



ADILSON DALBEN

**FATORES ASSOCIADOS À PROFICIÊNCIA EM
LEITURA E MATEMÁTICA: UMA APLICAÇÃO DO
MODELO LINEAR HIERÁRQUICO COM DADOS
LONGITUDINAIS DO PROJETO GERES**

CAMPINAS

2014



UNIVERSIDADE ESTADUAL DE CAMPINAS FACULDADE DE EDUCAÇÃO

ADILSON DALBEN

**FATORES ASSOCIADOS À PROFICIÊNCIA EM LEITURA E
MATEMÁTICA: UMA APLICAÇÃO DO MODELO LINEAR
HIERÁRQUICO COM DADOS LONGITUDINAIS DO
PROJETO GERES**

Orientador(a): Prof. Dr. Luiz Carlos de Freitas

Co-Orientador: Prof. Dr. Dalton Francisco de Andrade

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Educação da Faculdade de Educação da Universidade Estadual de Campinas para obtenção do título de Doutor em Educação, na área de concentração de Ensino e Práticas Culturais.

**ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA TESE
DEFENDIDA PELO ALUNO ADILSON DALBEN, ORIENTADO PELO
PROF. DR. LUIZ CARLOS DE FREITAS**

Assinatura do Orientador

A handwritten signature in blue ink, appearing to be "Luiz Carlos de Freitas", is written over a horizontal line. The signature is stylized and cursive.

CAMPINAS

2014

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Faculdade de Educação
Gildenir Carolino Santos - CRB 8/5447

D15f Dalben, Adilson, 1965-
Fatores associados à proficiência em leitura e matemática : uma aplicação do modelo linear hierárquico com dados longitudinais do Projeto GERES / Adilson Dalben. – Campinas, SP : [s.n.], 2014.

Orientador: Luiz Carlos de Freitas.
Coorientador: Dalton Francisco de Andrade.
Tese (doutorado) – Universidade Estadual de Campinas, Faculdade de Educação.

1. Avaliação educacional. 2. Eficácia no ensino. 3. Modelos de valor agregado. 4. Modelos lineares hierárquicos. 5. Estudos longitudinais. I. Freitas, Luiz Carlos de, 1947-. II. Andrade, Dalton Francisco de, 1951-. III. Universidade Estadual de Campinas. Faculdade de Educação. IV. Título.

Informações para Biblioteca Digital

Título em outro idioma: Factors associated with proficiency in reading and mathematics : an application of hierarchical linear models with longitudinal data of the GERES Project

Palavras-chave em inglês:

Educational evaluation

Teaching effectiveness

Value added models

Hierarchical linear models

Longitudinal studies

Área de concentração: Ensino e Práticas Culturais

Titulação: Doutor em Educação

Banca examinadora:

Luiz Carlos de Freitas [Orientador]

José Francisco Soares

Odival Faccenda

Ocimar Munhoz Alavarse

Dirceu da Silva

Data de defesa: 24-02-2014

Programa de Pós-Graduação: Educação

UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE EDUCAÇÃO

TESE DE DOUTORADO

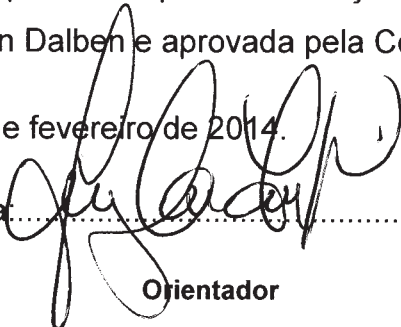
FATORES ASSOCIADOS À PROFICIÊNCIA EM LEITURA E
MATEMÁTICA: UMA APLICAÇÃO DO MODELO LINEAR
HIERÁRQUICO COM DADOS LONGITUDINAIS DO PROJETO
GERES

Autor : Adilson Dalben
Orientador: Prof. Dr. Luiz Carlos de Freitas
Co-orientador: Prof. Dr. Dalton Francisco de Andrade

Este exemplar corresponde à redação final da Tese defendida
por Adilson Dalben e aprovada pela Comissão Julgadora.

Data: 24 de fevereiro de 2014.

Assinatura:.....



Orientador

COMISSÃO JULGADORA:

Prof. Dr. Dalton Francisco de Andrade
Prof. Dr. Luiz Carlos de Freitas
Prof. Dr. [Signature]
Prof. Dr. [Signature]

RESUMO

Esta pesquisa é um estudo sobre a eficácia e equidade escolar que tem ganhado atenção especial nos países que usam as avaliações em larga escala a serviço da gestão do sistema educativo. No Brasil, que desde a década de 1990 colocou a avaliação educacional como recurso central em suas políticas educacionais, mas coletando dados seccionais, que são muito frágeis para essa finalidade. Essa fragilidade decorre da alta associação que os fatores extraescolares, sobretudo o nível socioeconômico do aluno, têm sobre as medidas de proficiência. Diante disso, foram usados dados longitudinais e a análise foi feita por meio de modelos lineares hierárquicos. Esta pesquisa teve como objetivo principal desenvolver um modelo estatístico capaz de identificar tais fatores para a realidade brasileira, considerando que a aprendizagem é um processo complexo, isto é, ela é influenciada simultaneamente por múltiplos fatores. Foram desenvolvidos modelos de valor agregado que não só identificam tais variáveis, como também caracterizam sua influência em alunos com distintas proficiências no início de cada período de escolarização. A base de dados utilizada nesses modelos foi fornecida pelo Projeto GERES, que, no período de 2005 a 2008, coletou dados dos mesmos alunos de 1ª a 4ª séries de uma amostra de 312 escolas em cinco grandes cidades brasileiras. Foram medidas as proficiências em Leitura e Matemática de 35.538 alunos e coletadas informações de contexto desses alunos, seus familiares, professores, diretores e escola. Após a redução do grande número de informações disponibilizadas pelo Projeto GERES, feita por meio da Análise Fatorial Exploratória (AFE), as variáveis resultantes foram reorganizadas em três arquivos usados para análise em modelos lineares hierárquicos de três níveis. Os resultados encontrados evidenciam uma significativa instabilidade nos efeitos que as variáveis têm sobre a proficiência, tanto em leitura quanto em matemática. Ao final da pesquisa, são encontrados alguns fatores que influenciam positivamente e negativamente a proficiência em Leitura e Matemática e outros que afetam especificamente cada uma dessas áreas, indicando que podem colaborar para o aumento da eficácia e da equidade das escolas. No entanto, constatam-se também algumas variáveis que têm comportamentos incoerentes com o esperado e outras com comportamentos opostos nas duas áreas. Assim, dos achados das pesquisas, comprova-se que, com base nos dados utilizados, procedimentos metodológicos e modelos estatísticos adotados, os modelos de valor agregado melhoram a confiabilidade das análises em comparação aos modelos que usam dados seccionais, mas ainda são inviáveis como ferramentas para a gestão do sistema educativo, sobretudo para o uso meritocrático de seus resultados. Dessa forma, esta pesquisa corrobora os achados de outras realizadas no âmbito internacional e permite afirmar que a qualidade da modelagem estatística depende da qualidade dos dados que busca modelar, podendo gerar distorções, estabelecer relações inesperadas ou levar a conclusões equivocadas. Em contrapartida, trata-se de recursos que podem ser usados no sistema educativo, fornecendo dados importantes para a orientação das políticas públicas numa perspectiva de avaliação formativa, com vistas ao melhoramento da qualidade de ensino oferecido pelas escolas e à melhor formação dos profissionais docentes e não-docentes que nelas trabalham.

Palavras-chave: Eficácia Escolar; Equidade Escolar; Modelos de Valor Agregado; Modelos Lineares Hierárquicos; Dados longitudinais.

ABSTRACT

This research is a study on school effectiveness and equality in Brazil, adding up to a number of other researches that have drawn special attention in countries that use large-scale evaluations at the service of the education system management. In the Brazil has regarded the educational evaluation as a central resource in national education policies, but using cross-sectional data, which are far more fragile for such purpose. This fragility has derived from the great influence that extra-school factors, particularly the students' socioeconomic status, exerts on proficiency measures. Longitudinal data was used in the analyses with hierarchical linear models. The main objective of this research was to develop a statistical model to identify such factors in the Brazilian reality, considering that learning is a complex process, i.e. it is simultaneously influenced by multiple factors. Value-added models were developed not only to identify such variables, but also to characterize their influence on students showing different proficiencies at the beginning of every school term. The data base used in those models was provided by the GERES Project, which collected data of the same students from the 1st to the 4th grade from a sample of 312 schools in five Brazilian cities from 2005 to 2008. Proficiencies of 35,538 students were measured, and information about these students' context, family, teachers, principals and school were gathered. After the reduction of the great amount of information made available by the GERES Project by means of Exploratory Factor Analysis (EFA), the resulting variables were reorganized in three files used for analysis in three-level hierarchical linear models. The results evidenced significant instability in the effects that the variables have on proficiency both in Reading and in Mathematics. At the end of the research, some factors that influence Reading and Mathematics proficiency either positively or negatively, as well as other factors that specifically affect one of those areas, were found, thus indicating that they may contribute to increased school effectiveness and equality. However, some variables whose behavior was inconsistent with the one expected, and others with opposite behaviors in the two areas were also found. Therefore, from the research findings, based on the data used, the methodological procedures and the statistical models adopted, it has been evidenced that value-added models improve the analysis reliability in comparison with models that use cross-sectional data, but they are still impracticable as tools for education system management, particularly for meritocratic use of their results. Hence, this research has corroborated the findings of other studies carried out over the world and has enabled us to state that the quality of the statistical modeling depends on the quality of data that it attempts to model, and it may generate distortions, establish unexpected relationships or lead to misleading conclusions. On the other hand, these resources may be used in the education system by providing important data for guiding public policies in a educative evaluation perspective, aiming at improving the quality of teaching offered by schools, teachers and other professionals that work in the school setting.

Keywords: School Effectiveness; School Equality; Value-Added Models; Hierarchical Linear Models; Longitudinal Data.

SUMÁRIO

APRESENTAÇÃO.....	1
CAPÍTULO 1 - O CAMPO DE INSERÇÃO DESTA PESQUISA	5
1.1) O CAMPO DE ESTUDOS DA EFICÁCIA ESCOLAR	8
1.2) O PROJETO GERES.....	18
CAPÍTULO 2 – DADOS LONGITUDINAIS E HLM	25
2.1) USO DE DADOS LONGITUDINAIS	28
2.2) USO DE MODELOS LINEARES HIERÁRQUICOS	31
2.3) MODELOS LINEARES HIERÁRQUICOS	43
2.4) MODELOS LINEARES HIERÁRQUICOS AJUSTADOS ANO A ANO.....	61
CAPÍTULO 3 - A PREPARAÇÃO DAS VARIÁVEIS E DOS ARQUIVOS	85
3.1) TRATAMENTO DOS DADOS	87
3.2) PREPARAÇÃO DOS ARQUIVOS.....	95
3.3) DESCRIÇÃO QUANTITATIVA DOS REGISTROS RESULTANTES EM CADA UM DOS ARQUIVOS.....	112
3.4) OUTRAS OBSERVAÇÕES.....	116
CAPÍTULO 4 – UM MODELO PARA A COMPLEXIDADE	119
4.1) PROCEDIMENTOS PARA O AJUSTE DOS MODELOS LINEARES HIERÁRQUICOS	122
4.2) O MODELO PARA ESTUDO COM BASE NA PROFICIÊNCIA EM LEITURA	125
4.3) O MODELO PARA ESTUDO COM BASE NA PROFICIÊNCIA EM MATEMÁTICA	141
CAPÍTULO 5 - ANÁLISE DOS EFEITOS DAS VARIÁVEIS.....	159
5.1) INTERPRETAÇÃO CONJUNTA DOS RESULTADOS DO MODELO NULO EM LEITURA E MATEMÁTICA	162
5.2) INTERPRETAÇÃO DOS RESULTADOS DO MODELO DE NÍVEL 1 EM LEITURA E EM MATEMÁTICA.....	164
5.3) INTERPRETAÇÃO DOS RESULTADOS DO MODELO DE NÍVEL 2	174
5.4) INTERPRETAÇÃO DOS RESULTADOS DO MODELO DE NÍVEL 3	184
CAPÍTULO 6 - ANÁLISE INDIVIDUAL DAS VARIÁVEIS.....	227
6.1) MODELO DE REFERÊNCIA EM LEITURA	231
6.2) MODELO DE REFERÊNCIA EM MATEMÁTICA.....	237
6.3) ANÁLISE DAS VARIÁVEIS	241
CONCLUSÕES	269
REFERÊNCIAS BIBLIOGRÁFICAS	293
APÊNDICES.....	303
ANEXOS.....	407

Aos professores, pela sua importância, representando a mais potente, senão única, possibilidade de transformação para a minimização das dificuldades enfrentadas pela sociedade. Àqueles que apesar das duras marcas não abandonam o caminho, não perdem a esperança no ser humano.

Agradecimentos

Susi, por todos os momentos de tolerância e incondicional apoio, tornando-se um exemplo de esposa, parceira e Mulher.

Susi (novamente!), mas agora por todos os momentos de intolerância, nunca me permitindo reduzir a vida ao desenvolvimento desta tese, sempre me forçando a lembrar de meus papéis de pai, filho, esposo, amigo, professor e profissional.

Lucas e Clara, meus filhos, por estarem o tempo todo ao meu lado e ajudando, como ninguém conseguiria fazer melhor, a marcar os momentos de tolerância e intolerância.

Noêmia e Luiz, meus pais, por toda base familiar, garantia inicial para a sustentação desta trajetória.

Prof. Luiz Carlos, orientador, exemplo de educador, pesquisador e cidadão, cuja trajetória influenciou a minha. Nunca existirão palavras que darão conta de meu eterno agradecimento.

Faccenda que, durante o desenvolvimento desta pesquisa, inicia como um parceiro de trabalho e termina como um grande amigo.

Luana, Clara Ede, Geisa e Geraldo, colegas de LOED, amigos na vida. Que alegria e sorte a minha em tê-los ao meu lado. Esse agradecimento daria um capítulo.

Colegas e professores do LOED pelo apoio acadêmico e parcerias.

Sérgio Bonjikian, médico homeopata, garantiu as condições para a continuidade na pesquisa, com remédios, conselhos e principalmente escuta, combinações que me colocaram em pé no trabalho e na vida.

Prof. Dalton, coorientador, pela disposição em estar presente e colocar, incondicionalmente, seus conhecimentos a serviço desta pesquisa.

Membros da banca, pelas contribuições e orientações, não só nos momentos de exame, mas também por toda a produção científica que fundamentou esta pesquisa.

Colegas de trabalho (Colégio Notre Dame de Campinas, Colégio Objetivo de Sorocaba, ESAMC, UNISAL, Instituto Internacional de Planejamento Educacional, Secretaria Municipal de Educação de Campinas, Faculdade SESI-SP de Educação), sempre compreendendo minhas ausências e colaborando para a superação das dificuldades inerentes ao trabalho.

Lista de Gráficos

Gráfico 1 - Proficiência média de alunos que permaneceram na mesma escola em função do NSE nos quatro primeiros anos de escolarização.....	29
Gráfico 2 - Associação matemática entre diferentes variáveis; exemplo hipotético.....	33
Gráfico 3 - Exemplo hipotético da análise de regressão linear	34
Gráfico 4 – Proficiência de alunos de quatro escolas; um exemplo	45
Gráfico 5 – Dispersão da proficiência em relação à média geral; exemplo	46
Gráfico 6 - Dispersão das proficiências dos alunos em relação à média de cada escola; exemplo	46
Gráfico 7 – Resíduos das médias das proficiência de cada escola em relação à média Geral; exemplo hipotético	47
Gráfico 8 - Representação gráfica do exemplo de Modelo de Nível 1	78
Gráfico 9 - Taxas de aprovação em cada ano de escolarização	100
Gráfico 10 - Taxas de aprovação em cada ano de escolarização	101
Gráfico 11 - Número de alunos que estão defasados no momento da aplicação do teste	101
Gráfico 12 - Partição da variância entre os três níveis em Leitura	127
Gráfico 13 - Retas dos diversos anos de escolarização - Leitura	131
Gráfico 14 - Valor agregado em Leitura por ano de escolarização em função da proficiência inicial	132
Gráfico 15 - Partição da variância entre os três níveis em Matemática	143
Gráfico 16 - Retas dos diversos anos de escolarização – Matemática	147
Gráfico 17 - Valor agregado em matemática por ano de escolarização em função da proficiência inicial	148
Gráfico 18 - Distribuição das proficiências medidas em Leitura e Matemática	162
Gráfico 19 - Representação gráfica das equações das retas, via HLM, dos diversos anos de escolarização tendo a Proficiência Final em função da Inicial	165
Gráfico 20 - Estimação da proficiência final em Leitura para diferentes proficiências iniciais	167
Gráfico 21 - Estimação da proficiência final em Matemática para diferentes proficiências iniciais	168
Gráfico 22 - Valor agregado médio em Leitura com base no resultado do Modelo de Nível 1	169
Gráfico 23 - Histograma da proficiência inicial em Leitura	170
Gráfico 24 - Dispersão dos ganhos em relação à proficiência inicial em Leitura.....	171
Gráfico 25 - Valor agregado médio em Matemática com base nos resultados do Modelo de Nível 1	171
Gráfico 26 - Distribuição de frequência da proficiência inicial em Matemática	172
Gráfico 27 - Dispersão dos ganhos em relação à proficiência inicial em Matemática	173
Gráfico 28 - Valor agregado por unidade da variável explicativa do Nível 2 em Leitura	179
Gráfico 29 - Valor agregado por unidade da variável explicativa do Nível 2 em Matemática	183
Gráfico 30 - Valor agregado por unidade da variável explicativa do Grupo 1 do Nível 3 em Leitura	188
Gráfico 31 - Valor agregado médio devido ao nível socioeconômico médio das escolas na proficiência final em Leitura	190

Gráfico 32 - Valor agregado por unidade da variável explicativa do Grupo 2 do Nível 3 em Leitura	193
Gráfico 33 - Valor agregado por unidade da variável explicativa do Grupo 3 do Nível 3 em Leitura	198
Gráfico 34 - Valor agregado por unidade da variável explicativa do Grupo 4 do Nível 3 em Leitura	202
Gráfico 35 - Valor agregado por unidade da variável explicativa do Grupo 1 do Nível em Matemática	209
Gráfico 36 - Valor agregado por unidade da variável explicativa do Grupo 1 do Nível 3 em Matemática	213
Gráfico 37 - Valor agregado por unidade da variável explicativa do Grupo 3 do Nível 3 em Matemática	218
Gráfico 38 - Valor agregado por unidade da variável explicativa do Grupo 4 do Nível 3 em Matemática	222
Gráfico 39 - Influência do MNSE sobre a proficiência final em Leitura.....	236
Gráfico 40 - Valor agregado em Leitura por unidade da variável MNSE	237
Gráfico 41 - Influência do MNSE sobre a proficiência final em Matemática	240
Gráfico 42 - Valor agregado em Matemática por unidade da variável MNSE	241
Gráfico 43 - Valor agregado na proficiência final quando o aluno fez pré-escola	250
Gráfico 44 - Valor agregado negativo que a indisciplina provoca na proficiência final	250
Gráfico 45 - Impacto negativo na proficiência final devido a outras atividades do diretor ...	251
Gráfico 46 - Valor agregado na proficiência final devido à ênfase na resolução de problemas	252
Gráfico 47 - Valor agregado na proficiência final em Leitura ligada à infraestrutura de Laboratório de Informática e acesso à Internet	252
Gráfico 48 - Valor agregado na proficiência final em Leitura devido à experiência do Diretor em Educação (FED25)	253
Gráfico 49 - Valor agregado na proficiência final em Leitura devido ao Ensino de Geografia, História e Ciências da Natureza (FEP49).....	253
Gráfico 50 - Valor agregado na proficiência final em Leitura ligado ao tempo usado em outras disciplinas	254
Gráfico 51 - Valor agregado na proficiência final em Leitura ligada à metodologia de ensino	254
Gráfico 52 - Valor agregado na proficiência final em Leitura devido ao Estímulo extrínseco para estudo (FA07).....	255
Gráfico 53 - Valor agregado na proficiência final em Leitura devido às expectativas negativas quanto à aprendizagem do aluno (FEP34)	256
Gráfico 54 - Valor agregado na proficiência final em Matemática devido às expectativas positivas quanto à aprendizagem do aluno (FEP35)	257
Gráfico 55 - Valor agregado na proficiência final em Leitura devido às interrupções das aulas por necessidade administrativa (FEP51).....	258
Gráfico 56 - Valor agregado na proficiência final em Matemática devido às interrupções das aulas por necessidade administrativa (FEP51)	258
Gráfico 57 - Valor agregado na proficiência final em Matemática devido às condições da Biblioteca (FEE05).....	259
Gráfico 58 - Valor agregado na proficiência final em Matemática quando o diretor trabalha em apenas uma escola (FED22)	259

Gráfico 59 - Valor agregado na proficiência final em Matemática devido ao dever de casa (FEP33)	260
Gráfico 60 - Valor agregado negativo que a variável que controla a existência da Sala de Leitura (FEE04) provoca na proficiência final	262
Gráfico 61 - Valor agregado negativo que a variável que controla o estímulo interno do aluno para fazer a tarefa de casa (FA11) provoca na proficiência final.....	262
Gráfico 62 - Valor agregado negativo que as variáveis associadas à autoestima (FA12) e (FA13)	263
Gráfico 63 - Variáveis com agregado negativo em Leitura	264
(FEE03- FEP28 - FA09)	264
Gráfico 64 - Variáveis com agregado negativo em Matemática (FED25 - FEP29 - FEP52 - FA14)	265
Gráfico 65 - Proficiências médias iniciais e finais de cada ano de escolarização da amostra usada nos Modelos Lineares Hierárquicos da pesquisa	276
Gráfico 66 - Variância dos Níveis 2 e 3 dos modelos em Leitura e em Matemática	278
Gráfico 67 - Valor agregado correspondente a uma unidade no nível socioeconômico do aluno (FA06)	279
Gráfico 68 - Valor agregado correspondente a uma unidade no nível socioeconômico médio da escola (FEA01).....	280
Gráfico 69 - Dispersão das escolas em função de seu valor agregado médio e da proficiência final média.....	282

Lista de Quadros

Quadro 1 - Aplicação dos instrumentos da pesquisa por onda	20
Quadro 2 - Níveis de habilidade e proficiência características	23
Quadro 3 - Atribuição de Valores para proficiência final e inicial de cada ano	65
Quadro 4 -Variáveis do Nível 1 - Ano de Escolarização	69
Quadro 5 - Valores atribuídos às variáveis indicadoras em cada ano de escolarização.....	69
Quadro 6 - Valores atribuídos aos coeficientes da reta de cada um dos anos de escolarização	74
Quadro 7 - Estimativas hipotéticas para exemplo do Modelo de Nível 1	76
Quadro 8 -Informações sobre os questionários aplicados.....	88
Quadro 9 - Comandos PASW para a criação da variável que controla a retenção	99
Quadro 10 - Variáveis do arquivo Nível 1 - Ano de Escolarização.....	105
Quadro 11 - Exemplo de transformação de uma variável categórica em variáveis indicadoras dicotômicas.....	106
Quadro 12 - Variáveis do arquivo do Nível 2 usadas no HLM.....	108
Quadro 13 - Variáveis do Nível 3 pertencente ao grupo 1	110
Quadro 14 - Variáveis do Nível 3 pertencente ao grupo 2.....	110
Quadro 15 - Variáveis do Nível 3 pertencente ao grupo 3.....	111
Quadro 16 - Variáveis do Nível 3 pertencente ao grupo 4.....	111
Quadro 17 - Resultados do Modelo Nulo em Leitura	126
Quadro 18 - Resultados Modelo de Nível 1 em Leitura	129
Quadro 19 - Resultados Modelo de Nível 1 em Leitura	130
Quadro 20 - Equação de regressão de cada ano em Leitura.....	131
Quadro 21 - Variáveis significantes no Modelo Nível 2 - Leitura.....	134
Quadro 22 - Estimativas das variâncias e do <i>deviance</i> no Modelo Nível 2 de Leitura.....	135
Quadro 23 - Variáveis significantes no Modelo Nível 3 - Leitura pertencentes ao Grupo 1..	137
Quadro 24 - Variáveis significantes no Modelo Nível 3 - Leitura pertencentes ao Grupo 2..	137
(continua)	137
Quadro 25 - Variáveis significantes no Modelo Nível 3 - Leitura pertencentes ao Grupo 3..	138
(continua)	138
(conclusão)	139
Quadro 26 – Variáveis significantes no Modelo Nível 3 – Leitura pertencentes ao Grupo 4	139
(continuação).....	140
Quadro 27 - Estimativas das variâncias e do <i>deviance</i> no Modelo Nível 3 em Leitura.....	140
Quadro 28 - Interpretação final das variáveis do Grupo 4 do Nível 3 em Leitura.....	141
Quadro 29 - Resultados do Modelo Nulo em Matemática.....	142
Quadro 30 - Resultados Modelo de Nível 1 em Matemática	144
Quadro 31 - Resultados Modelo de Nível 1 em Matemática	146
Quadro 32 -Equação de regressão de cada ano em Matemática	147
Quadro 33 - Variáveis significantes no Modelo Nível 2 – Matemática.....	149
Quadro 34 - Estimativas das variâncias e do <i>deviance</i> no Modelo Nível 2 de Matemática ...	150
Quadro 35 - Variáveis significantes no Modelo de Nível 3 - Matemática pertencentes ao Grupo 1.....	152
Quadro 36 - Variáveis significantes no Modelo de Nível 3 - Matemática pertencentes ao Grupo 2.....	153

Quadro 37 - Variáveis significantes no Modelo de Nível 3 - Matemática pertencentes ao Grupo 3.....	154
Quadro 38 - Variáveis significantes no Modelo de Nível 3 - Matemática pertencentes ao Grupo 4.....	155
Quadro 39 - Estimativas das variâncias e do <i>deviance</i> no Modelo Nível 3 de Matemática ...	156
Quadro 40 - Interpretação final das variáveis do Grupo 2 do Nível 3 em Matemática	157
Quadro 41 - Possíveis tipos de influência das variáveis explicativas as diferentes proficiências iniciais	175
Quadro 42 - Relação entre as variáveis explicativas do Nível 2 em Leitura e os anos de escolarização que elas influenciaram	176
Quadro 43 - Anos em que as variáveis de Nível 2 em Leitura impactam.....	176
Quadro 44 - Valores de proficiência finais e valor agregado por variável do Nível 2 em Leitura	177
Quadro 45 - Tipo de impacto da variável do Nível 2 em Leitura para cada ano	179
Quadro 46 - Relação entre as variáveis explicativas do Nível 2 em Matemática e os anos de escolarização que elas influenciaram	180
Quadro 47 - Anos em que as variáveis de Nível 2 em Matemática impactam	181
Quadro 48 - Valores de proficiência finais e valor agregado por variável do Nível 2 em Matemática.....	181
Quadro 49 - Tipo de impacto da variável do Nível 2 em Matemática a cada ano	183
Quadro 50 - Anos em que as variáveis do Grupo 1 do Nível 3 em Leitura impactam	186
Quadro 51 - Relação entre as variáveis explicativas do Grupo 1 do Nível 3 em Leitura e os anos de escolarização que elas influenciaram	186
Quadro 52 - Valores de proficiência finais e valor agregado por variável explicativa do Grupo 1 do Nível 3 em Leitura.....	187
Quadro 53 - Tipo de impacto das variáveis do Grupo 1 do Nível 3 em Leitura a cada ano ...	189
Quadro 54 - Interpretação final das variáveis do Grupo 1 do Nível 3 em Leitura.....	189
Quadro 55 - Relação entre as variáveis explicativas do Grupo 2 do Nível 3 em Leitura e os anos de escolarização que elas influenciaram	190
Quadro 56 - Anos em que as variáveis do Grupo 2 do Nível 3 em Leitura impactam	191
Quadro 57 - Valores de proficiência finais e valor agregado por variável explicativa do Grupo 2 do Nível 3 em Leitura.....	191
Quadro 58 - Tipo de impacto das variáveis do Grupo 2 do Nível 3 em Leitura a cada ano ...	194
Quadro 59 - Interpretação final das variáveis do Grupo 2 do Nível 3 em Leitura.....	194
Quadro 60 - Relação entre as variáveis explicativas do Grupo 3 do Nível 3 em Leitura e os anos de escolarização que elas influenciaram	195
Quadro 61 - Anos em que as variáveis do Grupo 3 do Nível 3 impactam em Leitura	195
Quadro 62 - Valores de proficiência finais e valor agregado por variável explicativa do Grupo 3 do Nível 3 em Leitura.....	196
Quadro 63 - Tipo de impacto das variáveis do Grupo 3 do Nível 3 em Leitura a cada ano ...	198
Quadro 64 - Interpretação final das variáveis do Grupo 2 do Nível 3 em Leitura.....	199
Quadro 65 - Relação entre as variáveis explicativas do Grupo 4 do Nível 3 em Leitura e os anos de escolarização que elas influenciaram	199
Quadro 66 - Anos em que as variáveis do Grupo 4 do Nível 4 em Leitura impactam	200
Quadro 67 - Valores de proficiência finais e valor agregado por variável explicativa do Grupo 4 do Nível 3 em Leitura.....	200
Quadro 68 - Tipo de impacto das variáveis do Grupo 4 do Nível 3 em Leitura a cada ano ..	203

Quadro 69 - Interpretação final das variáveis do Grupo 4 do Nível 3 em Leitura.....	203
Quadro 70 – Variáveis do Nível 3 que têm impacto na proficiência em Leitura no Ano 1 e com coeficiente diferente em outro ano	204
Quadro 71 - Variáveis que têm impacto no Ano 1 e têm o mesmo coeficiente em outro Ano (significa que a variável não foi significativa nesse ano).....	205
Quadro 72 - Relação entre as variáveis explicativas do grupo 1 do Nível 3 em Matemática e os anos de escolarização que elas influenciaram	206
Quadro 73 - Anos em que as variáveis do Grupo 1 do Nível 3 em Matemática impactam	206
Quadro 74 - Valores de proficiência finais e valor agregado por variável explicativa do Grupo 1 do Nível 3 em Matemática	207
Quadro 75 - Tipo de impacto das variáveis do Grupo 1 do Nível 3 em Matemática a cada ano	209
Quadro 76 - Interpretação final das variáveis do Grupo 1 do Nível 3 em Leitura.....	210
Quadro 77 - Relação entre as variáveis explicativas do Grupo 2 do Nível 3 em Matemática e os anos de escolarização que elas influenciaram	210
Quadro 78 - Anos em que as variáveis do Grupo 2 do Nível 3 em Matemática impactam	211
Quadro 79 - Valores de proficiência finais e valor agregado por variável explicativa do Grupo 2 do Nível 3 em Matemática	211
Quadro 80 - Tipo de impacto das variáveis do Grupo 2 do Nível 3 em Matemática a cada ano	214
Quadro 81 - Interpretação final das variáveis do Grupo 2 do Nível 3 em Leitura.....	214
Quadro 82 - Relação entre as variáveis explicativas do Grupo 3 do Nível 3 em Matemática e os anos de escolarização que elas influenciaram	215
Quadro 83 - Anos em que as variáveis do Grupo 3 do Nível 3 em Matemática impactam	215
Quadro 84 - Valores de proficiência finais e valor agregado por variável explicativa do Grupo 3 do Nível 3 em Matemática	216
Quadro 85 - Tipo de impacto das variáveis do Grupo 3 do Nível 3 em Matemática a cada ano	218
Quadro 86 - Interpretação final das variáveis do Grupo 2 do Nível 3 em Leitura.....	219
Quadro 87 - Relação entre as variáveis explicativas do Grupo 4 do Nível 3 em Matemática e os anos de escolarização que elas influenciaram	219
Quadro 88 - Anos em que as variáveis do Grupo 4 do Nível 3 em Matemática impactam	220
Quadro 89 - Valores de proficiência finais e valor agregado por variável explicativa do Grupo 3 do Nível 3 em Matemática	220
Quadro 90 - Tipo de impacto das variáveis do Grupo 4 do Nível 3 em Matemática a cada ano	223
Quadro 91 - Interpretação final das variáveis do Grupo 4 do Nível 3 em Matemática	223
Quadro 92 - Variáveis do Nível 3 que têm impacto na proficiência em Matemática no Ano 1 e com coeficiente diferente em outro ano	224
Quadro 93 - Variáveis que têm impacto no Ano 1 e têm o mesmo coeficiente em outro Ano (significa que a variável não foi significativa nesse ano).....	225
Quadro 94 - Síntese das variáveis significantes nos dois modelos	226
Quadro 95 - Resultados do Modelo de Nível 3 em Leitura	235
Quadro 96 - Resultados do Modelo de Nível 3 em Matemática	239
Quadro 97 - Variáveis com efeitos consistentes em todos os anos de acordo com o esperado	249

Quadro 98 - Variáveis com efeitos consistentes em todos os anos, mas contrários ao esperado	261
Quadro 99 - Variáveis sem significância estatística.....	266
Quadro 100 - Comparativo entre os resultados em Leitura e Matemática com o Modelo de Referência.....	288
Quadro 101 - Síntese das variáveis do arquivo de Nível 3	405

Lista de Tabelas

Tabela 1 Número de alunos em cada estrato por NSE.....	30
Tabela 2- Exemplo de associação matemática entre diferentes variáveis.....	32
Tabela 3 - Análise de regressão linear; exemplo hipotético.....	34
Tabela 4 - Resíduos em uma regressão linear simples; exemplo hipotético.....	36
Tabela 5 - Análise de regressão linear múltipla; exemplo hipotético.....	37
Tabela 6 -Resíduos com análise de regressão linear múltipla; exemplo hipotético.....	39
Tabela 7 - Proficiência dos alunos; um exemplo.....	45
Tabela 8 - Número de escolas com menos de quatro turmas.....	64
Tabela 9 - Número de variáveis e registros dos arquivos GERES.....	95
Tabela 10 - Frequência das variáveis que indicam a retenção.....	100
Tabela 11 - Número de alunos que trocaram de escola e participaram das cinco ondas.....	103
Tabela 12- Número de registros nos arquivos para análise no HLM.....	112
Tabela 13- Número de registros comuns no Nível 1 em Leitura e Matemática.....	113
Tabela 14 - Distribuição das escolas por polos e dependências administrativas.....	113
Tabela 15 - Número de alunos da amostra.....	114
Tabela 16 - Comparação das proficiências em Leitura.....	115
Tabela 17 - Comparação das proficiências em Matemática.....	116
Tabela 18 - Valor agregado em Leitura como diferença.....	132
da proficiência final estimada e da inicial.....	132
Tabela 19 - Valor agregado como diferença da proficiência final estimada e da inicial.....	148
Tabela 20 – Estimativas dos Componentes de Variância dos Modelos Nulos.....	162
Tabela 21- Correlação de Spearman entre medidas de proficiência sucessivas.....	163
Tabela 22- Resultado do teste de Normalidade das proficiências.....	164
Tabela 23- Parâmetros das retas obtidas com o Modelo de Nível 1 em leitura e Matemática.....	166
Tabela 24- Descritiva da proficiência Inicial em Leitura e Matemática.....	166
Tabela 25- Proficiência Final a cada ano, tendo como proficiência inicial a final do ano anterior.....	167
Tabela 26- Valor agregado médio nos diferentes anos de escolarização e o acumulado nos quatro anos.....	169
Tabela 27 - Valor agregado em Matemática em função do MNSE e da proficiência inicial.....	241
Tabela 28 - Valor Agregado Médio.....	274

Lista de Figuras

Figura 1– Desempenho médio da escola em função do valor agregado	14
Figura 2 -Agrupamento dos dados em níveis hierárquicos	44
Figura 3- Relação entre proficiência e nível socioeconômico de duas escolas hipotéticas	51
Figura 4 - Representações gráficas dos modelos para diferentes status da variância com variável explicativa centralizada	52
Figura 5 - Representações gráficas dos modelos para diferentes status da variância com variável explicativa não centralizada	54
Figura 6 - Diferentes medidas cognitivas ao longo do tempo (Nível 1) dentro de alunos (Nível 2) dentro de escolas (Nível 3)	64
Figura 7 - Representação esquemática das retas para cada ano de escolarização obtidas no Modelo de Nível 1	75
Figura 8 - Procedimentos adotados e arquivos relacionados na preparação dos arquivos utilizados nos modelos lineares hierárquicos	97
Figura 9 - Esquema de Reestruturação do arquivo Nível 1	98
Figura 10 - Esquema da definição das variáveis que armazenam as proficiências finais, proficiências iniciais e a defasagem em cada ano de escolarização	102
Figura 11- Exemplo de reestruturação de arquivo no PASW	104
Figura 12 - Número de registros nos arquivos de Nível 1 de Leitura e Matemática	112
Figura 13 - Representação gráfica dos possíveis tipos de influência das variáveis explicativas as diferentes proficiências iniciais	175

Lista de Siglas

AFE – Análise Fatorial Exploratória
HLM – Hierarchical Linear Model
VAM – Value added Model
VA – Valor agregado
NSE – nível socioeconômico
MNSE – Nível socioeconômico médio da escola

Lista de Apêndices

Apêndice 1 - Arquivos resultantes da AFE e itens que compõem cada constructo	305
Apêndice 2 - Análise descritiva das variáveis usadas nos modelos	357

Lista de Anexos

ANEXO 01 - Interpretação pedagógica dos níveis de proficiência em Leitura.....	409
ANEXO 02 - Interpretação pedagógica dos níveis de proficiência em Matemática	413
ANEXO 03 - Relatório HLM – Leitura – Modelo Nulo.....	415
ANEXO 04 - Relatório HLM – Leitura – Modelo de Referência.....	419
ANEXO 05 - Relatório HLM – Leitura – Modelo Nível 2.....	423
ANEXO 06 - Relatório HLM – Leitura – Modelo Nível 3	431
ANEXO 07 - Relatório HLM – Matemática – Modelo Nulo	447
ANEXO 08 - Relatório HLM – Matemática – Modelo Referência	451
ANEXO 09 - Relatório HLM – Matemática – Modelo Nível 2.....	455
ANEXO 10 - Relatório HLM - Matemática- Modelo Nível 3.....	463
ANEXO 11 - Relatório HLM – Modelo de Referência em Leitura.....	475
ANEXO 12 - Relatório HLM – Modelo de Referência em Matemática	479

APRESENTAÇÃO

Esta pesquisa faz parte do conjunto de outras desenvolvidas no Laboratório de Observação e Estudos Descritivos (LOED) da Faculdade de Educação da Unicamp que, dentre outros temas, têm estudado as questões ligadas à Avaliação Educacional.

A atenção do LOED nessa área deve-se aos impactos da crescente implementação da avaliação educacional para a definição e avaliação de políticas públicas, ocorrida no Brasil após a década de 1990 (e, em algumas outras partes do mundo, três décadas antes), e, mais recentemente, aos processos de responsabilização vertical (*accountability*).

Parte das reflexões e dos debates desse conjunto de pesquisas do LOED está registrada em uma vasta produção acadêmica tratando da avaliação de sistema, da avaliação institucional e da avaliação do ensino e da aprendizagem (FREITAS, SORDI, MALAVASI, FREITAS, 2009), sempre buscando contribuir para o aperfeiçoamento das práticas avaliativas nos mais diversos espaços onde elas se efetivam.

Identificando a efetivação dos sistemas avaliativos pelas políticas educacionais, desde o início dos anos 2000, foram realizados encontros que contaram com a participação de especialistas de diferentes centros universitários para colocar em evidência as diferentes concepções da produção na área da avaliação educacional. As análises foram retiradas do âmbito exclusivamente técnico e colocadas no âmbito da Educação, destacando-se seus limites e possibilidades, dadas as metodologias usadas nos processos avaliativos. Os debates ocorridos nesses encontros estão registrados em Freitas (2002), Freitas, Gatti e Souza (2003), Freitas, Belloni e Soares (2003), Malavazi, Bertagna e Freitas (2006) e Freitas *et al.* (2012).

A preocupação com o uso dos modelos estatísticos que permeiam processos avaliativos desde a coleta dos dados até a divulgação de seus resultados “suscita também uma nova agenda para a pesquisa em avaliação educacional, em estatística e, em particular, em estatística educacional, nos termos genericamente definidos pela *Encyclopedia of Statistical Sciences*” (FERRÃO e KLEIN, 2013, p.7). Por esse motivo, o

LOED também tem desenvolvido pesquisas quantitativas¹, visando à apropriação do conhecimento acerca da Teoria de Resposta ao Item, Análise por Envoltória de Dados, Análise Exploratória de Dados e Modelos Lineares Hierárquicos, que são alguns dos recursos estatísticos usados na produção de dados educacionais oriundos das avaliações de larga escala.

Esta pesquisa insere-se nesse contexto com o objetivo de identificar, por meio de modelos lineares hierárquicos, os fatores intraescolares e extraescolares que interferem na proficiência dos alunos das séries iniciais do Ensino Fundamental, explorando o conceito de valor agregado e efeito-escola e usando dados longitudinais. Com esse objetivo, também se insere no Projeto GERES, do qual o LOED participou colocando foco sobre a apropriação dos resultados das avaliações de larga escala pela comunidade interna das escolas, do que resultaram várias dissertações e teses produzidas no LOED² (DALBEN, 2008; BETINI, 2009; ALMEIDA, 2009; CANGUSSU, 2010; MORAES, 2010; COSTA, 2012; e ALMEIDA, 2014).

Assim, ao mesmo tempo em que esta pesquisa é apenas mais uma que pretende colaborar com essa produção de conhecimento, é também fortemente influenciada por ela, atribuindo a devida relevância aos resultados das avaliações educacionais e também aos achados de pesquisa. Considerando que a especificidade da linguagem da estatística, cujas simbologia e denotações carregam um ou mais conceitos, pode gerar alguma dificuldade no entendimento deste trabalho, buscou-se usar a linguagem mais simples possível, com a intenção de ajudar aqueles que, apesar de não acostumados com textos dessa natureza, têm interesse em apropriar-se, minimamente, da noção de como os dados são tratados e analisados com os modelos usados nesta pesquisa.

Este texto, bastante extenso por conta do grande número de tabelas e gráficos³ que o compõe, está organizado em seis capítulos, onde a expressão “variável explicativa” deve ser entendida como fator que interfere na aprendizagem.

1 Dentre essas pesquisas, podem ser citadas as de Ferrão (2003), Miranda (2006), Stevão (2008), Rodrigues (2005), Miranda (2008), Poli (2007) e Faccenda, Dalben e Freitas (2011).

2 O LOED é um dos cinco centros de pesquisas que se envolveram no desenho e desenvolvimento do Projeto GERES. A descrição de algumas das produções científicas desses centros de pesquisa são encontradas em Brooke e Bonamino (2011).

3 Todos os elementos inseridos no texto, tais como tabelas, quadros, gráficos e figuras, que não possuem a indicação da fonte foram produzidos nesta pesquisa. Portanto, as fontes só foram especificadas quando esses elementos foram produzidos em outros trabalhos.

No Capítulo 1, é feita uma breve apresentação do contexto em que os modelos de valor agregado estão inseridos e é apresentada a problemática da pesquisa e também alguns conceitos centrais, tais como eficácia escolar, equidade escolar, valor agregado, efeito-escola, modelagem estatística, proficiência e sua relação com a medida de conhecimento. O Projeto GERES é minuciosamente apresentado também nesse capítulo.

O Capítulo 2 visa a apresentar, no contexto da estatística educacional, os conceitos ligados à regressão linear e aos modelos lineares hierárquicos (ou modelos multinível, como também são conhecidos). São apresentadas a metodologia e a lógica usada para o ajuste dos modelos estatísticos usados nesta pesquisa, ao mesmo tempo em que se justificam:

- a maior adequação do uso dos modelos lineares hierárquicos em relação à regressão linear;
- a necessidade do uso de modelos de valor agregado na busca dos fatores intraescolares e extraescolares que interferem na proficiência dos alunos;
- e
- a necessidade de uso de dados longitudinais.

A primeira parte deste segundo capítulo busca evitar o uso da simbologia inerente à estatística para desenvolver as noções básicas, em detrimento da linguagem técnica⁴, feito que não é possível na mesma intensidade na segunda parte, quando o uso dessa simbologia é inevitável.

O terceiro capítulo apresenta todo o processo de tratamento dos dados fornecidos pelo Projeto GERES e a preparação dos arquivos, com as variáveis resultantes desse tratamento. Os dados foram preparados para a análise dos fatores usando os modelos lineares hierárquicos.

No Capítulo 4, apresenta-se o processo de ajuste dos modelos lineares hierárquicos usados na identificação das variáveis estatisticamente significantes na proficiência final em leitura e em matemática nos quatro anos de escolarização abarcados pelo Projeto GERES, considerando a inter-relação entre todas essas variáveis simultaneamente. Essa proposição metodológica foi feita justamente para verificar os efeitos das variáveis, tendo em vista que nenhuma delas influencia isoladamente a

4 As equações foram somente numeradas quando referenciadas em outras partes do texto.

proficiência, uma vez que o fenômeno educacional é complexo. Ao final do capítulo, é apresentada a relação dessas variáveis.

No Capítulo 5, são analisados não só os efeitos das variáveis relacionadas ao final do Capítulo 4, mas também a intensidade dessa associação sobre a proficiência final de alunos com diferentes proficiências iniciais. As estimativas dessas influências foram feitas sem a especificação do intervalo de confiança. Para uma interpretação pedagógica dos efeitos de cada variável, deve ser analisada sua composição pela Análise Fatorial Exploratória, descrita no Apêndice 01, e considerada sua análise descritiva, registrada no Apêndice 02.

Dada a instabilidade das análises percebidas nos dois capítulos anteriores, o sexto capítulo apresenta o ajuste dos modelos que permitem a análise das variáveis isoladamente, tomando-se apenas o nível socioeconômico do aluno e o nível socioeconômico médio da escola. Essas análises são importantes para fundamentar algumas conclusões acerca do uso dos modelos estatísticos, mais especificamente, dos modelos de valor agregado.

O texto é encerrado com a análise pedagógica dos efeitos encontrados nos diversos modelos e com a apresentação de subsídios para o adequado uso das avaliações de larga escala no âmbito das políticas educacionais e da escola.

**CAPÍTULO 1 -
O CAMPO DE INSERÇÃO DESTA PESQUISA**

Desde a década de 1960, quando foram publicados os resultados da pesquisa americana conduzida por James S. Coleman constatando que “as diferenças entre as escolas somente são responsáveis por uma pequena fração das diferenças no desempenho dos alunos” (COLEMAN *et al*, 1966, p. 131) e que as diferenças socioeconômicas entre esses alunos era o fator responsável pelas diferenças em seus desempenhos, muitas pesquisas com foco na eficácia e equidade escolar vêm sendo realizadas com finalidades acadêmicas e governamentais.

O Relatório Coleman colocou em xeque o conceito de igualdade de oportunidades educacionais vigente na sociedade americana. Tal conceito levava à crença de que, se os alunos, independentemente de sua característica social e étnica, tivessem acesso a escolas com bons recursos humanos, físicos e materiais, seus desempenhos seriam semelhantes aos dos demais alunos, ou seja, “as escolas seriam vistas como provedoras de igualdade de oportunidades educacionais quando diferentes alunos em diferentes escolas tivessem desempenhos semelhantes” (MADAUS, AIRASIAN e KELLAGHAN, 2008, p. 75).

O Relatório Coleman frustrou as expectativas com relação às escolas, pois

muitos acreditavam que eram os *inputs* ou insumos escolares, em termos dos equipamentos e outras condições de funcionamento da escola, que determinavam seus *outputs* ou resultados, segundo um modelo de entrada-saída que tratava a escola como uma caixa-preta e igualava o conceito de qualidade da escola ao da qualidade ou quantidade dos seus insumos. Aparentemente, a sabedoria popular estava errada, mas ao mesmo tempo, não havia tanta diferença entre as escolas que pudesse determinar o que aconteceria em casos de diferenças maiores. A polêmica estava instalada (BROOKE e SOARES, 2008, p. 15).

A reação política e acadêmica provocada pelo Relatório Coleman, que resultou de uma pesquisa sobre a Igualdade de Oportunidades Educacionais, colocou-o como um marco para a sociologia da educação, por ter sido publicado em uma época em que o processo da análise crítica de mitos e crenças sociais assumia um importante papel na vida nacional americana.

O processo foi especialmente evidente nos Estados Unidos, nos anos de 1950 e de 1960, como resultado de duas grandes forças. Uma combinação de eventos políticos, econômicos e sociais levou a acordos em um nível nacional nunca vistos antes para reestruturar a sociedade superando as injustiças e instabilidades associadas aos problemas históricos de pobreza e raça. Ao mesmo tempo, a metodologia das ciências sociais avançou rapidamente devido, principalmente, ao desenvolvimento de computadores de alta velocidade que permitiam que investigações complexas de

grande escala pudessem se realizar em um grau nunca visto antes (MOSTELLER e MOYNIHAN, 2008, p.34).

Com o grande impacto que gerou, o relatório passou a ser centro de atenções, recebendo críticas nessas duas diferentes perspectivas. De acordo com o que sintetizam Bonamino e Franco (1999, p. 103), uma dessas perspectivas é a política, pois

centrava-se no questionamento das políticas de educação compensatória. O argumento principal era o de que o fracasso das crianças de minorias étnicas e de camadas populares relacionava-se com a desvalorização de culturas específicas, promovida tanto pelos currículos escolares quanto pelos instrumentos de pesquisa. De acordo com esta linha de argumentação, as políticas de educação compensatória seriam apenas novos instrumentos de reafirmação de certas culturas, em detrimento de outras.

No âmbito acadêmico, o Relatório Coleman recebeu críticas metodológicas por erros mecânicos e processuais no tratamento dos dados coletados, abordagem dada ao problema de dados faltantes, composição da amostra com viés seletivo e, principalmente, erros nos métodos de análise de dados (MADAUS, AIRASIAN e KELLAGHAN, 2008). Assim, apesar de não mostrar outros resultados, a crítica na perspectiva metodológica,

os críticos argumentavam que as mencionadas pesquisas não possuíam instrumentos adequados para a mensuração de detalhes relevantes das variáveis intraescolares. De acordo com esta linha de argumentação, e dada a tendência de que investimentos educacionais ocorram em maior grau em escolas frequentadas por crianças de nível socioeconômico mais alto, a não mensuração explícita desses investimentos e sua conseqüente não inclusão nos modelos de regressão causaria, segundo os críticos, uma superestimação do papel do nível socioeconômico como fator de sucesso na escolarização (BONAMINO e FRANCO, 1999, p. 103)

1.1) O campo de estudos da eficácia escolar

A busca da superação dessas críticas fomentou o aumento das pesquisas acadêmicas nessa área, criando uma área de estudos denominada Eficácia Escolar, e, já no início da década de 1970, foi possível mensurar de maneira mais robusta as variáveis intraescolares e extraescolares que podem interferir no desempenho dos alunos. Ferrão (2003) aponta como exemplo desse conjunto de pesquisas o Relatório Plowden (PLOWDEN COMMITTEE, 1967), na Inglaterra; o estudo longitudinal desenvolvido pelo Institut National d'Études Démographiques que examinou amostra representativa em escala nacional composta por 17.641 alunos cujas trajetórias escolares foram acompanhadas desde a conclusão do último

ano da escola primária, em 1962, até o ano de 1972 na França (FORQUIN, 1995); a pesquisa *Inequality: a Reassessment of the Effect of Family in Schooling in America* (JENCKS *et al.*, 1972; JENCKS, 2008), nos Estados Unidos, bem como os estudos *Fifteen Thousand Hours* (RUTTER *et al.*, 1979) e *School Matters* (MORTIMORE *et al.*, 1988). Enquanto os primeiros estudos apresentaram resultados semelhantes aos do Relatório Coleman, os dois últimos mostraram que “as escolas variam consideravelmente quanto à sua eficácia em promover o progresso cognitivo e não cognitivo dos alunos” (FERRÃO, 2003, p.15).

Desde então, “inúmeros estudos travam debates com as teses defendidas pelos autores que, apesar de polêmicas, não foram ainda completamente refutadas” (ALVES e SOARES, 2008, p. 26), afinal, em alguma medida, as escolas interferem na aprendizagem de seus alunos. A dificuldade das pesquisas educacionais em refutar a tese de Coleman parece ser coerente e até previsível quando considerada a complexidade do fenômeno educacional, uma vez que qualquer síntese implica uma redução da realidade, por ser inevitável a perda de algumas nuances da associação das variáveis, ao mesmo tempo em que alguns vieses podem ser inseridos na modelagem usada que permite tal síntese.

Outro aspecto que estimulou as pesquisas nessa área deve-se ao fato de que, ainda na década de 1970, a avaliação educacional deixa de ser fonte de dados apenas para pesquisas acadêmicas acerca da eficácia escolar e se torna recurso central das políticas educacionais. No contexto político-econômico mundial, atendendo a demandas de políticas transnacionais, em muitos países, o Estado estava migrando de uma condição que buscava garantir os direitos educacionais (dentre outros) de seus cidadãos para Estado regulador (AFONSO, 2000; BARROSO, 2005), evidenciando serem indissociáveis os debates políticos e técnicos subjacentes a essa temática.

Nesse contexto, Goldstein e Woodhouse (2008), fundamentados nas pesquisas realizadas por Pring (1995), Elliot (1996), Hamilton (1996), Gibson e Asthhana (1998) e Willmott (1999), ao analisarem a incorporação dos estudos de eficácia escolar na elaboração das políticas públicas, observam que há um conjunto de críticas ligadas ao abuso por parte do governo britânico (seja ele Conservador ou Trabalhista). Segundo eles⁵, “o governo assumiu

5 Segundo Goldstein e Woodhouse (2008), ainda há três outros grupos de críticas: a demasiada simplificação das complexas casualidades associadas ao ensino e ao desvio da atenção para as tabelas de desempenho, a “teoria” no trabalho com os estudos e eficácia escolar é pouco mais do que a reificação das relações empíricas, e muitas pesquisas em eficácia escolar são simplesmente de baixa qualidade.

a eficácia escolar porque ela enfatiza a responsabilidade que as escolas têm sobre os resultados educacionais em lugar do governo” (GOLDSTEIN e WOODHOUSE, 2008, p. 412), isto é, as pesquisas mostram que as políticas governamentais usam os estudos em eficácia escolar para rotular as escolas como responsáveis únicas de seu sucesso ou fracasso. Além disso,

Existem muitas pesquisas sobre a eficácia escolar que reconhecem a verdadeira complexidade da educação, que está relacionada a processos mais sutis em vez de resultados de exames e testes; e muitas outras que são confessadamente críticas em relação ao governo de subverter as pesquisas em seu próprio interesse (GOLDSTEIN e WOODHOUSE, 2008, p. 414).

Essa crítica permite entender a observação feita por Murillo (2008, p. 468) de que

O termo “eficácia escolar”, assim como a linha de pesquisa que leva o seu nome, tem uma importante conotação negativa em grande parte de nossos países. Do nosso ponto de vista, isso foi gerado, em grande medida, por uma confusão conceitual, talvez interessada, que fez com que se considerassem como estudos de eficácia escolar trabalhos enquadrados na linha de produtividade escolar. A diferença entre ambos é radical. Enquanto estudos sobre a produtividade têm raízes e desenvolvimento estritamente economicistas e buscam otimizar os insumos para conseguir os produtos (o que se entende por eficiência), os trabalhos de eficácia escolar são estudos puramente pedagógicos, aos quais interessa analisar quais processos redundam em melhores objetivos (ou seja, eficácia).

Assim, com a conjunção de interesses acadêmicos, políticos e econômicos, as pesquisas que sucederam o Relatório Coleman, ao buscarem contrapontos aos seus achados empíricos ou com o objetivo de identificar os fatores que podem ser potencializados por meio de políticas públicas, visando ao aumento da associação das escolas no desempenho dos alunos, colocam a avaliação educacional e o uso de seus resultados no centro de inúmeros debates, trazendo informações essenciais para a elaboração de políticas.

O sentido de poder que tem a avaliação explica o grande protagonismo que ela exhibe hoje em quase todos os âmbitos de intervenção social. É necessário insistir nessa afirmação para deixar claro que a avaliação não é simplesmente coisa de escola, nem somente da educação. Ela ultrapassa largamente esses âmbitos e cada vez mais vem ampliando seus alcances e seus efeitos na economia e na política. Os países modernos, ao menos nos últimos vinte e pouco anos, estão empenhados em produzir profundas mudanças estruturais, principalmente no que se refere às transformações na produção e àquilo que geralmente é conhecido como “reformas do Estado”. No centro dessas transformações, como instrumento fundamental e motor das transformações, está a avaliação (DIAS SOBRINHO, 2002, p. 39).

Diante dessa conjuntura, percebe-se que o conceito de eficácia escolar “não é neutro e muito menos consensual, pois a definição da eficácia de uma escola prende-se a valores e critérios que são muitas vezes competitivos” (FERRÃO, 2003, p. 15) e se inserem em um campo⁶ caracterizado por certo grau de conflito. Já para Murillo (2008, p. 468),

é possível distinguir dois grandes objetivos sobre a eficácia escolar sob o ponto de vista técnico: 1- estimar a magnitude dos efeitos escolares e analisar suas propriedades científicas (consistência entre áreas, estabilidade, eficácia diferencial e permanência); 2- identificar os fatores da sala de aula, da escola e do contexto que fazem com que uma escola seja eficaz.

O conceito de eficácia escolar usado pelo autor é diferente do que atualmente é vigente, de forma preponderante, na política educacional brasileira. Como as escolas são ordenadas segundo a média da proficiência de seus alunos em exames seccionais, a medida de sua eficácia passa a ser essa média, que desconsidera a associação do contexto de vida do aluno antes mesmo de sua entrada na escola, efeito em que a escola não tem qualquer associação. Por esse motivo, para esse autor, a escola eficaz é aquela que consegue o desenvolvimento integral de todos e de cada um de seus alunos além do que seria esperado, considerando seu rendimento prévio e o contexto socioeconômico em que estão inseridos (MURILLO, 2001, 2003, 2011). Essa definição é compatível com a de Mortimore (1991, 1998), o precursor da definição desse conceito, que afirma ser escola eficaz aquela em que o progresso do aluno vai além do que seria esperado, levando em consideração as suas características ao entrar na escola. Ao comparar a definição desses dois autores, Ferrão (2003) conclui que

as definições são semelhantes – a ambas está subjacente a noção de valor agregado, como critério de eficácia, e ambas estabelecem a importância de considerar as características do aluno ao entrar na escola, para que a escola possa ser aferida, também por promover a equidade. Murillo enuncia explicitamente quais são as características dos alunos que devem ser consideradas – o status socioeconômico e cultural, bem como o conhecimento prévio acumulado pelo aluno até a sua entrada na escola (FERRÃO, 2003, p. 15).

6 Ver Lahire (2002).

Assim, diante de diferentes conceitos de eficácia, diferentes modelos estatísticos são requeridos para responder as perguntas que orientam as pesquisas sobre eficácia escolar. Para o National Research Council and National Academy of Education (2010), há quatro tipos de modelos estatísticos que podem ser usados em estudos de eficácia escolar com base em proficiências medidas em testes padronizados:

- Modelos de medidas seccionais. Estes modelos baseiam-se em medidas de proficiência coletadas em um momento da escolarização dos alunos. Nesses modelos, a eficácia escolar é medida por meio da média das proficiências obtidas pelos alunos agrupados segundo uma unidade de análise, que pode ser turma, professor, escola, rede de escolas, municípios, estados, regiões ou países;

- Modelos de coortes. Nesses modelos, a eficácia escolar é analisada por meio da comparação da média dos desempenhos dos alunos de diferentes amostras de uma mesma série e em momentos diferentes;

- Modelos de curva de crescimento. A eficácia escolar é analisada por meio do desempenho médio da mesma amostra de alunos em distintos momentos da escolarização. Com esse modelo, é possível estimar o progresso na proficiência dos alunos de um período para outro. Este modelo permite a determinação do valor agregado por meio da diferença entre os dois momentos distintos em que as proficiências foram medidas;

- Modelos de valor agregado. São modelos geralmente mais complexos que têm por objetivo estimar a fração do crescimento do aluno ao longo do tempo que se deve a cada um dos fatores controlados e comparar desempenho de diferentes unidades de análise (tais como professores, turmas, escolas, municípios, etc.) com a média obtida pelas demais unidades de análise. Nesses modelos, o valor agregado pode ser definido como sendo a “medida do progresso médio dos alunos, durante o período de tempo em que ele é exposto a determinado ambiente educativo, comparativamente com que ele obteria em outras escolas em que poderia estar estudando” (FERRÃO, 2003, p. 16).

Evidentemente, a coleta de dados seccionais usada apenas no primeiro tipo de modelo é muito mais simples e menos onerosa. No entanto, traz inúmeras restrições por não controlar a proficiência de entrada do aluno, caracterizando-se como uma limitação técnica, inerente aos estudos de eficácia escolar que utilizam modelos dessa natureza, como aponta Franco (2001, p.129):

Conforme já enfatizado por diversos especialistas, estudos que testam apenas uma vez os alunos apresentam importantes limitações para a explicação do efeito das escolas na aprendizagem. A razão disso reside no fato de que a medida de proficiência é uma medida do aprendizado dos alunos ao longo de muitos anos. Portanto, essa medida não pode ser explicada em função do passado recente dos alunos [...] Nos estudos longitudinais, a medida de proficiência prévia é usada para filtrar o efeito do aprendizado prévio do aluno. Deste modo, a segunda proficiência mensurada (após a filtragem do aprendizado prévio) representa o aprendizado ao longo do ano. Por certo, esse aprendizado ao longo do ano pode ser explicado pelas características da escola freqüentada.

Assim, a opção metodológica (que é, antes de tudo, política e, portanto, ideológica) de associar os resultados dos testes à eficácia escolar como recurso central para a tomada de decisões carrega dois aspectos problemáticos. O primeiro deles é que as escolas e os trabalhadores que nela atuam passaram a ser avaliados pela proficiência dos alunos, um indicador sobre o qual têm associação apenas em parte, uma vez que a proficiência é influenciada por fatores extraescolares e, de maneira especial, pelo nível socioeconômico dos estudantes antes de sua entrada na escola e durante o período de escolarização. O segundo refere-se ao fato de que as escolas e seus trabalhadores têm influência somente no período em que o aluno permanece na escola.

Destaca-se então que, quantitativamente, apenas parte da proficiência de um aluno é afetada pelo processo de escolarização, sendo, portanto, sobre essa parte, denominada efeito-escola, que todo o sistema educativo, no qual se inserem o professor e a escola, pode e deve ser responsabilizado. Para essa concepção, o efeito-escola deve ser entendido como “a porcentagem da variação no rendimento dos alunos devido às características processuais da escola onde estão matriculados. E é medida como uma porcentagem da variância” (MURILLO, 2005, p. 32).

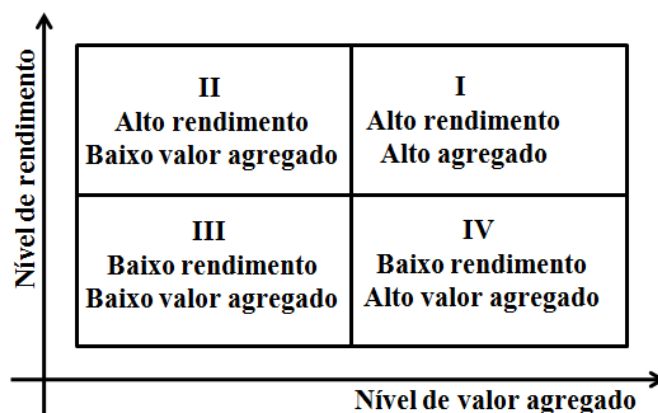
O estudo da eficácia escolar com base em dados longitudinais, o que permite determinar o valor agregado, favorece uma análise mais robusta e pode instrumentalizar melhor cada uma das unidades de análise (professor, escola, município, estado, país) para realizar de maneira mais efetiva seu trabalho em busca da equidade e da eficácia.

No entanto, a análise com valor agregado não é suficiente, pois

Algumas das escolas onde os alunos atingem níveis de desempenho elevado pouco contribuem para esse desempenho: tendencialmente escolas frequentadas por alunos com elevado nível de desempenho à entrada (escolas seletivas) e provenientes de grupos sociais favorecidos. Em contrapartida, as escolas que recebem alunos com baixo desempenho e onde, no final do período sob avaliação, o desempenho se mantém aquém da meta educativa estabelecida, podem, apesar disso, ter contribuído fortemente para alavancar a aprendizagem desses alunos e, portanto, apresentam elevado valor agregado. A sua contribuição não está refletida no indicador de status mas sim no indicador de valor agregado (FERRÃO e COUTO, 2013, p. 139).

A Figura 1 permite a visualização desse contexto. As escolas são posicionadas no gráfico segundo o desempenho médio de seus alunos em função do valor agregado médio no período.

Figura 1– Desempenho médio da escola em função do valor agregado



Fonte: National Research Council and National Academy of Education, 2010, p. 7.

Nesse sentido, a avaliação com base na análise dos ganhos de pontuação no teste dos alunos, isto é, no valor agregado na proficiência inicial, é uma perspectiva interessante que tem se difundido no cenário mundial, sobretudo nos Estados Unidos, que acumula vasta experiência. Para essa análise, são usados como recurso estatístico os modelos de valor agregado (*Value-added models - VAMs*). Esses modelos requerem dados que acompanham o crescimento acadêmico individual dos alunos em um período específico de escolarização, marcado por medidas de proficiência no início e no final. Com essas duas medidas, é possível estimar as contribuições dos diversos fatores relacionados ao sistema educativo para esse ganho de proficiência.

Porém, apesar do entusiasmo que esses modelos geraram entre pesquisadores e nos formuladores de políticas em diversas partes do mundo, dadas as suas possibilidades, há várias

revisões técnicas acerca dos modelos de valor agregado. Isso se refere a um conjunto de preocupações devido às limitações técnicas inerentes aos modelos estatísticos de qualquer natureza, tais como os problemas ligados ao equilíbrio entre a complexidade e a transparência, viés, não-aleatoriedade dos dados coletados, precisão, estabilidade, qualidade dos dados e interpretações causais (NATIONAL RESEARCH COUNCIL AND NATIONAL ACADEMY OF EDUCATION, 2010).

Para Braun (2005, p. 3), “a implementação de tais modelos e os usos propostos dos resultados levantam uma série de questões práticas, técnicas e até mesmo filosóficas”. Segundo esse mesmo autor, a “preocupação fundamental é que, se o objetivo é estabelecer uma relação causal, nenhum modelo estatístico, mesmo que complexo, e nenhum método de análise, mesmo que sofisticado, pode compensar totalmente a falta de randomização” (BRAUN, 2005. p.8).

Quanto ao problema da distribuição não-aleatória de alunos entre as escolas e até mesmo entre as turmas dentro de uma mesma escola, a modelagem de valor agregado permite ajustar as diferenças preexistentes entre os alunos, utilizando-se os resultados dos testes anteriores e outras características dos alunos e da escola. No entanto, dependendo do tipo de modelo selecionado, bem como do número e das características das variáveis que serão analisadas, os efeitos podem ser superestimados ou subestimados. Os modelos de valor agregado tendem a resumir a eficácia de uma unidade de análise (professor, escola, etc.) em um único número, mas considerando que a qualidade dessa unidade de análise é multidimensional. Assim, com frequência, não é possível dizer que uma unidade de análise é melhor do que outra.

Além disso, há o problema da precisão e da estabilidade. Pesquisas sobre a precisão das estimativas mostram grandes erros devido à amostragem (NATIONAL RESEARCH COUNCIL AND NATIONAL ACADEMY OF EDUCATION, 2010). Amostras pequenas representam um problema particular quando se estimam os efeitos das unidades de análise. Quando o número de alunos por unidade de análise é pequeno e apenas alguns alunos apresentam baixo desempenho, há uma substancial diminuição na eficácia dessa unidade de análise, ocorrendo o contrário quando apenas alguns alunos obtêm desempenho muito alto. Outras causas de instabilidade podem ocorrer devido às diferenças no desempenho de uma unidade de análise de ano para ano, ou em decorrência de alterações no contexto onde se

inserir, tais como liderança escolar, efeitos de políticas pontuais e mobilidade dos estudantes (NATIONAL RESEARCH COUNCIL AND NATIONAL ACADEMY OF EDUCATION, 2010).

A relevância desses problemas depende do objetivo do estudo. Se a finalidade é meritocrática, seu uso é inadequado. Se sua finalidade é identificar pontos de apoio a essa unidade de análise para que melhore a qualidade da oferta de ensino, sua utilização é de grande valia.

Além disso, os resultados das avaliações externas são apenas indicadores e como tal devem ser tratados (SORDI, 2002). Esses resultados não são verdades, são apenas dados quantitativos complementares à reflexão sobre a escola e o professor em processos de avaliação institucional. Afinal,

Explicar o desempenho de uma escola implica ter alguma familiaridade e proximidade com o seu dia-a-dia, o que não é possível para os sistemas de larga escala realizados pela Federação ou pelos estados, distantes da escola (FREITAS, 2007, p. 979).

No Brasil, onde é alta a correlação entre o nível socioeconômico e cultural dos alunos e as condições escolares associadas à eficácia escolar (FLETCHER, 1998; CÉSAR e SOARES, 2001; FERRÃO *et al.*, 2001; SOARES, CÉSAR e MAMBRINI, 2001; FERRÃO e FERNANDES, 2001; SOARES, 2004; OECD, 2004; SOARES e COLLARES, 2006; SOARES e ANDRADE, 2006; FRANCO *et al.*, 2007; SOARES e CANDIAN, 2007; FREITAS, 2007), as preocupações com a relação entre a avaliação educacional, feita em larga escala, e as políticas públicas ficam mais acentuadas no meio acadêmico na década de 1990, como registrado em Freitas (2002).

Em meio a essa discussão e desconsiderando, senão contrariando, tal contexto, a partir da década de 1980 (no Brasil, uma década depois), a avaliação educacional deixa de ser referência para pesquisas educacionais com dados seccionais e torna-se um recurso central e determinístico nas políticas públicas.

Se há uma política que avançou no Brasil, nos últimos 15 anos, foi a implantação dos sistemas de avaliação educacional. Neste período, inúmeras iniciativas deram forma a um robusto e eficiente sistema de avaliação em todos os níveis e modalidades de ensino, consolidando uma efetiva política de avaliação educacional. Considerada hoje uma das mais abrangentes e eficientes do mundo, a política de avaliação

engloba diferentes programas, tais como o Sistema Nacional de Avaliação da Educação Básica – Saeb, o Exame Nacional do Ensino Médio – Enem, [...], a Prova Brasil e o Índice de Desenvolvimento da Educação Básica – Ideb. Em conjunto, estes sistemas, ao lado da Avaliação da Pós-Graduação da Capes – o mais antigo sistema de avaliação do país no setor educacional –, configuram um macrossistema de avaliação da qualidade da educação brasileira (CASTRO, 2009, p.5).

Especificamente na Educação Básica, essas políticas vinculam a eficácia da escola ao desempenho médio de seus alunos em testes padronizados em algumas disciplinas. Os resultados desses exames, associados ao fluxo escolar, compõem indicadores que recentemente são entendidos de maneira recorrente como medidas da qualidade educacional oferecida por escolas, redes de ensino, municípios e estados.

A mera publicização desses indicadores em relatórios e páginas da Internet, sem qualquer ação mais contundente na orientação de como esses resultados devem ser interpretados, favorece a simples comparação entre as escolas, municípios e estados por meio de listas ordenadas, comumente noticiadas pela imprensa. Como consequência, na prática, a escola é “tomada como microssistema educacional, ela é responsabilizada pela construção do ‘sucesso escolar’, cabendo ao poder público a aferição da produtividade, por meio de aplicação de provas de rendimento aos alunos” (SOUZA, 1997, p. 265).

O grande desígnio da qualidade da educação tem, assim, justificado o interesse crescente na implantação de sistemas de avaliação que permitem encontrar os pontos fortes e fracos do(s) sistema(s) educativos(s) e, conseqüentemente, monitorar as mudanças ao longo do tempo. A monitoração é impossível sem indicadores. Contudo, reduzir os sistemas de avaliação à produção e disseminação de *rankings* elaborados com base em um ou outro indicador, sem cuidar da informação contida nesse indicador e do quão bem ele representa a realidade para a mensuração da qual foi concebido, pode ter efeitos catastróficos na educação. A comparação de desempenho pode ser bastante produtiva para os envolvidos quando é bem feita. Quando mal feita, pode ser custosa, e não apenas inútil, mas prejudicial e até mesmo destrutiva (BIRD *et al.*, 2005). [...] Dependendo do uso e conseqüências de tais sistemas, a omissão dessas limitações pode representar, na atualidade, forte adversidade à consecução do grande desígnio (FERRÃO, 2012, p. 457).

Nesse processo, sem a devida discussão, são consolidados pragmaticamente diversos conceitos, tais como qualidade e eficácia escolar, deixando aberto o caminho para que aconteça o mesmo com o conceito de responsabilização (*accountability*).

Segundo Ximenes (2012, p. 364 apud GOMES, 2008, p. 12), a Lei de Responsabilidade Educacional em trâmite no legislativo brasileiro “emana de um conjunto de direitos já então definidos na Constituição e na Lei”; assim sendo,

Para que o conceito de responsabilidade educacional seja vivido, é preciso o reconhecimento de três aspectos: 1 – A educação é um direito [...], 2- Se esse direito é subtraído de alguém, haverá responsáveis por atos e omissões. 3 – se existem responsáveis, eles devem ter os seus atos e omissões tipificados clara e concretamente e, assim, penalizados (XIMENES, 2012, p. 364 apud GOMES, 2008, p. 11).

Segundo o mesmo autor,

é essa a perspectiva que predomina nas proposições em discussão no Legislativo, que buscam redefinir a responsabilidade educacional para incluir como objeto de responsabilização o fracasso nos testes e como sujeitos responsáveis, em última instância, os trabalhadores da educação. A concepção de responsabilidade deixa de ser predominantemente política (do agente político) e coletiva para ser funcional (do funcionário) e individual, enfocando, em última instância, a responsabilização dos trabalhadores em educação e das escolas (XIMENES, 2012, p. 374).

É nesse contexto que a presente pesquisa se insere com o objetivo de desenvolver um modelo linear hierárquico capaz de identificar os fatores intraescolares e extraescolares que interferem na proficiência final em Leitura e Matemática nos quatro primeiros anos de escolarização⁷, usando os dados longitudinais disponibilizados pelo Projeto GERES.

1.2) O Projeto GERES

O Projeto GERES teve início em 2004 e foi desenvolvido em uma ação conjunta de seis universidades brasileiras⁸, contando com financiamento do Programa de Educação da Fundação Ford e do programa de financiamento a núcleos de excelência acadêmica do Ministério da Ciência e Tecnologia Seu desenho metodológico é um estudo longitudinal de painel, pois envolve “a coleta de dados, ao longo do tempo, de uma mesma amostra de respondentes” (BABBIE, 2005, p. 103), que ocorreu no período de 2005 a 2008, seguindo os mesmos alunos de uma amostra de 312 escolas em cinco grandes cidades brasileiras. Os idealizadores do Projeto GERES consideraram que,

7 No período em que foi realizada a coleta de dados, o Ensino Fundamental brasileiro era organizado em oito anos; assim, os quatro primeiros anos de escolarização correspondem ao período compreendido entre a primeira e a quarta série, hoje denominados 2º e 5º anos.

8 Laboratório de Avaliação da Educação da PUC-Rio, Grupo de Avaliação de Medidas Educacionais (GAME) da Universidade Federal de Minas Gerais (UFMG), Laboratório de Observação e Estudos Descritivos (LOED) da Universidade Estadual de Campinas (UNICAMP), Linha de Pesquisa de Avaliação da Educação do Programa de Pós-graduação da Universidade Federal da Bahia (UFBA), Centro de Avaliação de Educação (CAEd) da Universidade Federal de Juiz de Fora (UFJF) e Universidade Estadual de Mato Grosso do Sul (UEMS).

Trabalhando-se com um conjunto de medidas que se repetem no curso de um tempo determinado, a análise de séries temporais de informações consideradas relevantes permite indicar tendências e fazer projeções a respeito de intervenções possíveis na formulação e implementação de políticas educacionais. A partir de dados longitudinais de rendimento escolar de alunos, pode-se investigar, por exemplo, o progresso acadêmico, a alteração no rendimento escolar de um momento a outro, a influência da escola e da classe no rendimento de seus alunos, a aprendizagem de alunos durante um tempo determinado, ou seja, pode-se estudar a aprendizagem propriamente dita, realçando-se a responsabilidade do professor e da escola no processo educacional. Sem desqualificar a pesquisa transversal, já consolidada como modelagem adequada quando se pretende documentar o rendimento escolar em um momento específico, é importante realçar-se que os dados de tal pesquisa não são suficientes para estudos de eficácia e equidade escolares (OLIVEIRA, FRANCO e SOARES, 2007, p. 155).

O Projeto GERES teve como objetivo central investigar quais práticas educativas e quais condições escolares contribuem para a promoção da eficácia escolar e da equidade intraescolar. Pretendeu identificar:

- (i) as características escolares que maximizam a aprendizagem dos alunos e que minimizam o impacto da origem social sobre a aprendizagem;
- (ii) os fatores escolares que diminuem a probabilidade de repetência dos alunos;
- (iii) os fatores escolares que diminuem o absenteísmo; e
- (iv) os fatores escolares que maximizam a autoestima dos alunos e sua motivação para o estudo.

