliadas pela Prova Brasil.** Brasília: MEC, 2006. Disponível em <http://www.inep.gov.br/download/imprensa/2006/aprova_brasil.pdf> . Acesso em: jul. 2008.

BRESSOUX, P. Les recherches sur les effets-écoles et les effets-maîtres. **Revue Française de Pédagogie**, 108, p. 91-137, 1994.

BROOKE, N.; SOARES, J. F. **Pesquisa em eficácia escolar: origem e trajetórias.** Tradução: Viamundi Idiomas e Traduções; Cleusa Aguiar Brooke; Rômulo Monte-Alto. Belo Horizonte: Editora UFMG, 2008. 552 p.

BRYK, A. S.; LEE, V. E.; HOLLAND, P. B. **Catholic schools and the common good.** Harvard University Press, 1993.

CARVALHO, M. Quem é negro, quem é branco: desempenho escolar e classificação racial de alunos. **Revista Brasileira de Educação**, n.28, p. 77-95, jan./abr., 2005.

CASTRO, C. de M. Ensino privado ou público: eis a (falsa) questão. **Ensaio - Avaliação e Políticas Públicas em Educação.** Rio de Janeiro: Cesgranrio, v.5, n.17, p.423-452, out./dez. 1997.

CAVALLI-SFORZA, L. L. **Genes, povos e línguas.** São Paulo: Companhia das Letras, 2003.

COLEMAN, J. S. *et al.* **Equality of Educational Opportunity.** Washington: Office of Education and Welfare, 1966.

COLEMAN, J. S.; HOFFER, T. **Public and private high schools: The impact of communities**. Basic Books, 1987.

CURY, C. R. J. O público e o privado na educação brasileira contemporânea: posições e tendências. **Cadernos de Pesquisa**. São Paulo, n. 81, p. 33-44, 1992.

DUBET, F. O que é uma escola justa? **Cadernos de Pesquisa**, v. 34, n. 123, p. 539-555, set./dez. 2004.

FERRÃO, M. E. **Introdução aos modelos de Regressão Multinível em Educação**. Campinas: Editora Komedi, 2003. 106 p.

FLETCHER, P. R. A Teoria da Resposta ao Item: medidas invariantes do desempenho escolar. **Ensaio: Avaliação e Políticas Públicas em Educação**. Rio de Janeiro: Cesgranrio, 1(2), p. 21 – 28, jan./mar. 1994.

_____. A demografia do desenvolvimento da educação no Brasil. **Revista Brasileira de Política e Administração da Educação**. Rio de Janeiro, v. 21, n. 1/2 , p. 32-60. jan./dez. 2005.

FORQUIN, J. C. Sociologia das desigualdades de acesso à educação: principais orientações, principais resultados desde 1965. In: FORQUIN, J. C. (org.) **Sociologia da educação: dez anos de pesquisas**. Petrópolis: Vozes, 1995.

FRANCO, C. O SAEB: potencialidades, problemas e desafios. **Revista Brasileira de Educação**, n.17, p. 127-132, maio/ago., 2001.

_____. Quais as contribuições da avaliação para as políticas educacionais?. In: BONAMINO, A., BESSA, N., FRANCO, C. (org.) **Avaliação da educação básica – pesquisa e gestão**. Rio de Janeiro: Ed. PUC-Rio; São Paulo: Loyola, 2004.

FRANCO, C.; BONAMINO, A. Iniciativas recentes de avaliação da qualidade da educação no Brasil. In: FRANCO, C. (org.) **Avaliação, ciclos e promoção na educação**. Porto Alegre: Artmed Editora, 2001.

FUKUDA, C. C. **O ensino eficaz na educação básica: um modelo descritivo dos fatores de eficácia**. 2003. Tese de doutorado. Universidade de Brasília, Brasília, 2003.

GANZEBOOM, H. B. G; TREIMAN, D. J. Internationally comparable measures of occupational status for the 1988 International Standard Classification of Occupation. **Social Science Research**. n.25, p. 201-239, 1996.

GOLDSTEIN, H. Modelos da realidade: Novas abordagens para a compreensão de processos educacionais. In: FRANCO, C. (org.) **Avaliação, ciclos e promoção na educação**. Porto Alegre: Artmed Editora, 2001.

GOLDSTEIN, H.; THOMAS, S. Using examination results as indicators of school and college performance. **Journal of the Royal Statistical Society, A**, v. 159, n. 3, p. 149-163, 1996.

GOLDSTEIN, H.; SPIEGELHALTER, D. J. League tables and their limitations: Statistical issues in comparisons of institutional performance. **Journal of the Royal Statistical Society**, A. v. 159, n. 3, p. 385-443, 1996.

GREMAUD, A. P.; FELÍCIO, F.; Biondi, R.L. **Indicador de Efeito Escola: Uma Metodologia para a identificação dos sucessos escolares a partir dos dados da Prova Brasil**. In: III Congresso Internacional em Avaliação Educacional, 2006, Fortaleza. Anais do III Congresso Internacional em Avaliação Educacional, 2006.

GUIMARÃES, A. S. A. Raça e os estudos de relações raciais no Brasil. **Novos Estudos CEBRAP**. São Paulo, n.54, jul., p. 147-156, 1999.

HADJI, C. **A avaliação, regras do jogo: Das intenções aos instrumentos**. Tradução de Júlia L. Ferreira e José Manuel Cláudio. Porto: Porto Editora, 1994. 190 p. Título original: L'évaluation, règles du jeu.

HADJI, C. **A avaliação desmistificada**. Tradução de Patrícia C. Ramos. Porto Alegre: ARTMED Editora, 2001. 136 p. Título original: L'évaluation démystifiée.

HAMBLETON, R. K.; SWAMINATHAN, H.; ROGERS, H. J. **Fundamentals of Item Response Theory**. Newbury Park, Califórnia: Sage Publications Inc., 1991.

HAMBLETON, R. K. Principles and selected applications of Item Response Theory. In: LINN, R. L. (Ed.) **Educational measurement**. 3.ed. Washington, DC: American Council on Education and the National Council on Measurement in Education, p. 147-200, 1993.

HASENBALG, C. A. **Discriminação e desigualdades raciais no Brasil**. Rio de Janeiro: Edições Graal, 1979.

JESUS, G. R. de. **Fatores que afetam o desempenho em português: um estudo multinível com dados do SAEB 2001**. 2004. Dissertação de mestrado. Universidade de Brasília, Brasília, 2004.

KAMEL, A. **Não somos racistas: uma reação aos que querem nos transformar numa nação bicolor**. Rio de Janeiro: Nova Fronteira, 2006. 143 p.

KLEIN, R. . Utilização da Teoria de Resposta ao Item no Sistema Nacional de Avaliação da Educação Básica. **Ensaio: Avaliação e Políticas Públicas em Educação**. Rio de Janeiro: Cesgranrio, v. 11, n. 40, p. 283-296, 2003.

_____. Como está a educação no Brasil? O que fazer? **Ensaio: Avaliação e Políticas Públicas em Educação**. Rio de Janeiro: Cesgranrio, v.14, n.51, p. 139-172, abr./jun. 2006.

KLEIN, R.; FONTANIVE, N. S. Avaliação em Larga escala: uma proposta inovadora. **Em aberto**, Brasília, ano 15, n. 66, p. 29-35, abr./jun. 1995.

LEE, V. E. Using hierarchical linear modeling to study social contexts: the case of school effects. **Educational Psychologist**, 35, p. 125-141, 2000.

_____. What are multilevel questions, and how might we explore them with quantitative methods. **Estudos em Avaliação Educacional**, São Paulo: Fundação Carlos Chagas, n. 24, p. 31-45, 2001a.

_____. Using multilevel methods to investigate research questions that involve nested data: examples from education. **Estudos em Avaliação Educacional**, São Paulo: Fundação Carlos Chagas, n. 24, p. 47-68, 2001b.

LINDEMAN, R. H. **Medidas educacionais - Testes objetivos e outros instrumentos de medida para a avaliação da aprendizagem**. Tradução de Leonel Vallandro. Porto Alegre: Editora Globo, 1976. 176 p.

LOCATELLI, I. **O sistema de avaliação da educação brasileira**. In: Simposio Iberoamericano de Investigación Sobre Eficacia Escolar y Mejora de la Escuela, 1., 2002, Havana.

LORD, F. M. **A theory of test scores**. Psychometric Monograph. Iowa: Psychometric Society, 7, 1952.

MALUF, M. M. B. Sistema Nacional de Avaliação da Educação Básica no Brasil: análise e proposições. **Estudos em Avaliação Educacional**, São Paulo: Fundação Carlos Chagas, 14, p. 5-38, 1996.

MATO GROSSO DO SUL. Secretaria de Estado de Educação de Mato Grosso do Sul. **Boletim Pedagógico: Sistema de Avaliação da Estadual da Educação Básica – Língua Portuguesa e Matemática**. Juiz de Fora: CAEd, 2003.

MEDEIROS, A. **A termometria: de Galileu a Fahrenheit**. Recife: Editora Liber, 1999. 78 p.

MENEZES FILHO, N.; SOARES, J. F.; TERRA, R. **Definição de metas de desempenho para as escolas estaduais paulistas de ensino básico**. 2008. Manuscrito.

MESSICK, S. Validity. In: LINN, R. L. (Ed.) **Educational measurement**. 3.ed. Washington, DC: American Council on Education and the National Council on Measurement in Education, p. 13-103, 1993.

MINAS GERAIS. Secretaria de Estado de Educação de Minas Gerais. **Boletim Pedagógico: PROEB 2002 – Língua Portuguesa**. Juiz de Fora: CAEd, 2002.

_____. **Boletim Pedagógico: Resultados do SIMAVE/PROEB 2003 – Matemática**. Juiz de Fora: CAEd, 2003.

_____. **Atlas da Educação de Minas Gerais**. Belo Horizonte: Fundação João Pinheiro. 2005.

MISLEVY, R. J.; BOCK, R. D. **Bilog: Item analysis and test scoring with binary logistic models**. 2. ed. USA: Scientific Software, Inc., 1990.

MORETTO, V. P. **Prova: um momento privilegiado de estudo - não um acerto de contas**. Rio de Janeiro: DP&A Editora, 2001. 150 p.

MORTIMORE, P. *et al.* **School matters**. California/USA: University of California Press, 1988.

MOSTELLER, F.; MOYNIHAN, D. P. *et al.* **On equality of educational opportunity**. New York: Random House, 1972.

NATIS, L. Modelos lineares hierárquicos. **Estudos em Avaliação Educacional**, São Paulo: Fundação Carlos Chagas, n. 23, p. 3-29, 2001.

NOGUEIRA, M. A. A sociologia da educação do final dos anos 60/ início dos anos 70: o nascimento do paradigma da reprodução. **Em aberto**, 9 (46), p. 49-59, 1990.

NOGUEIRA, M. A.; CATANI, A. (org.) **Escritos de educação**. 2ed. Petrópolis: Vozes, 1999. 251 p.

OLIVEIRA, J. B. A.; CHADWICK, C. **Aprender e ensinar**. 7ed. Belo Horizonte: Alfa Educativa, 2006. 394 p.

OPDENAKKER, M-C.; DAMME, J. V. Differences between secondary schools: a study about school context, group composition, school practice and school effects with special attention to public and Catholic schools and types of schools. **School Effectiveness and School Improvement**. v. 17, n. 1, p. 87-117, 2006.

OSORIO, R. G. **O sistema classificatório de “cor ou raça” do IBGE**. Texto para discussão n.996. Brasília: Instituto de Pesquisa Econômica Aplicada (IPEA), 2003. 50 p.

OSORIO, R. G.; SOARES, S. A geração 80: um documentário estatístico sobre a produção das diferenças educacionais entre negros e brancos. In: SOARES, S. *et al.* (org.) **Os mecanismos de discriminação racial nas escolas brasileiras**. Rio de Janeiro: IPEA, 2005.

PENA, S. D. J.; CARVALHO-SILVA, D. R.; ALVES-SILVA, J.; PRADO, V. F.; SANTOS, F. R. O retrato molecular do Brasil. **Ciência Hoje**. Rio de Janeiro: SBPC, p. 17-25. Abril, 2000.

PERRENOUD, P. **Avaliação: Da excelência à regulação das aprendizagens - entre duas lógicas**. Tradução de Patrícia C. Ramos. Porto Alegre: Editora Artes Médicas, 1999. 183 p. Título original: L'évaluation des élèves: de la fabrication de l'excellence à la regulation des apprentissages.

PESTANA, M. I. O Sistema de avaliação brasileiro. **R. Bras. Est. Pedag.**, Brasília, 79 (191), p. 65 - 73, jan./abr. 1998.

RASCH, G. **Probabilistic models for some intelligence and attainment tests**. Copenhagen: Danish Institute for Educational Research, 1960.

RAUDENBUSH, S. W.; BRYK, A. S. **Hierarchical linear models**. Thousand Oaks, CA: Sage, 2002.

RAUDENBUSH, S. W.; WILLMS, J. D. The estimation of school effects. **Journal of Educational and Behavioral Statistics**. v. 20, n. 4, p. 307-335, 1995.

RUTTER, M. *et al.* **Fifteen thousand hours: secondary schools and their effects on children.** Cambridge: Harvard University Press, 1979.

SAMEJIMA, F. A. Estimation of latent ability using a response pattern of graded scores. **Psychometrika**, New York, v.17, 1969. Monograph Supplement, n.4, Part 2.

SAMMONS, P. *et al.* Continuity of school effects: A longitudinal analysis of Primary and Secondary school effects on GCEE performance. *School Effectiveness and School Improvement*, v.6, n.4, p. 285-307, 1995.

SILVA, C. S. **Medidas e avaliação em educação.** Petrópolis: Editora Vozes, 1992. 207 p.

SINGER, J. D. Using SAS PROC MIXED to fit multilevel models, hierarchical models and individual growth models. **Journal of Educational and Behavioral Statistics**, v.24, n. 4, p. 323-355. 1998.

SOARES, J. F. O efeito da escola no desempenho cognitivo de seus alunos. **REICE - Revista Eletrônica Iberoamericana sobre Calidad, Eficácia y Cambio en Educación.** 2 (2), p. 83-104, 2004. Disponível em <<http://www.ice.deusto.es/rinace/reice/vol2n2/Soares.pdf>> Acesso em: mar. 2005.

SOARES, J. F. Measuring cognitive achievement gaps and inequalities: The case of Brazil. **International Journal of Education Research.** v.45, n.3, p. 176-187. 2006.

SOARES, J. F. **Avaliação da qualidade da educação escolar brasileira.** 2008. Manuscrito.

SOARES, J. F.; ANDRADE, R. J. **Escala de posicionamento socioeconômico.** In: VII Encontro de Pesquisa em Educação da Região Sudeste, 2005, Belo Horizonte.

_____. Nível socioeconômico, qualidade e equidade das escolas de Belo Horizonte. **Ensaio - Avaliação e Políticas Públicas em Educação.** Rio de Janeiro: Cesgranrio, v.14, n.50, p.107 - 126, jan./mar. 2006.

SOARES, J. F.; CANDIAN, J. O efeito da escola básica brasileira: as evidências do Pisa e do Saeb. **Revista Contemporânea de Educação,** Rio de Janeiro: UFRJ, n.4, 2007.

SOARES, J. F.; ALVES, M. T. G.; OLIVEIRA, R. M. O efeito de 248 escolas de nível médio no vestibular da UFMG nos anos de 1998, 1999 e 2000. **Estudos em Avaliação Educacional,** São Paulo: Fundação Carlos Chagas, n. 24, p. 69-123, 2001.

SOARES, J. F.; CÉSAR, C. C.; MAMBRINI, J. Determinantes de desempenho dos alunos do ensino básico brasileiro: evidências do SAEB de 1997. In: FRANCO, C. (org.) **Avaliação, ciclos e promoção na educação.** Porto Alegre: Artmed Editora, 2001.

SOARES, J. F.; ALVES, M. T. G. Desigualdades raciais no sistema brasileiro de educação básica. **Educação e Pesquisa,** São Paulo, v. 29, n. 1, p. 147-165, jan./jun. 2003

SOARES, J. F.; ALVES, M. T. G.; MARI, F. Avaliação de escolas de ensino básico. In: FRANCO, C. (org.) **Promoção, ciclos e avaliação educacional**. Porto Alegre: ArtMed Editora, 2003.

SOARES, J. F.; MAMBRINI, J.; PEREIRA, L. R.; ANDRADE, R. J. **Fatores associados ao desempenho em Língua Portuguesa e Matemática: a evidência do SAEB 2003**. Relatório técnico. Universidade Federal de Minas Gerais: Belo Horizonte, 2004.

SOARES, R. M. **Classificação racial e desempenho escolar**. 2006. Dissertação de mestrado. Universidade Federal de Minas Gerais, Belo Horizonte, 2006.

SOARES, S.; BELTRÃO, K. I.; BARBOSA, M. L. de O.; FERRÃO, M. E. (org.) **Os mecanismos de discriminação racial nas escolas brasileiras**. Rio de Janeiro: IPEA, 2005. p. 202.

SOARES, T. M.; PEREIRA, D. R. M. **Estudo de critérios de ajuste para modelos da teoria da resposta ao item (TRI)**. In: Simpósio Brasileiro de Pesquisa Operacional, 33., 2001, Campos do Jordão. p. 403 – 413.

SOBRINHO, J. D. Campo e caminhos da avaliação: a avaliação da educação superior no Brasil. In: FREITAS, L. C. (org.) **Avaliação: construindo o campo e a crítica**. Florianópolis: Editora Insular, 2002. p. 13 - 38.

TEDDLIE, C.; REYNOLDS, D. **The international handbook of school effectiveness research**. Nova York: Falmer Press, 2000. 411p.

TEDESCO, J. C. Alguns aspectos da privatização da educação na América Latina. **Estudos Avançados**. São Paulo, v.5, n.12, p.23-44. 1991.

THISSEN, D. **Multilog: multiple, categorical item analysis and test scoring: Using Item Response Theory**. Chicago: Scientific Software, Inc., 1991.

TYLER, R. W. **Princípios básicos de currículo e ensino**. Porto Alegre: Editora Globo, 1981 apud SILVA, C. S. Medidas e avaliação em educação. Petrópolis: Editora Vozes, 1992. 207 p.

VALLE, R. C. Teoria de resposta ao item. **Estudos em Avaliação Educacional**, São Paulo, n. 21, p. 7-91, 2000.

_____. A construção e a interpretação das escalas de conhecimento – Considerações gerais e uma visão do que vem sendo feito no SARESP. **Estudos em Avaliação Educacional**, São Paulo: Fundação Carlos Chagas. n. 23, p. 71-92, 2001.

VIANNA, H. M. Avaliação de sistemas: Implementação de políticas públicas. **Estudos em Avaliação Educacional**, São Paulo: Fundação Carlos Chagas. n. 22, p. 119-134, 2000.

WILLMS, J. D. **Monitoring School Performance: a guide for educators**. London: The Falmer Press, 1992.

WILLMS, J. D.; RAUDENBUSH, S. W. A longitudinal hierarchical linear model for estimating school effects and their stability. **Journal of Educational Measurement**, 26, p. 209-232, 1989.

WILLMS, J. D.; SOMERS, M. A. Family, classroom and school effects on children's educational outcomes in Latin America. **School Effectiveness and School Improvement**, 12(4), p. 409-445, 2000.

ZIBAS, D. M. L. Escola pública versus escola privada: o fim da história? **Cadernos de Pesquisa**. São Paulo, n. 100, p. 57-77, 1997.

ZIMOWSKI, M. F.; MURAKI, E.; MISLEVY, R. J. & BOCK, R. D. BILOG-MG: **Multiple-Group IRT analysis and test maintenance for binary items**. Chicago: Scientific Software, Inc, 1996.

7 – Anexos

7.1 Metodologia e síntese do plano amostral de cada um dos ciclos pesquisados do SAEB

- SAEB 1995

A edição do SAEB/95 foi realizada pelo Inep/MEC com a participação das Secretarias Estaduais e Municipais de Educação, entre os dias 6 a 10 de novembro.