Para a coleta das médias cognitivas, o Projeto GERES utilizou como instrumento testes de Leitura e de Matemática aplicados em diferentes ondas, focalizando habilidades básicas tipicamente demandadas de alunos das séries iniciais pela escola. As duas primeiras coletas ocorreram em março e novembro de 2005 para tornar possível a medida dos ganhos nas proficiências durante o primeiro ano de escolarização. Para o cálculo do ganho nos demais anos, foram feitas as medidas no mês de novembro de 2006, 2007 e 2008, totalizando cinco ondas de aplicação de testes.

Além das medidas cognitivas, foram coletados dados de contexto do aluno, de sua família, turma, professores, diretor e escola. Essas informações são importantes não apenas para viabilizar o controle estatístico dos resultados, mas também para gerar conhecimento sobre as características que interferem positivamente ou negativamente na proficiência dos

alunos, independentemente das influências econômicas e culturais a que estão submetidos. Para essa coleta, foram usados os seguintes instrumentos:

- Escola, com foco sobre a infraestrutura da escola;
- Diretores, visando às informações sobre sua experiência e formação, sobre as características do financiamento da escola, tais como anuidades, bolsas de estudos, etc. quando a escola pertencia à rede privada, sobre os processos de nomeação do diretor nas escolas públicas, sobre o relacionamento da comunidade e sobre os critérios de formação de turmas;
- Professores, buscando dados acerca de sua previsão quanto ao desempenho da turma nos testes, sua percepção sobre o estilo e liderança do diretor, suas práticas em sala de aula, sua formação e práticas culturais;
- Pais, visando à caracterização sociodemográfica da família; e
- Alunos, também visando à caracterização sociodemográfica da família.

Segundo Brooke e Bonamino,

Os conceitos ou as variáveis que mais interessam aos pesquisadores sociais nem sempre podem ser observados diretamente. Muitas vezes o interesse está focalizado nos que oferecem “riqueza de significados”, tais como nível socioeconômico, preconceitos, sentimentos, satisfação etc. Os conceitos evocam diferentes imagens e noções para diferentes pesquisadores, por isso são classificados como latentes. O pesquisador precisa especificar as manifestações evocadas por eles e, ao fazer isso, estará medindo aquilo que ele convencionou ser seu conceito, ou variável latente (2011, p. 50).

O Quadro 1 apresenta os momentos da pesquisa em que os instrumentos foram aplicados.

Quadro 1 - Aplicação dos instrumentos da pesquisa por onda

Instrumento	1ª Onda	2ª Onda	3ª Onda	4ª Onda	5ª Onda
Testes de Leitura	*	*	*	*	*
Testes de Matemática	*	*	*	*	*
Questionário da Escola*	*				
Questionário do(a) Diretor(a)*	*		*		*
Questionário da Turma	*				
Questionário do(a) Professor(a)		*	*	*	*
Questionário do(a) Aluno(a)				(pré-teste)	*
Questionário de Pais		*		*	*

*Algumas escolas entraram na pesquisa após a 1ª Onda, sendo então aplicado o instrumento.

Fonte: Brooke e Bonamino (2011, p. 57).

Visando a aferir certos constructos teóricos, o questionário dos professores é composto por um conjunto de itens elaborados segundo uma temática que reflete o constructo, resultando nas seguintes escalas, divulgadas oficialmente em 2011:

- Escala 1: Percepção do professor sobre a liderança do diretor;
- Escala 2: Percepção do professor sobre a colaboração entre seus pares;
- Escala 3: Percepção do professor sobre a responsabilidade coletiva para a aprendizagem;
- Escala 4: Percepção do professor sobre os obstáculos que impedem o melhoramento da escola;
- Escala 5: Frequência de uso de certos recursos pedagógicos disponíveis na escola;
- Escala 6: Percepção do professor sobre a frequência de interrupção da aula;
- Escala 7: Frequência com que o professor realiza determinadas práticas nas suas aulas de língua portuguesa;
- Escala 8: Frequência com que o professor realiza determinadas práticas nas suas aulas de matemática;
- Escala 9: Frequência de leitura do professor;
- Escala 10: Frequência de participação em atividades culturais;
- Escala 11: Frequência com que o professor realiza procedimentos pedagógicos promotores da motivação intrínseca;
- Escala 12: Frequência com que o professor realiza procedimentos pedagógicos promotores da motivação extrínseca com regulação introjetada;
- Escala 13: Frequência com que o professor realiza procedimentos pedagógicos promotores da motivação extrínseca com regulação externa.

Ainda segundo Brooke e Bonamino (2011, p. 50),

A Teoria da Resposta ao Item não Paramétrica foi empregada de modo exploratório e confirmatório para extrair as escalas com boas propriedades de escalonidade. Os coeficientes da teoria clássica dos testes e a modelagem por equações estruturais por meio da análise fatorial confirmatória foram empregados para validar as escalas por intermédio da verificação do ajuste. E finalmente, a Teoria da Resposta ao Item Paramétrica foi utilizada para o cálculo dos parâmetros do modelo e dos escores das escalas.

A Teoria de Resposta ao Item Paramétrica (TRI) também foi utilizada para tornar possível a equalização das escalas de proficiência nos diferentes anos de escolarização, independentemente de as escolas terem organização curricular seriada ou por ciclos de formação. Após a coleta das medidas cognitivas, que ocorre, portanto, em cinco ondas de aplicação, as proficiências foram estimadas novamente e equalizadas, permitindo a determinação da curva de crescimento da proficiência de cada aluno, o que resultou em um banco de dados composto com informações de 35.538 alunos, porém nem todos com as medidas coletadas em todas as ondas de aplicação.

É importante destacar que há inadequado reducionismo existente no parágrafo anterior ao considerar o desenvolvimento cognitivo como a diferença entre duas proficiências em uma ou outra determinada área de conhecimento, em geral, a leitura na língua materna e a matemática. Obviamente, o desenvolvimento cognitivo é mais amplo do que as proficiências estimadas em determinadas áreas de conhecimento por meio de modelos estatísticos que envolvem a Teoria de Resposta ao Item – TRI - (FLETCHER, 1994; ANDRADE, TAVARES e VALLE, 2000). Não analisaremos aqui a restrição curricular implícita no processo, uma consequência da valorização das avaliações de larga escala que consiste na tendência de transformar a matriz avaliativas dessas avaliações no currículo escolar, uma vez que vamos considerar que as proficiências dos alunos são devidamente estimadas pelo processo estatístico, já que:

A Teoria da Resposta ao Item é um instrumento poderoso, que vem sendo progressivamente aplicado nos processos quantitativos na área de avaliação educacional. A TRI propõe modelos de variáveis latentes para representar a relação entre a probabilidade de um indivíduo apresentar determinada resposta a um item e seus traços latentes ou habilidades na área do conhecimento avaliada, os quais não podem ser observados diretamente (ALEXANDRE, 2002, p. 131).

Tomando por correta a relação entre a proficiência e o conhecimento do aluno, necessário salientar, mais uma vez, que parte desse desenvolvimento cognitivo, estimado pelo valor agregado no período, não é resultante exclusivamente das interferências da escola, pois é preciso considerar também que o aluno aprende fora dela nesse mesmo período.

O Quadro 2 apresenta os níveis de habilidade e proficiência características em Leitura e em Matemática, em função da pontuação que os alunos obtêm em cada uma das escalas.

Quadro 2 - Níveis de habilidade e proficiência características

Proficiência em	Nível	Aquém das habilidades do nível	Auge do desenvolvimento das habilidades do nível	Em via de consolidação das habilidades do nível	Habilidades do nível consolidadas
Leitura	1	<53	53 a 109	109 a 129	>= 129
	2	<109	109 a 127	123 a 143	>= 143
	3	<133	133 a 153	153 a 166	>=166
	4	<155	155 a 172	172 a 182	>=182
	5	<171	171 a 188	188 a 201	>=201
	6	<190	190 a 211	211 a 222	>=222
Matemática	1	<36	36 a 80	80 a 105	>105
	2	<85	85 a 131	131 a 159	>159
	3	<131	131 a 176	176 a 201	>201
	4	<185	185 a 205	205 a 339	>339
	5	<225	225 a 283	283 a 311	>311
	6	<288	288 a 362	362 a 387	>387

Fonte: Brooke e Bonamino (2011, p. 139).

As habilidades correspondentes a cada um dos níveis, em Leitura e em Matemática, estão descritas nos Anexos 1 e 2.

Para o Projeto GERES, o nível socioeconômico é um constructo teórico que não pode ser observado diretamente; por isso, assim como ocorreu nas informações sobre os professores, foi feita uma medida resultante da agregação de informações sobre a escolaridade, a renda e a ocupação dos pais.

Segundo Alves e Soares (2009, p.2),

O nível socioeconômico aparece em inúmeros estudos como variável explicativa ou de controle para a análise de diversos fenômenos sociais. No entanto, não há um consenso na literatura sobre sua conceituação, bem como sobre como medi-lo nas pesquisas empíricas. Há vários aspectos relacionados ao NSE que vêm merecendo debate entre os cientistas sociais. Por exemplo, a sua base de conceituação teórica, o tipo de medida - se contínua ou categórica -, os fatores a serem considerados na produção da medida - se variáveis isoladas ou em forma de um índice -, a definição de um esquema de classificação das variáveis consideradas e a importância relativa do pai e da mãe para a definição do NSE das famílias.

Para a elaboração dessa escala, mais uma vez, a Teoria de Resposta ao Item foi usada para agregar as informações obtidas pelos itens desses três constructos (escolaridade, a renda e a ocupação dos pais).

Segundo Buchmann (2002), o efeito da mãe na escolaridade dos filhos é superior à do pai. Por essa razão, para o constructo escolaridade, foi considerada apenas a escolaridade da mãe, descartando-se a do pai. Já para o constructo ocupação, foram consideradas as

informações sobre a principal ocupação e as atividades realizadas tanto pela mãe quanto pelo pai. O constructo renda foi composto por itens que verificam os itens de conforto existentes na casa do aluno.

**CAPÍTULO 2 –
DADOS LONGITUDINAIS E HLM**

O capítulo anterior trouxe o contexto no qual estão inseridos os estudos sobre eficácia e equidade escolar, assim como o controverso debate do uso meritocrático dado à avaliação educacional de larga escola a eles associados, justificando o objetivo desta pesquisa que é desenvolver um modelo linear hierárquico capaz de identificar os fatores intraescolares e extraescolares que interferem na proficiência final em Leitura e Matemática nos quatro primeiros anos de escolarização, usando os dados longitudinais disponibilizados pelo Projeto GERES e considerando a complexidade de tais dados, não só quanto à sua coleta e tratamento, mas também na inter-relação existente entre eles na associação com a proficiência dos alunos.

Este capítulo tem dupla finalidade. A primeira é justificar a decisão metodológica de se trabalhar com dados longitudinais e analisá-los por meio de modelos lineares hierárquicos e a segunda é apresentar os conceitos necessários para buscar maior entendimento dessa metodologia, uma vez que os recursos estatísticos inerentes são, por si só, significativamente complexos. Esta segunda atende a proposta feita por Braun (2005, p. 5), diante existente acerca dos modelos de valor agregado:

Há um contraste marcante entre o entusiasmo daqueles que aceitam as reivindicações feitas sobre os modelos de valor agregado e gostariam de usá-los e, por outro lado, as reservas expressas por aqueles que estudaram seus méritos e limites técnicos. Esta disjunção é motivo de preocupação. Porque os modelos de valor agregado contam com procedimentos estatísticos complexos, fazendo ser provável que os responsáveis políticos, funcionários da educação, professores e outros interessados poderiam se beneficiar de orientações básicas mais compreensíveis para as dúvidas acerca do uso dos modelos de valor agregado para avaliação de professores. (Embora exista também um considerável interesse para responsabilização das escolas).

Essa preocupação é corroborada com a denúncia feita por Strauss (2011)⁹,

Sejam elas ingênuas ou especialistas, os matemáticos precisam confrontar as pessoas que usam inadequadamente a matemática visando a intimidar os outros para que aceitem suas conclusões simplesmente porque são baseadas em certo conhecimento matemático. Ao contrário de muitos formuladores de políticas, os matemáticos não se enganam com a teoria subjacente aos modelos de valor agregado e devem manifestar-se claramente a respeito. Modelos matemáticos têm limitações. Por si só, não conferem autoridade às suas conclusões. Eles são ferramentas, não mágica. Usar a matemática para intimidar – para antecipar o debate sobre os objetivos da educação e medidas de sucesso – é prejudicial não só à educação, mas à própria matemática.

9 Provavelmente a autora se refere aos conhecimentos estatísticos quando se cita a matemática em sua escrita.

Com o objetivo de superar a recorrente dificuldade no entendimento da linguagem técnica, quando possível, a escrita deste capítulo explora uma linguagem mais textual, em detrimento dessa linguagem técnica, para que os leitores não familiarizados com os significados e símbolos da estatística e da matemática possam construir as ideias necessárias para o entendimento das possibilidades e limitações dos modelos de valor agregado, que usam os modelos lineares hierárquicos, um dos recursos estatísticos centrais desta pesquisa.

No desenvolvimento do texto, os conceitos serão apresentados em grau de complexidade crescente. As definições associadas aos conceitos, as nomenclaturas, a simbologia e as referências bibliográficas são inseridas para um melhor entendimento das informações contidas nos próximos capítulos.

No decorrer do texto, os exemplos cotidianos não pertencentes ao contexto educacional usados inicialmente são substituídos por outros com variáveis vinculadas à pesquisa. Próximo ao final do capítulo, ainda visando os leitores não acostumados com a área da matemática e da estatística, é apresentado um modelo com valores hipotéticos com exemplos que auxiliarão na interpretação dos capítulos seguintes.

Assim, este capítulo está organizado em três seções que abordam os conceitos e noções que caracterizam e justificam o uso de dados longitudinais, dos modelos lineares hierárquicos e do modelo escolhido nesta pesquisa.

2.1) Uso de dados longitudinais

A pesquisa feita pelo Projeto GERES, assim como a grande maioria das avaliações externas aplicadas nas redes escolares, buscam medir o conhecimento do aluno associando-o a uma escala de habilidades definidas segundo uma proficiência. Nessa concepção essa escala faz com que a proficiência, um dos traços latentes do conhecimento do aluno, é medida através da aplicação de instrumentos compostos por itens elaborados com base em uma matriz de habilidades, de maneira que quanto maior for o conhecimento do aluno, maior é a probabilidade de que ele acerte tais itens. Assim, a proficiência de um aluno é estimativa de seu conhecimento¹⁰.

Se for considerado que esse conhecimento não é construído apenas na escola, mas também fora dela e até mesmo antes do aluno ter ingressado nela, para identificar os fatores

10 Para uma aprofundamento sobre a Teoria de Resposta ao Item, consultar Andrade, Tavares e Valle (2000).

que interferem na proficiência dos alunos é essencial distinguir quais deles são associados às características da escola e quais estão associados ao perfil do aluno. A relevância dessa distinção deve-se ao fato de que “o desempenho anterior ao ingresso do aluno na escola sempre influencia o nível de desempenho final” (GRAY, 2008, p. 253).

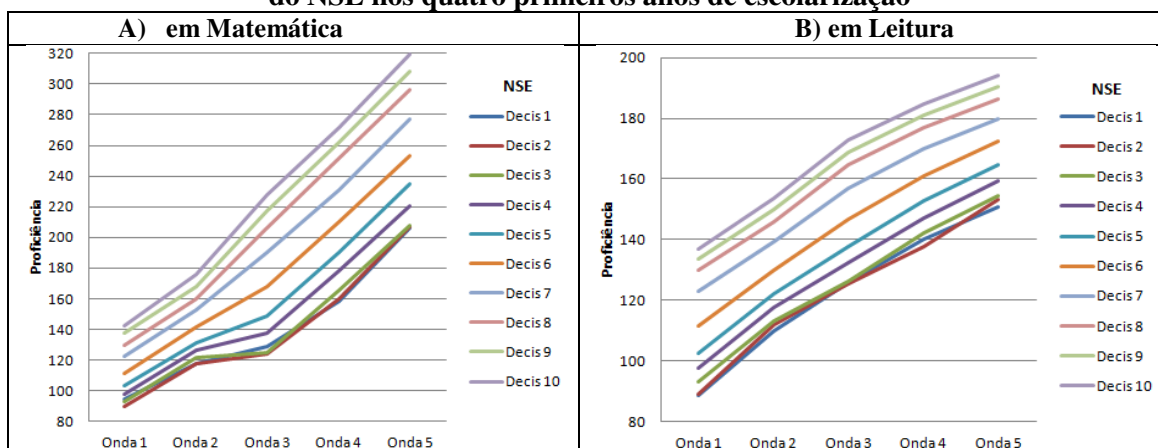
Também para Willms (2008, p. 270),

Ao tentarmos estimar os efeitos da escola, é crucial separar os efeitos sobre os resultados escolares associados às características iniciais dos alunos, que eles têm quando são matriculados na escola, dos efeitos associados às práticas internas das escolas que eles frequentam. Os alunos não são distribuídos aleatoriamente entre as escolas; eles são escolhidos por mecanismos formais e informais de seleção. As aparentes diferenças entre escolas quanto a seus resultados são explicadas em parte, por diferenças nas características que os alunos têm antes de entrarem na escola

Esse cuidado, como já apresentado no capítulo anterior, ganha destaque após a publicação do Relatório Coleman e de inúmeros estudos que o sucederam, evidencia a inadequação de qualquer análise sobre os efeitos da escola sobre o conhecimento do aluno de forma isolada de seu contexto social e bem como de seus alunos.

Os Gráficos 1A e 1B evidenciam a correlação existente entre as proficiências final e inicial e também a correlação de ambas com o nível socioeconômico e foram elaborados com dados longitudinais fornecidos pelo Projeto Geres, mostra a evolução da média da proficiência, distintamente em português e matemática, dos alunos que permaneceram na mesma escola durante os quatro anos de escolarização e que, em função do seu nível socioeconômico, foram divididos em dez estratos.

Gráfico 1 - Proficiência média de alunos que permaneceram na mesma escola em função do NSE nos quatro primeiros anos de escolarização



A Tabela 1 apresenta os limites e a frequência de cada um dos estratos, em função do nível socioeconômico. Para a elaboração desses estratos foi identificado o intervalo entre o menor e o maior NSE da amostra, em seguida, esse intervalo foi dividido em 10 partes iguais, definindo assim, cada um dos estratos, delimitados pelos seus limites inferior e superior.

Tabela 1 Número de alunos em cada estrato por NSE

Estrato	Limite do Estrato		Freq
	Inferior	Superior	
1	-1,50	-1,18	35
2	-1,18	-0,86	255
3	-0,86	-0,54	985
4	-0,54	-0,22	2013
5	-0,22	0,10	2665
6	0,10	0,42	1719
7	0,42	0,74	1322
8	0,74	1,06	439
9	1,06	1,38	379
10	1,38	1,70	456
		Total	10268

Com uma breve análise do Gráfico 1, uma primeira constatação é que quanto maior é o nível socioeconômico do agrupamento do aluno, maior é a proficiência média desse grupo, evidenciando a correlação existente entre a proficiência inicial e o nível socioeconômico.

A segunda constatação é que, com exceção dos três primeiros *estratos*, os quais as curvas de crescimento da proficiência média de cada um dos agrupamentos se cruzam, em ambas as disciplinas, que entram na escola (proficiências medidas na Onda 1, aplicada no início do primeiro ano de escolarização) em uma determinada posição, termina o período escolar na mesma posição, evidenciando a alta correlação entre a proficiência final e a inicial. Ou seja, a diferença apresentada na escala de proficiência entre os níveis socioeconômicos se mantém no tempo.

Os Gráficos 1A e 1B também permitem visualizar que nesse período de escolarização, em matemática, o valor agregado à proficiência inicial dos alunos com níveis socioeconômicos mais altos é maior do que aqueles com níveis socioeconômicos mais baixos, diferentemente do que ocorre em Leitura.

Segundo Gray (2008), medidas de valor agregado tornam possível comparar semelhantes de maneira que “se há dois alunos que sob todos os aspectos são idênticos, que diferença faz para seu crescimento se eles frequentarem escolas diferentes?” (p. 253). Com

esse questionamento de Gray (2008) pode-se inferir, em uma primeira análise, que se alunos com perfis idênticos tiverem valores agregados diferentes em suas proficiências, as escolas podem ter efeitos distintos sobre esses alunos.

Assim, é evidenciada uma das justificativas para o uso de dados longitudinais, uma vez que o cálculo de valor agregado na proficiência, que permite a identificação mais confiável dos fatores associados à escola, só é possível em estudos que disponibilizam medidas longitudinais em painel, isto é, medidas das proficiências dos mesmos alunos em momentos diferentes.

2.2) Uso de Modelos Lineares Hierárquicos

Em sua essência, essa pesquisa pretende identificar fatores que interferem na proficiência dos alunos ao final de um período de escolarização, com isso, é estabelecida a premissa de que a variação desses fatores interfere na variação da proficiência final.

Os modelos lineares hierárquicos, cuja denominação será explicada mais adiante, juntamente com os seus conceitos, são ajustados com um recurso derivado da análise de regressão linear e por isso, para o entendimento da primeira se faz necessário, antes, o entendimento da segunda.

2.2.1) Análise de regressão linear

A análise de regressão linear estabelece a relação existente entre duas variáveis e permite descobrir o efeito que uma variável (que possui várias denominações, podendo ser: variável controle, variável explicativa, variável independente ou preditora) exerce sobre outra, a variável resposta (ou variável dependente). Ao conhecer numericamente esse efeito, significa conhecer, com relativa precisão, o quanto a variável resposta se altera com relação a determinada variação da variável de controle. Assim a análise de regressão é um recurso estatístico que permite, a partir de uma base de dados compostas minimamente por duas

variáveis que tem alguma relação, estimar o efeito que a variação dessas variáveis tem sobre a outra.

Um exemplo cotidiano dessa relação entre diferentes variáveis são as tabelas existentes sobre os balcões das empresas que fazem fotocópias. Nessas tabelas é estabelecida a relação existente entre o número de cópias tiradas e o valor a ser pago pelo cliente, cujos valores são obtidos por meio de simples operações matemáticas. Dessa relação pode-se perceber que os valores da primeira variável explicam os valores da segunda, ou ainda, que os valores da segunda dependem dos valores da primeira. Por esse motivo, nos estudos quantitativos, a primeira variável é chamada de variáveis independentes e a segunda variável dependente¹¹.

Na mesma linha de pensamento, porém para maior aprofundamento, outro exemplo seria a relação existente entre o valor pago pelo abastecimento de combustível em um veículo, no qual uma das variáveis é a quantidade de combustível colocada no tanque do veículo (variável explicativa), aqui designada como variável x , e a outra o valor pago pelo cliente (variável resposta), designada variável y .

A Tabela 2 traz alguns valores para ambas as variáveis, considerando o valor hipotético de R\$ 2,00 a ser pago para cada unidade de combustível colocada no tanque.

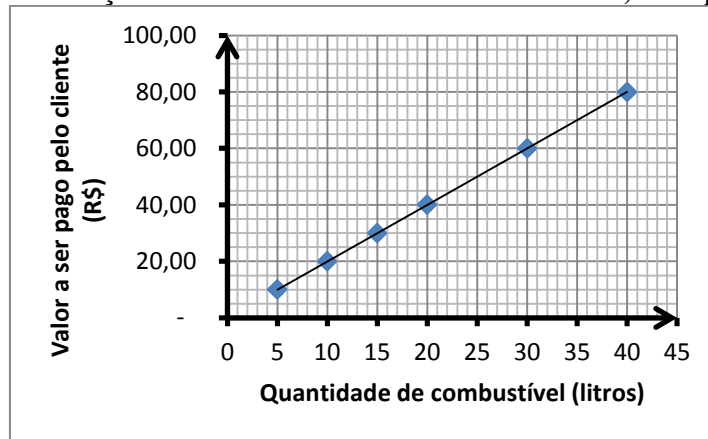
Tabela 2- Exemplo de associação matemática entre diferentes variáveis

Quantidade de combustível (litros)	Valor a ser pago (R\$)
5	10,00
10	20,00
15	30,00
20	40,00
30	60,00
40	80,00

O Gráfico 2 permite a visualização do alinhamento desses valores, que se encontram exatamente sobre uma reta que seria traçada se todos os inúmeros pontos entre 5 e 40 litros fossem representados, assim como outros fora desse limite também o fossem.

11 Nesse trabalho usaremos a denominação de variável resposta e variável explicativa.

Gráfico 2 - Associação matemática entre diferentes variáveis; exemplo hipotético



Da representação gráfica dessa reta que representa esse conjunto de possíveis valores é necessário destacar dois aspectos. O primeiro deles é o ponto que marca início dessa reta, que nesse exemplo é o valor de R\$ 0,00 quando zero litro de combustível é colocado no tanque. O segundo refere-se à inclinação dessa reta, que seria mais inclinada se o valor de cada litro de combustível fosse maior, ou menos inclinada se esse valor do combustível fosse menor. Qualquer reta é caracterizada por esses dois valores, chamados coeficientes da reta.

É importante destacar que nesses dois exemplos, nenhum ponto se desvia da reta porque uma das variáveis (nesses exemplos, o valor a ser pago pelo cliente) depende exclusivamente da outra (quantidade de cópias ou quantidade de combustível).

Outro destaque necessário, talvez mais importante para a finalidade desse texto, é que os valores expressos na tabela e em seguida representados no gráfico, foram obtidos a partir de um equacionamento elaborado com base em contexto hipoteticamente existente, ou seja, os pontos são encontrados a partir de uma simples equação matemática, que com maior rigor na linguagem, já pode ser considerado um Modelo Determinístico¹².

A análise de regressão linear faz um caminho inverso. É um recurso estatístico que permite fazer uma aproximação a uma reta, a partir de um conjunto de pontos dados. Encontrar uma reta significa encontrar os valores para os coeficientes da reta, isto é, o valor da variável resposta para quando o valor da variável explicativa for nula e a inclinação dessa reta.

12 A mais simples das representações de um sistema real, feita minimamente por uma equação, pode ser denominada como um modelo matemático, ou simplesmente modelo. O modelo é determinístico quando dado um conjunto de valores de entrada, o conjunto de valores que dele resulta são precisamente calculados.

O primeiro coeficiente é chamado coeficiente linear da reta (ou intercepto) e o segundo, chamado de coeficiente angular da reta (coeficiente de regressão).

Para exemplificar a aplicação da regressão linear, cabe a situação hipotética na qual se deseja verificar a relação existente entre a renda mensal de um funcionário e a quantidade de anos de escolarização, cujos valores, apresentados na Tabela 3 foram medidos empiricamente.

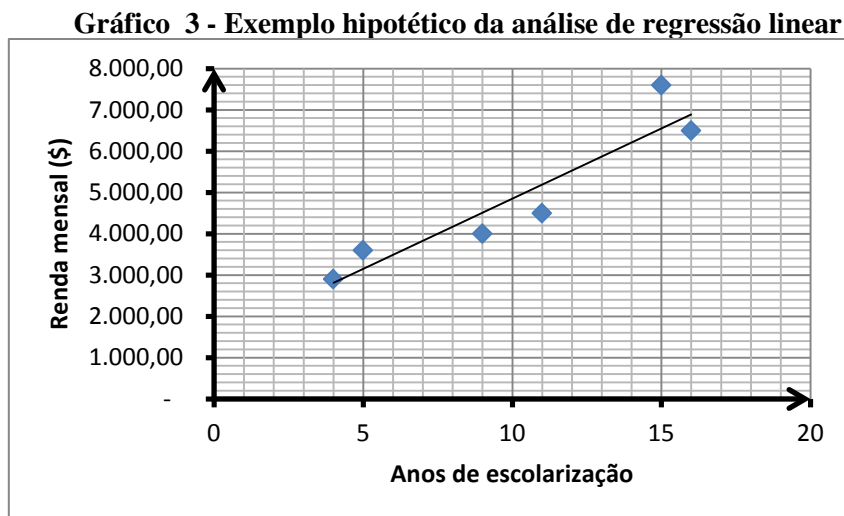
Tabela 3 - Análise de regressão linear; exemplo hipotético

Respondente	Escolarização em anos (X_i)	Renda Mensal em \$(Y_i)
A	11	4.500,00
B	5	3.600,00
C	16	6.500,00
D	4	2.900,00
E	15	7.600,00
F	9	4.000,00

Nessa tabela o conjunto de valores observados na variável explicativa é composto por $X_1=11$, $X_2=5$; $X_3=16$, $X_4=4$, $X_5=15$ e $X_6=9$, enquanto que para a variável resposta foram observados os valores $Y_1=4.500,00$, $Y_2=3.600,00$; $Y_3=6.500,00$, $Y_4=2.900,00$, $Y_5=7.600,00$ e $Y_6=4.000,00$.

Na linguagem estatística, as variáveis resposta e a explicativa são representadas, respectivamente, por X_i e Y_i , com $i= 1, 2, 3, \dots, 6$, de forma que $Y_i=f(X_i)$, isto é, Y_i depende de X_i , ou ainda, Y_i é uma função de X_i .

No Gráfico 3 estão representados cada um dos pontos (X_i, Y_i).



Nota-se que, diferentemente dos exemplos anteriores, os pontos não ficaram perfeitamente alinhados. Em situações reais, esse não alinhamento é esperado. Isso porque uma variável resposta é geralmente influenciada por mais que uma variável explicativa. No exemplo, é esperada a existência de outras variáveis influenciando na renda mensal, tais como idade, anos de experiência na função etc. Além disso, é também esperado erros de mensuração da variável resposta e, por vezes, também das variáveis explicativas.

Mas mesmo com esse não alinhamento, há uma reta que melhor se ajusta para representar esse conjunto de pontos (X_i, Y_i) que também pode ser observada no Gráfico 3. Essa reta (e, portanto, uma equação) é um modelo não determinístico¹³.

Com isso, esse conjunto de pontos pode ser representado pelo modelo estatístico $Y_i = \beta_0 + \beta_1 \times X_i + r_i$, em que β_0 é o intercepto (coeficiente linear da reta), β_1 é o coeficiente de regressão (coeficiente angular da reta) e r_i é o desvio existente entre o valor observado e aquele estimado pelo modelo, que se deve aos erros de mensuração e à associação de outras variáveis explicativas não inseridas no modelo.

Por esses dois motivos, o desvio é frequentemente denominado como resíduo ou, simplesmente erro. Por um lado, o termo desvio se adequa bem para dar a noção da distância entre o valor real e o estimado em cada observação, no entanto, o termo resíduo se adequa para dar a noção de que essa distância se deve a alguma variável não inserida no modelo, como explicitado no parágrafo anterior. De qualquer forma, esse distanciamento representa um erro na regressão (termo mais recorrente na literatura), que ao mesmo tempo se busca a sua diminuição, é o que permite encontrar as variáveis explicativas que permite diminuí-los, e consequentemente viabilizando os procedimentos estatísticos usados na regressão linear. Como pode ser percebido, há sutis diferenças conceitual entre erro e resíduo, mas para a finalidade de trabalhos dessa natureza, o foco deve estar no resíduo.

A variância da variável resposta, em uma abordagem inicial e, portanto superficial, é uma medida estatística que representa numericamente o quanto os valores observados estão dispersos em torno dessa reta estimada e normalmente é representada por σ^2 . Quanto mais os dados forem dispersos, maior será o valor estimado para σ^2 . Quando a variância é nula,

13 Modelos não determinísticos são aqueles, a partir de um conjunto de valores, levam a valores aproximados àqueles observados na realidade. Isso ocorre nos modelos que não controlam todas as variáveis dessa realidade ou existe algum efeito aleatório nesses valores (de entrada ou de resultado).

significa que todos os dados estão representados pela reta e não há desvios entre os valores estimados e os observados.

A regressão linear a partir de uma amostra faz as estimativas para os coeficientes β_1 e β_0 , geralmente desconhecidos de maneira que essa reta estimada é dada por $\hat{Y}_i = a + b \times X_i$, em que a é a estimativa para o coeficiente β_0 e b é a estimativa para o coeficiente β_1 .

Mesmo que com menor precisão, é possível no Gráfico 3 determinar visualmente que para 10 anos de escolarização a renda mensal é de aproximadamente \$ 4.800,00. Da mesma maneira, determinar aproximadamente que o valor da renda mensal para 15 anos (X_5) de escolarização é de \$ 6.600,00.

Nesse exemplo, as estimativas para os coeficientes β_0 e β_1 são respectivamente, iguais a 1454,84 e 339,52. Isso significa que renda mensal esperada para uma pessoa com zero ano de escolarização é de \$ 1.454,84 e que para cada ano de escolarização há uma estimativa de acréscimo no valor médio de \$ 339,52 na renda mensal.

Com essas estimativas a equação da reta que melhor se ajusta a esses dados é dada por $\hat{Y}_i = 1454,84 + 339,52 \times X_i$. Assim, usando a equação, o valor estimado da renda mensal para 15 anos de escolarização é de \$ 6.547,64 (\hat{Y}_5).

No entanto, como consta na Tabela 3, o valor hipoteticamente observado para 15 anos de escolarização é de \$ 7.600,00 (Y_5). Assim, o resíduo para esse ponto é de \$1.052,36 ($Y_5 - \hat{Y}_5$).

A Tabela 4 apresenta os valores observados, estimados e o resíduo observado em cada um dos pontos.

Tabela 4 - Resíduos em uma regressão linear simples; exemplo hipotético

Respondente	Escolarização em anos (X_i)	Renda Mensal em \$		Resíduo \$ (u_i)
		Observada (Y_i)	Estimada (\hat{Y}_i)	
A	11	4.500,00	5.189,56	- 689,56
B	5	3.600,00	3.152,44	447,56
C	16	6.500,00	6.887,16	- 387,16
D	4	2.900,00	2.812,92	87,08
E	15	7.600,00	6.547,64	1.052,36
F	9	4.000,00	4.510,52	- 510,52

Na tentativa de minimizar esses resíduos, isto é, minimizar a diferença entre os valores observados e os valores estimados, outras variáveis poderiam ser inseridas, desde que obviamente tenham sido medidas, resultando em outro modelo, dito mais ajustado .

Nesse caso, a regressão linear deixaria de ser bivariada ou simples (usada nesse exemplo por relacionar apenas duas variáveis) e passaria a ser usada a regressão linear múltipla, isto é, a variável resposta passaria a ser calculada por mais que uma variável explicativa. Nesse caso, o modelo estatístico com k variáveis explicativas ficaria assim definido:

$$Y_i = \beta_0 + \beta_1.X_{1i} + \beta_2.X_{2i} + \beta_3.X_{3i} + \dots + \beta_k.X_{ki} + r_i ,$$

Em que β_0 é o intercepto¹⁴ (coeficiente linear da reta), β_1 é o coeficiente de regressão referente à primeira variável explicativa, β_2 é o coeficiente de regressão referente à segunda variável explicativa, β_3 é o coeficiente de regressão referente à terceira variável explicativa, β_k é o coeficiente de regressão referente à *k-ésima* variável explicativa e r_i é o resíduo devido aos erros de mensuração e à associação de outras variáveis explicativas.

A Tabela 5 apresenta os dados hipotéticos dos seis respondentes do exemplo anterior, porém com outras duas variáveis explicativas.

Tabela 5 - Análise de regressão linear múltipla; exemplo hipotético

Respondente	Escolarização em anos (X_{1i})	Idade em anos (X_{2i})	Experiência em anos (X_{3i})	Renda Mensal em \$ (Y_i)
A	11	27	4	4.500,00
B	5	25	4	3.600,00
C	16	38	10	6.500,00
D	4	19	2	2.900,00
E	15	41	17	7.600,00
F	9	30	2	4.000,00

Nesse exemplo, como estão envolvidas três variáveis explicativas, a regressão linear múltipla busca estimar valores para os coeficientes β_0 , β_1 , β_2 e β_3 e usa o seguinte o modelo estatístico:

$$Y_i = \beta_0 + \beta_1.X_{1i} + \beta_2.X_{2i} + \beta_3.X_{3i} + r_i ,$$

Em que:

¹⁴ Deve ser ressaltado que $Y_i = \beta_0$, quando $X_{1i} = X_{2i} = X_{3i} = \dots = X_{ki} = 0$.

- Y_i representa a variável resposta e são os valores observados de renda final¹⁵;
- β_0 é o intercepto e o seu valor indica a renda mensal estimada para uma pessoa com valor nulo em todas as variáveis explicativas, isto é, não tem qualquer escolarização e experiência e, ainda, idade nula. Com isso, o valor estimado para esse coeficiente não teria qualquer significado real, portanto, sua importância se resume a apenas ajustar a reta que melhor representa esses dados;
- X_{1i} , X_{2i} e X_{3i} representam as variáveis explicativas envolvidas no modelo: escolarização, idade e experiência, respectivamente;
- β_1 é o coeficiente de regressão associado à variável X_{1i} (escolarização) e indica o valor estimado que deverá ser acrescentado à renda mensal para cada ano de escolarização;
- β_2 é o coeficiente de regressão associado à variável X_{2i} (idade) e indica o valor estimado que deverá ser acrescentado à renda mensal para cada ano de idade;
- β_3 é o coeficiente de regressão associado à variável X_{3i} (experiência) e indica o valor estimado que deverá ser acrescentado à renda mensal para cada ano de experiência;
- r_i é o resíduo.

No ajuste desse modelo aos dados da Tabela 5 foram estimados os seguintes valores para cada um dos coeficientes: $\beta_0= 1358,41$, $\beta_1= 136,33$, $\beta_2= 36,67$ e $\beta_3=158,19$. Com essas estimativas, a equação preditiva fica assim definida:

$$\hat{Y}_i = 1358,41 + 136,33 \cdot X_{1i} + 36,67 \cdot X_{2i} + 158,19 \cdot X_{3i}$$

Neste caso, a representação gráfica fica inviabilizada por existirem mais que três dimensões a serem representadas, restando apenas a interpretação das estimativas dos coeficientes α , β_1 , β_2 e β_3 dessa equação. Assim, com esses valores é possível afirmar que:

- a) a renda mensal para uma pessoa sem qualquer escolarização e experiência e, ainda, com idade nula seria de \$ 1.358,41. Como dito anteriormente, uma estimativa sem significado prático e que serve apenas para ajustar o modelo;

¹⁵ Essas estimativas são válidas apenas para os intervalos observados.

- b) Para cada ano de escolarização, a renda mensal estimada é acrescida de \$136,33;
- c) Para cada ano de idade, a renda mensal estimada é acrescida de \$ 36,67;
- d) Para cada ano de experiência, a renda mensal estimada é acrescida de \$158,19.

A Tabela 6 apresenta os resíduos encontrados com a inclusão das variáveis explicativas idade e experiência.

Tabela 6 -Resíduos com análise de regressão linear múltipla; exemplo hipotético

Respondente	Escolarização em anos (X_{1i})	Idade em anos (X_{2i})	Experiência em anos (X_{3i})	Renda Mensal em \$		Resíduos \$ (u_i)
				Observada (Y_i)	Estimada (\hat{Y}_i)	
A	11	27	4	4.500,00	4.480,89	19,11
B	5	25	4	3.600,00	3.589,57	0,43
C	16	38	10	6.500,00	6.515,05	15,05
D	4	19	2	2.900,00	2.916,84	16,84
E	15	41	17	7.600,00	7.596,06	3,94
F	9	30	2	4.000,00	4.001,86	1,86

Com a comparação entre esses resíduos e aqueles apresentados na Tabela 4 fica evidente a diminuição dos valores. Desde que estejam disponíveis, as variáveis explicativas devem ser inseridas nos modelos desde que consigam diminuir esses resíduos.

2.2.2) Pressupostos para o uso da análise de regressão linear múltipla

A regressão linear, seja ela simples ou múltipla, tem por objetivo encontrar as estimativas para os coeficientes (coeficiente de regressão de cada uma das variáveis explicativas e o intercepto) a partir de uma amostra representativa da população de maneira que os resíduos sejam os menores possíveis. Assim, um modelo é considerado ajustado quando as estimativas dos coeficientes resultem em resíduos não mais ser minimizados, isto é, as variáveis não mais conseguem explicar a variabilidade da variável resposta.

Na área educacional, sobretudo nos limites da sala de aula e da escola, a estatística mais utilizada é a média, que curiosamente, pode ser vista como sendo uma regressão linear cujo modelo estatístico não tenha qualquer variável explicativa. Nessa perspectiva, a média das notas obtidas pelos alunos de uma turma é a estimativa que melhor se ajusta àquela turma quando não se considera nenhuma das características individuais dos alunos além de sua nota. Nesse contexto, os resíduos que ficam, isto é, a distância entre as notas de cada um dos alunos e a média da turma é o que caracteriza a individualidade de cada um desses alunos em relação à turma que se encontra matriculado.

Dai a importância dos resíduos nos estudos com a regressão linear. Uma vez que, quando são encontradas variáveis explicativas tais como, por exemplo, se o aluno é retido ou não, horas de estudo, nível socioeconômico, autoestima, e estas são inseridas no modelo de regressão, é esperado que uma nova reta (um modelo estatístico) passe a representar melhor esse conjunto de notas do que a média, de maneira que os desvios entre essas notas e os valores estimados pelo modelo sejam minimizados, isto é, os resíduos sejam os menores possíveis.

Mas para que essas estimativas sejam confiáveis, algumas condições precisam ser atendidas. Essas condições, denominadas pressupostos do modelo, além de especificar que deva existir uma relação linear entre as variáveis explicativas e a variável resposta, também exigem que as observações das variáveis explicativas sejam fixas, isto é, não sejam valores aleatórios.

Para que esses pressupostos sejam atendidos é necessário que:

- a) os resíduos encontrados tenham média igual a zero. Como se pode observar no Gráfico 3, alguns pontos observados ficam acima da reta estimada, enquanto que outros ficam abaixo. Assim algumas das diferenças entre os valores observados e os estimados ($Y_i - \hat{Y}_i$), resultam em valores positivos e outros negativos, que, quando somados resultam em um valor nulo.
- b) os resíduos sejam homocedásticos, isto é, a variância dos resíduos em diferentes observações é constante;
- c) os resíduos devem apresentar distribuição normal, ou seja, uma distribuição que tenha aderência à curva de Gauss;
- d) os elementos que compõem o resíduo de diferentes observações sejam independentes entre si, isto é, as observações sejam selecionadas aleatoriamente, não existindo, portanto, qualquer viés de agrupamento.

Se esses pressupostos de linearidade, normalidade, homocedasticidade e independência são atendidos, significa que o resíduo é uma variável aleatória com média igual a zero e com variância constante. Na estatística, essa afirmativa é representada por $r_i \sim NID(0, \sigma^2)$. Como esses pressupostos são necessários para o uso da regressão linear, essa notação frequentemente aparece nos estudos que a utilizam para indicar que as pressuposições exigidas pelo modelo

foram atendidas. O símbolo σ^2 representa o valor que indica a variabilidade dos erros em torno da reta estimada e é chamado de variância residual.

O pressuposto de que os resíduos tenham distribuição normal (especificado no item c) é importante para que as estimativas encontradas a partir da amostra usada na análise de regressão possam ser generalizadas à população da qual a amostra foi extraída.

Já o pressuposto da independência dos elementos que compõem os desvios (especificado no item d), geralmente não é atendido em estudos cujos dados têm características de agrupamento ou hierarquização, como ocorre na área educacional. Em geral, o sistema educacional está organizado de forma que alunos são agrupados em turmas, turmas são agrupadas em escolas, escolas são agrupadas em redes etc., caracterizando, desta maneira, os agrupamentos e os níveis de hierarquização dos dados nos estudos com a análise de regressão.

Caso o objetivo do estudo seja estabelecer a relação entre a variável resposta e as explicativas no âmbito do aluno, este estudo com o modelo de regressão linear clássico exigiria “replicar o dado de cada escola tantas vezes quanto os alunos que estudam em cada uma delas. Isso conduz às estimativas de erro-padrão incorretas, com implicações nas inferências e conclusões daí decorrentes” (FERRÃO, 2003, pg. 30).

Em outra hipótese, caso o objetivo do estudo seja estabelecer a relação entre as variáveis de um agrupamento de alunos, por exemplo, no âmbito de escola, as características dos alunos teriam que ser resumidas em médias desse agrupamento e, com isso,

perde-se muita informação relativa à variabilidade intraescolar. Todavia são justamente o processo intraescolar e sua relação integral com o aluno que motivam a pesquisa da eficácia escolar. Há variadíssimas evidências empíricas que mostram que a variabilidade dos resultados escolares deve-se, principalmente, à variabilidade intraescolar. Portanto esta abordagem também nos convém. (FERRÃO, 2003, pg. 30).

Esta segunda hipótese é a mais recorrente nos estudos ligados à eficácia escolar, uma vez que o desempenho das escolas é avaliado a partir da média dos desempenhos de seus alunos nas avaliações externas.

Assim, em ambas as hipóteses, o pressuposto de independência dos elementos dos resíduos não é atendido.

Em estudos cujos dados têm essas características de agrupamento ou hierarquização, o uso da análise de regressão linear traria algumas limitações pois implicaria os seguintes problemas, apontados por Lee (2001, 2008):

- Viés de agregação. Ocorre quando se busca verificar um fenômeno que interfere de forma distinta em diferentes níveis. Por exemplo, o nível socioeconômico do aluno interfere de uma forma quando é considerado individualmente e de outra quando se verifica a sua média no nível de turma ou escola.
- Estimação de erros padrão. Ocorre devido ao tratamento estatístico num nível individual implicando numa (não) variância falsa das informações referentes aos grupos. Por exemplo, numa escola com 300 alunos, existirá uma variabilidade inerente aos dados individuais, no entanto, as variáveis referentes à escola se repetiriam 300 vezes.
- Heterogeneidade dos coeficientes de regressão. Consiste no fato de que mesmas características individuais podem ter influências diferentes em agrupamentos diferentes. Por exemplo, dois alunos pertencentes a um mesmo nível socioeconômico pode ter influências diferentes se eles estiverem em escolas com níveis socioeconômicos distintos.
- Diferenças entre agrupamentos. O tratamento em um único nível não permite o adequado reconhecimento das diferentes influências entre os diferentes agrupamentos, se a unidade de análise for aluno, ou, da mesma maneira, não permite identificar a influência das características individuais em diferentes agrupamentos, se a unidade de análise for a escola.

Não que esses problemas inviabilizem o uso da regressão, mas certamente faz com que a estimação dos coeficientes fique menos precisa.

A superação desses problemas é a principal justificativa metodológica do uso da regressão linear que considera esse agrupamento, onde, por exemplo, as variáveis explicativas inerentes aos alunos são inseridas em um nível, e aquelas associadas às escolas em outro, com isso, sendo eliminadas as repetições.

2.3) Modelos Lineares Hierárquicos

Até agora foram resgatados os principais conceitos estatísticos ligados à análise de regressão linear, que tem por objetivo identificar as variáveis explicativas e estimar o efeito que elas têm sobre a variável resposta. Para tal, com a análise de regressão, são estimados valores para os coeficientes de uma reta, denominada como um modelo estatístico, por representar um conjunto de dados, genericamente apresentado por:

$$Y_i = \beta_0 + \beta_1.X_{1i} + \beta_2.X_{2i} + \beta_3.X_{3i} + \dots + \beta_k.X_{ki} + r_i,$$

Esses coeficientes são basicamente um coeficiente de regressão (β_k) para cada variável explicativa (X_{ki}) associada ao modelo e um intercepto (β_0). Os valores para esses coeficientes são estimados de tal maneira que levam ao menor desvio (r_i) possível entre os valores observados e aqueles estimados pelo modelo estatístico.

Para que os valores estimados desses coeficientes sejam aceitáveis, é necessário que alguns pressupostos sejam atendidos, mas um deles, o da independência dos elementos dos desvios, geralmente não é atendido em estudos da área educacional por fazer agrupamento dos dados, devido à sua forma de organização.

Na regressão linear com modelos hierárquicos são estimados os coeficientes para cada um dos agrupamentos, de maneira que as variáveis explicativas que seriam repetitivas são trabalhadas em outro nível superior. Essa hierarquização em níveis distintos permite que os resíduos sejam separados e que as variáveis explicativas sejam inseridas em função de sua vinculação a cada um desses níveis (por exemplo, variáveis explicativas inerentes aos alunos, são inseridas no nível dos alunos, enquanto que aquelas inerentes às escolas sejam inseridas no nível da escola).

Com isso, na regressão linear com modelos lineares hierárquicos ou multiníveis, as estimativas ficam mais fidedignas, uma vez que os pressupostos da independência dos elementos que compõem os desvios não são assumidos inadequadamente.

Dessa maneira, a análise de regressão que usa modelos lineares hierárquicos incorpora

naturalmente, e de 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. Vale enfatizar que, no modelo de regressão linear clássico, tanto intercepto como os coeficientes de inclinação são fixos – este é o ponto crucial no qual as duas abordagens diferem. Ferrão (2003, p. 31)

A Figura 2 ilustra esse agrupamento e a respectiva hierarquização em níveis que adjetivam os modelos estatísticos com esses nomes.



Como descrito anteriormente, com a superação do problema da não independência dos dados com características hierárquicas, fazendo com que as informações repetitivas sejam colocadas em um nível superior, é possível estimar a parte da variância total dos dados (dispersão de todos os dados em torno da média geral) que se deve a cada um dos níveis e, conseqüentemente, possibilita a inserção de variáveis explicativas em cada um dos níveis, buscando minimizar os resíduos e aumentar a capacidade explicativa dos modelos.

2.3.1) Partição da variância total em diferentes níveis

Nos modelos lineares hierárquicos, a partição da variância total entre os diferentes níveis ocorre porque, enquanto na regressão linear é estimado o coeficiente de um único modelo, a regressão linear que usa modelos hierárquicos são estimados coeficientes da reta para cada um dos agrupamentos nos diferentes níveis.

O modelo hierárquico mais simples é aquele no qual nenhuma variável explicativa é inserida e, por isso, é estimado apenas o intercepto.

No caso hipotético de uma pesquisa, na qual se pretende verificar a proficiência final de alunos de quatro escolas distintas, sem qualquer variável explicativa, seriam estimados quatro distintos interceptos, um para cada escola.

A Tabela 7 traz um conjunto de dados para exemplificar a partição da variância total e a estimação dos coeficientes de cada uma das retas.

Tabela 7 - Proficiência dos alunos; um exemplo

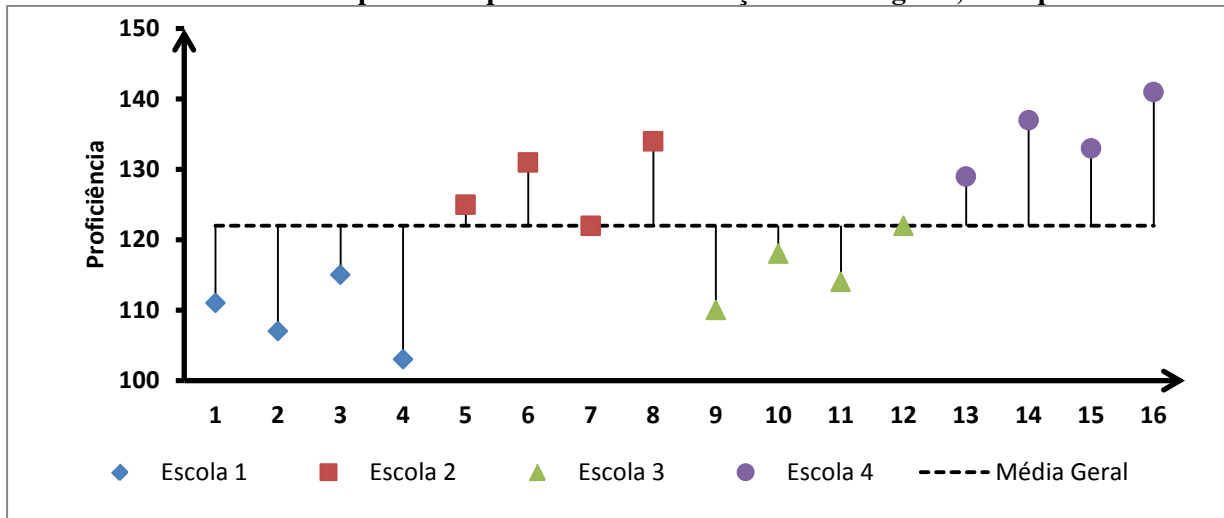
Escola	Aluno	Proficiência	Identificação Y_{ij}
1	1	111	Y_{11}
1	2	107	Y_{21}
1	3	115	Y_{31}
1	4	103	Y_{41}
2	5	129	Y_{12}
2	6	137	Y_{22}
2	7	133	Y_{32}
2	8	141	Y_{42}
3	9	110	Y_{13}
3	10	118	Y_{23}
3	11	114	Y_{33}
3	12	122	Y_{43}
4	13	125	Y_{14}
4	14	131	Y_{24}
4	15	122	Y_{34}
4	16	134	Y_{44}

Na regressão linear com dados hierarquizados, as observações são indexadas com um índice para cada nível. Como nesse exemplo há dois níveis, a variável resposta Y_{ij} é identificada por dois índices, em que i o aluno e j identifica a escola, em que ambos variam de 1 a 4. Com essa indexação, a observação Y_{24} é a proficiência do segundo aluno da quarta escola, enquanto que a observação Y_{42} é a proficiência do quarto aluno da segunda escola. As 16 observações (medidas de proficiência), de alunos de quatro escolas distintas, apresentadas nessa tabela estão representados no Gráfico 4.

Gráfico 4 – Proficiência de alunos de quatros escolas; um exemplo

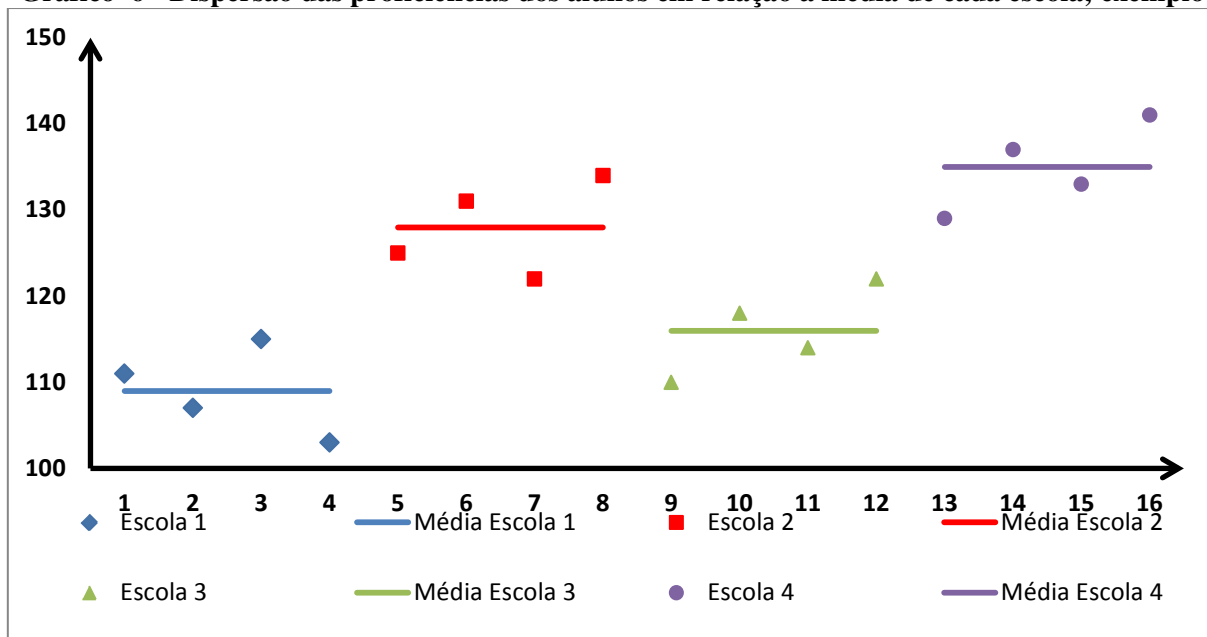
Como pode se observar cada uma das escolas tem um agrupamento bem distinto em relação às demais. O Gráfico 5 apresenta a média geral das proficiências medidas e o desvio de cada uma delas em relação a essa média, o que dá uma noção da dispersão dos dados e, portanto, da variância total (σ^2).

Gráfico 5 – Dispersão da proficiência em relação à média geral; exemplo



Como já descrito anteriormente, em um modelo linear hierárquico sem qualquer variável explicativa, é estimado apenas o intercepto de cada agrupamento, que nesse exemplo, representa a média do desempenho de cada um de seus alunos. No Gráfico 6 é apresentada a dispersão das medidas dos alunos em torno da média de cada escola.

Gráfico 6 - Dispersão das proficiências dos alunos em relação à média de cada escola; exemplo



Esse gráfico permite a visualização da dispersão das medidas dos alunos em relação a cada uma das escolas, de maneira que o desvio (resíduo) ali apresentado se dá pela distância

entre a medida observada e a média da escola que o aluno pertence. Assim, genericamente o modelo estatístico nesse nível, que por ser o primeiro é denominado Nível 1, é dado por:

$$Y_{ij} = \beta_{0j} + r_{ij}, \text{ com } r_{ij} \sim \text{NID} (0 ; \sigma^2)$$

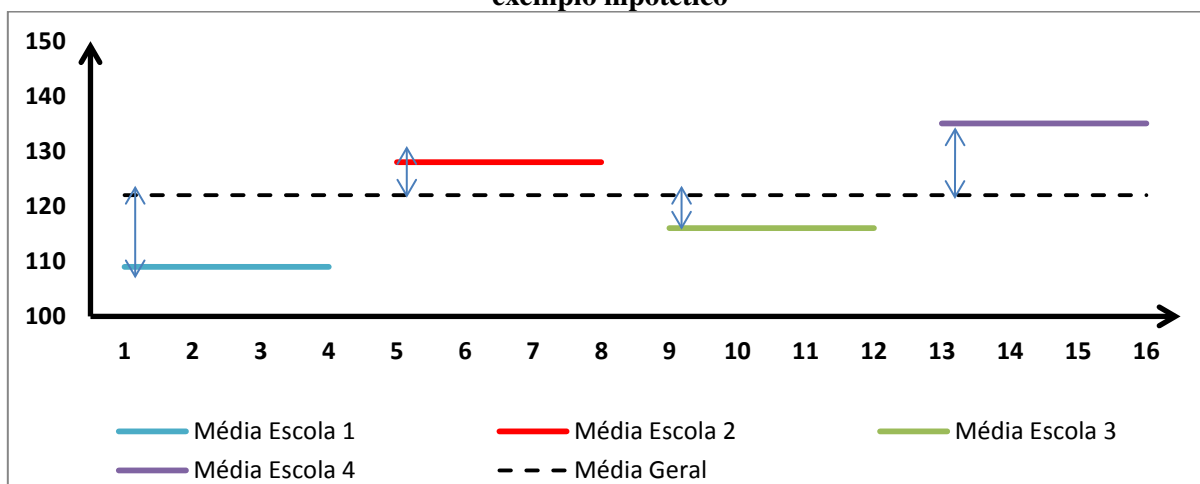
Em que:

- Y_{ij} é a proficiência observada do aluno i na escola j ;
- β_{0j} é o intercepto da escola j (a média estimada dos desempenhos dos alunos da escola);
- r_{ij} é o resíduo estimado para a observação Y_{ij} .

Nesse exemplo, os resíduos seguem a distribuição normal com média igual a zero e variância σ^2 , estimada com o valor de 27,5. Como essa variância refere-se à dispersão dos alunos de uma mesma escola, é denominada variância intraescolar¹⁶ e, assim, é estimada a parte a variância total que se refere ao nível dos alunos.

A parte da variância total que se refere às diferenças entre as escolas é encontrada a partir da dispersão da média das proficiências de cada escola em relação à média geral, estimada a partir de todos os alunos que compõem a amostra. O Gráfico 7 apresenta a dispersão das escolas.

Gráfico 7 – Resíduos das médias das proficiências de cada escola em relação à média Geral; exemplo hipotético



16 Os valores estimados para os interceptos são $\beta_{01} = 109,654$; $\beta_{02} = 134,346$, $\beta_{03} = 116,302$ e $\beta_{04} = 127,698$.

Com essa lógica, o modelo estatístico deste nível é dado por:

$$\beta_{0j} = \gamma_{00} + u_{0j}, \text{ com } u_{0j} \sim \text{NID} (0 ; \tau_{00})$$

Em que:

- β_{0j} é o intercepto da escola j (a média do desempenho dos alunos da escola);
- γ_{00} é a média geral (média da proficiência de todos os alunos da amostra);
- u_{0j} é o resíduo para a observação β_{0j} ;
- u_{0j} e r_{ij} independentes entre si.

Assim, como no nível anterior, os resíduos nesse nível seguem a distribuição normal com média igual a zero, porém, agora, com variância τ_{00} , por tratar da dispersão entre as escolas, que nesse exemplo é estimada com valor de 129,8.

Desta maneira, fica constatado que a variância total é de 157,3 (27,5+129,8), dos quais 27,5 devem-se às diferenças entre os alunos de uma mesma escola e 129,8 deve-se às diferenças entre as escolas. Com isso, a proporção da variância entre as escolas corresponde a 0,825 da variância total, o equivalente a 82,5%.

Raundebush e Bryk (2002, p. 24) denominam esse valor como coeficiente de correlação intraclasse e é calculado por:

$$\rho = \frac{\tau_{00}}{\tau_{00} + \sigma^2}$$

Na forma decimal, o coeficiente de correlação intraclasse pode variar de 0 a 1. Para a interpretação desse coeficiente, Ferrão (2003, p. 34)¹⁷ esclarece que:

Quando seu valor é nulo, significa que as escolas são homogêneas entre si e que o desempenho escolar do aluno independe da escola que ele frequenta, ou seja, toda a variabilidade dos resultados escolares deve-se apenas às diferenças entre os alunos. Na situação extrema, quando o coeficiente de correlação intraescolar tem valor 1, toda a variabilidade no desempenho deve-se a diferença entre as escolas e, em tal situação hipotética, as características individuais do aluno em nada afetam o seu desempenho escolar, ficando este a dever-se inteiramente às características da escola que ele frequenta.

17 A autora denomina por coeficiente intraescolar a mesma medida que Raundebush e Bryk (2002) denominam por coeficiente de correlação intraclasse. Em uma nota, essa autora observa que esse termo é também designado por Kreft e de Leeuw (1998, p. 9) como o coeficiente intraclasse e por Goldstein (1995, p. 9) e por Plewis (1997, p. 36) por coeficiente de correlação intraunidade.

A análise dessa variabilidade permite verificar a viabilidade da inserção de variáveis explicativas, que consiste nos procedimentos para o melhor ajuste do modelo. Quando a inserção de variáveis explicativas não mais diminui a variabilidade (medida pela variância) o modelo é considerado ajustado. Um modelo ajustado significa que é aquele que melhor representa o fenômeno estudado.

2.3.2) Modelos Lineares Hierárquicos em dois níveis

Do exposto anteriormente, um modelo linear hierárquico é composto por diferentes equações que se vinculam em níveis diferentes, e um modelo sem a inserção de qualquer variável explicativa¹⁸, é genericamente dado por:

$$\begin{aligned} \text{Nível 1: } Y_{ij} &= \beta_{0j} + r_{ij}, \text{ com } r_{ij} \sim \text{NID} (0 ; \sigma^2) \\ \text{Nível 2: } \beta_{0j} &= \gamma_{00} + u_{0j}, \text{ com } u_{0j} \sim \text{NID} (0 ; \tau_{00}) \end{aligned}$$

Com essa modelagem, onde o índice i identifica o aluno da escola j e que permite estimar separadamente a variância de cada um dos níveis (σ^2 e τ_{00}), se estrutura de maneira que as variáveis explicativas podem ser inseridas isoladamente em cada um dos níveis até que essas variâncias sejam minimizadas.

Com a inserção de uma variável explicativa (X_{1ij}) no nível aluno, o modelo fica assim descrito:

$$\begin{aligned} \text{Nível 1: } Y_{ij} &= \beta_{0j} + \beta_{1j} \cdot X_{1ij} + r_{ij}, \text{ com } r_{ij} \sim \text{NID} (0 ; \sigma^2) \\ \text{Nível 2: } \beta_{0j} &= \gamma_{00} + u_{0j}, \text{ com } u_{0j} \sim \text{NID} (0 ; \tau_{00}) \\ \beta_{1j} &= \gamma_{10} + u_{1j}, \text{ com } u_{1j} \sim \text{NID} (0 ; \tau_{10}) \end{aligned}$$

¹⁸ Raudenbush e Bryk (2002, p. 23) indicam que este é o mais simples modelo hierárquico possível, e são comumente denominados como Modelo Nulo (FERRÃO, 2003; SOARES, 2005; CRUZ, 2010; dentre outros).

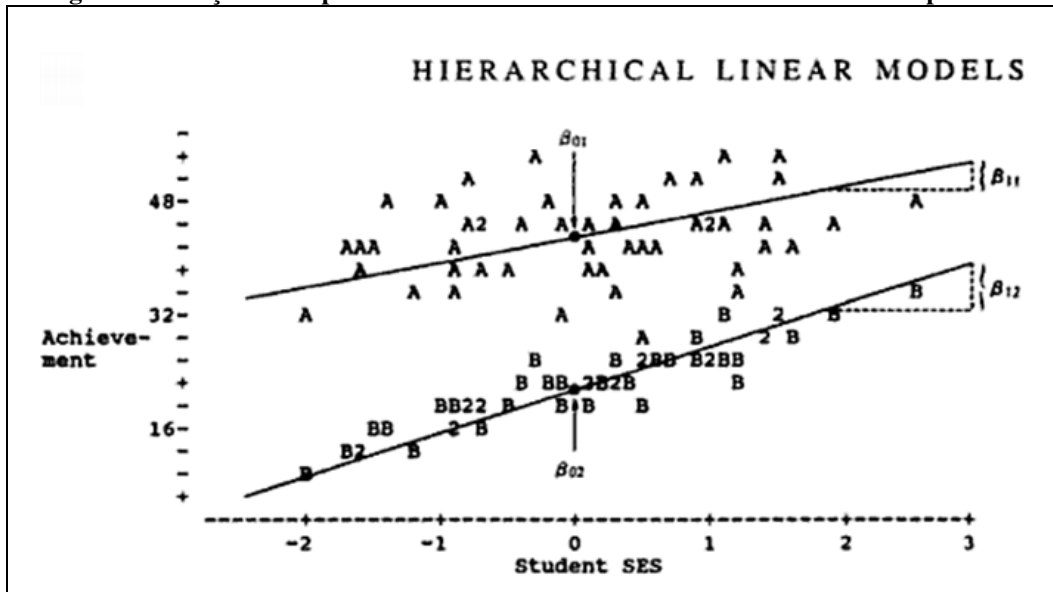
Em que:

- β_{0j} é o intercepto da escola j (a média dos desempenhos dos alunos da escola para $X_{1ij}=0$);
- X_{1ij} é a variável explicativa do Nível 1, e portanto associada ao perfil do aluno;
- β_{1j} é o coeficiente de regressão associado à variável X_{1ij} (indica o quanto o desempenho médio da escola j aumenta para cada unidade de X_{1ij});
- γ_{00} é a média geral de todos os alunos da amostra;
- γ_{10} é o coeficiente de regressão considerando todos os alunos da amostra;
- u_{0j} é o resíduo entre β_{0j} e γ_{00} ;
- u_{1j} é o resíduo entre β_{1j} e γ_{10} .

Nesse modelo, em relação ao anterior, é mantido o coeficiente que representa o intercepto para cada uma das escolas (β_{0j}) e aquele que representa o desvio (u_{0j}) desse intercepto em torno da média geral (γ_{00}), porém é inserido o coeficiente β_{1j} , que é o coeficiente de regressão para a variável explicativa X_{1ij} de cada escola j . Essas equações modelam a variabilidade do intercepto e do coeficiente de regressão, das escolas, por isso estão em outro nível e são identificadas como equações do Nível 2.

Usando outro exemplo hipotético, a Figura 3, onde os desempenhos (*Achievement*) de alunos de duas escolas (A e B) são representados graficamente em função de seus respectivos níveis socioeconômicos (*socioeconomic status – Student SES*), é útil para ilustrar a ideia fundamental empregada nos modelos lineares hierárquicos com dois níveis (aluno e escola) e com uma única variável explicativa. Nela, onde a variável resposta (Y_{ij}) é o desempenho e a variável explicativa (X_{1ij}) é o nível socioeconômico, é evidenciada a reta de regressão das duas escolas e seus coeficientes, β_{01} e β_{11} para a escola A e β_{01} e β_{12} para a escola B.

Figura 3- Relação entre proficiência e nível socioeconômico de duas escolas hipotéticas



Fonte: Raudenbush e Bryk (2002, p. 18)

Com esse equacionamento, são estimados os valores para os coeficientes de duas retas, como mostra a Figura 3, onde estão explicitados dois diferentes interceptos, uma para cada escola (β_{01} e β_{02}) e também dois distintos coeficientes de regressão (β_{11} e β_{12}) permitindo, assim, um ajuste específico para cada uma das escolas. Neste exemplo são apresentados dados hipotéticos de duas escolas, mas em situações nas quais há um número maior de escolas, o processo é o mesmo, ou seja, para cada escola são estimados os coeficientes lineares e de regressão.

Mas mesmo com o ajuste de retas específicas para cada escola, pode-se observar que os dados levantados empiricamente têm certo distanciamento dessas retas (dado por r_{ij}), indicando que ainda existe uma variância σ^2 .

Ainda da Figura 3 pode-se notar que os interceptos das diferentes retas têm valores diferentes e, portanto, se distanciam da média geral (γ_{00}). Essa distância (desvio) é dada por u_{0j} , com uma variância τ_{00} . De maneira análoga ao que ocorre com a estimativa do intercepto (β_{0j}), o modelo evidencia que as estimativas para β_{1j} tem um desvio u_{1j} em torno do coeficiente de regressão geral γ_{10} , isto significa que se a variância de β_{1j} não é nula, cada escola é influenciada diferentemente pela variável explicativa X_{1ij} .

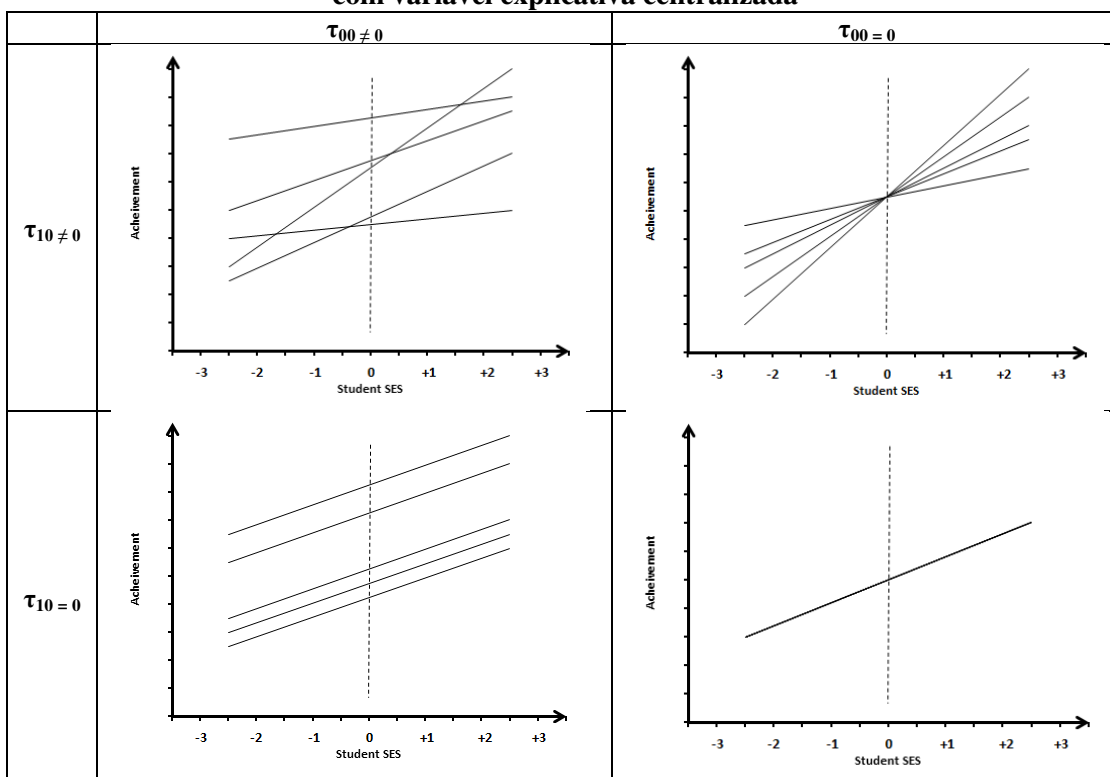
Dessas afirmativas algumas propriedades do modelo podem ser destacadas:

- a) se existirem resíduos no Nível 1 (no exemplo, nível dos alunos), a variância σ^2 será diferente de zero e isso significa que há diferença entre os alunos da mesma escola;
- b) no Nível 2 (nesse exemplo, nível das escolas), se existirem diferenças entre os interceptos das escolas, a variância τ_{00} será diferente de zero, e;
- c) ainda no nível 2, se as variáveis explicativas interferem diferentemente em cada escola, isto é, o coeficiente de regressão é diferente em cada escola, τ_{10} será diferente de zero.

No exemplo apresentado na Figura 3, as diferentes posições dos interceptos β_{01} e β_{02} indicam que τ_{00} é diferente de zero, assim como as diferentes inclinações das retas das escolas A e B, representados pelos coeficientes β_{11} e β_{12} , indicam que a variância τ_{10} é diferente de zero.

Na Figura 4 estão representados graficamente os modelos de cinco escolas, em que são combinadas as diferentes possibilidades das variâncias do intercepto (τ_{00}) e do coeficiente de regressão (τ_{10}), sendo iguais ou diferentes de zero.

Figura 4 - Representações gráficas dos modelos para diferentes status da variância com variável explicativa centralizada



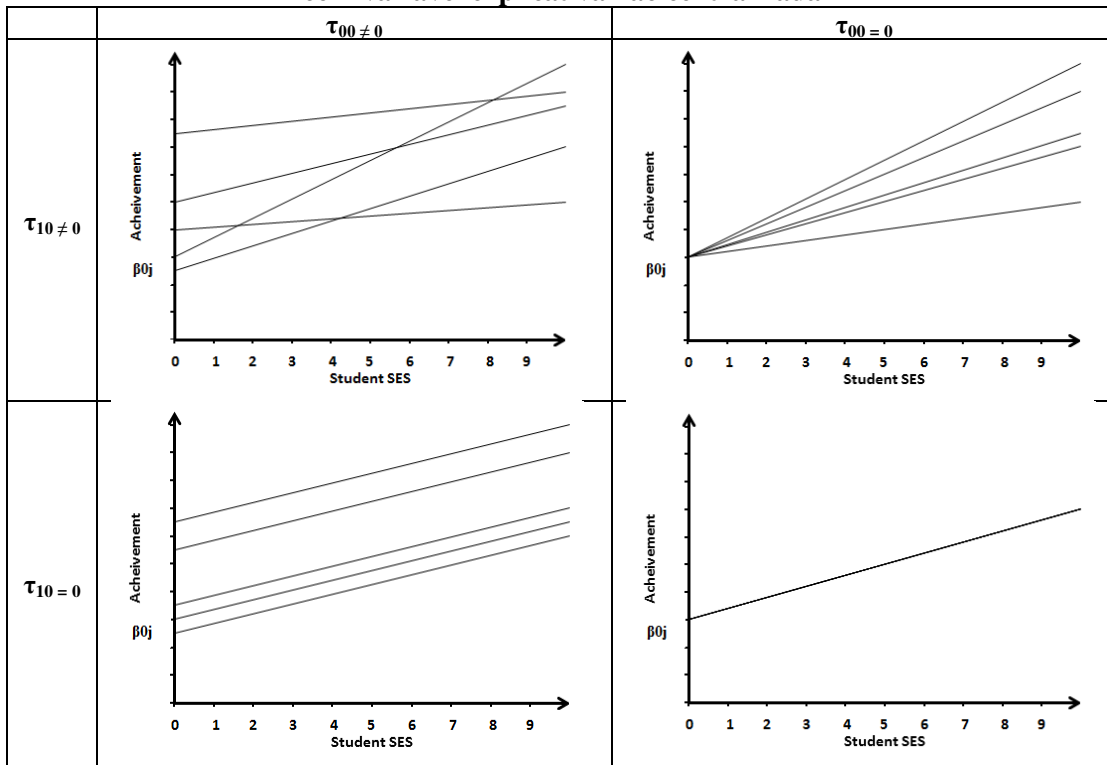
Essas representações gráficas permitem as seguintes interpretações:

- quando $\tau_{00} \neq 0$ e $\tau_{10} \neq 0$, significa que há variação entre os interceptos e também entre os coeficientes de regressão e, portanto, é interessante encontrar variáveis explicativas para as duas equações do Nível 2. Esta situação é a mesma daquela representada na Figura 4;
- quando $\tau_{00} = 0$ e $\tau_{10} \neq 0$, significa que há variação apenas entre os coeficientes de regressão e, portanto, todas as escolas têm desempenhos médio iguais para os alunos com nível socioeconômico igual a zero. Nessa situação é interessante encontrar variáveis explicativas apenas para os coeficientes de regressão do Nível 2;
- quando $\tau_{00} \neq 0$ e $\tau_{10} = 0$, significa que há variação apenas entre os interceptos das diferentes escolas e que as inclinações das retas são as mesmas. Portanto, o efeito do nível socioeconômico sobre o desempenho médios dos alunos das diferentes escolas é o mesmo. Nessa situação é interessante encontrar variáveis explicativas apenas para serem inseridas na equação do Nível 2 referente ao intercepto;
- quando $\tau_{00} = 0$ e $\tau_{10} = 0$, significa que não há variação nem nos interceptos e nem nas inclinações das retas e conseqüentemente todas as escolas são representadas pelos mesmos coeficientes da reta. Logo, não há diferenças entre as escolas e, conseqüentemente, toda a variabilidade existentes entre os alunos pode ser explicada com as variáveis vinculadas ao Nível 1. Nesta situação, o coeficiente de correlação intraclasse (ρ) é igual a 0.

É importante ressaltar que nos esboços dos gráficos colocados na Figura 4, os interceptos (destacados com a linha tracejada) encontra-se centralizado porque a variável explicativa usada tem como valor central o zero. Isso ocorre sempre que a variável explicativa é padronizada, como frequentemente ocorre com as medidas de nível socioeconômico.

Caso a variável explicativa não seja centralizada, as interpretações são as mesmas, porém os esboços dos gráficos ficam diferentes, como é apresentado na Figura 5.

Figura 5 - Representações gráficas dos modelos para diferentes status da variância com variável explicativa não centralizada



Assim, o processo de ajuste dos modelos lineares hierárquicos (da mesma maneira que ocorre na regressão linear) consiste em inserir as variáveis explicativas (evidentemente, desde que tenham sido medidas) até que a variância em cada um dos níveis seja significativamente diminuída.

Por esse motivo, durante o processo de ajuste dos modelos a análise das estimativas dos interceptos e dos coeficientes de regressão de cada variável explicativa, nos diferentes níveis denominados efeitos fixos, são menos importantes do que a análise da variância de cada um desses coeficientes (σ^2 , τ_{00} e τ_{10}), denominados efeitos aleatórios.

Genericamente, um modelo linear hierárquico de dois níveis, no qual são inseridas variáveis explicativas em cada nível é dado por:

$$\text{Nível 1: } Y_{ij} = \beta_{0j} + \beta_{1j} \cdot X_{1ij} + r_{ij}, \text{ com } r_{ij} \sim \text{NID}(0; \sigma^2)$$

$$\text{Nível 2: } \beta_{0j} = \gamma_{00} + \gamma_{01} \cdot W_{1j} + u_{0j}, \text{ com } u_{0j} \sim \text{NID}(0; \tau_{00})$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} \cdot W_{1j} + u_{1j}, \text{ com } u_{1j} \sim \text{NID}(0; \tau_{10})$$

Em que, com relação ao modelo especificado anteriormente, foram inseridos:

- W_{1j} é a variável explicativa do Nível 2 e, portanto, associada às características da escola;
- γ_{01} é o coeficiente de regressão associado à variável W_{1j} e que interfere no intercepto β_{0j} , ou seja, é o coeficiente que indica o quanto cada unidade da variável explicativa W_{1j} incrementa em β_{0j} ;
- γ_{11} é o coeficiente de regressão associado à variável W_{1j} e que interfere no coeficiente de regressão β_{1j} , ou seja, é o coeficiente que indica o quanto cada unidade da variável explicativa W_{1j} incrementa em β_{1j} .

Para cada variável explicativa inserida na equação do Nível 1 é gerada uma nova equação correspondente ao coeficiente de regressão a ela associada. Assim, se uma equação do Nível 1 tiver 10 variáveis explicativas, o Nível 2 terá uma equação para modelar o coeficiente de regressão de cada uma das variáveis explicativas além de uma que modela o intercepto.

2.3.3) Modelos Lineares Hierárquicos em três níveis

Caso um conjunto de dados requeira mais níveis para uma modelagem mais adequada, a mesma lógica de dois níveis é aplicada.

Em uma situação exemplo, na qual o objetivo é estudar a proficiência de alunos agrupados em turmas e essas turmas são agrupadas em escolas, um modelo linear hierárquico com três níveis deverá ser utilizado, exigindo o uso de três índices:

- k identifica a escola, e pode assumir valores entre 1 a K , em que K indica o número total de escolas. Resumidamente essa informação é representada por ($k = 1, \dots, K$ escolas);
- j identifica a turma, podendo assumir valores entre 1 a J_k , em que J_k representa o número total de turmas da escola k . Resumidamente essa informação é representada por ($j = 1, \dots, J_k$ turmas);
- i identifica o aluno, podendo assumir valores entre 1 a n_{jk} , em que n_{jk} representa o número total de alunos da turma j da escola k . Resumidamente essa informação é representada por ($i = 1, \dots, n_{jk}$ alunos).

Um modelo linear hierárquico com três níveis, sem qualquer variável explicativa, o Modelo Nulo¹⁹, é dado por:

$$\text{Nível 1: } Y_{ijk} = \pi_{0jk} + e_{ijk}, \text{ com } e_{ijk} \sim \text{NID} (0; \sigma^2)$$

$$\text{Nível 2: } \pi_{0jk} = \beta_{00k} + r_{0jk}, \text{ com } r_{0jk} \sim \text{NID} (0; \tau_\pi)$$

$$\text{Nível 3: } \beta_{00k} = \gamma_{000} + u_{00k}, \text{ com } u_{00k} \sim \text{NID} (0; \tau_\beta)$$

Em que:

- Y_{ijk} é a proficiência final do i -ésimo aluno, matriculado na j -ésima turma da k -ésima escola;
- π_{0jk} é o intercepto do Nível 1 e representa a média das proficiências dos n_{jk} alunos da turma j da escola k ;
- e_{ijk} é o resíduo do Nível 1, isto é, é o desvio da proficiência do aluno i da turma j da escola k em relação à média das proficiências dos n_{jk} alunos da turma j da escola k . Estes resíduos são assumidos como normalmente distribuídos com média zero e variância σ^2 ;
- β_{00k} é o intercepto do Nível 2 e representa a média das proficiências das J_k turmas da escola k ;
- r_{0jk} é o resíduo do Nível 2, isto é, é o desvio da proficiência média da turma j da escola k em relação à média das proficiências médias das J_k turmas da escola k . Estes resíduos são assumidos como normalmente distribuídos com média zero e variância τ_π . Além disso, dentro de cada uma das K escolas, as variâncias entre as turmas são assumidas sendo as mesmas;
- γ_{000} é a média geral da proficiência de todos os alunos que compõem a amostra;
- u_{00k} é a componente aleatória do Nível 3, isto é, é o desvio da proficiência média da escola k em relação à média geral da proficiência de todos os alunos. Estes resíduos são assumidos como normalmente distribuídos com média zero e variância τ_β .

¹⁹ Raudenbush e Bryk (2002, p. 229) o denominam como Modelo Totalmente Incondicional.

Como a hierarquização dos dados foram agrupados em três níveis, de maneira análoga ao Modelo Nulo com dois níveis, a partição da variância total é dada por:

- $\rho_1 = \frac{\sigma^2}{\sigma^2 + \tau_\pi + \tau_\beta}$ é a variância do Nível 1, dentro das turmas; [01]

- $\rho_2 = \frac{\tau_\pi}{\sigma^2 + \tau_\pi + \tau_\beta}$ é a variância do Nível 2, entre as turmas; [02]

- $\rho_3 = \frac{\tau_\beta}{\sigma^2 + \tau_\pi + \tau_\beta}$ é a variância do Nível 3, entre as escolas. [03]

Buscando a minimização dessas variâncias, em cada um dos níveis podem ser inseridas variáveis explicativas.

Com a inserção de uma variável explicativa em cada um dos níveis o modelo fica assim constituído:

Nível 1: $Y_{ijk} = \pi_{0jk} + \pi_{1jk} \cdot X_{ijk} + e_{ijk}$, com $e_{ijk} \sim \text{NID}(0; \sigma^2)$

Nível 2: $\pi_{0jk} = \beta_{00k} + \beta_{01k} \cdot Z_{jk} + r_{0jk}$, com $r_{0jk} \sim \text{NID}(0; \tau_{\pi 0})$

$\pi_{1jk} = \beta_{10k} + \beta_{11k} \cdot Z_{jk} + r_{1jk}$, com $r_{1jk} \sim \text{NID}(0; \tau_{\pi 1})$

Nível 3: $\beta_{00k} = \gamma_{000} + \gamma_{010} \cdot W_k + u_{00k}$, com $u_{00k} \sim \text{NID}(0; \tau_{\beta 00})$

$\beta_{01k} = \gamma_{010} + \gamma_{011} \cdot W_k + u_{01k}$, com $u_{01k} \sim \text{NID}(0; \tau_{\beta 01})$

$\beta_{10k} = \gamma_{100} + \gamma_{101} \cdot W_k + u_{10k}$, com $u_{10k} \sim \text{NID}(0; \tau_{\beta 10})$

$\beta_{11k} = \gamma_{110} + \gamma_{111} \cdot W_k + u_{11k}$, com $u_{11k} \sim \text{NID}(0; \tau_{\beta 11})$

E, em relação ao modelo anterior, tem-se que:

- X_{ijk} , Z_{ijk} e W_{ijk} são as variáveis explicativas dos Níveis 1, 2 e 3, respectivamente;

- π_{1jk} é o coeficiente de regressão associada à variável explicativa do Nível 1;

- β_{0jk} e β_{1jk} são os coeficientes de regressão associados à variável explicativa do Nível 2 e indicam o quanto cada unidade dessa variável incrementa o intercepto e o coeficiente de regressão do Nível 1;

- γ_{010} e γ_{011} são os coeficientes de regressão associados à variável explicativa do Nível 3 que interferem nas equações do Nível 2 que, por sua vez, interferem no intercepto da equação do Nível 1;

- γ_{101} e γ_{111} são os coeficientes de regressão associados à variável explicativa do Nível 3 que interferem nas equações do Nível 2 que, por sua vez, interferem no coeficiente de regressão da equação do Nível 1.

Notadamente, a complexidade do modelo é significativamente aumentada sempre que novas variáveis são associadas a cada um dos níveis, uma vez que, a inserção de uma variável em um nível, uma ou mais equações são criadas nos níveis superiores.

2.3.4) Significância estatística da variável explicativa – *p-value*

Mas diante de um número expressivo de variáveis explicativas disponíveis para que sejam inseridas no modelo, deve ser considerada a possibilidade de que algumas delas não sejam capazes de diminuir a variabilidade da variável resposta. Segundo Soares (2005, p. 80)

O processo básico mais utilizado na construção de um modelo hierárquico é do tipo *bottom-up*, isto é, parte-se do modelo nulo e vai-se incluindo as variáveis segundo uma heurística definida pelo especialista, que neste estudo se baseia na verificação da significância dos coeficientes (parâmetros fixos e aleatórios) para cada modelo.

Esta citação destaca dois aspectos importantes no procedimento de ajuste do modelo que são a ordem pela qual as variáveis explicativas podem ser inseridas no modelo, assunto a ser analisado no Capítulo 4, e a verificação da significância dos coeficientes.

Cotidianamente, o termo *significante* é usado para indicar que algo é importante, no entanto, na estatística esse termo indica que um valor estimado é provavelmente verdadeiro. Essa informação é importante, pois em modelos não determinísticos, como é esse trabalho nesta pesquisa, todo valor estimado leva a uma aproximação daquilo que é real.

Uma primeira verificação consiste na análise da significância estatística das estimativas dos coeficientes associados a uma variável explicativa inserida no modelo. Essa verificação é feita através do teste estatístico de Wald, que estima um valor denominado *p-value*. Quando um *p-value* estimado é maior que 0,05, implica que o valor estimado para o coeficiente não deve ser considerado diferentes de zero, de maneira que:

- Se o coeficiente é um intercepto, este deve ser interpretado como tendo valor nulo;

- Se o coeficiente é um coeficiente de regressão, a variável explicativa a ele associada deve ser excluída do modelo, uma vez que o coeficiente nulo (ou muito próximo de zero) significa que essa variável não tem qualquer efeito (ou um efeito não importante) sobre a variável resposta;

- Se o coeficiente for uma variância, significa que os desvios entre os valores reais (observados) e os estimados não mais existem e, portanto, o modelo deve ser considerado ajustado, não sendo mais necessária a inserção de quaisquer outras variáveis explicativas.

Assim, “a estatística de Wald constitui um bom teste durante a triagem inicial das variáveis (análises univariadas), servindo para apontar, nesta etapa, quais as variáveis que deverão compor os modelos multivariados” (CRUZ, 2010, p. 120)

No entanto, mesmo quando os coeficientes estimados são considerados estatisticamente significativos, a verificação da significância deve perpassar por outra dimensão que é a análise da magnitude do efeito que a variável explicativa tem sobre a variável resposta. No caso do valor estimado para o coeficiente de regressão indicar que o efeito é pequeno, a manutenção da variável explicativa deve ser avaliada, dado que a inserção de uma variável explicativa implica a criação de outras equações nos níveis superiores tornando a interpretação dos resultados do modelo muito complexa.

2.3.5) O grau de ajuste do Modelo Linear Hierárquico - *deviance*

A busca do modelo linear hierárquico que melhor representa os dados observados consiste na inserção de variáveis explicativas “segundo uma heurística definida pelo pesquisador” (SOARES, 2009, p. 80) até que as variâncias sejam minimizadas ou não sejam mais diferentes de zero de maneira significativa. Essa verificação é feita em função de cada variável explicativa, seja ela inserida individualmente ou em blocos.

Há outro teste estatístico que estima o grau de ajuste do modelo como um todo chamado *deviance*, que consiste em uma razão de verossimilhança. Quanto menor for o valor estimado da *deviance* do modelo, maior será o seu ajuste aos dados observados.

Normalmente esse teste é aplicado para comparar dois distintos modelos. Pode ocorrer que modelos com diferentes quantidades de variáveis explicativas e, portanto, com diferentes graus de complexidade, tenham o mesmo grau de ajuste. Para essa análise é feito outro teste estatístico, denominado quiquadrado, que verifica se a diferença entre a *deviance*

dos dois modelos é significativa. Neste teste também é estimado um *p-value* que deve ser inferior a 0,05 para que essa diferença seja considerada significativa.

2.3.6) Justificativa para o uso dos modelos lineares

Ferrão (2003, p. 31) destaca as seguintes vantagens da utilização do modelo de regressão linear hierárquico:

- Possibilita a obtenção de estimativas eficientes dos coeficientes de regressão;
- Leva em conta a informação sobre o agrupamento dos dados, tornando as medidas referentes ao erro-padrão, intervalos de confiança e testes de hipóteses corretos;
- Permite o uso de variáveis explicativas mensuradas em cada nível da hierarquia, permitindo ao analista a exploração detalhada do impacto e da contribuição de cada nível para a variabilidade da variável resposta;
- Permite estabelecer listas comparativas do desempenho institucional, que vem ganhando relevo na educação evidenciando a impossibilidade técnica de estabelecer listas ordenadas das escolas tal como elas têm sido apresentadas em alguns países²⁰.

Outra possibilidade com os modelos lineares hierárquicos é a determinação da variabilidade dos dados no interior de cada uma das escolas e, devido à hierarquização dos dados (ou agrupamento), também a variabilidade entre as escolas.

Apesar de Raundenbush e Bryk (2002) evidenciarem que os modelos hierárquicos são discutidos, mesmo que com outras denominações, em diversas áreas do conhecimento desde a década de 1960, a distorção das estimativas devido à falta da independência provocada pelo agrupamento de observações individuais foi demonstrada pela primeira vez por Aitkin e Longford (1986). Graças ao avanço tecnológico que viabilizam os cálculos, desde então a regressão linear com modelos hierárquicos são usados cada vez mais frequentemente quando se deseja estudar fenômenos cuja compreensão depende não apenas de características de observações individuais, mas também das características específicas de seus agrupamentos.

Assim, modelos multiníveis e modelos lineares hierárquicos são diferentes nomenclaturas para o mesmo recurso estatístico, derivada da regressão linear quando as

20 Essa observação feita por Ferrão (2003) refere-se ao problema demonstrado por Aitkin & Longford (1986), acerca dos erros devidos ao não atendimento ao pressuposto da independência dos dados. No entanto, a observação é feita por ela ainda é controversa dado que para avaliar eficácia através de um único indicador (unidimensional) é incompatível com a complexidade (multidimensionalidade) inerente às escolas, seja ela ligada aos resultados encontrados nas avaliações, seja em suas próprias condições de funcionamento.

informações não se referem ao mesmo nível. A primeira nomenclatura, modelos multiníveis, é usada pelo pesquisador inglês Goldstein (1995) e está associada ao programa computacional MLWIN, enquanto que a segunda denominação é usada pelos americanos Raudenbush e Bryk (2002) e está associado ao programa computacional HLM.

Lee (2010a) sintetiza essas características na melhor adequação dos modelos lineares hierárquicos à área educacional quando destaca que:

- O processo educacional é multinível, porque os estudantes estão agrupados em salas de aula e escolas.
- As hipóteses levantadas nas pesquisas sobre o efeito-escola são de caráter multinível, pois procuram identificar de que maneira as características da escola influenciam os alunos.
- Pesquisas que não levam em consideração essa natureza multinível da escola quase sempre apresentam resultados com “viés” (usamos o termo “viés” no mundo dos pesquisadores, mas o que realmente isso significa é “erro”).
- A direção do viés é quase sempre a de subestimar o efeito-escola.
- Geralmente, o método multinível – a metodologia *Hierarchical Linear Model* (HLM), ou seja, Modelo Linear Hierárquico, é uma delas – divide a variância de um resultado (o desempenho do aluno, por exemplo) em duas partes: a primeira, que varia sistematicamente entre as escolas, e a segunda, que varia entre os alunos de uma mesma escola.
- Somente a componente da variância referente ao desempenho entre escolas, é a que pode ser influenciada pelas características da instituição. (LEE, 2010a, p. 475)

2.4) Modelos Lineares Hierárquicos ajustados ano a ano

Esta pesquisa, para identificar os fatores que interferem na proficiência final dos alunos, pretende considerar a complexidade do fenômeno educacional que transcende os limites da própria escola e o tempo que o aluno nela permanece. Considera-se também que inúmeros fatores influenciam na sua aprendizagem mesmo antes de seu ingresso na escola. Para Lee (2008, p. 273)

o aprendizado infantil é fortemente influenciado pelo contexto educacional que ele está inserido. Quando as crianças são muito jovens, a família é o mais importante contexto para o aprendizado. Com o passar do tempo, as atividades educacionais das crianças acontecem mais e mais em ambientes formais.

As informações disponibilizadas pelo Projeto GERES, após o devido tratamento, são consideradas como variáveis explicativas, permitindo assim, verificar sua associação com a proficiência final dos alunos e conseqüentemente identificar alguns dos fatores citados anteriormente, objetivo desta pesquisa.

Desse contexto e considerando o que foi exposto no Capítulo 2, foram definidos as seguintes premissas para a elaboração do modelo:

a) ser multinível para atender aos pressupostos para o uso da análise de regressão linear;

b) ter até três níveis hierárquicos, limite do programa computacional selecionado para o estudo²¹;

c) explorar os conceitos de valor agregado e de efeito-escola aproveitando a disponibilização de dados longitudinais e, para tal, buscar eliminar as diferenças entre as proficiências dos alunos antes de sua entrada na escola;

d) considerar o peso amostral durante o processamento, para que os resultados encontrados possam ser usados de forma inferencial para as demais escolas que compõem o universo²²;

e) aproveitar a maior quantidade possível de dados disponibilizados pelo Projeto GERES, para evitar perdas adicionais àquelas inerentes à mobilidade escolar captada pelos estudos longitudinais;

f) manter os alunos que não permaneceram na mesma escola, mas permaneceram em escolas participantes do Projeto GERES;

g) permitir a inserção do maior número possível de variáveis explicativas para considerar a complexidade do fenômeno educacional;

h) processar, os dados dos diferentes anos de escolarização em um único modelo, mas que permita a identificação dos fatores e da influência de cada um deles, nos distintos anos de escolarização;

i) identificar a influência dos fatores encontrados na proficiência final em alunos com diferentes proficiências iniciais;

j) trabalhar com as proficiências de leitura e matemática separadamente dado que as proficiências foram elaboradas em escalas diferentes.

Dadas essas premissas, o modelo proposto em três níveis ficou assim definido:

21 O programa computacional utilizado foi o HLM, versão 6.08, fornecido pela Scientific Software International - SSI. Atualmente há uma versão que suporta quatro níveis hierárquicos e dados longitudinais, mas não estava disponível nos primeiros anos do desenvolvimento desta pesquisa.

22 Para maior aprofundamento, ver Miranda e Andrade (2005) e Brooke e Bonamino (2011).

- Nível 1, denominado Medidas Cognitivas no Tempo, onde se encontram as informações sobre proficiência final, proficiência inicial, variáveis indicadoras²³ para o controle do ano de escolarização e outra variável *indicadora* para o controle sobre uma possível retenção do aluno ao longo do processo de escolarização²⁴. Para a identificação do registro no arquivo associado a esse nível são usados o código da escola, do aluno e o ano de escolarização.

- Nível 2, denominado Alunos, onde se encontram todas as informações que caracterizam o perfil do aluno, identificação dos registros desse nível são usados o código da escola e do aluno;

- Nível 3, denominado Escolas, no qual se encontram todas as informações específicas de cada escola, identificadas pelo código da escola.

Seria desejável a inclusão de quarto nível hierárquico, entre o aluno e a escola, afinal, os alunos são agrupados em turmas e essas são agrupadas em escolas. No entanto, essa proposição precisou ser descartada pelos seguintes motivos:

- A limitação do programa computacional em três níveis;
- O grande número de escolas com menos de quatro turmas, como mostra a Tabela 8;
- A mobilidade dos alunos entre as turmas, provocando uma perda maior de informações.

As informações da Tabela 8 mostram que a variabilidade entre as turmas poderia ser considerada praticamente como sendo à da escola (quando tendo uma turma) ou não permite o cálculo da variância entre as turmas de uma mesma escola, o que levaria ao descarte dessa

23 Na análise de regressão, as variáveis quantitativas são facilmente correlacionadas com as demais, não ocorrendo o mesmo com a de natureza qualitativa. Assim, uma forma de quantificar os atributos qualitativos dessa variável é construir variáveis artificiais para assumirem valores binários. Na literatura relacionada à regressão linear esses variáveis artificiais são chamadas de variáveis *dummies*, mas que nesse trabalho serão denominadas como variáveis indicadoras.

24 Durante a aplicação dos testes usados para medir as proficiências dos alunos, aqueles que haviam sido retidos ao final da série anterior e que permaneciam na escola participante do Projeto GERES eram deslocados de sua turma (em uma série anterior) para aquela que estava sendo aplicado o teste. Assim, alunos que cursavam séries diferentes eram submetidos ao mesmo teste.

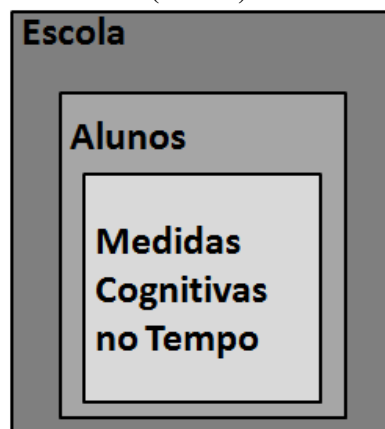
opção, mesmo que existisse a possibilidade do programa e não existisse a mobilidade de alunos entre as turmas²⁵.

Tabela 8 - Número de escolas com menos de quatro turmas

	Freq.	Total	%
Ano 1	155	223	69,5%
Ano 2	156	223	70,0%
Ano 3	63	223	28,3%
Ano 4	95	228	41,7%
Ano 5	129	229	56,3%

Assim, na Figura 6 está representada essa proposta de modelagem, idêntico ao desenho da quinta estratégia analítica para os estudos de eficácia e equidade escolar indicada por Brooke e Bonamino (2011, p. 168):

Figura 6 - Diferentes medidas cognitivas ao longo do tempo (Nível 1) dentro de alunos (Nível 2) dentro de escolas (Nível 3)



Nessa modelagem foi considerada como variável resposta a proficiência final de cada ano e todas as demais informações disponíveis no banco de dados utilizado foram processadas como variáveis explicativas, ou seja, o modelo coloca a proficiência final em função de todas as outras variáveis disponíveis, em diferentes níveis, para que com o processamento estatístico, sejam identificadas as variáveis explicativas, significantes, e estimados seus coeficientes.

25 A opção metodológica para a superação deste aspecto traria outra implicação, uma vez que, caso fossem selecionadas escolas com maior número de turmas, as escolas com menor número de alunos deixariam de participar do estudo.

Considerando o desenho metodológico usado para a coleta de dados, para a viabilização dessa modelagem, foi atribuído à proficiência inicial de um determinado ano o valor da proficiência final do ano anterior, com exceção, obviamente, do primeiro ano de escolarização que teve uma onda de testes aplicada no início do ano letivo. O Quadro 3 explicita essa atribuição de valores.

Quadro 3 - Atribuição de Valores para proficiência final e inicial de cada ano

	Proficiência	
	Proficiência Final	Proficiência Inicial
Ano1	Valor da Onda 2	Valor da Onda 1
Ano2	Valor da Onda 3	Valor da Onda 2
Ano3	Valor da Onda 4	Valor da Onda 3
Ano4	Valor da Onda 5	Valor da Onda 4

Desta forma, o valor da proficiência inicial no primeiro ano de escolarização é aquele estimado na onda de testes aplicada em março 2005; do segundo ano, estimado em novembro de 2005; do terceiro ano, em novembro de 2006 e, finalmente, para o quarto ano o valor estimado na onda aplicada em novembro de 2007. Os valores atribuídos para as proficiências finais de cada ano de escolarização provêm das coletas realizadas ao final dos anos de 2005, 2006, 2007 e 2008.

2.4.1) O modelo sem variáveis explicativas – O Modelo Nulo

É o modelo ajustado apenas com a proficiência sem incluir nenhuma variável explicativa. A finalidade do ajuste deste modelo é estimar a partição da variância total entre os três níveis hierárquicos, isto é, visa estimar a parte da variância total que se deve às diferenças entre as medidas feitas ao longo dos quatro anos de escolarização de cada um dos alunos, às diferenças entre os alunos de uma mesma escola e às diferenças entre as escolas.

Com essas estimativas serão verificadas as seguintes hipóteses:

- Hipótese 1: Há variabilidade significativa entre as medidas da proficiência dos alunos nos diferentes anos de escolarização.
- Hipótese 2: Há variabilidade significativa nas medidas da proficiência média entre os alunos de uma mesma escola nos diferentes anos de escolarização.

▪ Hipótese 3: Há variabilidade significativa entre as proficiências médias das diferentes escolas nos diferentes anos de escolarização.

O Modelo Nulo, com esta estruturação é composto pelas seguintes equações:

Nível 1 – Medidas Cognitivas ao longo do tempo:

$$PR_FI_{ijk} = \pi_{0jk} + e_{ijk}, \text{ com } e_{ijk} \sim \text{NID}(0, \sigma^2)$$

Em que,

- PR_FI_{ijk} é a proficiência final do ano de escolarização i (que varia de 1 a 4 anos) do aluno j (que varia de 1 a J_k alunos, sendo que J_k é o número total de alunos da escola k) da escola k (que varia de 1 a K , sendo K o número total de escolas);
- π_{0jk} é a média das proficiências finais do aluno j , na escola k ;
- e_{ijk} é o desvio entre cada uma das proficiências finais do ano i do aluno j matriculado na escola k e a média de suas proficiências finais;
- σ^2 variância de e_{ijk} .

Nível 2 – Alunos)

$$\pi_{0jk} = \beta_{00k} + r_{0jk}, \text{ com } r_{0jk} \sim \text{NID}(0, \tau_\pi)$$

Em que:

- β_{00k} é a média das proficiências finais da escola k calculada a partir das proficiências médias de cada um dos J_k alunos da escola k ;
- r_{0jk} é o desvio entre a média das proficiências finais do aluno j da escola k e a média das proficiências finais da escola k ;
- τ_π é a variância de r_{0jk} .

Nível 3 – Escolas)

$$B_{0jk} = \gamma_{000} + u_{00k}, \text{ com } u_{00k} \sim \text{NID}(0, \tau_\beta)$$

Em que:

- γ_{000} a média das proficiências finais de todos os alunos que compõem o banco de dados;
- u_{00k} é o desvio entre proficiência final média da escola k e a média das proficiências finais de todos os alunos que compõem o banco de dados;
- τ_{β} é a variância devida de u_{00k} .

Com essa modelagem é possível estimar a parte da variância total que se deve a cada um dos níveis hierárquicos, mostram as equações a seguir, que se referem às Equações [01], [02] e [03], apresentadas anteriormente.

- $\rho_1 = \frac{\sigma^2}{\sigma^2 + \tau_{\pi} + \tau_{\beta}} \times 100$ é a variância do Nível 1, isto é, variância das proficiências finais em relação à média das proficiências finais de cada aluno;

- $\rho_2 = \frac{\tau_{\pi}}{\sigma^2 + \tau_{\pi} + \tau_{\beta}} \times 100$ é a variância do Nível 2, isto é, variância entre as proficiências finais médias de cada aluno em relação à média das proficiências finais das escolas na qual esses alunos são matriculados;

- $\rho_3 = \frac{\tau_{\beta}}{\sigma^2 + \tau_{\pi} + \tau_{\beta}} \times 100$ é a variância do Nível 3, isto é, variância entre as proficiências finais médias de cada escola em relação à média das proficiências finais de todos os alunos.

Os valores de ρ_1 , ρ_2 e ρ_3 são valores percentuais que indicam a parte da variância total que é explicada pelo Nível 1, Nível 2 e Nível 3, respectivamente.

2.4.2) Pressupostos teóricos e a lógica proposta de Nível 1

Depois de confirmadas as hipóteses da existência da variância em todos os níveis, o próximo passo consiste na inserção de variáveis explicativas no modelo.

Dado que o objetivo central desta pesquisa é identificar os fatores que interferem na proficiência final dos alunos e considerando, pelo exposto no início deste capítulo, que a proficiência final é altamente correlacionada à proficiência inicial, uma das principais especificidades desta pesquisa é a busca da elaboração de um modelo estatístico que coloque a proficiência final em função da proficiência inicial.

Não se trata, portanto, de um modelo de curva de crescimento (*Linear Growth Model*²⁶) nos quais o coeficiente de regressão representa a taxa de crescimento em um determinado período de tempo.

Essa opção metodológica só é exequível graças à disponibilidade dos dados longitudinais e sua relevância é dada pelo fato de que, a princípio, com essa modelagem as diferenças entre os alunos ao entrarem na escola são significativamente minimizadas. Além disso, essa modelagem permite verificar se os efeitos desses fatores são os mesmos para os alunos que entram com diferentes proficiências nessas escolas, assim como as premissas apresentadas para a elaboração desse modelo, elencadas anteriormente, fossem atendidas.

Com essa inserção, a equação do Nível 1 passa a ser descrita:

$$PR_FI_{ijk} = \pi_{0jk} + \pi_{1jk} \cdot PR_IN_{ijk} + e_{ijk} \quad [04]$$

No entanto, essa modelagem provocaria um erro de estimativa demasiado grande, já que é esperado e observável que a proficiência final seja distinta e crescente a cada ano de escolarização que o aluno tenha vivenciado. Para fazer essa distinção e evitar esse erro de concepção, foram criadas variáveis indicadoras para indexar as proficiências finais e iniciais ao respectivo ano em que foram coletadas. Essa decisão metodológica, além de provocar a referida distinção, cria coeficientes na equação que permite a inserção de variáveis nos níveis superiores e, assim, a das influências dessas variáveis contemplando todo o período, mas mantendo as especificidades de cada ano.

Por último, também é inserida uma variável explicativa que controla se o aluno ao participar de uma determinada onda havia sido retido em alguma das séries anteriores. Esse cuidado é relevante, pois todo aluno que foi retido no processo, era chamado a participar do teste, independentemente dele ter ou não cursado a série que estava sendo avaliada. Assim, uma determinada medida de proficiência pode se distinguir das demais pelo fato dela ter sido coletada em um ano de escolarização na qual o aluno foi retido, sendo que nos anos anteriores esse mesmo aluno não tinha esse *status*, fazendo com que esse atributo caracterize uma proficiência e não um aluno.

26 Para um aprofundamento na temática sobre Curvas de crescimento ver Raudenbush e Bryk (2002, p. 160-204).

O Quadro 4, mostra as variáveis explicativas inseridas na Equação [04] do Nível 1, Medidas Cognitivas no Tempo:

Quadro 4 - Variáveis do Nível 1 - Ano de Escolarização

Variável	Descrição
T2	Variável indicadora que recebe o valor 1 quando o registro for do segundo ano de escolarização e zero para os demais anos.
T3	Variável indicadora que recebe o valor 1 quando o registro for do terceiro ano de escolarização e zero para os demais anos.
T4	Variável indicadora que recebe o valor 1 quando o registro for do quarto ano de escolarização e zero para os demais anos.
PR_IN	Proficiência inicial para os quatro anos de escolarização.
PR_INxT2	Variável indicadora que recebe o valor de PR_IN quando o registro for do segundo ano de escolarização e zero para os demais anos.
PR_INxT3	Variável indicadora que recebe o valor de PR_IN quando o registro for do terceiro ano de escolarização e zero para os demais anos.
PR_INxT4	Variável indicadora que recebe o valor de PR_IN quando o registro for do quarto ano de escolarização e zero para os demais anos.
RET	Variável que recebe o valor 1 quando o aluno sofreu alguma retenção no processo por isso cursa alguma série diferente daquela em que o teste foi aplicado.

As variáveis T2, T3 e T4, são variáveis indicadoras citadas anteriormente para fazer a distinção entre os quatro anos de escolarização e também permitir a interação com a variável de proficiência inicial, através do produto entre elas, produzindo outras variáveis Pr_INxT2, Pr_INxT3 e Pr_INxT4. O Quadro 5 explicita a atribuição de valores a essas variáveis em cada um dos anos de escolarização.

Quadro 5 - Valores atribuídos às variáveis indicadoras em cada ano de escolarização

Ano de escolarização	T2	T3	T4	PR_INxT2	PR_INxT3	PR_INxT4
1	0	0	0	0	0	0
2	1	0	0	PR_IN_ano2	0	0
3	0	1	0	0	PR_IN_ano3	0
4	0	0	1	0	0	PR_IN_ano4

Com a inserção dessas variáveis explicativas, o Nível 1, nível das Medidas Cognitivas no Tempo é dado por:

Nível 1 – Medidas Cognitivas no Tempo:

$$PR_FI_{ijk} = \pi_{0jk} + \pi_{1jk} \cdot T2_{ijk} + \pi_{2jk} \cdot T3_{ijk} + \pi_{3jk} \cdot T4_{ijk} + \pi_{4jk} \cdot PR_IN_{ijk} + \pi_{5jk} \cdot PR_IN_{ijk} \cdot T2 + \pi_{6jk} \cdot PR_IN_{ijk} \cdot T3 + \pi_{7jk} \cdot PR_IN_{ijk} \cdot T4 + \pi_{8jk} \cdot RET_{ijk} + e_{ijk}$$

Em que:

- PR_FI_{ijk} é a proficiência final do ano i do aluno j da escola k ;
- π_{0kj} representa a proficiência final média dos J_k alunos da escola k , que tiraram zero de proficiência inicial²⁷;
- π_{1kj} é o valor a ser adicionado a π_{0kj} para representar a proficiência final média dos J_k alunos da escola k , no segundo ano de escolarização, de maneira que essa proficiência final média é dada por $(\pi_{0kj} + \pi_{1kj})$;
- π_{2kj} e π_{3kj} são os valores a serem adicionados a π_{0kj} para representar a proficiência final média dos J_k alunos da escola k no terceiro e quarto ano, respectivamente. De forma análoga ao que ocorre no segundo ano, a soma de cada um desses coeficientes a π_{0kj} , definem os interceptos da reta relativa ao segundo e terceiro ano de escolarização e são dados por $(\pi_{0kj} + \pi_{2kj})$ e $(\pi_{0kj} + \pi_{3kj})$;
- PR_IN_{ijk} é proficiência inicial no ano de escolarização i do aluno j , da escola k ;
- RET_{ijk} é variável que indica se o aluno j da escola k fez a prova na condição de retido no ano de escolarização i ;
- π_{4kj} é o incremento para cada unidade da proficiência inicial (PR_IN_{ijk}) na proficiência final do aluno j , da escola k , no primeiro ano de escolarização (PR_FI_{ijk});
- π_{5kj} é o acréscimo médio a ser feito em π_{4kj} para se estimar o incremento para cada unidade da proficiência inicial na proficiência final do aluno j , da escola k , no segundo ano. Essa proficiência média é dada por $(\pi_{4kj} + \pi_{5kj})$;
- π_{6kj} e π_{7kj} são os acréscimos médios a serem feitos em π_{4kj} para cada unidade da proficiência inicial (PR_IN_{ijk}) na proficiência final do aluno j , da escola k , no terceiro e quarto ano, matematicamente dados por $(\pi_{4kj} + \pi_{6kj})$ e $(\pi_{4kj} + \pi_{7kj})$, respectivamente.
- π_{8kj} é o valor a ser adicionado ao intercepto de cada ano, caso o aluno tenha participado da onda de teste sem defasagem, considerando os diferentes anos de escolarização;
- e_{ijk} é o desvio entre o valor estimado para a proficiência final do aluno j matriculado na escola k no ano de escolarização i e o valor observado (PR_FI_{ijk});

27 Este é um valor sem significado pedagógico uma vez que não há alunos com proficiência inicial nula. É, portanto, um valor de referência inerente ao processo de regressão linear necessário para outras estimativas. Mais esclarecimentos sobre esse aspecto é dado a seguir.

- σ^2 variância de e_{ijk} .

No Nível 2, nível do modelo hierárquico no qual as variáveis explicativas associadas ao perfil do alunos poderão ser inseridas, são criadas uma equação para cada coeficiente existente na equação do Nível 1, mas ainda sem a inserção de cada variável explicativa é dado por:

Nível 2 – Alunos:

$$\mu_{jk} = \beta_{00k} + \gamma_{0jk};$$

$$\pi_{1jk} = \beta_{10k};$$

$$\pi_{2jk} = \beta_{20k};$$

$$\pi_{3jk} = \beta_{30k};$$

$$\pi_{4jk} = \beta_{40k} + \gamma_{4jk};$$

$$\pi_{5jk} = \beta_{50k};$$

$$\pi_{6jk} = \beta_{60k};$$

$$\pi_{7jk} = \beta_{70k};$$

$$\pi_{8jk} = \beta_{80k};$$

Em que:

- β_{00k} é a média das proficiências finais de todos os alunos da escola k no primeiro ano de escolarização;
- β_{10k} , β_{20k} e β_{30k} são os valores a serem adicionados a β_{00k} para compor a média das proficiências finais de todos os alunos da escola k , referentes ao segundo, terceiro e quarto anos, respectivamente. Estes são valores que são adicionados para ajustar o intercepto dos demais anos de escolarização em relação ao primeiro;
- β_{40k} é o incremento médio na média da proficiência final para cada unidade da proficiência inicial de todos os alunos da escola k no primeiro ano de escolarização;
- β_{50k} , β_{60k} e β_{70k} são os valores a serem adicionados a β_{40k} para compor os incrementos (para cada unidade da proficiência inicial considerando todos os alunos da escola k) na proficiência média final dos demais anos de escolarização. Dessa maneira, o incremento no primeiro ano de escolarização é dado por β_{40k} enquanto que para o segundo ano, o incremento no segundo ano é dado por $(\beta_{40k} + \beta_{50k})$. Para o terceiro e quarto anos, os incrementos são dados, respectivamente, por $(\beta_{40k} + \beta_{60k})$ e $(\beta_{40k} + \beta_{70k})$;

- β_{80k} é o valor a ser adicionado ao intercepto de cada ano, caso o aluno tenha participado da onda de teste sem defasagem, considerando todos os alunos da escola k ;
- r_{0jk} é o desvio entre a média das proficiências finais do aluno j da escola k e a média da proficiências finais de todos os alunos dessa escola;
- r_{4jk} é o desvio entre a média dos valores a serem adicionados a β_{40k} dos alunos de cada uma das escolas e o valor médio da escola onde esse aluno está matriculado;
- $\tau_{\pi 00}$ é a variância do erro de estimativa r_{0jk} ;
- $\tau_{\pi 40}$ é a variância do erro de estimativa r_{4jk} .

Como nessa modelagem estão previstas a existência da variabilidade apenas nos coeficientes β_{00k} e β_{40k} é assumido que:

- β_{00k} e β_{40k} podem ser diferentes entre os alunos, isto é, o intercepto (média das proficiências finais para os alunos com proficiência inicial igual a zero) e o coeficiente de regressão (incremento na proficiência final para cada unidade de proficiência inicial) podem ser diferentes;
- os incrementos a β_{00k} (β_{10k} , β_{20k} , β_{30k} e β_{80k}) e β_{40k} (β_{50k} , β_{60k} e β_{70k}) referentes aos demais anos de escolarização e a medida do aluno ser retido em um determinado ano são iguais entre os diferentes alunos.

Sabe-se que na verdade, esses incrementos podem ser diferentes, mas esse controle será feito com a inserção de variáveis explicativas.

Nível 3 – Escolas

$$\beta_{00k} = \gamma_{000} + u_{0k};$$

$$\beta_{10k} = \gamma_{100};$$

$$\beta_{20k} = \gamma_{200};$$

$$\beta_{30k} = \gamma_{300};$$

$$\beta_{40k} = \gamma_{400} + u_{4k};$$

$$\beta_{50k} = \gamma_{500};$$

$$\beta_{60k} = \gamma_{600};$$

$$\beta_{70k} = \gamma_{700};$$

$$\beta_{80k} = \gamma_{800}.$$

Em que:

- γ_{000} é a média de todas as proficiências finais de todas as escolas no primeiro ano de escolarização;
- γ_{100} , γ_{200} e γ_{300} são os valores a serem adicionados a γ_{000} para compor a proficiência final média de todas as escolas nos demais anos de escolarização;
- γ_{400} é o incremento médio na proficiência final média para cada unidade da proficiência inicial de todas as escolas k no primeiro ano de escolarização;
- γ_{500} , γ_{600} e γ_{700} são os valores a serem adicionados a γ_{400} para compor os incrementos na média das proficiências finais, nos demais anos de escolarização, para cada unidade da proficiência inicial considerando todas as escolas;
- γ_{800} é estimativa do valor a serem adicionado ao intercepto de cada ano, caso o aluno tenha participado da onda de teste sem defasagem, considerando todas as escolas;
- u_{00k} é o desvio entre média das proficiências finais estimadas de cada escola em relação à média de todas as escolas;
- u_{40k} é o desvio entre a média dos valores estimados para serem adicionados a γ_{400} e a média desses valores de todas as escolas;
- $\tau_{\beta 000}$ é a variância de u_{00k} ;
- $\tau_{\beta 400}$ é a variância de u_{40k} .

De maneira análoga ao que foi assumido no Nível 2, no Nível 3 também é assumido a existência da variabilidade apenas nos coeficientes γ_{000} e γ_{400} . Assim deve ser considerado que:

- γ_{000} e γ_{400} podem ser diferentes entre as escolas, isto é, o intercepto (média das proficiências finais das escolas, dentre os alunos com proficiência inicial igual a zero) e o coeficiente de regressão (incremento na proficiência final para cada unidade de proficiência inicial) podem ser diferentes;
- os incrementos aos coeficientes γ_{000} (γ_{100} , γ_{200} , γ_{300} e γ_{800}) e γ_{400} (γ_{500} , γ_{600} e γ_{700}) referentes aos demais anos de escolarização são iguais entre as diferentes escolas.

E igualmente ao que pode ocorrer no nível dos alunos, a inserção de variáveis explicativas referentes às características específicas de cada escola permitira a distinção entre elas nesse modelo.

A opção metodológica em manter os efeitos aleatórios apenas nos coeficientes do primeiro ano de escolarização elimina a sobreposição de efeitos aleatórios sobre outros efeitos aleatórios, uma vez que os coeficientes para intercepto e inclinação dos diferentes anos são somados aos do primeiro ano, o que dificultaria a convergência na estimação dos coeficientes e também interpretação dos resultados.

Como consequência, as declividades das retas não são iguais, tampouco são iguais os valores dos interceptos, cujos valores numéricos não têm qualquer significado, uma vez que, como alertado anteriormente, não há alunos com proficiência inicial igual a zero (valor nulo este que define o intercepto), como mostra os intervalos de proficiência inicial dos diferentes anos de escolarização indicados na Figura 7.

A interação entre as variáveis indicadoras com a proficiência inicial de cada um dos anos, apresentada no Quadro 5, faz com que os coeficientes π_{0jk} e π_{4jk} (intercepto e coeficiente de regressão, respectivamente) sejam os coeficientes da reta para o primeiro ano de escolarização e os demais coeficientes, π_{1jk} e π_{5jk} , π_{2jk} e π_{6jk} , π_{3jk} e π_{7jk} , sejam incrementos referentes a cada um dos anos subsequentes.

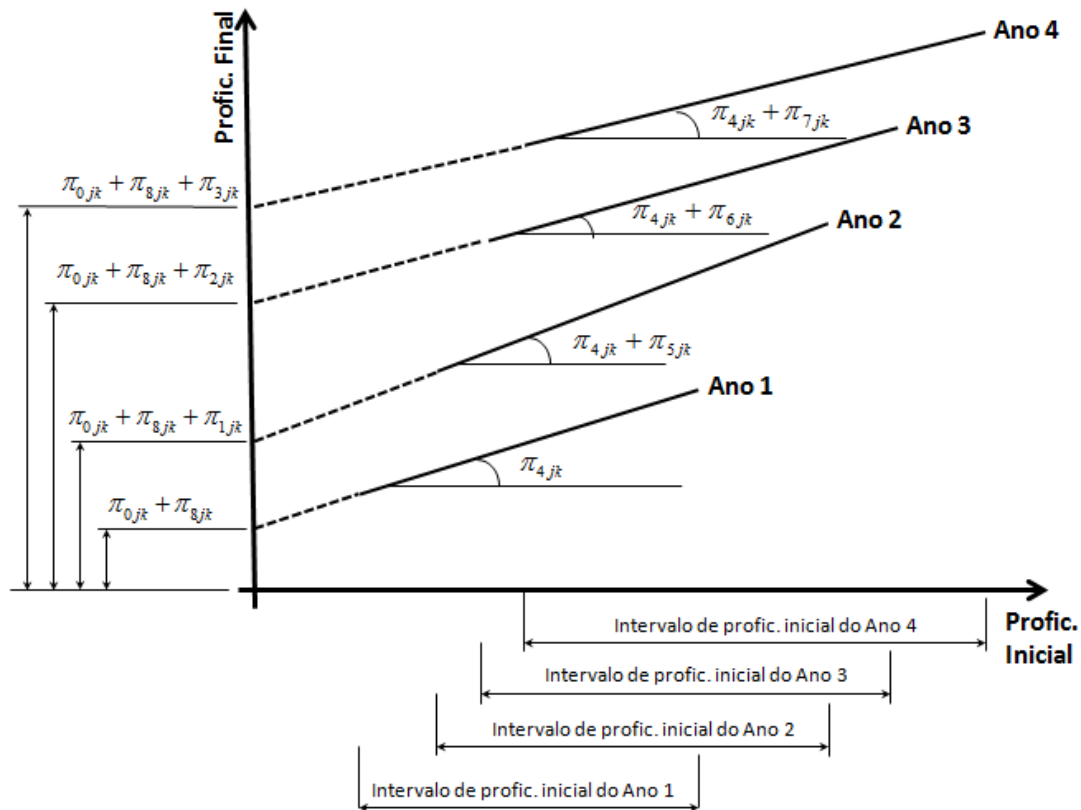
O Quadro 6 apresenta esses coeficientes em cada um dos anos de escolarização usados para a representação gráfica da Figura 7.

Quadro 6 - Valores atribuídos aos coeficientes da reta de cada um dos anos de escolarização

Ano de escolarização	T2	T3	T4	T2*LeitI	T3*LeitI	T4*LeitI	Intercepto da reta	Coefficiente de Inclinação
1	0	0	0	0	0	0	π_{0jk}	π_{4jk}
2	1	0	0	LeitI	0	0	$\pi_{0jk} + \pi_{1jk}$	$\pi_{4jk} + \pi_{5jk}$
3	0	1	0	0	LeitI	0	$\pi_{0jk} + \pi_{2jk}$	$\pi_{4jk} + \pi_{6jk}$
4	0	0	1	0	0	LeitI	$\pi_{0jk} + \pi_{3jk}$	$\pi_{4jk} + \pi_{7jk}$

A Figura 7 traz uma representação esquemática das retas que representam cada um dos anos de escolarização e indica significado de cada um dos coeficientes estimados pelo Modelo de Nível 1. Nesta figura, o coeficiente β_{80k} não foi representado para a simplificação do esquema.

Figura 7 - Representação esquemática das retas para cada ano de escolarização obtidas no Modelo de Nível 1.



Com essa modelagem é possível verificar as seguintes hipóteses:

- Hipótese 4: As medidas da proficiência dos alunos em função dos anos de escolarização podem ser representadas linearmente.
- Hipótese 5: Há variabilidade significativa nas proficiências finais dos alunos nos diferentes anos de escolarização, sendo possível determinar o quanto desta variabilidade foi explicada pelas variáveis controladas.

Dessa maneira, fica constituído um modelo denominado Modelo de Nível 1, pois nele se encontram os coeficientes, cujas estimativas permitem identificar as variáveis explicativas cujos efeitos que são significantes, levando ao cumprimento do objetivo da pesquisa, uma vez que, atende a expectativa para a qual foi concebido, permite identificar os coeficientes da reta de cada um dos anos de escolarização, ou seja, permite identificar quais variáveis explicativas são significantes em cada um dos anos de escolarização, sejam elas associadas ao perfil dos alunos ou às características das escolas.

2.4.3) Interpretação para o Modelo de Nível 1

Para exemplificar as interpretações do Modelo de Nível 1, encontra-se no Quadro 7 estimativas hipotéticas para os coeficientes associados apenas às variáveis explicativas do Nível 1 – Medidas Cognitivas no Tempo.

Quadro 7 - Estimativas hipotéticas para exemplo do Modelo de Nível 1

Coeficiente	Estimativas*
Fixos	
Intercepto Ano 1, γ_{000} ($= \beta_{00k} = \pi_{0jk}$)	+60
Incremento ao intercepto Ano 2, γ_{100} ($= \beta_{10k} = \pi_{1jk}$)	-32
Incremento ao intercepto Ano 3, γ_{200} ($= \beta_{20k} = \pi_{2jk}$)	-7
Incremento ao intercepto Ano 4, γ_{300} ($= \beta_{30k} = \pi_{3jk}$)	+25
Coefficiente de regressão Ano 1, γ_{400} ($= \beta_{40k} = \pi_{4jk}$)	+0,3
Incremento ao coeficiente de regressão Ano 2, γ_{500} ($= \beta_{50k} = \pi_{5jk}$)	+0,6
Incremento ao coeficiente de regressão Ano 3, γ_{600} ($= \beta_{60k} = \pi_{6jk}$)	+0,5
Incremento ao coeficiente de regressão Ano 4, γ_{700} ($= \beta_{70k} = \pi_{7jk}$)	+0,4
Incremento ao intercepto-Retido, γ_{800} ($= \beta_{80k} = \pi_{8jk}$)	+6
Aleatórios	
σ^2 (variância de e_{ijk} – Nível 1)	185
$\tau_{\pi 00}$ (variância de r_{0jk} – Nível 2)	110
$\tau_{\pi 40}$ (variância de r_{4jk} – Nível 2)	0,03
$\tau_{\beta 000}$ (variância de u_{00k} – Nível 3)	250
$\tau_{\beta 400}$ (variância de u_{40k} – Nível 3)	0,05

* Como se trata apenas de um exemplo, os erros-padrão das estimativas não foram citados.

Com essas estimativas, a equação do Nível 1 para esses dados hipotéticos seria:

$$\widehat{PR_FI}_{ijk} = 60 - 32.T2_{ijk} - 7.T3_{ijk} + 25.T4_{ijk} + 0,3.PR_IN_{ijk} + 0,6.PR_IN_{ijk} \times T2 + 0,5.PR_IN_{ijk} \times T3 + 0,4.PR_IN_{ijk} \times T4 + 6.RET_{ijk} + e_{ijk}.$$

Assim, para o primeiro ano de escolarização, quando o valor zero é atribuído às variáveis indicadoras T2, T3, T4, PR_IN_{ijk} × T2, PR_IN_{ijk} × T3 e PR_IN_{ijk} × T4, a equação fica:

$$\widehat{PR_FI}_{ijk} = 60 + 0,3.PR_IN_{ijk} + 6.RET_{ijk} + e_{ijk}.$$

Com essa equação, nesse primeiro ano de escolarização, pode-se interpretar que:

- a proficiência final média esperada para os alunos que entram na escola com proficiência inicial nula e que não foram retidos²⁸ é 66 (60 + 6). Esse valor, como já especificado anteriormente, não tem qualquer significado prático, já que essa possibilidade da proficiência nula não existe;

28 Quanto o aluno é retido nesse ano de escolarização é atribuído o valor 0 na variável RET_{ijk}, caso contrário, é atribuído a ela, o valor 1.

- para cada unidade de proficiência inicial, a proficiência inicial deve ser aumentada em 0,5 unidades. Logo, a proficiência final média esperada dos alunos com 100 unidades de proficiência inicial e que não foram retidos será de 96 (60 + 30 + 6);

Para o segundo ano de escolarização, o valor zero é atribuído às variáveis indicadoras T3, T4, PR_IN_{ijk}xT3 e PR_IN_{ijk}xT4, enquanto à variável T2 é atribuído o valor 1 e à variável PR_IN_{ijk}xT2 é atribuído o valor PR_IN_{ijk}. Com essas atribuições, as estimativas dos coeficientes π_{0jk} e π_{1jk} (60 e -32) são somados, sendo que o mesmo ocorre com as estimativas dos coeficientes π_{4jk} e π_{5jk} (0,3 e 0,6). Consequentemente, a equação para esse segundo ano de escolarização resulta em:

$$\widehat{PR_FI}_{ijk} = 28 + 0,9 \cdot PR_IN_{ijk} + 6 \cdot RET_{ijk} + e_{ijk}$$

Com esse equacionamento, de maneira análogo ao primeiro ano de escolarização, as seguintes interpretações podem ser feitas:

- a proficiência final média esperada para os alunos que entram na escola com proficiência inicial nula e não retidos é 34 (28 + 6).

- para cada unidade de proficiência inicial, a proficiência inicial deve ser aumentada em 0,9 unidades. Logo, a proficiência final média esperada dos alunos com 120 unidades de proficiência inicial e que não foram retidos nesse ano de escolarização será de 142 (28 + 108 + 6);

Também deve ser salientado que é esperado que as proficiência iniciais médias nos diferentes anos de escolarização sejam diferentes.

Para o terceiro ano de escolarização, o valor zero é atribuído às variáveis indicadoras T2, T4, PR_IN_{ijk}xT2 e PR_IN_{ijk}xT4, enquanto à variável T3 é atribuído o valor 1 e à variável PR_IN_{ijk}xT3 é atribuído o valor PR_IN_{ijk}. Com essas atribuições, as estimativas dos coeficientes π_{0jk} e π_{2jk} (60 e -7) são somados, sendo que o mesmo ocorre com as estimativas dos coeficientes π_{4jk} e π_{6jk} (0,3 e 0,5). Consequentemente, a equação para esse terceiro ano de escolarização resulta em:

$$\widehat{PR_FI}_{ijk} = 53 + 0,8 \cdot PR_IN_{ijk} + 6 \cdot RET_{ijk} + e_{ijk}$$

Com esse equacionamento, podem ser feitas as seguintes interpretações para o terceiro ano de escolarização:

- a proficiência final média esperada para os alunos que entram na escola com proficiência inicial nula e não retidos é 59 (53 + 6);

- para cada unidade de proficiência inicial, a proficiência inicial deve ser aumentada em 0,8 unidades. Assim, a proficiência final média esperada dos alunos com proficiência inicial de 140 unidades e que não foram retidos nesse ano de escolarização será de 171 (53 + 112 + 6).

Para o quarto ano de escolarização, o valor zero é atribuído às variáveis indicadoras T2, T3, PR_IN_{ijk}xT2 e PR_IN_{ijk}xT3, enquanto à variável T4 é atribuído o valor 1 e à variável PR_IN_{ijk}xT4 é atribuído o valor PR_IN_{ijk}. Assim, são somadas as estimativas dos coeficientes π_{0jk} e π_{3jk} (60 e 25) e também as estimativas dos coeficientes π_{4jk} e π_{7jk} (0,3 e 0,4). Consequentemente, a equação para esse segundo ano de escolarização resulta em:

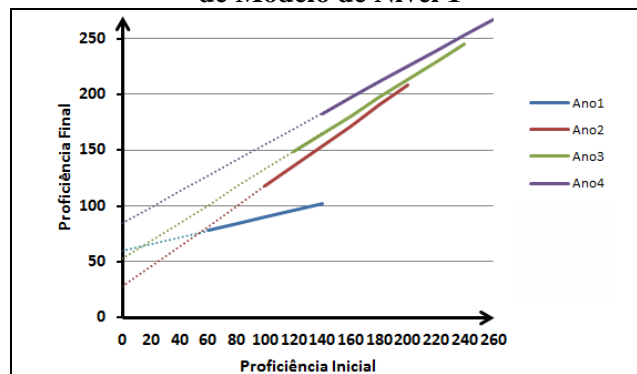
$$\widehat{PR_FI}_{ijk} = 85 + 0,7 \cdot PR_IN_{ijk} + 6 \cdot RET_{ijk} + e_{ijk}$$

As interpretações para o quarto ano de escolarização são:

- a proficiência final média esperada para os alunos que entram na escola com proficiência inicial nula e não retidos é 91 (85 + 6);
- para cada unidade de proficiência inicial, a proficiência inicial deve ser aumentada em 0,8 unidades. Assim, a proficiência final média esperada dos alunos com proficiência inicial de 160 unidades e que não foram retidos nesse ano de escolarização será de 203 (85 + 112 + 6).

O Gráfico 8 representa essas diferentes equações dos quatro anos de escolarização:

Gráfico 8 - Representação gráfica do exemplo de Modelo de Nível 1



Para esta representação gráfica, foram adotadas hipoteticamente os seguintes intervalos de proficiência inicial:

- para o primeiro ano de escolarização, de 60 a 140;
- para o segundo ano de escolarização, de 100 a 180;
- para o terceiro ano de escolarização, de 140 a 220;

d) para o quarto ano de escolarização, de 180 a 260;

Um alerta importante a ser dado com relação a esse modelo, e que, diferentemente de outros modelos lineares hierárquicos, assim como os de regressão simples, suas interpretações não devem ser feitas diretamente a partir dos valores estimados aos coeficientes. Uma situação notória neste exemplo é que, apesar das estimativas para os interceptos das equações referentes ao segundo e terceiro anos de escolarização serem iguais a -32 e -7, não implica que as proficiências médias esperadas para esses anos sejam menores que as do primeiro ano. Isso ocorre porque, nos diferentes anos, as proficiências iniciais e os coeficientes de regressão são diferentes. Neste exemplo, uma constatação desse fato é que para uma proficiência inicial de 160 unidades, implica uma proficiência final igual a 172,181 e 197, para o segundo, terceiro e quarto anos, respectivamente. Na verdade, essa discrepância ocorreria para a proficiência inicial nula, o que de fato, não ocorre.

2.4.4) Modelo de Nível 1

Com essa modelagem podem ser estimados os coeficientes da reta que melhor se ajusta a cada um dos quatro anos de escolarização. Genericamente, o Modelo de Nível 1 é definido por:

Nível 1 – Medidas Cognitivas no Tempo:

$$PR_FI_{ijk} = \pi_{0jk} + \pi_{1jk} \cdot T2_{ijk} + \pi_{2jk} \cdot T3_{ijk} + \pi_{3jk} \cdot T4_{ijk} + \pi_{4jk} \cdot PR_IN_{ijk} + \pi_{5jk} \cdot PR_IN_{ijk} \times T2 + \pi_{6jk} \cdot PR_IN_{ijk} \times T3 + \pi_{7jk} \cdot PR_IN_{ijk} \times T4 + \pi_{8jk} \cdot RET_{ijk} + e_{ijk}, \text{ com } e_{ijk} \sim NID(0, \sigma^2).$$

Nível 2 – Alunos:

$$\begin{aligned} \pi_{0jk} &= \beta_{00k} + r_{0jk}, \text{ com } r_{0jk} \sim NID(0, \tau_{\pi 00}) \\ \pi_{4jk} &= \beta_{40k} + r_{4jk}, \text{ com } r_{4jk} \sim NID(0, \tau_{\pi 40}) \\ \pi_{tjk} &= \beta_{t0k}, \text{ com } t = 1, 2, 3, 5, 6, 7 \text{ ou } 8 \end{aligned}$$

Nível 3 – Escolas:

$$\begin{aligned} \beta_{00k} &= \gamma_{000} + u_{0jk}, \text{ com } u_{0jk} \sim NID(0, \tau_{\beta 000}) \\ \beta_{40k} &= \gamma_{400} + u_{4jk}, \text{ com } u_{4jk} \sim NID(0, \tau_{\beta 400}) \\ \beta_{t0k} &= \gamma_{t00}, \text{ com } t = 1, 2, 3, 5, 6, 7 \text{ ou } 8 \end{aligned}$$

Essa modelagem permite a definição de uma equação da regressão para cada ano de escolarização²⁹, isto é, podem ser estimados os coeficientes da reta que melhor se ajusta a cada um dos quatro anos de escolarização. Esse ajuste é feito com o uso do teste da significância de cada uma das diversas variáveis explicativas de aluno e da escola, separadamente em cada ano. Dessa forma, os coeficientes gerados por essas variáveis indicadoras permitem a estimação do intercepto e da inclinação da proficiência final, em função da proficiência inicial, separadamente para cada ano de escolarização, tomando como base o primeiro ano de escolarização³⁰.

Dos resultados estimados por esses valores, é importante destacar que quando os *p-value* associados a essas variâncias forem menor que 0,05, indicam a possibilidade da inserção de variáveis explicativas, de forma que:

- σ^2 há variáveis do Nível 1 que ainda podem ser inseridas, apesar de todas disponíveis já estarem inseridas;
- $\tau_{\pi00}$ há variáveis do Nível 2, associadas ao perfil do aluno, que ainda podem ser inseridas e interferem no intercepto da reta de cada ano de escolarização;
- $\tau_{\pi40}$ há variáveis do Nível 2, associadas ao perfil do aluno, que ainda podem ser inseridas e interferem na inclinação da reta de cada ano de escolarização;
- $\tau_{\beta000}$ há variáveis do Nível 3, associadas ao perfil do aluno, que ainda podem ser inseridas e interferem no intercepto da reta de cada ano de escolarização;
- $\tau_{\beta400}$ há variáveis do Nível 3, associadas ao perfil do aluno, que ainda podem ser inseridas e interferem na inclinação da reta de cada ano de escolarização.

Evidentemente essa diversidade de possibilidades de inserção das variáveis conferem um significativo grau de complexidade ao modelo, atendendo assim à expectativa inicial deste projeto de pesquisa, sobretudo, por considerar a quantidade de variáveis associadas ao Nível 2 e 3, especificados a seguir.

Outro aspecto a ser enfatizado é que, apesar de inúmeras pesquisas apontarem que o nível socioeconômico do aluno e da escola, as variáveis que controla essa duas características foram inseridas como as demais para, justamente, verificar o seu efeito na interação com as

29 Resumidamente, os diferentes interceptos são definidos pelos coeficientes π_{0jk} , π_{1jk} , π_{2jk} , π_{3jk} e π_{8jk} e a inclinação da reta é definida pelos coeficientes π_{4jk} , π_{5jk} , π_{5jk} e π_{7jk} .

30 Qualquer um dos anos poderia ser selecionado para ser o ano base, mas optou-se pelo primeiro ano para maior facilidade de interpretação.

demais variáveis disponíveis do contexto do aluno e da escola. Essa verificação é relevante, dada à também alta correlação existente entre a proficiência inicial e o nível socioeconômico.

2.4.5) Modelo final usado na pesquisa

Com a inserção das variáveis explicativas associadas ao Nível 2 e Nível 3, o modelo teórico, revestido do significado pertinente a esta pesquisa, na qual as medidas das proficiências obtidas numa sequência temporal são agrupadas por alunos e estes o são por escolas, em que é composto pelas Equações [05], [06] e [07] correspondentes a cada um dos níveis hierárquicos, nas quais i ($i = 1, 2, 3$ ou 4 anos) indica o ano de escolarização ao qual as medidas de proficiência estão associadas e aninhada em cada dos j alunos ($j = 1, \dots, J_k$ alunos) das k escolas ($k = 1, \dots, K$ escolas).

De maneira genérica, tendo usando a simbologia proposta por Raudenbush e Bryk (2002, p. 231- 233), o modelo teórico genérico é dado por:

Nível 1 – Medidas Cognitivas no Tempo

$$Y_{ijk} = \pi_{0jk} + \sum_{p=1}^P \pi_{pjk} X_{pijk} + e_{ijk} \quad [05]$$

Em que:

- Y_{ijk} é a proficiência final do i -ésimo ano de escolarização, pelo j -ésimo aluno, pertencente a k -ésima escola;
- π_{0jk} é o intercepto do Nível 1;
- X_{pijk} representa a p -ésima variável explicativa associada ao Nível 1 do i -ésimo ano de escolarização, do j -ésimo aluno da escola k ;
- π_{pjk} ($p = 1, \dots, P$) são os coeficientes de regressão do Nível 1 associadas às P variáveis explicativas desse nível;
- e_{ijk} é a componente aleatória do Nível 1. Assume-se que o termo aleatório $e_{ijk} \sim \text{NID}(0, \sigma^2)$.

Nível 2 - Alunos)

$$\pi_{pjk} = \beta_{p0k} + \sum_{q=1}^{Q_p} \beta_{pqk} Z_{qjk} + r_{pjk} \quad [06]$$

Em que:

- β_{p0k} é o intercepto para o aluno j associado ao Nível 2;

- β_{pqk} ($q = 1 \dots, Q_p$) são os coeficientes de regressão do Nível 2;
- Z_{qjk} representa a q-ésima variável explicativa associada ao Nível 2 do j-ésimo aluno da escola k;
- r_{pjk} é a componente aleatória do Nível 2. Assume-se que para cada unidade j, o vetor $(r_{0jk}, r_{1jk}, \dots, r_{pjk})$ apresenta uma distribuição normal multivariada, em que cada elemento r_{pjk} tenha média zero e variância τ_{pp} . Para cada par de efeitos aleatórios p e p' $Cov(r_{pjk}, r_{p'jk}) = \tau_{pp'}$.

Nível 3 - Escolas)

$$\beta_{pqk} = \gamma_{pq0} + \sum_{s=1}^{S_{pq}} \gamma_{pqs} W_{sk} + u_{pqk} \quad [07]$$

Em que:

- γ_{pq0} é o intercepto do Nível 3;
- γ_{pqs} ($s = 1, \dots, S_{pq}$) são os coeficientes de regressão do Nível 3;
- W_{sk} representa a s-ésima variável explicativa associada ao Nível 3 da escola k;
- u_{pqk} é a componente aleatória do Nível 3. Assume-se que para cada variável de Nível 3, o vetor de efeitos aleatórios (os termos u_{pqk}) apresenta distribuição normal multivariada, com média zero e matriz de covariância T_β , cuja dimensão máxima depende do número de coeficientes especificados com os efeitos aleatórios e é de: $\sum_{p=0}^p (Q_p + 1) \times \sum_{p=0}^p (Q_p + 1)$.

Considerando as variáveis explicativas disponível nos Níveis 2 e 3, descritas anteriormente, o Modelo Final aplicado nesta pesquisa é:

Nível 1 – Medidas Cognitivas no Tempo)

$$PR_FI_{ijk} = \pi_{0jk} + \pi_{1jk} \cdot T2_{ijk} + \pi_{2jk} \cdot T3_{ijk} + \pi_{3jk} \cdot T4_{ijk} + \pi_{4jk} \cdot PR_IN_{ijk} + \pi_{5jk} \cdot PR_IN_{ijk} \times T2 + \pi_{6jk} \cdot PR_IN_{ijk} \times T3 + \pi_{7jk} \cdot PR_IN_{ijk} \times T4 + \pi_{8jk} \cdot RET_{ijk} + e_{ijk}$$

Nível 2 – Aluno

$$\pi_{0jk} = \beta_{00k} + \sum_{q=1}^{11} \beta_{0qk} \cdot FA_{qjk} + r_{0jk}, \text{ com } r_{0jk} \sim NID(0, \tau_{\pi00})$$

$$\pi_{tjk} = \beta_{t0k} + \sum_{q=1}^{11} \beta_{tqk} \cdot FA_{qjk}, \text{ com } t = 1, 2 \text{ e } 3$$

$$\pi_{4jk} = \beta_{40k} + \sum_{q=1}^{11} \beta_{4qk} \cdot FA_{qjk} + r_{4jk}, \text{ com } r_{4jk} \sim NID(0, \tau_{\pi40})$$

$$\pi_{tjk} = \beta_{t0k} + \sum_{q=1}^{11} \beta_{tqk} \cdot FA_{qjk}, \text{ com } t = 5, 6, 7 \text{ e } 8$$

Nível 3 – Escola

$$\beta_{00k} = \gamma_{000} + \sum_{s=1}^{59} \gamma_{00s} \cdot FE_s + u_{00k}, \text{ com } u_{00k} \sim \text{NID}(0, \tau_{000})$$

$$\beta_{v0k} = \gamma_{v00} + \sum_{s=1}^{59} \gamma_{v0s} \cdot FE_s, \text{ com } 1 \leq v \leq 3,$$

$$\beta_{40k} = \gamma_{400} + \sum_{s=1}^{59} \gamma_{40s} \cdot FE_s + u_{40k}, \text{ com } u_{40k} \sim \text{NID}(0, \tau_{400})$$

$$\beta_{v0k} = \gamma_{v00} + \sum_{s=1}^{59} \gamma_{v0s} \cdot FE_s, \text{ com } 5 \leq v \leq 8,$$

$$\beta_{v0k} = \gamma_{vz0}, \text{ com } 0 \leq v \leq 8 \text{ e com } 1 \leq z^* \leq 11,$$

Em que ,

- FA_{qjk} são as 11 variáveis explicativas do Nível 2, fatores associados ao perfil dos alunos;
- FE_s são as 59 variáveis explicativas do Nível 3, fatores associados às características das escolas;
- z^* indica a posição das variáveis significantes no Nível 2.

Essas variáveis explicativas estão disponível para a inserção no modelo, no entanto, aquelas que não são significantes ou que têm efeito desprezível não são mantidas no modelo.

Todas as variáveis explicativas consideradas significantes, nos diferentes níveis hierárquicos e seus respectivos anos de escolarização, podem ser consideradas, segundo esse modelo, como sendo fatores que interferem significativamente na proficiência final dos respectivos anos de escolarização.

**CAPÍTULO 3 -
A PREPARAÇÃO DAS VARIÁVEIS E DOS ARQUIVOS**

O objetivo deste capítulo é apresentar os procedimentos usados no tratamento dos dados das variáveis e na preparação do banco de dados, composto por três arquivos, um para cada nível do Modelo Linear Hierárquico usado nesta pesquisa.

Essa apresentação é feita em duas partes, sendo que a primeira consiste em uma síntese do tratamento de dados das variáveis explicativas feitas a partir dos dados fornecidos pelo Projeto GERES. Os procedimentos usados nesse tratamento consistem basicamente na aplicação da Análise Fatorial Exploratória como objetivo de reduzir o número de variáveis explicativas a serem usadas no modelo.

Na segunda parte são apresentadas as etapas de preparação³¹ dos arquivos utilizados para o ajuste dos modelos lineares hierárquicos, que têm como origem quatro diferentes arquivos fornecidos pelo Projeto GERES.

As informações de cada um desses arquivos, inicialmente separadas por aluno, professor, diretor e escola, passaram por um tratamento (análise descritiva, seleção, cálculos). Em seguida, esses arquivos foram reestruturados e resultaram em três arquivos, um para cada nível hierárquico do modelo.

3.1) Tratamento dos dados

O projeto GERES, como já apresentado no Capítulo 1, é um estudo longitudinal de painel, que mediu a proficiência em Leitura e Matemática de 35538 alunos, em cinco ondas de aplicação de testes padronizados nos primeiros anos de escolarização, (hoje equivalente ao período do 2º ao 5º anos, antes denominados de primeira a quarta séries), quando foram também coletados, através de observação e de questionários, dados sobre os alunos, seus familiares, professores, diretores e da escola.

Os instrumentos contextuais estão descritos no Quadro 8 onde são apresentadas as informações sobre as escolas, salas de aula e famílias que foram obtidas a partir de questionários aplicados a diretores, professores, pais e alunos, além de instrumento de observação preenchido pelo supervisor da aplicação dos testes em visitas à escola. O

31 Para essa preparação foi usado o programa estatístico PASW, versão 17.

questionário dos pais e a primeira versão do questionário dos alunos, aplicado na 4ª onda, serviram para identificar o nível educacional, profissão dos pais e os objetos de consumo presentes na casa do aluno. Este arquivo não foi usado neste trabalho, pois buscou informações que permitiram o cálculo do nível socioeconômico, cujo resultado já se encontra compilado na variável “nível socioeconômico do aluno” e consta da base de dados Geres.

Quadro 8 -Informações sobre os questionários aplicados.

Fonte	Descrição
Escola	Aplicado em Março de 2005. Informações básicas sobre a infraestrutura da escola.
Diretores	Aplicado em Março de 2005 e em novembro de 2007. Informações sobre sua experiência e formação. Sobre características do financiamento da escola (anuidades, bolsas de estudo etc.) em caso de escola privada. Sobre processos de nomeação do diretor (escolas públicas). Sobre relacionamento com comunidade. Critérios de formação de turmas. Reaplicado em 2008.
Professores	Aplicado em março de 2005 e em todos os anos da pesquisa em novembro. Previsão do professor sobre desempenho da turma nos testes. Percepção do estilo e liderança do diretor. Práticas de sala de aula. Formação. Práticas culturais.
Pais	Aplicado entre o segundo semestre de 2005 e o segundo semestre de 2007. Caracterização sociodemográfica da família. Aplicados aos pais de alunos novos incorporados à pesquisa a partir de 2007.
Alunos	Aplicado em novembro de 2007 (versão reduzida) e em novembro de 2008 (versão ampliada). Caracterização sociodemográfica da família.

Fonte: FRANCO, BROOKE e ALVES (2008)

Considerando-se que todo fenômeno educativo é complexo, se faz necessária a investigação do maior número possível de variáveis para que se possa estabelecer uma relação minimamente adequada entre os dados contextuais e os dados cognitivos medidos por meio da proficiência do aluno. Porém, ao mesmo tempo, essa grande quantidade de variáveis traz dificuldades de processamento e interpretativas e, por isso, é necessária a redução dessas informações, preservando-se sua capacidade explicativa.

Neste estudo, por ser uma pesquisa longitudinal com um desenho específico, seus dados permitem o cálculo do valor agregado no nível do aluno, o que favorece estudos que outras avaliações transversais não permitem, implicando em uma quantidade de informações ainda maiores.

Essa redução foi conseguida através de recursos da estatística descritiva e da análise fatorial exploratória (AFE) cujos resultados estão descritos a seguir, sendo que, com a primeira buscou-se o conhecimento geral de cada uma das variáveis (frequência, média,

desvio-padrão) e com a segunda buscou-se cumprir o objetivo último deste trabalho, que é substituir o conjunto inicial de variáveis por outro com uma menor quantidade de variáveis, porém, *sem perder capacidade explicativa*. As variáveis resultantes desse processo serão usadas na análise dessas variáveis envolvendo modelos lineares hierárquicos, para a determinação dos fatores associados às proficiências em leitura e matemática.

A AFE foi aplicada com a finalidade de se determinar um número mínimo de fatores que retenha uma parcela “razoável” da variabilidade devida às variáveis originais.

À medida que o número de variáveis a serem consideradas em técnicas multivariadas aumenta, há uma necessidade proporcional de maior conhecimento da estrutura e das inter-relações das variáveis. [...] A análise fatorial pode ser utilizada para examinar os padrões ou relações latentes para um grande número de variáveis e determinar se a informação pode ser condensada ou resumida a um conjunto menor de fatores ou componentes. (HAIR *et al.*, 2009, p. 100)

Para a redução do número de variáveis foi usado o programa computacional PASW, versão 17, da SPSS Inc., e o processo a que foi submetido cada um dos quatro arquivos de dados contextuais fornecidos pelo Projeto GERES será descrito a seguir.

A primeira tarefa foi realizar a seleção das variáveis presentes nos arquivos citados e que seriam submetidas à análise fatorial. Considera-se que cada “item” dos questionários dos quatro diferentes arquivos corresponde a uma variável. Como este trabalho foca a redução das variáveis de contexto, foram mantidas apenas as variáveis com essa natureza (a de fornecer informações de contexto) e que:

- a) trouxessem informações relevantes ao trabalho;
- b) não tivessem mais que 95% dos respondentes apontando para uma única e mesma resposta, quando o item tem apenas duas alternativas de resposta (item dicotômico);
- c) não ter mais que 90% apontando para uma única e mesma resposta, quando o item tem mais que duas alternativas de resposta (item politômico).

Do ponto de vista estatístico, isso se justifica pelo fato de estarmos interessados em “variabilidade”, ainda que do ponto de vista educativo todas as informações – mesmo as que se caracterizem por uma constante - sejam relevantes.