Amostra – Considerando as perdas de campo, a amostra efetiva ficou constituída de 90.499 alunos matriculados nas 4ª e 8ª séries do Ensino Fundamental e 2ª e 3ª séries do Ensino Médio, freqüentando 3.453 turmas de 2.839 escolas públicas e particulares do País. Obteve também informações de 6.292 professores de Matemática e de Língua Portuguesa e de 2.718 diretores das escolas pesquisadas. Particularmente em relação à 2ª série do Ensino Médio, devido a perdas não compensadas na ponderação, os dados foram excluídos da análise.

As 27 unidades da Federação (UF) e 9 regiões metropolitanas (RM) foram tratadas como estratos independentes. A amostra foi planejada para produzir estimativas agregadas e estimativas de várias subpopulações em cada série, com precisões aproximadamente iguais para cada âmbito do levantamento.

Testes - Para a avaliação das proficiências dos alunos brasileiros foram aplicados 1.236 itens de teste de Matemática e de Leitura. A maioria dos itens foi de múltipla escolha. Somente 33 dos itens aplicados aos alunos da 8ª série e 3ª série do Ensino Médio solicitaram que o aluno construísse e escrevesse sua resposta. Os itens de teste foram reunidos em 13 blocos e em 26 cadernos por série e por disciplina, de modo que foi confeccionado um total de 208 cadernos distintos. Os cadernos de teste foram compostos seguindo um planejamento de blocos balanceados incompletos, sendo que cada bloco aparece o mesmo número de vezes em cada posição no caderno e cada par de blocos aparece somente uma vez nos cadernos. Um caderno foi composto de três blocos contendo entre 10 a 13 questões cada um, portanto cada aluno respondeu cerca de 30 a 39 questões. Com a finalidade de construir as escalas de proficiências comuns a todas as séries avaliadas, foram aplicadas questões comuns às quatro séries: 10 questões de leitura e 10 de matemática comuns a 4ª, 8ª e 3ª série do Ensino Médio; 26 de leitura e 22 de matemática comuns somente a 8ª série e 3ª série do Ensino Médio; 30 de matemática e 30 de leitura comuns somente a 4ª e 8ª série; 13 de leitura e 13 de matemática comuns a 8ª e 3ª série do Ensino Médio [sic]. A correção dos itens foi procedida por bancas de especialistas, que estabeleceram os padrões de resposta esperados, os critérios de correção e atribuição de grau.

Questionários – Foram utilizados instrumentos para coletar dados sobre as características de infra-estrutura e de disponibilidades da unidade escolar, o perfil do diretor e os mecanismos de gestão escolar, o perfil do professor, sua prática pedagógica, as características socioculturais e os hábitos de estudo do aluno.

Análise - O SAEB/95 utilizou a Teoria de Reposta ao Item para realizar análises para cálculo das proficiências. Os modelos utilizados foram:

- da logística de 3 parâmetros – para itens de múltipla escolha
- da logística de 2 parâmetros, com c igual a 0 - para itens de resposta construída corrigidos como certo ou errado.

- *de crédito parcial generalizado - para itens de resposta construída corrigidos com graduação de 0 a 2 ou de 0 a 3. Os parâmetros dos itens foram estimados conjuntamente para todos os itens – múltipla escolha e resposta construída. (Brasil/MEC/INEP, 2006, p. 10-11)*

- **SAEB 1997**

A edição do SAEB/97 foi realizada pelo Inep/MEC com a participação das Secretarias Estaduais e Municipais de Educação, entre os dias 27 e 31 de outubro.

Amostra – *O universo considerado nesta pesquisa consistiu de todos os alunos matriculados na 4ª e na 8ª séries do Ensino Fundamental e 3ª série do Ensino Médio em todo o Brasil. Duas exclusões foram feitas do marco da amostragem: as escolas federais, por serem em número muito reduzido, e as escolas rurais da região Norte, pela dificuldade de acesso. Também foram excluídas escolas profissionalizantes e turmas multisseriadas. A amostra prevista para o SAEB/97, contemplava 220.000 alunos distribuídos nas três séries anteriormente citadas, cobrindo todas as Unidades da Federação (UF). A distribuição desses 220.000 alunos entre as séries considerou tanto a proporção de matrículas em cada série como o número de disciplinas a serem avaliadas em cada uma das séries (3 na 4ª, 3 na 8ª e 5 na 3ª). Assim, a amostra foi formada por 93.000 alunos da 4ª série, 57.000 alunos de 8ª série do Ensino Fundamental e 70.000 alunos de 3ª série do Ensino Médio, distribuídos igualmente entre as 27 Unidades da Federação. O plano de amostragem previa [sic] três etapas: seleção de municípios, de escolas e de turmas, nas três séries avaliadas. Considerando as diferenças de desempenho observadas no SAEB/95, referentes à zona urbana e rural, às escolas estaduais, municipais e particulares, a amostra do SAEB/97 foi estratificada por esses dois fatores: localização e dependência administrativa, à exceção da região Norte. Um terceiro fator de estratificação - Capital e Interior - foi introduzido apenas por conveniência administrativa.*

Testes – *Nesta edição, foram avaliadas as proficiências dos alunos de 4ª e de 8ª série do Ensino Fundamental nas disciplinas: Língua Portuguesa, Matemática e Ciências. Os alunos de 3ª série do Ensino Médio foram avaliados em Língua Portuguesa, Matemática, Física, Química e Biologia. Ao todo, foram aplicados 1.737 itens de teste, sendo 416 itens de 4ª série, 476 de 8ª e 845 de 3ª série do Ensino Médio. Os cadernos de testes, por série e disciplina, estruturam-se em três blocos cada um, contendo entre 10 a 13 itens; portanto cada aluno respondeu entre 30 a 39 itens. Foram organizados 26 cadernos por série/disciplina, com blocos de itens comuns entre séries e disciplinas. A maioria dos itens foi de múltipla escolha. Somente 33 dos itens aplicados aos alunos da 8ª série e 3ª série do Ensino Médio solicitaram que o aluno construísse e escrevesse sua resposta. Os itens de teste foram reunidos em 13 blocos e em 26 cadernos por série e por disciplina, de modo que foi confeccionado um total de 208 cadernos distintos. Os cadernos de teste foram compostos seguindo um planejamento de blocos balanceados incompletos, sendo que cada bloco aparece o mesmo número de vezes em cada posição no caderno e cada par de blocos aparece somente uma vez nos cadernos. Com a finalidade de construir as escalas de proficiências comuns a todas as séries avaliadas, foram aplicadas questões comuns às quatro séries: 10 questões de leitura e 10 de matemática comuns a 4ª, 8ª e 3ª série do Ensino Médio; 26 de leitura e 22 de matemática comuns somente a 8ª série e 3ª série do Ensino Médio; 30 de matemática e 30 de leitura comuns somente a 4ª e 8ª série; 13 de leitura e 13 de matemática comuns a 8ª e 3ª série do Ensino Médio [sic].*

Questionários - *Os alunos responderam a um questionário dividido em duas partes, colocadas respectivamente no início e no final do caderno de teste. A primeira continha perguntas sobre seu ambiente socioeconômico e cultural; a segunda, sobre seus hábitos de estudo. Os professores também responderam*

a um questionário dividido em duas partes: na primeira 48 questões abordavam aspectos da formação e experiência profissional; a segunda trazia 31 questões sobre a prática em sala de aula e atitudes em relação à escola e aos alunos. O questionário dos diretores continha 68 questões que abordavam aspectos da formação e experiência profissional, situação funcional, organização e funcionamento da escola. O questionário da escola, preenchido pelo aplicador juntamente com o Secretário da escola, continha 37 questões referentes às condições físicas e materiais da escola.

Análise - O SAEB/97 utilizou a Teoria de Reposta ao Item para realizar análises para cálculo das proficiências. Os modelos utilizados foram os de 3 parâmetros (modelo de Lord 1980). Assim, a probabilidade de acertar o item é uma função do índice de discriminação do item, do grau de dificuldade e da probabilidade de acerto casual. (Brasil/MEC/INEP, 2006, p. 10-12)

- SAEB 1999

A edição do SAEB/99 foi realizada pelo Inep/MEC com a participação das Secretarias Estaduais e Municipais de Educação, entre os dias 25 e 29 de outubro.

Amostra – O universo considerado nesta pesquisa consistiu de todos os alunos matriculados na 4ª e na 8ª séries do Ensino Fundamental e 3ª série do Ensino Médio em todo o Brasil. Três exclusões foram feitas do marco da amostragem: as escolas federais, por serem em número muito reduzido, as escolas rurais da região Norte, pela dificuldade de acesso e as turmas multisseriadas, pela dificuldade de aplicação dos testes. A amostra prevista para o SAEB/99, contemplava 360.451 alunos distribuídos nas três séries anteriormente citadas, cobrindo todas as Unidades da Federação (UF). A distribuição desses alunos entre as séries considerou tanto a proporção de matrículas em cada série como o número de disciplinas a serem avaliadas em cada uma das séries (5 na 4ª, 5 na 8ª e 8 na 3ª). Assim, a amostra foi formada por 133.143 alunos da 4ª série, 114.516 alunos de 8ª série do Ensino Fundamental e 112.792 alunos de 3ª série do Ensino Médio, distribuídos igualmente entre as 27 Unidades da Federação. A amostra do SAEB/99 foi estratificada pela localização (urbana e rural) e dependência administrativa (estadual, municipal, particular). O desenho amostral previu a exclusão dos alunos das escolas rurais da 8ª série do Ensino Fundamental e da 3ª série do Ensino Médio, tendo em vista o tamanho reduzido desse universo de escolas e o pequeno impacto de tal exclusão nas médias de desempenho das populações de interesse. Apenas no caso da 4ª série do Ensino Fundamental, nos estados onde a matrícula nas escolas da zona rural representa uma parcela significativa da matrícula total, os alunos dessas escolas foram incluídos na amostra. Além disso, a avaliação na área rural somente atingiu a Região Nordeste e os estados de Minas Gerais e Mato Grosso do Sul. As demais regiões e estados, na 4ª série do ensino fundamental, mostraram médias de desempenho no Saeb/97 muito semelhantes àquelas estimadas para as áreas urbanas (diferença inferior a 1%), levando a que fossem excluídas do Saeb/99, sob o argumento de não causarem impacto nas médias finais observadas para os estados correspondentes. Algumas sub-populações, por apresentarem número muito reduzido de elementos, não permitem abertura dos estratos para o nível de Dependência Administrativa, por exemplo. Esse é o caso dos chamados estratos especiais, que correspondem aos estados do Acre, Roraima e Amapá, na 4ª série do ensino fundamental, e aos quais se somam, na 8ª série do ensino fundamental e na 3ª série do ensino médio, os estados de Rondônia e Tocantins. O plano de amostragem previa três etapas: seleção de municípios, de escolas e de turmas, nas três séries avaliadas.

Testes – Nesta edição, foram avaliadas as proficiências dos alunos de 4ª e de 8ª série do Ensino Fundamental nas disciplinas: Língua Portuguesa, Matemática, Ciências, História e Geografia. Os alunos de 3ª série do Ensino

Médio foram avaliados em Língua Portuguesa, Matemática, Física, Química, Biologia, História e Geografia. Os cadernos de testes, por série e disciplina, estruturaram-se em três blocos cada um, contendo 13 itens; portanto cada aluno respondeu 39 itens. Foram organizados 26 cadernos por série/disciplina, com blocos de itens comuns entre séries e disciplinas, seguindo um planejamento de blocos balanceados incompletos, sendo que cada bloco aparece o mesmo número de vezes em cada posição no caderno e cada par de blocos aparece somente uma vez nos cadernos.

Questionários - Os alunos responderam a um questionário dividido em duas partes, colocadas respectivamente no início e no final do caderno de teste. A primeira continha perguntas sobre seu ambiente socioeconômico e cultural; a segunda, sobre seus hábitos de estudo. Os professores também responderam a um questionário dividido em duas partes: na primeira abordavam aspectos da formação e experiência profissional; a segunda sobre a prática em sala de aula e atitudes em relação à escola e aos alunos. O questionário dos diretores abordavam [sic] aspectos da formação e experiência profissional, situação funcional, organização e funcionamento da escola. O questionário da escola, preenchido pelo aplicador juntamente com o Secretário da escola, continha questões referentes às condições físicas e materiais da escola.

Análise - O SAEB/99 utilizou a Teoria de Reposta ao Item para realizar análises para cálculo das proficiências. Os modelos utilizados foram os de 3 parâmetros (modelo de Lord 1980). Assim, a probabilidade de acertar o item é uma função do índice de discriminação do item, do grau de dificuldade e da probabilidade de acerto casual. (Brasil/MEC/INEP, 2006, p. 10-12)

- SAEB 2001

A edição do SAEB/2001 foi realizada pelo MEC com a participação das Secretarias Estaduais e Municipais de Educação, entre os dias 22 e 26 de outubro.

Amostra - Coletou e processou informações de 288.331 alunos participantes de uma amostra probabilística de todos os alunos das escolas no Brasil, matriculados nas 4ª e 8ª séries do Ensino Fundamental e 3ª série do Ensino Médio, freqüentando 11.846 turmas em 6.907 escolas públicas e particulares do País. Obteve também informações de 23.503 professores de Matemática e de Língua Portuguesa e de 6.831 diretores das escolas pesquisadas. A população de referência foi composta por alunos de 4ª e 8ª séries do Ensino Fundamental e alunos de 3ª série do Ensino Médio, matriculados em 2001, nas escolas urbanas constantes no Censo Escolar de 2000, excetuando-se os alunos das escolas federais, rurais ou de turmas multisseriadas. Para o universo da 4ª série, foram mantidos os alunos das escolas rurais do Nordeste, de Minas Gerais e do Mato Grosso do Sul. A população de referência das escolas foi dividida em subpopulações independentes: séries, unidades da Federação, dependência administrativa (estadual, municipal, particular), localização (capital, interior) e tamanho das escolas segundo o número de turmas. A combinação desses critérios resulta em um determinado número de estratos de interesse dentre os quais foram selecionadas as amostras de alunos, por série. (Consultar plano amostral e resultados relatório final da amostra nos anexos)

Testes - Para a avaliação das proficiências dos alunos brasileiros foram aplicados itens de múltipla escolha, das áreas de Matemática e de Leitura. Os testes foram constituídos por 169 itens para cada uma das séries e disciplinas avaliadas. Esse total foi distribuído em 13 blocos de 13 itens, organizados de três a três, por meio de um delineamento denominado Blocos Incompletos Balanceados (BIB), possibilitando a composição de 26 cadernos de provas diferentes de Língua Portuguesa e outros 26 de Matemática, para cada uma das séries (4ª e 8ª do Ensino Fundamental e 3ª do Ensino Médio). Como cada

caderno contém três blocos com 13 itens, significa que cada aluno responde, no máximo, a 39 itens.

Questionários - Os alunos responderam a um questionário dividido em duas partes, colocadas respectivamente no início e no final do caderno de teste. A primeira continha perguntas sobre seu ambiente socioeconômico e cultural; a segunda, sobre seus hábitos de estudo. Os professores também responderam a um questionário dividido em duas partes: na primeira abordava aspectos da formação e experiência profissional; a segunda sobre a prática em sala de aula e atitudes em relação à escola e aos alunos. O questionário dos diretores abordava aspectos da formação e experiência profissional, situação funcional, organização e funcionamento da escola. O questionário da escola, preenchido pelo aplicador juntamente com o Secretário da escola, continha questões referentes às condições físicas e materiais da escola.

Análise - O SAEB/2001 utilizou a Teoria de Reposta ao Item para realizar análises para cálculo das proficiências. O modelo utilizado foi o da logística de 3 parâmetros – para itens de múltipla escolha. (Brasil/MEC/INEP, 2006, p. 10-11)

- SAEB 2003

A edição do SAEB/2003 foi realizada pelo MEC com a participação das Secretarias Estaduais e Municipais de Educação, entre os dias 03 e 07 de novembro.

Amostra - Coletou e processou informações de 218.521 alunos participantes de uma amostra probabilística de todos os alunos das escolas no Brasil, distribuídos conforme as tabelas abaixo.

**Número de Alunos SAEB 2003
Previsto e Realizado, por Dependência Administrativa e Série**

SAEB					
SÉRIE	DEP. ADM.	PREVISTO	PRESENTES	AUSENTES	% AUSÊNCIA
4ª EF	Estadual	39425	32620	6805	17,3%
	Federal	635	600	35	5,5%
	Municipal	39345	33224	6121	15,6%
	Particular	28077	25754	2323	8,3%
4ª Total		107482	92198	15284	14,2%
8ª EF	Estadual	40254	30055	10199	25,3%
	Federal	717	623	94	13,1%
	Municipal	26495	20670	5825	22,0%
	Particular	24916	22569	2347	9,4%
8ª Total		92382	73917	18465	20,0%
3ª EM	Estadual	39206	26995	12211	31,1%
	Federal	1289	958	331	25,7%
	Municipal	1002	740	262	26,1%
	Particular	30128	23713	6415	21,3%
3ª EM Total		71625	52406	19219	26,8%
Total geral		271489	218521	52968	19,5%

Dezessete mil, trezentos e setenta e seis professores brasileiros das séries iniciais, de Língua Portuguesa e de Matemática, participaram do SAEB/2003

respondendo a um questionário de dados pessoais, formação profissional e opiniões/ posicionamentos sobre o cotidiano escolar, inclusive sobre sua situação como profissional (Burnout, síndrome da desistência). Em 2003, 8.096 professores de 4ª série EF responderam ao questionário, de um total previsto de 8.286, correspondendo a uma perda de 2,3%. Na 8ª série, de uma previsão de 5.896 professores, participaram 5.590 docentes, com uma perda de 5,2%, enquanto que na 3ª série EM foram obtidos dados de 3.690 professores de um total previsto de 3.932, o que significa uma perda de 6,2%.

Questionário do Professor SAEB 2003
Previsto e Realizado, por Série e Dependência Administrativa

SÉRIE	DEPENDÊNCIA	PREVISTO	PRESENTES	AUSENTES	% AUSÊNCIA
4ª EF	Estadual	2728	2654	74	2,7%
	Federal	46	42	4	8,7%
	Municipal	2730	2677	53	1,9%
	Particular	2782	2723	59	2,1%
4ª EF Total		8286	8096	190	2,3%
8ª EF	Estadual	2410	2266	144	6,0%
	Federal	44	42	2	4,5%
	Municipal	1710	1617	93	5,4%
	Particular	1732	1665	67	3,9%
8ª EF Total		5896	5590	306	5,2%
3ª EM	Estadual	2166	2043	123	5,7%
	Federal	80	74	6	7,5%
	Municipal	58	56	2	3,4%
	Particular	1628	1517	111	6,8%
3ª EM Total		3932	3690	242	6,2%
Total geral		18114	17376	738	4,1%

Os resultados de 2003 apontam que dos 6.270 Diretores previstos, 6.180 responderam ao questionário, com uma perda portanto de apenas 4%. No SAEB 2003 foram pesquisadas 6179 escolas por meio de um instrumento específico (lista de verificação) preenchido pelos aplicadores (Anexo 3)

Amostra - Como no ciclo de 2001, a população de referência foi constituída considerando o universo dos alunos de 4ª e 8ª séries do Ensino Fundamental e alunos de 3ª série do Ensino Médio. Para o SAEB-2003 foram feitas algumas demandas específicas visando ampliar a cobertura da população de alunos das séries de interesse. Entre elas destacaram-se a intenção de cobrir a população de alunos de 4ª série do ensino fundamental em escolas rurais, e a inclusão de alunos das escolas federais para alunos das três séries. Foi então refeita a avaliação das exclusões de alunos e a definição do universo de investigação. Os resultados dessa avaliação estão descritos em Silva et al (2003a). Assim, a população de referência no SAEB/2003 foi assim constituída:

- Todos os alunos matriculados em 2003 nas escolas urbanas e constantes no Censo Escolar de 2002 em uma das três séries de interesse, excetuando-se os alunos das turmas multisseriadas e de aceleração.