A segunda etapa foi fazer o agrupamento de itens politômicos para a composição de constructos³², sendo que os dicotômicos estão sendo tratados de forma diferenciada. Para tal, os itens atendem aos seguintes critérios:

a) tratar de um mesmo assunto, *denominado aqui de constructo*;

b) ter as mesmas alternativas de respostas que os demais itens que compõem o constructo, ou mais claramente, ter as mesmas alternativas de múltipla escolha.

Na terceira etapa, fez-se a determinação dos fatores que resumiriam as respostas dos itens de questionários a eles relacionados, sem perder significativamente sua capacidade explicativa e variabilidade, através de AFE. Neste processo foi usado o método de extração das componentes principais, com a rotação ortogonal *Varimax*³³. Cada uma das etapas desse processo, assim como os valores de referência usados em cada uma delas, será pormenorizada mais adiante.

Como quarta etapa, procedeu-se à denominação dos constructos e dos fatores que o compõem. Essa denominação é feita em função do conceito medido em cada um dos itens de questionário que o compõe. Como cada fator resultante do processo da AFE será uma variável do arquivo resultante desse processo, a denominação dada ao fator identificará essa variável. Nesta etapa, quando se busca atribuir um significado ao fator, Hair *et al.* (2009, p. 136) sugere que “todas as cargas fatoriais significantes tipicamente são utilizadas no processo de interpretação. Variáveis com cargas fatoriais maiores influenciam mais a seleção de nome ou rótulo para representar o fator”.

A quinta etapa envolveu a verificação da confiabilidade. Confiabilidade “é uma avaliação do grau de consistência entre múltiplas medidas de uma variável” (HAIR *et al.*, 2009, p.126). Para essa verificação foram medidas a correlação inter-itens, a correlação item-fator e a consistência interna do fator. Em seu trabalho, Hair *et al.* (2009), apoiado em Robinson *et al.* (1991), sugere que as correlações inter-itens excedam a 0,3 e que as correlações item-fator devem ser maiores que 0,5. A consistência interna do fator, medida pelo “limite inferior para o Alfa de Cronbach geralmente aceito é de 0,7 - apesar de poder diminuir para 0,6 em pesquisa exploratória” (HAIR *et al.*, *idem*) - e ainda, o Alfa de Cronbach

32 Para itens dicotômicos é recomendado o uso de outras técnicas tais como a Análise Fatorial de Informação Plena (SOARES, 2005), ou Análise Fatorial Booleana (HAIR *et al.*, 2009).

33 Segundo Hair *et al.* (2009), o método de rotação ortogonal *Varimax* é o mais popular, “concentrando-se na simplificação das colunas em uma matriz fatorial” (p. 102) e “maximiza a soma de variâncias de cargas exigidas da Matriz fatorial” (p. 118).

não deve aumentar se algum item for eliminado do fator. Uma atenção especial foi dada a essas medidas quando o número de itens por fator foi inferior a três. Nestes casos, o Alfa de Cronbach é consideravelmente diminuído enquanto as correlações item-fator são aumentadas. Nesta etapa, Hair *et al.*, (2009, p. 126) faz duas ressalvas importantes: (i) “como nenhum item isolado é uma medida perfeita de um conceito, devemos confiar em várias medidas diagnósticas para avaliar consistência interna e (ii) “como o aumento do número de itens, mesmo com grau igual de intercorrelação, aumenta o valor de confiabilidade do Alfa de Cronbach, os pesquisadores devem fazer exigências mais severas para as escalas com muitos itens”. Ou seja, a verificação da confiabilidade deve ser feita de maneira global envolvendo mais que uma medida de confiabilidade, sobretudo quando os fatores são compostos por menos itens e os pesquisadores precisam ser menos severos na medida do alfa de Cronbach.

A sexta etapa foi dedicada à validação da análise fatorial. Após verificar que os fatores gerados estavam de acordo com sua definição conceitual e atendiam aos critérios de confiabilidade, foi avaliado o grau de generalidade dos resultados para a população. “A questão da generalidade é crítica para todo método multivariado, mas é especialmente relevante nos métodos de interdependência, pois eles descrevem uma estrutura de dados que também deve ser representativa da população” (HAIR *et al.*, 2009, p. 123). Neste trabalho, essa verificação foi feita de forma empírica através da divisão da amostra em subamostras distintas e aleatórias, com 50% do total dos registros em cada subamostra. O modelo fatorial de cada constructo foi considerado estável quando o estimado em cada subamostra coincidia com o da amostra. Caso contrário, o processo de AFE era retomado reconsiderando a alteração no número de fatores ou na relação de itens que compunham o constructo.

Finalmente, foi feita a atribuição de valor ao fator. Como um fator é composto geralmente por mais que um dos itens de questionário, o valor que é atribuído ao fator corresponde à média aritmética simples dos valores que constam em cada um dos itens do questionário que compõe o respectivo fator.

É importante registrar que o processo de análise fatorial, apesar de se apoiar em testes estatísticos e em critérios bem delineados, não apresenta uma solução única ao pesquisador. Por isso, a AFE deve ser tomada como ferramenta cujo uso deve associar *critérios semânticos e conceituais* a resultados de testes estatísticos inerentes ao processo. Ou seja, a realização do trabalho deve apoiar-se no princípio da interpretabilidade (quando o pesquisador, de forma

subjetiva, agrupa os itens em um constructo e decide sobre quais e quantos fatores deverão compô-lo e no princípio da parcimônia (que busca explicar as correlações entre as variáveis observadas com o menor número de fatores possível). Há, portanto, certo grau de subjetividade na organização das variáveis que entram em um estudo, o que nem sempre é percebido e se encontra detalhadamente descrito neste trabalho.

Para a composição dos constructos, cada um dos fatores e itens que o compõe deve atender a um conjunto de critérios, de forma que o processo descrito a seguir deve ser repetido um número necessário de vezes, excluindo itens ou alterando o número de fatores de um constructo, até que os critérios sejam atendidos.

Dentre esses critérios o mais importante foi a busca da adequada distribuição da carga fatorial dos itens dentre os fatores de um constructo, afinal a análise fatorial por componentes principais é feita de forma que:

A primeira combinação linear é o primeiro componente principal. Este tem a propriedade de ter a maior variância possível. O segundo componente principal tem a propriedade de ter a maior variância possível e sendo não correlacionado com o primeiro componente. Os outros componentes principais são definidos similarmente, com o i -ésimo componente principal tendo a maior variância possível dado que ele é não correlacionado com os $i-1$ primeiros componentes principais. (MANLY, 2008, p. 100).

O resultado final desse processo é um constructo constituído de fatores que, por sua vez, é composto por um determinado número de itens. Nessa questão da relação de número de itens por fator, Worthington e Whittaker (2006) salientam que:

Quanto maior o número de itens em um fator, maior será a confiança que se poderá ter no uso desse fator em estudos futuros. Assim, com algumas ressalvas menores, alguns autores têm recomendado contra a retenção de fatores com menos de três itens (TABACHNICK e FIDELL, 2001). É possível manter um fator com apenas dois itens desde que estes sejam altamente correlacionados entre si (isto é, $r > 0,70$) e tenham baixa correlação com os outros fatores. Sob essas condições, pode ser apropriado considerar outros critérios (por exemplo, interpretabilidade) para decidir se mantém o factor ou descartá-lo. (p.821, tradução nossa)

As etapas desse processo e os valores das medidas estatísticas a elas associadas são os seguintes:

- 1) Determinação dos fatores de um constructo. A AFE agrupa todos os itens de um constructo em fatores de forma que haja a maior correlação possível entre os itens de um mesmo fator e ao mesmo tempo, se garanta a menor correlação possível entre os diferentes fatores. Para verificar a qualidade dessas correlações é usada a medida de adequação da amostragem de *Kaiser-Meyer-Olkin*, conhecido por teste KMO e o teste de esfericidade de *Bartlett*. O teste KMO compara as correlações simples com as demais parciais observadas, ou seja, verifica as correlações entre os itens quando o efeito linear dos demais itens é eliminado. Os valores do KMO oscilam entre 0 e 1, mas são desejáveis os valores mais próximos de 1, tendo 0,5 como limite mínimo de adequabilidade. Para uma interpretação mais detalhada, Hair *et al.* (2009, p. 110) apresenta a seguinte escala de interpretabilidade: no intervalo de 0 a 0,40, inadequado; 0,50 a 0,59, ruim; 0,60 a 0,69, fraco; 0,70 a 0,79, mediano; 0,80 e 0,89, bom e entre 0,90 e 1,0, excelente. Para Miranda (2006, p. 36),

O teste KMO apresenta a variância das variáveis e é necessário que o resultado seja acima de 0,50 para que os dados sejam úteis na AFE. O teste de Bartlett's conclui se a matriz é idêntica, ou seja, sendo idênticos, os dados não seriam relacionados. Por isso, espera-se que a hipótese deste teste seja negada, resultando um sintótico menor do que 0,05.

- 2) A avaliação da qualidade das correlações é feita pelo teste de esfericidade de Bartlett, que verifica se há correlação entre os fatores. Neste teste, valores menores que 0,05 indicam que o processo de AFE é adequado.
- 3) Verificação da comunalidade do item, ou seja, a proporção da variância que é por ele explicada deve ser aceitável. Para Hair *et al.* (2009, p.121) “o pesquisador identificaria todos os itens com comunalidade menores que 0,50 como não tendo explicação suficiente”. Há outros autores, como por exemplo, Worthington e Whittaker (2006) que propõem que valores aceitáveis são aqueles superiores a 0,4.
- 4) Exame da adequada distribuição da carga fatorial de cada um dos itens entre os distintos fatores que compõem um determinado constructo. Neste trabalho, os valores são orientados por Hair *et al.* (2009) e Worthington e Whittaker (2006) e são os seguintes:

- a. Pelo critério de Kaiser, um fator é considerado adequado quando seu valor próprio seja *igual ou superior a 1* (*eigenvalue* ≥ 1). Em alguns casos, em que essa condição não é atendida e não se deseja excluir o item, dada sua importância contextual e também por não se tratar de uma análise confirmatória, mas sim exploratória, pode-se optar pelo critério do *Screen Plot* onde se verifica a não variação entre os valores próprios dos diferentes fatores;
- b. Carga fatorial dos itens *igual ou superior a 0,5* (*factor loadings* $\geq 0,5$). Segundo Hair *et al.* (2009):

para uma amostra grande, maior que 350, uma carga fatorial maior que 0,30 pode ser considerada estatisticamente significativa. Do ponto de vista prático cargas fatoriais com valor absoluto maiores que 0,5 são tidas como significativas (p. 120).

- c. Inexistência de carga fatorial cruzada de um item entre diferentes fatores. Um item tem carga fatorial cruzada quando apresenta *carga fatorial superior ou igual a 0,32* (*factor loadings* $\geq 0,32$) em mais que um fator e ainda, quando a diferença entre essas cargas for inferior a 0,15 (*cross loadings* $\leq 0,15$). Essas orientações são dadas por Worthington e Whittaker (2006) quando afirmam que “pesquisadores devem excluir os itens com cargas fatoriais menores que 0,32 ou com cargas cruzada com diferenças inferiores a 0,15 entre as maiores cargas” (p. 823, tradução nossa).
 - d. Um item deve ser excluído quando apresentar cargas fatoriais cruzadas mesmo no método de extração de componentes principais com rotação ortogonal (*Varimax*).
- 5) Verificação da capacidade explicativa da variância. Após a redução da quantidade do número de itens em alguns fatores, é necessário que esses fatores expliquem no mínimo 40% da variabilidade de todos os itens juntos antes da redução.

Finalmente, cabe assinalar que muitas variáveis dos arquivos de origem são de identificação e não de caracterização e, por isso, não foram submetidas à AFE. Outros itens do questionário dos alunos que se destinavam a medir o nível socioeconômico, também não foram trabalhados uma vez que já foram tratados por ALVES e SOARES (2009) e constam nos arquivos fornecidos na base de dados GERES. A variável com essa informação consta na base de dados GERES com a designação de “NSE”, nível socioeconômico do aluno.

Tabela 9 - Número de variáveis e registros dos arquivos GERES

Arquivo	Número de variáveis	Números de variáveis submetidas à AFE	Números de itens excluídos no processo	Número de Registros	Numero de fatores resultantes da AFE
Alunos	145	21	4	35538	5
Professores	244	149	17	2918	34
Diretores	205	42	15	669	8
Escolas	67	53	26	312	10

Como se pode constatar, muitas variáveis não foram submetidas à AFE, isso porque, contêm informações com função de identificação e não caracterização do perfil. Por isso, foram mantidas e, a opção pela exclusão, deverá ocorrer segundo as necessidades do estudo.

Do total das 661 variáveis disponibilizadas, 396 não foram submetidas ao processo, por terem apresentado respostas sem a mínima variabilidade necessária, por serem usadas diretamente no banco de dados (12 delas) ou ainda por serem usadas na estimativa do nível socioeconômico também disponibilizado no arquivo. As 265 variáveis resultantes foram submetidas ao processo. O resultado desse processo de redução de variáveis foi bastante satisfatório uma vez que das 265 variáveis submetidas ao processo, 203 (76,6%) delas foram reduzidas a 57 variáveis utilizadas como variáveis explicativas nos modelos lineares hierárquicos. Das 62 variáveis excluídas do processo, 49 delas o foram por apresentarem cargas cruzadas no processo, indicando que a variabilidade dos dados tem uma correlação inter-itens semelhante em diferentes fatores do constructo, ou seja, o item propiciou interpretação diferente entre os respondentes revelando má adequabilidade do mesmo.. Os demais itens foram excluídos por conterem informações não adequadas à pesquisa.

3.2) Preparação dos arquivos

Na preparação dos arquivos, as informações dos quatro arquivos foram trabalhadas de forma que resultam em três outros diferentes arquivos:

- O primeiro arquivo, denominado Nível 1 – Medidas Cognitivas no Tempo, é composto por informações inerentes ao ano de escolarização de cada aluno tais como a

proficiência inicial e final, a informação se ele foi aprovado ou retido no processo. Cada registro contém a medida da proficiência inicial e final de um ano de escolarização de um único aluno, que é identificado por chave composta, ordenadamente, pelo código que identifica a escola onde o aluno mais permaneceu durante o processo de escolarização, pela identificação do aluno e pelo ano de escolarização;

- No segundo arquivo, denominado de Nível 2 – Alunos, estão as informações que definem o perfil do aluno tais como nível socioeconômico, indicadores da valorização que ele atribui à escola e aos estudos, estímulos para o estudo recebidos por pais e professores, além de informações que indicam sua autoestima. O arquivo é indexado pelo código que identifica a escola onde o aluno permaneceu naquele ano de escolarização³⁴ e identificação do aluno;

- No terceiro arquivo, denominado Nível 3 – Escolas, são armazenadas as informações que caracterizam a escola tais como sua estrutura física, o perfil da sua gestão, do seu diretor e de seus professores. Nesse arquivo a chave de indexação é a identificação da escola.

Cada um desses arquivos resulta das informações disponíveis nos arquivos fornecidos pelo Projeto GERES e de sua finalidade para o HLM.

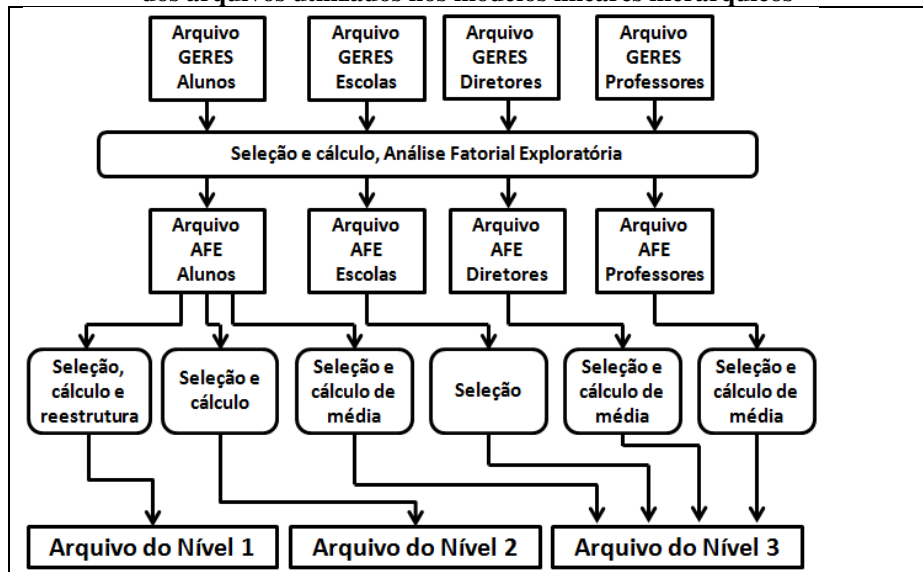
Todas as etapas percorridas na preparação desses três diferentes arquivos são descritas a seguir, que tomam como origem os arquivos preparados após a aplicação da AFE.

A preparação dos arquivos do Nível 1 (Medidas Cognitivas no Tempo) e do Nível 2 (Alunos) foi feita a partir apenas das informações contidas no arquivo que continham as informações sobre alunos, fornecido pelo Projeto GERES. Já para a preparação do arquivo do Nível 3 (Escolas) foram usadas as informações sobre os alunos, professores e escolas.

A Figura 8 traz um esquema que sintetiza os procedimentos adotados para a preparação dos três arquivos que serão usados nos modelos lineares hierárquicos.

34 Esta opção foi a escolhida por oferecer menor discrepância na vinculação dos dados entre os diferentes níveis. Essa problemática e as decisões metodológicas adotadas para minimizar seus efeitos são abordadas mais adiante quando são descritos detalhadamente os procedimentos utilizados na preparação dos arquivos.

Figura 8 - Procedimentos adotados e arquivos relacionados na preparação dos arquivos utilizados nos modelos lineares hierárquicos



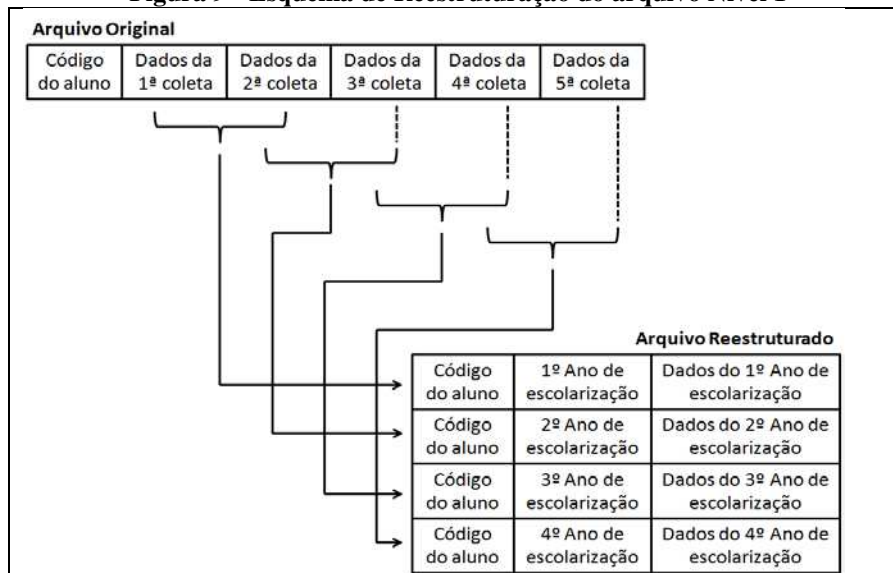
Todos os procedimentos aplicados na preparação desses arquivos estão descritos detalhadamente a seguir.

3.2.1) Preparação do arquivo Nível 1 – Medidas Cognitivas no Tempo

O arquivo do Nível 1, Ano de Escolarização, foi preparado a partir da seleção, cálculos de variáveis, criação de variáveis indicadoras, atribuições de valores e reestruturação contidas em arquivo que continha informações sobre alunos fornecido pelo Projeto GERES. A reestruturação é o procedimento que cria quatro registros diferentes para cada aluno, sendo um para cada ano de escolarização, a partir de um único registro que contém as informações das cinco coletas. Essa reestruturação está ilustrada na Figura 9.

A criação de variáveis indicadoras foi necessária para tornar possível o controle por ano de escolarização e para a atribuição do valor da proficiência inicial de cada um desses anos de escolarização, uma vez que, o modelo linear hierárquico foi concebido de forma que a proficiência final de cada ano de escolarização será considerada como variável resposta e a proficiência inicial como variável explicativa.

Figura 9 - Esquema de Reestruturação do arquivo Nível 1



3.2.1.1) Definição da variável que indica se o aluno foi retido em algum dos quatro primeiros anos de escolarização.

A inclusão dessa variável foi necessária para discriminar o aluno que, por ter sido retido, tenha participado do teste num determinado ano de escolarização mesmo que tenha cursado uma série anterior. Esse cuidado é necessário uma vez que, mesmo estando defasado, o aluno participa do teste que contém as mesmas questões dos demais alunos que não foram retidos. A princípio, entende-se que esse aluno retido, por não ter sido submetido às intervenções pedagógicas do ano em que foi medida sua proficiência, tenha desempenho inferior aos demais alunos.

Para a identificação dos alunos que ficaram retidos foram comparadas as séries de cada uma das ondas. É atribuído o valor 1 quando o aluno não foi retido no processo e 0 quando foi retido, ou seja, o aluno terá o valor 1 atribuído a essa série quando não foi retido em qualquer um dos anos anteriores ao ano em que a variável está sendo atribuída.

O Quadro 9 explicita os comandos usados no PASW para a obtenção desses resultados.

Quadro 9 - Comandos PASW para a criação da variável que controla a retenção

Função:	COMPARA AS SÉRIES ENTRE DUAS ONDAS CONSECUTIVAS
Comandos	COMPUTE Retido21= ~ (Série2=Série1). COMPUTE Retido32= ~ (Série3=Série2). COMPUTE Retido43= ~ (Série4=Série3). COMPUTE Retido54= ~ (Série5=Série4).
Função:	INVERTE E AJUSTA OS VALORES ATRIBUÍDOS ÀS VARIÁVEIS
Comandos	RECODE Retido21 (0=0) (1=0) (MISSING=0) INTO Ret_Ano1_INV. RECODE Retido32 (0=1) (1=0) (MISSING=0) INTO Ret_Ano2_INV. RECODE Retido43 (0=1) (1=0) (MISSING=0) INTO Ret_Ano3_INV. RECODE Retido54 (0=1) (1=0) (MISSING=0) INTO Ret_Ano4_INV.
Função:	VERIFICA SE O ALUNO FOI RETIDO EM ALGUMA SERIE ANTERIOR AO RESPECTIVO ANO
Comandos	COMPUTE Ret_Ac_Ano1_INV=((Ret_Ano1_INV=1)). COMPUTE Ret_Ac_Ano2_INV=((Ret_Ano1_INV=1) (Ret_Ano2_INV=1)). COMPUTE Ret_Ac_Ano3_INV=((Ret_Ano1_INV=1) (Ret_Ano2_INV=1) (Ret_Ano3_INV=1)). COMPUTE Ret_Ac_Ano4_INV=((Ret_Ano1_INV=1) (Ret_Ano2_INV=1) (Ret_Ano3_INV=1) (Ret_Ano4_INV=1)).
Função:	ATRIBUI O VALOR 0 PARA O ALUNO QUE FOI RETIDO NO PROCESSO E 1 PARA OS DEMAIS
Comandos	RECODE Ret_Ac_Ano1_INV (0=1) (1=0) INTO DEFAS_Ano1. RECODE Ret_Ac_Ano2_INV (0=1) (1=0) INTO DEFAS_Ano2. RECODE Ret_Ac_Ano3_INV (0=1) (1=0) INTO DEFAS_Ano3. RECODE Ret_Ac_Ano4_INV (0=1) (1=0) INTO DEFAS_Ano4.

Deste processamento, a primeira variável construída é aquela que indica se o aluno cursou a mesma série em duas ondas consecutivas o que indica se houve a retenção. A Tabela 10 apresenta os valores encontrados nessas variáveis.

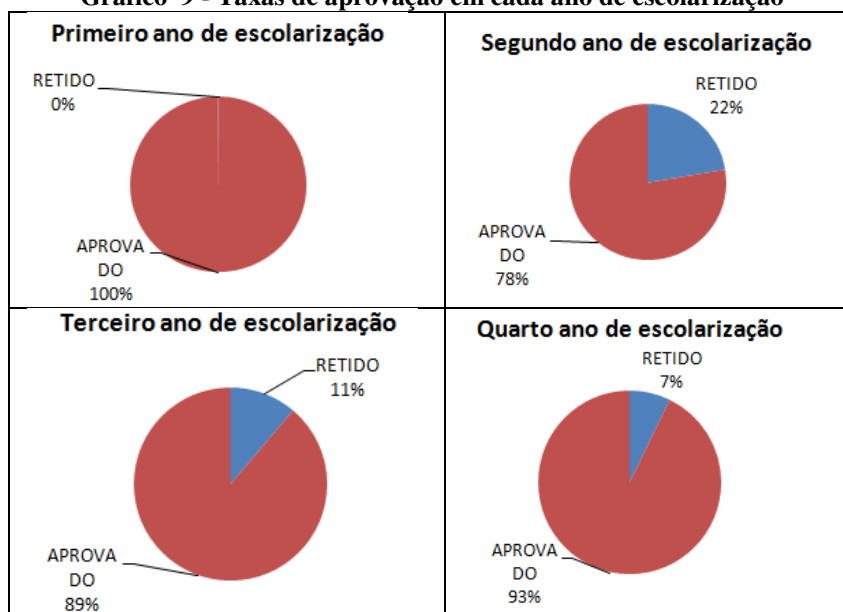
Apesar da variável Retido21 não indicar a retenção, uma vez que as ondas 1 e 2 foram aplicadas no mesmo ano e por isso, evidentemente, espera-se que a série seja a mesma, ela foi mantida no estudo, pois essas ondas são usadas para a verificação do valor agregado no primeiro ano de escolarização. Com a manutenção desta variável pode-se constatar que havia uma inconsistência nos dados em 24 alunos. Em uma investigação mais aprofundada sobre esses casos verificou-se que eles indicavam que os alunos estavam cadastrados na primeira onda na segunda série e na segunda onda na primeira série, caracterizando assim, um equívoco em tal cadastramento. Como pode ser observado na sequência de comandos, esse equívoco é corrigido, já que não é possível ter um aluno defasado no primeiro ano de escolarização.

Tabela 10 - Frequência das variáveis que indicam a retenção

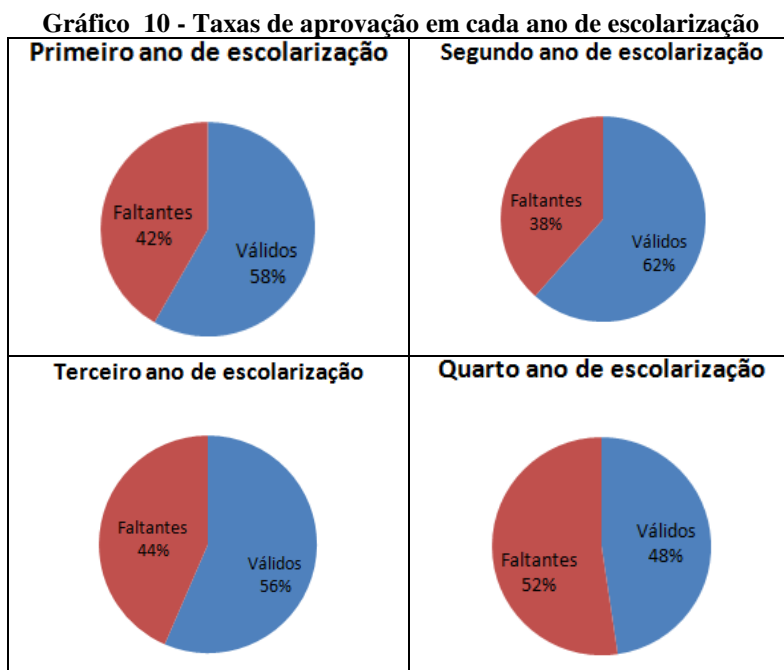
	Status	Frequência	(%)
Retido21	RETIDO	24	0,1
	APROVADO	20695	58,2
	Dado Faltante	14819	41,7
	Total	35538	100
Retido32	RETIDO	4880	13,7
	APROVADO	16991	47,8
	Dado Faltante	13667	38,5
	Total	35538	100
Retido43	RETIDO	2246	6,3
	APROVADO	17828	50,2
	Dado Faltante	15464	43,5
	Total	35538	100
Retido54	RETIDO	1216	3,4
	APROVADO	15720	44,2
	Dado Faltante	18602	52,3
	Total	35538	100

Os sectogramas que compõem o Gráfico 9 sintetizam as informações acerca das retenções e aprovações de cada ano de escolarização.

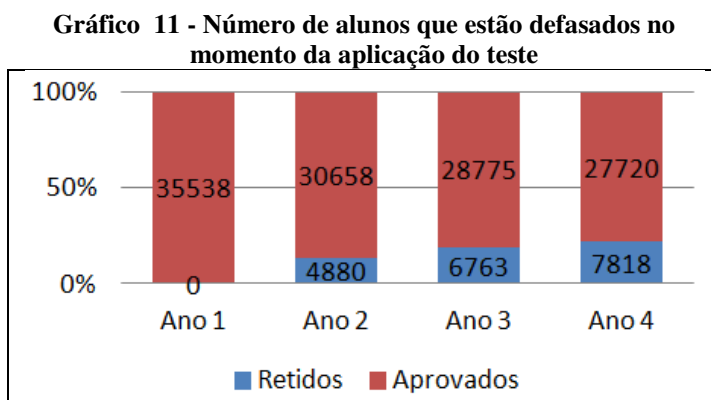
Gráfico 9 - Taxas de aprovação em cada ano de escolarização



Da Tabela 10 constata-se também que os dados faltantes (*missing data system*) em cada um dos anos oscilam entre 38,5% a 52,3%, o que indica a quantidade de registros que não possuem duas medidas consecutivas em cada ano, como é evidenciado no Gráfico 10.



No entanto, essa informação que indica a retenção ou aprovação do aluno em cada ano não é suficiente, uma vez que, se supostamente ele seja retido somente no primeiro ano de escolarização, ele estará defasado dos demais alunos da turma até o final do processo. Para esse controle, foram criadas variáveis indicadoras que permitem o processamento que resulta em uma variável que indica essa defasagem do aluno em relação a sua turma inicial, como mostra o Gráfico 11.

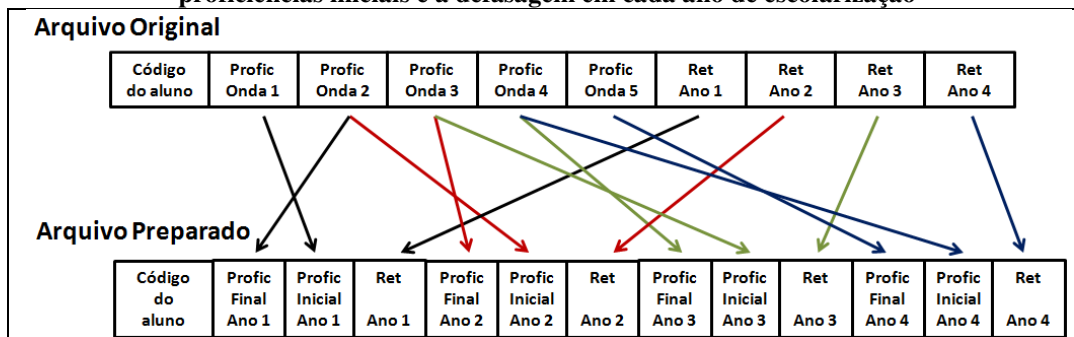


3.2.1.2) Inclusão de variáveis indicadoras para permitir a especificação em inicial e final das variáveis que serão inserida no arquivo do Nível 1.

A organização do arquivo do Nível 1 se caracteriza pela intenção de se colocar a proficiência final como variável resposta e como variáveis explicativas, proficiência inicial e controle da retenção.

Os procedimentos necessários para a preparação desse arquivo é esquematizado na Figura 10.

Figura 10 - Esquema da definição das variáveis que armazenam as proficiências finais, proficiências iniciais e a defasagem em cada ano de escolarização



3.2.1.3) Inclusão de variáveis auxiliares para identificação da escola onde o aluno mais permaneceu

Dada a estruturação proposta para o arquivo de Nível 1, onde a chave de cada registro deve ser “escola+aluno+ano_de_escolarização”, a existência de alunos que podem ter frequentado diferentes escolas gera um problema de vinculação dos registros deste arquivo com os do arquivo de Nível 2, cuja chave é “escola+aluno”. Vinculação essa que garante o aninhamento dos dados.

O arquivo original tem como a variável chave o registro “escola” cujo conteúdo é a última escola na qual o aluno estudou. A manutenção desse critério para a estruturação do arquivo pode trazer distorções, uma vez que, mesmo quando o aluno muda todas as informações acumuladas ao longo dos diferentes anos que seriam atribuídas a essa última escola.

Um exemplo hipotético para elucidar esse problema seria a situação na qual um aluno permaneceu em uma mesma escola durante os seus quatro anos de escolarização e tenha se transferido para outra escola um mês antes da aplicação da última onda. No processamento,

todos os fatores da escola de origem que interferiram na proficiência deste aluno seriam creditados à última escola, a qual, devido ao tempo, pouco o influenciou.

Nesse contexto, duas diferentes situações poderiam ocorrer:

a) Caso seja atribuído o conteúdo na variável “escola” do Nível 2 também no Nível 1 (identificação da escola na qual o aluno estava matriculado na última onda onde o aluno participou) , os desempenhos desse aluno seriam associados à última escola, o que representaria uma distorção na análise;

b) Caso se mantenha na variável da escola do arquivo do Nível 2, a identificação da última escola, e no Nível 1, identificação da escola onde o aluno efetivamente participou de cada uma das ondas, seriam perdidos o desempenho de três anos.

Assim, na preparação dos arquivos de Nível 1 e 2, atribuiu-se à variável “escola” a identificação da escola que o aluno mais permaneceu no decorrer dos quatro anos de escolarização. Evidentemente, este procedimento minimiza as distorções no momento da análise dos resultados, porém não os erradica.

A Tabela 11 mostra que este cuidado se restringe a aproximadamente 4% dos registros dos alunos que participaram das cinco ondas.

Tabela 11 - Número de alunos que trocaram de escola e participaram das cinco ondas

Categoria	Freq	%
Escola2 = Escola3 ≠ Escola4 ≠ Escola5	1	0,01%
Escola2 = Escola3 ≠ Escola4 = Escola5	182	1,70%
Escola2 = Escola3 = Escola4 ≠ Escola5	161	1,51%
Escola3 = Escola4 ≠ Escola2 = Escola5	2	0,02%
Escola3 ≠ Escola2 = Escola4 = Escola5	5	0,05%
Escola2 ≠ Escola3 = Escola4 = Escola5	53	0,50%
Escola4 ≠ Escola 2 = Escola3 = Escola5	4	0,04%
Escola2 = Escola 3 = Escola4 = Escola5	10277	96,18%
Subtotal	10685	
Dados faltantes	24853	
Total	35538	

3.2.1.4) Reestruturação do arquivo

Como já apresentado anteriormente, é necessário que cada registro do arquivo Nível 1 contenha as informações inerentes de um determinado ano de escolarização.

Por isso, o procedimento para a criação desse arquivo requer buscar as informações de cada aluno, que se encontram em um único registro no arquivo Origem, e as coloquem em diferentes registros no arquivo do Nível 1, sendo um para cada ano de escolarização.

Criadas as variáveis indicadas nos passos anteriores, para proceder essa transferência é necessário reestruturar o arquivo origem, criando o arquivo do Nível 1.

A Figura 11 a seguir representa essa transferência:

Figura 11- Exemplo de reestruturação de arquivo no PASW

Registro das informações de um aluno no arquivo origem										
Aluno	Escola Onda 1	Escola Onda 2	Escola Onda 3	Escola Onda 4	Escola Onda 5	Profic Onda 1	Profic Onda 2	Profic Onda 3	Profic Onda 4	Profic Onda 5
X	A	A	B	C	C	101,17	122,20	130,33	148,34	160,51

Registro das informações de um aluno no arquivo Nível 1					
Aluno	Ano de Escolarização	Escola Inicial	Escola Final	Proficiência Final	Proficiência Inicial
X	1	A	A	101,17	122,20
X	2	A	B	122,20	130,33
X	3	B	C	130,33	148,34
X	4	C	C	148,34	160,51

Nesta figura encontra-se um exemplo com as informações de um aluno e envolvendo apenas as variáveis “escola” e “proficiência”.

Para efetivamente fazer essa reestruturação, foi usado o comando “*Reestructure*” do PASW de forma que as informações de cada uma das ondas sobre escola, turma, proficiência e indicador de retenção ou aprovação, que antes estavam em um único registro, foram transferidos para diferentes registros, sendo um para cada ano de escolarização e devidamente colocados de forma que sejam indicadas as situações iniciais e finais de cada uma dessas variáveis.

3.2.1.5) Inclusão das variáveis indicadoras para o controle do ano de escolarização

Esse passo consiste na inclusão das variáveis *T2*, *T3* e *T4* e atribuição do valor a cada uma dessas variáveis indicadoras em função do ano de escolarização, como descrito no Quadro 5 do Capítulo 2.

3.2.1.6) Inclusão das variáveis indicadoras para o incremento da proficiência inicial

Esse passo consiste na inclusão das variáveis *PR_IN*T2*, *PR_IN*T3* e *PR_IN*T4* e atribuição do valor a cada uma delas em função do ano de escolarização, como descrito no Quadro 6 do Capítulo 2.

Ao final da preparação o arquivo de Nível 1 é composto pelas seguintes variáveis:

Quadro 10 - Variáveis do arquivo Nível 1 - Ano de Escolarização

Variável	Descrição
<i>IDEscola</i>	Código que identifica a escola
<i>IDAluno</i>	Código que identifica o aluno
<i>Ano</i>	Ano de escolarização
<i>PR_FI</i>	Proficiência final para os quatro anos de escolarização.
<i>T2</i>	Variável indicadora que recebe o valor 1 quando o registro for do segundo ano de escolarização e zero para os demais anos.
<i>T3</i>	Variável indicadora que recebe o valor 1 quando o registro for do terceiro ano de escolarização e zero para os demais anos.
<i>T4</i>	Variável indicadora que recebe o valor 1 quando o registro for do quarto ano de escolarização e zero para os demais anos.
<i>PR_IN</i>	Proficiência inicial para os quatro anos de escolarização.
<i>PR_INxT2</i>	Variável indicadora que recebe o valor de <i>PR_IN</i> quando o registro for do segundo ano de escolarização e zero para os demais anos.
<i>PR_INxT3</i>	Variável indicadora que recebe o valor de <i>PR_IN</i> quando o registro for do terceiro ano de escolarização e zero para os demais anos.
<i>PR_INxT4</i>	Variável indicadora que recebe o valor de <i>PR_IN</i> quando o registro for do quarto ano de escolarização e zero para os demais anos.
<i>RET</i>	Variável que recebe o valor 1 quando o aluno sofreu alguma retenção no processo por isso cursa alguma série diferente daquela em que o teste foi aplicado.

3.2.2) Preparação do arquivo Nível 2 – Alunos

A preparação do arquivo do Nível 2 foi feita a partir do arquivo com informações sobre alunos e seguiu as etapas descritas na sequência.

3.2.2.1) Associação da identificação do aluno à da escola

O primeiro passo nessa preparação foi associar a identificação do aluno à escola, respeitando o cuidado explicitado anteriormente, no item 1.1.4 (na preparação do arquivo de Nível 1).

3.2.2.2) Inclusão da variável que controla a defasagem idade série

Essa variável é determinada a partir da diferença entre a idade do aluno e a idade que é esperada para a série, tomando como base a última onda de aplicação, quando essa informação foi coletada.

3.2.2.3) Associação da identificação da cor da pele do aluno

O programa HLM somente processa variáveis numéricas e, por isso, quando a variável cujo conteúdo tenha mais que duas categorias e que não seja possível, a princípio, dicotomizar essa variável, um procedimento com mais etapas precisa ser feito.

Neste caso, quando uma variável tiver como conteúdo n possibilidades, devem ser criadas (n-1) variáveis indicadoras de forma que a combinação booleana dessas variáveis indique que o resultado seja atribuído a cada uma das categorias existentes.

Um exemplo para essa situação é cor, cujo conteúdo pode conter as informações: branco, negro, oriental e indígena. Como no arquivo de alunos essa variável tem cinco categorias, é necessária a criação de quatro variáveis indicadoras cujos conteúdos podem ser zero ou um, de forma que quando uma delas receber o valor um, as demais ficam zeradas. A combinação dessas variáveis resulta na identificação de uma cor, como mostra o Quadro 11, a seguir.

Quadro 11 - Exemplo de transformação de uma variável categórica em variáveis indicadoras dicotômicas

Cor	Variáveis indicadoras			
	Cor1	Cor2	Cor3	Cor4
branco	0	0	0	0
pardo	1	0	0	0
negro	0	1	0	0
oriental	0	0	1	0
indígena	0	0	0	1

3.2.2.4) Determinação da variável que controla o peso amostral

O uso do peso amostral busca assegurar que as estimativas feitas possam ser comparadas à população de referência. No entanto, durante o período de quatro anos, quando

ocorreu a coleta de dados, algumas alterações em relação ao plano amostral ocorreram, uma vez que as amostras de escolas e de alunos se alteraram.

Além disso, para o processamento dos modelos lineares hierárquicos as seguintes restrições foram colocadas, de forma que permaneceram na amostra estudada apenas: (a) os alunos que tiveram duas medidas consecutivas; (b) os alunos com dados completos no arquivo do Nível 2 e (c) as escolas com no mínimo quatro alunos e com dados de contexto completos.

O plano amostral adotado pelo projeto GERES consiste de um desenho de amostra estratificada de conglomerados em três estágios e encontra-se descrito em Miranda e Andrade (2004)

A ponderação da amostra do GERES foi feita respeitando-se o máximo possível a estrutura do plano amostral utilizado. Foram atribuídos valores a cada uma das unidades de amostragem, com exceção de turmas, devido a não constância desta unidade na janela temporal. Primeiramente, as escolas da amostra selecionada foram classificadas em seus respectivos estratos de origem. Toda a ponderação de escolas e alunos em cada estrato foi efetuada independentemente dos demais estratos. As contagens populacionais para calibração dos pesos em cada estrato foram obtidas do cadastro de expansão derivado do Censo Escolar 2003. (BROOKE e BONAMINO, 2011, p. 35)

O cálculo do peso amostral dos alunos é dado por $P_{hij} = \frac{N_{hi}}{n_{hi}}$, de forma que o peso do aluno j da escola i é condicionado ao estrato h , onde N_{hi} e n_{hi} são, respectivamente, o número de alunos no universo e na amostra pertencente à escola i e ao estrato h .

O peso semi-calibrado do aluno j da escola i pertencente ao estrato h é dado por:

$$P_{hij}^{SC} = P_{hij} P_{hi}^C = P_{hij} \frac{N_L}{\sum_{heP} \sum_{ieh} P_{hi}} P_{hi}.$$

O peso calibrado do aluno j da escola i pertencente ao estrato h é dado por:

$$P_{hij}^C = \frac{N_L}{\sum_{heP} \sum_{ieh} P_{hij}^{SC}} P_{hij}^{SC},$$

Em que N_L é o número total de alunos no estrato de calibração L segundo o cadastro de referência.

O peso calibrado P_{hij}^C é atribuído à variável de cada aluno, cujo valor permite estimar o total de alunos no estrato de calibração L e determinar exatamente o total conhecido no cadastro de referência de alunos (N_L).

O Quadro 12 apresenta as variáveis que compõem o arquivo do Nível 2, resultantes desse processo.

Quadro 12 - Variáveis do arquivo do Nível 2 usadas no HLM

Variável	Descrição
<i>IDEscola</i>	Código que identifica a escola
<i>IDAluno</i>	Código que identifica o aluno
<i>FA01</i>	Sexo
<i>FA02</i>	Cor
<i>FA02_cor1</i>	Variável indicadora que indica quando o aluno é negro
<i>FA02_cor2</i>	Variável indicadora que indica quando o aluno é oriental
<i>FA02_cor3</i>	Variável indicadora que indica quando o aluno é negro
<i>FA02_cor4</i>	Variável indicadora que indica quando o aluno é negro
<i>FA03</i>	Anos de defasagem idade-série
<i>FA04</i>	Indica se o aluno fez a pré-escola
<i>FA05</i>	Indica se o aluno foi retido em uma das séries
<i>FA06</i>	Nível socioeconômico dos alunos GERES
<i>FA07</i>	Estímulo extrínseco para estudo
<i>FA08</i>	Estímulo intrínseco para o estudo
<i>FA09</i>	Percepção do aluno quanto ao incentivo do professor
<i>FA10</i>	Gosto pela escola
<i>FA11</i>	Estímulo interno para lição de casa
<i>FA12</i>	Meus professores pensam que sou um aluno lento
<i>FA13</i>	Sou tipo de pessoa que sempre será capaz de conseguir, se tentar
<i>FA14</i>	Algumas vezes eu sinto que não sou bom o bastante
<i>PESOAL</i>	Peso amostral

3.2.3) Preparação do arquivo Nível 3 – Escolas

O arquivo de Nível 3 é composto por 59 variáveis que se originam dos quatro arquivos resultantes da Análise Fatorial Exploratória que reduziu a quantidade de variáveis disponibilizadas nos quatro arquivos correspondentes do Projeto GERES.

A descrição dessa preparação serão feitas a seguir.

3.2.3.1) Inclusão da variável que controla o nível socioeconômico médio da escola

Nessa variável (FEA01) foi atribuída a média entre os valores que representam o nível socioeconômico dos alunos da escola.

3.2.3.2) Inclusão das variáveis que controlam as características da escola

Essas variáveis (FEE02 a FEE11) receberam diretamente os valores encontrados no arquivo Escolas resultante da AFE.

3.2.3.3) Inclusão das variáveis que controlam as características da escola referente às características associadas ao diretor

Segundo Brooke e Bonamino (2011, pg. 87), foram recolhidas informações “em 300 escolas entre os anos de 2005 (293 escolas), 2006 (280 escolas) e 2008 (96 escolas). Por isso, nos casos em que uma escola teve mais que um diretor durante período de coleta de dados, optou-se por calcular a média, em cada uma das variáveis, das diferentes respostas dos diretores e atribuir essa média às variáveis correspondentes no arquivo do Nível 3. Esse procedimento foi adotado nas variáveis (FED12 a FED26).

As variáveis FED12 a FED19 são resultados da AFE, enquanto que as variáveis FED20 a FED26 são retiradas diretamente dos itens de questionário.

3.2.3.4) Inclusão das variáveis que controlam as características da escola referente às características associadas ao professor

Assim como ocorre com diretores, é certo que muitos são os professores dos alunos em uma escola. Por isso, o mesmo procedimento foi adotado nas variáveis FEP27 a FEP59, ou seja, a cada uma dessas variáveis foi atribuído o valor médio das respostas dos diferentes professores que passaram pelos alunos nos diferentes anos de escolarização.

Certamente, algo indesejável para um estudo que considera a hierarquia dos dados, mas esta é uma opção metodológica dada devido às limitações já apontadas anteriormente, que se referem ao número de níveis possíveis no modelo e à quantidade de dados disponíveis para esse possível outro nível.

As variáveis FEP27 a FE54 são resultados da AFE, enquanto que as variáveis FEP55 a FEP59 são retiradas diretamente dos itens de questionário.

Ao final da preparação o arquivo de Nível 3 é composto por 59 variáveis que para a melhor organização do texto e facilitar a análise foram distribuídas em 4 grupos:

a) Grupo 1, encontram-se a variável que traz o nível socioeconômico da escola, determinado pela média do nível socioeconômico de seus alunos e pelas variáveis que foram sintetizadas dos questionários de contexto sobre a escola;

Quadro 13 - Variáveis do Nível 3 pertencente ao grupo 1

Variável	Descrição
FEA01	NSE médio da escola a partir do NSE dos alunos
FEE02	Conservação e limpeza supraestrutura
FEE03	Conservação e limpeza área de alimentação
FEE04	Espaço sala de leitura
FEE05	Condições da Biblioteca
FEE06	Laboratório de informática
FEE07	Sala de artes e laboratório de ciências
FEE08	Disponibilidade de computador e internet para professor
FEE09	Existência de equipamentos para fotocópia e escaneamento
FEE10	Existência de material concreto para matemática
FEE11	Existência de terrário e aquário

b) Grupo 2, estão as variáveis originadas dos questionários de contexto aplicado nos diretores;

Quadro 14 - Variáveis do Nível 3 pertencente ao grupo 2

Variável	Descrição
FED12	Interferência por intimidação, violência, discriminação e drogas
FED13	Interferência do perfil do professor
FED14	Interferência do perfil do aluno e da família
FED15	Interferência da orientação pedagógica
FED16	Presença física do diretor
FED17	Organização de eventos
FED18	Frequência de encontros com a comunidade
FED19	Frequência da presença na comunidade
FED20*	Proporção de diretores que possuem Ens. Sup. em Pedagogia
FED21*	Proporção de Diretores que possui outra atividade que contribui para a renda pessoal
FED22*	Proporção de diretores que trabalham em apenas uma escola
FED23*	Carga horária de trabalho do(a) diretor(a) na escola
FED24*	Tempo que o diretor trabalha na escola ininterruptamente na função de direção
FED25*	Tempo que o(a) diretor (a) trabalha em anos na educação
FED26*	Idade do(a) diretor(a) em anos

c) Grupo 3, estão as variáveis sintetizadas a partir do questionário de contexto dos professores e estão ligadas às práticas pedagógicas realizadas na escola.

Quadro 15 - Variáveis do Nível 3 pertencente ao grupo 3

Variável	Descrição
FEP33	Frequência do dever de casa
FEP36	Acesso a atividades culturais
FEP37	Proposição de leitura
FEP38	Proposição de redação
FEP39	Proposição de cópia ou caligrafia
FEP40	Proposição de leitura individual é precedida da explicação do professor
FEP41	Leitura em prática coletiva
FEP42	Leitura individual e produção de texto individual
FEP43	Uso da metodologia de resolução de problemas
FEP44	Ensino com ênfase em memorização
FEP45	Adequação do livro ao ensino
FEP46	Grau de importância inerente a indicação externa do livro
FEP47	Tempo usado nas disciplinas relacionadas com as proficiências
FEP48	Tempo usado nas disciplinas não relacionadas com as proficiências
FEP49	Tempo usado em disciplinas de formação complementar
FEP52	Uso de recursos audiovisuais
FEP53	Uso de laboratório e multimídia
FEP54	Uso de materiais concretos de matemática e geografia

d) Grupo 4, estão as variáveis sintetizadas a partir do questionário de contexto dos professores e caracterizam a dinâmica e o perfil da escola.

Quadro 16 - Variáveis do Nível 3 pertencente ao grupo 4

Variável	Descrição
FEP27	Relação com o diretor
FEP28	Boa interação entre a equipe
FEP29	Pouco trabalho coletivo
FEP30	Boa relação entre professores
FEP31	Unidade de propósito entre os professores
FEP32	Violência no ambiente escolar
FEP34	Grau de expectativas negativa do aprendizado
FEP35	Grau de expectativas positiva do aprendizado
FEP50	Interrupção das aulas por indisciplina
FEP51	Interrupção das aulas por necessidade administrativa
FEP55*	Proporção de professores do sexo feminino
FEP56*	Idade média dos professores
FEP57*	Renda Familiar média dos professores (salários mínimos)
FEP58*	Tempo de exercício médio na ocupação

3.3) Descrição Quantitativa dos registros resultantes em cada um dos arquivos

A descrição a seguir refere-se aos três arquivos resultantes do processo de preparação feito a partir dos quatro arquivos fornecidos pelo Projeto GERES.

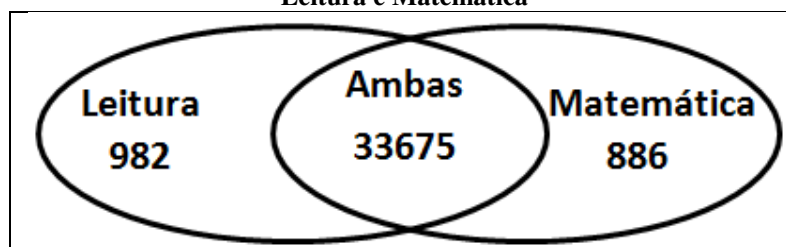
A Tabela 12 apresenta o número de registros de cada um desses arquivos.

Tabela 12- Número de registros nos arquivos para análise no HLM

Arquivo	Área do conhecimento	Registros
Nível 1 - Medidas Cognitivas no Tempo	Leitura	34657
	Matemática	34561
Nível 2 - Alunos	Leitura	12180
	Matemática	12153
Nível 3 - Escolas	Leitura	229
	Matemática	230

Como as provas de Leitura e Matemática são aplicadas em dias diferentes, a não presença dos alunos interfere diretamente no número de registros apresentados. A Figura 12 apresenta o número de registros do arquivo do Nível 1 que possuem medidas apenas em Leitura, apenas em matemática e em ambos.

Figura 12 - Número de registros nos arquivos de Nível 1 de Leitura e Matemática



Como pode ser observado, aproximadamente 97% dos registros comuns em Leitura e em Matemática são comuns aos dois arquivos, ou seja, apenas 3% deles têm observações em apenas umas das áreas.

A Tabela 13 apresenta os resultados dessa comparação em cada um dos anos de escolarização.

Tabela 13- Número de registros comuns no Nível 1 em Leitura e Matemática

	Registros			% dos registros	
	só em leitura	em ambas	só em matemática	só em leitura	só em matemática
Ano 1	230	7023	196	96,8%	97,3%
Ano 2	235	6998	177	96,8%	97,5%
Ano 3	218	8413	192	97,5%	97,8%
Ano 4	299	11241	321	97,4%	97,2%

Com base nesses dados, constata-se que o número de alunos que possuem medidas de proficiências nas duas áreas é praticamente o mesmo nos diferentes anos de escolarização.

A Tabela 14 apresenta o número de escolas que permanecem na amostra após o tratamento que implicou a exclusão e alguns registros. Como pode ser observado, uma das escolas privadas do pólo de Campinas permanece na amostra em matemática, mas não em leitura.

Tabela 14 - Distribuição das escolas por polos e dependências administrativas

Pólo	Rede escolar (dependência administrativa)				Total
	Leitura / Matemática				
	Especial	Estadual	Municipal	Privada	
Belo Horizonte	1/1	20/20	20/20	18/18	59/59
Campinas	0/0	16/16	21/21	18/19	55/56
Campo Grande	0/0	18/18	20/20	12/12	50/50
Rio de Janeiro	8/8	0/0	33/33	24/24	65/65
Salvador	0/0	0/0	0/0	0/0	0/0
Total	9/9	54/54	94/94	72/73	229/230

Como as escolas do pólo Salvador não participaram da última onda de aplicação de testes, esse pólo ficou sem as informações de contexto dos alunos o que implicou a exclusão da amostra. A diferença entre o número de escolas em Leitura e em Matemática na amostra ocorreu devido à permanência uma escola privada do pólo de Campinas, provavelmente por conta da falta de alguma medida em leitura.

Outra diferença a ser destacada pode ser percebida na distribuição do número de registros nos diferentes anos de escolarização como apresenta a Tabela 15 na qual está registrado o número de alunos na amostra que participaram pelo menos de um ano, que tenha duas medidas consecutivas, na coleta de dados e que permitem o cálculo do valor agregado.

Tabela 15 - Número de alunos da amostra

Alunos com dados completos no:	Leitura		Matemática	
	Quantidade	%	Quantidade	%
Ano 1	107	0,9	105	0,9
Ano 2	53	0,4	56	0,6
Ano 1 e 2	408	3,3	412	3,4
Ano 3	15	0,1	3	0,1
Ano 2 e 3	3	0,1	2	0,1
Ano 1, 2 e 3	54	0,4	13	0,1
Ano 4	2416	19,9	2392	19,2
Ano 1 e 4	565	4,7	583	4,8
Ano 3 e 4	1844	15,1	1895	15,6
Ano 2, 3 e 4	596	4,9	586	4,9
Ano 1, 2, 3 e 4	6119	50,2	6106	50,3
Total	12180		12153	

Dessa tabela pode-se também constatar que, dentro da amostra usada nos modelos, aproximadamente metade dos alunos se transferiram de ou para escola não participante do Projeto GERES no decorrer dos primeiros quatro anos de escolarização.

Outro aspecto importante a ser destacado é que, com a exclusão do pólo Salvador, 6613 dos 35538 registros de alunos são excluídos, restando 28925 registros, implicando também a redução de 312 para 257 escolas.

Assim, a amostra usada no modelo é composta por aproximadamente 42% (12180 alunos para Leitura e 12153 alunos para Matemática) dos registros de alunos e com 89% (229 escolas para Leitura e 230 para Matemática) do total fornecido pelo Projeto GERES, excluindo as informações do pólo Salvador.

Dessa maneira, fica claro que os arquivos gerados são uma parte da amostra de dados fornecida pelo Projeto GERES, dado que mais de 50% dos alunos dos quatro polos são excluídos. A Tabela 16 permite a comparação entre a média e o desvio padrão das proficiências em Leitura de cada uma das cinco ondas, estimados a partir da base de dados fornecida pelo projeto GERES, e as proficiências iniciais e finais de cada ano de escolarização, estimados a partir da amostra de dados a serem usadas nos modelos multinível, resultante do tratamento de dados. Essa comparação tem por objetivo verificar em que medida as proficiências finais e iniciais tem o mesmo comportamento estatístico da amostra de que se originaram.

Tabela 16 - Comparação das proficiências em Leitura

Proficiência	Número	Mínimo	Máximo	Média	Desvio Padrão
Onda 1	15596	47,8	159,9	107,3	26,5
Inicial Ano 1	7253	47,8	159,9	111,6	25,4
Onda 2	15934	68,6	183,9	125,7	23,2
Final Ano 1	7253	68,7	183,9	130,2	22,0
Inicial Ano 2	7233	68,6	183,9	129,9	22,1
Onda 3	16312	85,7	214,5	140,0	27,3
Final Ano 2	7233	85,7	214,5	146,9	26,3
Inicial Ano 3	8631	85,7	214,5	145,0	26,3
Onda 4	18024	87,8	214,2	153,1	27,1
Final Ano 3	8631	87,8	214,2	160,5	25,4
Inicial Ano 4	11540	87,8	214,2	157,2	25,6
Onda 5	17557	89,4	223,0	164,7	25,3
Final Ano 4	11540	90,2	223,0	169,5	23,7

Esta tabela possibilita estimar o valor agregado médio em cada um dos anos de escolarização, tomando como referência duas distintas amostras:

- a) pela diferença entre as médias das proficiências de duas ondas sucessivas;
- b) pela diferença entre as médias das proficiências finais e iniciais.

No caso mais extremo, se constata que a maior diferença ocorre entre as médias das proficiências medidas na Onda 3 (140,0) e na Onda 4 (153,1) indica que no terceiro ano de escolarização é agregado um valor médio de 13,1 unidades. Este é um valor bastante próximo de 15,5 unidades quando considerada a diferença entre as proficiências final e inicial, neste mesmo ano de escolarização. No entanto, a diferença entre a média das proficiências da Onda 4 e média da proficiência final do Ano 3 é de 7,4, evidenciando que a exclusão dos registros altera a significativamente a média.

O teste T realizado com as amostra evidencia que a diferença entre as médias é significativa.

Para a mesma análise, a Tabela 17 traz os valores estimados da média e o desvio padrão das proficiências em Matemática, que levaram à mesma constatação que os valores em Leitura.

Tabela 17 - Comparação das proficiências em Matemática

Proficiência	Número	Mínimo	Máximo	Média	Desvio Padrão
Onda 1	15625	19,5	173,0	107,6	29,9
Inicial Ano 1	7219	19,5	173,0	112,5	27,6
Onda 2	15869	54,7	217,1	136,3	32,7
Final Ano 1	7219	54,7	217,1	142,0	31,2
Inicial Ano 2	7175	54,7	217,1	141,7	31,1
Onda 3	16290	39,1	290,4	154,0	57,2
Final Ano 2	7175	39,1	290,4	168,7	54,8
Inicial Ano 3	8605	39,1	290,4	164,8	54,6
Onda 4	17944	45,5	356,0	193,1	63,0
Final Ano 3	8605	45,5	356,0	209,0	59,8
Inicial Ano 4	11562	45,5	356,0	202,6	59,8
Onda 5	17460	57,1	396,5	235,2	65,0
Final Ano 4	11562	57,4	396,5	247,0	62,9

No entanto, é importante salientar que essas médias são alteradas no processamento com os modelos multinível, uma vez que as variáveis são inseridas no modelo com pesos amostrais associados a cada aluno altera.

3.4) Outras observações

Estes procedimentos de redução de variáveis foi realizado independentemente das escalas e constructos propostos pelo Projeto GERES, uma vez que eles foram divulgados apenas em 2011, quando este trabalho já estava concluído. Certamente há algumas perdas com esses processo, no entanto, duas vantagens podem ser destacadas.

A primeira delas é comparando os resultados apresentados neste capítulo com os descritos em Brooke e Bonamino (2011) pode-se perceber que, em linhas gerais, há grande semelhança nos constructos encontrados, dando assim maior confiabilidade a ambos os estudos.

A segunda vantagem está ligada ao fato de que, como o processo desenvolvido, desvelou uma série de processos inerentes a essa ferramenta estatística, explicitando assim, as informações dessa caixa-preta, que consistia essa fase da pesquisa. Esses processos evidenciam algo transparente quando se recebe os resultados das avaliações educacionais, tais como a subjetividade inerente ao processo e outros procedimentos técnicos tais como a imputação de dados ou o método de redução das variáveis.

Ficou também evidente a necessidade de um futuro aprofundamento acerca da interdependência entre os fatores, pois se o foco está na busca de fatores que interferem no

desempenho do aluno através da proficiência espera-se que o fator estudado apresente alta correlação com as proficiências, porém baixa correlação entre os fatores o que nem sempre ocorre, principalmente quando se inclui no estudo um número muito grande de fatores. O método de análise de componentes principais corrige esta questão para cada construto, pois a mesma foi aplicada separadamente. Como são muitos construtos considerados separadamente, o problema pode persistir.

De forma geral, a análise de componentes principais caracteriza-se com um potente instrumento na determinação de fatores, possibilitando uma simplificação bastante útil da informação contida nos questionários de contexto e que a composição dos fatores passaram pelo teste de validade de construto. Fato que nos habilita generalizar os resultados para a população estudada.

De forma mais específica observa-se que muitos construtos e fatores poderiam ser melhorados. Os construtos em cada nível que se pretende estudar poderiam ser mais bem definidos para que o conjunto de questões (variáveis) possa captar melhor o que se pretende medir. Desta forma se evitaria variáveis com baixa comunalidade, (proporção da variância de cada variável explicada pelos fatores comuns). Outra medida a ser observada seria estabelecer um número mínimo de variáveis para cada fator dentro do construto, isto melhoraria alguns indicadores como alfa de Cronbach

Também é importante destacar que os resultados em pesquisas de eficácia e equidade escolar, assim como qualquer outro que requerem esses procedimentos, não dependem apenas dos modelos multinível usados para a análise. Ao contrário dependem significativamente de como as variáveis de contexto são tratadas.

Mais claramente, os resultados dessas pesquisas dependem diretamente da escolha feitas por quem realiza os procedimentos, podem ocorrer erros acidentais e até mesmo intencionalmente tendenciosos, sobretudo quando os resultados são destinados para uso meritocráticos.

Assim, é de fundamental importância que todo e qualquer estudo dessa natureza tenha também divulgado relatórios técnicos e além dito sejam, de alguma forma, auditados evitando, assim, esses dois possíveis tipos de erros que podem levar a decisões menos acertadas e até, muitas vezes, injustas.

CAPÍTULO 4 – UM MODELO PARA A COMPLEXIDADE

Neste capítulo são apresentados os procedimentos usados para o ajuste dos modelos que permitiram encontrar dentre um conjunto de variáveis explicativas, composto por 11 associadas ao perfil do aluno e por 59 associadas às características da escola, quais são estatisticamente significantes, a cada ano de escolarização, e considerando a influência de todas simultaneamente.

A lógica deste modelo, assim como os conceitos envolvidos e premissas que a orientam, é apresentada detalhadamente, com uma linguagem menos técnica, no Capítulo 2.

O resultado do ajuste são os modelos compostos pelas variáveis com significância estatística, isto é, por aquelas que têm a capacidade de explicar a variância da proficiência final. Os procedimentos para esse ajuste foram feitos separadamente para Leitura e Matemática, e seguem a seguinte sequência:

- análise do Modelo Nulo, tendo apenas a proficiência final como variável resposta;
- ajuste do Modelo no Nível 1, com a inserção das variáveis explicativas inerentes às medidas cognitivas no tempo;
- ajuste do modelo Nível 2, com a inserção das variáveis explicativas correspondentes ao perfil dos alunos;
- ajuste do modelo Nível 3, com a inserção das variáveis explicativas relacionadas ao âmbito escolar.
- análise do efeito de cada variável explicativa sobre a proficiência final, em cada um dos diferentes anos de escolarização.

Para essa análise, as premissas para a elaboração desse modelo, e conseqüentemente as decisões metodológicas deles decorrentes, têm implicação direta na identificação desses fatores. Por isso, mesmo que já apresentadas no Capítulo 2, são resgatadas aqui, uma vez que especificam aquilo que é esperado do modelo para o cumprimento dos objetivos desta pesquisa. Tais premissas fazem com que o modelo:

- a) seja multinível, para atender aos pressupostos para o uso da análise de regressão linear;
- b) tenha até três níveis hierárquicos, limite do programa computacional selecionado para o estudo;

c) explore os conceitos de valor agregado e de efeito-escola aproveitando a disponibilização de dados longitudinais e, para tal, buscar eliminar as diferenças entre as proficiências dos alunos antes de sua entrada na escola;

d) considere o peso amostral durante o processamento, para que os resultados encontrados possam ser usados de forma inferencial para as demais escolas que compõem o universo;

e) aproveite a maior quantidade possível de dados disponibilizados pelo Projeto GERES, para evitar perdas adicionais àquelas inerentes à mobilidade escolar captada pelos estudos longitudinais;

f) mantenha os alunos que não permaneceram na mesma escola, mas permaneceram em escolas participantes do Projeto GERES;

g) permita a inserção do maior número possível de variáveis explicativas para considerar a complexidade do fenômeno educacional;

h) processe os dados dos diferentes anos de escolarização em um único modelo, mas que permita a identificação dos fatores e da influência de cada um deles, nos distintos anos de escolarização;

i) aponte a influência dos fatores identificados na proficiência final em alunos com diferentes proficiências iniciais;

j) trabalhe com as proficiências de leitura e matemática separadamente dado que as proficiências não pertencem à mesma escala.

Tais premissas atribuem, inevitavelmente, significativa complexidade ao modelo e, por conseguinte, também ao trabalho para o seu ajuste, que é descrito detalhadamente a seguir.

4.1) Procedimentos para o ajuste dos modelos lineares hierárquicos

Os procedimentos de ajuste do modelo consistem basicamente na identificação de um conjunto de variáveis explicativas que minimizam, de forma significativa, a variância da variável resposta. Essencialmente, consistem na inserção dessas variáveis no modelo e na decisão por sua manutenção, tomada em função da diminuição da variância da variável resposta que ela provoca. Quando o *p-value* indica que essa variável é significativa, ela é mantida no modelo e, caso contrário, é excluída.

Nesta pesquisa, as variáveis explicativas são consideradas significantes tanto no aspecto estatístico, em um primeiro momento, quanto na perspectiva pedagógica (quando tem sentido interpretativo) são os fatores que interferem na proficiência final em Leitura e em Matemática, num momento posterior. Conseqüentemente, quando o modelo estiver ajustado, significa que foram encontradas um conjunto de variáveis explicativas que minimizam a variabilidade das proficiências finais de cada um dos anos de escolarização e, portanto, estão identificados um conjunto de fatores que interferem nessa proficiência final.

A escolha desses passos foi orientada por Snijders e Bosker (1999), Raundenbush e Bryk (2002), Hox (2010), Lee (2008) e Laros e Marciano (2008), dentre outros. Os trabalhos desses autores, mesmo sendo exemplos de modelos lineares hierárquicos de apenas dois níveis, mostram que a análise multinível deve começar pelo Modelo Nulo, para determinar a partição da variância total da variável resposta entre os diferentes níveis e, posteriormente, sejam inseridas as variáveis explicativas em cada um dos níveis. Esse processo deve ser iniciado por um nível inferior, partindo para o superior, somente depois que este foi estabilizado. Isto é, somente depois de que as variáveis explicativas de um nível estejam definidas deve começar a inserção das variáveis do nível imediatamente superior.

Segundo Lee (2008, p. 281),

Existem três fases na construção de um modelo típico de MLH para o estudo do efeito escola. Na primeira fase, o pesquisador divide a variância de uma variável dependente, que é medida em alunos individuais, em duas partes – a proporção da variância encontrada entre os alunos de uma mesma escola (compartilhadas dentro de várias escolas) e a proporção da variância na variável dependente que é encontrada sistematicamente entre as escolas. [...] A segunda fase em uma análise MLH de efeitos da escola envolve a estimativa de um modelo dentro da escola ou de modelo Nível 1. Nesta fase, investigamos as características de alunos individuais que estão associadas ao resultado. [...] Somente na fase final (ou seja, no modelo MLH de Nível 2) estimamos os efeitos da escola. Neste nível, o resultado, ajustado para as características do aluno no modelo de Nível 1, é explorado como uma função das características escolares.

Proposta análoga é feita por Laros e Marciano (2008, p. 268)

Passo 1. Analisa-se um modelo sem nenhuma variável explicativa. Esse modelo, dito modelo somente de intercepto ou modelo vazio [...] O modelo vazio é útil porque proporciona uma estimativa da correlação intraclasse. [...] Passo 2. Analisa-se um modelo com todas as variáveis explicativas fixas do nível mais baixo. [...] A decisão de inserir primeiramente as variáveis do nível mais baixo deve-se ao maior número de observações disponíveis neste nível, gerando coeficientes mais acurados. [...] Passo 3. Acrescentam-se as variáveis explicativas do nível da escola.

Após a análise do Modelo Nulo, foram inseridas as variáveis explicativas no Nível 1 (Medidas Cognitivas no Tempo), por ser o nível com a maior variância e também por definir a reta de regressão de cada ano de escolarização. Com isso, é criado no modelo, um parâmetro que define o intercepto e outro para a inclinação da reta (coeficiente de regressão) para cada ano de escolarização.

Os dois passos seguintes, quando foram inseridas as variáveis explicativas nos demais níveis, mostraram-se problemáticos. Neles foi constatado que a ordem pela qual elas são inseridas levava a diferentes resultados. Isto é, diferentes sequências de inserção de variáveis explicativas resultavam em modelos compostos por diferentes variáveis explicativas, mas com a mesma capacidade explicativa.

Diante dessa problemática, a opção por priorizar a inserção das variáveis consideradas pela literatura existente como sendo as mais significativas, também teve que ser descartada. Isso porque, dadas as características do modelo, que o difere dos demais modelos existentes (tais como a regressão múltipla, ou curvas de crescimento, por exemplo), leva à necessidade de se verificar se o comportamento dessas variáveis é compatível ou não com tais literaturas.

O exemplo mais notável para esta situação é a variável explicativa que controla o nível socioeconômico. Como neste estudo a proficiência inicial foi controlada como variável explicativa no Nível 1 e a sua correlação com o nível socioeconômico é alta, as diferenças entre os alunos devido ao nível socioeconômico são empiricamente anuladas na linha de base do estudo. Daí a necessidade de se verificar se a variável que controla o nível socioeconômico continua a influenciar após a entrada do aluno na escola.

Diante desse impasse, optou-se por inserir todas as variáveis explicativas e verificar, em primeira instância, quais delas se mostrariam mais significantes tomando como critério apenas sua significância estatística. Ou seja, optou-se pelo procedimento *backward*³⁵, no qual todas as variáveis são inseridas no modelo e, posteriormente, são sucessivamente excluídas aquelas estatisticamente menos significantes, em função do seu *p-value*.

35 Draper e Smith (1981) apresentam dois métodos para a inserção de variáveis. O primeiro deles é o *forward*, que consiste em colocar as variáveis explicativas, uma a uma, para verificar a sua significância e depois são colocadas todas juntas para verificar a significância de todas elas no modelo. A outra forma é o método *backward*, no qual são colocadas todas as variáveis no modelo e depois são excluídas, uma a uma, também em função de sua significância estatística.

Assim, foi eliminado o efeito da ordem de inserção das variáveis, deixando para ser excluídas as variáveis que apresentavam *p-value* menos aceitáveis, e considerando a interação entre todas as variáveis inseridas no modelo.

Com esses procedimentos, foi possível verificar, em cada um dos quatro anos de escolarização, quais eram as variáveis explicativas que se mostravam significantes, tendo como critério apenas os resultados numéricos, para posterior análise de seu significado, relevância pedagógica e sua associação com a proficiência final. Inicialmente foram inseridas todas as variáveis explicativas do Nível 2, que depois de estabilizado, foram inseridas as do Nível 3.

Como último passo, depois de ajustados os modelos, foram analisados os efeitos de cada uma das variáveis explicativas mantidas no modelo, descritos no próximo capítulo. Essa análise foi feita considerando três diferentes proficiências iniciais de cada um dos anos de escolarização: média (proficiência média), baixa (proficiência média menos um desvio padrão) e alta (proficiência média mais um desvio padrão).

Todos esses cinco passos foram feitos separadamente para as proficiências em Matemática e em Leitura. Ao final desta etapa foram identificadas as variáveis que, apesar de significantes, teriam pouco impacto na proficiência final.

4.2) O modelo para estudo com base na proficiência em Leitura

Na sequência são apresentados os quatro modelos ajustados com base nas proficiências em Leitura, seguindo cada uma das etapas descritas no Capítulo 2, a lembrar: O Modelo Nulo, Modelo de Nível 1 que busca diminuir variância devido às diferenças das proficiências nos diferentes anos de escolarização, Nível 2 que conta com a inserção das variáveis associadas ao perfil do aluno e o último, o Modelo de Nível 3 quando são inseridas as variáveis associadas às características da escola.

4.2.1) Modelo Nulo para Leitura

O resultado do processamento do Modelo Nulo, composto pelas equações a seguir, encontra-se no relatório de saída do HLM³⁶ (Anexo 03), cujos valores foram estimados com base nas seguintes equações³⁷:

Nível 1 – Medidas Cognitivas no Tempo:

$$PR_FI_{ijk} = \pi_{0jk} + e_{ijk}, \text{ com } e_{ijk} \sim \text{NID}(0, \sigma^2).$$

Nível 2 – Alunos:

$$\pi_{0jk} = \beta_{00k} + r_{0jk}, \text{ com } r_{0jk} \sim \text{NID}(0, \tau_{\pi 00})$$

Nível 3 – Escolas:

$$\beta_{00k} = \gamma_{000} + u_{0jk}, \text{ com } u_{0jk} \sim \text{NID}(0, \tau_{\beta 000})$$

Os resultados das estimativas feitas com as proficiências dos alunos da amostra GERES usando este modelo estão descritas no Quadro 17 e permitem identificar a variabilidade de cada um dos níveis, cuja partição é definida pelas Equações [01], [02] e [03] descritas no Capítulo 2.

Quadro 17 - Resultados do Modelo Nulo em Leitura

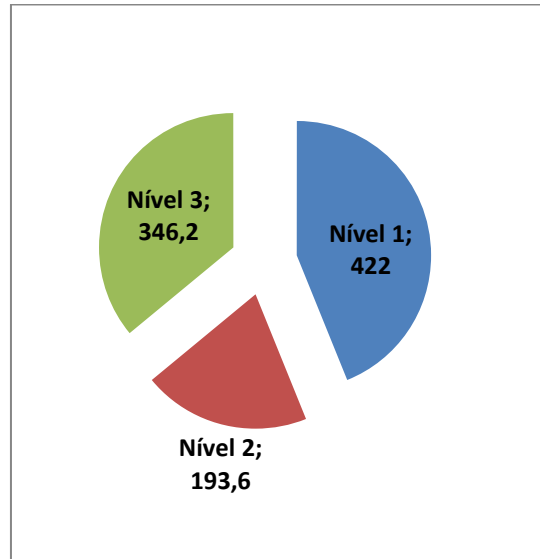
Resultados do relatório de saída		Valor estimado	Erro Padrão	p-value
Efeitos fixos	π_{0jk}	153,0	4,0	<0,001
Efeitos aleatórios	$\sigma^2 =$	Componente de Variância		
	$\tau_{\pi 00} =$	422,0		<0,001
	$\tau_{\beta 000} =$	193,6		<0,001
Deviance				318744
Número de parâmetros estimados				4

No Gráfico 12 está representado a proporção da variabilidade de cada um dos níveis. Desta forma podemos concluir que, da variância total das proficiências finais constantes do banco de dados ($422,0 + 193,6 + 346,2 = 961,8$), 43,9% (422,0) deve-se às diferenças de medidas entre os anos de escolarização, 20,1% (193,6) deve-se às diferenças entre alunos e 36,0% (346,43) deve-se às diferenças entre as escolas.

36 Neste trabalho, os coeficientes encontrados no relatório não são copiados para o texto devido à grande quantidade de informações, que aqui são sintetizadas visando o melhor entendimento do leitor.

37 Não haverá a preocupação em definir cada uma das variáveis e coeficientes que compõem as equações dos modelos apresentados a partir desta parte do texto até o momento em que tais definições se tornam importantes para o entendimento do modelo. Essa supressão das definições é possível por já terem sido feitas no Capítulo 2.