- Foram incluídos também no universo da quarta série do ensino fundamental os alunos das escolas não federais rurais com 10 ou mais alunos na série.

A população de referência das escolas foi dividida em subpopulações independentes: séries, unidades da Federação, dependência administrativa (estadual, municipal, particular), localização (capital, interior) e tamanho das escolas segundo o número de turmas. A combinação desses critérios resulta em um determinado número de estratos de interesse dentre os quais foram

selecionadas as amostras de alunos, por série. (Consultar plano amostral e resultados relatório final da amostra nos anexos)

Testes - Para a avaliação das proficiências dos alunos brasileiros foram aplicados itens de múltipla escolha, das áreas de Matemática e de Leitura. Os testes foram constituídos por 169 itens para cada uma das séries e disciplinas avaliadas. Esse total foi distribuído em 13 blocos de 13 itens, organizados de três a três, por meio de um delineamento denominado Blocos Incompletos Balanceados (BIB), possibilitando a composição de 26 cadernos de provas diferentes de Língua Portuguesa e outros 26 de Matemática, para cada uma das séries (4ª e 8ª do Ensino Fundamental e 3ª do Ensino Médio). Como cada caderno contém três blocos com 13 itens, significa que cada aluno responde, no máximo, a 39 itens.

Questionários - Os alunos responderam a um questionário dividido em duas partes, colocadas respectivamente no início e no final do caderno de teste. A primeira continha perguntas sobre seu ambiente socioeconômico e cultural; a segunda, sobre seus hábitos e ambiente de estudos. Os professores também responderam a um questionário dividido em duas partes: na primeira abordava aspectos da formação e experiência profissional; a segunda sobre a prática em sala de aula e atitudes em relação à escola e aos alunos e suas sensações e expectativas como trabalhador da educação (Burnout, síndrome da desistência). O questionário dos diretores abordava aspectos da formação e experiência profissional, situação funcional, organização e funcionamento da escola. O questionário da escola, preenchido pelo aplicador juntamente com o Secretário da escola, continha questões referentes às condições físicas e materiais da escola.

Análise - O SAEB/2003 utilizou a Teoria de Reposta ao Item para realizar análises para cálculo das proficiências. O modelo utilizado foi o da logística de 3 parâmetros – para itens de múltipla escolha. (Brasil/MEC/INEP, 2006, p. 10-13)

7.2 Transformação linear de escalas

1. $\theta^* = dp \cdot \theta + m$
2. $b^* = dp \cdot b + m$
3. $a^* = a/dp$
4. $P(X_i = 1 | \theta) = P(X_i = 1 | \theta^*)$

Onde:

- θ^* é a proficiência na nova escala;
- dp é o desvio padrão da nova escala;
- θ é a proficiência na escala antiga;
- m é a média da nova escala;
- b^* é o parâmetro de dificuldade de cada item na escala nova;

- b é o parâmetro de dificuldade de cada item na escala antiga;
- a^* é o parâmetro de discriminação de cada item na escala nova;
- a é o parâmetro de discriminação de cada item na escala antiga;
- e a sentença 4 significa que a probabilidade de um aluno com habilidade θ acertar o item i é a mesma nas duas escalas.

7.3 Modelo de regressão multinível no formato PROC MIXED

O modelo a seguir foi detalhado no item 2.4.2 e está representado em dois níveis na equação (5) desta tese.

$$Y_{ij} = [\gamma_{00} + \gamma_{01} MNSE_j + \gamma_{02} REDE_j + \gamma_{03} MATRASO_j + \gamma_{10} NSE_{ij} + \gamma_{20} SEXO_{ij} + \gamma_{30} PARDO_{ij} + \gamma_{40} PRETO_{ij} + \gamma_{50} ATRASO_{ij} + \gamma_{60} DISCIPLINA_{ij} + \gamma_{70} ATITUDE_{ij}] + [u_{0j} + r_{ij}]$$

$$\text{Onde } r_{ij} \sim N(0, \sigma^2) \text{ e } \begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{pmatrix} \right]$$

A *syntax* a seguir foi utilizada para rodar, no SPSS 13.0, este modelo base. Entre outras coisas, pode-se perceber a importância de que as análises dos dados do SAEB sejam sempre feitas com a utilização de um peso.

```
MIXED
profic WITH mnse rede_1 nse_cmg sexo preto pardo atraso matraso disc atitude
/FIXED = mnse rede_1 matraso nse_cmg sexo pardo preto atraso disc atitude /
SSTYPE(3)
/PRINT = SOLUTION TESTCOV
/RANDOM INTERCEPT nse_cmg | SUBJECT(id) COVTYPE(UN)
/REGWGT = peso_r .
```

A seguir estão todas as tabelas de resultados, para as 14 bases do SAEB, geradas pelo SPSS ao rodar a *syntax* acima.

7.4 Saídas do SPSS para o Modelo Base

Mixed Model Analysis 1995 - 8ª série

Warnings

The covariance structure for random effect with only one level will be changed to Identity.

Model Dimension(a,b)

		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1	Identity	1	id
	mnse	1		1	
	rede_1	1		1	
	matraso	1		1	
	nse_cmg	1		1	
	sexo	1		1	
	pardo	1		1	
	preto	1		1	
	atraso	1		1	
	disc	1		1	
	atitude	1		1	
Random Effects	Intercept	1	Identity	1	id
Residual				1	
Total		12		13	

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Information Criteria(a,b)

-2 Restricted Log Likelihood	259978,96
Akaike's Information Criterion (AIC)	259982,96
Hurvich and Tsai's Criterion (AICC)	259982,96
Bozdogan's Criterion (CAIC)	260001,23
Schwarz's Bayesian Criterion (BIC)	259999,23

The information criteria are displayed in smaller-is-better forms.

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Fixed Effects

Type III Tests of Fixed Effects(a,b)

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	1847,812	57246,640	,000
mNSE	1	2347,004	370,354	,000
rede_1	1	1188,058	1,607	,205
matrasso	1	2640,073	11,907	,001
nse_cmG	1	24156,809	81,155	,000
sexo	1	24761,240	111,737	,000
pardo	1	25161,530	4,532	,033
preto	1	25009,099	65,795	,000
atraso	1	24154,043	873,566	,000
disc	1	24180,507	2,439	,118
atitude	1	24150,414	276,139	,000

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Estimates of Fixed Effects(a,b)

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	265,36159	1,109081	1847,812	239,263	,000	263,186417	267,536782
mNSE	17,795611	,924708	2347,004	19,245	,000	15,982282	19,608939
rede_1	2,428762	1,916178	1188,058	1,268	,205	-1,330708	6,188233
matrasso	-1,757051	,509203	2640,073	-3,451	,001	-2,755529	-,758573
nse_cmG	2,803997	,311257	24156,809	9,009	,000	2,193914	3,414079
sexo	-5,438499	,514493	24761,240	-10,571	,000	-6,446937	-4,430062
pardo	-1,212021	,569302	25161,530	-2,129	,033	-2,327887	-,096156
preto	-7,776189	,958676	25009,099	-8,111	,000	-9,655249	-5,897128
atraso	-6,066228	,205244	24154,043	-29,556	,000	-6,468519	-5,663937
disc	-,778829	,498647	24180,507	-1,562	,118	-1,756208	,198550
atitude	12,589322	,757598	24150,414	16,617	,000	11,104383	14,074260

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Covariance Parameters

Estimates of Covariance Parameters(a,b)

Parameter	Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Residual	1513,0281	13,787521	109,739	,000	1486,244996	1540,293960
Intercept [subject = id] Variance	240,75850	13,983421	17,217	,000	214,853912	269,786376

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Mixed Model Analysis

1995 - 11ª série

Warnings

The covariance structure for random effect with only one level will be changed to Identity.

Model Dimension(a,b)

		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1	Identity	1	
	mnse	1		1	
	rede_1	1		1	
	matraso	1		1	
	nse_cmg	1		1	
	sexo	1		1	
	pardo	1		1	
	preto	1		1	
	atraso	1		1	
	disc	1		1	
	atitude	1		1	
	Random Effects	Intercept		1	
Residual				1	
Total		12		13	

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Information Criteria(a,b)

-2 Restricted Log Likelihood	111706,00
Akaike's Information Criterion (AIC)	111710,00
Hurvich and Tsai's Criterion (AICC)	111710,00
Bozdogan's Criterion (CAIC)	111726,56
Schwarz's Bayesian Criterion (BIC)	111724,56

The information criteria are displayed in smaller-is-better forms.

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Fixed Effects

Type III Tests of Fixed Effects(a,b)

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	1725,437	19753,024	,000
mnse	1	969,147	193,158	,000
rede_1	1	706,181	5,699	,017
matraso	1	1568,301	9,621	,002
nse_cmg	1	10148,714	8,454	,004
sexo	1	10718,420	370,790	,000
pardo	1	10720,833	3,217	,073
preto	1	10582,657	6,047	,014
atraso	1	10132,765	232,319	,000
disc	1	10723,310	,035	,852
atitude	1	10238,869	107,565	,000

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Estimates of Fixed Effects(a,b)

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	300,78810	2,140148	1725,437	140,545	,000	296,590546	304,985662
mnse	20,380012	1,466386	969,147	13,898	,000	17,502355	23,257669
rede_1	5,441581	2,279357	706,181	2,387	,017	,966454	9,916708
matraso	-2,771128	,893385	1568,301	-3,102	,002	-4,523482	-1,018773
nse_cmg	1,401508	,482006	10148,714	2,908	,004	,456680	2,346336
sexo	-16,31273	,847155	10718,420	-19,256	,000	-17,973318	-14,652157
pardo	-1,608434	,896763	10720,833	-1,794	,073	-3,366255	,149387
preto	-4,217365	1,715021	10582,657	-2,459	,014	-7,579130	-,855601
atraso	-4,426384	,290407	10132,765	-15,242	,000	-4,995639	-3,857129
disc	,192885	1,035979	10723,310	,186	,852	-1,837825	2,223595
atitude	10,964095	1,057152	10238,869	10,371	,000	8,891870	13,036321

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Covariance Parameters

Estimates of Covariance Parameters(a,b)

Parameter	Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Residual	1547,1247	21,781913	71,028	,000	1505,016580	1590,410945
Intercept [subject = id] Variance	319,75574	23,597479	13,550	,000	276,694800	369,518096

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Mixed Model Analysis

1997 - 4ª série

Warnings

The covariance structure for random effect with only one level will be changed to Identity.

Model Dimension(a,b)

		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1	Identity	1	
	mnse	1		1	
	rede_1	1		1	
	matraso	1		1	
	nse_cmg	1		1	
	sexo	1		1	
	pardo	1		1	
	preto	1		1	
	atraso	1		1	
	disc	1		1	
	atitude	1		1	
	Random Effects	Intercept		1	
Residual				1	
Total		12		13	

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Information Criteria(a,b)

-2 Restricted Log Likelihood	421311,58
Akaike's Information Criterion (AIC)	421315,58
Hurvich and Tsai's Criterion (AICC)	421315,58
Bozdogan's Criterion (CAIC)	421334,80
Schwarz's Bayesian Criterion (BIC)	421332,80

The information criteria are displayed in smaller-is-better forms.

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Fixed Effects

Type III Tests of Fixed Effects(a,b)

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	4322,821	63343,758	,000
mnse	1	6893,207	685,919	,000

rede_1	1	2941,891	349,171	,000
matraso	1	7272,137	2,147	,143
nse_cmg	1	38448,314	41,830	,000
sexo	1	39446,343	,239	,625
pardo	1	40236,096	6,947	,008
preto	1	39781,445	122,648	,000
atraso	1	38438,381	227,590	,000
disc	1	38585,145	197,571	,000
atitude	1	38475,869	736,932	,000

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Estimates of Fixed Effects(a,b)

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	181,23243	,720085	4322,821	251,682	,000	179,820700	182,644173
mnse	15,292964	,583923	6893,207	26,190	,000	14,148296	16,437632
rede_1	20,470414	1,095487	2941,891	18,686	,000	18,322414	22,618413
matraso	,565182	,385704	7272,137	1,465	,143	-,190910	1,321275
nse_cmg	1,467487	,226899	38448,314	6,468	,000	1,022760	1,912214
sexo	,179116	,366471	39446,343	,489	,625	-,539175	,897407
pardo	1,056811	,400958	40236,096	2,636	,008	,270924	1,842698
preto	-7,488491	,676182	39781,445	-11,075	,000	-8,813824	-6,163158
atraso	-2,648960	,175589	38438,381	-15,086	,000	-2,993120	-2,304800
disc	5,032370	,358023	38585,145	14,056	,000	4,330637	5,734104
atitude	21,751711	,801272	38475,869	27,146	,000	20,181198	23,322224

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Covariance Parameters

Estimates of Covariance Parameters(a,b)

Parameter	Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Residual	1277,5577	9,233947	138,354	,000	1259,587150	1295,784768
Intercept [subject = id] Variance	196,50037	8,452338	23,248	,000	180,613205	213,785025

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Mixed Model Analysis 1997 - 8ª série

Warnings

The covariance structure for random effect with only one level will be changed to Identity.

Model Dimension(a,b)

		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1	Identity	1	id
	mnse	1		1	
	rede_1	1		1	
	matraso	1		1	
	nse_cmg	1		1	
	sexo	1		1	
	pardo	1		1	
	preto	1		1	
	atraso	1		1	
	disc	1		1	
	atitude	1		1	
Random Effects	Intercept	1		1	
Residual				1	
Total		12		13	

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Information Criteria(a,b)

-2 Restricted Log Likelihood	331014,75
Akaike's Information Criterion (AIC)	331018,75
Hurvich and Tsai's Criterion (AICC)	331018,75
Bozdogan's Criterion (CAIC)	331037,46
Schwarz's Bayesian Criterion (BIC)	331035,46

The information criteria are displayed in smaller-is-better forms.

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Fixed Effects

Type III Tests of Fixed Effects(a,b)

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	3347,566	78056,060	,000
mnse	1	4328,725	413,395	,000
rede_1	1	2144,815	173,992	,000
matraso	1	5719,395	2,425	,120
nse_cmg	1	30039,303	21,011	,000
sexo	1	30896,101	81,596	,000

pardo	1	31398,635	16,641	,000
preto	1	31102,153	59,924	,000
atraso	1	30018,355	589,990	,000
disc	1	30091,831	16,575	,000
atitude	1	30023,121	474,282	,000

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Estimates of Fixed Effects(a,b)

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	258,05741	,923662	3347,566	279,385	,000	256,246411	259,868409
mnse	13,333845	,655802	4328,725	20,332	,000	12,048137	14,619553
rede_1	17,180149	1,302452	2144,815	13,191	,000	14,625948	19,734350
matraso	-,696603	,447376	5719,395	-1,557	,120	-1,573629	,180423
nse_cmg	1,215220	,265114	30039,303	4,584	,000	,695586	1,734854
sexo	-4,142727	,458620	30896,101	-9,033	,000	-5,041640	-3,243813
pardo	-2,012735	,493394	31398,635	-4,079	,000	-2,979806	-1,045664
preto	-8,100785	1,046470	31102,153	-7,741	,000	-10,151908	-6,049663
atraso	-5,119341	,210762	30018,355	-24,290	,000	-5,532443	-4,706239
disc	1,799399	,441983	30091,831	4,071	,000	,933093	2,665705
atitude	14,212136	,652591	30023,121	21,778	,000	12,933030	15,491241

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Covariance Parameters

Estimates of Covariance Parameters(a,b)

Parameter	Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Residual	1515,6670	12,404298	122,189	,000	1491,549006	1540,175046
Intercept [subject = id] Variance	202,34096	10,229907	19,779	,000	183,252106	223,418264

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Mixed Model Analysis

1997 - 11ª série

Warnings

The covariance structure for random effect with only one level will be changed to Identity.

Model Dimension(a,b)

		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1	Identity	1	id
	mNSE	1		1	
	rede_1	1		1	
	matraso	1		1	
	nse_cmG	1		1	
	sexo	1		1	
	pardo	1		1	
	preto	1		1	
	atraso	1		1	
	disc	1		1	
	atitude	1		1	
Random Effects	Intercept	1		1	
Residual				1	
Total		12		13	

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Information Criteria(a,b)

-2 Restricted Log Likelihood	151251,85
Akaike's Information Criterion (AIC)	151255,85
Hurvich and Tsai's Criterion (AICC)	151255,85
Bozdogan's Criterion (CAIC)	151272,97
Schwarz's Bayesian Criterion (BIC)	151270,97

The information criteria are displayed in smaller-is-better forms.

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Fixed Effects

Type III Tests of Fixed Effects(a,b)

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	2551,465	37833,678	,000
mNSE	1	3801,347	204,267	,000
rede_1	1	1458,003	111,131	,000
matraso	1	5717,626	48,024	,000
nse_cmG	1	13207,101	9,597	,002
sexo	1	13895,881	190,160	,000
pardo	1	14079,072	,700	,403
preto	1	13948,077	8,133	,004
atraso	1	13199,813	261,293	,000

disc	1	13244,865	365,974	,000
atitude	1	13196,082	179,851	,000

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Estimates of Fixed Effects(a,b)

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	299,78873	1,541260	2551,465	194,509	,000	296,766487	302,810985
mNSE	12,798409	,895482	3801,347	14,292	,000	11,042737	14,554080
rede_1	18,953061	1,797886	1458,003	10,542	,000	15,426342	22,479780
matrasso	-4,965835	,716577	5717,626	-6,930	,000	-6,370598	-3,561072
nse_cmG	1,259229	,406480	13207,101	3,098	,002	,462470	2,055988
sexo	-10,32527	,748759	13895,881	-13,790	,000	-11,792937	-8,857602
pardo	-,681175	,814311	14079,072	-,837	,403	-2,277333	,914983
preto	-5,313578	1,863185	13948,077	-2,852	,004	-8,965671	-1,661485
atraso	-5,189828	,321062	13199,813	-16,165	,000	-5,819156	-4,560500
disc	13,464023	,703801	13244,865	19,130	,000	12,084472	14,843573
atitude	13,641157	1,017172	13196,082	13,411	,000	11,647354	15,634960

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Covariance Parameters

Estimates of Covariance Parameters(a,b)

Parameter	Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Residual	1714,1307	21,185450	80,911	,000	1673,106886	1756,160445
Intercept [subject = id] Variance	343,38777	21,513702	15,961	,000	303,707780	388,252035

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Mixed Model Analysis

1999 - 4ª série

Warnings

The covariance structure for random effect with only one level will be changed to Identity.

Model Dimension(a,b)

	Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects Intercept	1		1	
mNSE	1		1	

	rede_1	1		1	
	matraso	1		1	
	nse_cmg	1		1	
	sexo	1		1	
	pardo	1		1	
	preto	1		1	
	atraso	1		1	
	disc	1		1	
	atitude	1		1	
Random Effects	Intercept	1	Identity	1	id
Residual				1	
Total		12		13	

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Information Criteria(a,b)

-2 Restricted Log Likelihood	377900,12
Akaike's Information Criterion (AIC)	377904,12
Hurvich and Tsai's Criterion (AICC)	377904,12
Bozdogan's Criterion (CAIC)	377923,16
Schwarz's Bayesian Criterion (BIC)	377921,16

The information criteria are displayed in smaller-is-better forms.