Gráfico 12 - Partição da variância entre os três níveis em Leitura



Se comparada apenas as variâncias que se deve aos alunos e às escolas, constata-se que 39,9% da variabilidade dos dados podem ser explicadas pelas diferenças entre os perfis dos alunos enquanto que 64,1% podem ser explicados pelas características das escolas.

Considerando que (i) é esperada a diferença entre as proficiências finais no decorrer dos quatro anos de escolarização, que (ii) há escolas que pertencem a diferentes redes de ensino nas quais os alunos não são alocados aleatoriamente nas escolas, mesmo nas escolas públicas, há famílias que selecionam a escola, quando no caso contrário, mesmo que de forma branda, as escolas selecionam seus alunos, não sendo exclusivamente em função do seu nível socioeconômico, o resultado da partição da variância encontrada no Modelo Nulo está dentro das expectativas.

Esse resultado também indicou a possibilidade da continuação da exploração do modelo, buscando variáveis que pudessem explicar a variância encontrada nos três níveis, uma vez que as três primeiras hipóteses descritas no Capítulo 2 foram satisfeitas. (Hipótese 1: Há variabilidade significativa entre as medidas da proficiência dos alunos nos diferentes anos de escolarização; Hipótese 2: Há variabilidade significativa nas medidas da proficiência média entre os alunos de uma mesma escola nos diferentes anos de escolarização; Hipótese 3: Há variabilidade significativa entre as proficiências médias das diferentes escolas nos diferentes anos de escolarização).

Como já apresentado no Capítulo 2, a *deviance* é um teste estatístico que verifica o ajuste global do modelo aos dados. Nesse Modelo Nulo, para a estimação dos 4 parâmetros (o intercepto π_{0jk} , a variância dos resíduos a nível de escola - $\tau_{\pi 00}$, a variância dos resíduos a nível de alunos - $\tau_{\beta 000}$ e a variância dos resíduos a nível de ano de escolarização do aluno - σ^2) o valor da deviance é de 318744.

4.2.2) Modelo de Nível 1 para Leitura – Medidas Cognitivas no Tempo

Neste modelo foram introduzidas como variáveis explicativas:

- as três variáveis indicadoras (T2, T3 e T4) que identificam os anos de escolarização, cujos coeficientes indicam os acréscimos no intercepto em função de cada um desses anos. Quando as estimativas forem positivas indicam que o intercepto terá um valor superior ao do primeiro ano, em caso contrário, evidentemente, esse valor será inferior;
- a proficiência inicial (PR_IN), cujo coeficiente indicará o quanto a proficiência final é aumentada para cada unidade de proficiência inicial;
- as três variáveis indicadoras PR_INxT2, PR_INxT3 e PR_INxT4, cujos coeficientes indicam o acréscimo que deve ser feito ao coeficiente de PR_IN e assim fazer a distinção entre os diferentes anos de escolarização; e,
- a variável *RET*, que identifica se o aluno ao participar do teste estava cursando a série que foi testada no respectivo ano de escolarização.

Com a inserção dessas variáveis, o Modelo de Nível 1 composto pelas seguintes equações, cujos coeficientes são estimados no relatório de saída do HLM da seguinte forma:

Nível 1 – Medidas Cognitivas no Tempo:

$$PR_FI_{ijk} = \pi_{0jk} + \pi_{1jk} \cdot T2_{ijk} + \pi_{2jk} \cdot T3_{ijk} + \pi_{3jk} \cdot T4_{ijk} + \pi_{4jk} \cdot PR_IN_{ijk} + \pi_{5jk} \cdot PR_INxT2_{ijk} + \pi_{6jk} \cdot PR_INxT3_{ijk} + \pi_{7jk} \cdot PR_INxT4_{ijk} + \pi_{8jk} \cdot RET_{ijk} + e_{ijk}, \text{ com } e_{ijk} \sim NID(0, \sigma^2).$$

Nível 2 – Alunos:

$$\pi_{0jk} = \beta_{00k} + r_{0jk}, \text{ com } r_{0jk} \sim NID(0, \tau_{\pi 00})$$

$$\pi_{1jk} = \beta_{10k}; \pi_{2jk} = \beta_{20k}; \pi_{3jk} = \beta_{30k}$$

$$\pi_{4jk} = \beta_{40k} + r_{4jk}, \text{ com } r_{4jk} \sim NID(0, \tau_{\pi 40})$$

$$\pi_{5jk} = \beta_{50k}; \pi_{6jk} = \beta_{60k}; \pi_{7jk} = \beta_{70k}; \pi_{8jk} = \beta_{80k}$$

Nível 3 – Escolas:

$$\beta_{00k} = \gamma_{000} + u_{0jk}, \text{ com } u_{0jk} \sim NID(0, \tau_{\beta 000})$$

$$\beta_{10k} = \gamma_{100}; \beta_{20k} = \gamma_{200}; \beta_{30k} = \gamma_{300}$$

$$\beta_{40k} = \gamma_{400} + u_{4jk}, \text{ com } u_{4jk} \sim NID(0, \tau_{\beta 400})$$

$$\beta_{50k} = \gamma_{500}; \beta_{60k} = \gamma_{600}; \beta_{70k} = \gamma_{700}; \beta_{80k} = \gamma_{800}$$

Os resultados destas estimativas estão descritos no Quadro 18, cujos valores foram retirados do relatório de saída do HLM que se encontra no Anexo 04.

Quadro 18 - Resultados Modelo de Nível 1 em Leitura

		Estimativa	E. P.	<i>p-value</i>	
		Efeitos Fixos	Intercepto	π_{0jk}	63,5
π_{1jk} (T2)	-27,2			3,3	<0,001
π_{2jk} (T3)	-8,8			2,8	0,002
π_{3jk} (T4)	-1,5			3,7	0,685
Coefic. de inclinação	π_{4jk} (PR_IN)		0,539	0,026	<0,001
	π_{5jk} (PR_INxT2)		0,264	0,026	<0,001
	π_{6jk} (PR_INxT3)		0,151	0,022	<0,001
	π_{7jk} (PR_INxT4)		0,113	0,026	<0,001
	π_{8jk} (RET)	5,6	1,5	0,001	
Efeitos Aleatórios	Componente de Variância			<i>p-value</i>	
	$\sigma^2 =$	186,5			
	$\tau_{\pi00} =$	97,4		0,010	
	$\tau_{\pi40} =$	0,00282		0,008	
	$\tau_{\beta000} =$	244,7		<0,001	
	$\tau_{\beta400} =$	0,00504		<0,001	
Deviance: 282226		Número de parâmetros estimados: 16			

O sinal positivo na estimativa para o coeficiente dos retidos é positivo porque nessa variável foi atribuído 0 quando o aluno é retido e 1 quando não foi retido no processo. Como pode ser observado nesse quadro, o *p-value* referente ao coeficiente π_{3jk} é maior que 0,05 que essa estimativa de -1,5 não pode ser considerada diferente de zero, ou seja, o intercepto das retas do primeiro e do quarto ano são os mesmos.

Essa constatação implica a necessidade de excluir a variável indicadora T4 do modelo e, diante dela, é importante retomar o fato de que os interceptos do primeiro e do quarto anos serem iguais não tem significado pedagógico, por não existir proficiência igual a zero. Igualmente importante é destacar que as retas não serão as mesmas, uma vez que os coeficientes de inclinação (π_{3jk} e $\pi_{3jk} + \pi_{6jk}$) são diferentes, além de que os intervalos de proficiência inicial desses dois anos de escolarização são distintos.

Após a exclusão da variável indicadora T4 o Modelo Nível 1 para leitura ficou assim ajustado:

Nível 1 – Medidas Cognitivas no Tempo:

$$PR_FI_{ijk} = \pi_{0jk} + \pi_{1jk} \cdot T2_{ijk} + \pi_{2jk} \cdot T3_{ijk} + \pi_{3jk} \cdot PR_IN_{ijk} + \pi_{4jk} \cdot PR_IN \times T2_{ijk} + \pi_{5jk} \cdot PR_IN \times T3_{ijk} + \pi_{6jk} \cdot PR_IN \times T4_{ijk} + \pi_{7jk} \cdot RET_{ijk} + e_{ijk}, \text{ com } e_{ijk} \sim NID(0, \sigma^2).$$

Nível 2 – Alunos:

$$\pi_{0jk} = \beta_{00k} + r_{0jk}, \text{ com } r_{0jk} \sim NID(0, \tau_{\pi00})$$

$$\pi_{1jk} = \beta_{10k}; \pi_{2jk} = \beta_{20k}$$

$$\pi_{3jk} = \beta_{30k} + r_{3jk}, \text{ com } r_{3jk} \sim NID(0, \tau_{\pi30})$$

$$\pi_{4jk} = \beta_{40k}; \pi_{5jk} = \beta_{50k}; \pi_{6jk} = \beta_{60k}; \pi_{7jk} = \beta_{70k}$$

Nível 3 – Escolas:

$$\beta_{00k} = \gamma_{000} + u_{0jk}, \text{ com } u_{0jk} \sim NID(0, \tau_{\beta000})$$

$$\beta_{10k} = \gamma_{100}; \beta_{20k} = \gamma_{200};$$

$$\beta_{30k} = \gamma_{300} + u_{3jk}, \text{ com } u_{3jk} \sim NID(0, \tau_{\beta300})$$

$$\beta_{40k} = \gamma_{400}; \beta_{50k} = \gamma_{500}; \beta_{60k} = \gamma_{600}; \beta_{70k} = \gamma_{700}$$

O Quadro 19 traz as estimativas desses coeficientes.

Quadro 19 - Resultados Modelo de Nível 1 em Leitura

Efeitos Fixos	Intercept to		Estimativa	E. P.	p-value
			π_{0jk}	63,6	3,4
		π_{1jk} (T2)	-26,5	2,7	<0,001
		π_{2jk} (T3)	-7,9	2,7	0,002
Coefc. de inclinação		π_{3jk} (PR_IN)	0,545	0,018	<0,001
		π_{4jk} (PR_INxT2)	0,258	0,021	<0,001
		π_{5jk} (PR_INxT3)	0,144	0,020	<0,001
		π_{6jk} (PR_INxT4)	0,102	0,0586	<0,001
		π_{7jk} (RET)	5,7	1,5	0,001
Efeitos Aleatórios	Componente de Variância				p-value
		$\sigma^2 =$	186,3		
		$\tau_{\pi00} =$	98,8		0,010
		$\tau_{\pi40} =$	0,00285		0,007
		$\tau_{\beta000} =$	238,5		<0,001
	$\tau_{\beta400} =$	0,006979		<0,001	
Deviance: 282181		Número de parâmetros estimados: 15			

Para o cumprimento da próxima etapa, que tem por objetivo identificar as variáveis que são significativas e por perceber que os efeitos aleatórios indicam que há variabilidade a ser explicada, a variável T4 foi mantida no modelo para justamente possibilitar a inserção das variáveis nesse ano, mas no Modelo de Nível 1, descrito mais adiante, usado para a análise individual dos efeitos das variáveis, ela foi excluída.

Em ambos os quadros fica caracterizado que nos diversos anos de escolarização as proficiências finais são maiores que as proficiências iniciais e que os dados se ajustam

adequadamente à regressão linear. Além disso, está evidenciado que há diferença significativa nas proficiências entre os diferentes anos de escolarização, confirmando, assim, a Hipótese 4 (As medidas da proficiência dos alunos em função dos anos de escolarização podem ser representadas linearmente) e também a Hipótese 5 (Há variabilidade significativa nas proficiências finais dos alunos nos diferentes anos de escolarização, sendo possível determinar o quanto desta variabilidade foi explicada pelas variáveis controladas).

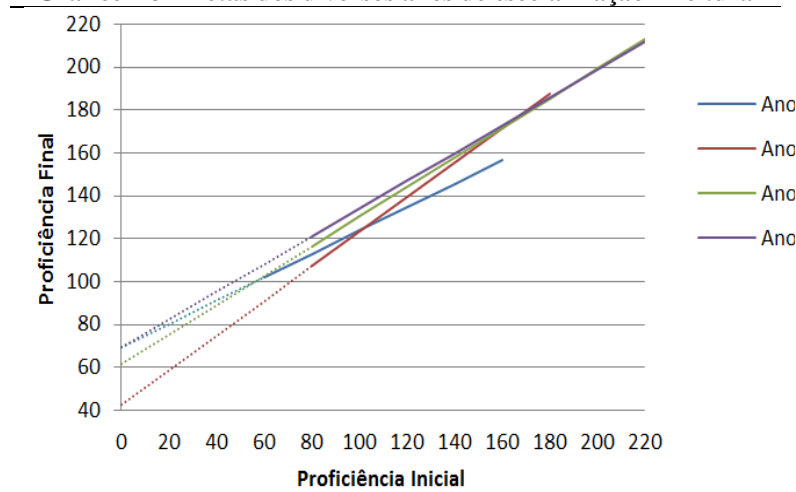
Com base nas estimativas apresentadas no Quadro 19, os interceptos e os coeficientes de inclinação para cada um dos anos está descrito no Quadro 20.

Quadro 20 - Equação de regressão de cada ano em Leitura

Ano		Intercepto	Coefficiente de inclinação
1	Composição	$\pi_{0jk} + \pi_{7jk}$	π_{3jk}
	Valor	69,3	0,545
2	Composição	$\pi_{0jk} + \pi_{1jk} + \pi_{7jk}$	$\pi_{3jk} + \pi_{4jk}$
	Valor	42,8	0,803
3	Composição	$\pi_{0jk} + \pi_{2jk} + \pi_{7jk}$	$\pi_{3jk} + \pi_{5jk}$
	Valor	61,4	0,689
4	Composição	$\pi_{0jk} + \pi_{7jk}$	$\pi_{3jk} + \pi_{6jk}$
	Valor	69,3	0,647

O gráfico 13 representa a equação de regressão de cada um dos anos de escolarização³⁸.

Gráfico 13 - Retas dos diversos anos de escolarização - Leitura



38 Este é o gráfico é produzido com as estimativas reais encontradas para proficiência em leitura, correspondente à Figura 7 que mostra teoricamente essas retas e também ao Gráfico 8 com dados hipotéticos, apresentados no Capítulo 2

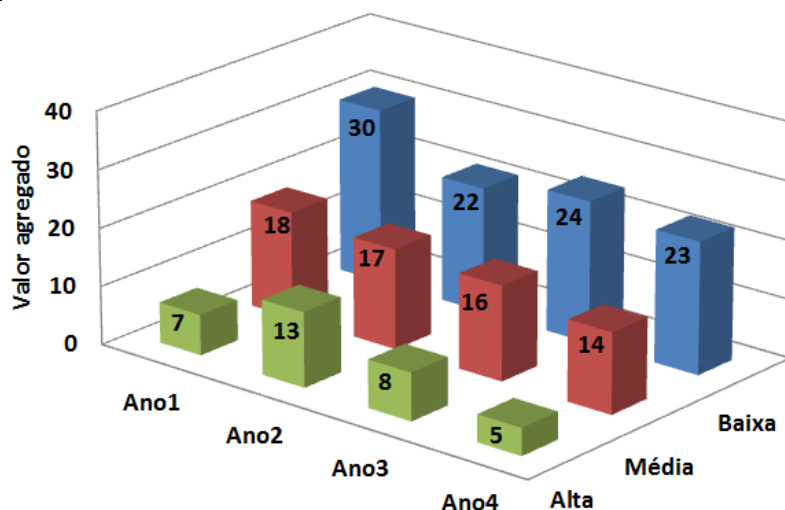
Através da diferença entre a média das proficiências finais e a média das proficiências iniciais, o valor agregado médio seria de 18 para o primeiro ano de escolarização, 17 para o segundo, 15 para o terceiro e 12 para o quarto. Mas, com essa modelagem torna-se possível analisar o comportamento do valor agregado na proficiência, em cada um dos anos de escolarização, para os diferentes alunos que compõem a amostra em função de sua proficiência inicial, como mostra a Tabela 18.

Tabela 18 - Valor agregado em Leitura como diferença da proficiência final estimada e da inicial

		Proficiência				Valor agregado			
		Ano1	Ano2	Ano3	Ano4	Ano1	Ano2	Ano3	Ano4
Baixa	Inicial	86	108	119	132	30	22	24	23
	Final estimada	116	130	143	155				
Média	Inicial	112	130	145	157	18	17	16	14
	Final estimada	130	147	161	171				
Alta	Inicial	137	152	171	183	7	13	8	5
	Final estimada	144	165	179	188				

O Gráfico 14 representa esses valores.

Gráfico 14 - Valor agregado em Leitura por ano de escolarização em função da proficiência inicial



Para esta estimativa, sem a preocupação de estimar valores com um determinado intervalo de confiança, a proficiência inicial média foi obtida a partir da média das

proficiências iniciais de cada um dos anos de escolarização separadamente. A proficiência baixa foi colocada a um desvio padrão abaixo da média e a proficiência inicial alta.

Esses valores agregados estimados por meio das equações de regressão, cujos valores foram apresentados anteriormente no Quadro 20, mostram que, em média, os alunos com menor proficiência inicial têm um valor agregado estimado maior no período de escolarização do que aqueles com proficiência inicial alta.

4.2.3) Modelo de Nível 2 para Leitura – Alunos

Nesta etapa do processo, as variáveis do Nível 1 são mantidas e são inseridas as 11 variáveis explicativas investigadas do Nível 2 que estão vinculadas ao perfil do aluno.

Como descrito anteriormente, o ajuste do modelo foi feito por meio do método *backward* que consiste na inserção simultânea de todas as variáveis e, em seguida, são excluídas, sucessivamente, uma a uma, aquelas que se apresentam estatisticamente menos significantes até que todas fossem significantes.

São consideradas estatisticamente menos significantes aquelas que apresentam maior *p-value* no momento do processamento. O *p-value* mede a significância estatística dos parâmetros do modelo, ou seja, verifica o grau de correlação entre a variável explicativa e a proficiência final em leitura do aluno. Na medida em que o *p-value* diminui aumenta a correlação entre as duas variáveis. Neste trabalho, se adotou como significativo quando o teste apresentava *p-value* menor que 0,05. Há casos em que foram tolerados valores marginais, para um *p-value* <0,10, quando a *deviance* diminuía significativamente, mostrando que a capacidade explicativa do modelo é maior.

Ao final desse processo, o modelo ajustado (ANEXO 05) ficou composto pelas variáveis que se mostraram significante que estão indicadas no Quadro 21.

Quadro 21 - Variáveis significantes no Modelo Nível 2 - Leitura

Variável	Descrição	Intercepto					Coeficiente de inclinação			
		Ano1	Ano2	Ano3	Ano4	Ret	Ano1	Ano2	Ano3	Ano4
FA01	Sexo	*				*				
FA04	Indica se o aluno fez a pré-escola	*	**					**		
FA05	Indica se o aluno foi retido em uma das séries									
FA06	Nível socioeconômico do aluno				**		*	**	*	**
FA07	Estímulo extrínseco para estudo	*	*		**					**
FA08	Estímulo intrínseco para o estudo						*	**	**	**
FA09	Percepção dos alunos quanto ao incentivo do professor	*	**		*			**		**
FA10	Gosto pela escola		**					**		
FA11	Estímulo interno para lição de casa			**		*				
FA12	Meus professores pensam que sou um aluno lento			**			**			
FA13	Sou tipo de pessoa que sempre será capaz de conseguir, se tentar	*					***			
FA14	Algumas vezes eu sinto que não sou bom o bastante	**					**			***

Obs.: * indica *p-value* < 0,001; ** indica *p-value* < 0,05; *** indica *p-value* < 0,1

Essas variáveis são mantidas para o processo de inclusão das variáveis explicativas de Nível 3, descrito no próximo item.

O Quadro 22 apresenta os valores estimados para as variâncias, *deviance* e o teste quiquadrado. As demais estimativas do Modelo de Nível 2 encontra-se no Anexo 06.

Quadro 22 - Estimativas das variâncias e do *deviance* no Modelo Nível 2 de Leitura

Componente de Variância		<i>p-value</i>
$\sigma^2 =$	171,0	
$\tau_{\pi 00} =$	87,0	>0,35
$\tau_{\pi 40} =$	0,00248	>0,33
$\tau_{\beta 000} =$	175,3	<0,001
$\tau_{\beta 400} =$	0,00410	<0,001
Deviance: 279718	Número de parâmetros estimados: 49	
Teste quiquadrado: 3340,5	Graus de liberdade: 31	<0,001

Em uma análise global dos resultados do Modelo de Nível 2 verificou-se que a inclusão de variáveis explicativas ligadas ao perfil do aluno provocaram uma redução significativa na *deviance* em relação ao Modelo de Nível 1, ou seja, as variáveis explicativas do Nível 2 contribuíram para um maior ajuste do modelo aos dados.

Os resultados dos efeitos aleatórios do Modelo de Nível 2 também são menores do que aqueles encontrados do Modelo de Nível 1, o que indica que essas variáveis explicam parte da variância total da proficiência final a cada ano. Em especial, com as informações do Quadro 22, constata-se que:

- a redução de σ^2 foi de 8,3%, diminuindo o valor de 186,5 para 171,0;
- a redução de $\tau_{\pi 00}$ foi de 10,7%, diminuindo o valor de 97,4 para 87,0;
- a redução de $\tau_{\pi 40}$ foi de 12,1% , pois o valor diminuiu de 0,00282 para 0,00248;
- a redução de $\tau_{\beta 000}$ foi de 28,4%, pois o valor diminuiu de 244,7 para 175,3;
- a redução de $\tau_{\beta 400}$ foi de 18,7%, pois o valor diminuiu de 0,00504 para 0,0041.

Além disso, os valores de *p-value* > 0,05 associados às variâncias $\tau_{\pi 00}$ e $\tau_{\pi 40}$ mostram que não podem ser consideradas diferentes de zero e, portanto, significa que toda a variabilidade da proficiência final já foi explicada com as variáveis de Nível 2 inseridas no modelo.

Assim, com os resultados apresentados nesse Modelo constata-se que as variáveis explicativas do Nível 2 deste nível são capazes de explicar a variabilidade da proficiência final em cada um dos anos, confirmando, também, a Hipótese 5 (Há variabilidade significativa nas

proficiências finais dos alunos nos diferentes anos de escolarização, sendo possível determinar o quanto desta variabilidade foi explicada pelas variáveis controladas).

Todas essas análises foram feitas exclusivamente na perspectiva da estatística, de forma que essas variáveis, quando mantidas, significam que são fatores que interferem na proficiência final em leitura.

Em uma análise de seu significado pedagógico, pode-se afirmar que, dentre elas:

a) têm comportamento de acordo com o esperado as variáveis FA01, FA04, FA06, FA07, FA08 e FA11.

b) têm comportamento incoerente com o esperado as variáveis FA09 e FA13.

c) têm comportamento instável, dificultando a interpretação, as variáveis FA10, FA12 e FA14.

4.2.4) Modelo de Nível 3 para Leitura – Escolas

De maneira análoga ao que foi feito nos modelos anteriores, as variáveis explicativas consideradas significantes nos níveis 1 e 2 foram mantidas para a inserção de todas as variáveis associadas ao Nível 3. Seguindo o método *backward*, as variáveis menos significantes (segundo o *p-value* estimado) foram sendo retiradas até que permanecessem apenas aquelas com significância estatística. Como já sinalizado no capítulo anterior, ao se perceber que a ordem da inserção das variáveis interferem no resultado, optou-se por esse método, para que se verifique, quais são as variáveis que permaneceriam no modelo, independentemente do seu significado pedagógico, segundo os dados disponíveis, para posterior interpretação pedagógica das variáveis que permaneceram.

Os quadros a seguir apresentam as variáveis associadas às características da escola que se mostraram significantes. É importante lembrar que, quando uma variável é significativa no primeiro ano, seja no intercepto, seja na inclinação, significa que ela tem impacto em todos os anos de escolarização.

Quadro 23 - Variáveis significantes no Modelo Nível 3 - Leitura pertencentes ao Grupo 1

Variável	Descrição	Intercepto				Coeficiente de inclinação			
		Ano 1	Ano2	Ano3	Ano4	Ano 1	Ano2	Ano3	Ano4
FEA01	NSE médio da escola a partir do NSE dos alunos	*	**	**	**	*	***	**	**
FEE02	Conservação e limpeza supraestrutura	*	*	**	*	*	*	**	
FEE03	Conservação e limpeza área de alimentação		**	**			**		
FEE04	Espaço sala de leitura	*			**	*			**
FEE05	Condições da Biblioteca	**							
FEE06	Laboratório de informática		*						**
FEE07	Sala de artes e laboratório de ciências		*		*				*
FEE08	Disponibilidade de computador e internet para professor	*				*			
FEE09	Existência de equipamentos para fotocópia e escaneamento								**
FEE10	Existência de material concreto para matemática		***				***	**	
FEE11	Existência de terrário e aquário				**			*	**

Obs.: * indica p-value < 0,001; ** indica p-value < 0,05; *** indica p-value < 0,1

Todas as variáveis desse grupo tiveram alguma associação com na proficiência final, com destaque à variável (FEA01) que controla o nível socioeconômico médio da escola (MNSE) e à variável “Conservação e limpeza da supra estrutura” (FEE01) que influenciam em todos os anos de escolarização e de forma distinta em cada um deles.

Quadro 24 - Variáveis significantes no Modelo Nível 3 - Leitura pertencentes ao Grupo 2 (continua)

Variável	Descrição	Intercepto				Coeficiente de inclinação			
		Ano 1	Ano2	Ano3	Ano4	Ano 1	Ano2	Ano3	Ano4
FED12	Interferência por intimidação, violência, discriminação e drogas	**		*		**		*	***
FED13	Interferência do perfil do professor	**			***	**			**
FED14	Interferência do perfil do aluno, família	*	*	**	*	*	*	**	*
FED15	Interferência da orientação pedagógica			**				**	
FED16	Presença física do diretor								
FED17	Organização de eventos				*				*
FED18	Frequência de encontros com a comunidade				**			**	
FED19	Frequência da presença na comunidade		**	*			**	*	**

(conclusão)

Variável	Descrição	Intercepto				Coeficiente de inclinação			
		Ano 1	Ano2	Ano3	Ano4	Ano 1	Ano2	Ano3	Ano4
FED20	Proporção de diretores que possuem Ens. Sup. em Pedagogia		**				**		
FED21	Proporção de Diretores que possui outra atividade que contribui para a renda pessoal		**		**				
FED22	Proporção de diretores que trabalham em apenas uma escola	*		**	**	*		**	**
FED23	Carga horária de trabalho do(a) diretor(a) na escola	*			*	*	**		
FED24	Tempo que o diretor trabalha na escola ininterruptamente na função de direção		**	**			**		
FED25	Tempo que o(a) diretor (a) trabalha em anos na educação	**	**		**	**	**	**	**
FED26	Idade do(a) diretor(a) em anos		**	**			**		

Obs.: * indica p-value < 0,001; ** indica p-value < 0,05; *** indica p-value < 0,1

Neste grupo a variável “Presença Física do Diretor” (FED16) não teve qualquer influencia em todos os anos. Efeito contrário têm as variáveis “Interferência do perfil do aluno” (FED15) e “Tempo que o(a) diretor (a) trabalha em anos na educação“ (FED25) que influenciam em todos os anos de escolarização e de forma distinta em cada um deles.

Quadro 25 - Variáveis significantes no Modelo Nível 3 - Leitura pertencentes ao Grupo 3

(continua)

Variável	Descrição	Intercepto				Coeficiente de inclinação			
		Ano 1	Ano2	Ano3	Ano4	Ano 1	Ano2	Ano3	Ano4
FEP33	Frequência do dever de casa	**	*	**		*	**		*
FEP36	Acesso a atividades culturais	**			*	**			*
FEP37	Proposição de leitura	*			*	*			*
FEP38	Proposição de redação	*	**		**	*	**		**
FEP39	Proposição de cópia ou caligrafia	*	**	**	**	*	**	**	**
FEP40	Proposição de leitura individual é precedida da explicação do professor	*	**		**	**	**	**	
FEP41	Leitura em prática coletiva	**			**				**
FEP42	Leitura individual e produção de texto individual	**		***		**			
FEP43	Uso da metodologia de resolução de problemas	*		**				***	
FEP45	Adequação do livro ao ensino		*			**	*		**

(conclusão)

Variável	Descrição	Intercepto				Coeficiente de inclinação			
		Ano 1	Ano2	Ano3	Ano4	Ano 1	Ano2	Ano3	Ano4
FEP46	Grau de importância inerente a indicação externa do livro	*			*	*			*
FEP47	Tempo usado nas disciplinas relacionadas com as proficiências	*							*
FEP48	Tempo usado nas disciplinas não relacionadas com as proficiências	**	**		**	**	**		**
FEP49	Tempo usado em disciplinas de formação complementar			*	**		**	*	**
FEP52	Uso de recursos audiovisuais		**		**		**		
FEP53	Uso de laboratório e multimídia		**		**		**		**
FEP54	Uso de materiais concretos de matemática e geografia	**	**			**	**		

Obs.: * indica p-value < 0,001; ** indica p-value < 0,05; *** indica p-value < 0,1

Neste grupo a variável que controla a “Proposição de cópia ou caligrafia” (FEP39) é a única que influencia em todos os anos de escolarização. Mas todas as demais tem alguma associação com a proficiência final.

Quadro 26 – Variáveis significantes no Modelo Nível 3 – Leitura pertencentes ao Grupo 4

(continua)

Variável	Descrição	Intercepto				Coeficiente de inclinação			
		Ano 1	Ano2	Ano3	Ano4	Ano 1	Ano2	Ano3	Ano4
FEP27	Relação com o diretor								
FEP28	Boa interação entre a equipe				**				**
FEP29	Pouco trabalho coletivo	**	**		*	**	***		*
FEP30	Boa relação entre professores						*		
FEP31	Unidade de propósito entre os professores	*	*	*	**	*	*	*	**
FEP32	Violência no ambiente escolar	*	**	*	**		**	*	**
FEP34	Grau de expectativas negativa do aprendizado	*	**	*	**	*		*	*
FEP35	Grau de expectativas positiva do aprendizado	*	**		**	*	**		
FEP50	Interrupção das aulas por indisciplina		**						*
FEP51	Interrupção das aulas por necessidade administrativa	**		**	*				*
FEP55	Proporção de professores do sexo feminino		*		*				**
FEP56	Idade média dos professores								

(continuação)

Variável	Descrição	Intercepto				Coeficiente de inclinação			
		Ano 1	Ano2	Ano3	Ano4	Ano 1	Ano2	Ano3	Ano4
FEP57	Renda Familiar média dos professores (salários mínimos)			**			**	**	
FEP58	Tempo de exercício médio na ocupação	*	*	*	*	*	*	*	**
FEP59	Número médio de pessoas da família que mora com o professor								

Obs.: * indica p-value < 0,001; ** indica p-value < 0,05; *** indica p-value < 0,1

Neste grupo, os resultados mostraram que as variáveis “Relação com o diretor” (FED16 – apontadas sob a perspectiva do professor), “Idade média dos professores” (FEP56*) e “Número médio de pessoas da família que mora com o professor” (FEP59) não tiveram qualquer influencia em todos os anos. Efeito contrário tiveram as variáveis “Unidade de propósito entre os professores” (FEP31) e “Tempo de exercício médio na ocupação” (FEP58) que influenciam em todos os anos de escolarização e de forma distinta em cada um deles.

O Quadro 27 apresenta os valores estimados para as variâncias e a *deviance*, obtidos com o modelo, cujas equações estão no Anexo 06. As demais estimativas do Modelo de Nível 2 encontra-se no Anexo 05.

Quadro 27 - Estimativas das variâncias e do *deviance* no Modelo Nível 3 em Leitura

Componente de Variância		<i>p-value</i>
$\sigma^2 =$	167,6	
$\tau_{\pi00} =$	60,4	>0,38
$\tau_{\pi40} =$	0,00150	>0,30
$\tau_{\beta000} =$	10,6	<0,001
$\tau_{\beta400} =$	0,00044	<0,001
Deviance: 277884	Número de parâmetros estimados: 276	

Uma primeira constatação com os resultados apresentados neste quadro é que este modelo teve um melhor ajuste do que o Modelo de Nível 2. Um resultado esperado, dado o número de variáveis que foram inseridas.

Mas também foi notável a redução da variância total, sobretudo no Nível 3, onde as variáveis explicativas foram inseridas. Dos valores que constam do Quadro 27 podemos afirmar que:

- a redução de σ^2 foi de 2,0%, pois o valor diminuiu de 171,0 para 167,6;
- a redução de $\tau_{\pi 00}$ foi de 30,6%, pois o valor diminuiu de 87,0 para 60,4;
- a redução de $\tau_{\pi 40}$ foi de 39,5% , pois o valor diminuiu de 0,00248 para 0,00150;
- a redução de $\tau_{\beta 000}$ foi de 94,0%, pois o valor diminuiu de 175,3 para 10,6;
- a redução de $\tau_{\beta 400}$ foi de 89,3%, pois o valor diminuiu de 0,00410 para 0,00044.

Desta maneira, os modelos construídos a partir dos dados disponíveis, mostram que são confirmadas as Hipóteses levantadas e também já indicam quais são as variáveis explicativas que, de forma significativa, interferem na proficiência final em Leitura dos alunos em cada um dos anos de escolarização.

O Quadro 28 apresenta uma breve análise³⁹ sobre o comportamento das variáveis do Nível 3, associadas, às características das escolas, na perspectiva pedagógica.

Quadro 28 - Interpretação final das variáveis do Grupo 4 do Nível 3 em Leitura

Comportamento	Variáveis
Coerente como esperado	FEA01, FEE08, FEE09, FED19, FED21, FED24*, FEP28, FEP43, FEP48.
Incoerente com o esperado	FEE04, FEE05, FEE06, FED15, FED18, FEP30.
Instável, dificultando a interpretação	FEE02, FEE03, FEE07, FEE10, FEE11, FED12, FED13, FED14, FED17, FED20, FED22*, FED23*, FED25*, FED26*, FEP29, FEP31, FEP32, FEP33, FEP34, FEP35, FEP36, FEP36, FEP37, FEP38, FEP39, FEP40, FEP41, FEP42, FEP43, FEP44, FEP45, FEP46, FEP47, FEP49, FEP52, FEP53, FEP54, FEP51, FEP55*, FEP57*, FEP58*.

4.3) O modelo para estudo com base na proficiência em Matemática

A apresentação do processo de ajuste dos modelos usando a proficiência em matemática seguem os mesmos passos dos modelos que analisaram as variáveis explicativas com base nas proficiências em Leitura.

Uma diferença essencial é que nesta análise para a proficiência em matemática foram feitas duas distintas análises das variáveis de Nível 3, cujos motivos serão descritos de forma concomitante à descrição do processo.

Assim essa seção apresenta o ajuste do Modelo Nulo, do Modelo de Nível 1 e em seguida descreve o ajuste do Nível 3.

³⁹ O detalhamento dessas análises assim como o impacto de cada uma as variáveis serão descritas no próximo capítulo.

4.3.1) Modelo Nulo para Matemática

O Modelo Nulo no qual é colocada a proficiência final de matemática como variável resposta, composto pelas Equações [08], [09] e [107], é descrito no relatório de saída do HLM por:

Nível 1 – Medidas Cognitivas no Tempo:
 $PR_FI_{ijk} = \pi_{0jk} + e_{ijk}$, com $e_{ijk} \sim NID(0, \sigma^2)$. [08]

Nível 2 – Alunos:
 $\pi_{0jk} = \beta_{00k} + r_{0jk}$, com $r_{0jk} \sim NID(0, \tau_{\pi 00})$ [09]

Nível 3 – Escolas:
 $\beta_{00k} = \gamma_{000} + u_{0jk}$, com $u_{0jk} \sim NID(0, \tau_{\beta 000})$ [10]

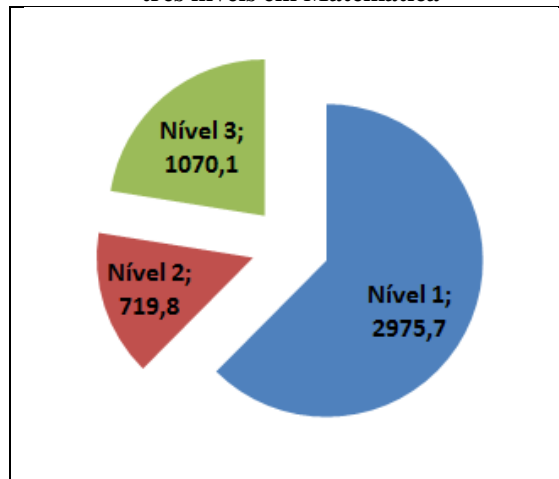
Os resultados das estimativas feitas com esse modelo estão descritas no Quadro 29 e permitem identificar a variabilidade de cada um dos níveis, cuja partição é definida pelas Equações [01], [02] e [03], apresentadas no Capítulo 2. Os valores desse Quadro são retirados do relatório HLM que se encontra no Anexo 07.

Quadro 29 - Resultados do Modelo Nulo em Matemática

Resultados do relatório de saída		Estimativa	Erro Padrão	p-value
Efeitos fixos	π_{0jk}	200,7	4,1	<0,001
Efeitos aleatórios	$\sigma^2 =$	Componente de Variância		
	$\tau_{\pi 00} =$	2975,7		<0,001
	$\tau_{\beta 000} =$	719,8		<0,001
Deviance				381628
Número de parâmetros estimados				4

O Gráfico 15 representa a proporção da variabilidade de cada um dos níveis e com esses resultados podemos concluir que a variância total das proficiências finais constantes do banco de dados ($2975,7 + 719,8 + 1070,1 = 4765,6$), 62,4% (2975,7) deve-se às diferenças de medidas entre os anos de escolarização, 15,1% (719,8) resultam das diferenças entre alunos e 22,5% (1070,1) são devidas às diferenças entre as escolas.

Gráfico 15 - Partição da variância entre os três níveis em Matemática



Em Matemática, a diferença entre as proficiências finais no decorrer dos quatro anos de escolarização representam aproximadamente 62%, ao mesmo tempo em que as diferenças entre as escolas são responsáveis por uma parte maior da variabilidade total que as diferenças entre os alunos e são confirmadas, também em matemática, as Hipóteses 1, 2 e 3.

Com essa partição das variâncias, comparando-as apenas entre ao que se deve ao perfil do aluno e a que se deve às características da escola, a distribuição fica de 40,2% e 59,8 %, respectivamente.

Nesse Modelo Nulo o valor da *deviance* é de 381628 com a estimação de 4 parâmetros.

4.3.2) Modelo de Nível 1 para Matemática – Medidas Cognitivas no Tempo

Assim como ocorreu com o modelo desenvolvido para Leitura, nesse modelo foram introduzidas como variáveis explicativas as três variáveis indicadoras (T2, T3 e T4); a proficiência inicial (PR_IN), as três variáveis indicadoras PR_INxT2, PR_INxT3 e PR_INxT4 e a variável RET⁴⁰.

Como as variáveis têm as mesmas denominações, as equações que compõem o modelo são:

40 As definições dessas variáveis já foram feitas no modelo que analisou os efeitos sobre a proficiência em leitura, no seção 2.2 desse capítulo.

Nível 1 – Medidas Cognitivas no Tempo:

$$PR_FI_{ijk} = \pi_{0jk} + \pi_{1jk} \cdot T2_{ijk} + \pi_{2jk} \cdot T3_{ijk} + \pi_{3jk} \cdot T4_{ijk} + \pi_{4jk} \cdot PR_IN_{ijk} + \pi_{5jk} \cdot PR_IN \times T2_{ijk} + \pi_{6jk} \cdot PR_IN \times T3_{ijk} + \pi_{7jk} \cdot PR_IN \times T4_{ijk} + \pi_{8jk} \cdot RET_{ijk} + e_{ijk}, \text{ com } e_{ijk} \sim NID(0, \sigma^2).$$

Nível 2 – Alunos:

$$\pi_{0jk} = \beta_{00k} + r_{0jk}, \text{ com } r_{0jk} \sim NID(0, \tau_{\pi 00})$$

$$\pi_{1jk} = \beta_{10k};$$

$$\pi_{2jk} = \beta_{20k};$$

$$\pi_{3jk} = \beta_{30k}$$

$$\pi_{4jk} = \beta_{40k} + r_{4jk}, \text{ com } r_{4jk} \sim NID(0, \tau_{\pi 40})$$

$$\pi_{5jk} = \beta_{50k};$$

$$\pi_{6jk} = \beta_{60k};$$

$$\pi_{7jk} = \beta_{70k};$$

$$\pi_{8jk} = \beta_{80k}$$

Nível 3 – Escolas:

$$\beta_{00k} = \gamma_{000} + u_{0jk}, \text{ com } u_{0jk} \sim NID(0, \tau_{\beta 000})$$

$$\beta_{10k} = \gamma_{100};$$

$$\beta_{20k} = \gamma_{200};$$

$$\beta_{30k} = \gamma_{300}$$

$$\beta_{40k} = \gamma_{400} + u_{4jk}, \text{ com } u_{4jk} \sim NID(0, \tau_{\beta 400})$$

$$\beta_{50k} = \gamma_{500};$$

$$\beta_{60k} = \gamma_{600};$$

$$\beta_{70k} = \gamma_{700};$$

$$\beta_{80k} = \gamma_{800}$$

O Quadro 30 apresenta as estimativas encontradas com base na proficiência final em matemática, extraídos do relatório que se encontra no Anexo 08.

Quadro 30 - Resultados Modelo de Nível 1 em Matemática

		Estimativa	E. P.	p-value	
Efeitos Fixos	Intercepto	π_{0jk}	59,7	4,9	<0,001
		π_{1jk} (T2)	-81,6	5,0	<0,001
		π_{2jk} (T3)	4,2	4,3	0333
		π_{3jk} (T4)	16,6	3,9	<0,001
	Coefic. de inclinação	π_{4jk} (PR_IN)	0,589	0,034	<0,001
		π_{5jk} (PR_INxT2)	0,628	0,037	<0,001
		π_{6jk} (PR_INxT3)	0,186	0,033	<0,001
		π_{7jk} (PR_INxT4)	0,174	0,031	<0,001
		π_{8jk} (RET)	17,0	13,1	0,001
	Efeitos Aleatórios	Componente de Variância			p-value
$\sigma^2 =$		971,3			
$\tau_{\pi 00} =$		96,3		>0,50	
$\tau_{\pi 40} =$		0,00020		0,001	
$\tau_{\beta 000} =$		191,2		<0,001	
$\tau_{\beta 400} =$	0,00434		<0,001		
Deviance: 336871		Número de parâmetros estimados: 16			

Duas estimativas apresentadas nesse quadro requerem destaque:

- o *p-value* associado ao coeficiente π_{2jk} aponta que a variável indicadora T3, não é significativa, pois possui valor maior que 0,05. Isso significa que o valor estimado de 4,2 não pode ser considerado diferente de zero. Como consequência, os interceptos para os anos 1 e 3 são os mesmos⁴¹. Neste caso, assim como na proficiência em leitura, optou-se por não excluí-lo tornando possível a inserção das variáveis explicativas separadamente para esses anos.

- O *p-value* do efeito aleatório $\tau_{\pi 00}$ evidencia que as diferenças entre os interceptos das proficiências finais em matemática dos alunos de uma mesma escola não são significativamente diferentes, daí a inferência de que as variáveis explicativas associadas ao perfil do aluno, a ser inseridas no intercepto do modelo construído no próximo passo, níveis superiores, não teriam a capacidade explicativa sobre o intercepto, mas sim na inclinação.

Após a exclusão da variável indicadora T3 o Modelo Nível 1 para leitura ficou assim ajustado:

Nível 1 – Medidas Cognitivas no Tempo:

$$PR_FI_{ijk} = \pi_{0jk} + \pi_{1jk} \cdot T2_{ijk} + \pi_{2jk} \cdot T4_{ijk} + \pi_{3jk} \cdot PR_IN_{ijk} + \pi_{4jk} \cdot PR_IN \times T2_{ijk} + \pi_{5jk} \cdot PR_IN \times T3_{ijk} + \pi_{6jk} \cdot PR_IN \times T4_{ijk} + \pi_{7jk} \cdot RET_{ijk} + e_{ijk}, \text{ com } e_{ijk} \sim NID(0, \sigma^2).$$

Nível 2 – Alunos:

$$\pi_{0jk} = \beta_{00k} + r_{0jk}, \text{ com } r_{0jk} \sim NID(0, \tau_{\pi 00})$$

$$\pi_{1jk} = \beta_{10k};$$

$$\pi_{2jk} = \beta_{20k};$$

$$\pi_{3jk} = \beta_{30k} + r_{3jk}, \text{ com } r_{3jk} \sim NID(0, \tau_{\pi 30})$$

$$\pi_{4jk} = \beta_{40k};$$

$$\pi_{5jk} = \beta_{50k};$$

$$\pi_{6jk} = \beta_{60k};$$

$$\pi_{7jk} = \beta_{70k}$$

Nível 3 – Escolas:

$$\beta_{00k} = \gamma_{000} + u_{0jk}, \text{ com } u_{0jk} \sim NID(0, \tau_{\beta 000})$$

$$\beta_{10k} = \gamma_{100}; \beta_{20k} = \gamma_{200}$$

$$\beta_{30k} = \gamma_{300} + u_{3jk}, \text{ com } u_{3jk} \sim NID(0, \tau_{\beta 300})$$

$$\beta_{40k} = \gamma_{400}; \beta_{50k} = \gamma_{500}; \beta_{60k} = \gamma_{600}; \beta_{70k} = \gamma_{700}$$

O quadro 31 traz as estimativas desses coeficientes.

41 Mais uma vez é importante destacar nessa situação que o fato dos interceptos do primeiro e do terceiro anos serem iguais, não tem qualquer significado pedagógico, uma vez que não existe proficiência igual a zero. Essa igualdade nos interceptos também não implicam retas iguais, uma vez que os coeficientes de inclinação (π_{3jk} e $\pi_{3jk} + \pi_{5jk}$) são diferentes, além de que os intervalos de proficiência inicial desses dois anos de escolarização são distintos.

Quadro 31 - Resultados Modelo de Nível 1 em Matemática

		Estimativa	E. P.	<i>p-value</i>	
Efeitos Fixos	Intercepto	π_{0jk}	62,9	4,2	<0,001
		π_{1jk} (T2)	-84,4	4,2	<0,001
		π_{2jk} (T4)	+13,6	3,0	<0,001
	Coefc. de inclinação	π_{3jk} (PR_IN)	0,564	0,021	<0,001
		π_{4jk} (PR_INxT2)	0,651	0,030	<0,001
		π_{5jk} (PR_INxT3)	0,218	0,010	<0,001
		π_{6k} (PR_INxT4)	0,200	0,021	<0,001
		π_{7jk} (RET)	16,7	3,1	<0,001
Efeitos Aleatórios	Componente de Variância			<i>p-value</i>	
	$\sigma^2=$	971,7			
	$\tau_{\pi00}=$	7,5		>0,50	
	$\tau_{\pi40}=$	0,00024		0,001	
	$\tau_{\beta000}=$	195,9		<0,001	
	$\tau_{\beta400}=$	0,00431		<0,001	
Deviance: 336875		Número de parâmetros estimados: 15			

Como o objetivo da próxima etapa é identificar as variáveis que são significantes e, mesmo com o efeito aleatório associado à variável T3 indicar não existir variabilidade a ser explicada, ela foi mantida no modelo para justamente possibilitar a inserção de variáveis nesse ano. No entanto, ela foi excluída no modelo usado para a análise individual dos efeitos das variáveis sobre a proficiência final, denominado Modelo de Nível 1, descrito mais adiante.

Em ambos os quadros fica caracterizado que nos diversos anos de escolarização as proficiências finais são maiores que as proficiências iniciais e que os dados se ajustam adequadamente à regressão linear. Além disso, está evidenciado que há diferença significativa nas proficiências entre os diferentes anos de escolarização, confirmando, assim, a Hipótese 4 (As medidas da proficiência dos alunos em função dos anos de escolarização podem ser representadas linearmente) e também a Hipótese 5 (Há variabilidade significativa nas proficiências finais dos alunos nos diferentes anos de escolarização, sendo possível determinar o quanto desta variabilidade foi explicada pelas variáveis controladas).

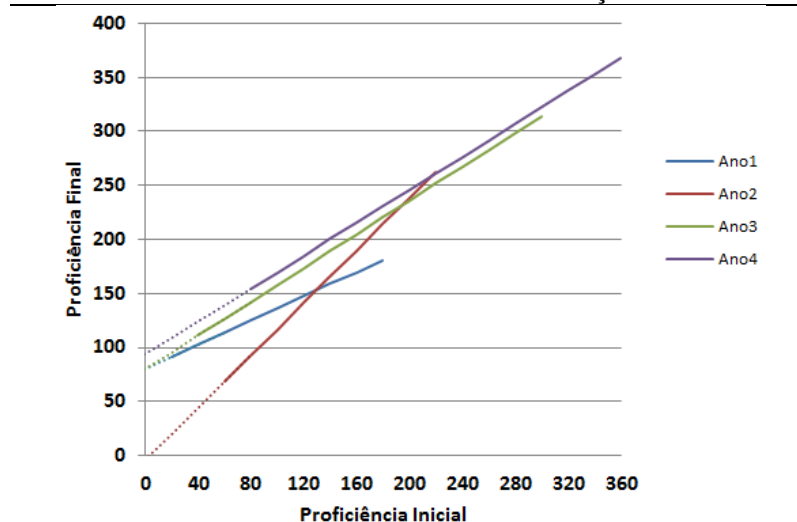
Com base nas estimativas apresentadas no Quadro 31, os interceptos e os coeficientes de inclinação para cada um dos anos está descrito no Quadro 32.

Quadro 32 -Equação de regressão de cada ano em Matemática

Ano		Intercepto	Coefficiente de inclinação
1	Composição	$\pi_{0jk} + \pi_{7jk}$	π_{3jk}
	Valor	79,6	0,564
2	Composição	$\pi_{0jk} + \pi_{1jk} + \pi_{7jk}$	$\pi_{3jk} + \pi_{4jk}$
	Valor	-4,8	1,215
3	Composição	$\pi_{0jk} + \pi_{2jk} + \pi_{7jk}$	$\pi_{3jk} + \pi_{5jk}$
	Valor	79,6	0,782
4	Composição	$\pi_{0jk} + \pi_{7jk}$	$\pi_{3jk} + \pi_{6jk}$
	Valor	93,2	0,764

O gráfico 16 representa a equação de regressão de cada um dos anos de escolarização⁴².

Gráfico 16 - Retas dos diversos anos de escolarização – Matemática



Através da diferença entre a média das proficiências finais e a média das proficiências iniciais, o valor agregado médio é de 29 para o primeiro ano de escolarização, 26 para o segundo, 44 para o terceiro e, também, de 44 para o quarto. Mas, com essa modelagem torna-se possível analisar o comportamento do valor agregado na proficiência, em cada um dos anos de escolarização, para os diferentes alunos que compõem a amostra em função de sua proficiência inicial, como mostra a Tabela 19.

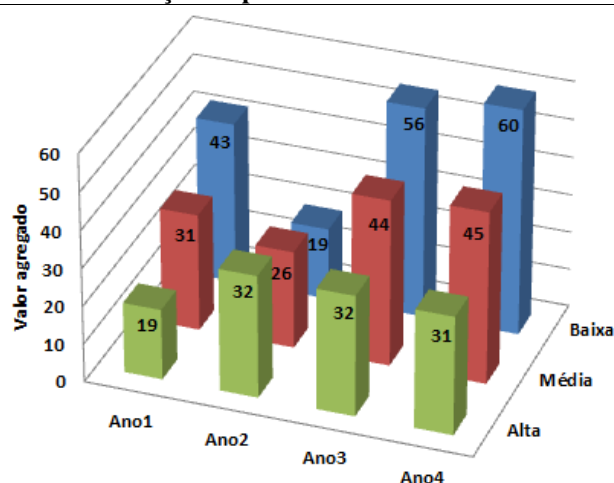
42 Este é o gráfico é produzido com as estimativas reais encontradas para proficiência em Matemática, correspondente à Figura 7 que mostra teoricamente essas retas e também ao Gráfico 8 com dados hipotéticos, apresentados no Capítulo 2.

Tabela 19 - Valor agregado como diferença da proficiência final estimada e da inicial

		Proficiência				Valor agregado			
		Ano1	Ano2	Ano3	Ano4	Ano1	Ano2	Ano3	Ano4
Baixa	Inicial	85	110	110	142	43	19	56	60
	Final estimada	128	129	166	202				
Média	Inicial	112	141	164	202	31	26	44	45
	Final estimada	143	167	208	247				
Alta	Inicial	140	172	219	262	19	32	32	31
	Final estimada	159	204	251	293				

O Gráfico 17 representa esses valores.

Gráfico 17 - Valor agregado em matemática por ano de escolarização em função da proficiência inicial



Para esta estimativa sem a preocupação de colocar os valores dentro de um intervalo de confiança. A proficiência inicial média foi obtida a partir da média das proficiências iniciais de cada um dos anos de escolarização separadamente. A proficiência baixa foi colocada a um desvio padrão abaixo da média e a proficiência inicial alta.

Esses valores agregados estimados por meio das equações de regressão, cujos valores se encontram no Quadro 32, mostram que, em média, os alunos com menor proficiência inicial têm um valor agregado estimado⁴³ maior no período de escolarização do que aqueles com proficiência inicial alta.

⁴³ É importante lembrar, mais uma vez, que estas estimativas foram feitas sem qualquer preocupação em estabelecer intervalos de confiança.

4.3.3) Modelo de Nível 2 para Matemática – Alunos

Esta etapa do processo seguiu os mesmos procedimentos para a inserção das variáveis de Nível 2, cujo modelo ajustado se encontra no Anexo 09, é composto pelas variáveis elencadas no Quadro 33.

Quadro 33 - Variáveis significantes no Modelo Nível 2 – Matemática

Variável	Descrição	Intercepto					Coeficiente de inclinação			
		Ano 1	Ano2	Ano3	Ano4	Ret	Ano 1	Ano2	Ano3	Ano4
FA01	Sexo				**		*			
FA04	Indica se o aluno fez a pré-escola	**	**				**	**		
FA06	Nível socioeconômico do aluno									*
FA07	Estímulo extrínseco para estudo			**			*	**	*	*
FA08	Estímulo intrínseco para o estudo						**			
FA09	Percepção dos alunos quanto ao incentivo do professor				*	*		**		
FA10	Gosto pela escola									
FA11	Estímulo interno para lição de casa					*				
FA12	Meus professores pensam que sou um aluno lento	**	*			**		**	**	*
FA13	Sou tipo de pessoa que sempre será capaz de conseguir, se tentar	*	**							*
FA14	Algumas vezes eu sinto que não sou bom o bastante	**	**		**	*				

Obs.: * indica $p\text{-value} < 0,001$; ** indica $p\text{-value} < 0,05$; *** indica $p\text{-value} < 0,1$

Essas variáveis são mantidas no processo de inclusão das variáveis explicativas de Nível 3, descrito no próximo item.

O Quadro 34 apresenta os valores estimados para as variâncias, deviance e o teste qui-quadrado. As demais estimativas do Modelo de Nível 2 encontram-se no Anexo 09.

Quadro 34 - Estimativas das variâncias e do deviance no Modelo Nível 2 de Matemática

Componente de Variância		<i>p-value</i>
$\sigma^2 =$	921,18	
$\tau_{\pi00} =$	7,31	>0,5
$\tau_{\pi40} =$	0,002	<0,001
$\tau_{\beta000} =$	132,9	<0,001
$\tau_{\beta400} =$	0,00375	<0,001
Deviance: 334970	Número de parâmetros estimados: 46	
Teste qui-quadrado: 1901,5	Graus de liberdade: 31	<0,001

Em uma análise global dos resultados do Modelo de Nível 2 verificou-se que a inclusão de variáveis explicativas ligadas ao perfil do aluno apresentaram uma redução significativa na *deviance* em relação ao Modelo de Nível 1 reduzindo de 336871 para 334970, redução considerada significativa pelo teste qui-quadrado, cujo *p-value* resultou em um valor muito próximo de zero. Logo, as variáveis explicativas do nível de aluno contribuíram para um maior ajuste do modelo aos dados.

Outra constatação desse Modelo é o *p-value* <0,001 do coeficiente P2, associado à variável T3, tornando-a significativa, diferentemente do que apontava o Modelo de Nível 1. No entanto, optou-se por não interferir nesse nível.

Os efeitos aleatórios do Modelo de Nível 2 indicam uma redução na variância total dos modelos pois diminuíram significativamente no Nível 1 (Medidas Cognitivas no Tempo) e no Nível 3 (Escolas), mas indicaram um aumento na variância no Nível 2 (alunos). As alterações encontradas são as seguintes:

- redução de σ^2 foi de 5,2%, diminuindo o valor de 971,4 para 921,2;
- aumento de $\tau_{\pi00}$ foi de 16,6%, aumentando o valor de 6,7 para 7,3;
- manutenção do valor de R4 em 0,0002;
- a redução de $\tau_{\beta000}$ foi de 30,5%, diminuindo o valor de 191,2 para 132,9;
- a redução de $\tau_{\beta400}$ foi de 12,8%, diminuindo o valor de 0,0043 para 0,0038

Assim, com os resultados apresentados nesse Modelo, constata-se que a Hipótese 5 aparentemente não é confirmada, dado que as variáveis explicativas do Nível 2, associadas aos alunos, são capazes de explicar a variabilidade da proficiência final em cada um dos anos,

uma vez que a variabilidade constatada no Modelo Nulo é praticamente eliminada com a inserção das variáveis de Nível 1.

O termo “aparentemente” foi usado porque a proficiência inicial deve ser entendida como uma das características inerentes ao aluno. Logo a Hipótese 5 (Há variabilidade significativa nas proficiências finais dos alunos nos diferentes anos de escolarização, sendo possível determinar o quanto desta variabilidade foi explicada pelas variáveis controladas), deve ser mantida como verdadeira.

Com esse resultado, para posterior comparação, o passo 4 que consiste na inserção das variáveis explicativas do Nível 3 será descrito a seguir.

4.3.4) Modelo de Nível 3 para Matemática – Escolas

Da mesma maneira do que foi feito nos modelos anteriores, as variáveis explicativas consideradas significantes nos níveis 1 e 2 foram mantidas e, seguindo o método *backward*, foram inseridas simultaneamente todas as variáveis associadas ao Nível 3, e em seguida, aquelas menos significantes (segundo o *p-value* estimado) foram sendo excluídas até que todas apresentassem *p-value* adequado.

Além disso, é importante destacar que quando essa estimativa é significativa no primeiro ano de escolarização, tanto no intercepto, quanto na inclinação, tem o mesmo impacto nos demais anos. Com essa observação, quando um coeficiente é negativo em um determinado ano, não significa necessariamente que seu efeito é negativo, mas podendo representar apenas que é menor que no primeiro ano.

Por isso, a análise do efeito das variáveis explicativas requerem um outro modelo interpretativo onde todos os coeficientes da variáveis sejam considerados.

Os quadros a seguir apresentam as variáveis associadas às características da escola que se mostraram significantes tanto no intercepto quanto na inclinação da reta de cada um dos anos. Eles evidenciam os anos em que a estimativa do coeficiente associado à variável explicativa foi significativa.

Quadro 35 - Variáveis significantes no Modelo de Nível 3 - Matemática pertencentes ao Grupo 1

Variável	Descrição	Intercepto				Coeficiente de inclinação			
		Ano1	Ano2	Ano3	Ano4	Ano1	Ano2	Ano3	Ano4
FEA01	NSE médio da escola a partir do NSE dos alunos	*				**	*	*	*
FEE02	Conservação e limpeza supraestrutura	*				*	**	**	**
FEE03	Conservação e limpeza área de alimentação		**			**		**	
FEE04	Espaço sala de leitura		**	**				**	
FEE05	Condições da Biblioteca								
FEE06	Laboratório de informática								
FEE07	Sala de artes e laboratório de ciências	**				**			
FEE08	Disponibilidade de computador e internet para professor			**				**	
FEE09	Existência de equipamentos para fotocópia e escaneamento				**				**
FEE10	Existência de material concreto para matemática				**				
FEE11	Existência de terrário e aquário		**			**	**		

Obs.: * indica p-value < 0,001; ** indica p-value < 0,05; *** indica p-value < 0,1

Novamente é evidenciado que o nível socioeconômico médio da escola (FEA01) continua influenciando durante todo o período de escolarização e, ainda, de forma diferente para alunos com distintas proficiências iniciais, o mesmo ocorrendo com a variável que controla a “Conservação e limpeza da supraestrutura” (FEE02). Além delas, somente a variável que controla o uso de “sala de arte e laboratório de ciências” influencia, porém igualmente, em todos os anos.

As variáveis “Condições da Biblioteca” (FEE05) e “Laboratório de informática” (FEE06) não tiveram qualquer associação com a proficiência final em matemática.

Das onze variáveis desse grupo apenas quatro (FEA01, FEE02, FEE03 e FEE07) têm influência em todos os anos de escolarização.

Quadro 36 - Variáveis significantes no Modelo de Nível 3 - Matemática pertencentes ao Grupo 2

Variável	Descrição	Intercepto				Coeficiente de inclinação			
		Ano1	Ano2	Ano3	Ano4	Ano1	Ano2	Ano3	Ano4
FED12	Interferência por intimidação, violência, discriminação e drogas	**							
FED13	Interferência do perfil do professor					**			**
FED14	Interferência do perfil do aluno, família								
FED15	Interferência da orientação pedagógica								
FED16	Presença física do diretor				**				*
FED17	Organização de eventos				**				
FED18	Frequência de encontros com a comunidade	**				**		**	**
FED19	Frequência da presença na comunidade				**				**
FED20	Proporção de diretores que possuem Ens. Sup. em Pedagogia				*				*
FED21	Proporção de Diretores que possui outra atividade que contribui para a renda pessoal					***	**	**	**
FED22	Proporção de diretores que trabalham em apenas uma escola		**				**		
FED23	Carga horária de trabalho do(a) diretor(a) na escola		**						
FED24	Tempo que o diretor trabalha na escola ininterruptamente na função de direção								
FED25	Tempo que o(a) diretor (a) trabalha em anos na educação	**	**			**			
FED26	Idade do(a) diretor(a) em anos								

Obs.: * indica p-value < 0,001; ** indica p-value < 0,05; *** indica p-value < 0,1

Das quinze variáveis que compõem esse grupo, quatro (FED14, FED15, FED24 e FED26) não tem qualquer influência, cinco (FED12, FED13, FED18, FED21 e FED25) têm alguma influência em todos os anos e as demais têm alguma influência em algum dos anos de escolarização.

Quadro 37 - Variáveis significantes no Modelo de Nível 3 - Matemática pertencentes ao Grupo 3

Variável	Descrição	Intercepto				Coeficiente de inclinação			
		Ano1	Ano2	Ano3	Ano4	Ano1	Ano2	Ano3	Ano4
FEP33	Frequência do dever de casa				*				**
FEP36	Acesso a atividades culturais								
FEP37	Proposição de leitura					**			
FEP38	Proposição de redação		**						
FEP39	Proposição de cópia ou caligrafia	*				*			
FEP40	Proposição de leitura individual é precedida da explicação do professor			**	*				
FEP41	Leitura em prática coletiva	**	**		**		**		
FEP42	Leitura individual e produção de texto individual								
FEP43	Uso da metodologia de resolução de problemas								
FEP44	Ensino com ênfase em memorização								
FEP45	Adequação do livro ao ensino		**				**		
FEP46	Grau de importância inerente a indicação externa do livro								
FEP47	Tempo usado nas disciplinas relacionadas com as proficiências		**			*	**		
FEP48	Tempo usado nas disciplinas não relacionadas com as proficiências	**							**
FEP49	Tempo usado em disciplinas de formação complementar	**	**	*	**		**	**	
FEP52	Uso de recursos audiovisuais								
FEP53	Uso de laboratório e multimídia		**				**		**
FEP54	Uso de materiais concretos de matemática e geografia	***							

Obs.: * indica p-value < 0,001; ** indica p-value < 0,05; *** indica p-value < 0,1

Neste grupo, das dezoito variáveis, seis não tem qualquer influência (FEP36, FEP42, FEP43, FEP44, FEP46 e FEP52), sete tem alguma influência em todos os anos (FEP37, FEP39, FEP41, FEP47, FEP48, FEP49 e FEP54) e as demais têm alguma influência em algum dos anos de escolarização.

O fato de uma variável aparecer em apenas um dos anos, como por exemplo ocorre com a variável “Uso de materiais concretos de matemática e geografia” não significa que ela não seja importante, interpretação equivocada que a organização desses quadros podem induzir. No caso dessa variável, apesar de ser significativa apenas no intercepto do primeiro

ano de escolarização, cada unidade que é acrescida influencia em quase dois pontos na proficiência final, o equivalente a 10% do valor agregado médio nos distintos anos.

Quadro 38 - Variáveis significantes no Modelo de Nível 3 - Matemática pertencentes ao Grupo 4

Variável	Descrição	Intercepto				Coeficiente de inclinação			
		Ano1	Ano2	Ano3	Ano4	Ano1	Ano2	Ano3	Ano4
FEP27	Relação com o diretor						**		
FEP28	Boa interação entre a equipe								
FEP29	Pouco trabalho coletivo					*			
FEP30	Boa relação entre professores								
FEP31	Unidade de propósito entre os professores								
FEP32	Violência no ambiente escolar	**	**			*	**		
FEP33	Frequência do dever de casa				*				**
FEP34	Grau de expectativas negativa do aprendizado	**	**	*		**	**	*	*
FEP35	Grau de expectativas positiva do aprendizado					**			
FEP50	Interrupção das aulas por indisciplina			**	*			**	
FEP51	Interrupção das aulas por necessidade administrativa				*	**	**	**	
FEP55	Proporção de professores do sexo feminino							**	**
FEP56	Idade média dos professores		**						
FEP57	Renda Familiar média dos professores (salários mínimos)								
FEP58	Tempo de exercício médio na ocupação		**						
FEP59	Número médio de pessoas da família que mora com o professor								

Obs.: * indica p-value < 0,001; ** indica p-value < 0,05; *** indica p-value < 0,1

Desse grupo, cinco variáveis não tiveram qualquer associação com a proficiência final dos alunos em matemática em todos os anos. São elas:

- FEP28, “Boa interação entre a equipe”;
- FEP30, “Boa relação entre professores”;
- FEP31, “Unidade de propósito entre os professores”;
- FEP57, “Renda Familiar média dos professores (salários mínimos)”, e;
- FEP59, “Número médio de pessoas da família que mora com o professor”.

Com exceção das duas últimas, esse resultado traz significativo estranhamento, por ser evidente a importância de cada uma delas e, por isso, colocando em xeque os resultados do modelo. No entanto, deve-se considerar como essas variáveis tiveram baixa variabilidade nos questionários de contexto, não sendo, portanto, capazes explicar a variabilidade da proficiência final⁴⁴.

Ainda desse grupo, das 16 variáveis, apenas cinco tiveram algum impacto nos diferentes anos de escolarização (FEP29, FEP32, FEP34, FEP35 e FEP51) .

O Quadro 39 apresenta os valores estimados para as variâncias e a *deviance*, obtidos com o modelo, cujas equações, assim como as estimativas encontram-se no relatório emitido pelo programa HLM no Anexo 10. As demais estimativas do Modelo de Nível 2 encontra-se no Anexo 09.

Quadro 39 - Estimativas das variâncias e do *deviance* no Modelo Nível 3 de Matemática

Componente de Variância		<i>p-value</i>
$\sigma^2=$	892,9	
$\tau_{\pi00}=$	18,5	>0,50
$\tau_{\pi40}=$	0,0006	0,033
$\tau_{\beta000}=$	24,0	<0,001
$\tau_{\beta400}=$	0,00112	<0,001
Deviance: 333593	Número de parâmetros estimados: 153	
Teste quiquadrado: 1276,9	Graus de liberdade: 107	<0,001

Este Modelo apresentou *deviance* menor de forma que no Modelo de Nível 2, diminuindo para 333593 o valor, que era de 334970, uma redução significativa (*p-value* <0,001, no teste quiquadrado). Mas, apesar disso, as componentes das variâncias do Nível 2 aumentaram significativamente, mesmo com as demais diminuindo, como mostram os valores a seguir:

- σ^2 diminuiu de 921,18 para 892,94 (redução de 3,1%);
- $\tau_{\pi00}$ aumentou de 7,31 para 18,50 (aumento de 153,1%);
- $\tau_{\pi40}$ aumentou de 0,0002 para 0,0006 (aumento de 200,0%);
- $\tau_{\beta000}$ reduziu de 132,9 para 24,03 (redução de 81,9%);
- $\tau_{\beta400}$ reduziu de 0,00375 para 0,00112 (redução de 70,1%).

44 FEP28 tem média de 3,21 e desvio padrão de 0,39. FEP30 tem média de 3,35 e o desvio padrão de 0,33, FEP31 tem média de 3,04 e o desvio padrão de 0,57. Todas num intervalo de 0 a 4.

Esses resultados confirmam a interpretação dada no Modelo de Nível 2, quando foi indicado que as variáveis do Nível 2 não têm capacidade explicativa da variância entre alunos de uma mesa escola, ao contrário do que foi sinalizado no Modelo Nulo (15% da variância seria explicada pela diferença entre alunos).

Das 59 variáveis associadas às características da escola, pode-se perceber 17 delas não foram significantes. São elas: FEE05, FEE06, FED14, FED15, FED24*, FED26*, FEP28, FEP30, FEP31, FEP36, FEP42, FEP43, FEP44, FEP46, FEP52, FEP57*, FEP59*

O Quadro 40 apresenta uma breve análise sobre o comportamento das variáveis do Nível 3, associadas, às características das escolas, na perspectiva pedagógica.

Quadro 40 - Interpretação final das variáveis do Grupo 2 do Nível 3 em Matemática

Comportamento	Variáveis
Coerente como esperado	FEA01, FEE08, FEE09, FED19, FED21, FED24*, FEP28, FEP43, FEP48.
Incoerente com o esperado	FEE04, FEE05, FEE06, FED15, FED18, FEP30.
Instável, dificultando a interpretação	FEE02, FEE03, FEE07, FEE10, FEE11, FED12, FED13, FED14, FED17, FED20, FED22*, FED23*, FED25*, FED26*, FEP29, FEP31, FEP32, FEP33, FEP34, FEP35, FEP36, FEP36, FEP37, FEP38, FEP39, FEP40, FEP41, FEP42, FEP43, FEP44, FEP45, FEP46, FEP47, FEP49, FEP52, FEP53, FEP54, FEP51, FEP55*, FEP57*, FEP58*.

Os resultados encontrados comprovam que a proficiência final dos alunos não só é influenciada por diversos fatores, mas também que há uma forte interação entre esses fatores, evidenciando, quantitativamente a complexidade do fenômeno educacional.

Na análise feita até o momento, não teve como foco a verificação da intensidade com que ocorre essa influência, e se limitou a verificar a forma pela qual ela se manifesta, em função das estimativas apresentadas. O objetivo foi evidenciar que essa múltipla influência existe e que se modelos interpretativos forem criados com a possibilidade de incorporar essa complexidade ela por ser verificada quantitativamente.

Mesmo com a instabilidade percebida em parte das variáveis, pode se constatar ainda que, sobretudo pelo fato de que o banco de dados usado incorpora informações de escolas e alunos de diferentes redes de ensino (municipais, estaduais, privadas e especiais). O nível socioeconômico do aluno, dada sua alta correlação com a proficiência inicial, é a variável que mais influência tem sobre a proficiência inicial, o mesmo ocorrendo com a proficiência média

da escola. Essa influência é forte, não só antes da entrada do aluno na escola, mas também durante todo seu período de escolarização.

Outra constatação importante é que o número de variáveis que se mostraram significantes no modelo de matemática é menor do que foi encontrado em Leitura, cuja proficiência final é menor. Essa constatação é importante por evidenciar que em um ano de escolarização, a variabilidade é suficiente para encontrar variáveis explicativas.

É importante registrar também que uma das características relevantes desse modelo é a possibilidade de analisar a influência das variáveis explicativas sobre a proficiência final de alunos com proficiências iniciais diferentes.

Um aspecto importante a ser destacado aqui, é o fato de que essas estimativas não foram feitas com intervalo de confiança, o que dificultaria ainda mais a análise, dada a faixa pequena de valores que elas trabalham. Dentro da mesma limitação técnica, também existentes, é que essas interpretações são feitas para intervalos de proficiência inicial entre um desvio-padrão abaixo e um acima da média. Se esse intervalo fosse menor, a instabilidade seria menos percebida, ocorrendo o contrário se esse intervalo fosse aumentado.

Essa análise dos efeitos dessas variáveis será feita no próximo capítulo.

CAPÍTULO 5 - ANÁLISE DOS EFEITOS DAS VARIÁVEIS

No capítulo anterior foram apresentados os procedimentos usados para o ajuste dos modelos que permitem identificar variáveis que interferem nas proficiências finais em Leitura e em Matemática em cada um dos anos de escolarização, ou seja, permitindo identificar o rol dos fatores associados às proficiências finais dos alunos.

Este capítulo tem por objetivo trazer as interpretações acerca da forma que essas variáveis explicativas (fatores) interferem na proficiência final de cada um dos anos de escolarização, considerando as diferentes proficiências iniciais dos alunos. Seu objetivo é apresentar o comportamento das diversas variáveis, considerando a interação entre elas, em Leitura e Matemática, dos Modelos de Nível 2 – Alunos e dos Modelos de Nível 3 – Escola, culminando em uma síntese comparativas entre esses modelos para Leitura e Matemática, apresentadas no final do capítulo, para também destacar as especificidades das variáveis em cada um deles.

Para essa análise de resultados, é importante resgatar as quatro principais características dos modelos:

a) ao colocar a proficiência final como variável resposta e a proficiência inicial como variável explicativa no Nível 1 dos modelos, os resultados são estimativas dos parâmetros de equações das retas que se ajustam a cada um dos anos de escolarização envolvidos no estudo;

b) as demais variáveis explicativas que compõem o modelo, sejam elas do Nível 2 ou do Nível 3, podem ser estatisticamente significantes:

- apenas no intercepto (coeficiente linear);
- apenas na inclinação (coeficiente angular);
- em ambos;

c) com a criação e inserção das variáveis indicadoras a equação da reta do primeiro ano é a referência para os demais anos;

d) permitem verificar a influência das variáveis explicativas em alunos com proficiências iniciais diferentes.

Dessa última característica decorre que se uma variável é significativa:

a) apenas no primeiro ano, ela tem influência também nos demais anos;

b) no primeiro ano e também em algum dos outros anos, ela tem influência nesse outro ano, mas de forma diferente do que é no primeiro;

c) em algum dos outros anos, e não no primeiro, ela influencia a proficiência nesses outros anos e não nos demais.

5.1) Interpretação conjunta dos resultados do Modelo Nulo em Leitura e Matemática

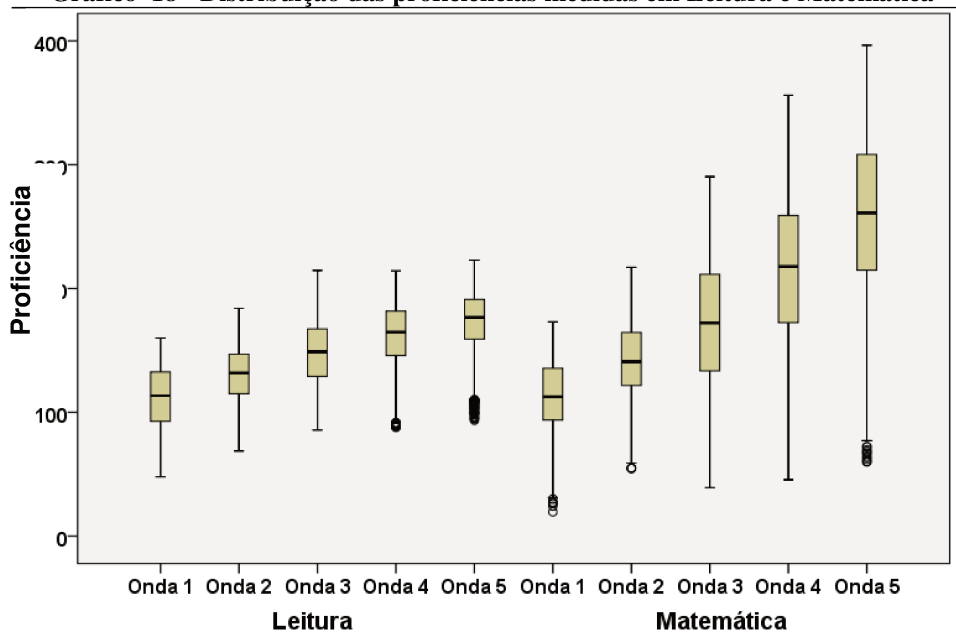
O Modelo Nulo é necessário para a estimação da partição da variância total em cada um dos níveis, como já analisado no capítulo anterior. A Tabela 20 reapresenta os resultados encontrados.

Tabela 20 – Estimativas dos Componentes de Variância dos Modelos Nulos

	Leitura		Matemática	
	Valor	%	Valor	%
σ^2	422,7	44,0%	2975,7	62,4%
$\tau_{\pi00}$	191,6	19,9%	719,8	15,1%
$\tau_{\beta000}$	346,4	36,1%	1070,1	22,5%
Total	960,7		4765,6	

Para além das considerações feitas no capítulo anterior, a primeira constatação é a maior variância das medidas em Matemática do que em Leitura, que é representada no Gráfico 18.

Gráfico 18 - Distribuição das proficiências medidas em Leitura e Matemática



Essa diferença na variabilidade das proficiências entre Leitura e Matemática pode ser gerada por diversos fatores, mas dentre eles está a distinção entre as amplitudes dos valores das escalas de proficiência, sobre as quais os itens dos testes aplicados nas diferentes ondas foram elaborados. Como visto no Capítulo 1, segundo Brooke e Bonamino (2011, p. 139) em Leitura, a escala de proficiência é interpretada no intervalo de 53 a 222, enquanto a de Matemática é interpretada numa escala com intervalo compreendido entre 36 e 387. Esse é outro aspecto que corrobora com a afirmação de que as proficiências iniciais dos alunos devem ser vistas como uma das características inerentes do perfil do aluno, juntamente com aquela que controla se o aluno foi retido no processo. São elas que carregam todas as diferenças entre os alunos quando entram em um determinado ano de escolarização.

Outra consideração é que a proficiência inicial tem alta capacidade explicativa da variância da proficiência final, dado que ambas são altamente correlacionadas como mostra a Tabela 21.

Tabela 21- Correlação de Spearman entre medidas de proficiência sucessivas

Entre proficiência medidas nas			Coefficiente de Correlação	p-value	N
Leitura	Onda 1	e Onda 2	0,772	< 0,001	7253
	Onda 2	e Onda 3	0,804	< 0,001	7233
	Onda 3	e Onda 4	0,845	< 0,001	8631
	Onda 4	e Onda 5	0,836	< 0,001	11540
Matemática	Onda 1	e Onda 2	0,694	< 0,001	7220
	Onda 2	e Onda 3	0,770	< 0,001	7175
	Onda 3	e Onda 4	0,845	< 0,001	8587
	Onda 4	e Onda 5	0,829	< 0,001	11494

A opção por determinar a correlação pelo método não paramétrico (Sperman) deve-se ao fato de que as proficiências dos alunos selecionados na amostra não têm uma distribuição Normal, em qual foi usado o teste de Kolmogorov- Smirnof, como mostra a Tabela 22.

Uma observação importante neste aspecto é que, nos modelos que estudaram a proficiência em matemática, a proficiência inicial explicou quase toda a variabilidade dos dados referentes ao aluno, como pode ser visto no Quadro 34 do Capítulo 4, fazendo com que as demais variáveis associadas ao seu não se mostraram significantes. Com essa constatação em matemática, pode-se afirmar que, descontando as características do aluno ao entrar em um

determinado ano de escolarização, toda a variação da sua proficiência final depende das características da escola, diferentemente do que ocorre em Leitura.

Tabela 22- Resultado do teste de Normalidade das proficiências

	Medida	Estatística	Graus de Liberdade	<i>p-value</i>
Leitura	Onda 1	,054	5991	<0,001
	Onda 2	,022	5991	,000
	Onda 3	,025	5991	,000
	Onda 4	,044	5991	,000
	Onda 5	,051	5991	,000
Matemática	Onda 1	,045	5991	,000
	Onda 2	,030	5991	,000
	Onda 3	,025	5991	,000
	Onda 4	,034	5991	,000
	Onda 5	,036	5991	,000

Essa análise é aprofundada mais adiante, quando são analisados os resultados dos Modelos de Nível 1.

5.2) Interpretação dos resultados do Modelo de Nível 1 em Leitura e em Matemática

No capítulo anterior, alguns resultados desses modelos já foram destacados, visto que são úteis para a construção dos modelos de Nível 2 e Nível 3. Esses destaques foram:

- a) a análise dos Componentes da Variância;
- b) a descrição dos diferentes intervalos das proficiências iniciais a cada ano de escolarização que permitiu a análise dos resultados da significância de alguns dos coeficientes ($P3$ no Modelo de Nível 1 de Leitura e $P2$ e σ^2 no de Matemática).

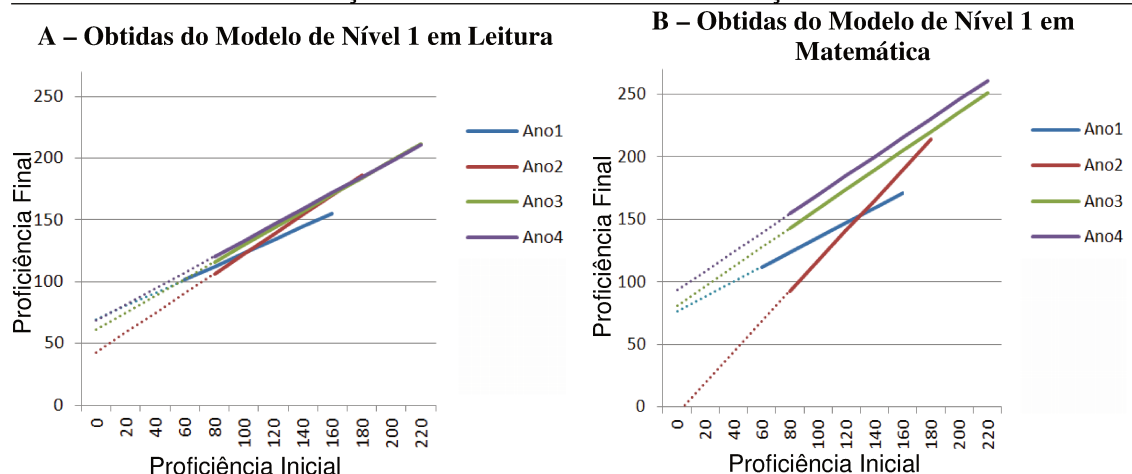
Os resultados encontrados nos coeficientes (efeitos fixos) dos Modelos de Nível 1 são valores estimados para os parâmetros das equações das retas que se ajustam a cada ano, e que permite, em termos de valores médios, analisar a evolução da proficiência final de cada um dos anos de escolarização.

Além disso, com base nesses valores encontrados, permitem uma análise também do valor agregado em cada um dos anos.

Inicialmente é importante explicitar as sobreposições das retas estimadas em função da sobreposição dos intervalos da proficiência inicial.

Como os Modelos de Nível 1 são elaborados com a proficiência final em função da inicial, a representação gráfica das retas permite uma interpretação bastante simples, como mostram os Gráficos 19A e 19B.

Gráfico 19 - Representação gráfica das equações das retas, via HLM, dos diversos anos de escolarização tendo a Proficiência Final em função da Inicial



Estes gráficos entre outras coisas permite: (i) evidenciar a diferença na estimação usando a regressão linear simples e aquela usando o modelo linear hierárquico; (ii) explicitar os intervalos das proficiências iniciais em cada um dos anos (trecho contínuo das retas) e (iii) representar geometricamente a posição de cada uma das retas, destacando o intercepto (valor da proficiência final para um aluno com o valor fictício da proficiência inicial igual a zero, indicado pelo prolongamento tracejado da reta, tecnicamente chamado de coeficiente linear) e a inclinação (tecnicamente chamado de coeficiente angular).

Essas retas são produzidas a partir de coeficientes lineares (a) e angulares (b), cujos valores estão sintetizados na Tabela 23 e foram obtidos dos Quadros 20 e 32 (para Leitura e Matemática, respectivamente) do Capítulo 4.

Ao considerar que dessa amostra de alunos, 6619 têm medidas de proficiências de leitura nas cinco ondas e 6106 em matemática, ou seja, em ambas as disciplinas, aproximadamente metade da amostra são alunos que a proficiência final de um ano é a inicial de outro. Outra possibilidade que pode ser explorada com os valores obtidos no Modelo de

Nível 1 é estimar a proficiência final do quarto ano de escolarização tendo a proficiência inicial do primeiro ano e, assim, analisar o valor agregado médio para cada ano.

Tabela 23- Parâmetros das retas obtidas com o Modelo de Nível 1 em leitura e Matemática

	Ano	proficiência	
		Coefficiente Linear (a)	Coefficiente Angular (b)
Leitura	1	69,6	0,535
	2	42,9	0,795
	3	61,3	0,683
	4	68,6	0,646
Matemática	1	76,7	0,589
	2	-5,0	1,217
	3	80,8	0,775
	4	93,3	0,763

As retas de cada um dos anos, usando esses parâmetros, estão representadas nos Gráficos 19A e 19B, nos quais as proficiências finais são calculadas a partir dos valores contidos nos intervalos das proficiências iniciais de cada um dos anos de escolarização, indicados na Tabela 24.

Tabela 24- Descritiva da proficiência Inicial em Leitura e Matemática

	Ano	Número de alunos	proficiência			Desvio Padrão
			Mínimo	Máximo	Média	
Leitura	1	7253	47,8	159,9	111,6	25,4
	2	7233	68,6	183,9	129,9	22,1
	3	8631	85,7	214,5	145,0	26,3
	4	11540	87,8	214,2	157,2	25,6
Matemática	1	7824	19,5	173,0	112,2	27,5
	2	7944	54,7	217,1	141,3	31,1
	3	9075	39,1	290,4	164,5	54,6
	4	11536	45,5	356,0	202,8	59,7

Para essa estimativa são usados os mesmos coeficientes das retas descritas na Tabela 24, de forma que, para o cálculo da proficiência final do primeiro ano, são colocados três valores de proficiência inicial na equação do primeiro. Para o cálculo da proficiência final do segundo ano são colocadas as proficiências finais calculadas no primeiro ano, e assim sucessivamente para os demais anos.

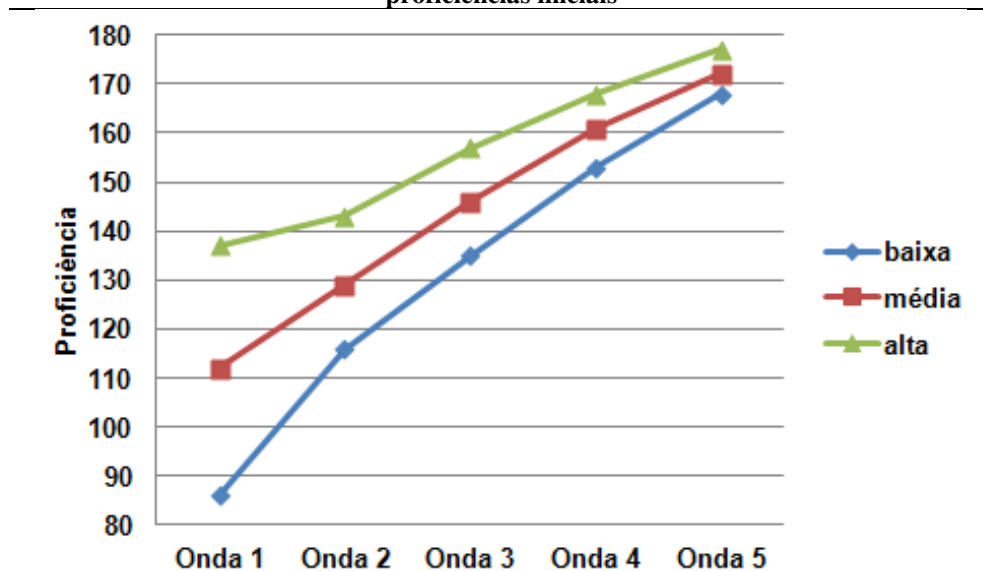
Na Tabela 25 são apresentados os valores estimados seguindo estes procedimentos para as proficiências finais a cada ano de escolarização para alunos que tenham diferentes proficiências iniciais no primeiro ano de escolarização. São escolhidos três valores de referência para proficiência inicial do primeiro ano⁴⁵: baixa ($\bar{x} - \sigma$), média (\bar{x}) e alta ($\bar{x} + \sigma$). Os valores para a média (\bar{x}) e desvio padrão (σ) são obtidos da Tabela 24.

Tabela 25- Proficiência Final a cada ano, tendo como proficiência inicial a final do ano anterior

	Proficiência Inicial	Proficiência Final				
		Ano 1	Ano2	Ano 3	Ano 4	
Leitura	baixa	86	116	135	153	168
	média	112	129	146	161	172
	alta	137	143	157	168	177
Matemática	baixa	85	127	149	176	218
	média	112	143	169	200	236
	alta	140	159	188	224	255

O Gráfico 20 representa os dados dessa tabela e permite a visualização da evolução das proficiências em leitura

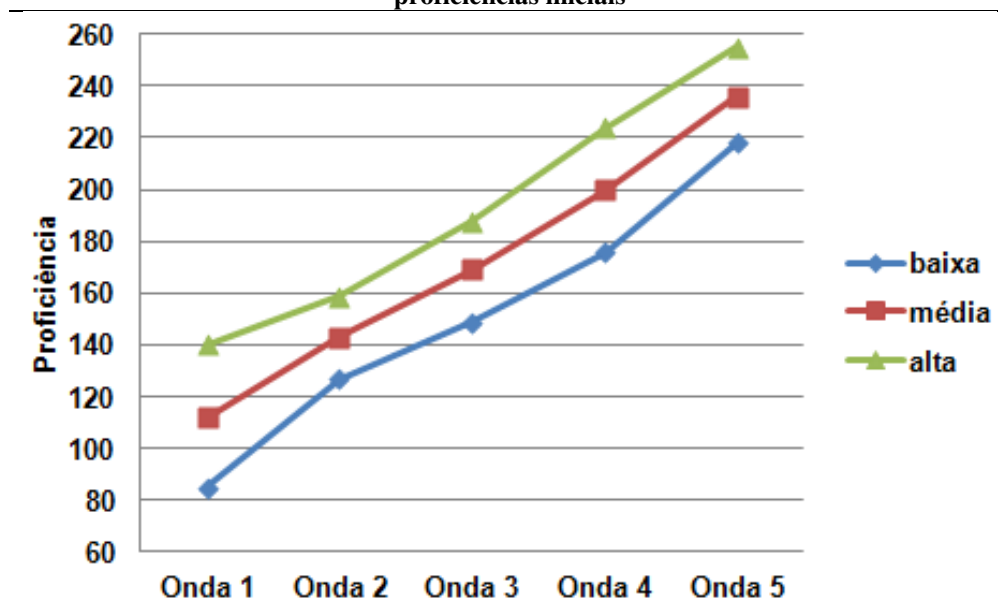
Gráfico 20 - Estimação da proficiência final em Leitura para diferentes proficiências iniciais



45 Inicialmente foram usadas cinco proficiências, incluindo proficiência que ficavam a 1,96 desvios-padrão, no entanto esses valores caíam fora do intervalo e provocavam distorções nas estimativas.

Esse gráfico evidencia que, em média, a diferença entre as proficiências finais estimadas com diferentes proficiências iniciais vai diminuindo durante os anos de escolarização, ou ainda, os alunos com proficiência iniciais baixas agregam mais que alunos com proficiência inicial alta no decorrer dos quatro primeiros anos de escolarização. O mesmo ocorre com as proficiências em matemática, porém de forma muito menos acentuada, como mostra o Gráfico 21.

Gráfico 21 - Estimação da proficiência final em Matemática para diferentes proficiências iniciais



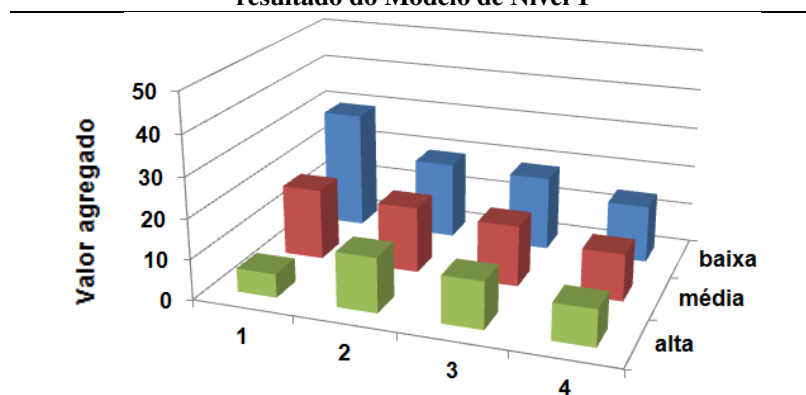
Essas evoluções nas proficiências finais estimadas permitem explorar os valores agregados a cada ano de escolarização como mostram os dados apresentados na Tabela 26, onde fica evidenciado que em Leitura, os alunos com proficiência inicial baixa têm um valor agregado médio, acumulado nos quatro anos, é 105% superior aos alunos com proficiência inicial alta, enquanto que essa diferença em Matemática é 15,7%.

Tabela 26- Valor agregado médio nos diferentes anos de escolarização e o acumulado nos quatro anos

		Ano 1	Ano2	Ano 3	Ano 4	Acumulado nos quatro anos
Leitura	baixa	30	19	19	14	82
	média	18	16	15	12	61
	alta	6	14	12	9	40
Matemática	baixa	42	22	27	41	133
	média	31	26	32	36	124
	alta	19	29	36	30	115

O Gráfico 22 é a representação dos valores agregados em Leitura da Tabela acima para diferentes proficiência iniciais e nos diferentes anos de escolarização.

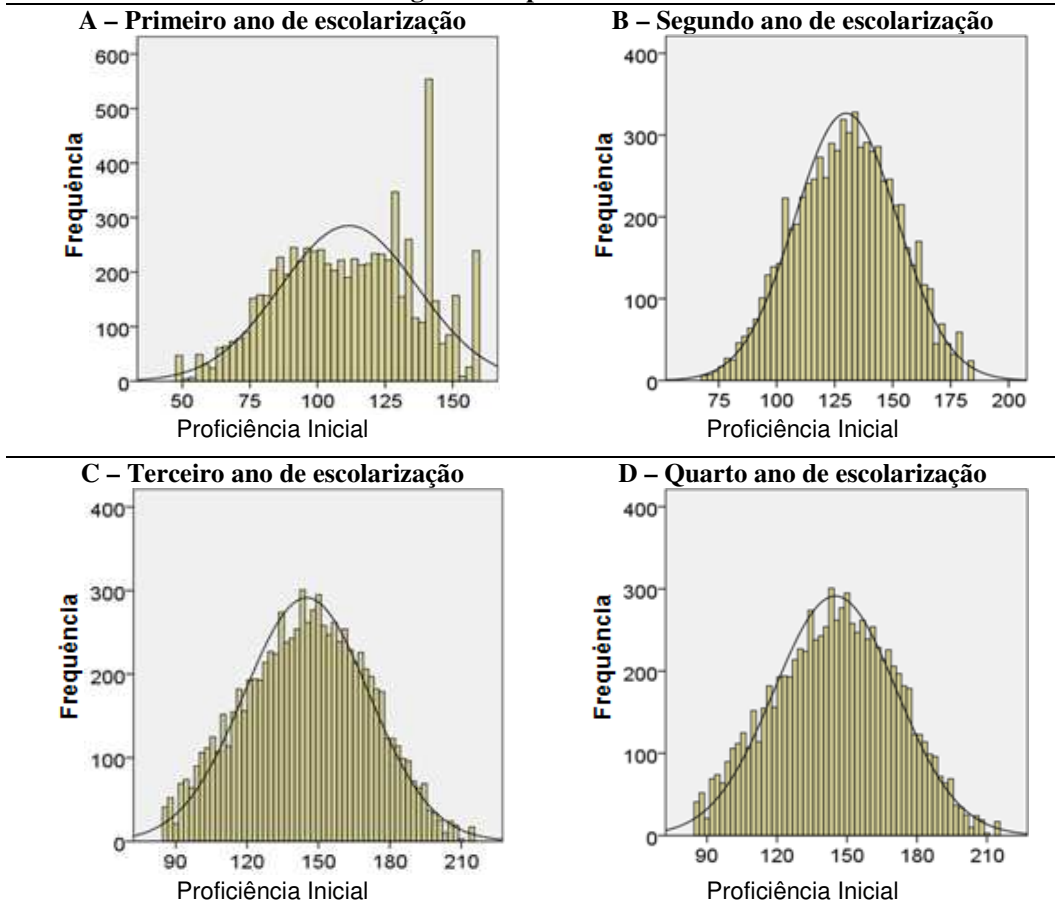
Gráfico 22 - Valor agregado médio em Leitura com base no resultado do Modelo de Nível 1



O Gráfico 22 evidencia, ainda, que em leitura, exceto para os alunos com proficiência inicial alta no primeiro ano de escolarização, o valor agregado diminui a cada ano. Ao mesmo tempo, dentro de cada ano, o valor agregado dos alunos com proficiência inicial mais baixa tem um valor agregado maior do que aqueles com proficiência inicial mais alta.

Buscando entender essa discrepância no valor agregado do primeiro ano de escolarização dos alunos com proficiência alta, nota-se alguns valores que fogem da regularidade esperada na medida da primeira proficiência, como mostram os Gráfico 23A, 23B, 23C e 23D, onde estão representadas as distribuições das proficiências iniciais em cada um dos anos de escolarização.

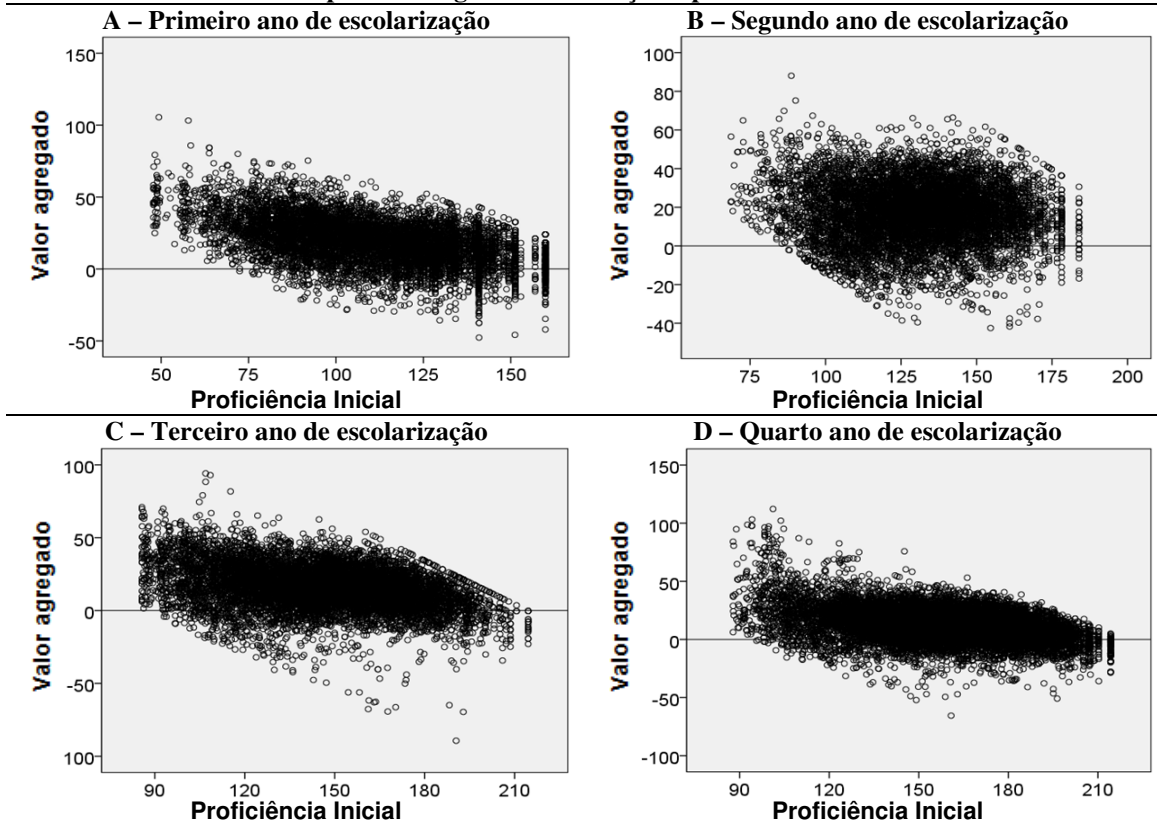
Gráfico 23 - Histograma da proficiência inicial em Leitura



Desta forma pode-se observar que há uma discrepância na distribuição do primeiro ano em relação aos demais.

Como o Modelo de Nível 1 coloca a proficiência final em função da inicial e, conseqüentemente permite o cálculo do valor agregado em função da proficiência inicial, foram elaborados os Gráficos 24A, 24B, 24C e 24D para mostrar a relação existente entre essas duas variáveis.

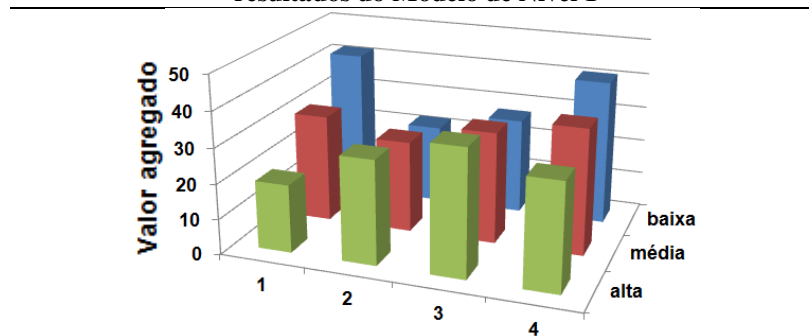
Gráfico 24 - Dispersão dos ganhos em relação à proficiência inicial em Leitura



Desses gráficos pode-se observar que há também ganhos negativos e que estes se concentram nas faixas com proficiência iniciais em Leitura mais altas, mostrando uma tendência para valores agregados menores quando as proficiências iniciais são mais altas.

Já em Matemática, o comportamento do valor agregado não é tão linear como os de Leitura como mostra o Gráfico 25, que representa os valores agregados em Matemática para diferentes proficiências iniciais e nos diferentes anos de escolarização. Esses valores também são retirados da Tabela 25.

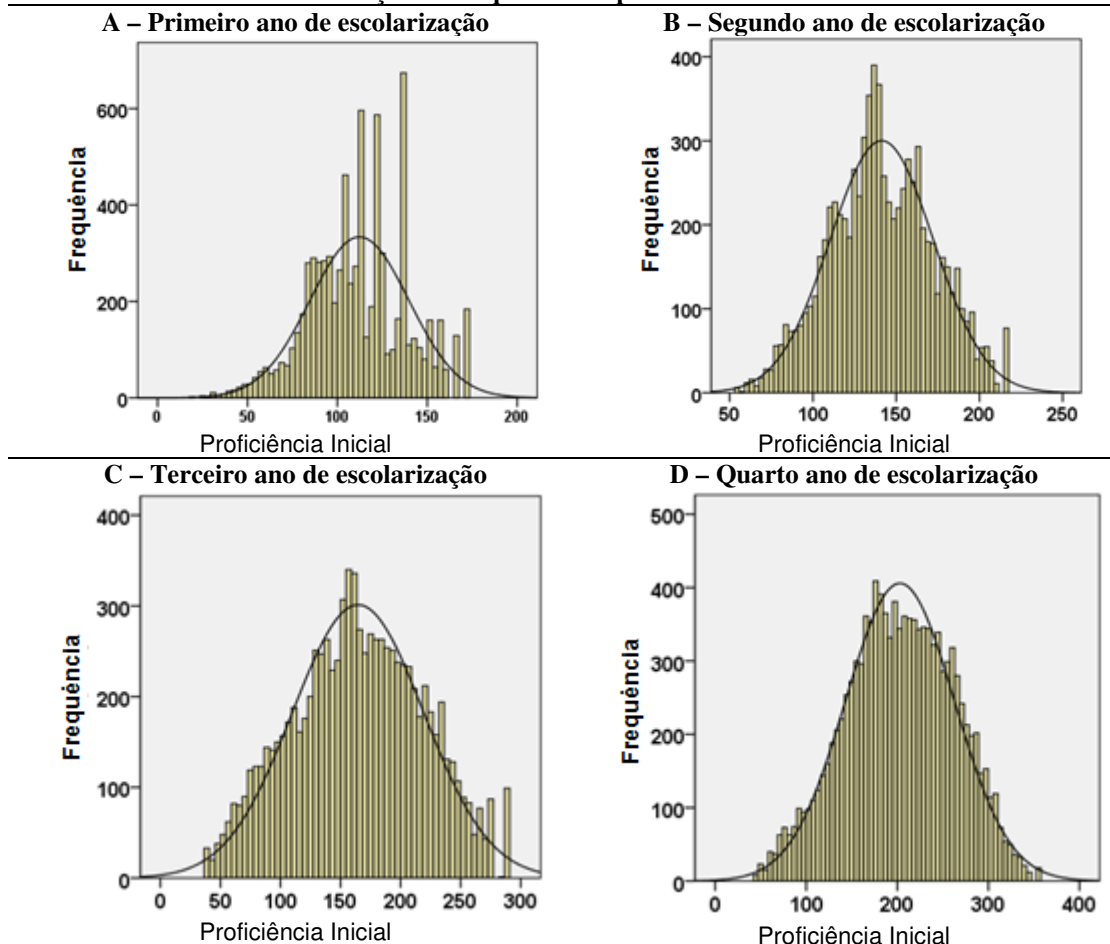
Gráfico 25 - Valor agregado médio em Matemática com base nos resultados do Modelo de Nível 1



Observa-se nesse gráfico que para os alunos com proficiência inicial baixa e média, o valor agregado decai do primeiro para o segundo ano e cresce nos demais, sendo que para os alunos com proficiência inicial alta, o valor agregado cresce nos três primeiros anos de escolarização e decai no último.

Também buscando entender esse resultado dos valores agregados em Matemática, foram criados os Gráficos 26A, 26B, 26C e 26D, onde estão representadas as distribuições das proficiências iniciais em Matemática para cada ano de escolarização.

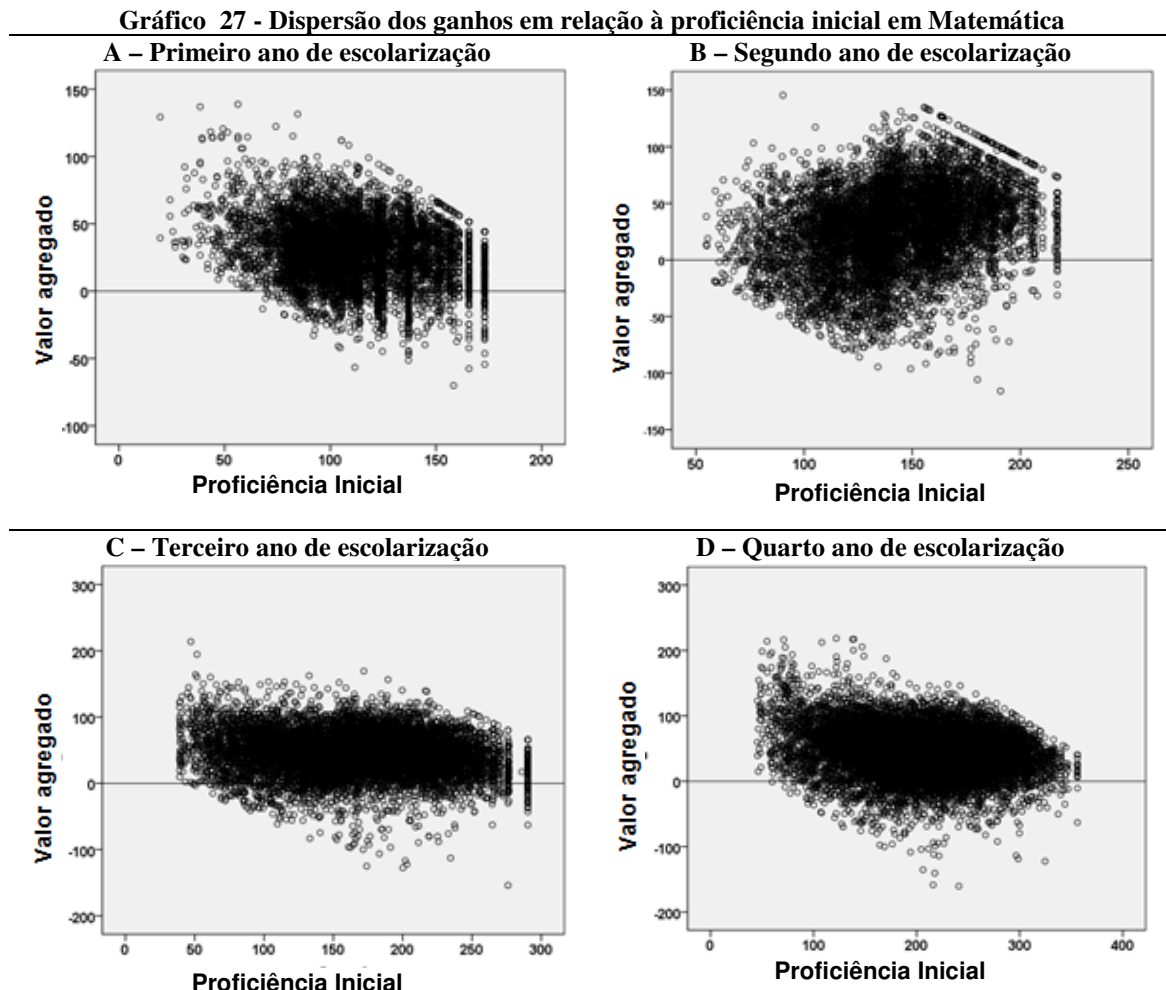
Gráfico 26 - Distribuição de frequência da proficiência inicial em Matemática



Pode-se observar que há diferença na distribuição de frequências do primeiro ano em relação aos demais, que levará a estimativas para o valor agregado em um comportamento também diferente nos demais anos. Ou seja, dado o modelo que coloca o primeiro ano como

referência para os demais anos, é razoável considerar que haja coeficientes estatisticamente significativos nas variáveis desses demais anos.

Para facilitar a análise do comportamento do valor agregado em função da proficiência inicial, foram elaborados os Gráficos 27A, 27B, 27C e 27D.



Apesar dessas dispersões não considerarem o peso amostral, pode-se observar que somente no primeiro ano há uma tendência para diminuir o valor agregado para as proficiências mais altas. Já nos demais anos, os valores agregados negativos se concentram nas proficiências iniciais intermediárias. Diante disso, mais uma vez fica evidenciado que as proficiências finais dos alunos dependem da proficiência inicial.

5.3) Interpretação dos resultados do Modelo de Nível 2

A interpretação dos valores obtidos nessa etapa traz algumas limitações uma vez que a representação gráfica já não é mais tão simples assim dado que a análise passa a ser multivariada.

Além disso, a análise do quanto e da forma com que as variáveis explicativas impactam na proficiência final não tem relevância nesse momento, uma vez que esses valores são alterados quando as variáveis explicativas de Nível 3 são inseridas, apenas é possível identificar os anos em que essas variáveis tem algum impacto.

Um último aspecto a destacar é que, assim como nos Modelos Nulos e de Referência, alguns aspectos relevantes desse modelo foram analisados no Capítulo anterior, pois eram necessários para a construção dos Modelos de Nível 3.

Diante desse contexto, os coeficientes encontrados no Modelo de Nível 2 serão explorados aqui para explicar como serão analisadas as variáveis explicativas que compõem o Modelo de Nível 3, afinal é nele que o quanto e a forma pelas quais as variáveis explicativa interferem na proficiência final dos alunos poderá ser estimada, considerando toda a complexidade e a interrelação entre elas, sejam de Nível 1, 2 ou 3.

Para essa análise foram percorridas quatro etapas:

a) são calculados três valores de referência para a proficiência final atribuindo o valor nulo a todas as variáveis explicativas e três distintos valores para a proficiência inicial considerando os intervalos de cada ano: baixa ($\bar{x} - \sigma$), média (\bar{x}) e alta ($\bar{x} + \sigma$). Os valores para a média (\bar{x}) e desvio padrão (σ) são obtidos da Tabela 24;

b) em seguida, uma nova proficiência final é calculada, porém atribuindo o valor de uma unidade para cada variável explicativa que permaneceram no modelo, uma a uma;

c) é calculado o valor agregado (positivo ou negativo) à proficiência inicial para cada unidade da variável explicativa para as diferentes proficiências iniciais;

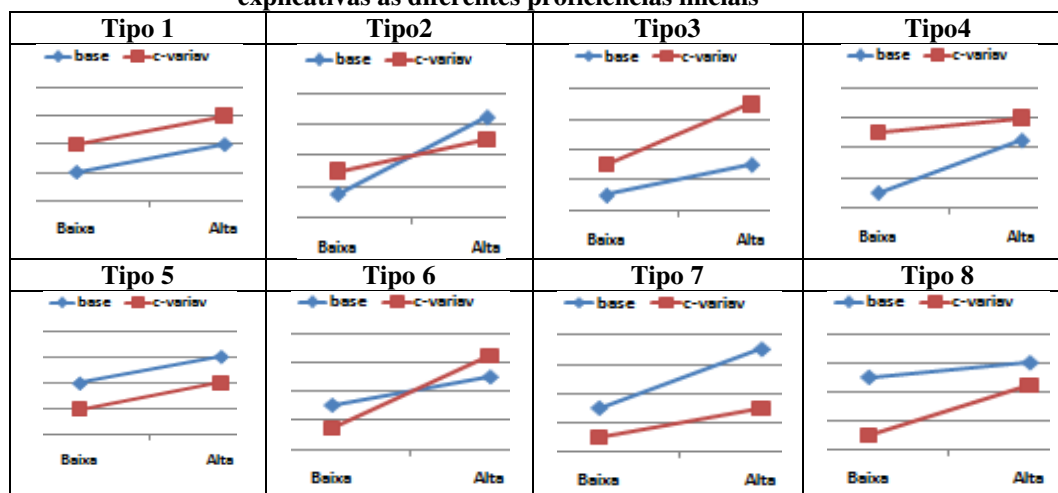
d) são comparados os valores do impacto dessa variável às proficiências alta e baixa e essa influência é classificado segundo o Quadro 41.

Quadro 41 - Possíveis tipos de influência das variáveis explicativas as diferentes proficiências iniciais

Tipo	Interpretação
1	Ganho igual independentemente da proficiência de entrada
2	Ganho quando proficiência é baixa e perda quando a proficiência é alta
3	Ganho menor na proficiência mais baixa do que na proficiência mais alta
4	Ganho maior na proficiência mais baixa do que na proficiência mais alta
5	Perda igual independentemente da proficiência de entrada
6	Perda quando proficiência é baixa e ganho quando a proficiência é alta
7	Perda menor na proficiência mais baixa do que na proficiência mais alta
8	Perda maior na proficiência mais baixa do que na proficiência mais alta

A Figura 13 traz a representação gráfica de cada uma dessas possibilidades descritas no Quadro 42 e que permite comparar a proficiência final considerando a variável explicativa quando lhe é atribuída uma unidade (*c-variav*) e a proficiência final quando o valor dessa mesma variável explicativa é nula (*base*).

Figura 13 - Representação gráfica dos possíveis tipos de influência das variáveis explicativas as diferentes proficiências iniciais



Do exposto até agora, fica evidente a forma pela qual a variável explicativa interfere na proficiência final depende da proficiência inicial e os valores apresentados indicam quanto cada unidade da variável explicativa impacta sobre a proficiência final.

5.3.1) Interpretação dos resultados do Modelo de Nível 2 em Leitura.

Nesta etapa do processo é construído o terceiro modelo, quando são mantidas as variáveis do Nível 1 e são inseridas as 11 variáveis explicativas do Nível 2 que estão vinculadas ao perfil do aluno.

Com base nos coeficientes apresentados no Capítulo 4, pode-se verificar quais são as variáveis explicativas do Nível 2 que foram significantes no Modelo. O Quadro 42 traz um agrupamento dessas variáveis em função dos anos em que seus coeficientes foram considerados estatisticamente significativos e ou tiveram algum impacto devido à significância dessa variável no primeiro ano de escolarização (ano base).

Quadro 42 - Relação entre as variáveis explicativas do Nível 2 em Leitura e os anos de escolarização que elas influenciaram

Teve coeficientes significativos em todos os anos de escolarização, e conseqüentemente impactaram em todos esses anos	FA06, FA08
Teve coeficientes significativos no primeiro e segundo ano, mas tem impacto em todos os anos, por ser estatisticamente significante no primeiro ano	FA04
Teve coeficientes significativos no primeiro, segundo e quarto ano, mas tem impacto em todos os anos, por ser estatisticamente significante no primeiro ano	FA07, FA09
Teve coeficientes significativos no primeiro e no terceiro, mas tem impacto em todos os anos, por ser estatisticamente significante no primeiro ano	FA11, FA13
Teve coeficientes significativos no primeiro e no quarto ano, mas tem impacto em todos os anos, por ser estatisticamente significante no primeiro ano	FA14
Teve coeficiente significativo apenas no primeiro, mas tem impacto em todos os anos, por ser estatisticamente significante no primeiro ano	FA01, FA12
Teve coeficiente significativo apenas no segundo e, portanto, impacta apenas nesse ano	FA10

O Quadro 43 traz uma síntese gráfica dessas informações.

Quadro 43 - Anos em que as variáveis de Nível 2 em Leitura impactam

	Ano1	Ano2	Ano3	Ano4
FA01*	((#))	#	#	#
FA04	((#))	((#))	#	#
FA06	((#))	((#))	((#))	((#))
FA07	((#))	((#))	#	((#))
FA08	((#))	((#))	((#))	((#))
FA09	((#))	((#))	#	((#))
FA10		((#))		
FA11*	((#))	#	((#))	#
FA12	((#))	#	#	#
FA13	((#))	#	((#))	#
FA14	((#))	#	#	((#))

Legenda:

((#)) a variável impacta nesse ano e tem um coeficiente significativo

a variável impacta nesse ano, mas não tem coeficiente significativo

* também influencia no intercepto do Ano 1 em não retidos

Como pode ser observado, com exceção de FA10, todas as variáveis de Nível 2 tem alguma influência na proficiência final de todos os anos de escolarização.

Como dito anteriormente, foi criado um modelo analítico onde é calculada a proficiência final a partir de três diferentes proficiências iniciais, cujos valores são comparados com aquele assumido quando é atribuída uma unidade na variável explicativa que está sendo analisada. Esses valores são apresentados no Quadro 44.

Quadro 44 - Valores de proficiência finais e valor agregado por variável do Nível 2 em Leitura

(continua)

Ano1	Proficiência Final	Profic. Inicial	Baixa	Média	Alta	Valor Agregado	Profic. Inicial		
		Base	114,6	158,8	203,0		Baixa	Média	Alta
		FA01	115,6	159,8	204,0		1,0	1,0	1,0
FA04	116,6	160,8	205,0	2,0	2,0	2,0			
FA06	117,3	164,2	211,0	2,7	5,4	8,0			
FA07	119,1	163,3	207,5	4,5	4,5	4,5			
FA08	117,0	163,7	210,3	2,4	4,9	7,3			
FA09	112,3	156,5	200,7	-2,3	-2,3	-2,3			
FA10	114,6	158,8	203,0						
FA11	115,6	159,8	204,0	1,0	1,0	1,0			
FA12	118,4	160,4	202,4	3,8	1,6	-0,6			
FA13	113,8	157,2	200,5	-0,8	-1,6	-2,5			
FA14	113,3	160,2	207,0	-1,3	1,3	4,0			
Ano2	Proficiência Final	Profic. Inicial	Baixa	Média	Alta	Valor Agregado	Profic. Inicial		
		Base	115,4	176,9	238,4		Baixa	Média	Alta
		FA01	116,4	177,9	239,4		1,0	1,0	1,0
FA04	114,3	190,4	266,5	-1,1	13,5	28,1			
FA06	116,5	179,1	241,7	1,1	2,2	3,4			
FA07	122,3	183,8	245,3	7,0	7,0	7,0			
FA08	116,3	178,6	241,0	0,9	1,8	2,6			
FA09	109,8	178,0	246,2	-5,6	1,1	7,8			
FA10	118,5	171,5	224,4	3,1	-5,4				
FA11	116,4	177,9	239,4	1,0	1,0	1,0			
FA12	119,2	178,5	237,8	3,8	1,6	-0,6			
FA13	114,6	175,2	235,9	-0,8	-1,6	-2,5			
FA14	114,1	178,2	242,4	-1,3	1,4	4,0			

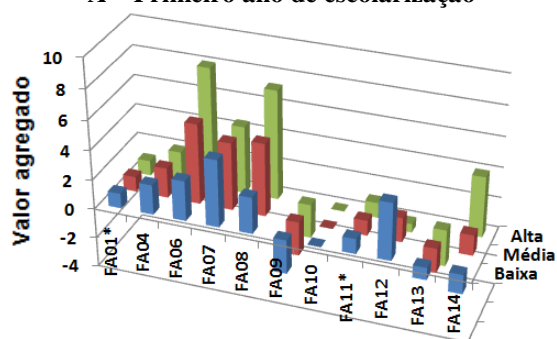
(continuação)

	Profic. Inicial	Baixa	Média	Alta	Valor Agregado	Profic. Inicial			
						Baixa	Média	Alta	
Ano3	Base	123,0	185,7	248,4					
	FA01	124,0	186,7	249,4		1,0	1,0	1,0	
	FA04	124,9	187,7	250,4		2,0	2,0	2,0	
	FA06	123,9	187,5	251,1		0,9	1,8	2,7	
	FA07	127,5	190,2	252,9		4,5	4,5	4,5	
	FA08	123,6	186,9	250,3		0,6	1,2	1,9	
	FA09	120,7	183,4	246,1		-2,3	-2,3	-2,3	
	FA10	123,0	185,7	248,4					
	FA11	125,9	188,7	251,4		3,0	3,0	3,0	
	FA12	126,8	187,3	247,9		3,8	1,6	-0,6	
	FA13	119,0	180,9	242,9		-3,9	-4,8	-5,6	
	FA14	121,7	187,0	252,4		-1,3	1,3	4,0	
	Ano4	Base	131,5	187,2	242,8				
		FA01	132,5	188,2	243,8		1,0	1,0	1,0
		FA04	133,5	189,1	244,8		2,0	2,0	2,0
FA06		135,5	188,7	242,0		4,0	1,6	-0,8	
FA07		139,9	190,3	240,7		8,4	3,2	-2,1	
FA08		132,8	189,6	246,4		1,2	2,4	3,7	
FA09		122,8	187,1	251,4		-8,8	-0,1	8,6	
FA10		131,5	187,2	242,8					
FA11		132,5	188,2	243,8		1,0	1,0	1,0	
FA12		135,3	188,8	242,2		3,8	1,6	-0,6	
FA13		130,7	185,5	240,3		-0,8	-1,6	-2,5	
FA14		129,6	187,2	244,9		-1,9	0,1	2,1	

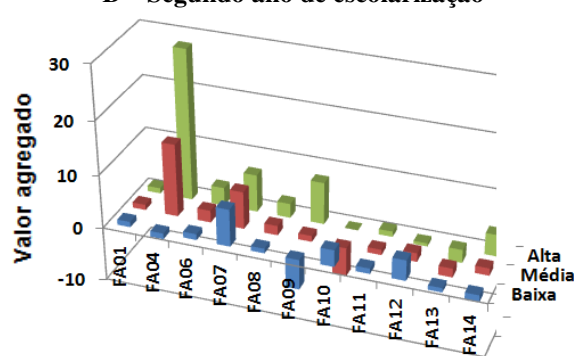
Os Gráfico 28A, 28B, 28C e 28D representam, para cada ano, o valor que cada unidade da variável explicativa de Nível 2 impacta na proficiência final de alunos com diferentes proficiências iniciais, isto é, representa o valor agregado devido a uma unidade da variável explicativa.

Gráfico 28 - Valor agregado por unidade da variável explicativa do Nível 2 em Leitura

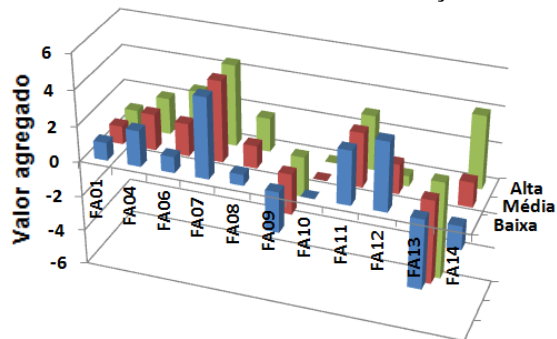
A – Primeiro ano de escolarização



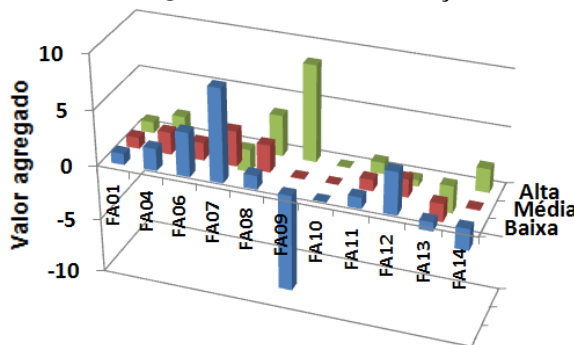
B – Segundo ano de escolarização



C – Terceiro ano de escolarização



D – Quarto ano de escolarização



O Quadro 45 traz uma síntese do impacto de cada uma das variáveis explicativas de Nível 2 na proficiência final em Leitura em cada um dos anos de escolarização.

Quadro 45 - Tipo de impacto da variável do Nível 2 em Leitura para cada ano

	Ano1	Ano2	Ano3	Ano4
FA01	Tipo 1	Tipo 1	Tipo 1	Tipo 1
FA04	Tipo 1	Tipo 6	Tipo 1	Tipo 1
FA06	Tipo 3	Tipo 3	Tipo 3	Tipo 2
FA07	Tipo 1	Tipo 1	Tipo 1	Tipo 2
FA08	Tipo 3	Tipo 3	Tipo 3	Tipo 3
FA09	Tipo 5	Tipo 6	Tipo 7	Tipo 6
FA10		Tipo 2		
FA11	Tipo 1	Tipo 1	Tipo 1	Tipo 1
FA12	Tipo 2	Tipo 2	Tipo 2	Tipo 2
FA13	Tipo 7	Tipo 7	Tipo 7	Tipo 7
FA14	Tipo 6	Tipo 6	Tipo 6	Tipo 6

Ainda desse quadro pode-se constatar que as variáveis FA01, FA08 e FA11 têm impactos esperados e semelhantes nos diferentes anos de escolarização. A FA13 tem impacto semelhante nos diferentes anos, mas, apesar da significância estatística encontrada, pedagogicamente seu resultado é incoerente em relação àquilo que é esperado. As demais variáveis influenciam diferentemente nas proficiências finais dos alunos com proficiência iniciais diferentes e, por isso, requerem um aprofundamento na análise de natureza qualitativa, não feita nesse estudo.

5.3.2) Interpretação dos resultados do Modelo de Nível 2 em Matemática.

Como visto no capítulo anterior, ao inserir as variáveis de Nível 1 para construir o Modelo de Nível 1, praticamente toda a variabilidade dos dados no Nível 2 foi explicada. Também foi visto que com a inserção das variáveis de Nível 2, como era de se esperar, não conseguiu só não explicar a variabilidade restante, e pelo contrário, chegou a aumentá-la, apesar de ter diminuído nos demais níveis.

Para efeito de análise, o Modelo de Nível 2 foi elaborado e seus resultados serão apresentados a seguir.

Quadro 46 - Relação entre as variáveis explicativas do Nível 2 em Matemática e os anos de escolarização que elas influenciaram

Teve coeficientes significativos em todos os anos de escolarização, impactando em todos esses anos	FA07, FA12
Teve coeficientes significativos apenas no primeiro ano, mas tem impacto em todos os anos, por ser estatisticamente significativa no primeiro ano	FA08, FA11, FA13
Teve coeficientes significativos no primeiro e segundo ano, mas tem impacto em todos os anos, por ser estatisticamente significativa no primeiro ano	FA04
Teve coeficientes significativos no primeiro e no quarto ano, mas tem impacto em todos os anos, por ser estatisticamente significativa no primeiro ano	FA01, FA14
Teve coeficientes significativos no primeiro, no segundo e quarto ano, mas tem impacto em todos os anos, por ser estatisticamente significativa no primeiro ano	FA09
Teve coeficiente significativo apenas no segundo e quarto anos, portanto, impactando apenas nesses anos	FA06
Não teve coeficiente significativo em nenhum dos anos e, portanto, não tem impactos	FA10

O Quadro 47 traz uma síntese gráfica dessas informações.

Quadro 47 - Anos em que as variáveis de Nível 2 em Matemática impactam

	Ano1	Ano2	Ano3	Ano4
FA01	((#))	#	#	((#))
FA04	((#))	((#))	#	#
FA06		((#))		((#))
FA07	((#))	((#))	((#))	((#))
FA08	((#))	#	#	#
FA09*	((#))	((#))	#	((#))
FA10				
FA11*	((#))	#	#	#
FA12*	((#))	((#))	((#))	((#))
FA13	((#))	#	#	#
FA14*	((#))	#	#	((#))

Legenda:

((#)) a variável impacta nesse ano e tem um coeficiente significativo

a variável impacta nesse ano mas não tem coeficiente

* também influencia no intercepto do Ano 1 em não retidos

Como pode ser observado, com exceção de FA10, todas as demais variáveis de Nível 2 tem alguma influência na proficiência final de todos os anos de escolarização.

Para a análise do quanto e de que forma essas variáveis explicativas interferem na proficiência final de Matemática, os resultados do modelo analítico descrito anteriormente estão apresentados no Quadro 48.

Quadro 48 - Valores de proficiência finais e valor agregado por variável do Nível 2 em Matemática

(continua)

Ano1	Proficiência Final	Profic. Inicial	Baixa	Média	Alta	Profic. Inicial		
		Base	118,7	130,8	142,9	Baixa	Média	Alta
		FA01	120,7	133,4	146,2	2,0	2,6	3,2
FA04	131,4	144,6	157,8	12,7	13,8	14,9		
FA06	118,7	130,8	142,9					
FA07	124,9	139,0	153,1	6,2	8,2	10,2		
FA08	119,5	131,9	144,3	0,8	1,1	1,3		
FA09*	113,8	125,9	138,0	-5,0	-5,0	-5,0		
FA10	118,7	130,8	142,9					
FA11*	115,9	128,0	140,1	-2,8	-2,8	-2,8		
FA12*	121,3	133,4	145,5	2,5	2,5	2,5		
FA13	114,0	126,1	138,2	-4,7	-4,7	-4,7		
FA14*	118,4	130,5	142,6	-0,4	-0,4	-0,4		

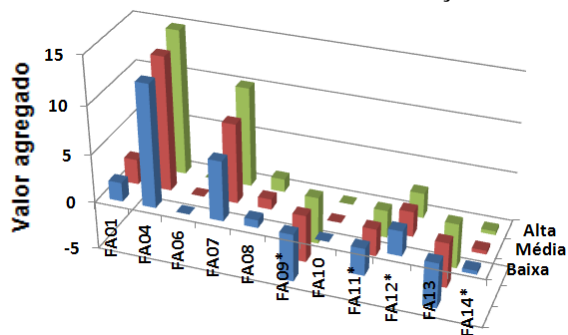
(Conclusão)

Ano2	Profic. Inicial	Baixa	Média	Alta	Valor Agregado	Profic. Inicial			
	Base	118,2	146,2	174,0		Baixa	Média	Alta	
Proficiência Final	FA01	120,8	149,5	178,0		2,5	3,3	4,0	
	FA04	129,5	164,2	198,8		11,2	18,0	24,8	
	FA06	125,2	153,2	181,0		7,0	7,0	7,0	
	FA07	160,1	194,4	228,5		41,8	48,2	54,6	
	FA08	119,3	147,5	175,6		1,0	1,3	1,6	
	FA09	106,5	137,9	169,0		-11,7	-8,3	-5,0	
	FA10	118,2	146,2	174,0					
	FA11	115,4	143,4	171,2		-2,8	-2,8	-2,8	
	FA12	126,0	155,4	184,7		7,8	9,2	10,7	
	FA13	113,5	141,5	169,3		-4,7	-4,7	-4,7	
	FA14	117,8	145,8	173,6		-0,4	-0,4	-0,4	
	Ano3	Profic. Inicial	Baixa	Média	Alta	Valor Agregado	Profic. Inicial		
		Base	150,1	192,5	234,8		Baixa	Média	Alta
	Proficiência Final	FA01	152,7	196,3	239,8		2,5	3,8	5,1
FA04		163,9	208,4	252,9		13,7	15,9	18,1	
FA06		150,1	192,5	234,8					
FA07		160,8	202,4	244,0		10,6	9,9	9,3	
FA08		151,2	194,0	236,8		1,0	1,6	2,1	
FA09		145,2	187,5	229,8		-5,0	-5,0	-5,0	
FA10		150,1	192,5	234,8					
FA11		147,3	189,6	231,9		-2,8	-2,8	-2,8	
FA12		155,2	198,7	242,3		5,0	6,3	7,5	
FA13		145,4	187,7	230,1		-4,7	-4,7	-4,7	
FA14		149,8	192,1	234,4		-0,4	-0,4	-0,4	
Ano4		Profic. Inicial	Baixa	Média	Alta	Valor Agregado	Profic. Inicial		
		Base	202,4	239,2	276,1		Baixa	Média	Alta
Proficiência Final		FA01	209,9	248,2	286,4		7,5	8,9	10,3
	FA04	217,4	256,7	295,9		15,0	17,5	19,9	
	FA06	205,5	243,6	281,8		3,1	4,4	5,7	
	FA07	208,3	247,6	287,0		5,9	8,4	10,9	
	FA08	203,8	241,2	278,6		1,4	1,9	2,5	
	FA09	190,7	227,6	264,4		-11,7	-11,7	-11,7	
	FA10	202,4	239,2	276,1					
	FA11	199,6	236,4	273,2		-2,8	-2,8	-2,8	
	FA12	208,7	247,2	285,6		6,4	7,9	9,5	
	FA13	197,7	234,5	271,4		-4,7	-4,7	-4,7	
	FA14	199,6	239,5	279,4		-2,8	0,3	3,3	

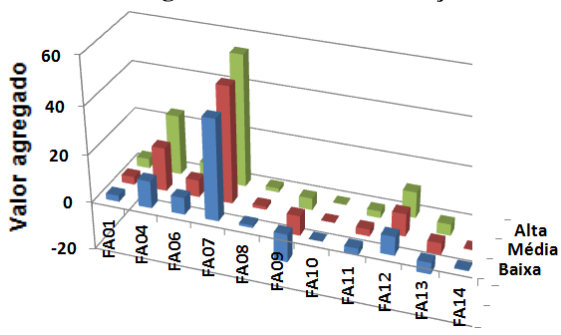
Os Gráficos 29A, 29B, 29C e 29D representam, para cada ano, o valor que cada unidade da variável explicativa de Nível 2 impacta na proficiência final de alunos com diferentes proficiências iniciais, isto é, representa o valor agregado devido a uma unidade da variável explicativa.

Gráfico 29 - Valor agregado por unidade da variável explicativa do Nível 2 em Matemática

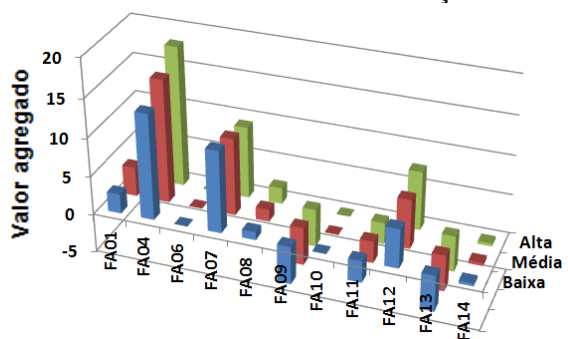
A – Primeiro ano de escolarização



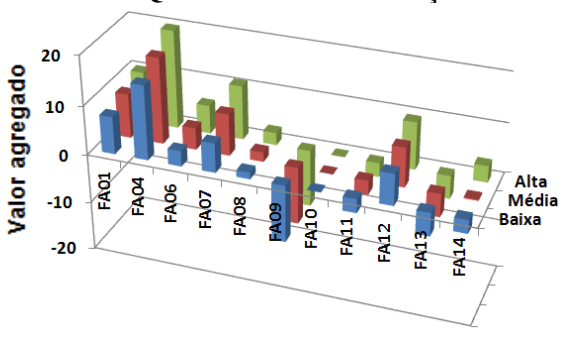
B – Segundo ano de escolarização



C – Terceiro ano de escolarização



D – Quarto ano de escolarização



O Quadro 49 traz uma síntese do impacto de cada uma das variáveis explicativas de Nível 2 na proficiência final em cada um dos anos de escolarização.

Quadro 49 - Tipo de impacto da variável do Nível 2 em Matemática a cada ano

	Ano1	Ano2	Ano3	Ano4
FA01	Tipo 3	Tipo 3	Tipo 3	Tipo 3
FA04	Tipo 3	Tipo 3	Tipo 3	Tipo 3
FA06		Tipo 1		Tipo 3
FA07	Tipo 3	Tipo 3	Tipo 4	Tipo 3
FA08	Tipo 3	Tipo 3	Tipo 3	Tipo 3
FA09*	Tipo 5	Tipo 8	Tipo 5	Tipo 5
FA10				
FA11*	Tipo 5	Tipo 8	Tipo 5	Tipo 8
FA12*	Tipo 1	Tipo 3	Tipo 3	Tipo 3
FA13	Tipo 5	Tipo 8	Tipo 5	Tipo 8
FA14*	Tipo 5	Tipo 8	Tipo 5	Tipo 6

Desse quadro pode-se constatar que as variáveis FA01 e FA04 têm impactos esperados e semelhantes nos diferentes anos de escolarização. As variáveis FA09, FA11, FA12 e FA13 têm impactos semelhantes nos diferentes anos, no entanto, apesar da significância estatística encontrada, pedagogicamente seus resultados não condizem com o que é esperado. A FA10 não teve qualquer impacto e as demais variáveis influenciam diferentemente as proficiências finais dos alunos com proficiência iniciais diferentes e por isso requerem um aprofundamento na análise, não feita neste estudo.

5.4) Interpretação dos resultados do Modelo de Nível 3

No Modelo de Nível 3 foram inseridas as 59 variáveis no intercepto e na inclinação de cada um dos anos de escolarização. Seguindo o método *backward*, as variáveis menos significantes (segundo o *p-value* estimado) foram sendo retiradas até que permanecessem apenas aquelas com significância estatística. Como já sinalizado no capítulo anterior, ao se perceber que a ordem da inserção das variáveis interferem no resultado, optou-se por esse método, para que se verifique, quais são as variáveis que permaneceriam no modelo, independentemente do seu significado pedagógico, segundo os dados disponíveis, para posterior interpretação pedagógica das variáveis que permaneceram.

O Modelo de Nível 3 é aquele que contém todas as variáveis explicativas, consideradas estatisticamente significantes nos três níveis, de maneira que, os valores estimados para os coeficientes a elas associados são os que melhor representam as influências dessas variáveis considerando todas às outras, isto é, o valor estimado para um coeficiente é único nesse contexto, pois certamente qualquer outra variável que não estivesse presente esse valor seria outro.

Para melhor organização do texto, dada a extensão das tabelas, a análise dos resultados dos Modelos de Nível 3 que se seguem, é feita separadamente em quatro grupos de variáveis explicativas.

a) Grupo 1, encontram-se a variável que traz o nível socioeconômico da escola, determinado pela média do nível socioeconômico de seus alunos e pelas variáveis que foram sintetizadas dos questionários de contexto sobre a escola;

b) Grupo 2, estão as variáveis originadas dos questionários de contexto aplicado nos diretores;

c) Grupo 3, estão as variáveis sintetizadas a partir do questionário de contexto dos professores e estão ligadas às práticas pedagógicas realizadas na escola.

d) Grupo 4, estão as variáveis sintetizadas a partir do questionário de contexto dos professores e caracterizam a dinâmica e o perfil da escola.

5.4.1) Interpretação dos resultados do Modelo de Nível 3 em Leitura

Como citado anteriormente, as análises a seguir são feitas com base em três diferentes proficiências iniciais considerando todas as variáveis explicativas do Nível 3 nulas, cujos valores são comparados, individualmente, com aquele assumido quando a cada uma delas é atribuído o valor de uma unidade. Os coeficientes que compõem esse modelo que auxilia a análise são retirados do modelo apresentado no Capítulo 4.

5.4.1.1) Análise das variáveis do Modelo de Nível 3 em Leitura pertencentes ao Grupo 1 (nível socioeconômico médio e outras características oriundas do questionário de contexto da escola)

Este grupo de variáveis de Nível 3 é composto por 11 das 59 variáveis existentes. O Quadro 50 apresenta o ano no qual essas variáveis foram significantes, explicitando aqueles nos quais elas tiveram seu coeficiente significativo.

Quadro 50 - Anos em que as variáveis do Grupo 1 do Nível 3 em Leitura impactam

	Coeficiente significativo e impacto			
	Ano1	Ano2	Ano3	Ano4
FEA01	((#))	((#))	((#))	((#))
FEE02	((#))	((#))	((#))	((#))
FEE03		((#))	((#))	
FEE04	((#))	#	#	((#))
FEE05	((#))	#	#	#
FEE06		((#))		((#))
FEE07		((#))		((#))
FEE08	((#))	#	#	#
FEE09				((#))
FEE10		((#))	((#))	
FEE11			((#))	((#))

Legenda:

((#)) a variável impacta nesse ano e tem um coeficiente significativo

a variável impacta nesse ano, mas não tem coeficiente

* também influencia no intercepto do Ano 1 em não retidos

Desse quadro, pode se observar que:

Quadro 51 - Relação entre as variáveis explicativas do Grupo 1 do Nível 3 em Leitura e os anos de escolarização que elas influenciaram

Teve coeficientes significativos em todos os anos de escolarização, impactando em todos esses anos	FEA01, FEE02
Teve coeficientes significativos apenas no primeiro ano, mas tem impacto em todos os anos, por ser estatisticamente significante no primeiro ano	FEE05, FEE08
Teve coeficientes significativos no primeiro e no quarto ano, mas tem impacto em todos os anos, por ser estatisticamente significante no primeiro ano	FEE04
Teve coeficiente significativo apenas no segundo e terceiro anos, portanto, impactando apenas nesses anos	FEE03, FEE10
Teve coeficiente significativo apenas no segundo e quarto anos, portanto, impactando apenas nesses anos	FEE06, FEE07
Teve coeficiente significativo apenas no terceiro e quarto anos, portanto, impactando apenas nesses anos	FEE11
Teve coeficiente significativo apenas no quarto ano, portanto, impactando apenas nesse ano	FEE09

Essas sínteses são produzidas a partir da análise do valor agregado que a variável provoca na proficiência final quando o valor de uma unidade é atribuído à variável explicativa, cujos valores são apresentados no Quadro 52.

Quadro 52 - Valores de proficiência finais e valor agregado por variável explicativa do Grupo 1 do Nível 3 em Leitura
(continua)

Ano1	Profic. Inicial	Baixa	Média	Alta	Valor Agregado	Profic. Inicial			
		Base	102,7	123,6		144,5	Baixa	Média	Alta
Ano1	Proficiência Final	FEA01	115,0	132,2	149,3	12,4	8,6	4,9	
		FEE02	104,5	123,2	142,0	1,8	-0,4	-2,5	
		FEE03	102,7	123,6	144,5				
		FEE04	100,9	122,2	143,6	-1,8	-1,4	-0,9	
		FEE05	102,2	123,1	143,9	-0,5	-0,5	-0,5	
		FEE06	102,7	123,6	144,5				
		FEE07	102,7	123,6	144,5				
		FEE08	104,9	125,3	145,6	2,2	1,7	1,2	
		FEE09	102,7	123,6	144,5				
		FEE10	102,7	123,6	144,5				
		FEE11	102,7	123,6	144,5				
		Ano2	Profic. Inicial	Baixa	Média	Alta	Valor Agregado	Profic. Inicial	
Base	131,2			147,0	162,8	Baixa		Média	Alta
FEA01	134,6			148,8	163,0	3,5		1,8	0,2
FEE02	129,5			147,4	165,3	-1,6		0,4	2,5
FEE03	132,3			147,5	162,7	1,1		0,5	-0,2
FEE04	129,7			145,9	162,2	-1,4		-1,0	-0,6
FEE05	130,6			146,4	162,3	-0,5		-0,5	-0,5
FEE06	130,2			146,1	161,9	-0,9		-0,9	-0,9
FEE07	133,1			149,0	164,8	2,0		2,0	2,0
FEE08	132,9			148,3	163,6	1,8		1,3	0,8
FEE09	131,2			147,0	162,8				
FEE10	130,4			146,8	163,1	-0,8		-0,2	0,3
FEE11	131,2	147,0	162,8						
Ano3	Profic. Inicial	Baixa	Média	Alta	Valor Agregado	Profic. Inicial			
		Base	155,8	166,9		177,9	Baixa	Média	Alta
		FEA01	161,6	171,6		181,6	5,8	4,7	3,7
		FEE02	156,2	166,6		177,0	0,3	-0,3	-1,0
		FEE03	155,3	166,3		177,4	-0,5	-0,5	-0,5
		FEE04	154,6	166,1		177,6	-1,2	-0,8	-0,3
		FEE05	155,3	166,3		177,4	-0,5	-0,5	-0,5
		FEE06	155,8	166,9		177,9			
		FEE07	155,8	166,9		177,9			
		FEE08	157,4	167,9		178,4	1,5	1,0	0,4
		FEE09	155,8	166,9		177,9			
		FEE10	155,1	166,0		176,9	-0,7	-0,9	-1,0
FEE11	154,7	165,5	176,3	-1,1	-1,4	-1,6			

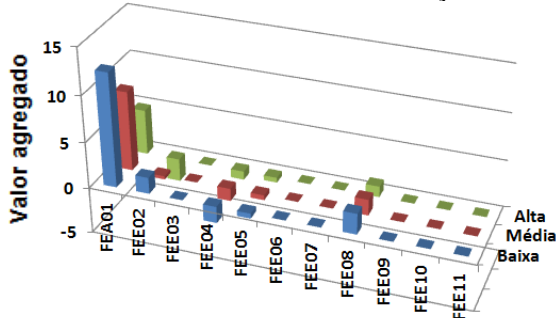
(continua)

Ano4	Profic. Inicial	Profic. Inicial	Baixa	Média	Alta	Profic. Inicial		
		Base	149,8	166,3	182,9	Baixa	Média	Alta
		FEA01	156,6	173,0	189,4	6,8	6,7	6,6
	FEE02	151,8	166,2	180,6	2,0	-0,1	-2,3	
	FEE03	149,8	166,3	182,9				
	FEE04	149,5	166,0	182,5	-0,3	-0,4	-0,4	
	FEE05	149,3	165,8	182,3	-0,5	-0,5	-0,5	
	FEE06	149,2	165,6	182,1	-0,6	-0,7	-0,8	
	FEE07	152,7	166,9	181,1	2,8	0,6	-1,7	
	FEE08	151,1	167,1	183,1	1,3	0,7	0,2	
	FEE09	150,4	167,1	183,7	0,6	0,7	0,8	
	FEE10	149,8	166,3	182,9				
	FEE11	150,7	166,2	181,7	0,8	-0,2	-1,2	

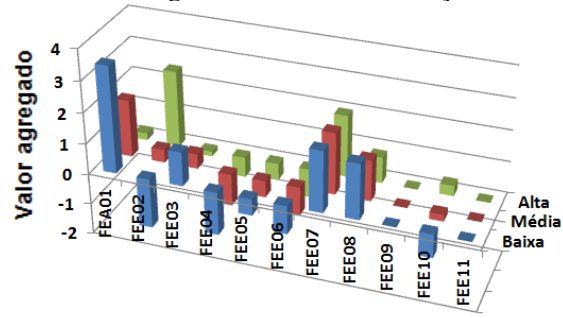
Os Gráficos 30A, 30B, 30C e 30D representam, para cada ano, o valor que cada unidade da variável explicativa de Nível 3 impacta na proficiência final de alunos com diferentes proficiências iniciais, isto é, representa o valor agregado devido a uma unidade da variável explicativa.

Gráfico 30 - Valor agregado por unidade da variável explicativa do Grupo 1 do Nível 3 em Leitura

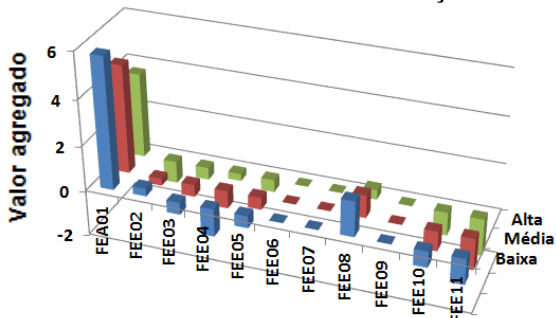
A – Primeiro ano de escolarização



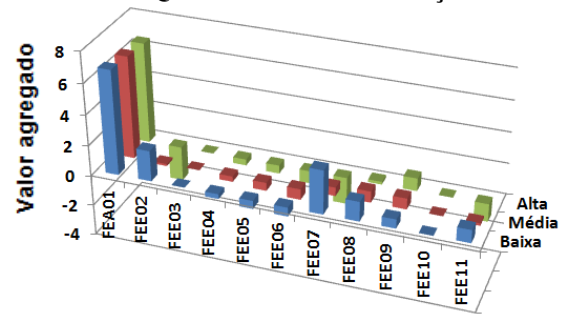
B – Segundo ano de escolarização



C – Terceiro ano de escolarização



D – Quarto ano de escolarização



O Quadro 53 traz uma síntese do impacto de cada uma das variáveis explicativas de Nível 3 na proficiência final em cada um dos anos de escolarização.

Quadro 53 - Tipo de impacto das variáveis do Grupo 1 do Nível 3 em Leitura a cada ano

	Ano1	Ano2	Ano3	Ano4
FEA01	Tipo 4	Tipo 4	Tipo 4	Tipo 4
FEE02	Tipo 2	Tipo 6	Tipo 2	Tipo 2
FEE03		Tipo 2	Tipo 5	
FEE04	Tipo 8	Tipo 8	Tipo 8	Tipo 7
FEE05	Tipo 5	Tipo 5	Tipo 5	Tipo 5
FEE06		Tipo 5		Tipo 7
FEE07		Tipo 1		Tipo 2
FEE08	Tipo 4	Tipo 4	Tipo 4	Tipo 4
FEE09				Tipo 3
FEE10		Tipo 6	Tipo 7	
FEE11			Tipo 7	Tipo 2

O Quadro 54 apresenta a síntese na interpretação do comportamento dessas variáveis.

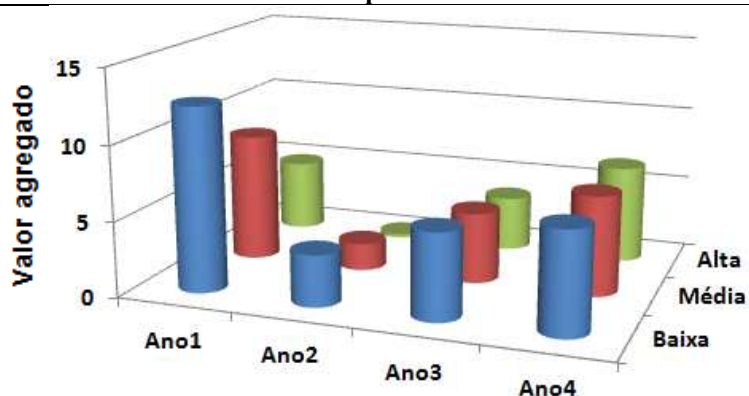
Quadro 54 - Interpretação final das variáveis do Grupo 1 do Nível 3 em Leitura

Comportamento	Variáveis
Coerente com relação ao esperado	FEA01, FEE08, FEE09
Incoerente com o esperado	FEE04, FEE05, FEE06,
Instável, dificultando a interpretação	FEE02, FEE03, FEE07, FEE10, FEE11

A FEA01, que controla o nível socioeconômico da escola requer uma atenção especial, pois indica a influência que tem na proficiência final em todos os anos e parece mostrar que as escolas, em média, têm uma tendência para ser equânimes.

O Gráfico 31 apresenta os valores que cada unidade do valor do nível socioeconômico médio da escola agrega à proficiência final dos alunos.

Gráfico 31 - Valor agregado médio devido ao nível socioeconômico médio das escolas na proficiência final em Leitura



5.4.1.2) Análise das variáveis do Modelo de Nível 3 em Leitura pertencentes ao Grupo 2 (características da escola extraídas do questionário de contexto do diretor)

Este grupo de variáveis de Nível 3 é composto por 15 das 59 variáveis existentes. Quadro 55 apresenta uma síntese do ano no qual essas variáveis foram significantes e o tipo de impacto que essa variável teve sobre a proficiência final em Leitura.

Quadro 55 - Relação entre as variáveis explicativas do Grupo 2 do Nível 3 em Leitura e os anos de escolarização que elas influenciaram

Teve coeficientes significativos em todos os anos de escolarização, impactando em todos esses anos	FED14, FED25*
Teve coeficientes significativos no primeiro e quarto anos, mas tem impacto em todos os anos, por ser estatisticamente significativa no primeiro ano	FED13
Teve coeficientes significativos no primeiro, no segundo e no quarto ano, mas tem impacto em todos os anos, por ser estatisticamente significativa no primeiro ano	FED23*
Teve coeficientes significativos no primeiro, no terceiro e no quarto ano, mas tem impacto em todos os anos, por ser estatisticamente significativa no primeiro ano	FED12, FED22*
Teve coeficiente significativo apenas no segundo, portanto, impactando apenas nesse ano	FED20
Teve coeficiente significativo apenas no terceiro, portanto, impactando apenas nesse ano	FED15
Teve coeficiente significativo apenas no quarto, portanto, impactando apenas nesse ano	FED17
Teve coeficiente significativo apenas no terceiro e no quarto anos, portanto, impactando apenas nesses anos	FED18
Teve coeficiente significativo apenas no segundo, no terceiro e quarto ano, portanto, impactando apenas nesses anos	FED19
Teve coeficiente significativo apenas no segundo e terceiro anos, portanto, impactando apenas nesses anos	FED24*, FED26*
Teve coeficiente significativo apenas no segundo e quarto anos, portanto, impactando apenas nesses anos	FED21
Não teve coeficiente significativo em nenhum dos anos e, portanto, não tem impactos	FED16

O Quadro 56 traz uma síntese gráfica dessas informações.