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Fixed Effects

Type III Tests of Fixed Effects(a,b)

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	9618,876	74521,082	,000
mnse	1	9285,618	931,526	,000
rede_1	1	5164,489	464,633	,000
matraso	1	19165,509	1,734	,188
nse_cmg	1	33438,297	8,217	,004
sexo	1	35872,114	13,887	,000
pardo	1	36905,595	,003	,958
preto	1	36500,853	220,037	,000
atraso	1	33373,014	390,562	,000
disc	1	33586,366	1085,856	,000
atitude	1	33408,626	452,120	,000

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Estimates of Fixed Effects(a,b)

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	170,56310	,624806	9618,876	272,985	,000	169,338353	171,787857
mnse	8,535647	,279666	9285,618	30,521	,000	7,987442	9,083853
rede_1	18,854273	,874691	5164,489	21,555	,000	17,139508	20,569037
matraso	,429838	,326417	19165,509	1,317	,188	-,209968	1,069644
nse_cmg	,589322	,205591	33438,297	2,866	,004	,186356	,992288
sexo	1,375613	,369141	35872,114	3,727	,000	,652087	2,099140
pardo	,021741	,408012	36905,595	,053	,958	-,777974	,821456
preto	-9,322362	,628460	36500,853	-14,834	,000	-10,554162	-8,090562
atraso	-3,369557	,170501	33373,014	-19,763	,000	-3,703746	-3,035368
disc	11,769789	,357176	33586,366	32,952	,000	11,069712	12,469867
atitude	13,318531	,626368	33408,626	21,263	,000	12,090829	14,546233

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Covariance Parameters

Estimates of Covariance Parameters(a,b)

Parameter	Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Residual	1146,6135	8,932833	128,359	,000	1129,238515	1164,255939
Intercept [subject = id] Variance	211,02633	7,865002	26,831	,000	196,160780	227,018449

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Mixed Model Analysis 1999 - 8ª série

Warnings

The covariance structure for random effect with only one level will be changed to Identity.

Model Dimension(a,b)

	Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects				
Intercept	1		1	
mnse	1		1	
rede_1	1		1	
matraso	1		1	
nse_cmg	1		1	
sexo	1		1	

	pardo	1		1	
	preto	1		1	
	atraso	1		1	
	disc	1		1	
	atitude	1		1	
Random Effects	Intercept	1	Identity	1	id
Residual				1	
Total		12		13	

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Information Criteria(a,b)

-2 Restricted Log Likelihood	320639,54
Akaike's Information Criterion (AIC)	320643,54
Hurvich and Tsai's Criterion (AICC)	320643,54
Bozdogan's Criterion (CAIC)	320662,23
Schwarz's Bayesian Criterion (BIC)	320660,23

The information criteria are displayed in smaller-is-better forms.

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Fixed Effects

Type III Tests of Fixed Effects(a,b)

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	6726,198	96988,697	,000
mNSE	1	6604,277	674,456	,000
rede_1	1	4029,437	187,495	,000
matraso	1	12720,786	17,093	,000
nse_cmG	1	28413,596	,096	,757
sexo	1	30340,121	56,046	,000
pardo	1	30917,542	26,458	,000
preto	1	30802,018	34,356	,000
atraso	1	28333,045	635,174	,000
disc	1	28464,492	1292,812	,000
atitude	1	28370,122	884,869	,000

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Estimates of Fixed Effects(a,b)

Paramete	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval
----------	----------	------------	----	---	------	-------------------------

r						Lower Bound	Upper Bound
Intercept	246,18159	,790488	6726,198	311,430	,000	244,631984	247,731196
mnse	9,134235	,351719	6604,277	25,970	,000	8,444753	9,823718
rede_1	14,335866	1,046957	4029,437	13,693	,000	12,283252	16,388480
matraso	-1,622452	,392436	12720,786	-4,134	,000	-2,391685	-,853219
nse_cmg	,071760	,231966	28413,596	,309	,757	-,382903	,526424
sexo	-3,377238	,451119	30340,121	-7,486	,000	-4,261450	-2,493026
pardo	-2,524366	,490764	30917,542	-5,144	,000	-3,486283	-1,562448
preto	-5,316204	,906989	30802,018	-5,861	,000	-7,093940	-3,538468
atraso	-5,628450	,223328	28333,045	-25,203	,000	-6,066183	-5,190718
disc	15,518630	,431604	28464,492	35,956	,000	14,672665	16,364594
atitude	18,741506	,630035	28370,122	29,747	,000	17,506607	19,976405

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Covariance Parameters

Estimates of Covariance Parameters(a,b)

Parameter	Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Residual	1413,6662	11,942064	118,377	,000	1390,452937	1437,267107
Intercept [subject = id] Variance	224,16690	9,601147	23,348	,000	206,117197	243,797222

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Mixed Model Analysis 1999 - 11^a série

Warnings

The covariance structure for random effect with only one level will be changed to Identity.

Model Dimension(a,b)

		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1		1	
	mnse	1		1	
	rede_1	1		1	
	matraso	1		1	
	nse_cmg	1		1	
	sexo	1		1	
	pardo	1		1	
	preto	1		1	
	atraso	1		1	

	disc	1		1	
	atitude	1		1	
Random Effects	Intercept	1	Identity	1	id
Residual				1	
Total		12		13	

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Information Criteria(a,b)

-2 Restricted Log Likelihood	220242,41
Akaike's Information Criterion (AIC)	220246,41
Hurvich and Tsai's Criterion (AICC)	220246,41
Bozdogan's Criterion (CAIC)	220264,32
Schwarz's Bayesian Criterion (BIC)	220262,32

The information criteria are displayed in smaller-is-better forms.

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Fixed Effects

Type III Tests of Fixed Effects(a,b)

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	6221,519	64334,538	,000
mnse	1	6774,404	673,387	,000
rede_1	1	3366,385	200,677	,000
matraso	1	13818,208	86,888	,000
nse_cmj	1	18553,386	,088	,767
sexo	1	20670,310	81,449	,000
pardo	1	20810,320	10,950	,001
preto	1	20611,386	26,617	,000
atraso	1	18509,759	423,686	,000
disc	1	18571,211	1069,411	,000
atitude	1	18510,184	645,504	,000

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Estimates of Fixed Effects(a,b)

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	282,21678	1,112656	6221,519	253,643	,000	280,035600	284,397978

mnse	10,966014	,422587	6774,404	25,950	,000	10,137611	11,794418
rede_1	16,847365	1,189278	3366,385	14,166	,000	14,515585	19,179145
matraso	-4,429040	,475150	13818,208	-9,321	,000	-5,360398	-3,497681
nse_cmj	,089389	,301385	18553,386	,297	,767	-,501352	,680131
sexo	-5,586391	,618999	20670,310	-9,025	,000	-6,799677	-4,373105
pardo	-2,196970	,663911	20810,320	-3,309	,001	-3,498287	-,895653
preto	-6,961432	1,349331	20611,386	-5,159	,000	-9,606227	-4,316637
atraso	-5,486352	,266540	18509,759	-20,584	,000	-6,008794	-4,963910
disc	18,905879	,578129	18571,211	32,702	,000	17,772693	20,039064
atitude	20,203584	,795205	18510,184	25,407	,000	18,644910	21,762258

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Covariance Parameters

Estimates of Covariance Parameters(a,b)

Parameter	Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Residual	1674,6447	17,522126	95,573	,000	1640,651779	1709,342067
Intercept [subject = id] Variance	374,04491	16,936537	22,085	,000	342,280252	408,757440

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Mixed Model Analysis

2001 - 4ª série

Warnings

The covariance structure for random effect with only one level will be changed to Identity.

Model Dimension(a,b)

		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1		1	
	mnse	1		1	
	rede_1	1		1	
	matraso	1		1	
	nse_cmj	1		1	
	sexo	1		1	
	pardo	1		1	
	preto	1		1	
	atraso	1		1	
	disc	1		1	
	atitude	1		1	
Random Effects	Intercept	1	Identity	1	id
Residual				1	

Total	12	13
-------	----	----

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Information Criteria(a,b)

-2 Restricted Log Likelihood	892018,67
Akaike's Information Criterion (AIC)	892022,67
Hurvich and Tsai's Criterion (AICC)	892022,67
Bozdogan's Criterion (CAIC)	892043,39
Schwarz's Bayesian Criterion (BIC)	892041,39

The information criteria are displayed in smaller-is-better forms.

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Fixed Effects

Type III Tests of Fixed Effects(a,b)

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	8590,335	127172,67	,000
mnse	1	8668,740	2598,512	,000
rede_1	1	5996,126	548,022	,000
matraso	1	14572,023	2,125	,145
nse_cmig	1	82061,478	133,202	,000
sexo	1	84160,024	108,268	,000
pardo	1	85832,383	5,044	,025
preto	1	85418,832	444,775	,000
atraso	1	82037,558	365,309	,000
disc	1	82094,656	2328,955	,000
atitude	1	82015,803	2900,839	,000

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Estimates of Fixed Effects(a,b)

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	169,21809	,474515	8590,335	356,613	,000	168,287933	170,148260
mnse	14,543158	,285296	8668,740	50,976	,000	13,983909	15,102406
rede_1	16,938198	,723549	5996,126	23,410	,000	15,519781	18,356615
matraso	,569901	,390955	14572,023	1,458	,145	-,196421	1,336222
nse_cmig	1,612744	,139736	82061,478	11,541	,000	1,338862	1,886626

sexo	2,683656	,257915	84160,024	10,405	,000	2,178146	3,189167
pardo	,655606	,291911	85832,383	2,246	,025	,083462	1,227749
preto	-9,425204	,446911	85418,832	-21,090	,000	-10,301144	-8,549263
atraso	-3,202193	,167540	82037,558	-19,113	,000	-3,530569	-2,873817
disc	12,144364	,251648	82094,656	48,259	,000	11,651135	12,637593
atitude	23,376353	,434025	82015,803	53,859	,000	22,525667	24,227039

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Covariance Parameters

Estimates of Covariance Parameters(a,b)

Parameter	Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Residual	1340,2398	6,647322	201,621	,000	1327,274429	1353,331863
Intercept [subject = id] Variance	186,42381	5,831746	31,967	,000	175,337144	198,211495

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Mixed Model Analysis

2001 - 8ª série

Warnings

The covariance structure for random effect with only one level will be changed to Identity.

Model Dimension(a,b)

		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1	Identity	1	id
	mnse	1		1	
	rede_1	1		1	
	matraso	1		1	
	nse_cmg	1		1	
	sexo	1		1	
	pardo	1		1	
	preto	1		1	
	atraso	1		1	
	disc	1		1	
	atitude	1		1	
	Random Effects	Intercept		1	
Residual				1	
Total		12		13	

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Information Criteria(a,b)

-2 Restricted Log Likelihood	892538,39
Akaike's Information Criterion (AIC)	892542,39
Hurvich and Tsai's Criterion (AICC)	892542,39
Bozdogan's Criterion (CAIC)	892563,09
Schwarz's Bayesian Criterion (BIC)	892561,09

The information criteria are displayed in smaller-is-better forms.

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Fixed Effects

Type III Tests of Fixed Effects(a,b)

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	6225,182	188783,19	,000
mnse	1	6397,626	1338,837	,000
rede_1	1	4614,328	386,045	,000
matraso	1	9550,532	15,499	,000
nse_cmig	1	82019,671	21,313	,000
sexo	1	83887,823	113,201	,000
pardo	1	84914,416	25,477	,000
preto	1	84856,307	218,245	,000
atraso	1	81958,768	2352,085	,000
disc	1	82007,186	1517,772	,000
atitude	1	81960,546	1668,097	,000

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Estimates of Fixed Effects(a,b)

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	250,04464	,575487	6225,182	434,492	,000	248,916488	251,172796
mnse	11,960416	,326876	6397,626	36,590	,000	11,319631	12,601202
rede_1	16,602691	,845005	4614,328	19,648	,000	14,946077	18,259305
matraso	-1,242653	,315642	9550,532	-3,937	,000	-1,861379	-,623927
nse_cmig	,678616	,146996	82019,671	4,617	,000	,390505	,966726
sexo	-2,926946	,275099	83887,823	-10,640	,000	-3,466138	-2,387755
pardo	-1,523824	,301898	84914,416	-5,047	,000	-2,115542	-,932105
preto	-8,244734	,558091	84856,307	-14,773	,000	-9,338587	-7,150881
atraso	-6,735145	,138874	81958,768	-48,498	,000	-7,007336	-6,462953

disc	10,450591	,268249	82007,186	38,959	,000	9,924826	10,976356
atitude	15,955041	,390649	81960,546	40,842	,000	15,189371	16,720711

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Covariance Parameters

Estimates of Covariance Parameters(a,b)

Parameter	Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Residual	1503,7736	7,442734	202,046	,000	1489,256716	1518,432156
Intercept [subject = id] Variance	168,04050	5,859164	28,680	,000	156,940361	179,925743

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Mixed Model Analysis

2001 - 11ª série

Warnings

The covariance structure for random effect with only one level will be changed to Identity.

Model Dimension(a,b)

		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1		1	
	mnse	1		1	
	rede_1	1		1	
	matraso	1		1	
	nse_cmg	1		1	
	sexo	1		1	
	pardo	1		1	
	preto	1		1	
	atraso	1		1	
	disc	1		1	
	atitude	1		1	
Random Effects	Intercept	1	Identity	1	id
Residual				1	
Total		12		13	

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Information Criteria(a,b)

-2 Restricted Log Likelihood	676960,62
------------------------------	-----------

Akaike's Information Criterion (AIC)	676964,62
Hurvich and Tsai's Criterion (AICC)	676964,62
Bozdogan's Criterion (CAIC)	676984,74
Schwarz's Bayesian Criterion (BIC)	676982,74

The information criteria are displayed in smaller-is-better forms.

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Fixed Effects

Type III Tests of Fixed Effects(a,b)

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	4648,919	100885,66	,000
mnse	1	5135,782	1154,498	,000
rede_1	1	3174,760	188,358	,000
matraso	1	7790,465	47,128	,000
nse_cmg	1	61358,375	,192	,662
sexo	1	62949,300	488,979	,000
pardo	1	63446,742	15,190	,000
preto	1	63176,176	61,608	,000
atraso	1	61324,129	2248,707	,000
disc	1	61348,698	3672,557	,000
atitude	1	61309,563	1996,388	,000

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Estimates of Fixed Effects(a,b)

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	282,37379	,889016	4648,919	317,625	,000	280,630900	284,116687
mnse	15,066384	,443417	5135,782	33,978	,000	14,197097	15,935670
rede_1	14,409222	1,049901	3174,760	13,724	,000	12,350669	16,467775
matraso	-2,938124	,427985	7790,465	-6,865	,000	-3,777090	-2,099157
nse_cmg	,076746	,175331	61358,375	,438	,662	-,266904	,420395
sexo	-7,710235	,348676	62949,300	-22,113	,000	-8,393641	-7,026829
pardo	-1,477285	,379038	63446,742	-3,897	,000	-2,220200	-,734369
preto	-5,772591	,735451	63176,176	-7,849	,000	-7,214076	-4,331105
atraso	-6,884083	,145171	61324,129	-47,421	,000	-7,168619	-6,599548
disc	20,326207	,335407	61348,698	60,602	,000	19,668809	20,983605
atitude	19,959959	,446722	61309,563	44,681	,000	19,084383	20,835535

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Covariance Parameters

Estimates of Covariance Parameters(a,b)

Parameter	Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Residual	1724,0573	9,857973	174,890	,000	1704,843913	1743,487264
Intercept [subject = id] Variance	269,18013	10,274216	26,200	,000	249,777820	290,089582

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Mixed Model Analysis

2003 - 4ª série

Warnings

The covariance structure for random effect with only one level will be changed to Identity.

Model Dimension(a,b)

		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1	Identity	1	id
	mnse	1		1	
	rede_1	1		1	
	matraso	1		1	
	nse_cmg	1		1	
	sexo	1		1	
	pardo	1		1	
	preto	1		1	
	atraso	1		1	
	disc	1		1	
	atitude	1		1	
Random Effects	Intercept	1	Identity	1	id
Residual				1	
Total		12		13	

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Information Criteria(a,b)

-2 Restricted Log Likelihood	857591,12
Akaike's Information Criterion (AIC)	857595,12
Hurvich and Tsai's Criterion (AICC)	857595,12

Bozdogan's Criterion (CAIC)	857615,76
Schwarz's Bayesian Criterion (BIC)	857613,76

The information criteria are displayed in smaller-is-better forms.

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Fixed Effects

Type III Tests of Fixed Effects(a,b)

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	7644,220	129947,15	,000
mnse	1	7152,237	2049,501	,000
rede_1	1	5195,123	381,536	,000
matraso	1	11679,533	13,016	,000
nse_cmg	1	79129,502	161,657	,000
sexo	1	80737,608	157,578	,000
pardo	1	82432,608	2,485	,115
preto	1	81703,112	481,711	,000
atraso	1	79144,540	947,719	,000
disc	1	79164,982	925,564	,000
atitude	1	79113,844	249,621	,000

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Estimates of Fixed Effects(a,b)

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	174,41675	,483843	7644,220	360,482	,000	173,468289	175,365221
mnse	14,158630	,312750	7152,237	45,271	,000	13,545548	14,771712
rede_1	16,067241	,822571	5195,123	19,533	,000	14,454655	17,679827
matraso	1,648090	,456810	11679,533	3,608	,000	,752666	2,543515
nse_cmg	1,854581	,145864	79129,502	12,714	,000	1,568688	2,140473
sexo	3,212443	,255910	80737,608	12,553	,000	2,710860	3,714025
pardo	,453659	,287783	82432,608	1,576	,115	-,110393	1,017712
preto	-10,15402	,462641	81703,112	-21,948	,000	-11,060794	-9,247246
atraso	-5,143526	,167079	79144,540	-30,785	,000	-5,470999	-4,816053
disc	7,575195	,248995	79164,982	30,423	,000	7,087166	8,063224
atitude	6,454451	,408525	79113,844	15,799	,000	5,653744	7,255157

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Covariance Parameters

Estimates of Covariance Parameters(a,b)

Parameter	Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Residual	1276,2043	6,423908	198,665	,000	1263,675641	1288,857304
Intercept [subject = id] Variance	191,91322	6,149126	31,210	,000	180,231792	204,351769

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Mixed Model Analysis 2003 - 8ª série

Warnings

The covariance structure for random effect with only one level will be changed to Identity.

Model Dimension(a,b)

		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1		1	
	mnse	1		1	
	rede_1	1		1	
	matraso	1		1	
	nse_cmg	1		1	
	sexo	1		1	
	pardo	1		1	
	preto	1		1	
	atraso	1		1	
	disc	1		1	
	atitude	1		1	
Random Effects	Intercept	1	Identity	1	id
Residual				1	
Total		12		13	

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Information Criteria(a,b)

-2 Restricted Log Likelihood	672979,55
Akaike's Information Criterion (AIC)	672983,55
Hurvich and Tsai's Criterion (AICC)	672983,55
Bozdogan's Criterion (CAIC)	673003,68
Schwarz's Bayesian Criterion (BIC)	673001,68

The information criteria are displayed in smaller-is-better forms.

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Fixed Effects

Type III Tests of Fixed Effects(a,b)

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	5257,840	139980,55	,000
mnse	1	4874,073	890,999	,000
rede_1	1	3522,078	264,630	,000
matraso	1	8264,666	1,973	,160
nse_cmg	1	61453,277	138,468	,000
sexo	1	62822,105	1,661	,197
pardo	1	63731,970	2,568	,109
preto	1	63502,915	186,515	,000
atraso	1	61388,026	2112,410	,000
disc	1	61478,505	2174,375	,000
atitude	1	61402,420	1322,538	,000

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Estimates of Fixed Effects(a,b)

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	243,29686	,650283	5257,840	374,140	,000	242,022040	244,571691
mnse	12,932584	,433258	4874,073	29,850	,000	12,083203	13,781965
rede_1	18,051287	1,109657	3522,078	16,267	,000	15,875652	20,226922
matraso	-,617495	,439567	8264,666	-1,405	,160	-1,479156	,244166
nse_cmg	2,184929	,185679	61453,277	11,767	,000	1,820997	2,548860
sexo	,421228	,326842	62822,105	1,289	,197	-,219383	1,061840
pardo	-,584020	,364469	63731,970	-1,602	,109	-1,298380	,130339
preto	-8,899128	,651615	63502,915	-13,657	,000	-10,176293	-7,621962
atraso	-8,006423	,174201	61388,026	-45,961	,000	-8,347857	-7,664990
disc	14,840080	,318250	61478,505	46,630	,000	14,216309	15,463852
atitude	16,294282	,448055	61402,420	36,367	,000	15,416093	17,172470

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Covariance Parameters

Estimates of Covariance Parameters(a,b)

Parameter	Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound

Residual		1581,1278	9,038826	174,926	,000	1563,510934	1598,943223
Intercept [subject = id]	Variance	185,76612	7,358496	25,245	,000	171,889387	200,763145

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Mixed Model Analysis 2003 - 11ª série

Warnings

The covariance structure for random effect with only one level will be changed to Identity.