Quadro 56 - Anos em que as variáveis do Grupo 2 do Nível 3 em Leitura impactam

	Ano1	Ano2	Ano3	Ano4
FED12	((#))	#	((#))	((#))
FED13	((#))	#	#	((#))
FED14	((#))	((#))	((#))	((#))
FED15			((#))	
FED16				
FED17				((#))
FED18			((#))	((#))
FED19		((#))	((#))	((#))
FED20		((#))		
FED21		((#))		((#))
FED22*	((#))	#	((#))	((#))
FED23*	((#))	((#))	#	((#))
FED24*		((#))	((#))	
FED25*	((#))	((#))	((#))	((#))
FED26*		((#))	((#))	

Legenda:

((#)) a variável impacta nesse ano e tem um coeficiente significativo

a variável impacta nesse ano mas não tem coeficiente

Essas sínteses são produzidas a partir da análise do valor agregado que a variável provoca na proficiência final quando o valor de uma unidade é atribuído à variável explicativa, cujos valores são apresentados no Quadro 57.

Quadro 57 - Valores de proficiência finais e valor agregado por variável explicativa do Grupo 2 do Nível 3 em Leitura
(continua)

Ano1	Profic. Inicial	Profic. Inicial	Profic. Inicial	Profic. Inicial	Profic. Inicial			
		Base	Baixa	Média	Alta	Baixa	Média	Alta
		102,7	123,6	144,5				
	FED12	100,4	122,5	144,7	-2,3	-1,1	0,2	
	FED13	105,0	124,8	144,6	2,3	1,2	0,2	
	FED14	105,0	123,6	142,2	2,3	0,0	-2,3	
	FED15	102,7	123,6	144,5	0,0	0,0	0,0	
	FED16	102,7	123,6	144,5	0,0	0,0	0,0	
	FED17	102,7	123,6	144,5	0,0	0,0	0,0	
	FED18	102,7	123,6	144,5	0,0	0,0	0,0	
	FED19	102,7	123,6	144,5	0,0	0,0	0,0	
	FED20*	102,7	123,6	144,5	0,0	0,0	0,0	
	FED21*	102,7	123,6	144,5	0,0	0,0	0,0	
	FED22*	107,4	126,4	145,3	4,7	2,8	0,8	
	FED23*	100,8	122,7	144,5	-1,9	-0,9	0,0	
	FED24*	102,7	123,6	144,5	0,0	0,0	0,0	
	FED25*	103,8	123,8	143,8	1,1	0,2	-0,7	
	FED26*	102,7	123,6	144,5	0,0	0,0	0,0	

(continua)

	Profic. Inicial	Baixa	Média	Alta	Profic. Inicial		
					Baixa	Média	Alta
Ano2	Base	131,2	147,0	162,8			
	FED12	129,9	146,8	163,7	-1,2	-0,2	0,9
	FED13	132,5	147,4	162,3	1,4	0,5	-0,5
	FED14	128,9	146,0	163,0	-2,3	-1,0	0,2
	FED15	131,2	147,0	162,8	0,0	0,0	0,0
	FED16	131,2	147,0	162,8	0,0	0,0	0,0
	FED17	131,2	147,0	162,8	0,0	0,0	0,0
	FED18	131,2	147,0	162,8	0,0	0,0	0,0
	FED19	131,1	147,7	164,2	0,0	0,7	1,4
	FED20*	129,6	146,6	163,6	-1,5	-0,3	0,8
	FED21*	128,6	144,5	160,3	-2,5	-2,5	-2,5
	FED22*	134,2	148,4	162,5	3,1	1,4	-0,3
	FED23*	129,3	145,8	162,3	-1,9	-1,2	-0,5
	FED24*	132,6	147,9	163,3	1,4	1,0	0,5
	FED25*	130,1	146,2	162,4	-1,0	-0,7	-0,4
FED26*	130,6	146,8	163,1	-0,6	-0,2	0,3	
Ano3	Base	155,8	166,9	177,9			
	FED12	156,9	165,9	174,9	1,0	-1,0	-3,0
	FED13	156,7	166,7	176,7	0,9	-0,2	-1,3
	FED14	155,4	166,0	176,7	-0,4	-0,8	-1,2
	FED15	154,9	166,9	178,9	-0,9	0,0	1,0
	FED16	155,8	166,9	177,9	0,0	0,0	0,0
	FED17	155,8	166,9	177,9	0,0	0,0	0,0
	FED18	155,4	166,3	177,3	-0,5	-0,6	-0,7
	FED19	157,3	167,3	177,4	1,5	0,5	-0,6
	FED20*	155,8	166,9	177,9	0,0	0,0	0,0
	FED21*	155,8	166,9	177,9	0,0	0,0	0,0
	FED22*	156,7	167,1	177,5	0,9	0,2	-0,5
	FED23*	155,2	167,2	179,3	-0,7	0,4	1,4
	FED24*	156,2	167,3	178,4	0,4	0,4	0,4
	FED25*	156,8	167,2	177,6	1,0	0,3	-0,4
FED26*	155,5	166,5	177,6	-0,3	-0,3	-0,3	

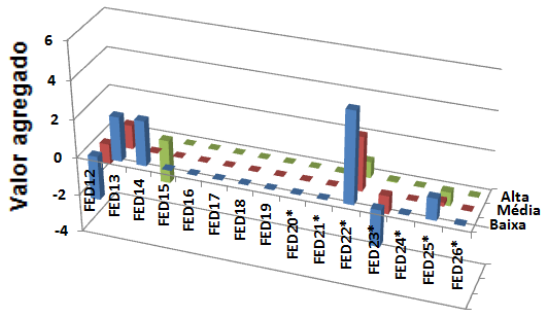
(conclusão)

Ano4	Proficiência Final	Profic. Inicial	Baixa	Média	Alta	Profic. Inicial		
		Base	149,8	166,3	182,9	Baixa	Média	Alta
		FED12	147,8	165,3	182,7	-2,0	-1,1	-0,2
FED13	151,1	167,7	184,3	1,3	1,4	1,4		
FED14	148,3	165,2	182,1	-1,5	-1,1	-0,7		
FED15	149,8	166,3	182,9	0,0	0,0	0,0		
FED16	149,8	166,3	182,9	0,0	0,0	0,0		
FED17	152,9	167,4	181,8	3,1	1,0	-1,0		
FED18	149,3	165,9	182,4	-0,5	-0,5	-0,5		
FED19	150,4	167,1	183,7	0,6	0,7	0,8		
FED20*	149,8	166,3	182,9	0,0	0,0	0,0		
FED21*	148,3	164,8	181,3	-1,5	-1,5	-1,5		
FED22*	151,1	167,4	183,6	1,3	1,0	0,8		
FED23*	148,0	165,5	183,0	-1,8	-0,8	0,2		
FED24*	149,8	166,3	182,9	0,0	0,0	0,0		
FED25*	148,9	165,8	182,8	-0,9	-0,5	-0,1		
FED26*	149,8	166,3	182,9	0,0	0,0	0,0		

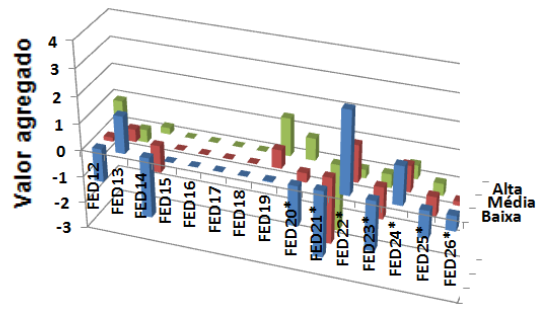
Os Gráficos 31A, 31B, 31C e 31D representam, para cada ano, o valor que cada unidade da variável explicativa de Nível 3 impacta na proficiência final de alunos com diferentes proficiências iniciais

Gráfico 32 - Valor agregado por unidade da variável explicativa do Grupo 2 do Nível 3 em Leitura

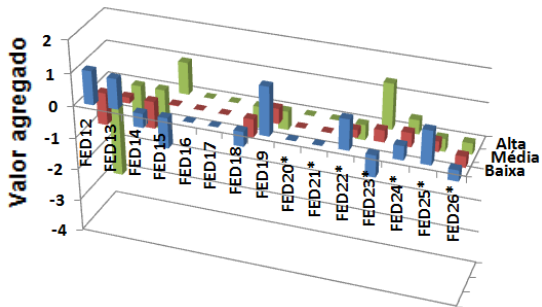
A – Primeiro ano de escolarização



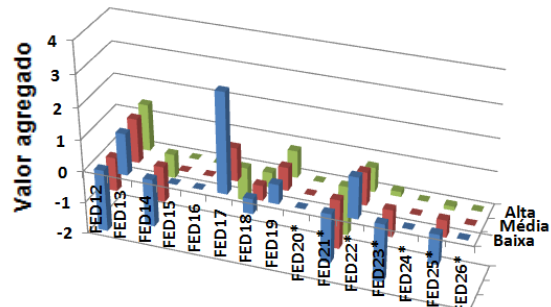
B – Segundo ano de escolarização



C – Terceiro ano de escolarização



D – Quarto ano de escolarização



O Quadro 58 traz uma síntese do impacto de cada uma das variáveis explicativas de Nível 3 na proficiência final em cada um dos anos de escolarização.

Quadro 58 - Tipo de impacto das variáveis do Grupo 2 do Nível 3 em Leitura a cada ano

	Ano1	Ano2	Ano3	Ano4
FED12	Tipo 6	Tipo 6	Tipo 2	Tipo 8
FED13	Tipo 4	Tipo 2	Tipo 2	Tipo 3
FED14	Tipo 2	Tipo 6	Tipo 7	Tipo 8
FED15			Tipo 6	
FED16				
FED17				Tipo 2
FED18			Tipo 7	Tipo 5
FED19		Tipo 6	Tipo 2	Tipo 3
FED20		Tipo 6		
FED21		Tipo 5		Tipo 5
FED22*	Tipo 4	Tipo 2	Tipo 2	Tipo 4
FED23*	Tipo 6	Tipo 8	Tipo 6	Tipo 6
FED24*		Tipo 4	Tipo 1	
FED25*	Tipo 2	Tipo 8	Tipo 2	Tipo 8
FED26*		Tipo 6	Tipo 5	

O Quadro 59 apresenta a síntese na interpretação do comportamento dessas variáveis.

Quadro 59 - Interpretação final das variáveis do Grupo 2 do Nível 3 em Leitura

Comportamento	Variáveis
Coerente com o esperado	FED19, FED21, FED24*
Incoerente com o esperado	FED15, FED18
Instável, dificultando a interpretação	FED12, FED13, FED14, FED17, FED20, FED22*, FED23*, FED25*, FED26*
Não influenciou	FED16

5.4.1.3) Análise das variáveis do Modelo de Nível 3 em Leitura pertencentes ao Grupo 3 (características da escola extraídas do questionário de contexto do professor)

Este grupo de variáveis de Nível 3 é composto por 18 das 50 variáveis existentes. O Quadro 60 apresenta uma síntese do ano no qual essas variáveis foram significantes.

Quadro 60 - Relação entre as variáveis explicativas do Grupo 3 do Nível 3 em Leitura e os anos de escolarização que elas influenciaram

Teve coeficientes significativos em todos os anos de escolarização, impactando em todos esses anos	FEP33, FEP39, FEP40
Teve coeficientes significativos no primeiro e segundo ano, mas tem impacto em todos os anos, por ser estatisticamente significante no primeiro ano	FEP54
Teve coeficientes significativos no primeiro e terceiro ano, mas tem impacto em todos os anos, por ser estatisticamente significante no primeiro ano	FEP42, FEP43
Teve coeficientes significativos no primeiro e quarto ano, mas tem impacto em todos os anos, por ser estatisticamente significante no primeiro ano	FEP36, FEP37, FEP41, FEP46, FEP47
Teve coeficientes significativos no primeiro, segundo e quarto ano, mas tem impacto em todos os anos, por ser estatisticamente significante no primeiro ano	FEP38, FEP45, FEP48
Teve coeficiente significativo apenas no segundo e quarto anos, portanto, impactando apenas nesses anos	FEP52, FEP53
Teve coeficiente significativo apenas no terceiro e quarto anos, portanto, impactando apenas nesses anos	FEP44
Teve coeficiente significativo apenas no segundo, terceiro e quarto anos, portanto, impactando apenas nesses anos	FEP49

O Quadro 61 traz uma síntese gráfica dessas informações.

Quadro 61 - Anos em que as variáveis do Grupo 3 do Nível 3 impactam em Leitura

	Ano1	Ano2	Ano3	Ano4
FEP33	((#))	((#))	((#))	((#))
FEP36	((#))	#	#	((#))
FEP37	((#))	#	#	((#))
FEP38	((#))	((#))	#	((#))
FEP39	((#))	((#))	((#))	((#))
FEP40	((#))	((#))	((#))	((#))
FEP41	((#))	#	#	((#))
FEP42	((#))	#	((#))	#
FEP43	((#))	#	((#))	#
FEP44			((#))	((#))
FEP45	((#))	((#))	#	((#))
FEP46	((#))	#	#	((#))
FEP47	((#))	#	#	((#))
FEP48	((#))	((#))	#	((#))
FEP49		((#))	((#))	((#))
FEP52		((#))		((#))
FEP53		((#))		((#))
FEP54	((#))	((#))	#	#

Legenda:

((#)) a variável impacta nesse ano e tem um coeficiente significativo

a variável impacta nesse ano mas não tem coeficiente

Essas sínteses são produzidas a partir da análise do valor agregado que a variável provoca na proficiência final quando o valor de uma unidade é atribuído à variável explicativa, cujos valores são apresentados no Quadro 62.

Quadro 62 - Valores de proficiência finais e valor agregado por variável explicativa do Grupo 3 do Nível 3 em Leitura
(continua)

	Ano1				Ano2			
	Profic. Inicial	Baixa	Média	Alta	Profic. Inicial	Baixa	Média	Alta
Proficiência Final	Base	102,7	123,6	144,5	Base	131,2	147,0	162,8
	FEP33	100,8	123,4	145,9	FEP33	131,9	146,0	160,2
	FEP36	104,1	124,1	144,1	FEP36	131,8	146,9	161,9
	FEP37	94,9	118,6	142,4	FEP37	125,8	144,1	162,4
	FEP38	107,5	125,9	144,2	FEP38	130,7	146,7	162,6
	FEP39	99,3	122,2	145,1	FEP39	130,5	146,5	162,6
	FEP40	100,4	122,2	144,0	FEP40	132,1	147,9	163,6
	FEP41	101,5	122,4	143,3	FEP41	130,0	145,8	161,6
	FEP42	106,0	126,2	146,3	FEP42	133,9	149,0	164,2
	FEP43	105,1	126,0	146,9	FEP43	133,6	149,4	165,2
	FEP44	102,7	123,6	144,5	FEP44	131,2	147,0	162,8
	FEP45	102,0	122,7	143,3	FEP45	135,2	148,0	160,8
	FEP46	98,7	121,0	143,4	FEP46	128,4	145,5	162,6
	FEP47	103,9	124,8	145,7	FEP47	132,4	148,2	164,0
	FEP48	100,7	122,1	143,4	FEP48	131,0	146,6	162,1
	FEP49	102,7	123,6	144,5	FEP49	130,6	146,3	162,0
	FEP52	102,7	123,6	144,5	FEP52	134,1	148,4	162,7
	FEP53	102,7	123,6	144,5	FEP53	127,5	144,7	162,0
FEP54	106,1	126,1	146,1	FEP54	135,2	148,9	162,7	
Valor Agregado								

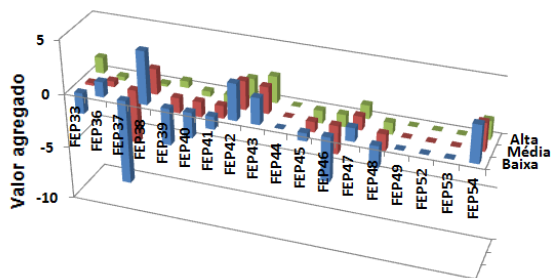
(conclusão)

	Profic. Inicial	Baixa	Média	Alta	Profic. Inicial		
					Baixa	Média	Alta
Ano3	Base	155,8	166,9	177,9			
	FEP33	152,9	165,7	178,5	-3,0	-1,2	0,5
	FEP36	156,1	166,3	176,4	0,3	-0,6	-1,5
	FEP37	151,7	165,7	179,7	-4,1	-1,2	1,8
	FEP38	157,4	165,9	174,3	1,6	-1,0	-3,7
	FEP39	157,3	169,2	181,1	1,5	2,3	3,2
	FEP40	153,5	165,2	177,0	-2,3	-1,7	-1,0
	FEP41	154,7	165,7	176,8	-1,2	-1,2	-1,2
	FEP42	157,1	167,4	177,7	1,3	0,5	-0,2
	FEP43	155,9	168,2	180,5	0,1	1,3	2,5
	FEP44	158,3	167,5	176,8	2,4	0,7	-1,1
	FEP45	154,8	165,7	176,5	-1,0	-1,2	-1,4
	FEP46	153,7	166,2	178,8	-2,2	-0,7	0,8
	FEP47	157,0	168,1	179,1	1,2	1,2	1,2
	FEP48	154,4	166,0	177,5	-1,4	-0,9	-0,4
	FEP49	157,1	166,9	176,7	1,3	0,1	-1,2
	FEP52	155,8	166,9	177,9	0,0	0,0	0,0
FEP53	155,8	166,9	177,9	0,0	0,0	0,0	
FEP54	158,1	168,3	178,5	2,3	1,4	0,5	
Ano4	Base	149,8	166,3	182,9			
	FEP33	146,8	164,2	181,6	-3,0	-2,1	-1,2
	FEP36	147,3	166,2	185,2	-2,5	-0,1	2,3
	FEP37	150,5	166,1	181,7	0,7	-0,2	-1,2
	FEP38	146,5	163,5	180,6	-3,3	-2,8	-2,3
	FEP39	149,6	166,9	184,1	-0,2	0,5	1,2
	FEP40	147,5	164,9	182,4	-2,4	-1,4	-0,5
	FEP41	151,3	166,6	182,0	1,4	0,3	-0,9
	FEP42	151,8	167,6	183,4	2,0	1,3	0,5
	FEP43	152,2	168,7	185,3	2,4	2,4	2,4
	FEP44	151,7	166,9	182,2	1,9	0,6	-0,7
	FEP45	149,9	166,4	183,0	0,1	0,1	0,1
	FEP46	146,4	164,0	181,7	-3,4	-2,3	-1,2
	FEP47	149,6	167,3	185,0	-0,3	0,9	2,1
	FEP48	149,6	165,1	180,7	-0,3	-1,2	-2,1
	FEP49	150,1	165,9	181,8	0,3	-0,4	-1,1
	FEP52	151,1	167,6	184,2	1,3	1,3	1,3
FEP53	148,4	165,9	183,3	-1,4	-0,5	0,4	
FEP54	151,7	167,3	183,0	1,8	1,0	0,1	

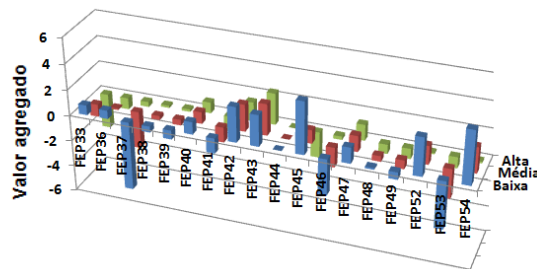
Os Gráficos 33A, 33B, 33C e 33D representam, para cada ano, o valor que cada unidade da variável explicativa de Nível 3 impacta na proficiência final de alunos com diferentes proficiências iniciais.

Gráfico 33 - Valor agregado por unidade da variável explicativa do Grupo 3 do Nível 3 em Leitura

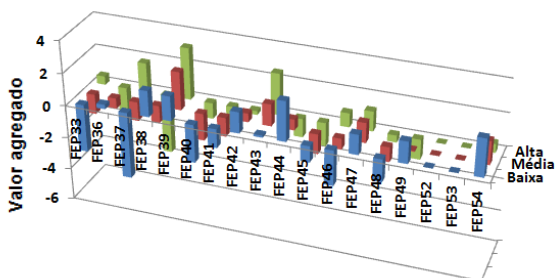
A – Primeiro ano de escolarização



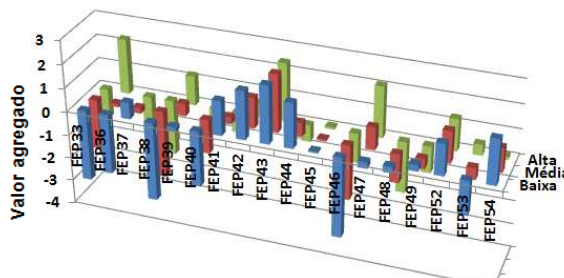
B – Segundo ano de escolarização



C – Terceiro ano de escolarização



D – Quarto ano de escolarização



O Quadro 63 traz uma síntese do impacto de cada uma das variáveis explicativas de Nível 3 na proficiência final em cada um dos anos de escolarização.

Quadro 63 - Tipo de impacto das variáveis do Grupo 3 do Nível 3 em Leitura a cada ano

	Ano1	Ano2	Ano3	Ano4
FEP33	Tipo 6	Tipo 2	Tipo 6	Tipo 8
FEP36	Tipo 2	Tipo 2	Tipo 2	Tipo 6
FEP37	Tipo 8	Tipo 8	Tipo 6	Tipo 2
FEP38	Tipo 2	Tipo 8	Tipo 2	Tipo 8
FEP39	Tipo 6	Tipo 8	Tipo 3	Tipo 6
FEP40	Tipo 8	Tipo 4	Tipo 8	Tipo 8
FEP41	Tipo 5	Tipo 5	Tipo 5	Tipo 2
FEP42	Tipo 4	Tipo 4	Tipo 2	Tipo 4
FEP43	Tipo 1	Tipo 1	Tipo 3	Tipo 1
FEP44			Tipo 2	Tipo 2
FEP45	Tipo 7	Tipo 2	Tipo 7	Tipo 3
FEP46	Tipo 8	Tipo 8	Tipo 6	Tipo 8
FEP47	Tipo 1	Tipo 1	Tipo 1	Tipo 6
FEP48	Tipo 8	Tipo 7	Tipo 8	Tipo 7
FEP49		Tipo 7	Tipo 2	Tipo 2
FEP52		Tipo 2		Tipo 1
FEP53		Tipo 8		Tipo 6
FEP54	Tipo 4	Tipo 2	Tipo 4	Tipo 4

O Quadro 64 apresenta a síntese na interpretação do comportamento dessas variáveis.

Quadro 64 - Interpretação final das variáveis do Grupo 2 do Nível 3 em Leitura

Comportamento	Variáveis
Coerente com o esperado	FEP43, FEP48
Instável, dificultando a interpretação	FEP33, FEP36, FEP37, FEP38, FEP39, FEP40, FEP41, FEP42, FEP43, FEP44, FEP45, FEP46, FEP47, FEP49, FEP52, FEP53, FEP54.

5.4.1.4) Análise das variáveis do Modelo de Nível 3 em Leitura pertencentes ao Grupo 4 (características da escola extraídas do questionário de contexto do professor)

Este grupo de variáveis de Nível 3 é composto por 15 das 50 variáveis existentes. O Quadro 65 apresenta uma síntese do ano no qual essas variáveis foram significantes e o tipo de impacto que essa variável teve sobre a proficiência final em Leitura.

Para efeito de análise, o Modelo de Nível 3 foi elaborado e seus resultados serão apresentados a seguir.

Quadro 65 - Relação entre as variáveis explicativas do Grupo 4 do Nível 3 em Leitura e os anos de escolarização que elas influenciaram

Teve coeficientes significativos em todos os anos de escolarização, impactando em todos esses anos	FEP31, FEP32, FEP34, FEP58*
Teve coeficientes significativos no primeiro, no segundo e quarto ano, mas tem impacto em todos os anos, por ser estatisticamente significativa no primeiro ano	FEP29, FEP35
Teve coeficientes significativos no primeiro, no terceiro e quarto ano, mas tem impacto em todos os anos, por ser estatisticamente significativa no primeiro ano	FEP51
Teve coeficiente significativo apenas no segundo ano, portanto, impactando apenas nesse ano	FEP30
Teve coeficiente significativo apenas no quarto ano, portanto, impactando apenas nesse ano	FEP28
Teve coeficiente significativo apenas no segundo e quarto anos, portanto, impactando apenas nesses anos	FEP50, FEP55*
Teve coeficiente significativo apenas no segundo e terceiro anos, portanto, impactando apenas nesses anos	FEP57*
Não teve coeficiente significativo em nenhum dos anos e, portanto, não tem impactos	EP27, FEP56*, FEP59

O Quadro 66 traz uma síntese gráfica dessas informações.

Quadro 66 - Anos em que as variáveis do Grupo 4 do Nível 4 em Leitura impactam

	Ano1	Ano2	Ano3	Ano4
FEP27				
FEP28				((#))
FEP29	((#))	((#))	#	((#))
FEP30		((#))		
FEP31	((#))	((#))	((#))	((#))
FEP32	((#))	((#))	((#))	((#))
FEP34	((#))	((#))	((#))	((#))
FEP35	((#))	((#))	#	((#))
FEP50		((#))		((#))
FEP51	((#))	#	((#))	((#))
FEP55*		((#))		((#))
FEP56*				
FEP57*		((#))	((#))	
FEP58*	((#))	((#))	((#))	((#))
FEP59				

Legenda:

((#)) a variável impacta nesse ano e tem um coeficiente significativo

a variável impacta nesse ano mas não tem coeficiente

Essas sínteses são produzidas a partir da análise do valor agregado que a variável provoca na proficiência final quando o valor de uma unidade é atribuído à variável explicativa, cujos valores são apresentados no Quadro 67.

Quadro 67 - Valores de proficiência finais e valor agregado por variável explicativa do Grupo 4 do Nível 3 em Leitura

(continua)

Ano1	Profic. Inic.	Baixa	Média	Alta	Profic Inicial		
	Base	155,8	166,9	177,9	Baixa	Média	Alta
Proficiência Final	FEP27	102,7	123,6	144,5	0,0	0,0	0,0
	FEP28	102,7	123,6	144,5	0,0	0,0	0,0
	FEP29	102,3	124,0	145,8	-0,4	0,5	1,3
	FEP30	102,7	123,6	144,5	0,0	0,0	0,0
	FEP31	104,7	124,2	143,6	2,0	0,6	-0,9
	FEP32	98,5	119,4	140,2	-4,2	-4,2	-4,2
	FEP34	105,2	124,6	143,9	2,5	1,0	-0,6
	FEP35	100,3	122,3	144,2	-2,3	-1,3	-0,3
	FEP50	102,7	123,6	144,5	0,0	0,0	0,0
	FEP51	101,1	121,9	142,8	-1,6	-1,6	-1,6
	FEP55*	102,7	123,6	144,5	0,0	0,0	0,0
	FEP56*	102,7	123,6	144,5	0,0	0,0	0,0
	FEP57*	102,7	123,6	144,5	0,0	0,0	0,0
	FEP58*	105,7	124,7	143,7	3,0	1,1	-0,8
	FEP59	102,7	123,6	144,5	0,0	0,0	0,0

(continua)

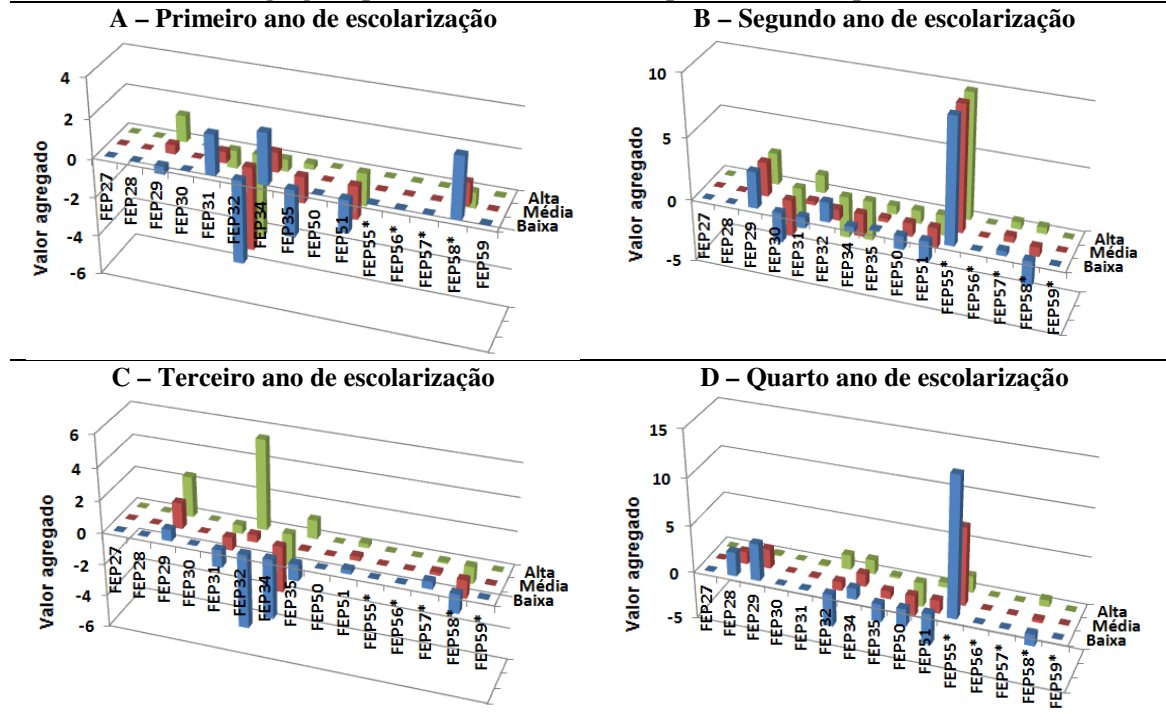
Ano2	Profic. Inic.	Baixa	Média	Alta	Valor Agregado	Profic Inicial		
						Baixa	Média	Alta
Proficiência Final	Base	157,0	168,1	179,1		Ano4	Ano4	Ano4
	Ano4	Ano4	Ano4	Ano4		Ano4	Ano4	Ano4
	FEP27	131,2	147,0	162,8		0,0	0,0	0,0
	FEP28	131,2	147,0	162,8		0,0	0,0	0,0
	FEP29	134,1	149,7	165,3		2,9	2,7	2,5
	FEP30	128,8	144,1	159,5		-2,4	-2,9	-3,4
	FEP31	130,3	147,3	164,2		-0,8	0,3	1,4
	FEP32	132,7	146,1	159,5		1,6	-0,9	-3,3
	FEP34	130,7	145,2	159,6		-0,5	-1,8	-3,2
	FEP35	131,2	146,7	162,2		0,1	-0,3	-0,7
	FEP50	130,1	145,9	161,7		-1,1	-1,1	-1,1
	FEP51	129,5	145,3	161,2		-1,6	-1,6	-1,6
	FEP55*	141,0	156,9	172,7		9,9	9,9	9,9
	FEP56*	131,2	147,0	162,8		0,0	0,0	0,0
	FEP57*	131,6	147,5	163,4		0,4	0,5	0,6
	FEP58*	129,2	146,3	163,3		-2,0	-0,7	0,5
FEP59*	131,2	147,0	162,8		0,0	0,0	0,0	
Ano3	Profic. Inic.	Baixa	Média	Alta	Valor Agregado	Profic Inicial		
						Baixa	Média	Alta
Proficiência Final	Base	153,7	168,4	183,0		Ano4	Ano4	Ano4
	FEP27	155,8	166,9	177,9		0,0	0,0	0,0
	FEP28	155,8	166,9	177,9		0,0	0,0	0,0
	FEP29	156,5	168,5	180,4		0,7	1,6	2,5
	FEP30	155,8	166,9	177,9		0,0	0,0	0,0
	FEP31	154,7	166,1	177,4		-1,1	-0,8	-0,5
	FEP32	151,2	167,3	183,4		-4,6	0,4	5,5
	FEP34	152,1	164,0	176,0		-3,7	-2,8	-2,0
	FEP35	154,8	166,9	179,1		-1,0	0,1	1,1
	FEP50	155,8	166,9	177,9		0,0	0,0	0,0
	FEP51	156,1	167,1	178,2		0,3	0,3	0,3
	FEP55*	155,8	166,9	177,9		0,0	0,0	0,0
	FEP56*	155,8	166,9	177,9		0,0	0,0	0,0
	FEP57*	156,3	167,1	177,8		0,5	0,2	-0,1
	FEP58*	154,7	165,8	176,9		-1,2	-1,1	-1,0
	FEP59*	155,8	166,9	177,9		0,0	0,0	0,0

(conclusão)

Ano4	Profic. Inic.				Profic Inicial		
	Base	Baixa	Média	Alta	Baixa	Média	Alta
		146,4	164,0	181,7	0,0	0,0	0,0
	FEP27	149,8	166,3	182,9	2,6	1,3	0,0
	FEP28	152,4	167,6	182,9	3,9	2,0	0,1
	FEP29	153,7	168,4	183,0	0,0	0,0	0,0
	FEP30	149,8	166,3	182,9	-0,1	-0,1	-0,2
	FEP31	149,8	166,2	182,7	-3,5	-1,0	1,5
	FEP32	146,4	165,4	184,4	1,2	1,3	1,5
	FEP34	151,0	167,7	184,4	-1,9	-0,8	0,2
	FEP35	147,9	165,5	183,1	-1,9	-2,2	-2,6
	FEP50	147,9	164,1	180,3	-3,3	-1,4	0,5
	FEP51	146,5	165,0	183,4	14,6	8,2	1,7
	FEP55*	164,4	174,5	184,6	0,0	0,0	0,0
	FEP56*	149,8	166,3	182,9	0,0	0,0	0,0
	FEP57*	149,8	166,3	182,9	-1,2	-0,3	0,7
	FEP58*	148,6	166,1	183,6	0,0	0,0	0,0
	FEP59*	149,8	166,3	182,9	0,0	0,0	0,0

Os Gráficos 34A, 34B, 34C e 34D representam, para cada ano, o valor que cada unidade da variável explicativa de Nível 3 impacta na proficiência final de alunos com diferentes proficiências iniciais, isto é, representa o valor agregado devido a uma unidade da variável explicativa.

Gráfico 34 - Valor agregado por unidade da variável explicativa do Grupo 4 do Nível 3 em Leitura



O Quadro 68 traz uma síntese do impacto de cada uma das variáveis explicativas de Nível 3 na proficiência final em cada um dos anos de escolarização.

Quadro 68 - Tipo de impacto das variáveis do Grupo 4 do Nível 3 em Leitura a cada ano

	Ano1	Ano2	Ano3	Ano4
FEP27				
FEP28				Tipo 4
FEP29	Tipo 6	Tipo 4	Tipo 3	Tipo 4
FEP30		Tipo 7		
FEP31	Tipo 2	Tipo 6	Tipo 8	Tipo 7
FEP32	Tipo 5	Tipo 2	Tipo 6	Tipo 6
FEP34	Tipo 2	Tipo 7	Tipo 8	Tipo 3
FEP35	Tipo 8	Tipo 2	Tipo 6	Tipo 6
FEP50		Tipo 5		Tipo 7
FEP51	Tipo 5	Tipo 5	Tipo 1	Tipo 6
FEP55*		Tipo 1		Tipo 4
FEP56*				
FEP57*		Tipo 3	Tipo 2	
FEP58*	Tipo 2	Tipo 6	Tipo 8	Tipo 6
FEP59				

O Quadro 69 apresenta a síntese na interpretação do comportamento dessas variáveis.

Quadro 69 - Interpretação final das variáveis do Grupo 4 do Nível 3 em Leitura

Comportamento	Variáveis
Coerente com o esperado	FEP28,
Incoerente com o esperado	FEP30,
Instável, dificultando a interpretação	FEP29, FEP31, FEP32, FEP34, FEP35, FEP36, FEP51, FEP55*, FEP57*, FEP58*
Não influenciou	FEP27, FEP56*, FEP59*

5.4.1.5) Síntese da análise das variáveis do Modelo de Nível 3 em Leitura

Para a interpretação do efeito de cada uma das variáveis do Nível 3, foram consideradas as estimativas dos coeficientes associados a cada uma das variáveis considerando as proficiência iniciais e finais de cada um dos diferentes anos de escolarização, porém, para os Anos 2, 3 e 4, tais coeficientes são incrementos, sejam eles positivos ou

negativos, aos coeficientes estimados para o Ano 1. Por isso, os coeficientes do Ano 1 são referência para a análise dos demais anos. Como consequência desta lógica, tem-se que:

- Para o Ano 1, as variáveis que apresentam coeficientes significativos para o intercepto, para a inclinação ou ainda para ambos, deve-se entender que essas variáveis têm impactos significativos na determinação da proficiência final, em relação à proficiência inicial, neste primeiro ano. Este impacto pode ser um ganho ou uma perda.

- Para os demais anos, as variáveis que apresentam coeficientes significativos, para o intercepto, para a inclinação ou ainda para ambos, deve-se entender que essa variável tem um valor para tais coeficientes significativamente diferente àqueles estimados para o Ano 1. Ou seja, esse coeficiente representa um valor que deve ser somado algebricamente ao valor estimado para esse coeficiente no Ano 1. Disso decorre que:

a) Quando uma variável é significativa no Ano 1 e não nos demais anos, significa que ela também tem igual impacto nesses outros anos. Vale a pena observar que, mesmo com coeficientes iguais, os impactos são distintos já que os valores das proficiências iniciais nesses diferentes anos de escolarização são diferentes. Esse comportamento foi observado nas variáveis FEE05, FEE08, FED14, FED25*, FEP31, FEP32, FEP33, FEP34, FEP39, FEP40 e FEP58*.

b) Quando uma variável é significativa no Ano 1 e também em qualquer um dos anos seguintes, significa que essa variável tem impacto em todos os anos, porém de forma diferenciada no ano em que a variável foi significativa no Ano 1 e no ano analisado, e ainda que esse impacto é diferentes nesses dois anos. No Quadro 70 são apresentadas as variáveis que têm efeitos sobre a proficiência final no primeiro ano e um efeito distinto nos demais anos.

Quadro 70 – Variáveis do Nível 3 que têm impacto na proficiência em Leitura no Ano 1 e com coeficiente diferente em outro ano

Anos	Variáveis
1 e 2	FEP54,
1 e 3	FEP42, FEP43,
1 e 4	FEE04 ,FED13, FEP36, FEP37, FEP41, FEP46, FEP47
1, 2 e 3	
1, 2 e 4	FED23* , FEP38, FEP45, FEP48 , FEP29, FEP35
1, 3 e 4	FED12 , FED22*, FEP51
1, 2, 3 e 4	FEA01, FEE02, FED14, FED25*

- c) Quando uma variável não é significativa no Ano1, mas o é em qualquer dos anos seguintes, significa que ela tem impacto somente neste ano em que a variável foi significativa. O Quadro 71 mostra as variáveis e os respectivos anos que apresentaram esse efeito.

Quadro 71 - Variáveis que têm impacto no Ano 1 e têm o mesmo coeficiente em outro Ano (significa que a variável não foi significativa nesse ano)

Anos	Variáveis
3 e 4	FEE11 ; FEP44
2 e 4	FEE06 ; FEE07 ; FED21; FEP52, FEP53 ; FEP50, FEP55*
2 e 3	FEE03 ; FEE10 ; FED24* ; FED26* ; FEP27*
4	FEE09 ; FED18 ; FEP28
3	FED15
2	FED20; FEP30;
2, 3 e 4	FED19

- d) Quando a variável não tem qualquer efeito significativo em qualquer um dos anos seus coeficientes são zerados, sendo eliminadas do modelo. São elas as variáveis FED16, FEP27, FEP49, FEP56* e FEP59.

5.4.2) Interpretação dos resultados do Modelo de Nível 3 em Matemática

A seguir serão apresentados os resultados dos Modelos de Nível 3 em Matemática, em uma sequência análoga ao que foi feito com o modelo em Leitura, porém em dois distintos modelos sendo o primeiro considerando as variáveis de Nível 2 e o segundo sem essas variáveis.

5.4.2.1) Interpretação dos resultados do Modelo de Nível 3 em Matemática

Da mesma maneira como foi apresentadas as análises no Modelo de Nível 3 em Leitura, os próximos três tópicos apresentam os efeitos que as variáveis explicativas do Nível 3 têm na proficiência final a partir de três diferentes proficiências iniciais distintas em cada um dos anos de escolarização. A análise do comportamento dessas variáveis é feita através da comparação entre a proficiência final com todas as variáveis nulas e quando a cada uma delas é atribuído o valor de uma unidade.

5.4.2.1.1) Análise das variáveis pertencentes ao Grupo 1 (nível socioeconômico médio e outras características oriundas do questionário de contexto da escola) do Modelo de Nível 3 em Matemática

Este grupo de variáveis de Nível 3 é composto por 11 das 59 variáveis existentes. O Quadro 72 apresenta uma síntese do ano no qual essas variáveis foram significantes e o tipo de impacto que essa variável teve sobre a proficiência final em Matemática.

Quadro 72 - Relação entre as variáveis explicativas do grupo 1 do Nível 3 em Matemática e os anos de escolarização que elas influenciaram

Teve coeficientes significativos em todos os anos de escolarização, impactando em todos esses anos	FEA01, FEE02
Teve coeficientes significativos apenas no primeiro ano, mas tem impacto em todos os anos, por ser estatisticamente significativa no primeiro ano	FEE07
Teve coeficientes significativos no primeiro e segundo ano, mas tem impacto em todos os anos, por ser estatisticamente significativa no primeiro ano	FEE11
Teve coeficientes significativos no primeiro, segundo e terceiro ano, mas tem impacto em todos os anos, por ser estatisticamente significativa no primeiro ano	FEE03
Teve coeficiente significativo apenas no segundo e terceiro anos, portanto, impactando apenas nesses anos	FEE04
Teve coeficiente significativo apenas no terceiro ano, portanto, impactando apenas nesse ano	FEE08
Teve coeficiente significativo apenas no quarto ano, portanto, impactando apenas nesse ano	FEE09, FEE10
Não teve coeficiente significativo em nenhum dos anos e, portanto, não tem impactos	FEE05, FEE06

O Quadro 73 traz uma síntese gráfica dessas informações.

Quadro 73 - Anos em que as variáveis do Grupo 1 do Nível 3 em Matemática impactam

	Ano1	Ano2	Ano3	Ano4
FEA01	((#))	((#))	((#))	((#))
FEE02	((#))	((#))	((#))	((#))
FEE03	((#))	((#))	((#))	#
FEE04		((#))	((#))	
FEE05				
FEE06				
FEE07	((#))	#	#	#
FEE08			((#))	
FEE09				((#))
FEE10				((#))
FEE11	((#))	((#))	#	#

Legenda:

((#)) a variável impacta nesse ano e tem um coeficiente significativo

a variável impacta nesse ano mas não tem coeficiente

Essas sínteses são produzidas a partir da análise do valor agregado que a variável provoca na proficiência final quando o valor de uma unidade é atribuído à variável explicativa, cujos valores são apresentados no Quadro 74.

Quadro 74 - Valores de proficiência finais e valor agregado por variável explicativa do Grupo 1 do Nível 3 em Matemática
(continua)

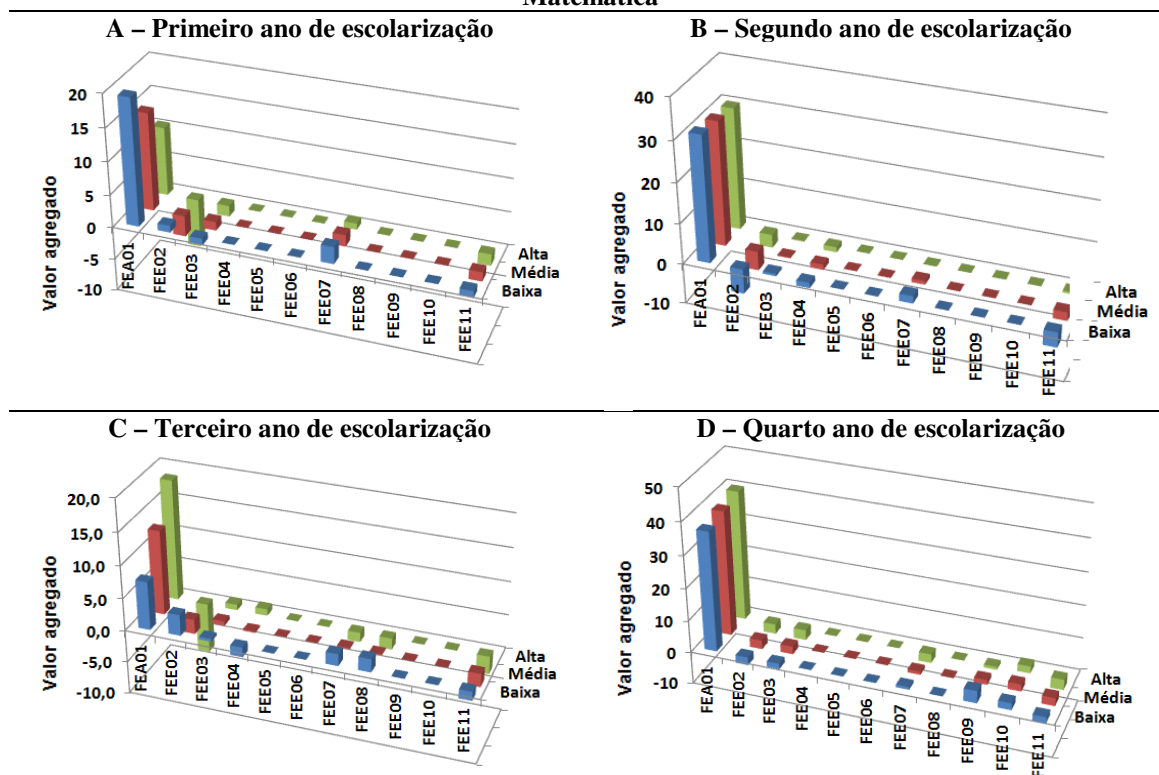
Ano1	Profic. Inic.	Baixa	Média	Alta	Valor Agregado	Profic. Inicial		
						Baixa	Média	Alta
Proficiência Final	Base	109,2	121,3	133,5		19,4	14,9	10,5
	FEA01	128,6	136,3	144,0		1,1	-3,2	-7,5
	FEE02	110,2	118,1	126,0		-1,0	-1,3	-1,7
	FEE03	108,1	120,0	131,8				
	FEE04	109,2	121,3	133,5				
	FEE05	109,2	121,3	133,5				
	FEE06	109,2	121,3	133,5				
	FEE07	111,7	123,1	134,4		2,6	1,8	0,9
	FEE08	109,2	121,3	133,5				
	FEE09	109,2	121,3	133,5				
	FEE10	109,2	121,3	133,5				
	FEE11	108,1	119,9	131,8		-1,0	-1,4	-1,7
Ano2	Profic. Inic.	Baixa	Média	Alta	Valor Agregado	Profic Inicial		
Proficiência Final	Base	120,0	177,6	234,9		31,2	30,6	30,1
	FEA01	151,1	208,2	265,0		-6,1	-4,7	-3,3
	FEE02	113,8	172,9	231,6		0,6	0,2	-0,1
	FEE03	120,6	177,9	234,8		-1,3	-1,3	-1,3
	FEE04	118,7	176,3	233,6				
	FEE05	120,0	177,6	234,9				
	FEE06	120,0	177,6	234,9				
	FEE07	121,8	178,5	234,9		1,8	0,9	-0,1
	FEE08	120,0	177,6	234,9				
	FEE09	120,0	177,6	234,9				
	FEE10	120,0	177,6	234,9				
	FEE11	116,2	175,5	234,4		-3,7	-2,1	-0,5

(conclusão)

	Profic. Inic.	Baixa	Média	Alta	Valor Agregado	Profic Inicial		
						Baixa	Média	Alta
Ano3	Base	166,8	198,8	230,8				
	FEA01	174,3	211,8	249,3		7,4	12,9	18,5
	FEE02	170,1	196,6	223,1		3,2	-2,3	-7,7
	FEE03	167,3	199,5	231,7		0,4	0,6	0,9
	FEE04	165,4	198,6	231,8		-1,5	-0,2	1,0
	FEE05	166,8	198,8	230,8				
	FEE06	166,8	198,8	230,8				
	FEE07	168,7	199,0	229,4		1,8	0,2	-1,5
	FEE08	168,9	199,0	229,1		2,0	0,1	-1,7
	FEE09	166,8	198,8	230,8				
	FEE10	166,8	198,8	230,8				
	FEE11	165,5	196,8	228,1		-1,4	-2,0	-2,7
Ano4	Base	225,2	246,3	267,4				
	FEA01	262,2	284,9	307,5		37,0	38,5	40,1
	FEE02	222,8	243,6	264,4		-2,5	-2,7	-3,0
	FEE03	223,5	243,9	264,3		-1,7	-2,4	-3,1
	FEE04	225,2	246,3	267,4				
	FEE05	225,2	246,3	267,4				
	FEE06	225,2	246,3	267,4				
	FEE07	226,1	245,4	264,7		0,8	-1,0	-2,8
	FEE08	225,2	246,3	267,4				
	FEE09	229,1	247,8	266,5		3,8	1,5	-0,9
	FEE10	227,4	248,5	269,6		2,2	2,2	2,2
	FEE11	223,5	243,8	264,2		-1,8	-2,5	-3,3

Os Gráficos 36A, 36B, 36C e 36D representam, para cada ano, o valor que cada unidade da variável explicativa de Nível 3 impacta na proficiência final de alunos com diferentes proficiências iniciais, isto é, representa o valor agregado devido a uma unidade da variável explicativa.

Gráfico 35 - Valor agregado por unidade da variável explicativa do Grupo 1 do Nível em Matemática



O Quadro 75 traz uma síntese do impacto de cada uma das variáveis explicativas de Nível 3 na proficiência final em cada um dos anos de escolarização.

Quadro 75 - Tipo de impacto das variáveis do Grupo 1 do Nível 3 em Matemática a cada ano

	Ano1	Ano2	Ano3	Ano4
FEA01	Tipo 4	Tipo 4	Tipo 3	Tipo 3
FEE02	Tipo 2	Tipo 8	Tipo 2	Tipo 7
FEE03	Tipo 7	Tipo 2	Tipo 3	Tipo 7
FEE04		Tipo 8	Tipo 6	
FEE05				
FEE06				
FEE07	Tipo 4	Tipo 2	Tipo 2	Tipo 2
FEE08			Tipo 2	
FEE09				Tipo 2
FEE10				Tipo 1
FEE11	Tipo 7	Tipo 8	Tipo 7	Tipo 7

O Quadro 76 apresenta a síntese na interpretação do comportamento dessas variáveis.

Quadro 76 - Interpretação final das variáveis do Grupo 1 do Nível 3 em Leitura

Comportamento	Variáveis
Coerente com o esperado	FEA01, FEE10
Incoerente com o esperado	FEE11
Instável, dificultando a interpretação	FEE02, FEE03, FEE04
Não influenciou	FEE05, FEE06, FEE07, FEE08, FEE09,

5.4.2.1.2) Análise das variáveis pertencentes ao Grupo 2 (características da escola extraídas do questionário de contexto do professor) do Modelo de Nível 3 em Matemática

Este grupo de variáveis de Nível 3 é composto por 15 das 59 variáveis existentes. O Quadro 77 apresenta uma síntese do ano no qual essas variáveis foram significantes e o tipo de impacto que essa variável teve sobre a proficiência final em Matemática.

Quadro 77 - Relação entre as variáveis explicativas do Grupo 2 do Nível 3 em Matemática e os anos de escolarização que elas influenciaram

Teve coeficientes significativos em todos os anos de escolarização, impactando em todos esses anos	FED21
Teve coeficientes significativos apenas no primeiro ano, mas tem impacto em todos os anos, por ser estatisticamente significativa no primeiro ano	FED12
Teve coeficientes significativos no primeiro e segundo ano, mas tem impacto em todos os anos, por ser estatisticamente significativa no primeiro ano	FED25*
Teve coeficientes significativos no primeiro e no quarto ano, mas tem impacto em todos os anos, por ser estatisticamente significativa no primeiro ano	FED13
Teve coeficientes significativos no primeiro, no terceiro e quarto ano, mas tem impacto em todos os anos, por ser estatisticamente significativa no primeiro ano	FED18
Teve coeficiente significativo apenas no quarto ano, portanto, impactando apenas nesse ano	FED16, FED17, FED19, FED20
Teve coeficiente significativo apenas no segundo ano, portanto, impactando apenas nesse ano	FED22*, FED23*
Não teve coeficiente significativo em nenhum dos anos e, portanto, não tem impactos	FED14, FED15, FED24*, FED26*

O Quadro 78 traz uma síntese gráfica dessas informações.

Quadro 78 - Anos em que as variáveis do Grupo 2 do Nível 3 em Matemática impactam

	Ano1	Ano2	Ano3	Ano4
FED12	((#))	#	#	#
FED13	((#))	#	#	((#))
FED14				
FED15				
FED16				((#))
FED17				((#))
FED18	((#))	#	((#))	((#))
FED19				((#))
FED20				((#))
FED21	((#))	((#))	((#))	((#))
FED22*		((#))		
FED23*		((#))		
FED24*				
FED25*	((#))	((#))	#	#
FED26*				

Legenda:

((#)) a variável impacta nesse ano e tem um coeficiente significativo

a variável impacta nesse ano mas não tem coeficiente

Essas sínteses são produzidas a partir da análise do valor agregado que a variável provoca na proficiência final quando o valor de uma unidade é atribuído à variável explicativa, cujos valores são apresentados no Quadro 79.

Quadro 79 - Valores de proficiência finais e valor agregado por variável explicativa do Grupo 2 do Nível 3 em Matemática
(continua)

Ano1	Profic. Inic.	Baixa	Média	Alta	Profic Inicial		
					Baixa	Média	Alta
	Base	109,2	121,3	133,5			
	FED12	104,8	117,0	129,2	-4,3	-4,3	-4,3
	FED13	107,5	119,1	130,8	-1,6	-2,2	-2,7
	FED14	109,2	121,3	133,5			
	FED15	109,2	121,3	133,5			
	FED16	109,2	121,3	133,5			
	FED17	109,2	121,3	133,5			
	FED18	108,2	121,4	134,6	-0,9	0,1	1,1
	FED19	109,2	121,3	133,5			
	FED20*	109,2	121,3	133,5			
	FED21*	111,4	124,3	137,2	2,2	3,0	3,7
	FED22*	109,2	121,3	133,5			
	FED23*	109,2	121,3	133,5			
	FED24*	109,2	121,3	133,5			
	FED25*	111,6	122,8	134,1	2,4	1,5	0,6
	FED26*	109,2	121,3	133,5			

(conclusão)

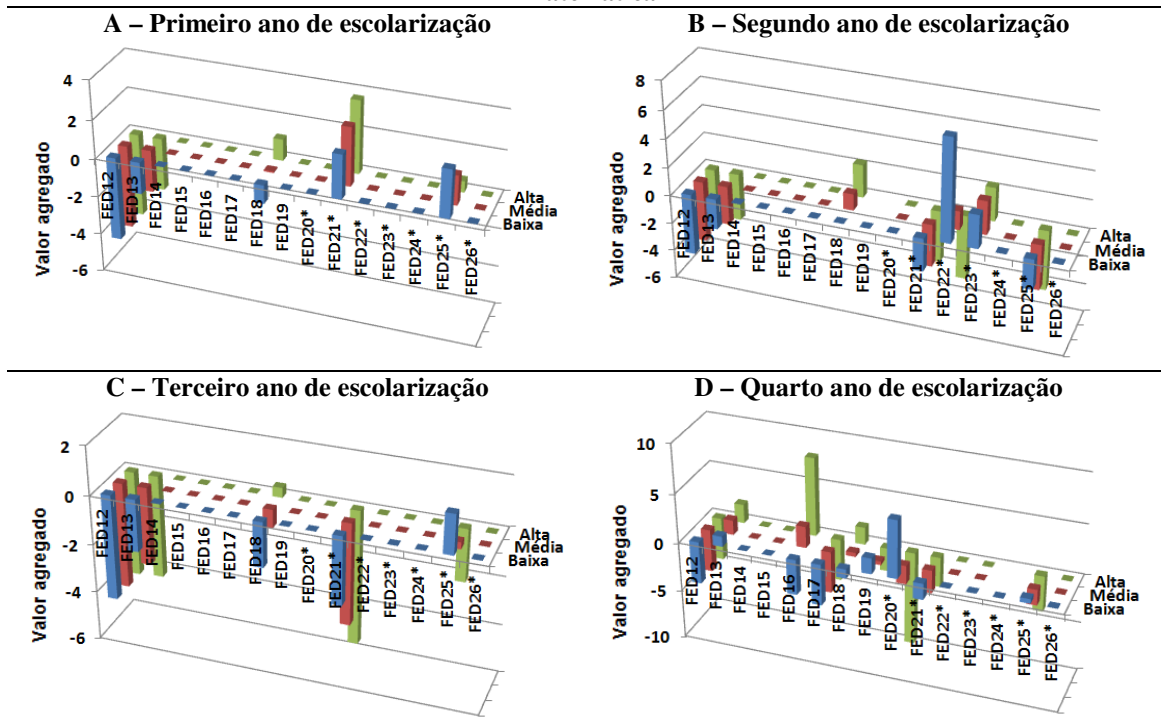
Ano2	Profic. Inic.	Baixa	Média	Alta	Profic Inicial			
					Baixa	Média	Alta	
Proficiência Final	Base	120,0	177,6	234,9				
	FED12	115,6	173,3	230,6	-4,3	-4,3	-4,3	
	FED13	117,8	174,9	231,6	-2,1	-2,7	-3,3	
	FED14	120,0	177,6	234,9				
	FED15	120,0	177,6	234,9				
	FED16	120,0	177,6	234,9				
	FED17	120,0	177,6	234,9				
	FED18	120,0	178,8	237,2	0,0	1,1	2,3	
	FED19	120,0	177,6	234,9				
	FED20*	120,0	177,6	234,9				
	FED21*	117,6	174,6	231,3	-2,3	-3,0	-3,6	
	FED22*	127,1	178,9	230,4	7,2	1,3	-4,5	
	FED23*	122,3	180,0	237,3	2,3	2,3	2,3	
	FED24*	120,0	177,6	234,9				
	FED25*	117,8	174,4	230,7	-2,1	-3,2	-4,2	
FED26*	120,0	177,6	234,9					
Ano3	Profic. Inic.	Baixa	Média	Alta	Profic Inicial			
					Baixa	Média	Alta	
	Proficiência Final	Base	166,8	198,8	230,8			
		FED12	162,5	194,5	226,5	-4,3	-4,3	-4,3
		FED13	164,7	195,6	226,6	-2,1	-3,2	-4,3
		FED14	166,8	198,8	230,8			
		FED15	166,8	198,8	230,8			
		FED16	166,8	198,8	230,8			
		FED17	166,8	198,8	230,8			
		FED18	165,0	198,1	231,2	-1,8	-0,7	0,4
		FED19	166,8	198,8	230,8			
		FED20*	166,8	198,8	230,8			
		FED21*	164,1	194,7	225,3	-2,8	-4,2	-5,6
		FED22*	166,8	198,8	230,8			
		FED23*	166,8	198,8	230,8			
FED24*		166,8	198,8	230,8				
FED25*	168,4	198,6	228,7	1,6	-0,3	-2,1		
FED26*	166,8	198,8	230,8					

(conclusão)

Ano4	Profic. Inic.	Base	Baixa	Média	Alta	Profic Inicial		
		Base	Baixa	Média	Alta	Baixa	Média	Alta
		FED12	220,9	242,0	263,1	-4,3	-4,3	-4,3
		FED13	226,3	247,8	269,3	1,0	1,5	1,9
		FED14	225,2	246,3	267,4			
		FED15	225,2	246,3	267,4			
		FED16	221,6	248,5	275,3	-3,6	2,1	7,9
		FED17	221,0	242,1	263,2	-4,2	-4,2	-4,2
		FED18	224,3	246,7	269,1	-0,9	0,4	1,7
		FED19	226,8	246,0	265,1	1,6	-0,4	-2,3
		FED20*	231,1	244,5	257,8	5,8	-1,9	-9,6
		FED21*	223,6	244,0	264,4	-1,6	-2,3	-3,0
		FED22*	225,2	246,3	267,4			
		FED23*	225,2	246,3	267,4			
		FED24*	225,2	246,3	267,4			
		FED25*	225,7	244,8	263,8	0,5	-1,6	-3,6
		FED26*	225,2	246,3	267,4			

Os Gráficos 37A, 37B, 37C e 37D representam, para cada ano, o valor que cada unidade da variável explicativa de Nível 3 impacta na proficiência final de alunos com diferentes proficiências iniciais.

Gráfico 36 - Valor agregado por unidade da variável explicativa do Grupo 1 do Nível 3 em Matemática



O Quadro 80 traz uma síntese do impacto de cada uma das variáveis explicativas de Nível 3 na proficiência final em cada um dos anos de escolarização.

Quadro 80 - Tipo de impacto das variáveis do Grupo 2 do Nível 3 em Matemática a cada ano

	Ano1	Ano2	Ano3	Ano4
FED12	Tipo 5	Tipo 5	Tipo 5	Tipo 5
FED13	Tipo 7	Tipo 7	Tipo 7	Tipo 3
FED14				
FED15				
FED16				Tipo 6
FED17				Tipo 5
FED18	Tipo 6	Tipo 6	Tipo 6	Tipo 6
FED19				Tipo 2
FED20				Tipo 2
FED21	Tipo 3	Tipo 7	Tipo 7	Tipo 7
FED22*		Tipo 2		
FED23*		Tipo 1		
FED24*				
FED25*	Tipo 4	Tipo 7	Tipo 2	Tipo 2
FED26*				

O Quadro 81 apresenta a síntese na interpretação do comportamento dessas variáveis.

Quadro 81 - Interpretação final das variáveis do Grupo 2 do Nível 3 em Leitura

Comportamento	Variáveis
Coerente com o esperado	FED12, FED18, FED23*,
Incoerente com o esperado	FED17,
Instável, dificultando a interpretação	FED13, FED16, FED19, FED20*, FED21*, FED22*, FED25*
Não influenciou	FED14, FED15, FED24*, FED26*

5.4.2.1.3) Análise das variáveis pertencentes ao Grupo 3 (características da escola extraídas do questionário de contexto do diretor) do Modelo de Nível 3 em Matemática

Este grupo de variáveis de Nível 3 é composto por 18 das 59 variáveis existentes. O Quadro 82 apresenta uma síntese do ano no qual essas variáveis foram significantes e o tipo de impacto que essa variável teve sobre a proficiência final em Matemática.

Quadro 82 - Relação entre as variáveis explicativas do Grupo 3 do Nível 3 em Matemática e os anos de escolarização que elas influenciaram

Teve coeficientes significativos em todos os anos de escolarização, impactando em todos esses anos	FEP49
Teve coeficientes significativos apenas no primeiro ano, mas tem impacto em todos os anos, por ser estatisticamente significativa no primeiro ano	FEP37, FEP39, FEP54
Teve coeficientes significativos no primeiro e segundo ano, mas tem impacto em todos os anos, por ser estatisticamente significativa no primeiro ano	FEP47
Teve coeficientes significativos no primeiro e no quarto ano, mas tem impacto em todos os anos, por ser estatisticamente significativa no primeiro ano	FEP48
Teve coeficientes significativos no primeiro, no segundo e quarto ano, mas tem impacto em todos os anos, por ser estatisticamente significativa no primeiro ano	FEP41
Teve coeficiente significativo apenas no segundo e quarto anos, portanto, impactando apenas nesses anos	FEP53
Teve coeficiente significativo apenas no terceiro e quarto anos, portanto, impactando apenas nesses anos	FEP40
Teve coeficiente significativo apenas no segundo ano, portanto, impactando apenas nesse ano	FEP38, FEP45
Teve coeficiente significativo apenas no quarto ano, portanto, impactando apenas nesse ano	FEP33
Não teve coeficiente significativo em nenhum dos anos e, portanto, não tem impactos	FEP36, FEP42, FEP43, FEP44, FEP46, FEP52

O Quadro 83 traz uma síntese gráfica dessas informações.

Quadro 83 - Anos em que as variáveis do Grupo 3 do Nível 3 em Matemática impactam

	Ano1	Ano2	Ano3	Ano4
FEP33				((#))
FEP36				
FEP37	((#))	#	#	#
FEP38		((#))		
FEP39	((#))	#	#	#
FEP40			((#))	((#))
FEP41	((#))	((#))	#	((#))
FEP42				
FEP43				
FEP44				
FEP45		((#))		
FEP46				
FEP47	((#))	((#))	#	#
FEP48	((#))	#	#	((#))
FEP49	((#))	((#))	((#))	((#))
FEP52				
FEP53		((#))		((#))
FEP54	((#))	#	#	#

Essas sínteses são produzidas a partir da análise do valor agregado que a variável provoca na proficiência final quando o valor de uma unidade é atribuído à variável explicativa, cujos valores são apresentados no Quadro 84.

Quadro 84 - Valores de proficiência finais e valor agregado por variável explicativa do Grupo 3 do Nível 3 em Matemática
(continua)

	Profic. Inic.	Baixa	Média	Alta		Profic Inicial		
						Baixa	Média	Alta
Ano1	Base	109,2	121,3	133,5				
	FEP33	109,2	121,3	133,5				
	FEP36	109,2	121,3	133,5				
	FEP37	111,5	124,4	137,3		2,3	3,1	3,9
	FEP38	109,2	121,3	133,5				
	FEP39	106,5	119,7	133,0		-2,7	-1,6	-0,5
	FEP40	109,2	121,3	133,5				
	FEP41	106,7	118,8	131,0		-2,5	-2,5	-2,5
	FEP42	109,2	121,3	133,5				
	FEP43	109,2	121,3	133,5				
	FEP44	109,2	121,3	133,5				
	FEP45	109,2	121,3	133,5				
	FEP46	109,2	121,3	133,5				
	FEP47	110,3	122,9	135,4		1,2	1,5	1,9
	FEP48	107,5	119,7	131,9		-1,6	-1,6	-1,6
	FEP49	106,8	119,0	131,2		-2,3	-2,3	-2,3
	FEP52	109,2	121,3	133,5				
	FEP53	109,2	121,3	133,5				
FEP54	111,0	123,1	135,3		1,8	1,8	1,8	
Ano2	Base	120,0	177,6	234,9				
	FEP33	120,0	177,6	234,9				
	FEP36	120,0	177,6	234,9				
	FEP37	123,0	181,5	239,7		3,0	3,9	4,8
	FEP38	114,3	172,0	229,3		-5,7	-5,7	-5,7
	FEP39	118,3	177,2	235,7		-1,7	-0,4	0,8
	FEP40	120,0	177,6	234,9				
	FEP41	125,6	179,3	232,6		5,6	1,6	-2,3
	FEP42	120,0	177,6	234,9				
	FEP43	120,0	177,6	234,9				
	FEP44	120,0	177,6	234,9				
	FEP45	127,1	181,3	235,2		7,1	3,7	0,3
	FEP46	120,0	177,6	234,9				
	FEP47	123,8	179,8	235,4		3,8	2,1	0,4
	FEP48	118,3	176,0	233,3		-1,6	-1,6	-1,6
	FEP49	117,0	177,3	237,3		-2,9	-0,3	2,4
	FEP52	120,0	177,6	234,9				
	FEP53	121,6	174,7	227,6		1,6	-2,9	-7,4
FEP54	121,8	179,4	236,7		1,8	1,8	1,8	

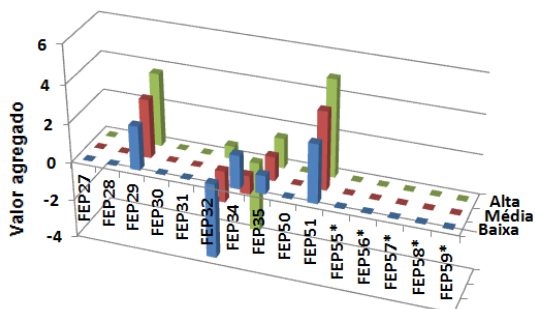
(conclusão)

	Profic. Inic.	Baixa	Média	Alta	Valor Agregado	Profic Inicial		
						Baixa	Média	Alta
Ano3	Base	166,8	198,8	230,8				
	FEP33	166,8	198,8	230,8				
	FEP36	166,8	198,8	230,8				
	FEP37	169,9	203,4	236,9	3,0	4,5	6,0	
	FEP38	166,8	198,8	230,8				
	FEP39	165,2	199,3	233,5	-1,7	0,5	2,7	
	FEP40	163,1	195,1	227,1	-3,7	-3,7	-3,7	
	FEP41	164,4	196,3	228,3	-2,5	-2,5	-2,5	
	FEP42	166,8	198,8	230,8				
	FEP43	166,8	198,8	230,8				
	FEP44	166,8	198,8	230,8				
	FEP45	166,8	198,8	230,8				
	FEP46	166,8	198,8	230,8				
	FEP47	168,3	201,1	233,8	1,5	2,2	3,0	
	FEP48	165,2	197,2	229,2	-1,6	-1,6	-1,6	
	FEP49	167,8	197,3	226,9	0,9	-1,5	-3,9	
	FEP52	166,8	198,8	230,8				
	FEP53	166,8	198,8	230,8				
	FEP54	168,7	200,6	232,6	1,8	1,8	1,8	
Ano4	Base	225,2	246,3	267,4				
	FEP33	216,3	242,4	268,5	-8,9	-3,9	1,1	
	FEP36	225,2	246,3	267,4				
	FEP37	229,2	251,9	274,6	3,9	5,6	7,2	
	FEP38	225,2	246,3	267,4				
	FEP39	224,9	248,3	271,8	-0,4	2,0	4,4	
	FEP40	220,5	241,6	262,7	-4,7	-4,7	-4,7	
	FEP41	230,4	251,5	272,6	5,2	5,2	5,2	
	FEP42	225,2	246,3	267,4				
	FEP43	225,2	246,3	267,4				
	FEP44	225,2	246,3	267,4				
	FEP45	225,2	246,3	267,4				
	FEP46	225,2	246,3	267,4				
	FEP47	227,2	249,1	271,0	2,0	2,8	3,6	
	FEP48	221,4	241,5	261,7	-3,9	-4,8	-5,7	
	FEP49	225,2	246,3	267,4	0,0	0,0	0,0	
	FEP52	225,2	246,3	267,4				
	FEP53	221,3	240,7	260,2	-4,0	-5,6	-7,3	
	FEP54	227,0	248,1	269,2	1,8	1,8	1,8	

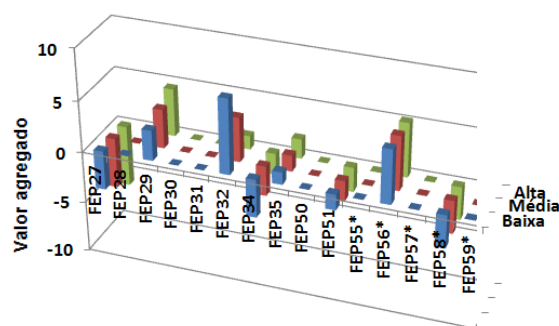
Para facilitar a interpretação do comportamento de cada uma dessas variáveis, eles são apresentados nos Gráficos 38A, 38B, 38C e 38D representam.

Gráfico 37 - Valor agregado por unidade da variável explicativa do Grupo 3 do Nível 3 em Matemática

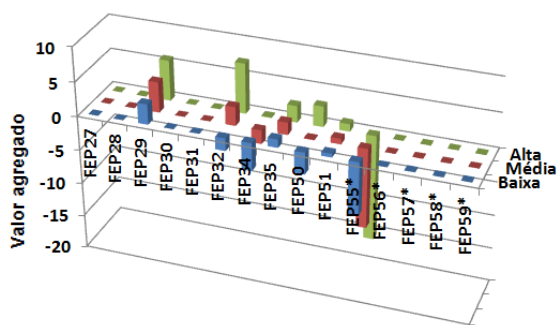
A – Primeiro ano de escolarização



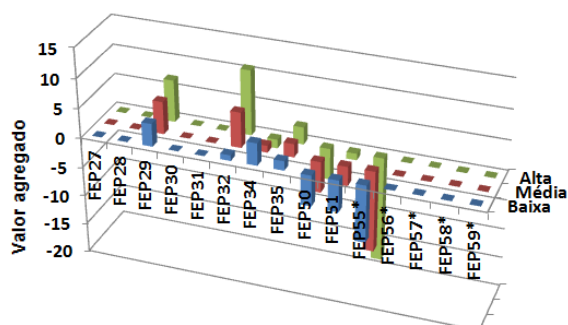
B – Segundo ano de escolarização



C – Terceiro ano de escolarização



D – Quarto ano de escolarização



O Quadro 85 traz uma síntese do impacto de cada uma das variáveis explicativas de Nível 3 na proficiência final em cada um dos anos de escolarização.

Quadro 85 - Tipo de impacto das variáveis do Grupo 3 do Nível 3 em Matemática a cada ano

	Ano1	Ano2	Ano3	Ano4
FEP33				Tipo 6
FEP36				
FEP37	Tipo 3	Tipo 3	Tipo 3	Tipo 3
FEP38		Tipo 5		
FEP39	Tipo 8	Tipo 6	Tipo 6	Tipo 6
FEP40			Tipo 5	Tipo 5
FEP41	Tipo 5	Tipo 2	Tipo 5	Tipo 1
FEP42				
FEP43				
FEP44				
FEP45		Tipo 4		
FEP46				
FEP47	Tipo 3	Tipo 4	Tipo 3	Tipo 3
FEP48	Tipo 5	Tipo 5	Tipo 5	Tipo 7
FEP49	Tipo 5	Tipo 6	Tipo 2	Tipo 1
FEP52				
FEP53		Tipo 2		Tipo 7
FEP54	Tipo 1	Tipo 1	Tipo 1	Tipo 1

O Quadro 86 apresenta a síntese na interpretação do comportamento dessas variáveis.

Quadro 86 - Interpretação final das variáveis do Grupo 2 do Nível 3 em Leitura

Comportamento	Variáveis
Coerente com o esperado	FEP37, FEP45, FEP47, FEP48, FEP54,
Incoerente com o esperado	FEP38, FEP40,
Instável, dificultando a interpretação	FEP33, FEP39, FEP41, FEP49, FEP53
Não influenciou	FEP36, FEP42, FEP43, FEP44, FEP46, FEP52

5.4.2.1.4) Análise das variáveis pertencentes ao Grupo 4 (características da escola extraídas do questionário de contexto do diretor) do Modelo de Nível 3 em Matemática

Este grupo de variáveis de Nível 3 é composto por 15 das 59 variáveis existentes. Quadro 87 apresenta uma síntese do ano no qual essas variáveis foram significantes e o tipo de impacto que essa variável teve sobre a proficiência final em Matemática .

Quadro 87 - Relação entre as variáveis explicativas do Grupo 4 do Nível 3 em Matemática e os anos de escolarização que elas influenciaram

Teve coeficientes significativos em todos os anos de escolarização, impactando em todos esses anos	FEP34, FEP51
Teve coeficientes significativos apenas no primeiro ano, mas tem impacto em todos os anos, por ser estatisticamente significativa no primeiro ano	FEP29, FEP35
Teve coeficientes significativos no primeiro e segundo ano, mas tem impacto em todos os anos, por ser estatisticamente significativa no primeiro ano	FEP32
Teve coeficientes significativos no terceiro e quarto ano, impactando apenas nesses anos	FEP50, FEP55*
Teve coeficiente significativo no segundo ano, portanto, impactando apenas nesse ano	FEP27, FEP56*, FEP58*
Não teve coeficiente significativo em nenhum dos anos e, portanto, não tem impactos	FEP28, FEP30, FEP31, FEP57*, FEP59

O Quadro 88 traz uma síntese gráfica dessas informações.

Quadro 88 - Anos em que as variáveis do Grupo 4 do Nível 3 em Matemática impactam

	Ano1	Ano2	Ano3	Ano4
FEP27		((#))		
FEP28				
FEP29	((#))	#	#	#
FEP30				
FEP31				
FEP32	((#))	((#))	#	#
FEP34	((#))	((#))	((#))	((#))
FEP35	((#))	#	#	#
FEP50			((#))	((#))
FEP51	((#))	((#))	((#))	((#))
FEP55*			((#))	((#))
FEP56*		((#))		
FEP57*				
FEP58*		((#))		
FEP59*				

Essas sínteses são produzidas a partir da análise do valor agregado que a variável provoca na proficiência final quando o valor de uma unidade é atribuído à variável explicativa, cujos valores são apresentados no Quadro 89.