Model Dimension(a,b)

		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1		1	
	mnse	1		1	
	rede_1	1		1	
	matraso	1		1	
	nse_cmg	1		1	
	sexo	1		1	
	pardo	1		1	
	preto	1		1	
	atraso	1		1	
	disc	1		1	
	atitude	1		1	
Random Effects	Intercept	1	Identity	1	id
Residual				1	
Total		12		13	

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Information Criteria(a,b)

-2 Restricted Log Likelihood	474095,80
Akaike's Information Criterion (AIC)	474099,80
Hurvich and Tsai's Criterion (AICC)	474099,80
Bozdogan's Criterion (CAIC)	474119,21
Schwarz's Bayesian Criterion (BIC)	474117,21

The information criteria are displayed in smaller-is-better forms.

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Fixed Effects

Type III Tests of Fixed Effects(a,b)

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	3006,528	77200,787	,000
mnse	1	3595,840	651,174	,000
rede_1	1	2187,751	153,777	,000
matraso	1	6929,381	41,158	,000
nse_cmg	1	42826,718	23,819	,000
sexo	1	43746,775	152,169	,000
pardo	1	44394,469	14,585	,000
preto	1	44155,025	55,327	,000
atraso	1	42758,223	1540,236	,000
disc	1	42802,849	2473,303	,000
atitude	1	42777,106	1617,921	,000

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Estimates of Fixed Effects(a,b)

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	284,34721	1,023383	3006,528	277,850	,000	282,340615	286,353818
mnse	14,950352	,585872	3595,840	25,518	,000	13,801677	16,099027
rede_1	17,293924	1,394593	2187,751	12,401	,000	14,559060	20,028788
matraso	-4,049472	,631208	6929,381	-6,415	,000	-5,286833	-2,812111
nse_cmg	1,075211	,220311	42826,718	4,880	,000	,643397	1,507024
sexo	-5,124373	,415411	43746,775	-12,336	,000	-5,938585	-4,310160
pardo	-1,802834	,472068	44394,469	-3,819	,000	-2,728095	-,877573
preto	-6,439260	,865698	44155,025	-7,438	,000	-8,136043	-4,742477
atraso	-8,180428	,208441	42758,223	-39,246	,000	-8,588976	-7,771880
disc	19,997707	,402107	42802,849	49,732	,000	19,209570	20,785844
atitude	19,725968	,490410	42777,106	40,223	,000	18,764755	20,687182

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

Covariance Parameters

Estimates of Covariance Parameters(a,b)

Parameter	Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Residual	1754,3989	12,011578	146,059	,000	1731,013982	1778,099914
Intercept [subject = id] Variance	291,64145	13,428513	21,718	,000	266,474729	319,185012

a Dependent Variable: Proficiência.

b Residual is weighted by Peso padronizado por rede.

7.5 Saídas do HLM para o Modelo Base – racial

- raça_1995_08

Program: HLM 5 Hierarchical Linear and Nonlinear Modeling
Authors: Stephen Raudenbush, Tony Bryk, & Richard Congdon
Publisher: Scientific Software International, Inc. (c) 2000
techsupport@ssicentral.com
www.ssicentral.com

Module: HLM2.EXE (5.04.21205.1)
Date: 29 May 2004, Saturday
Time: 12:58:16

SPECIFICATIONS FOR THIS HLM2 RUN Sat May 29 12:58:16 2004

Problem Title: NO TITLE

The data source for this run = RAÇA_1995_08.SSM
The command file for this run = C:\Renato\UFMG\Doutorado\Computacional\Efeito raça\HLM\Modelo base - racial II.hlm
Output file name = C:\RENATO\UFMG\DOCTORADO\COMPUTACIONAL\EFEITO RAÇA\HLM\RAÇA_1995_08.OUT
The maximum number of level-2 units = 1056
The maximum number of iterations = 100
Method of estimation: restricted maximum likelihood

Weighting Specification

Weight
Variable
Weighting? Name Normalized?
Level 1 yes PESO_R no
Level 2 no no
Generalizations are at level-1

The outcome variable is PROFIC

The model specified for the fixed effects was:

Level-1 Level-2
Coefficients Predictors

INTRCPT1, B0 INTRCPT2, G00
 NSE_M, G01
 REDE_1, G02
 MATRASO, G03
DISC slope, B1 INTRCPT2, G10

```

# SEXO slope, B2 INTRCPT2, G20
# NSE_CMG slope, B3 INTRCPT2, G30
  PRETO slope, B4 INTRCPT2, G40
  PARDO slope, B5 INTRCPT2, G50
# ATRASO slope, B6 INTRCPT2, G60
# ATITUDE slope, B7 INTRCPT2, G70

```

'#' - The residual parameter variance for this level-1 coefficient has been set to zero.

The model specified for the covariance components was:

Sigma squared (constant across level-2 units)

Tau dimensions
 INTRCPT1
 PRETO slope
 PARDO slope

Summary of the model specified (in equation format)

Level-1 Model

$$Y = B0 + B1*(DISC) + B2*(SEXO) + B3*(NSE_CMG) + B4*(PRETO) + B5*(PARDO) + B6*(ATRASO) + B7*(ATITUDE) + R$$

Level-2 Model

$$\begin{aligned}
 B0 &= G00 + G01*(NSE_M) + G02*(REDE_1) + G03*(MATRASO) + U0 \\
 B1 &= G10 \\
 B2 &= G20 \\
 B3 &= G30 \\
 B4 &= G40 + U4 \\
 B5 &= G50 + U5 \\
 B6 &= G60 \\
 B7 &= G70
 \end{aligned}$$

The outcome variable is PROFIC

Final estimation of fixed effects:

Fixed Effect	Standard Coefficient	Standard Error	Approx. T-ratio	d.f.	P-value

For INTRCPT1, B0					
INTRCPT2, G00	265.180149	1.142153	232.176	1052	0.000
NSE_M, G01	23.056607	1.140239	20.221	1052	0.000
REDE_1, G02	-5.498139	2.034467	-2.702	1052	0.007
MATRASO, G03	-3.648199	0.556907	-6.551	1052	0.000
For DISC slope, B1					
INTRCPT2, G10	0.122492	0.416643	0.294	28847	0.769
For SEXO slope, B2					
INTRCPT2, G20	-4.393129	0.430655	-10.201	28847	0.000
For NSE_CMG slope, B3					
INTRCPT2, G30	1.892353	0.263078	7.193	28847	0.000
For PRETO slope, B4					
INTRCPT2, G40	-5.821077	0.905746	-6.427	1055	0.000
For PARDO slope, B5					
INTRCPT2, G50	-0.538329	0.510776	-1.054	1055	0.292
For ATRASO slope, B6					
INTRCPT2, G60	-4.635355	0.164793	-28.128	28847	0.000
For ATITUDE slope, B7					
INTRCPT2, G70	10.002496	0.638976	15.654	28847	0.000

The outcome variable is PROFIC

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Standard Coefficient	Standard Error	Approx. T-ratio	d.f.	P-value

For INTRCPT1, B0					
INTRCPT2, G00	265.180149	1.182309	224.290	1052	0.000
NSE_M, G01	23.056607	1.137738	20.265	1052	0.000
REDE_1, G02	-5.498139	2.117731	-2.596	1052	0.010
MATRASO, G03	-3.648199	0.595572	-6.126	1052	0.000
For DISC slope, B1					
INTRCPT2, G10	0.122492	0.494829	0.248	28847	0.805
For SEXO slope, B2					
INTRCPT2, G20	-4.393129	0.522295	-8.411	28847	0.000
For NSE_CMG slope, B3					
INTRCPT2, G30	1.892353	0.320819	5.899	28847	0.000
For PRETO slope, B4					
INTRCPT2, G40	-5.821077	0.895403	-6.501	1055	0.000
For PARDO slope, B5					
INTRCPT2, G50	-0.538329	0.496445	-1.084	1055	0.279
For ATRASO slope, B6					
INTRCPT2, G60	-4.635355	0.299399	-15.482	28847	0.000
For ATITUDE slope, B7					
INTRCPT2, G70	10.002496	0.764063	13.091	28847	0.000

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, U0	15.57291	242.51564	776	2380.97885	0.000
PRETO slope, U4	7.11524	50.62668	779	789.03503	0.394
PARDO slope, U5	3.18165	10.12289	779	742.45375	>.500
level-1, R	39.61358	1569.23598			

Note: The chi-square statistics reported above are based on only 780 of 1056 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Statistics for current covariance components model

Deviance = 295995.967774
Number of estimated parameters = 7

- raça_1995_11

Program: HLM 5 Hierarchical Linear and Nonlinear Modeling
Authors: Stephen Raudenbush, Tony Bryk, & Richard Congdon
Publisher: Scientific Software International, Inc. (c) 2000
techsupport@ssicentral.com
www.ssicentral.com

Module: HLM2.EXE (5.04.21205.1)
Date: 29 May 2004, Saturday
Time: 13:24:23

SPECIFICATIONS FOR THIS HLM2 RUN Sat May 29 13:24:22 2004

Problem Title: NO TITLE

The data source for this run = raça_1995_11.ssm
The command file for this run = C:\Renato\UFMG\Doutorado\Computacional\Efeito raça\HLM - modelo racial II\Modelo base - racial II - 1995_11.hlm
Output file name = C:\Renato\UFMG\Doutorado\Computacional\Efeito raça\HLM - modelo racial II\raça_1995_11.out
The maximum number of level-2 units = 681
The maximum number of iterations = 100
Method of estimation: restricted maximum likelihood

Weighting Specification

Weight

	Variable	Weighting?	Name	Normalized?
Level 1	yes	PESO_R	no	
Level 2	no		no	

Generalizations are at level-1

The outcome variable is PROFIC

The model specified for the fixed effects was:

```

-----
Level-1          Level-2
Coefficients     Predictors
-----
      INTRCPT1, B0  INTRCPT2, G00
                   NSE_M, G01
                   REDE_1, G02
                   MATRASO, G03
#  DISC slope, B1  INTRCPT2, G10
#  SEXO slope, B2  INTRCPT2, G20
#  NSE_CMG slope, B3  INTRCPT2, G30
  PRETO slope, B4  INTRCPT2, G40
  PARDO slope, B5  INTRCPT2, G50
#  ATRASO slope, B6  INTRCPT2, G60
#  ATITUDE slope, B7  INTRCPT2, G70

```

'#' - The residual parameter variance for this level-1 coefficient has been set to zero.

The model specified for the covariance components was:

Sigma squared (constant across level-2 units)

Tau dimensions
 INTRCPT1
 PRETO slope
 PARDO slope

Summary of the model specified (in equation format)

Level-1 Model

$$Y = B0 + B1*(DISC) + B2*(SEXO) + B3*(NSE_CMG) + B4*(PRETO) + B5*(PARDO) + B6*(ATRASO) + B7*(ATITUDE) + R$$

Level-2 Model

$$B0 = G00 + G01*(NSE_M) + G02*(REDE_1) + G03*(MATRASO) + U0$$

$$B1 = G10$$

$$B2 = G20$$

$$B3 = G30$$

$$B4 = G40 + U4$$

$$B5 = G50 + U5$$

$$B6 = G60$$

$$B7 = G70$$

Final estimation of fixed effects:

Fixed Effect	Standard Coefficient	Approx. Error	T-ratio	d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	302.203630	2.026729	149.109	677	0.000
NSE_M, G01	23.782644	1.620344	14.678	677	0.000
REDE_1, G02	-2.287814	2.210657	-1.035	677	0.301
MATRASO, G03	-4.088641	0.977696	-4.182	677	0.000
For DISC slope, B1					
INTRCPT2, G10	-3.329927	0.558391	-5.963	17957	0.000
For SEXO slope, B2					
INTRCPT2, G20	-8.095584	0.620684	-13.043	17957	0.000
For NSE_CMG slope, B3					
INTRCPT2, G30	0.816629	0.342451	2.385	17957	0.017
For PRETO slope, B4					
INTRCPT2, G40	-3.314265	1.257666	-2.635	680	0.009
For PARDO slope, B5					
INTRCPT2, G50	-0.387330	0.670156	-0.578	680	0.563
For ATRASO slope, B6					
INTRCPT2, G60	-3.587043	0.196158	-18.287	17957	0.000
For ATITUDE slope, B7					
INTRCPT2, G70	9.061435	0.808343	11.210	17957	0.000

The outcome variable is PROFIC

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Standard Coefficient	Approx. Error	T-ratio	d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	302.203630	1.835968	164.602	677	0.000
NSE_M, G01	23.782644	1.869612	12.721	677	0.000
REDE_1, G02	-2.287814	2.320786	-0.986	677	0.325
MATRASO, G03	-4.088641	0.962840	-4.246	677	0.000
For DISC slope, B1					
INTRCPT2, G10	-3.329927	0.726049	-4.586	17957	0.000
For SEXO slope, B2					
INTRCPT2, G20	-8.095584	0.682673	-11.859	17957	0.000
For NSE_CMG slope, B3					
INTRCPT2, G30	0.816629	0.365983	2.231	17957	0.026
For PRETO slope, B4					
INTRCPT2, G40	-3.314265	1.159617	-2.858	680	0.005
For PARDO slope, B5					
INTRCPT2, G50	-0.387330	0.654775	-0.592	680	0.554
For ATRASO slope, B6					
INTRCPT2, G60	-3.587043	0.246624	-14.545	17957	0.000
For ATITUDE slope, B7					
INTRCPT2, G70	9.061435	0.836156	10.837	17957	0.000

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, U0	17.65489	311.69513	431	1552.32870	0.000
PRETO slope, U4	3.38168	11.43574	434	406.47146	>.500
PARDO slope, U5	2.18902	4.79180	434	405.11397	>.500
level-1, R	42.52332	1808.23235			

Note: The chi-square statistics reported above are based on only 435 of 681 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Statistics for current covariance components model

Deviance = 186859.591654
Number of estimated parameters = 7

- raça_1997_04

Program: HLM 5 Hierarchical Linear and Nonlinear Modeling
Authors: Stephen Raudenbush, Tony Bryk, & Richard Congdon
Publisher: Scientific Software International, Inc. (c) 2000
techsupport@ssicentral.com
www.ssicentral.com

Module: HLM2.EXE (5.04.21205.1)
Date: 29 May 2004, Saturday
Time: 13:34:10

SPECIFICATIONS FOR THIS HLM2 RUN Sat May 29 13:34:10 2004

Problem Title: NO TITLE

The data source for this run = raça_1997_04.ssm
The command file for this run = C:\Renato\UFMG\Doutorado\Computacional\Efeito raça\HLM - modelo racial II\Modelo base - racial II - 1997_04.hlm
Output file name = C:\Renato\UFMG\Doutorado\Computacional\Efeito raça\HLM - modelo racial II\raça_1997_04.out
The maximum number of level-2 units = 2538
The maximum number of iterations = 100
Method of estimation: restricted maximum likelihood

Weighting Specification

Weight
Variable

	Weighting?	Name	Normalized?
Level 1	yes	PESO_R	no
Level 2	no		no

Generalizations are at level-1

The outcome variable is PROFIC

The model specified for the fixed effects was:

```

-----
Level-1          Level-2
Coefficients      Predictors
-----
      INTRCPT1, B0  INTRCPT2, G00
                   NSE_M, G01
                   REDE_1, G02
                   MATRASO, G03
#  DISC slope, B1  INTRCPT2, G10
#  SEXO slope, B2  INTRCPT2, G20
#  NSE_CMG slope, B3  INTRCPT2, G30
      PRETO slope, B4  INTRCPT2, G40
      PARDO slope, B5  INTRCPT2, G50
#  ATRASO slope, B6  INTRCPT2, G60
#  ATITUDE slope, B7  INTRCPT2, G70

```

'#' - The residual parameter variance for this level-1 coefficient has been set to zero.

The model specified for the covariance components was:

Sigma squared (constant across level-2 units)

Tau dimensions
 INTRCPT1
 PRETO slope
 PARDO slope

Summary of the model specified (in equation format)

Level-1 Model

$$Y = B0 + B1*(DISC) + B2*(SEXO) + B3*(NSE_CMG) + B4*(PRETO) + B5*(PARDO) + B6*(ATRASO) + B7*(ATITUDE) + R$$

Level-2 Model

$$\begin{aligned}
 B0 &= G00 + G01*(NSE_M) + G02*(REDE_1) + G03*(MATRASO) + U0 \\
 B1 &= G10 \\
 B2 &= G20 \\
 B3 &= G30 \\
 B4 &= G40 + U4 \\
 B5 &= G50 + U5 \\
 B6 &= G60 \\
 B7 &= G70
 \end{aligned}$$

Final estimation of fixed effects:

Fixed Effect	Standard Coefficient	Approx. Error	T-ratio	d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	180.328174	0.750951	240.133	2534	0.000
NSE_M, G01	23.605448	0.754404	31.290	2534	0.000
REDE_1, G02	11.786527	1.167804	10.093	2534	0.000
MATRASO, G03	3.101784	0.435831	7.117	2534	0.000
For DISC slope, B1					
INTRCPT2, G10	5.197584	0.357580	14.535	40385	0.000
For SEXO slope, B2					
INTRCPT2, G20	0.169806	0.365999	0.464	40385	0.642
For NSE_CMG slope, B3					
INTRCPT2, G30	1.458195	0.226292	6.444	40385	0.000
For PRETO slope, B4					
INTRCPT2, G40	-7.935218	0.741350	-10.704	2537	0.000
For PARDO slope, B5					
INTRCPT2, G50	1.267690	0.470755	2.693	2537	0.007
For ATRASO slope, B6					
INTRCPT2, G60	-2.684151	0.170755	-15.719	40385	0.000
For ATITUDE slope, B7					
INTRCPT2, G70	21.636228	0.801331	27.000	40385	0.000

The outcome variable is PROFIC

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Standard Coefficient	Approx. Error	T-ratio	d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	180.328174	0.811168	222.307	2534	0.000
NSE_M, G01	23.605448	0.815007	28.963	2534	0.000
REDE_1, G02	11.786527	1.356650	8.688	2534	0.000
MATRASO, G03	3.101784	0.458520	6.765	2534	0.000
For DISC slope, B1					
INTRCPT2, G10	5.197584	0.470528	11.046	40385	0.000
For SEXO slope, B2					
INTRCPT2, G20	0.169806	0.458478	0.370	40385	0.711
For NSE_CMG slope, B3					
INTRCPT2, G30	1.458195	0.296296	4.921	40385	0.000
For PRETO slope, B4					
INTRCPT2, G40	-7.935218	0.744745	-10.655	2537	0.000
For PARDO slope, B5					
INTRCPT2, G50	1.267690	0.468952	2.703	2537	0.007
For ATRASO slope, B6					
INTRCPT2, G60	-2.684151	0.240294	-11.170	40385	0.000
For ATITUDE slope, B7					
INTRCPT2, G70	21.636228	1.073655	20.152	40385	0.000

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, U0	14.70841	216.33737	1707	3530.67908	0.000
PRETO slope, U4	10.87337	118.23012	1710	1490.49670	>.500
PARDO slope, U5	10.07192	101.44348	1710	1704.06697	>.500
level-1, R	35.26283	1243.46699			

Note: The chi-square statistics reported above are based on only 1711 of 2538 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Statistics for current covariance components model

Deviance = 405822.311230
Number of estimated parameters = 7

- raça_1997_08

Program: HLM 5 Hierarchical Linear and Nonlinear Modeling
Authors: Stephen Raudenbush, Tony Bryk, & Richard Congdon
Publisher: Scientific Software International, Inc. (c) 2000
techsupport@ssicentral.com
www.ssicentral.com

Module: HLM2.EXE (5.04.21205.1)
Date: 29 May 2004, Saturday
Time: 13:39:17

SPECIFICATIONS FOR THIS HLM2 RUN Sat May 29 13:39:17 2004

Problem Title: NO TITLE

The data source for this run = raça_1997_08.ssm
The command file for this run = C:\Renato\UFMG\Doutorado\Computacional\Efeito raça\HLM - modelo racial II\Modelo base - racial II - 1997_08.hlm
Output file name = C:\Renato\UFMG\Doutorado\Computacional\Efeito raça\HLM - modelo racial II\raça_1997_08.out
The maximum number of level-2 units = 1871
The maximum number of iterations = 100
Method of estimation: restricted maximum likelihood

Weighting Specification

Weight
Variable

	Weighting?	Name	Normalized?
Level 1	yes	PESO_R	no
Level 2	no		no

Generalizations are at level-1

The outcome variable is PROFIC

The model specified for the fixed effects was:

```

-----
Level-1          Level-2
Coefficients     Predictors
-----
      INTRCPT1, B0  INTRCPT2, G00
                   NSE_M, G01
                   REDE_1, G02
                   MATRASO, G03
#  DISC slope, B1  INTRCPT2, G10
#  SEXO slope, B2  INTRCPT2, G20
#  NSE_CMG slope, B3  INTRCPT2, G30
      PRETO slope, B4  INTRCPT2, G40
      PARDO slope, B5  INTRCPT2, G50
#  ATRASO slope, B6  INTRCPT2, G60
#  ATITUDE slope, B7  INTRCPT2, G70

```

'#' - The residual parameter variance for this level-1 coefficient has been set to zero.

The model specified for the covariance components was:

Sigma squared (constant across level-2 units)

Tau dimensions
 INTRCPT1
 PRETO slope
 PARDO slope

Summary of the model specified (in equation format)

Level-1 Model

$$Y = B0 + B1*(DISC) + B2*(SEXO) + B3*(NSE_CMG) + B4*(PRETO) + B5*(PARDO) + B6*(ATRASO) + B7*(ATITUDE) + R$$

Level-2 Model

$$\begin{aligned}
 B0 &= G00 + G01*(NSE_M) + G02*(REDE_1) + G03*(MATRASO) + U0 \\
 B1 &= G10 \\
 B2 &= G20 \\
 B3 &= G30 \\
 B4 &= G40 + U4 \\
 B5 &= G50 + U5 \\
 B6 &= G60 \\
 B7 &= G70
 \end{aligned}$$

Final estimation of fixed effects:

Fixed Effect	Standard Coefficient	Standard Error	Approx. T-ratio	d.f.	P-value

For INTRCPT1, B0					
INTRCPT2, G00	256.093819	0.902799	283.667	1867	0.000
NSE_M, G01	19.650269	0.792347	24.800	1867	0.000
REDE_1, G02	9.501983	1.365078	6.961	1867	0.000
MATRASO, G03	-0.648185	0.475886	-1.362	1867	0.173
For DISC slope, B1					
INTRCPT2, G10	1.326410	0.332620	3.988	37625	0.000
For SEXO slope, B2					
INTRCPT2, G20	-2.724911	0.346810	-7.857	37625	0.000
For NSE_CMG slope, B3					
INTRCPT2, G30	0.729064	0.199330	3.658	37625	0.000
For PRETO slope, B4					
INTRCPT2, G40	-4.541480	0.775132	-5.859	1870	0.000
For PARD0 slope, B5					
INTRCPT2, G50	-1.210180	0.381132	-3.175	1870	0.002
For ATRASO slope, B6					
INTRCPT2, G60	-3.180507	0.149653	-21.253	37625	0.000
For ATITUDE slope, B7					
INTRCPT2, G70	8.981122	0.483110	18.590	37625	0.000

The outcome variable is PROFIC

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Standard Coefficient	Standard Error	Approx. T-ratio	d.f.	P-value

For INTRCPT1, B0					
INTRCPT2, G00	256.093819	0.901946	283.935	1867	0.000
NSE_M, G01	19.650269	0.833787	23.567	1867	0.000
REDE_1, G02	9.501983	1.523297	6.238	1867	0.000
MATRASO, G03	-0.648185	0.469020	-1.382	1867	0.167
For DISC slope, B1					
INTRCPT2, G10	1.326410	0.393581	3.370	37625	0.001
For SEXO slope, B2					
INTRCPT2, G20	-2.724911	0.375607	-7.255	37625	0.000
For NSE_CMG slope, B3					
INTRCPT2, G30	0.729064	0.223914	3.256	37625	0.002
For PRETO slope, B4					
INTRCPT2, G40	-4.541480	0.709513	-6.401	1870	0.000
For PARD0 slope, B5					
INTRCPT2, G50	-1.210180	0.351117	-3.447	1870	0.001
For ATRASO slope, B6					
INTRCPT2, G60	-3.180507	0.194553	-16.348	37625	0.000
For ATITUDE slope, B7					
INTRCPT2, G70	8.981122	0.598159	15.015	37625	0.000

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, U0	14.05929	197.66365	943	2379.29396	0.000
PRETO slope, U4	2.98900	8.93411	946	824.85753	>.500
PARDO slope, U5	1.46027	2.13238	946	831.43918	>.500
level-1, R	39.40794	1552.98592			

Note: The chi-square statistics reported above are based on only 947 of 1871 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Statistics for current covariance components model

Deviance = 385461.932164
Number of estimated parameters = 7

- raça_1997_11

Program: HLM 5 Hierarchical Linear and Nonlinear Modeling
Authors: Stephen Raudenbush, Tony Bryk, & Richard Congdon
Publisher: Scientific Software International, Inc. (c) 2000
techsupport@ssicentral.com
www.ssicentral.com

Module: HLM2.EXE (5.04.21205.1)
Date: 29 May 2004, Saturday
Time: 13:43:50

SPECIFICATIONS FOR THIS HLM2 RUN Sat May 29 13:43:50 2004

Problem Title: NO TITLE

The data source for this run = raça_1997_11.ssm
The command file for this run = C:\Renato\UFMG\Doutorado\Computacional\Efeito raça\HLM - modelo racial II\Modelo base - racial II - 1997_11.hlm
Output file name = C:\Renato\UFMG\Doutorado\Computacional\Efeito raça\HLM - modelo racial II\raça_1997_11.out
The maximum number of level-2 units = 1239
The maximum number of iterations = 100
Method of estimation: restricted maximum likelihood

Weighting Specification

	Weight Variable	Weighting?	Name	Normalized?
Level 1	yes	PESO_R	no	

Level 2 no no
Generalizations are at level-1

The outcome variable is PROFIC

The model specified for the fixed effects was:

Level-1 Level-2
Coefficients Predictors

 INTRCPT2, G00
INTRCPT1, B0 NSE_M, G01
 REDE_1, G02
 MATRASO, G03
DISC slope, B1 INTRCPT2, G10
SEXO slope, B2 INTRCPT2, G20
NSE_CMG slope, B3 INTRCPT2, G30
 PRETO slope, B4 INTRCPT2, G40
 PARDO slope, B5 INTRCPT2, G50
ATRASO slope, B6 INTRCPT2, G60
ATITUDE slope, B7 INTRCPT2, G70

'#' - The residual parameter variance for this level-1 coefficient has been set to zero.

The model specified for the covariance components was:

Sigma squared (constant across level-2 units)

Tau dimensions
INTRCPT1
PRETO slope
PARDO slope

Summary of the model specified (in equation format)

Level-1 Model

$$Y = B0 + B1*(DISC) + B2*(SEXO) + B3*(NSE_CMG) + B4*(PRETO) + B5*(PARDO) + B6*(ATRASO) + B7*(ATITUDE) + R$$

Level-2 Model

$$B0 = G00 + G01*(NSE_M) + G02*(REDE_1) + G03*(MATRASO) + U0$$
$$B1 = G10$$
$$B2 = G20$$
$$B3 = G30$$
$$B4 = G40 + U4$$
$$B5 = G50 + U5$$
$$B6 = G60$$
$$B7 = G70$$

Final estimation of fixed effects:

Fixed Effect	Standard Coefficient	Approx. Error	T-ratio	d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	302.987946	1.577369	192.084	1235	0.000
NSE_M, G01	20.435654	1.186527	17.223	1235	0.000
REDE_1, G02	8.858458	1.883336	4.704	1235	0.000
MATRASO, G03	-6.331178	0.861198	-7.352	1235	0.000
For DISC slope, B1					
INTRCPT2, G10	8.844134	0.528795	16.725	16272	0.000
For SEXO slope, B2					
INTRCPT2, G20	-6.446125	0.573362	-11.243	16272	0.000
For NSE_CMG slope, B3					
INTRCPT2, G30	0.799140	0.309385	2.583	16272	0.010
For PRETO slope, B4					
INTRCPT2, G40	-2.040278	1.444835	-1.412	1238	0.158
For PARDO slope, B5					
INTRCPT2, G50	0.141311	0.657447	0.215	1238	0.830
For ATRASO slope, B6					
INTRCPT2, G60	-3.137977	0.221794	-14.148	16272	0.000
For ATITUDE slope, B7					
INTRCPT2, G70	8.863466	0.770545	11.503	16272	0.000

The outcome variable is PROFIC

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Standard Coefficient	Approx. Error	T-ratio	d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	302.987946	1.642181	184.503	1235	0.000
NSE_M, G01	20.435654	1.231873	16.589	1235	0.000
REDE_1, G02	8.858458	2.055169	4.310	1235	0.000
MATRASO, G03	-6.331178	0.874366	-7.241	1235	0.000
For DISC slope, B1					
INTRCPT2, G10	8.844134	0.916580	9.649	16272	0.000
For SEXO slope, B2					
INTRCPT2, G20	-6.446125	0.685172	-9.408	16272	0.000
For NSE_CMG slope, B3					
INTRCPT2, G30	0.799140	0.307801	2.596	16272	0.010
For PRETO slope, B4					
INTRCPT2, G40	-2.040278	1.373306	-1.486	1238	0.137
For PARDO slope, B5					
INTRCPT2, G50	0.141311	0.644248	0.219	1238	0.827
For ATRASO slope, B6					
INTRCPT2, G60	-3.137977	0.290340	-10.808	16272	0.000
For ATITUDE slope, B7					
INTRCPT2, G70	8.863466	0.964515	9.190	16272	0.000

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, U0	17.89951	320.39257	379	1075.50696	0.000
PRETO slope, U4	5.29647	28.05257	382	361.68089	>.500
PARDO slope, U5	3.69662	13.66497	382	388.14578	0.403
level-1, R	41.88058	1753.98270			

Note: The chi-square statistics reported above are based on only 383 of 1239 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Statistics for current covariance components model

Deviance = 169198.526262
Number of estimated parameters = 7

- raça_1999_04

Program: HLM 5 Hierarchical Linear and Nonlinear Modeling
Authors: Stephen Raudenbush, Tony Bryk, & Richard Congdon
Publisher: Scientific Software International, Inc. (c) 2000
techsupport@ssicentral.com
www.ssicentral.com

Module: HLM2.EXE (5.04.21205.1)
Date: 29 May 2004, Saturday
Time: 13:48:58

SPECIFICATIONS FOR THIS HLM2 RUN Sat May 29 13:48:58 2004

Problem Title: NO TITLE

The data source for this run = raça_1999_04.ssm
The command file for this run = C:\Renato\UFMG\Doutorado\Computacional\Efeito raça\HLM - modelo racial II\Modelo base - racial II - 1999_04.hlm
Output file name = C:\Renato\UFMG\Doutorado\Computacional\Efeito raça\HLM - modelo racial II\raça_1999_04.out
The maximum number of level-2 units = 4929
The maximum number of iterations = 100
Method of estimation: restricted maximum likelihood

Weighting Specification

Weight
Variable

	Weighting?	Name	Normalized?
Level 1	yes	PESO_R	no
Level 2	no		no

Generalizations are at level-1

The outcome variable is PROFIC

The model specified for the fixed effects was:

```

-----
Level-1          Level-2
Coefficients     Predictors
-----
      INTRCPT1, B0  INTRCPT2, G00
                   NSE_M, G01
                   REDE_1, G02
                   MATRASO, G03
#  DISC slope, B1  INTRCPT2, G10
#  SEXO slope, B2  INTRCPT2, G20
#  NSE_CMG slope, B3  INTRCPT2, G30
      PRETO slope, B4  INTRCPT2, G40
      PARDO slope, B5  INTRCPT2, G50
#  ATRASO slope, B6  INTRCPT2, G60
#  ATITUDE slope, B7  INTRCPT2, G70

```

'#' - The residual parameter variance for this level-1 coefficient has been set to zero.

The model specified for the covariance components was:

Sigma squared (constant across level-2 units)

Tau dimensions
 INTRCPT1
 PRETO slope
 PARDO slope

Summary of the model specified (in equation format)

Level-1 Model

$$Y = B0 + B1*(DISC) + B2*(SEXO) + B3*(NSE_CMG) + B4*(PRETO) + B5*(PARDO) + B6*(ATRASO) + B7*(ATITUDE) + R$$

Level-2 Model

$$\begin{aligned}
 B0 &= G00 + G01*(NSE_M) + G02*(REDE_1) + G03*(MATRASO) + U0 \\
 B1 &= G10 \\
 B2 &= G20 \\
 B3 &= G30 \\
 B4 &= G40 + U4 \\
 B5 &= G50 + U5 \\
 B6 &= G60 \\
 B7 &= G70
 \end{aligned}$$

Final estimation of fixed effects:

Fixed Effect	Standard Coefficient	Standard Error	Approx. T-ratio	d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	171.198913	0.631151	271.249	4925	0.000
NSE_M, G01	11.833465	0.322161	36.732	4925	0.000
REDE_1, G02	13.695745	0.885049	15.475	4925	0.000
MATRASO, G03	1.259204	0.362676	3.472	4925	0.001
For DISC slope, B1					
INTRCPT2, G10	8.484294	0.283696	29.906	43017	0.000
For SEXO slope, B2					
INTRCPT2, G20	0.830243	0.296011	2.805	43017	0.005
For NSE_CMG slope, B3					
INTRCPT2, G30	0.419285	0.163974	2.557	43017	0.011
For PRETO slope, B4					
INTRCPT2, G40	-6.169979	0.506666	-12.178	4928	0.000
For PARDO slope, B5					
INTRCPT2, G50	0.090550	0.335821	0.270	4928	0.787
For ATRASO slope, B6					
INTRCPT2, G60	-2.362022	0.125799	-18.776	43017	0.000
For ATITUDE slope, B7					
INTRCPT2, G70	9.330620	0.488355	19.106	43017	0.000

The outcome variable is PROFIC

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Standard Coefficient	Standard Error	Approx. T-ratio	d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	171.198913	0.667725	256.391	4925	0.000
NSE_M, G01	11.833465	0.336128	35.205	4925	0.000
REDE_1, G02	13.695745	0.958895	14.283	4925	0.000
MATRASO, G03	1.259204	0.378751	3.325	4925	0.001
For DISC slope, B1					
INTRCPT2, G10	8.484294	0.323174	26.253	43017	0.000
For SEXO slope, B2					
INTRCPT2, G20	0.830243	0.303595	2.735	43017	0.007
For NSE_CMG slope, B3					
INTRCPT2, G30	0.419285	0.171987	2.438	43017	0.015
For PRETO slope, B4					
INTRCPT2, G40	-6.169979	0.466874	-13.216	4928	0.000
For PARDO slope, B5					
INTRCPT2, G50	0.090550	0.326596	0.277	4928	0.782
For ATRASO slope, B6					
INTRCPT2, G60	-2.362022	0.137475	-17.181	43017	0.000
For ATITUDE slope, B7					
INTRCPT2, G70	9.330620	0.528270	17.663	43017	0.000

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, U0	14.40172	207.40944	2132	3734.20151	0.000
PRETO slope, U4	3.11822	9.72330	2135	1751.05062	>.500
PARDO slope, U5	2.55152	6.51028	2135	1800.00936	>.500
level-1, R	34.17734	1168.09050			

Note: The chi-square statistics reported above are based on only 2136 of 4929 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Statistics for current covariance components model

Deviance = 430092.327066
Number of estimated parameters = 7

- raça_1999_08

Program: HLM 5 Hierarchical Linear and Nonlinear Modeling
Authors: Stephen Raudenbush, Tony Bryk, & Richard Congdon
Publisher: Scientific Software International, Inc. (c) 2000
techsupport@ssicentral.com
www.ssicentral.com

Module: HLM2.EXE (5.04.21205.1)
Date: 29 May 2004, Saturday
Time: 13:53:57

SPECIFICATIONS FOR THIS HLM2 RUN Sat May 29 13:53:57 2004

Problem Title: NO TITLE

The data source for this run = raça_1999_08.ssm
The command file for this run = C:\Renato\UFMG\Doutorado\Computacional\Efeito raça\HLM - modelo racial II\Modelo base - racial II - 1999_08.hlm
Output file name = C:\Renato\UFMG\Doutorado\Computacional\Efeito raça\HLM - modelo racial II\raça_1999_08.out
The maximum number of level-2 units = 3413
The maximum number of iterations = 100
Method of estimation: restricted maximum likelihood

Weighting Specification

Weight
Variable

	Weighting?	Name	Normalized?
Level 1	yes	PESO_R	no
Level 2	no		no

Generalizations are at level-1

The outcome variable is PROFIC

The model specified for the fixed effects was:

```

-----
Level-1          Level-2
Coefficients      Predictors
-----
      INTRCPT1, B0  INTRCPT2, G00
                    NSE_M, G01
                    REDE_1, G02
                    MATRASO, G03
#   DISC slope, B1  INTRCPT2, G10
#   SEXO slope, B2  INTRCPT2, G20
#   NSE_CMG slope, B3  INTRCPT2, G30
      PRETO slope, B4  INTRCPT2, G40
      PARDO slope, B5  INTRCPT2, G50
#   ATRASO slope, B6  INTRCPT2, G60
#   ATITUDE slope, B7  INTRCPT2, G70

```

'#' - The residual parameter variance for this level-1 coefficient has been set to zero.

The model specified for the covariance components was:

Sigma squared (constant across level-2 units)

Tau dimensions
 INTRCPT1
 PRETO slope
 PARDO slope

Summary of the model specified (in equation format)

Level-1 Model

$$Y = B0 + B1*(DISC) + B2*(SEXO) + B3*(NSE_CMG) + B4*(PRETO) + B5*(PARDO) + B6*(ATRASO) + B7*(ATITUDE) + R$$

Level-2 Model

$$\begin{aligned}
 B0 &= G00 + G01*(NSE_M) + G02*(REDE_1) + G03*(MATRASO) + U0 \\
 B1 &= G10 \\
 B2 &= G20 \\
 B3 &= G30 \\
 B4 &= G40 + U4 \\
 B5 &= G50 + U5 \\
 B6 &= G60 \\
 B7 &= G70
 \end{aligned}$$

Final estimation of fixed effects:

Fixed Effect	Standard Coefficient	Approx. Error	T-ratio	d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	249.580461	0.783593	318.508	3409	0.000
NSE_M, G01	11.619236	0.399873	29.057	3409	0.000
REDE_1, G02	8.964417	1.068700	8.388	3409	0.000
MATRASO, G03	-3.425570	0.406851	-8.420	3409	0.000
For DISC slope, B1					
INTRCPT2, G10	12.255796	0.362977	33.765	35780	0.000
For SEXO slope, B2					
INTRCPT2, G20	-2.705690	0.381694	-7.089	35780	0.000
For NSE_CMG slope, B3					
INTRCPT2, G30	0.100887	0.194951	0.517	35780	0.604
For PRETO slope, B4					
INTRCPT2, G40	-3.694716	0.770269	-4.797	3412	0.000
For PARDO slope, B5					
INTRCPT2, G50	-1.640500	0.420245	-3.904	3412	0.000
For ATRASO slope, B6					
INTRCPT2, G60	-4.325014	0.175287	-24.674	35780	0.000
For ATITUDE slope, B7					
INTRCPT2, G70	15.087792	0.529956	28.470	35780	0.000

The outcome variable is PROFIC

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Standard Coefficient	Approx. Error	T-ratio	d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	249.580461	0.761015	327.957	3409	0.000
NSE_M, G01	11.619236	0.438116	26.521	3409	0.000
REDE_1, G02	8.964417	1.080883	8.294	3409	0.000
MATRASO, G03	-3.425570	0.407658	-8.403	3409	0.000
For DISC slope, B1					
INTRCPT2, G10	12.255796	0.419601	29.208	35780	0.000
For SEXO slope, B2					
INTRCPT2, G20	-2.705690	0.391619	-6.909	35780	0.000
For NSE_CMG slope, B3					
INTRCPT2, G30	0.100887	0.204894	0.492	35780	0.622
For PRETO slope, B4					
INTRCPT2, G40	-3.694716	0.729502	-5.065	3412	0.000
For PARDO slope, B5					
INTRCPT2, G50	-1.640500	0.406174	-4.039	3412	0.000
For ATRASO slope, B6					
INTRCPT2, G60	-4.325014	0.201181	-21.498	35780	0.000
For ATITUDE slope, B7					
INTRCPT2, G70	15.087792	0.583869	25.841	35780	0.000

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, U0	14.77882	218.41346	1297	2106.36093	0.000
PRETO slope, U4	4.55648	20.76153	1300	1129.52110	>.500
PARDO slope, U5	2.41240	5.81967	1300	1155.04417	>.500
level-1, R	37.91305	1437.39911			

Note: The chi-square statistics reported above are based on only 1301 of 3413 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Statistics for current covariance components model

Deviance = 364799.255457
Number of estimated parameters = 7

- raça_1999_11

Program: HLM 5 Hierarchical Linear and Nonlinear Modeling
Authors: Stephen Raudenbush, Tony Bryk, & Richard Congdon
Publisher: Scientific Software International, Inc. (c) 2000
techsupport@ssicentral.com
www.ssicentral.com

Module: HLM2.EXE (5.04.21205.1)
Date: 29 May 2004, Saturday
Time: 13:58:27

SPECIFICATIONS FOR THIS HLM2 RUN Sat May 29 13:58:27 2004

Problem Title: NO TITLE

The data source for this run = raça_1999_11.ssm
The command file for this run = C:\Renato\UFMG\Doutorado\Computacional\Efeito raça\HLM - modelo racial II\Modelo base - racial II - 1999_11.hlm
Output file name = C:\Renato\UFMG\Doutorado\Computacional\Efeito raça\HLM - modelo racial II\raça_1999_11.out
The maximum number of level-2 units = 2988
The maximum number of iterations = 100
Method of estimation: restricted maximum likelihood

Weighting Specification

Weight
Variable

	Weighting?	Name	Normalized?
Level 1	yes	PESO_R	no
Level 2	no		no

Generalizations are at level-1

The outcome variable is PROFIC

The model specified for the fixed effects was:

```

-----
Level-1          Level-2
Coefficients      Predictors
-----
      INTRCPT1, B0  INTRCPT2, G00
                    NSE_M, G01
                    REDE_1, G02
                    MATRASO, G03
#  DISC slope, B1  INTRCPT2, G10
#  SEXO slope, B2  INTRCPT2, G20
#  NSE_CMG slope, B3  INTRCPT2, G30
      PRETO slope, B4  INTRCPT2, G40
      PARDO slope, B5  INTRCPT2, G50
#  ATRASO slope, B6  INTRCPT2, G60
#  ATITUDE slope, B7  INTRCPT2, G70

```

'#' - The residual parameter variance for this level-1 coefficient has been set to zero.

The model specified for the covariance components was:

Sigma squared (constant across level-2 units)

Tau dimensions
 INTRCPT1
 PRETO slope
 PARDO slope

Summary of the model specified (in equation format)

Level-1 Model

$$Y = B0 + B1*(DISC) + B2*(SEXO) + B3*(NSE_CMG) + B4*(PRETO) + B5*(PARDO) + B6*(ATRASO) + B7*(ATITUDE) + R$$

Level-2 Model

$$\begin{aligned}
 B0 &= G00 + G01*(NSE_M) + G02*(REDE_1) + G03*(MATRASO) + U0 \\
 B1 &= G10 \\
 B2 &= G20 \\
 B3 &= G30 \\
 B4 &= G40 + U4 \\
 B5 &= G50 + U5 \\
 B6 &= G60 \\
 B7 &= G70
 \end{aligned}$$

Final estimation of fixed effects:

Fixed Effect	Standard Coefficient	Approx. Error	T-ratio	d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	288.226971	1.137639	253.355	2984	0.000
NSE_M, G01	14.614002	0.514106	28.426	2984	0.000
REDE_1, G02	8.659987	1.189095	7.283	2984	0.000
MATRASO, G03	-7.023047	0.539677	-13.013	2984	0.000
For DISC slope, B1					
INTRCPT2, G10	15.722431	0.497214	31.621	23653	0.000
For SEXO slope, B2					
INTRCPT2, G20	-4.521789	0.536123	-8.434	23653	0.000
For NSE_CMG slope, B3					
INTRCPT2, G30	0.015781	0.258717	0.061	23653	0.952
For PRETO slope, B4					
INTRCPT2, G40	-4.648486	1.155181	-4.024	2987	0.000
For PARDO slope, B5					
INTRCPT2, G50	-1.008024	0.577443	-1.746	2987	0.080
For ATRASO slope, B6					
INTRCPT2, G60	-4.302193	0.209069	-20.578	23653	0.000
For ATITUDE slope, B7					
INTRCPT2, G70	16.660979	0.688628	24.194	23653	0.000

The outcome variable is PROFIC

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Standard Coefficient	Approx. Error	T-ratio	d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	288.226971	1.078051	267.359	2984	0.000
NSE_M, G01	14.614002	0.563037	25.956	2984	0.000
REDE_1, G02	8.659987	1.187282	7.294	2984	0.000
MATRASO, G03	-7.023047	0.507800	-13.830	2984	0.000
For DISC slope, B1					
INTRCPT2, G10	15.722431	0.551659	28.500	23653	0.000
For SEXO slope, B2					
INTRCPT2, G20	-4.521789	0.560604	-8.066	23653	0.000
For NSE_CMG slope, B3					
INTRCPT2, G30	0.015781	0.269234	0.059	23653	0.954
For PRETO slope, B4					
INTRCPT2, G40	-4.648486	1.098930	-4.230	2987	0.000
For PARDO slope, B5					
INTRCPT2, G50	-1.008024	0.546980	-1.843	2987	0.065
For ATRASO slope, B6					
INTRCPT2, G60	-4.302193	0.218050	-19.730	23653	0.000
For ATITUDE slope, B7					
INTRCPT2, G70	16.660979	0.722089	23.073	23653	0.000

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, U0	18.77541	352.51616	731	1313.89674	0.000
PRETO slope, U4	5.91820	35.02513	734	694.00550	>.500
PARDO slope, U5	2.70088	7.29474	734	652.93168	>.500
level-1, R	41.06403	1686.25419			

Note: The chi-square statistics reported above are based on only 735 of 2988 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Statistics for current covariance components model

Deviance = 245635.808298
Number of estimated parameters = 7

- raça_2001_04

Program: HLM 5 Hierarchical Linear and Nonlinear Modeling
Authors: Stephen Raudenbush, Tony Bryk, & Richard Congdon
Publisher: Scientific Software International, Inc. (c) 2000
techsupport@ssicentral.com
www.ssicentral.com

Module: HLM2.EXE (5.04.21205.1)
Date: 29 May 2004, Saturday
Time: 14: 3: 7

SPECIFICATIONS FOR THIS HLM2 RUN Sat May 29 14:03:07 2004

Problem Title: NO TITLE

The data source for this run = raça_2001_04.ssm
The command file for this run = C:\Renato\UFMG\Doutorado\Computacional\Efeito raça\HLM - modelo racial II\Modelo base - racial II - 2001_04.hlm
Output file name = C:\Renato\UFMG\Doutorado\Computacional\Efeito raça\HLM - modelo racial II\raça_2001_04.out
The maximum number of level-2 units = 5133
The maximum number of iterations = 100
Method of estimation: restricted maximum likelihood

Weighting Specification

Weight
Variable

	Weighting?	Name	Normalized?
Level 1	yes	PESO_R	no
Level 2	no		no

Generalizations are at level-1

The outcome variable is PROFIC

The model specified for the fixed effects was:

```

-----
Level-1          Level-2
Coefficients      Predictors
-----
      INTRCPT1, B0  INTRCPT2, G00
                   NSE_M, G01
                   REDE_1, G02
                   MATRASO, G03
#  DISC slope, B1  INTRCPT2, G10
#  SEXO slope, B2  INTRCPT2, G20
#  NSE_CMG slope, B3  INTRCPT2, G30
      PRETO slope, B4  INTRCPT2, G40
      PARDO slope, B5  INTRCPT2, G50
#  ATRASO slope, B6  INTRCPT2, G60
#  ATITUDE slope, B7  INTRCPT2, G70

```

'#' - The residual parameter variance for this level-1 coefficient has been set to zero.

The model specified for the covariance components was:

Sigma squared (constant across level-2 units)

Tau dimensions
 INTRCPT1
 PRETO slope
 PARDO slope

Summary of the model specified (in equation format)

Level-1 Model

$$Y = B0 + B1*(DISC) + B2*(SEXO) + B3*(NSE_CMG) + B4*(PRETO) + B5*(PARDO) + B6*(ATRASO) + B7*(ATITUDE) + R$$

Level-2 Model

$$\begin{aligned}
 B0 &= G00 + G01*(NSE_M) + G02*(REDE_1) + G03*(MATRASO) + U0 \\
 B1 &= G10 \\
 B2 &= G20 \\
 B3 &= G30 \\
 B4 &= G40 + U4 \\
 B5 &= G50 + U5 \\
 B6 &= G60 \\
 B7 &= G70
 \end{aligned}$$

Final estimation of fixed effects:

Fixed Effect	Standard Coefficient	Approx. Error	T-ratio	d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	169.013335	0.451888	374.016	5129	0.000
NSE_M, G01	18.379584	0.322455	56.999	5129	0.000
REDE_1, G02	13.179113	0.748611	17.605	5129	0.000
MATRASO, G03	0.881973	0.395363	2.231	5129	0.026
For DISC slope, B1					
INTRCPT2, G10	7.554804	0.173817	43.464	114288	0.000
For SEXO slope, B2					
INTRCPT2, G20	1.667155	0.179271	9.300	114288	0.000
For NSE_CMG slope, B3					
INTRCPT2, G30	0.733399	0.095065	7.715	114288	0.000
For PRETO slope, B4					
INTRCPT2, G40	-4.859038	0.316885	-15.334	5132	0.000
For PARDO slope, B5					
INTRCPT2, G50	0.719478	0.211617	3.400	5132	0.001
For ATRASO slope, B6					
INTRCPT2, G60	-1.419235	0.104894	-13.530	114288	0.000
For ATITUDE slope, B7					
INTRCPT2, G70	13.504526	0.290903	46.423	114288	0.000

The outcome variable is PROFIC

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Standard Coefficient	Approx. Error	T-ratio	d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	169.013335	0.492550	343.139	5129	0.000
NSE_M, G01	18.379584	0.321705	57.132	5129	0.000
REDE_1, G02	13.179113	0.771111	17.091	5129	0.000
MATRASO, G03	0.881973	0.458134	1.925	5129	0.054
For DISC slope, B1					
INTRCPT2, G10	7.554804	0.240404	31.425	114288	0.000
For SEXO slope, B2					
INTRCPT2, G20	1.667155	0.188759	8.832	114288	0.000
For NSE_CMG slope, B3					
INTRCPT2, G30	0.733399	0.106170	6.908	114288	0.000
For PRETO slope, B4					
INTRCPT2, G40	-4.859038	0.302246	-16.076	5132	0.000
For PARDO slope, B5					
INTRCPT2, G50	0.719478	0.195794	3.675	5132	0.000
For ATRASO slope, B6					
INTRCPT2, G60	-1.419235	0.110250	-12.873	114288	0.000
For ATITUDE slope, B7					
INTRCPT2, G70	13.504526	0.378314	35.697	114288	0.000

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, U0	14.37861	206.74433	3511	9895.94563	0.000
PRETO slope, U4	3.04433	9.26795	3514	3052.97575	>.500
PARDO slope, U5	1.56867	2.46071	3514	3079.50696	>.500
level-1, R	37.37378	1396.79933			

Note: The chi-square statistics reported above are based on only 3515 of 5133 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Statistics for current covariance components model

Deviance = 1158593.038770
Number of estimated parameters = 7

- raça_2001_08

Program: HLM 5 Hierarchical Linear and Nonlinear Modeling
Authors: Stephen Raudenbush, Tony Bryk, & Richard Congdon
Publisher: Scientific Software International, Inc. (c) 2000
techsupport@ssicentral.com
www.ssicentral.com

Module: HLM2.EXE (5.04.21205.1)
Date: 29 May 2004, Saturday
Time: 14: 8:54

SPECIFICATIONS FOR THIS HLM2 RUN Sat May 29 14:08:54 2004

Problem Title: NO TITLE

The data source for this run = raça_2001_08.ssm
The command file for this run = C:\Renato\UFMG\Doutorado\Computacional\Efeito raça\HLM - modelo racial II\Modelo base - racial II - 2001_08.hlm
Output file name = C:\Renato\UFMG\Doutorado\Computacional\Efeito raça\HLM - modelo racial II\raça_2001_08.out
The maximum number of level-2 units = 3891
The maximum number of iterations = 100
Method of estimation: restricted maximum likelihood

Weighting Specification

Weight
Variable

	Weighting?	Name	Normalized?
Level 1	yes	PESO_R	no
Level 2	no		no

Generalizations are at level-1

The outcome variable is PROFIC

The model specified for the fixed effects was:

```

-----
Level-1          Level-2
Coefficients      Predictors
-----
      INTRCPT1, B0  INTRCPT2, G00
                    NSE_M, G01
                    REDE_1, G02
                    MATRASO, G03
#  DISC slope, B1  INTRCPT2, G10
#  SEXO slope, B2  INTRCPT2, G20
#  NSE_CMG slope, B3  INTRCPT2, G30
      PRETO slope, B4  INTRCPT2, G40
      PARDO slope, B5  INTRCPT2, G50
#  ATRASO slope, B6  INTRCPT2, G60
#  ATITUDE slope, B7  INTRCPT2, G70

```

'#' - The residual parameter variance for this level-1 coefficient has been set to zero.

The model specified for the covariance components was:

Sigma squared (constant across level-2 units)

Tau dimensions
 INTRCPT1
 PRETO slope
 PARDO slope

Summary of the model specified (in equation format)

Level-1 Model

$$Y = B0 + B1*(DISC) + B2*(SEXO) + B3*(NSE_CMG) + B4*(PRETO) + B5*(PARDO) + B6*(ATRASO) + B7*(ATITUDE) + R$$

Level-2 Model

$$\begin{aligned}
 B0 &= G00 + G01*(NSE_M) + G02*(REDE_1) + G03*(MATRASO) + U0 \\
 B1 &= G10 \\
 B2 &= G20 \\
 B3 &= G30 \\
 B4 &= G40 + U4 \\
 B5 &= G50 + U5 \\
 B6 &= G60 \\
 B7 &= G70
 \end{aligned}$$

Final estimation of fixed effects:

Fixed Effect	Standard Coefficient	Standard Error	Approx. T-ratio	d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	251.168458	0.562166	446.787	3887	0.000
NSE_M, G01	15.463915	0.363039	42.596	3887	0.000
REDE_1, G02	10.478582	0.868129	12.070	3887	0.000
MATRASO, G03	-3.440144	0.321978	-10.684	3887	0.000
For DISC slope, B1					
INTRCPT2, G10	6.746153	0.198059	34.061	100778	0.000
For SEXO slope, B2					
INTRCPT2, G20	-1.436923	0.203366	-7.066	100778	0.000
For NSE_CMG slope, B3					
INTRCPT2, G30	0.520251	0.108913	4.777	100778	0.000
For PRETO slope, B4					
INTRCPT2, G40	-4.945785	0.436249	-11.337	3890	0.000
For PARDO slope, B5					
INTRCPT2, G50	-0.698210	0.227656	-3.067	3890	0.003
For ATRASO slope, B6					
INTRCPT2, G60	-3.616907	0.094855	-38.131	100778	0.000
For ATITUDE slope, B7					
INTRCPT2, G70	10.299158	0.289316	35.598	100778	0.000

The outcome variable is PROFIC

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Standard Coefficient	Standard Error	Approx. T-ratio	d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	251.168458	0.563634	445.623	3887	0.000
NSE_M, G01	15.463915	0.394276	39.221	3887	0.000
REDE_1, G02	10.478582	0.951648	11.011	3887	0.000
MATRASO, G03	-3.440144	0.334913	-10.272	3887	0.000
For DISC slope, B1					
INTRCPT2, G10	6.746153	0.265472	25.412	100778	0.000
For SEXO slope, B2					
INTRCPT2, G20	-1.436923	0.237904	-6.040	100778	0.000
For NSE_CMG slope, B3					
INTRCPT2, G30	0.520251	0.119052	4.370	100778	0.000
For PRETO slope, B4					
INTRCPT2, G40	-4.945785	0.425001	-11.637	3890	0.000
For PARDO slope, B5					
INTRCPT2, G50	-0.698210	0.208608	-3.347	3890	0.001
For ATRASO slope, B6					
INTRCPT2, G60	-3.616907	0.156276	-23.144	100778	0.000
For ATITUDE slope, B7					
INTRCPT2, G70	10.299158	0.383310	26.869	100778	0.000

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, U0	13.16148	173.22444	2659	7153.89175	0.000
PRETO slope, U4	3.70049	13.69362	2662	2683.04247	0.383
PARDO slope, U5	1.55956	2.43224	2662	2413.65328	>.500
level-1, R	39.50974	1561.01927			

Note: The chi-square statistics reported above are based on only 2663 of 3891 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Statistics for current covariance components model

Deviance = 1031753.280007
Number of estimated parameters = 7

- raça_2001_11

Program: HLM 5 Hierarchical Linear and Nonlinear Modeling
Authors: Stephen Raudenbush, Tony Bryk, & Richard Congdon
Publisher: Scientific Software International, Inc. (c) 2000
techsupport@ssicentral.com
www.ssicentral.com

Module: HLM2.EXE (5.04.21205.1)
Date: 29 May 2004, Saturday
Time: 14:35:10

SPECIFICATIONS FOR THIS HLM2 RUN Sat May 29 14:35:10 2004

Problem Title: NO TITLE

The data source for this run = raça_2001_11.ssm
The command file for this run = C:\Renato\UFMG\Doutorado\Computacional\Efeito raça\HLM - modelo racial II\Modelo base - racial II - 2001_11.hlm
Output file name = C:\Renato\UFMG\Doutorado\Computacional\Efeito raça\HLM - modelo racial II\raça_2001_11.out
The maximum number of level-2 units = 2690
The maximum number of iterations = 100
Method of estimation: restricted maximum likelihood

Weighting Specification

Weight
Variable

	Weighting?	Name	Normalized?
Level 1	yes	PESO_R	no
Level 2	no		no

Generalizations are at level-1

The outcome variable is PROFIC

The model specified for the fixed effects was:

```

-----
Level-1          Level-2
Coefficients     Predictors
-----
      INTRCPT1, B0  INTRCPT2, G00
                   NSE_M, G01
                   REDE_1, G02
                   MATRASO, G03
#  DISC slope, B1  INTRCPT2, G10
#  SEXO slope, B2  INTRCPT2, G20
#  NSE_CMG slope, B3  INTRCPT2, G30
      PRETO slope, B4  INTRCPT2, G40
      PARDO slope, B5  INTRCPT2, G50
#  ATRASO slope, B6  INTRCPT2, G60
#  ATITUDE slope, B7  INTRCPT2, G70

```

'#' - The residual parameter variance for this level-1 coefficient has been set to zero.

The model specified for the covariance components was:

Sigma squared (constant across level-2 units)

Tau dimensions
 INTRCPT1
 PRETO slope
 PARDO slope

Summary of the model specified (in equation format)

Level-1 Model

$$Y = B0 + B1*(DISC) + B2*(SEXO) + B3*(NSE_CMG) + B4*(PRETO) + B5*(PARDO) + B6*(ATRASO) + B7*(ATITUDE) + R$$

Level-2 Model

$$\begin{aligned}
 B0 &= G00 + G01*(NSE_M) + G02*(REDE_1) + G03*(MATRASO) + U0 \\
 B1 &= G10 \\
 B2 &= G20 \\
 B3 &= G30 \\
 B4 &= G40 + U4 \\
 B5 &= G50 + U5 \\
 B6 &= G60 \\
 B7 &= G70
 \end{aligned}$$

Final estimation of fixed effects:

Fixed Effect	Standard Coefficient	Approx. Error	T-ratio	d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	286.346556	0.922972	310.244	2686	0.000
NSE_M, G01	18.756872	0.515326	36.398	2686	0.000
REDE_1, G02	7.916180	1.060972	7.461	2686	0.000
MATRASO, G03	-5.757027	0.495807	-11.611	2686	0.000
For DISC slope, B1					
INTRCPT2, G10	12.635947	0.251889	50.165	72372	0.000
For SEXO slope, B2					
INTRCPT2, G20	-4.761575	0.261754	-18.191	72372	0.000
For NSE_CMG slope, B3					
INTRCPT2, G30	-0.025375	0.131270	-0.193	72372	0.847
For PRETO slope, B4					
INTRCPT2, G40	-2.907775	0.561480	-5.179	2689	0.000
For PARDO slope, B5					
INTRCPT2, G50	-0.593575	0.293992	-2.019	2689	0.043
For ATRASO slope, B6					
INTRCPT2, G60	-4.006522	0.104725	-38.258	72372	0.000
For ATITUDE slope, B7					
INTRCPT2, G70	12.465472	0.336082	37.091	72372	0.000

The outcome variable is PROFIC

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Standard Coefficient	Approx. Error	T-ratio	d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	286.346556	0.886693	322.938	2686	0.000
NSE_M, G01	18.756872	0.547298	34.272	2686	0.000
REDE_1, G02	7.916180	1.069264	7.403	2686	0.000
MATRASO, G03	-5.757027	0.490292	-11.742	2686	0.000
For DISC slope, B1					
INTRCPT2, G10	12.635947	0.462597	27.315	72372	0.000
For SEXO slope, B2					
INTRCPT2, G20	-4.761575	0.310833	-15.319	72372	0.000
For NSE_CMG slope, B3					
INTRCPT2, G30	-0.025375	0.148069	-0.171	72372	0.864
For PRETO slope, B4					
INTRCPT2, G40	-2.907775	0.530981	-5.476	2689	0.000
For PARDO slope, B5					
INTRCPT2, G50	-0.593575	0.276327	-2.148	2689	0.032
For ATRASO slope, B6					
INTRCPT2, G60	-4.006522	0.178997	-22.383	72372	0.000
For ATITUDE slope, B7					
INTRCPT2, G70	12.465472	0.491645	25.355	72372	0.000

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, U0	16.03631	257.16320	1675	5559.95445	0.000
PRETO slope, U4	3.11994	9.73404	1678	1547.49989	>.500
PARDO slope, U5	1.61810	2.61824	1678	1532.08080	>.500
level-1, R	42.52726	1808.56799			

Note: The chi-square statistics reported above are based on only 1679 of 2690 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Statistics for current covariance components model

Deviance = 751990.156430
Number of estimated parameters = 7

- raça_2003_04

Program: HLM 5 Hierarchical Linear and Nonlinear Modeling
Authors: Stephen Raudenbush, Tony Bryk, & Richard Congdon
Publisher: Scientific Software International, Inc. (c) 2000
techsupport@ssicentral.com
www.ssicentral.com

Module: HLM2.EXE (5.04.21205.1)
Date: 29 May 2004, Saturday
Time: 14:39:31

SPECIFICATIONS FOR THIS HLM2 RUN Sat May 29 14:39:31 2004

Problem Title: NO TITLE

The data source for this run = raça_2003_04.ssm
The command file for this run = C:\Renato\UFMG\Doutorado\Computacional\Efeito raça\HLM - modelo racial II\Modelo base - racial II - 2003_04.hlm
Output file name = C:\Renato\UFMG\Doutorado\Computacional\Efeito raça\HLM - modelo racial II\raça_2003_04.out
The maximum number of level-2 units = 4058
The maximum number of iterations = 100
Method of estimation: restricted maximum likelihood

Weighting Specification

Weight
Variable

	Weighting?	Name	Normalized?
Level 1	yes	PESO_R	no
Level 2	no		no

Generalizations are at level-1

The outcome variable is PROFIC

The model specified for the fixed effects was:

```

-----
Level-1          Level-2
Coefficients     Predictors
-----
      INTRCPT1, B0  INTRCPT2, G00
                   NSE_M, G01
                   REDE_1, G02
                   MATRASO, G03
#   DISC slope, B1  INTRCPT2, G10
#   SEXO slope, B2  INTRCPT2, G20
#   NSE_CMG slope, B3  INTRCPT2, G30
      PRETO slope, B4  INTRCPT2, G40
      PARDO slope, B5  INTRCPT2, G50
#   ATRASO slope, B6  INTRCPT2, G60
#   ATITUDE slope, B7  INTRCPT2, G70

```

'#' - The residual parameter variance for this level-1 coefficient has been set to zero.

The model specified for the covariance components was:

Sigma squared (constant across level-2 units)

Tau dimensions
 INTRCPT1
 PRETO slope
 PARDO slope

Summary of the model specified (in equation format)

Level-1 Model

$$Y = B0 + B1*(DISC) + B2*(SEXO) + B3*(NSE_CMG) + B4*(PRETO) + B5*(PARDO) + B6*(ATRASO) + B7*(ATITUDE) + R$$

Level-2 Model

$$\begin{aligned}
 B0 &= G00 + G01*(NSE_M) + G02*(REDE_1) + G03*(MATRASO) + U0 \\
 B1 &= G10 \\
 B2 &= G20 \\
 B3 &= G30 \\
 B4 &= G40 + U4 \\
 B5 &= G50 + U5 \\
 B6 &= G60 \\
 B7 &= G70
 \end{aligned}$$

Final estimation of fixed effects:

Fixed Effect	Standard Coefficient	Approx. Error	T-ratio	d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	177.170948	0.469012	377.754	4054	0.000
NSE_M, G01	17.919913	0.350669	51.102	4054	0.000
REDE_1, G02	10.216497	0.850618	12.011	4054	0.000
MATRASO, G03	0.559245	0.507364	1.102	4054	0.271
For DISC slope, B1					
INTRCPT2, G10	3.990896	0.172181	23.179	90763	0.000
For SEXO slope, B2					
INTRCPT2, G20	1.485213	0.179470	8.276	90763	0.000
For NSE_CMG slope, B3					
INTRCPT2, G30	1.025644	0.105069	9.762	90763	0.000
For PRETO slope, B4					
INTRCPT2, G40	-5.729792	0.355630	-16.112	4057	0.000
For PARDO slope, B5					
INTRCPT2, G50	0.478273	0.219591	2.178	4057	0.029
For ATRASO slope, B6					
INTRCPT2, G60	-2.113566	0.101941	-20.733	90763	0.000
For ATITUDE slope, B7					
INTRCPT2, G70	4.215642	0.282634	14.916	90763	0.000

The outcome variable is PROFIC

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Standard Coefficient	Approx. Error	T-ratio	d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	177.170948	0.543885	325.751	4054	0.000
NSE_M, G01	17.919913	0.391574	45.764	4054	0.000
REDE_1, G02	10.216497	0.965985	10.576	4054	0.000
MATRASO, G03	0.559245	0.674798	0.829	4054	0.407
For DISC slope, B1					
INTRCPT2, G10	3.990896	0.271236	14.714	90763	0.000
For SEXO slope, B2					
INTRCPT2, G20	1.485213	0.244962	6.063	90763	0.000
For NSE_CMG slope, B3					
INTRCPT2, G30	1.025644	0.121599	8.435	90763	0.000
For PRETO slope, B4					
INTRCPT2, G40	-5.729792	0.338435	-16.930	4057	0.000
For PARDO slope, B5					
INTRCPT2, G50	0.478273	0.209713	2.281	4057	0.023
For ATRASO slope, B6					
INTRCPT2, G60	-2.113566	0.309117	-6.837	90763	0.000
For ATITUDE slope, B7					
INTRCPT2, G70	4.215642	0.367131	11.483	90763	0.000

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, U0	13.95378	194.70789	2649	7845.51188	0.000
PRETO slope, U4	3.17122	10.05665	2652	2734.46704	0.129
PARDO slope, U5	1.78746	3.19501	2652	2632.67603	>.500
level-1, R	36.05786	1300.16951			

Note: The chi-square statistics reported above are based on only 2653 of 4058 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Statistics for current covariance components model

Deviance = 913604.960985
Number of estimated parameters = 7

- raça_2003_08

Program: HLM 5 Hierarchical Linear and Nonlinear Modeling
Authors: Stephen Raudenbush, Tony Bryk, & Richard Congdon
Publisher: Scientific Software International, Inc. (c) 2000
techsupport@ssicentral.com
www.ssicentral.com

Module: HLM2.EXE (5.04.21205.1)
Date: 29 May 2004, Saturday
Time: 14:44: 1

SPECIFICATIONS FOR THIS HLM2 RUN Sat May 29 14:44:01 2004

Problem Title: NO TITLE

The data source for this run = raça_2003_08.ssm
The command file for this run = C:\Renato\UFMG\Doutorado\Computacional\Efeito raça\HLM - modelo racial II\Modelo base - racial II - 2003_08.hlm
Output file name = C:\Renato\UFMG\Doutorado\Computacional\Efeito raça\HLM - modelo racial II\raça_2003_08.out
The maximum number of level-2 units = 2878
The maximum number of iterations = 100
Method of estimation: restricted maximum likelihood

Weighting Specification

Weight
Variable

	Weighting?	Name	Normalized?
Level 1	yes	PESO_R	no
Level 2	no		no

Generalizations are at level-1

The outcome variable is PROFIC

The model specified for the fixed effects was:

```

-----
Level-1          Level-2
Coefficients     Predictors
-----
      INTRCPT1, B0  INTRCPT2, G00
                   NSE_M, G01
                   REDE_1, G02
                   MATRASO, G03
#  DISC slope, B1  INTRCPT2, G10
#  SEXO slope, B2  INTRCPT2, G20
#  NSE_CMG slope, B3  INTRCPT2, G30
      PRETO slope, B4  INTRCPT2, G40
      PARDO slope, B5  INTRCPT2, G50
#  ATRASO slope, B6  INTRCPT2, G60
#  ATITUDE slope, B7  INTRCPT2, G70

```

'#' - The residual parameter variance for this level-1 coefficient has been set to zero.

The model specified for the covariance components was:

Sigma squared (constant across level-2 units)

Tau dimensions
 INTRCPT1
 PRETO slope
 PARDO slope

Summary of the model specified (in equation format)

Level-1 Model

$$Y = B0 + B1*(DISC) + B2*(SEXO) + B3*(NSE_CMG) + B4*(PRETO) + B5*(PARDO) + B6*(ATRASO) + B7*(ATITUDE) + R$$

Level-2 Model

$$\begin{aligned}
 B0 &= G00 + G01*(NSE_M) + G02*(REDE_1) + G03*(MATRASO) + U0 \\
 B1 &= G10 \\
 B2 &= G20 \\
 B3 &= G30 \\
 B4 &= G40 + U4 \\
 B5 &= G50 + U5 \\
 B6 &= G60 \\
 B7 &= G70
 \end{aligned}$$

Final estimation of fixed effects:

Fixed Effect	Standard Coefficient	Standard Error	Approx. T-ratio	d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	246.959967	0.642119	384.601	2874	0.000
NSE_M, G01	16.210024	0.489692	33.103	2874	0.000
REDE_1, G02	11.719864	1.163022	10.077	2874	0.000
MATRASO, G03	-3.596007	0.470714	-7.639	2874	0.000
For DISC slope, B1					
INTRCPT2, G10	9.047902	0.239638	37.757	72475	0.000
For SEXO slope, B2					
INTRCPT2, G20	0.310429	0.246766	1.258	72475	0.209
For NSE_CMG slope, B3					
INTRCPT2, G30	1.286314	0.139841	9.198	72475	0.000
For PRETO slope, B4					
INTRCPT2, G40	-5.132862	0.518418	-9.901	2877	0.000
For PARDO slope, B5					
INTRCPT2, G50	-0.292340	0.289007	-1.012	2877	0.312
For ATRASO slope, B6					
INTRCPT2, G60	-4.191436	0.121307	-34.552	72475	0.000
For ATITUDE slope, B7					
INTRCPT2, G70	10.369337	0.339313	30.560	72475	0.000

The outcome variable is PROFIC

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Standard Coefficient	Standard Error	Approx. T-ratio	d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	246.959967	0.640506	385.570	2874	0.000
NSE_M, G01	16.210024	0.506221	32.022	2874	0.000
REDE_1, G02	11.719864	1.229319	9.534	2874	0.000
MATRASO, G03	-3.596007	0.463103	-7.765	2874	0.000
For DISC slope, B1					
INTRCPT2, G10	9.047902	0.347507	26.037	72475	0.000
For SEXO slope, B2					
INTRCPT2, G20	0.310429	0.251405	1.235	72475	0.217
For NSE_CMG slope, B3					
INTRCPT2, G30	1.286314	0.163029	7.890	72475	0.000
For PRETO slope, B4					
INTRCPT2, G40	-5.132862	0.502472	-10.215	2877	0.000
For PARDO slope, B5					
INTRCPT2, G50	-0.292340	0.277677	-1.053	2877	0.293
For ATRASO slope, B6					
INTRCPT2, G60	-4.191436	0.197983	-21.171	72475	0.000
For ATITUDE slope, B7					
INTRCPT2, G70	10.369337	0.481192	21.549	72475	0.000

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, U0	13.78003	189.88911	1938	4936.29950	0.000
PRETO slope, U4	4.66295	21.74312	1941	2009.83767	0.135
PARDO slope, U5	2.15302	4.63552	1941	1772.38489	>.500
level-1, R	40.66784	1653.87305			

Note: The chi-square statistics reported above are based on only 1942 of 2878 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Statistics for current covariance components model

Deviance = 746378.710275
Number of estimated parameters = 7

- raça_2003_11

Program: HLM 5 Hierarchical Linear and Nonlinear Modeling
Authors: Stephen Raudenbush, Tony Bryk, & Richard Congdon
Publisher: Scientific Software International, Inc. (c) 2000
techsupport@ssicentral.com
www.ssicentral.com

Module: HLM2.EXE (5.04.21205.1)
Date: 29 May 2004, Saturday
Time: 14:47:41

SPECIFICATIONS FOR THIS HLM2 RUN Sat May 29 14:47:41 2004

Problem Title: NO TITLE

The data source for this run = raça_2003_11.ssm
The command file for this run = C:\Renato\UFMG\Doutorado\Computacional\Efeito raça\HLM - modelo racial II\Modelo base - racial II - 2003_11.hlm
Output file name = C:\Renato\UFMG\Doutorado\Computacional\Efeito raça\HLM - modelo racial II\raça_2003_11.out
The maximum number of level-2 units = 1879
The maximum number of iterations = 100
Method of estimation: restricted maximum likelihood

Weighting Specification

Weight
Variable

	Weighting?	Name	Normalized?
Level 1	yes	PESO_R	no
Level 2	no		no

Generalizations are at level-1

The outcome variable is PROFIC

The model specified for the fixed effects was:

```

-----
Level-1          Level-2
Coefficients     Predictors
-----
      INTRCPT1, B0  INTRCPT2, G00
                   NSE_M, G01
                   REDE_1, G02
                   MATRASO, G03
#  DISC slope, B1  INTRCPT2, G10
#  SEXO slope, B2  INTRCPT2, G20
#  NSE_CMG slope, B3  INTRCPT2, G30
  PRETO slope, B4  INTRCPT2, G40
  PARDO slope, B5  INTRCPT2, G50
#  ATRASO slope, B6  INTRCPT2, G60
#  ATITUDE slope, B7  INTRCPT2, G70

```

'#' - The residual parameter variance for this level-1 coefficient has been set to zero.

The model specified for the covariance components was:

Sigma squared (constant across level-2 units)

Tau dimensions
 INTRCPT1
 PRETO slope
 PARDO slope

Summary of the model specified (in equation format)

Level-1 Model

$$Y = B0 + B1*(DISC) + B2*(SEXO) + B3*(NSE_CMG) + B4*(PRETO) + B5*(PARDO) + B6*(ATRASO) + B7*(ATITUDE) + R$$

Level-2 Model

$$\begin{aligned}
 B0 &= G00 + G01*(NSE_M) + G02*(REDE_1) + G03*(MATRASO) + U0 \\
 B1 &= G10 \\
 B2 &= G20 \\
 B3 &= G30 \\
 B4 &= G40 + U4 \\
 B5 &= G50 + U5 \\
 B6 &= G60 \\
 B7 &= G70
 \end{aligned}$$

Final estimation of fixed effects:

Fixed Effect	Standard Coefficient	Standard Error	Approx. T-ratio	d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	288.412636	1.039481	277.458	1875	0.000
NSE_M, G01	21.129641	0.687272	30.744	1875	0.000
REDE_1, G02	8.306714	1.416253	5.865	1875	0.000
MATRASO, G03	-5.142916	0.754432	-6.817	1875	0.000
For DISC slope, B1					
INTRCPT2, G10	11.669933	0.302145	38.624	50836	0.000
For SEXO slope, B2					
INTRCPT2, G20	-3.171746	0.314246	-10.093	50836	0.000
For NSE_CMG slope, B3					
INTRCPT2, G30	0.720136	0.165556	4.350	50836	0.000
For PRETO slope, B4					
INTRCPT2, G40	-3.769744	0.672500	-5.606	1878	0.000
For PARDO slope, B5					
INTRCPT2, G50	-0.981593	0.377712	-2.599	1878	0.010
For ATRASO slope, B6					
INTRCPT2, G60	-4.613274	0.150518	-30.649	50836	0.000
For ATITUDE slope, B7					
INTRCPT2, G70	11.982680	0.367586	32.598	50836	0.000

The outcome variable is PROFIC

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Standard Coefficient	Standard Error	Approx. T-ratio	d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	288.412636	1.010342	285.460	1875	0.000
NSE_M, G01	21.129641	0.724286	29.173	1875	0.000
REDE_1, G02	8.306714	1.408549	5.897	1875	0.000
MATRASO, G03	-5.142916	0.760817	-6.760	1875	0.000
For DISC slope, B1					
INTRCPT2, G10	11.669933	0.509822	22.890	50836	0.000
For SEXO slope, B2					
INTRCPT2, G20	-3.171746	0.353243	-8.979	50836	0.000
For NSE_CMG slope, B3					
INTRCPT2, G30	0.720136	0.162541	4.430	50836	0.000
For PRETO slope, B4					
INTRCPT2, G40	-3.769744	0.626263	-6.019	1878	0.000
For PARDO slope, B5					
INTRCPT2, G50	-0.981593	0.366260	-2.680	1878	0.008
For ATRASO slope, B6					
INTRCPT2, G60	-4.613274	0.228582	-20.182	50836	0.000
For ATITUDE slope, B7					
INTRCPT2, G70	11.982680	0.517722	23.145	50836	0.000

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, U0	16.54493	273.73471	1258	4088.22733	0.000
PRETO slope, U4	3.57824	12.80379	1261	1149.12165	>.500
PARDO slope, U5	2.38982	5.71123	1261	1156.51582	>.500
level-1, R	42.81947	1833.50679			

Note: The chi-square statistics reported above are based on only 1262 of 1879 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Statistics for current covariance components model

Deviance = 528996.578186
Number of estimated parameters = 7