Quadro 89 - Valores de proficiência finais e valor agregado por variável explicativa do Grupo 3 do Nível 3 em Matemática (continua)

Ano1	Profic. Inic.	Baixa	Média	Alta	Profic Inicial		
					Baixa	Média	Alta
	Base	169,8	203,2	236,7			
	FEP27	109,2	121,3	133,5			
	FEP28	109,2	121,3	133,5			
	FEP29	111,4	124,3	137,2	2,3	3,0	3,7
	FEP30	109,2	121,3	133,5			
	FEP31	109,2	121,3	133,5			
	FEP32	105,2	119,7	134,1	-3,9	-1,6	0,7
	FEP34	110,9	120,4	129,9	1,7	-1,0	-3,6
	FEP35	110,1	122,5	135,0	0,9	1,2	1,5
	FEP50	109,2	121,3	133,5			
	FEP51	112,2	125,3	138,4	3,0	4,0	4,9
	FEP55*	109,2	121,3	133,5			
	FEP56*	109,2	121,3	133,5			
	FEP57*	109,2	121,3	133,5			
	FEP58*	109,2	121,3	133,5			
	FEP59	109,2	121,3	133,5			

(continua)

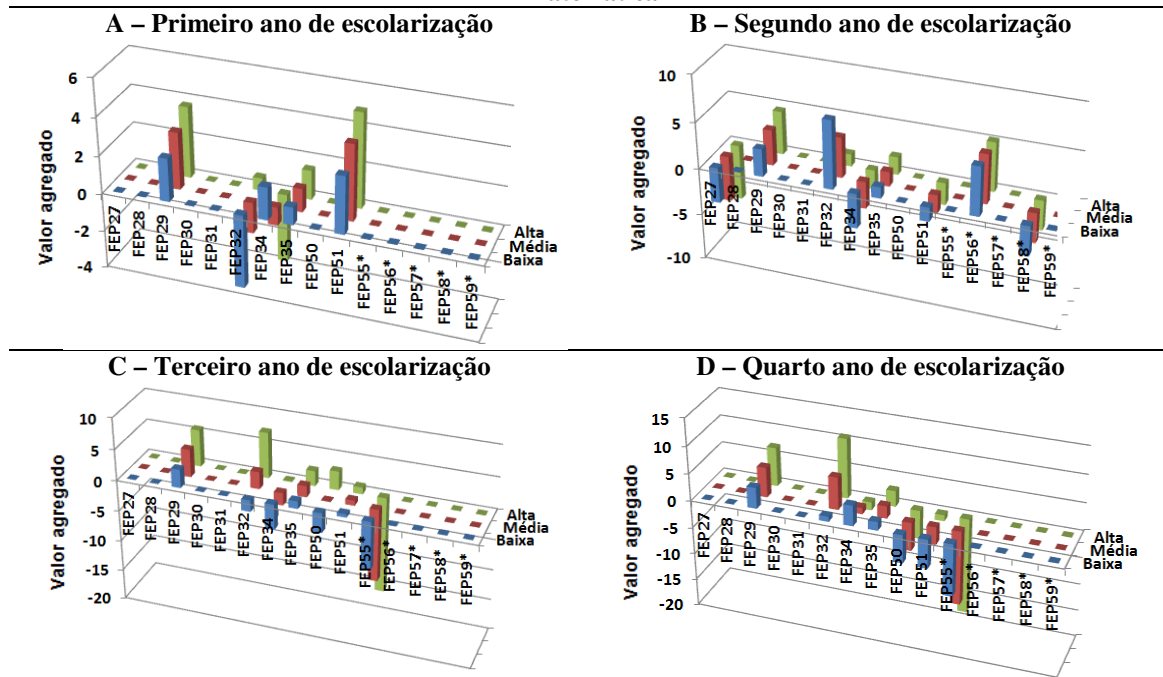
Ano2	Profic. Inic.	Baixa	Média	Alta	Valor Agregado	Profic Inicial		
						Baixa	Média	Alta
Ano2	Base	166,8	198,8	230,8				
	FEP27	116,1	172,7	228,9		-3,8	-4,9	-6,0
	FEP28	120,0	177,6	234,9				
	FEP29	122,9	181,4	239,5		2,9	3,8	4,6
	FEP30	120,0	177,6	234,9				
	FEP31	120,0	177,6	234,9				
	FEP32	127,3	181,9	236,2		7,3	4,3	1,3
	FEP34	116,2	174,7	232,9		-3,7	-2,9	-2,1
	FEP35	121,2	179,2	236,8		1,2	1,5	1,9
	FEP50	120,0	177,6	234,9				
	FEP51	118,4	175,7	232,5		-1,5	-1,9	-2,4
	FEP55*	120,0	177,6	234,9				
	FEP56*	125,2	182,9	240,1		5,2	5,2	5,2
	FEP57*	120,0	177,6	234,9				
	FEP58*	116,8	174,4	231,7		-3,2	-3,2	-3,2
	FEP59*	120,0	177,6	234,9				
Ano3	Profic. Inic.	Baixa	Média	Alta	Valor Agregado	Profic Inicial		
						Baixa	Média	Alta
Ano3	Base	225,2	246,3	267,4				
	FEP27	166,8	198,8	230,8				
	FEP28	166,8	198,8	230,8				
	FEP29	169,8	203,2	236,7		2,9	4,4	5,9
	FEP30	166,8	198,8	230,8				
	FEP31	166,8	198,8	230,8				
	FEP32	165,0	201,5	238,1		-1,8	2,7	7,3
	FEP34	162,8	196,8	230,9		-4,1	-2,0	0,1
	FEP35	168,0	200,6	233,2		1,2	1,8	2,4
	FEP50	163,7	198,7	233,8		-3,1	-0,1	3,0
	FEP51	167,4	199,6	231,9		0,5	0,8	1,1
	FEP55*	159,1	187,2	215,3		-7,8	-11,7	-15,5
	FEP56*	166,8	198,8	230,8				
	FEP57*	166,8	198,8	230,8				
	FEP58*	166,8	198,8	230,8				
	FEP59*	166,8	198,8	230,8				

(conclusão)

Ano4	Profic. Inic.	Baixa	Média	Alta	Profic Inicial		
					Baixa	Média	Alta
Proficiência Final	Base	225,2	246,3	267,4			
	FEP27	225,2	246,3	267,4			
	FEP28	225,2	246,3	267,4			
	FEP29	229,1	251,8	274,4	3,8	5,4	7,0
	FEP30	225,2	246,3	267,4			
	FEP31	225,2	246,3	267,4			
	FEP32	226,2	252,2	278,3	0,9	5,9	10,9
	FEP34	229,0	247,5	265,9	3,7	1,1	-1,5
	FEP35	226,8	248,5	270,3	1,6	2,2	2,9
	FEP50	220,1	241,1	262,2	-5,2	-5,2	-5,2
	FEP51	219,9	243,1	266,3	-5,3	-3,2	-1,1
	FEP55*	215,6	232,7	249,8	-9,6	-13,6	-17,7
	FEP56*	225,2	246,3	267,4			
	FEP57*	225,2	246,3	267,4			
	FEP58*	225,2	246,3	267,4			
FEP59*	225,2	246,3	267,4				

Os Gráficos 39A, 39B, 39C e 39D representam, para cada ano, o valor que cada unidade da variável explicativa de Nível 3 impacta na proficiência final de alunos com diferentes proficiências iniciais.

Gráfico 38 - Valor agregado por unidade da variável explicativa do Grupo 4 do Nível 3 em Matemática



O Quadro 90 traz uma síntese do impacto de cada uma das variáveis explicativas de Nível 3 na proficiência final em cada um dos anos de escolarização.

Quadro 90 - Tipo de impacto das variáveis do Grupo 4 do Nível 3 em Matemática a cada ano

	Ano1	Ano2	Ano3	Ano4
FEP27		Tipo 7		
FEP28				
FEP29	Tipo 3	Tipo 3	Tipo 3	Tipo 3
FEP30				
FEP31				
FEP32	Tipo 6	Tipo 4	Tipo 6	Tipo 3
FEP34	Tipo 2	Tipo 8	Tipo 6	Tipo 2
FEP35	Tipo 3	Tipo 3	Tipo 3	Tipo 3
FEP50			Tipo 6	Tipo 5
FEP51	Tipo 3	Tipo 7	Tipo 3	Tipo 8
FEP55*			Tipo 7	Tipo 7
FEP56*		Tipo 1		
FEP57*				
FEP58*		Tipo 5		
FEP59				

O Quadro 91 apresenta a síntese na interpretação do comportamento dessas variáveis.

Quadro 91 - Interpretação final das variáveis do Grupo 4 do Nível 3 em Matemática

Comportamento	Variáveis
Coerente com o esperado	FEP34, FEP35, FEP56*,
Incoerente com o esperado	FEP27, FEP29, FEP58*
Instável, dificultando a interpretação	FEP32, FEP50, FEP51, FEP55*, FEP57*, FEP59*
Não influenciou	FEP28, FEP30, FEP31,

5.4.2.1.5) Síntese da análise das variáveis do Modelo de Nível 3 em Matemática

Para a interpretação do efeito de cada uma das variáveis do Nível 3, foram consideradas as estimativas dos coeficientes associados a cada uma das variáveis considerando as proficiência iniciais e finais de cada um dos diferentes anos de escolarização, porém, para os Anos 2, 3 e 4, tais coeficientes são incrementos, sejam eles positivos ou negativos, aos coeficientes estimados para o Ano 1. Por isso, os coeficientes do Ano 1 são referência para a análise dos demais anos.

Como consequência desta lógica, tem-se que:

- Para o Ano 1, as variáveis que apresentam coeficientes significativos para o intercepto, para a inclinação ou ainda para ambos, deve-se entender que essas variáveis têm impactos significativos na determinação da proficiência final, em relação à proficiência inicial, neste primeiro ano. Este impacto pode ser um ganho ou uma perda.

- Para os demais anos, as variáveis que apresentam coeficientes significativos, para o intercepto, para a inclinação ou ainda para ambos, deve-se entender que essa variável tem um valor para tais coeficientes significativamente diferente àqueles estimados para o Ano 1. Ou seja, esse coeficiente representa um valor que deve ser somado algebricamente valor estimado para esse coeficiente no Ano 1. Disso decorre que:

a) Quando uma variável é significativa no Ano 1 e não dos demais anos, significa que ela também tem igual impacto nesses outros anos. Vale a pena observar que, mesmo com coeficientes iguais, os impactos são distintos já que os valores das proficiências iniciais nesses diferentes anos de escolarização. Esse comportamento foi observado nas variáveis FEE07, FED12, FEP29, FEP35, FEP37, FEP39, FEP54.

b) Quando uma variável é significativa no Ano 1 e também em qualquer um dos anos seguintes, significa que essa variável tem impacto em todos os anos, porém de forma diferenciada no ano em que a variável foi significativa no Ano 1 e no ano analisado, e ainda que esse impacto é diferentes nesses dois anos. O Quadro 92 mostra as variáveis que têm efeitos sobre a proficiência final no primeiro ano e um efeito distinto nos demais anos.

Quadro 92 - Variáveis do Nível 3 que têm impacto na proficiência em Matemática no Ano 1 e com coeficiente diferente em outro ano

Anos	Variáveis
1 e 2	FEE11, FED25*, FEP32, FEP47
1 e 3	
1 e 4	FED13, FEP48
1, 2 e 3	FEE03
1, 2 e 4	FEP41
1, 3 e 4	FED18
1, 2, 3 e 4	FEA01, FEE02, FEE21, FEP34, FEP49, FEP51

c) Quando uma variável não é significativa no Ano 1, mas o é em qualquer dos anos seguintes, significa que ela tem impacto somente neste ano em que a variável foi significativa. O Quadro 93 mostra as variáveis e os respectivos anos que apresentaram esse efeito.

Quadro 93 - Variáveis que têm impacto no Ano 1 e têm o mesmo coeficiente em outro Ano (significa que a variável não foi significativa nesse ano)

Anos	Variáveis
3 e 4	FEP40, FEP50, FEP55*
2 e 4	FEP53
2 e 3	FEE04
4	FEE09, FEE10, FED16, FED17, FED19, FED20, FEP33
3	FEE08
2	FED22*, FED23*, FEP38, FEP45, FEP27, FEP56*, FEP58*
2, 3 e 4	

d) Quando a variável não tem qualquer efeito significativo em qualquer um dos anos seus coeficientes são zerados, sendo eliminadas do modelo. As variáveis são: FEE05, FEE06, FED14, FED15, FED24*, FED26*, FEP28, FEP30, FEP31, FEP36, FEP42, FEP43, FEP44, FEP46, FEP52, FEP57*, FEP59.

5.4.3) Síntese do comparativa das variáveis de Nível 3 nos modelos de Leitura e de Matemática.

O Quadro 94 agrupa os resultados encontrados nos dois modelos criados, distribuindo as variáveis explicativas do Nível 3 em função do impacto que teve em cada um dos anos de escolarização.

Na comparação entre os dois modelos, apenas as variáveis FEA01 (NSE médio da escola a partir do NSE dos alunos) e FEP39 (Proposição de cópia ou caligrafia) foram igualmente significantes nos dois modelos. Já a variável FEP59* (Número médio de pessoas da família que mora com o professor) foi a única não significativa nos dois modelos. Essas variáveis estão destacadas em negrito no Quadro 94.

Quadro 94 - Síntese das variáveis significantes nos dois modelos

Anos de impacto	Modelo de Nível 3 em Leitura	Modelo de Nível 3 em Matemática
Igualmente em todos os anos	FEE05, FEE08, FED14, FED25*, FEP31, FEP32, FEP33, FEP34, FEP39 , FEP40 e FEP58*	FEE07, FED12, FEP29, FEP35, FEP37, FEP39 , FEP54
Diferentemente em todos os anos	FEA01 , FEE02 , FED14, FED25*	FEA01 , FEE02 , FEE21, FEP34, FEP49, FEP51
2	FED20; FEP30;	FED22*, FED23*, FEP38, FEP45, FEP27, FEP56*, FEP58*
3	FED15	FEE08
4	FEE09 ; FED18 ; FEP28	FEE09 , FEE10, FED16, FED17, FED19, FED20, FEP33
1 e 2	FEP54	FEE11, FED25*, FEP32, FEP47
1 e 3	FEP42, FEP43	
1 e 4	FEE04, FED13 , FEP36, FEP37, FEP41, FEP46, FEP47	FED13 , FEP48
1, 2 e 3		FEE03
1, 2 e 4	FED23* , FEP38, FEP45, FEP48, FEP29, FEP35	FEP41
1, 3 e 4	FED12 , FED22*, FEP51	FED18
3 e 4	FEE11 ; FEP44	FEP40, FEP50, FEP55*
2 e 4	FEE06 ; FEE07 ; FED21; FEP52, FEP53 ; FEP50, FEP55*	FEP53
2 e 3	FEE03 ; FEE10 ; FED24* ; FED26* ; FEP27*	FEE04
2, 3 e 4	FED19	
Não tiveram impacto	FED16, FEP27, FEP49, FEP56*, FEP59 *	FEE05, FEE06, FED14, FED15, FED24*, FED26*, FEP28, FEP30, FEP31, FEP36, FEP42, FEP43, FEP44, FEP46, FEP52, FEP57*, FEP59 *

Em um primeiro momento, considerando que os alunos, professores e escolas que compõem o banco de dados são praticamente os mesmos, era esperado que as diferenças na significância estatística entre as variáveis fossem constatadas apenas naquelas inerentes às diferenças entre Leitura e Matemática.

**CAPÍTULO 6 -
ANÁLISE INDIVIDUAL DAS VARIÁVEIS**

No capítulo anterior foram apresentados os efeitos das variáveis associadas ao perfil dos alunos e às características das escolas estimados em modelos multinível com todas elas inseridas simultaneamente, quando se constatou uma significativa instabilidade nos efeitos das variáveis e até mesmo resultados inesperados que vão de encontro ao que é esperado, chegando a dar indícios da existência de equívocos nos resultados.

Buscando melhorar essa interpretação e reconhecendo a influência que o nível socioeconômico individual (NSE) e que o nível socioeconômico médio da escola (MNSE) tem sobre a proficiência final, largamente evidenciada nos capítulos anteriores, optou-se por ajustar um Modelo de Referência, descrito nas próximas seções deste capítulo para analisar o efeito de cada uma das variáveis sobre a proficiência final desconsiderando os efeitos que as demais têm sobre ela.

Para o ajuste desse modelo foi usado novamente o método *backward*, mantendo a proficiência final como variável resposta, com a inserção das seguintes variáveis explicativas:

- Nível 1 – Medidas Cognitivas no tempo: proficiência inicial a cada ano de escolarização, variável de controle sobre a retenção do aluno e as variáveis que indicam o ano de escolarização;
- Nível 2 – Aluno: nível socioeconômico do aluno
- Nível 3 – Escola: nível socioeconômico médio da escola

Com a inserção dessas variáveis, segundo esse critério, para verificar onde essas variáveis ligadas ao nível socioeconômico seriam estatisticamente significativas com base nesses dados, o modelo no inicial ficou:

Nível 1 – Medidas Cognitivas no Tempo:

$$PR_FI_{ijk} = \pi_{0jk} + \pi_{1jk} \cdot T2_{ijk} + \pi_{2jk} \cdot T3_{ijk} + \pi_{3jk} \cdot T4_{ijk} + \pi_{4jk} \cdot PR_IN_{ijk} + \pi_{5jk} \cdot PR_IN \times T2_{ijk} + \pi_{6jk} \cdot PR_IN \times T3_{ijk} + \pi_{7jk} \cdot PR_IN \times T4_{ijk} + \pi_{8jk} \cdot RET_{ijk} + e_{ijk}, \text{ com } e_{ijk} \sim NID(0, \sigma^2).$$

Nível 2 – Alunos:

$$\pi_{0jk} = \beta_{00k} + \beta_{01k} \cdot NSE_{jk} + r_{0jk}, \text{ com } r_{0jk} \sim NID(0, \tau_{\pi 00})$$

$$\pi_{1jk} = \beta_{10k} + \beta_{11k} \cdot NSE_{ijk}$$

$$\pi_{2jk} = \beta_{20k} + \beta_{21k} \cdot NSE_{ijk}$$

$$\pi_{3jk} = \beta_{30k} + \beta_{31k} \cdot NSE_{ijk}$$

$$\pi_{4jk} = \beta_{40k} + \beta_{41k} \cdot NSE_{ijk} + r_{4jk}, \text{ com } r_{4jk} \sim NID(0, \tau_{\pi 40})$$

$$\pi_{5jk} = \beta_{50k} + \beta_{51k} \cdot NSE_{ijk}$$

$$\pi_{6jk} = \beta_{60k} + \beta_{61k} \cdot NSE_{ijk}$$

$$\pi_{7jk} = \beta_{70k} + \beta_{71k} \cdot NSE_{ijk}$$

$$\pi_{8jk} = \beta_{80k}$$

Nível 3 – Escolas:

$$\beta_{00k} = \gamma_{000} + \gamma_{001} \cdot \text{MNSE}_k + u_{0jk}, \text{ com } u_{0jk} \sim \text{NID}(0, \tau_{\beta 000})$$

$$\beta_{01k} = \gamma_{010}$$

$$\beta_{10k} = \gamma_{100} + \gamma_{101} \cdot \text{MNSE}_k$$

$$\beta_{11k} = \gamma_{110}$$

$$\beta_{20k} = \gamma_{200} + \gamma_{201} \cdot \text{MNSE}_k$$

$$\beta_{21k} = \gamma_{210}$$

$$\beta_{30k} = \gamma_{300} + \gamma_{301} \cdot \text{MNSE}_k$$

$$\beta_{31k} = \gamma_{310}$$

$$\beta_{40k} = \gamma_{400} + \gamma_{401} \cdot \text{MNSE}_k + u_{4jk}, \text{ com } u_{4jk} \sim \text{NID}(0, \tau_{\beta 400})$$

$$\beta_{41k} = \gamma_{410}$$

$$\beta_{50k} = \gamma_{500} + \gamma_{501} \cdot \text{MNSE}_k$$

$$\beta_{51k} = \gamma_{510}$$

$$\beta_{60k} = \gamma_{600} + \gamma_{601} \cdot \text{MNSE}_k$$

$$\beta_{61k} = \gamma_{610}$$

$$\beta_{70k} = \gamma_{700} + \gamma_{701} \cdot \text{MNSE}_k$$

$$\beta_{71k} = \gamma_{710}$$

$$\beta_{80k} = \gamma_{800}$$

Neste modelo inicial os coeficientes do Nível 1, nomeadamente, π_{0jk} , π_{1jk} , π_{2jk} e π_{3jk} definem o intercepto das retas de regressão de cada um dos anos de escolarização, enquanto que os coeficientes π_{4jk} , π_{5jk} , π_{6jk} e π_{7jk} definem o coeficiente de inclinação em função da proficiência inicial nessas retas. O coeficiente π_{8jk} , que também interfere no intercepto dessas retas faz o incremento na proficiência final quando o aluno não foi retido naquele ano em que sua proficiência foi medida.

Dadas as características específicas de modelos hierárquicos, deve-se notar que para essa modelagem cada um desses coeficientes, com exceção de π_{8jk} são influenciados pelas equações do Nível 2, com novos interceptos (β_{00k} , β_{10k} , β_{20k} e β_{30k}) e coeficientes de inclinação (β_{40k} , β_{50k} , β_{60k} e β_{70k}), em que esses últimos são determinados pelo nível socioeconômico individual do aluno (NSE)⁴⁶.

Nessa modelagem, para cada um dos interceptos das equações do Nível 2 é especificada uma nova equação no Nível 3, com intercepto e coeficiente de inclinação, agora colocado em função do nível socioeconômico médio da escola (MNES).

46 Visando a complexidade do modelo e como o foco deste estudo não foi a análise da influencia do nível socioeconômico sobre os alunos retidos as variáveis não foram inseridas nos níveis 2 e 3.

Essa modelagem é importante para viabilizar a análise da influência dessas duas variáveis após a entrada do aluno na escola, em cada um dos anos de escolarização, uma vez, dada alta correlação existente entre a proficiência inicial e o nível socioeconômico do aluno.

6.1) Modelo de Referência em Leitura

Após o ajuste do modelo, no qual as variáveis que controlam o nível socioeconômico do aluno (NSE) e o nível socioeconômico médio da escola (MNSE) permaneceram apenas nos anos em que foram estatisticamente significativas, ficou:

Nível 1 – Medidas Cognitivas no Tempo:

$$PR_FI_{ijk} = \pi_{0jk} + \pi_{1jk} \cdot T2_{ijk} + \pi_{2jk} \cdot T3_{ijk} + \pi_{3jk} \cdot PR_IN_{ijk} + \pi_{4jk} \cdot PR_IN \times T2_{ijk} + \pi_{5jk} \cdot PR_IN \times T3_{ijk} + \pi_{6jk} \cdot PR_IN \times T4_{ijk} + \pi_{7jk} \cdot RET_{ijk} + e_{ijk}, \text{ com } e_{ijk} \sim NID(0, \sigma^2).$$

Nível 2 – Alunos:

$$\pi_{0jk} = \beta_{00k} + \beta_{01k} \cdot NSE_{jk} + r_{0jk}, \text{ com } r_{0jk} \sim NID(0, \tau_{\pi 00})$$

$$\pi_{1jk} = \beta_{10k}$$

$$\pi_{2jk} = \beta_{20k} + \beta_{21k} \cdot NSE_{ijk}$$

$$\pi_{3jk} = \beta_{30k} + r_{3jk}, \text{ com } r_{3jk} \sim NID(0, \tau_{\pi 30})$$

$$\pi_{4jk} = \beta_{40k}; \pi_{5jk} = \beta_{50k}; \pi_{6jk} = \beta_{60k}; \pi_{7jk} = \beta_{70k};$$

Nível 3 – Escolas:

$$\beta_{00k} = \gamma_{000} + \gamma_{001} \cdot MNSE_k + u_{0jk}, \text{ com } u_{0jk} \sim NID(0, \tau_{\beta 000})$$

$$\beta_{01k} = \gamma_{010}$$

$$\beta_{10k} = \gamma_{100} + \gamma_{101} \cdot MNSE_k$$

$$\beta_{20k} = \gamma_{200}; \beta_{21k} = \gamma_{210}$$

$$\beta_{30k} = \gamma_{300} + \gamma_{301} \cdot MNSE_k + u_{3jk}, \text{ com } u_{3jk} \sim NID(0, \tau_{\beta 300})$$

$$\beta_{40k} = \gamma_{400}; \beta_{50k} = \gamma_{500}; \beta_{60k} = \gamma_{600}; \beta_{70k} = \gamma_{700}$$

Na equação do Nível 1 - Medidas Cognitivas no Tempo - tem como variáveis explicativas:

- PR_FI_{ijk} é a proficiência final em Leitura do ano i do aluno j da escola k ;
- $T2$, $T3$ e $T4$, são variáveis indicadoras usadas para identificar o ano de escolarização⁴⁷;
- PR_IN_{ijk} é proficiência inicial no ano de escolarização i do aluno j , da escola k ;

47 Explicações mais detalhada sobre essas variáveis encontram-se no Capítulo 2.

- RET_{ijk} é variável que indica se o aluno j da escola k fez a prova na condição de retido no ano de escolarização i ;

- Pr_INxT2 , Pr_INxT3 e Pr_INxT4 são variáveis indicadoras que armazenam a proficiência final de cada um dos anos de escolarização⁴⁸.

Nesta equação do Nível são coeficientes (estimadores):

- π_{0jk} representa a proficiência final média dos J_k alunos da escola k , que tiraram zero de proficiência inicial⁴⁹;

- π_{1jk} é o valor a ser adicionado a π_{0jk} para representar a proficiência final média dos J_k alunos da escola k , no segundo ano de escolarização, de maneira que essa proficiência final média é dada por $(\pi_{0jk} + \pi_{1jk})$;

- π_{2jk} é o valor a ser adicionado a π_{0jk} para representar a proficiência final média dos J_k alunos da escola k no terceiro ano, de maneira que o intercepto da reta do terceiro ano de escolarização é dados por $(\pi_{0jk} + \pi_{2jk})$;

- π_{3jk} é o incremento para cada unidade da proficiência inicial (PR_IN_{ijk}) na proficiência final do aluno j , da escola k , no primeiro ano de escolarização (PR_FI_{ijk})⁵⁰;

- π_{4jk} , π_{5jk} e π_{6jk} são os acréscimos médios a serem feitos em π_{3jk} para cada unidade da proficiência inicial (PR_IN_{ijk}) na proficiência final do aluno j , da escola k , referentes ao segundo, terceiro e quarto ano, definidos matematicamente por $(\pi_{3jk} + \pi_{4jk})$, $(\pi_{3jk} + \pi_{5jk})$ e $(\pi_{3jk} + \pi_{6jk})$, respectivamente;

- π_{7jk} é o valor a ser adicionado ao intercepto de cada ano, caso o aluno tenha participado da onda de teste sem defasagem, considerando os diferentes anos de escolarização;

- e_{ijk} é o desvio entre o valor estimado para a proficiência final do aluno j matriculado na escola k no ano de escolarização i e o valor observado (PR_FI_{ijk});

- σ^2 variância de e_{ijk} .

Na equação do Nível 2 – Alunos, tem-se:

48 Assim como nas variáveis T2, T3 e T4, há explicações mais detalhada sobre essas variáveis no Capítulo 2.

49 Este é um valor sem significado pedagógico uma vez que não há alunos proficiência inicial nula. É, portanto, um valor de referência inerente ao processo de regressão linear necessário para outras estimativas. Mais esclarecimentos sobre esse aspecto é dado a seguir.

50 Neste modelo, o valor estimado para este coeficiente referente ao quarto ano de escolarização não foi significativamente diferente de zero, ou seja, o intercepto das equações de regressão do primeiro e quarto ano são iguais. É importante lembrar que, como já descrito no capítulo anterior, esse valor não tem significado pedagógico pois não há proficiência inicial igual a zero. A distinção entre as equações destes dois anos se dá pelos diferentes coeficientes angulares e também pelos distintos intervalos de proficiências iniciais inerentes a esses dois anos de escolarização.

- NSE_{jk} é a variável que contém a informação sobre o nível socioeconômico do aluno j matriculado na escola k ;
- β_{00k} é a média das proficiências finais de todos os alunos da escola k no primeiro ano de escolarização, com nível socioeconômico médio, isto é, com $NSE_{jk} = 0$;
- β_{01k} é o acréscimo médio a ser feito em β_{00k} para cada unidade da variável que controla o nível socioeconômico (NSE_{jk});
- β_{10k} é o valor médio a ser adicionados a β_{00k} para compor a média das proficiências finais aos alunos da escola k , para ajustar o intercepto da equação de regressão referente ao segundo ano de escolarização;
- β_{20k} é o valor médio entre os alunos com $NSE_{jk} = 0$ a ser adicionado a β_{00k} para compor a média das proficiências finais de todos os alunos da escola k , para ajustar o intercepto da equação de regressão referente ao terceiro ano de escolarização⁵¹;
- β_{21k} é o acréscimo médio a ser feito em β_{20k} para cada unidade da variável que controla o nível socioeconômico (NSE_{jk});
- β_{30k} é o incremento médio na média da proficiência final para cada unidade da proficiência inicial de todos os alunos da escola k no primeiro ano de escolarização;
- β_{40k} , β_{50k} e β_{60k} são os valores a serem adicionados a β_{30k} para compor os incrementos para cada unidade da proficiência inicial considerando todos os alunos da escola k na proficiência média final dos demais anos de escolarização. Dessa maneira, o incremento no primeiro ano de escolarização é dado por β_{30k} enquanto que para o segundo ano, o incremento no segundo ano é dado por $(\beta_{30k} + \beta_{40k})$. Para o terceiro e quarto anos, os incrementos são dados, respectivamente, por $(\beta_{30k} + \beta_{50k})$ e $(\beta_{30k} + \beta_{60k})$;
- β_{70k} é o valor a ser adicionado ao intercepto de cada ano, caso o aluno tenha participado da onda de teste sem defasagem, considerando todos os alunos da escola k ;
- r_{0jk} é o desvio entre a média das proficiências finais do aluno j da escola k e a média das proficiências finais de todos os alunos dessa escola;
- r_{3jk} é o desvio entre a média dos valores a serem adicionados a β_{30k} dos alunos de cada uma das escolas e o valor médio da escola onde esse aluno está matriculado;

51 Neste modelo, deve ser notado que o nível socioeconômico do aluno afeta igualmente no intercepto das equações de regressão de todos os anos de escolarização, com exceção do terceiro ano em que a variável passou a ser significativa. Além disso, pelo fato dessa variável não ter sido estatisticamente significativa nos coeficientes de inclinação, significa que ela não interfere entre os alunos com proficiência iniciais distintas.

- $\tau_{\pi 00}$ é a variância do erro de estimativa r_{0jk} ;
- $\tau_{\pi 30}$ é a variância do erro de estimativa r_{3jk} .

Na equação do Nível 3 – Escolas, tem-se:

- $MNSE_k$ é a variável que contém a informação sobre o nível socioeconômico médio da escola k ;

- γ_{000} é a média de todas as proficiências finais de todas as escolas no primeiro ano de escolarização e que tenham nível socioeconômico médio igual à média de toda a amostra, isto é, com $MNSE_k = 0$;

- γ_{001} é o valor médio das proficiências finais das k escolas a ser acrescido a γ_{000} para cada unidade da variável que controla o nível socioeconômico médio das escolas ($MNSE_k$);

- γ_{010} é o acréscimo médio a ser feito em β_{00k} para cada unidade da variável que controla o nível socioeconômico (NSE_{jk}), referente ao primeiro ano de escolarização;

- γ_{100} é o valor médio das proficiências finais das escolas com $MNSE_k = 0$ a ser adicionado a β_{00k} para compor o intercepto da equação de regressão referente ao segundo ano de escolarização;

- γ_{101} é o valor médio a ser adicionado a γ_{100} para cada unidade da variável que controla o nível socioeconômico médio das escolas ($MNSE_k$) para compor o intercepto da equação de regressão referente ao segundo ano de escolarização;

- γ_{200} é incremento referente ao valor médio das proficiências finais das k escolas para compor o intercepto da equação de regressão referente ao terceiro ano de escolarização;

- γ_{210} é o acréscimo médio a ser feito em β_{00k} para cada unidade da variável que controla o nível socioeconômico (NSE_{jk}), referente ao terceiro ano de escolarização;

- γ_{300} é a proficiência média das escolas com $MNSE_k = 0$ a ser incrementada para cada unidade de proficiência inicial;

- γ_{301} é o valor médio das proficiências finais das k escolas a ser acrescido a γ_{300} para cada unidade da variável que controla o nível socioeconômico médio das escolas ($MNSE_k$);

- γ_{400} , γ_{500} e γ_{600} são os valores a serem adicionados a γ_{300} para compor os incrementos na média das proficiências finais, nos demais anos de escolarização, para cada unidade da proficiência inicial considerando todas as escolas;

- γ_{700} é o valor a ser adicionado ao intercepto de cada ano, caso o aluno tenha participado da onda de teste sem defasagem, considerando todas as escolas;

- u_{00k} é o desvio entre média das proficiências finais estimadas de cada escola em relação à média de todas as escolas;
- u_{30k} é o desvio entre a média dos valores estimados para serem adicionados a γ_{300} e a média desses valores de todas as escolas;
- $\tau_{\beta 000}$ é a variância de u_{00k} ;
- $\tau_{\beta 300}$ é a variância de u_{30k} .

Os valores estimados para cada um desses coeficientes estão indicados no Quadro 95 abaixo e foram retirado do relatório do programa HLM que se encontra no Anexo 11:

Quadro 95 - Resultados do Modelo de Nível 3 em Leitura

				Estimativa	E. P.	p-value	
Efeitos fixos	Intercepto	π_{0jk}	β_{00k} (Intercepto)	γ_{000}	65,873322	2,7996	<0,001
				γ_{001} (MNSE)	29,95538	5,8444	<0,001
			β_{01k}	γ_{010} (NSE)	2,213134	0,3275	<0,001
		π_{1jk} (T2)	β_{10k} (Intercepto)	γ_{100}	-31,077944	3,0342	<0,001
				γ_{101} (MNSE)	-3,603676	1,2161	0,004
		π_{2jk} (T3)	β_{20k} (Intercepto)	γ_{200}	-10,538071	2,7092	<0,001
	β_{21k}		γ_{210} (NSE)	-1,708899	0,5368	0,002	
	Coefficiente de inclinação	π_{3jk} (PR_IN)	β_{30k} (Intercepto)	γ_{300}	0,518063	0,0177	<0,001
				γ_{310} (MNSE)	-0,131387	0,0269	<0,001
		π_{4jk} (PR_INxT2)	β_{40k} (Intercepto)	γ_{400}	0,293439	0,0226	<0,001
		π_{5jk} (PR_INxT3)	β_{50k} (Intercepto)	γ_{500}	0,167611	0,0206	<0,001
		π_{6jk} (PR_INxT4)	β_{60k} (Intercepto)	γ_{600}	0,10931	0,0061	<0,001
	Inter.	π_{7jk} (RET)	β_{70k} (Intercepto)	γ_{700}	6,030269	1,5902	<0,001
	Efeitos Aleatórios				Componente de variância		
					$\sigma^2 =$	185,88	
$\tau_{\pi 00} =$					91,12	0,02	
$\tau_{\pi 30} =$					56	0,013	
$\tau_{\beta 000} =$					118,95	<0,001	
	$\tau_{\beta 300} =$	0,0032	<0,001				
Deviance:		281835	Número de parâmetros estimados:		20		
Teste quiquadrado:		345,85	Graus de liberdade:	5		<0,001	

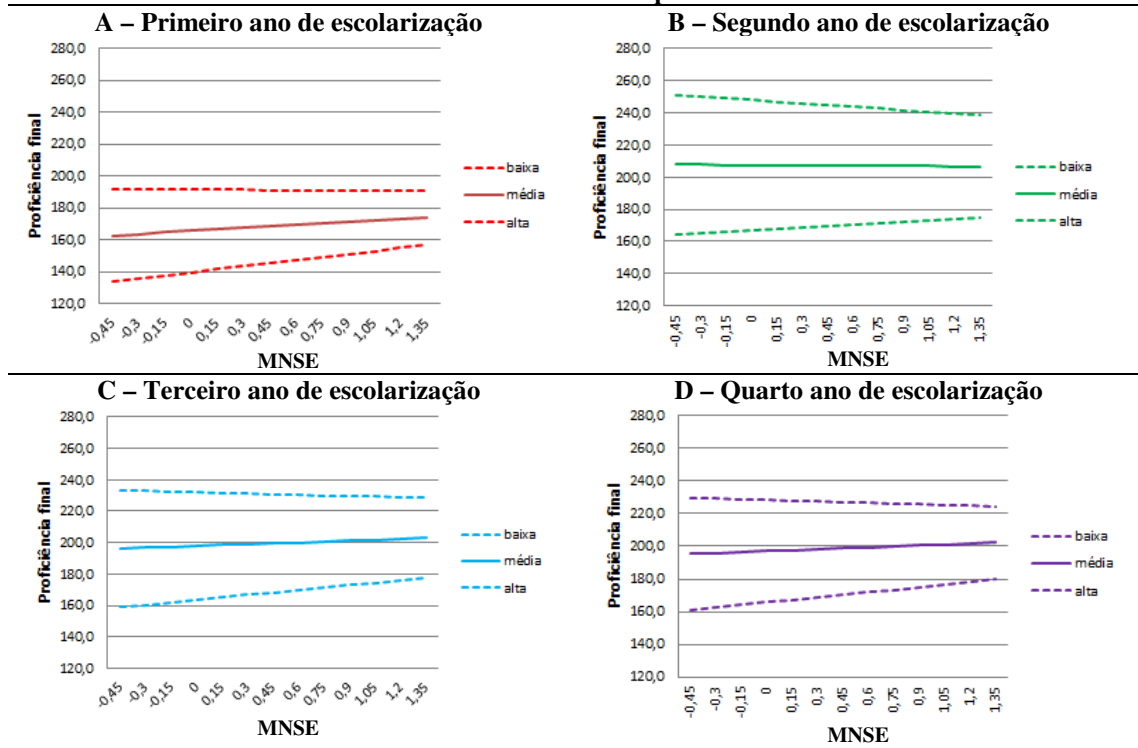
Analisando o *p-value* dos efeitos aleatórios $\tau_{\pi 00}$, $\tau_{\pi 40}$, $\tau_{\beta 000}$ e $\tau_{\beta 400}$ desse quadro, pode-se constatar que a inclusão dessas variáveis não foram suficientes para explicar toda a

variabilidade da proficiência final, indicando que este modelo pode ser usado para a inserção e novas variáveis.

Como já evidenciado ao longo deste trabalho, o nível socioeconômico médio da escola (MNSE) tem efeito diferente para alunos com distintas proficiências iniciais.

Os Gráficos 40A, 40B, 40C e 40D representam a estimativa para a proficiência final em todo o intervalo de valores encontrados para a variável MNSE na amostra dos dados em leitura.

Gráfico 39 - Influência do MNSE sobre a proficiência final em Leitura



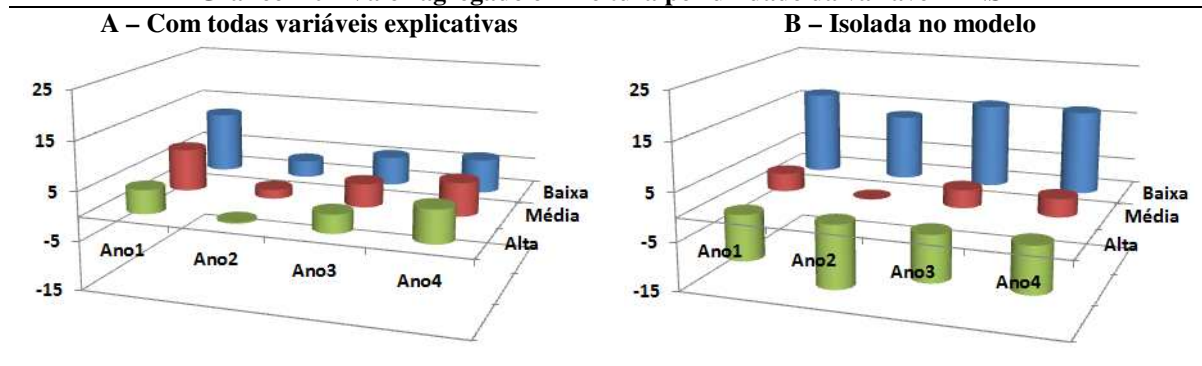
Para permitir a comparação do efeito do MNSE nos diferentes anos de escolarização, esses gráfico foram produzidos considerando como proficiência média, a média das proficiências iniciais; proficiência inicial baixa, a proficiência média subtraída de 50 unidades e alta, a proficiência média acrescida de 50 unidades. Esses valores foram adotados para evitar que as inclinações sejam influenciadas pelas diferenças entre o desvio padrão dos diferentes anos de escolarização.

Essa representação gráfica evidencia que, com essa modelagem, a variável MNSE tem um comportamento no primeiro ano de escolarização, e outro para nos demais anos.

Percebe-se que no primeiro ano de escolarização, a variável MNSE tem pouco impacto para os alunos com proficiência inicial alta e, para os demais alunos, quanto maior for o MNSE da escola, maior também será sua proficiência final. Nos demais anos de escolarização, o MNSE não interfere tanto para os alunos com proficiência inicial média, tem um efeito positivo para os alunos com proficiência inicial baixa e um efeito negativo para aqueles com proficiência inicial alta. Apesar de causar certo estranhamento inicial, este é um resultado esperado uma vez que alunos com NSE menor entram em escolas com MNSE alto e ocorre o contrário quando alunos com NSE mais alto do que MNSE da escola.

Outra constatação importante é a diferença do comportamento do valor agregado para cada unidade de MNSE no modelo onde se consideram todas as variáveis e nesse que é isolado, como mostram os Gráficos 41A e 41B.

Gráfico 40 - Valor agregado em Leitura por unidade da variável MNSE



Com isso se evidencia mais uma vez, os efeitos de uma variável depende diretamente do modelo a que ela é submetido.

6.2) Modelo de Referência em Matemática

Neste modelo, as variáveis que controlam o nível socioeconômico do aluno (NSE) e o nível socioeconômico médio da escola (MNSE) permaneceram apenas nos anos em que foram estatisticamente significativas. Este é o Modelo de Referência em Matemática usado para a análise dos efeitos das variáveis sobre a proficiência final:

Nível 1 – Medidas Cognitivas no Tempo:

$$PR_FI_{ijk} = \pi_{0jk} + \pi_{1jk} \cdot T2_{ijk} + \pi_{2jk} \cdot T3_{ijk} + \pi_{3jk} \cdot PR_IN_{ijk} + \pi_{4jk} \cdot PR_IN \times T2_{ijk} + \pi_{5jk} \cdot PR_IN \times T3_{ijk} + \pi_{6jk} \cdot PR_IN \times T4_{ijk} + \pi_{7jk} \cdot RET_{ijk} + e_{ijk}, \text{ com } e_{ijk} \sim NID(0, \sigma^2).$$

Nível 2 – Alunos:

$$\pi_{0jk} = \beta_{00k} + \beta_{01k} \cdot NSE_{jk} + r_{0jk}, \text{ com } r_{0jk} \sim NID(0, \tau_{\pi 00})$$

$$\pi_{1jk} = \beta_{10k}$$

$$\pi_{2jk} = \beta_{20k} + \beta_{21k} \cdot NSE_{ijk}$$

$$\pi_{3jk} = \beta_{30k} + r_{3jk}, \text{ com } r_{3jk} \sim NID(0, \tau_{\pi 30})$$

$$\pi_{4jk} = \beta_{40k}$$

$$\pi_{5jk} = \beta_{50k}$$

$$\pi_{6jk} = \beta_{60k}$$

$$\pi_{7jk} = \beta_{70k};$$

Nível 3 – Escolas:

$$\beta_{00k} = \gamma_{000} + \gamma_{001} \cdot MNSE_k + u_{0jk}, \text{ com } u_{0jk} \sim NID(0, \tau_{\beta 000})$$

$$\beta_{01k} = \gamma_{010}$$

$$\beta_{10k} = \gamma_{100} + \gamma_{101} \cdot MNSE_k$$

$$\beta_{20k} = \gamma_{200}$$

$$\beta_{21k} = \gamma_{210}$$

$$\beta_{30k} = \gamma_{300} + \gamma_{301} \cdot MNSE_k + u_{3jk}, \text{ com } u_{3jk} \sim NID(0, \tau_{\beta 300})$$

$$\beta_{40k} = \gamma_{400}$$

$$\beta_{50k} = \gamma_{500}$$

$$\beta_{60k} = \gamma_{600}$$

$$\beta_{70k} = \gamma_{700}$$

As variáveis e coeficientes de regressão têm as mesmas definições que o Modelo de Referência em Leitura com exceção à variável, PR_FI_{ijk} , que neste modelo é a proficiência final em Matemática. Outra diferença está nas variáveis e coeficientes associadas aos interceptos dos anos 3 e 4. Enquanto que no Modelo de Referência de Leitura o intercepto do ano 4 não foi estatisticamente significativo, neste de Matemática o mesmo ocorreu apenas com o ano 3. Assim, $T4$ é a variável indicadora usada para identificar o quarto ano de escolarização e está associada ao coeficiente π_{2jk} , que é o valor a ser adicionado a π_{0jk} para representar a proficiência final média dos J_k alunos da escola k no quarto ano de escolarização. Desta forma, o intercepto da reta do quarto ano de escolarização é dado por $(\pi_{0jk} + \pi_{2jk})$. Como nesse modelo o intercepto para o terceiro ano de escolarização é o mesmo do primeiro ano.

Os valores estimados para cada um dos coeficientes do Modelo de Referência de Matemática estão indicados no Quadro 96 a seguir e foram retirado do relatório do programa HLM que se encontra no Anexo 12:

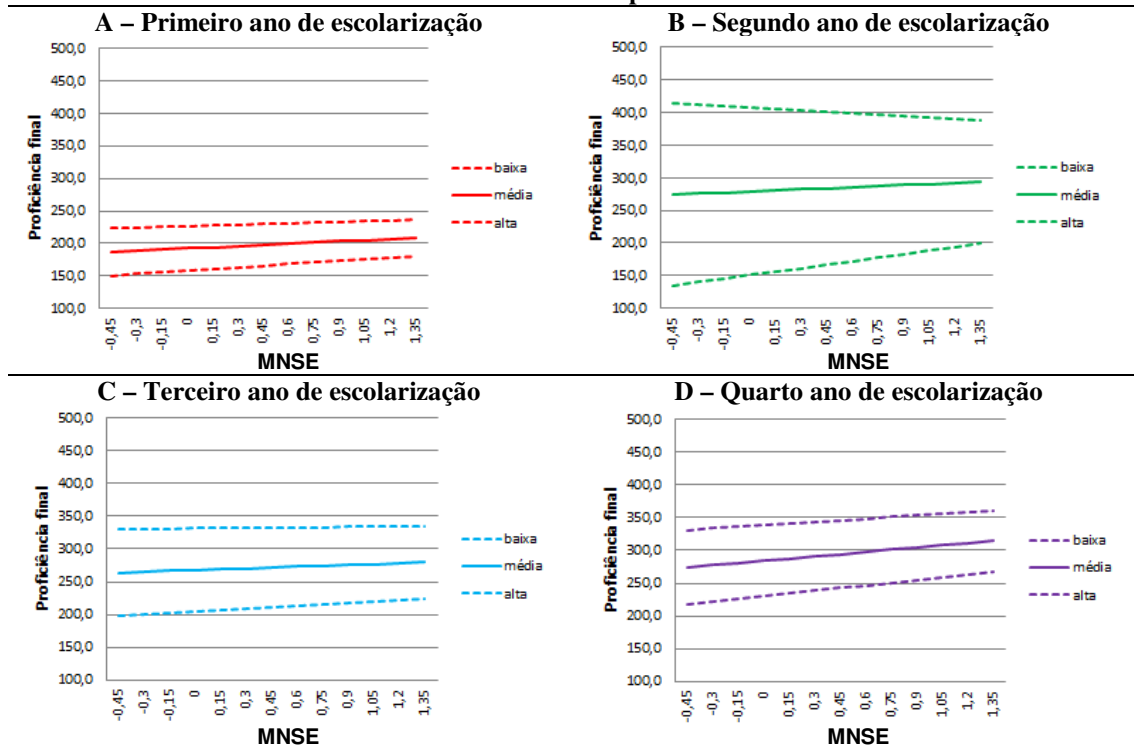
Quadro 96 - Resultados do Modelo de Nível 3 em Matemática

				Estimativa	E. P.	<i>p-value</i>	
Efeitos fixos	Intercepto	π_{0jk}	β_{00k} (Intercepto)	γ_{000}	61,85477	3,94473	<0,001
				γ_{001} (MNSE)	27,92391	3,79015	<0,001
		π_{1jk} (T2)	β_{10k} (Intercepto)	γ_{100}	-67,89886	4,34101	<0,001
				γ_{101} (MNSE)	35,30228	9,01404	<0,001
		π_{2jk} (T4)	β_{20k} (Intercepto)	γ_{200}	22,77123	3,31267	<0,001
				γ_{201} (MNSE)	14,86122	3,01572	<0,001
	Coeficiente de inclinação	π_{3jk} (PR_IN)	β_{30k} (Intercepto)	γ_{300}	0,578427	0,020938	<0,001
				γ_{301} (MNSE)	-0,079935	0,023608	0,001
			β_{31k}	γ_{310} (NSE)	0,014401	0,005035	0,005
		π_{4jk} (PR_INxT2)	β_{40k} (Intercepto)	γ_{400}	0,547525	0,031142	<0,001
				γ_{401} (MNSE)	-0,140216	0,052315	0,008
			π_{5jk} (PR_INxT3)	β_{50k} (Intercepto)	γ_{500}	0,213092	0,009862
	π_{6jk} (PR_INxT4)	β_{60k} (Intercepto)	γ_{600}	0,153125	0,021874	<0,001	
	Inter.	π_{7jk} (RET)	β_{70k} (Intercepto)	γ_{700}	16,41566	3,06374	<0,001
	Efeitos Aleatórios				Componente de variância		
					$\sigma^2 =$	965,28	
$\tau_{\pi00} =$					6,02	>0,500	
$\tau_{\pi30} =$					0,0002	0,001	
$\tau_{\beta000} =$					80,83	<0,001	
$\tau_{\beta300} =$					0,0038	<0,001	
Deviance:		336335	Número de parâmetros estimados:		21		
Teste quiquadrado:		19,27	Graus de liberdade:		8	0,014	

Analisando o *p-value* dos efeitos aleatórios $\tau_{\pi40}$, $\tau_{\beta000}$ e $\tau_{\beta400}$ constata-se que, mesmo com a inclusão dessas variáveis, ainda há variabilidade da proficiência final a ser explicada, indicando que este modelo pode ser usado para a inserção de novas variáveis. No entanto o mesmo não ocorre com $\tau_{\pi00}$, cujo *p-value*, por ser maior que 0,5, indica que não há mais variabilidade a ser explicada com a inserção de outras variáveis.

Os Gráficos 42A, 42B, 42C e 42D mostram a estimativa para a proficiência final em todo o intervalo de valores encontrados para a variável MNSE na amostra dos dados em Matemática.

Gráfico 41 - Influência do MNSE sobre a proficiência final em Matemática



Para representar a influência do MNSE sobre os alunos com proficiência inicial baixa e alta, foram considerados os valores acrescendo e subtraindo 50 unidades à proficiência inicial média de cada um dos anos de escolarização.

Percebe-se que somente para os alunos com proficiência inicial alta, o aumento do MNSE influencia negativamente a proficiência final em matemática.

E comparando a influência do MNSE sobre a proficiência final em Leitura (Gráfico 40) e de matemática (Gráfico 42) pode-se constatar que quanto maior for o MNSE das escolas, mais as proficiências finais dos alunos com alta proficiência inicial e baixa proficiência única se aproxima.

Mas essa ideia de que essa aproximação, sugerindo que as escolas se tornem mais equânimes, seja um resultado positivo é um equívoco pois essa aproximação se deve a uma influência mais negativa para os alunos com proficiência inicial alta, como mostra a Tabela 27.

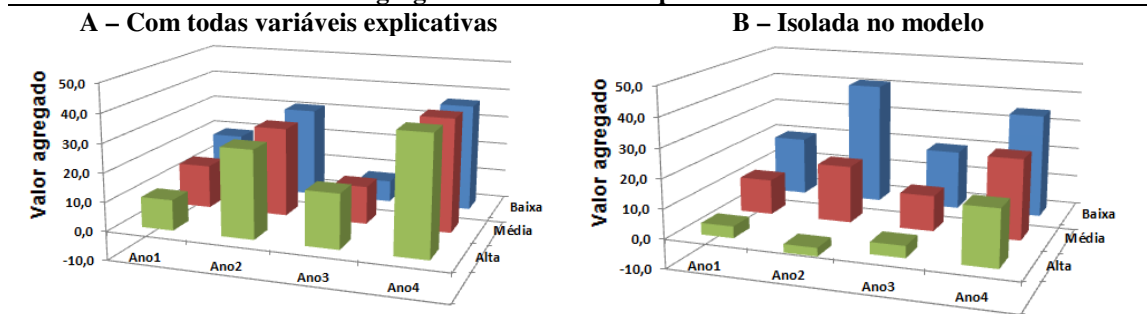
Tabela 27 - Valor agregado em Matemática em função do MNSE e da proficiência inicial

MNSE	Proficiência inicial											
	Ano 1			Ano 2			Ano 3			Ano 4		
	baixa	média	alta	baixa	média	alta	baixa	média	alta	baixa	média	alta
-0,45	12,7	-10,2	-33,1	9,9	35,6	61,2	38,3	24,4	10,5	41,0	23,6	6,3
0,45	27,9	0,8	-26,4	42,2	45,2	48,3	52,0	32,3	12,6	66,8	44,1	21,4
1,35	43,1	11,7	-19,6	74,4	54,9	35,4	65,7	40,2	14,7	92,7	64,6	36,5

Nessa tabela fica evidente que, nos diferentes anos de escolarização, o valor agregado para os alunos com proficiência inicial baixa em escolas com MNSE alto têm um valor agregado maior do que aqueles com proficiência inicial alta e estudam em escolas com MNSE baixo, sendo que no primeiro ano de escolarização, para esses últimos, o valor agregado estimado é negativo.

Assim como ocorreu no estudo em Leitura, observa-se a diferença do comportamento do valor agregado para cada unidade de MNSE entre o modelo no qual todas as variáveis estão inseridas e neste que a analisa isoladamente, como mostra os Gráficos 43A e 43B.

Gráfico 42 - Valor agregado em Matemática por unidade da variável MNSE



Com isso se evidencia mais uma vez que o efeito de uma variável depende diretamente do modelo a que ela é submetida.

6.3) Análise das variáveis

Para essa análise, cada uma das variáveis do Nível 2 e do Nível 3 foram inseridas nos modelos para identificar seus efeitos no intercepto e na inclinação da reta de cada um dos anos de escolarização.

Após essa inclusão, buscando o melhor ajuste do modelo, as variáveis foram sendo excluídas uma a uma, segundo o valor do *p-value* apresentado, ou seja, mais uma vez a metodológica *backward* foi aplicada para o ajuste do um modelo linear hierárquico.

Em seguida, com as estimativas dos coeficientes o modelo ajustado, é estimado o valor que cada unidade da variável analisada agrega à proficiência final, considerando também a proficiência inicial, o nível socioeconômico individual do aluno e o nível socioeconômico médio da escola.

A análise desse valor agregado é feita para três valores distintos de proficiência inicial em cada ano, seguindo o mesmo critério do capítulo anterior:

- proficiência inicial baixa, cujo valor corresponde ao valor de um desvio-padrão abaixo da média do respectivo ano ($\bar{x} - \sigma$);
- proficiência inicial média, que corresponde à media das proficiência iniciais (\bar{x}), e;
- proficiência inicial alta, correspondente ao valor de um desvio-padrão acima da média ($\bar{x} + \sigma$).

Para a apresentação dessa análise foram criados três agrupamentos de variáveis:

- a) Variáveis que tiveram os mesmos efeitos em leitura e em matemática;
- b) Variáveis que tiveram efeitos específicos em leitura;
- c) Variáveis que tiveram efeitos específicos em matemática.

6.3.1) Variáveis que tiveram os mesmos efeitos em leitura e em matemática

Das 67 variáveis analisadas, nove não foram estatisticamente significantes e sete tiveram efeito sobre a proficiência final com o mesmo comportamento.

Dessas sete variáveis quatro tiveram resultados são consistentes em todos os anos de escolarização e são coerentes com aquilo que é esperado. São elas as que controlam:

- Se o aluno fez a pré-escola (FA04);
- A Interrupção das aulas por indisciplina (FEP50);
- O uso da metodologia de resolução de problemas (FEP43);
- A proporção de Diretores que possui outra atividade que contribui para a renda pessoal (FED02).

As outras três, cujos resultados são consistentes, mas não são coerentes com o esperado, são as variáveis que controlam.

- Espaço sala de leitura (FEE04);
- Estímulo interno para lição de casa (FA11);
- Meus professores pensam que sou um aluno lento (FA12).

Neste terceiro agrupamento estão as variáveis que não foram estatisticamente significantes em Leitura e em Matemática:

- Existência de equipamentos para fotocópia e escaneamento (FEE09);
- Interferência do perfil do professor (FED13);
- Interferência da direção na orientação pedagógica (FED15);
- Organização de eventos (FED17);
- Carga horária de trabalho do(a) diretor(a) na escola (FED23);
- Proposição de leitura (FEP37);
- Proposição de redação (FEP38);
- Grau de importância inerente à indicação externa do livro (FEP46);
- Proporção de professores do sexo feminino (FEP55).

6.3.2) Variáveis que tiveram efeitos específicos em Leitura

A análise das variáveis com efeitos específicos sobre a proficiência final em Leitura resultou em seis agrupamentos específicos, apresentados a seguir:

a) Efeitos consistentes em todos os anos de acordo com o esperado:

- Laboratório de informática (FEE06);
- Disponibilidade de computador e internet para professor (FEE08);
- Tempo que o(a) diretor (a) trabalha em anos na educação (FED25);
- Grau de expectativas negativa do aprendizado (FEP34);
- Leitura individual e produção de texto individual (FEP42);
- Ensino com ênfase em memorização (FEP44);
- Tempo usado em disciplinas de formação complementar (FEP49);
- Interrupção das aulas por necessidade administrativa (FEP51);

- Estímulo extrínseco para estudo (FA07);
- b) Efeitos consistentes em parte dos anos e de acordo com o esperado:
- Boa relação entre professores (FEP30);
 - Tempo usado nas disciplinas relacionadas com as proficiências (FEP47);
 - Tempo usado nas disciplinas não relacionadas com as proficiências (FEP53);
 - Uso de laboratório e multimídia (FEP48);
 - Tempo de exercício médio do professor na ocupação (FEP58);
- c) Efeitos consistentes em todos os anos de acordo, mas com resultado contrário ao esperado:
- Conservação e limpeza área de alimentação (FEE03);
 - Existência de terrário e aquário (FEE11);
 - Boa interação entre a equipe (FEP28);
 - Percepção do aluno quanto ao incentivo do professor (FA09);
- d) Efeitos consistentes em parte dos anos de acordo, mas com resultado contrário ao esperado:
- Condições da Biblioteca (FEE05);
 - Existência de material concreto para matemática (FEE10);
 - Relação com o diretor (FEP27);
 - Pouco trabalho coletivo (FEP29);
- e) Efeitos instáveis ou com valores desprezíveis:
- Conservação e limpeza supraestrutura (FEE02);
 - Frequência da presença na comunidade (FED19);
 - Unidade de propósito entre os professores (FEP31);
 - Acesso a atividades culturais (FEP36);
 - Proposição de leitura individual é precedida da explicação do professor (FEP40);

- Adequação do livro ao ensino (FEP45);
- Uso de materiais concretos de matemática e geografia (FEP54);
- Renda Familiar média dos professores, em salários mínimos (FEP57);
- Número médio de pessoas da família que mora com o professor (FEP59);
- Estímulo intrínseco para o estudo (FA08);
- Gosto pela escola (FA10);
- Estímulo interno para lição de casa (FA11);
- Algumas vezes eu sinto que não sou bom o bastante (FA14).

f) não significativa:

- Sala de artes e laboratório de ciências (FEE07);
- Interferência por intimidação, violência, discriminação e drogas (FED12);
- Interferência do perfil do aluno, família (FED14);
- Presença física do diretor (FED16);
- Frequência de encontros com a comunidade (FED18);
- Proporção de diretores que possuem Ens. Sup. em Pedagogia (FED20);
- Proporção de diretores que trabalham em apenas uma escola (FED22);
- Tempo que o diretor trabalha na escola ininterruptamente na função de direção (FED24);
- Idade do(a) diretor(a) em anos (FED26);
- Violência no ambiente escolar (FEP32);
- Frequência do dever de casa (FEP33);
- Grau de expectativas positiva do aprendizado (FEP35);
- Proposição de cópia ou caligrafia (FEP39);
- Leitura em prática coletiva (FEP41);
- Uso de recursos audiovisuais (FEP52);
- Idade média dos professores (FEP56);

6.3.3) Variáveis que tiveram efeitos específicos em matemática

A análise dos efeitos específicos sobre a proficiência final em Matemática resultou nos mesmos seis agrupamentos de variáveis encontrado especificados anteriormente, no entanto, as variáveis que os compõem são diferentes.

Esses sub-agrupamentos são apresentados a seguir:

a) Efeitos consistentes em todos os anos de acordo com o esperado:

- Condições da Biblioteca (FEE05);
- Proporção de diretores que trabalham em apenas uma escola (FED22);
- Frequência do dever de casa (FEP33);

b) Efeitos consistentes em parte dos anos e de acordo com o esperado:

- Frequência da presença na comunidade (FED19);
- Grau de expectativas positiva do aprendizado (FEP35);
- Interrupção das aulas por necessidade administrativa (FEP51);

c) Efeitos consistentes em todos os anos de acordo, mas com resultado contrário ao esperado:

- Tempo que o(a) diretor (a) trabalha em anos na educação (FED25);
- Pouco trabalho coletivo (FEP29);
- Uso de recursos audiovisuais (FEP52);
- Algumas vezes eu sinto que não sou bom o bastante (FA14);

d) Efeitos consistentes em parte dos anos de acordo, mas com resultado contrário ao esperado:

- Conservação e limpeza supraestrutura (FEE02);
- Idade do(a) diretor(a) em anos (FED26);
- Proposição de leitura individual é precedida da explicação do professor (FEP40);

e) Efeitos instáveis ou com valores desprezíveis:

- Conservação e limpeza área de alimentação (FEE03);
- Laboratório de informática (FEE06);
- Sala de artes e laboratório de ciências (FEE07);
- Existência de terrário e aquário (FEE11);
- Interferência por intimidação, violência, discriminação e drogas (FED12);
- Interferência do perfil do aluno, família (FED14);
- Presença física do diretor (FED16);
- Frequência de encontros com a comunidade (FED18);
- Proporção de diretores que possuem Ens. Sup. em Pedagogia (FED20);
- Tempo que o diretor trabalha na escola ininterruptamente na função de direção (FED24);
- Grau de expectativas negativa do aprendizado (FEP34);
- Proposição de cópia ou caligrafia (FEP39);
- Leitura em prática coletiva (FEP41);
- Tempo usado nas disciplinas não relacionadas com as proficiências (FEP48);
- Uso de laboratório e multimídia (FEP53);
- Idade média dos professores (FEP56);
- Gosto pela escola (FA10);
- Estímulo interno para lição de casa (FA11);

f) Não significante:

- Disponibilidade de computador e internet para professor (FEE08);
- Existência de material concreto para matemática (FEE10);
- Relação com o diretor (FEP27);
- Boa interação entre a equipe (FEP28);
- Boa relação entre professores (FEP30);
- Unidade de propósito entre os professores (FEP31);
- Violência no ambiente escolar (FEP32);
- Acesso a atividades culturais (FEP36);
- Leitura individual e produção de texto individual (FEP42);

- Adequação do livro ao ensino (FEP45);
- Ensino com ênfase em memorização (FEP44);
- Tempo usado nas disciplinas relacionadas com as proficiências (FEP47);
- Tempo usado em disciplinas de formação complementar (FEP49);
- Uso de materiais concretos de matemática e geografia (FEP54);
- Renda Familiar média dos professores, em salários mínimos (FEP57);
- Tempo de exercício médio do professor na ocupação (FEP58);
- Número médio de pessoas da família que mora com o professor (FEP59);
- Estímulo extrínseco para estudo (FA07);
- Estímulo intrínseco para o estudo (FA08);
- Percepção do aluno quanto ao incentivo do professor (FA09).

6.3.4) Análise de alguns dos resultados encontrados

A análise desses resultados é extensa, não só pela quantidade de 67 variáveis que compõem essa parte do estudo e que se repetem em duas áreas do conhecimento, mas também, senão principalmente pela necessária retomada do processo que cada uma delas passou até culminar no resultado final.

Essa retomada consiste na verificação dos itens do questionário que compõe cada um dos fatores dos constructos descritos no Apêndice 01 e da análise descritiva da variável apresentada no Apêndice 02.

A análise a seguir está dividida em três partes. Na primeira, são analisadas as variáveis estatisticamente significantes e que cujos resultados vão de acordo com o esperado. Na segunda estão as variáveis também estatisticamente significantes, mas com efeitos contrários ao desejado e na última estão as variáveis que neste estudo com o Modelo de Referência proposto se mostraram não significantes.

6.3.4.1) Comentários acerca das variáveis estatisticamente significantes e com efeito de acordo com o esperado.

No Quadro 97 são apresentados as variáveis cujos efeitos foram consistentes em todos os anos e ainda corresponderam com é aquilo que é esperado⁵².

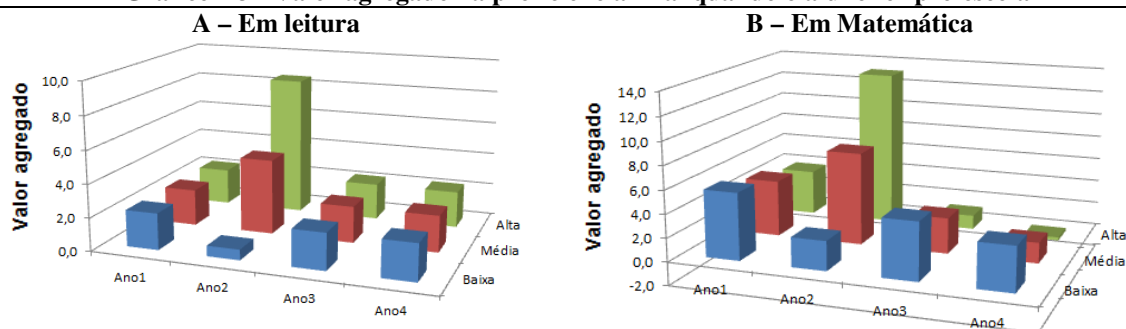
Quadro 97 - Variáveis com efeitos consistentes em todos os anos de acordo com o esperado

Efeito em:	Descrição
Leitura e Matemática	Indica se o aluno fez a pré-escola (FA04)
	A Interrupção das aulas por indisciplina (FEP50)
	O uso da metodologia de resolução de problemas (FEP43)
	A proporção de Diretores que possui outra atividade que contribui para a renda pessoal (FED02).
Leitura	Laboratório de informática (FEE06)
	Disponibilidade de computador e internet para professor (FEE08)
	Tempo que o(a) diretor (a) trabalha em anos na educação (FED25)
	Grau de expectativas negativa do aprendizado (FEP34)
	Leitura individual e produção de texto individual (FEP42)
	Ensino com ênfase em memorização (FEP44)
	Tempo usado em disciplinas de formação complementar (FEP49)
	Interrupção das aulas por necessidade administrativa (FEP51)
	Estímulo extrínseco para estudo (FA07)
Matemática	Condições da Biblioteca (FEE05)
	Proporção de diretores que trabalham em apenas uma escola (FED22)
	Frequência do dever de casa (FEP33)

Os efeitos encontrados na variável FA04 evidenciam que o fato do aluno ter frequentado a pré-escola antes de ingressar no ensino fundamental influencia positivamente na proficiência final dos alunos em Leitura em Matemática, indicando que é acertada a política educacional que ampliou o Ensino Fundamental no Brasil para 9 anos, tornando obrigatório a matrícula aos seis anos. Em Leitura o valor agregado é de 2,2 unidades para as diferentes proficiência iniciais, exceto para o segundo ano em que os alunos de proficiência inicial alta igual é de aproximadamente 8, enquanto que para os de proficiência inicial baixa é de aproximadamente nula. Essa excepcionalidade para o segundo ano de escolarização também é percebida em Matemática, pois em todos os anos os alunos com proficiência inicial baixa tem um valor agregado maior que os de alta, quando cursaram a pré-escola. Os Gráficos 44A e 44B apresentam o valor agregado na proficiência final referente a essa variável.

⁵² Caso haja o interesse de analisar outras variáveis basta a consulta aos Anexos 03 e 04.

Gráfico 43 - Valor agregado na proficiência final quando o aluno fez pré-escola



A interrupção das aulas por indisciplina, controlada pela variável FEP50, é resultante dos seguintes itens de questionário do professor⁵³, coletados em 2005 e 2006:

Quantas vezes em média a aula é interrompida por bagunça dos alunos?

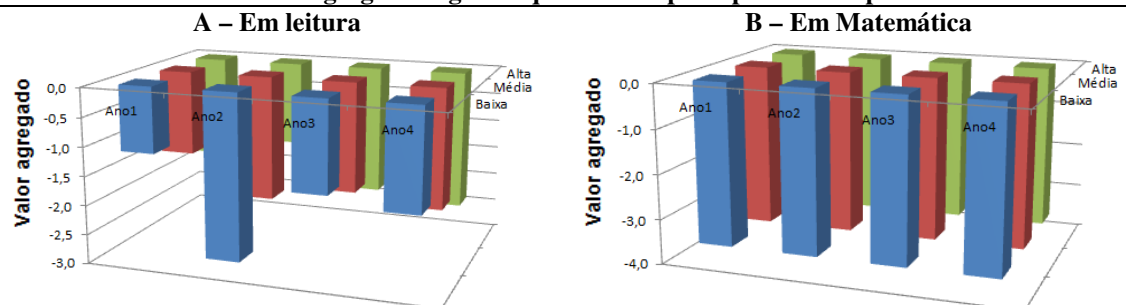
Quantas vezes em média a aula é interrompida por estudantes atrasados?

Quantas vezes em média a aula é interrompida por barulho no corredor?

Para cada um desses itens, o professor respondente, tinha como opções de resposta as seguintes opções: “Nenhuma” (0), “Uma vez” (1), “Duas vezes” (2), “três a quatro vezes” (3), “cinco a seis vezes” (4), e “sete ou mais vezes” (5). Ou seja, esta variável tem como intervalo 5 unidades.

E, como esperado, esse aspecto tem efeito negativo na proficiência final dos alunos, cujo valor agregado para cada unidade⁵⁴ dessa variável está apresentado nos Gráficos 44A e 44B.

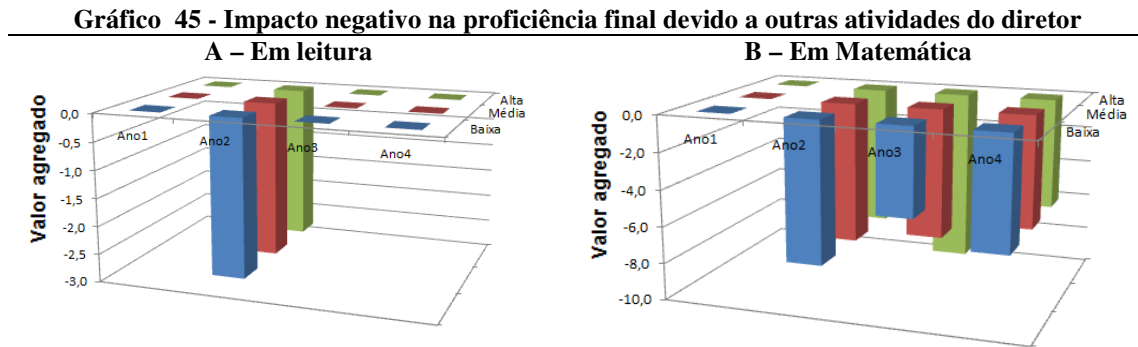
Gráfico 44 - Valor agregado negativo que a indisciplina provoca na proficiência final



53 Os dados da AFE desta variável estão apresentados nas Tabelas 58 e 59 do Apêndice 01.

54 A análise descritiva desta variável (PEP50), que se encontra no Apêndice 02 indica que a média é 1,7, com desvio-padrão de 0,7, implicando que, em 68% das escolas essas interrupções ocorrem em torno um ou dia duas vezes, fazendo com que os o efeito desta variável apresentado no Gráfico 45 seja dobrado.

Este estudo mostra que também tem efeito negativo quando o diretor da escola tem outra atividade para complementar sua renda. Esta variável foi retirada diretamente do questionário, não passando, portanto, pela Análise Fatorial Exploratória.

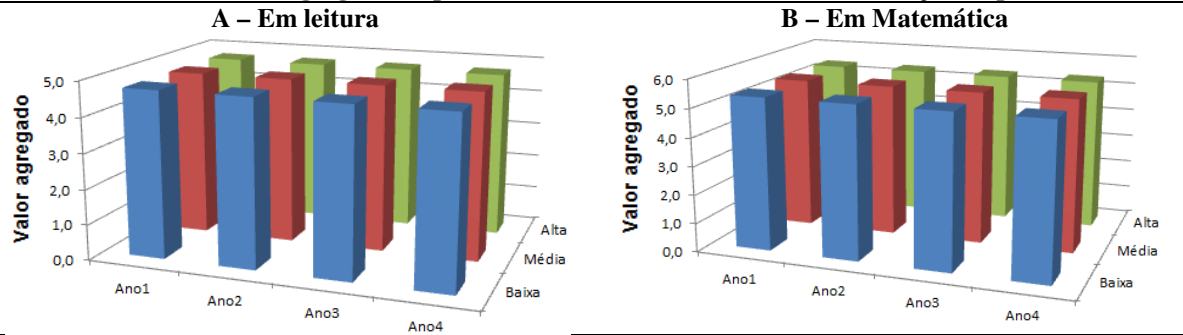


Resultado interessante, por ter aparecido pouco em estudos e eficácia escolar foi observado com a variável FEP43, que indica a ênfase na metodologia de resolução de problemas. Essa variável é resultante da AFE que reduziu a essa única variável, sete questões coletadas através do questionário dos professores nos anos de 2006 e 2007, ou seja, nas Ondas de aplicação 3 e 4, aplicadas nos finais dos segundo e terceiro anos de escolarização.

As alternativas de respostas nos itens do questionário e os valores correspondentes nos arquivos são: “nunca” (0), “raramente” (1), “algumas vezes no bimestre” (2), “cerca de uma vez por semana” (3) e “várias vezes por semana” (4) para a periodicidade com que os professores fazem propostas dentro desse enfoque aos seus alunos, ou seja é uma variável que tem como intervalo de respostas valores que variam de 0 a 4. O valor médio para essa variável é 3 com desvio padrão de 0,3, indicando que, em geral, os professores dão ênfase nesse trabalho uma vez por semana.

Os Gráficos 46A e 46B mostram o valor agregado que cada unidade dessa variável à proficiência final.

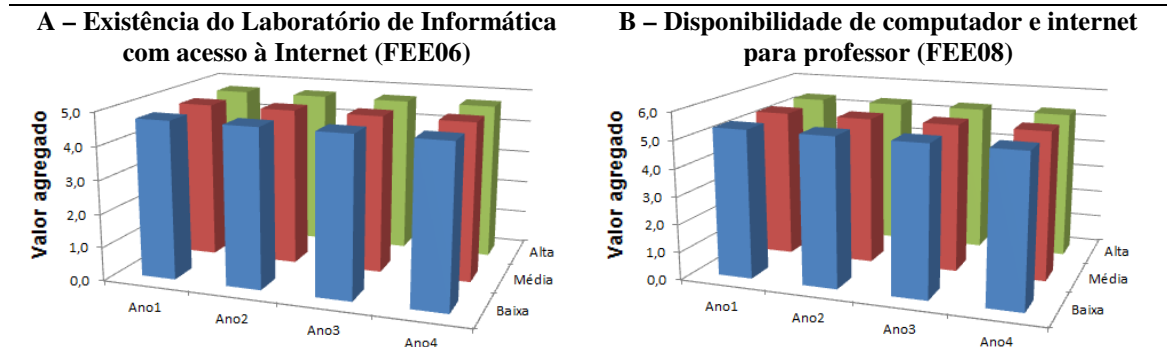
Gráfico 46 - Valor agregado na proficiência final devido à ênfase na resolução de problemas



É importante destacar o efeito das variáveis FEP43 (ênfase na resolução de problemas) e FEP50 (interrupção das aulas por indisciplina) têm os mesmos efeitos em Leitura e Matemática, de forma equitativa (pois são as mesmas para alunos com diferentes proficiências iniciais) e de maneira relevante, dado que seus valores relativamente provocam efeitos maiores, em valores absolutos, que o Nível socioeconômico após a entrada do aluno na escola.

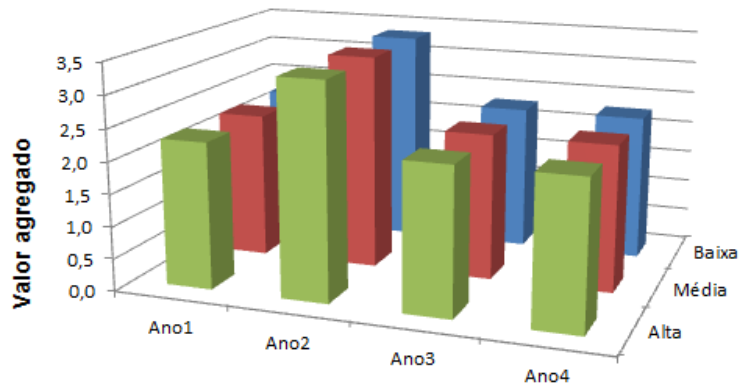
Especificamente em Leitura, a infraestrutura aparece como relevante em duas variáveis, cujos valores agregados estão apresentados nos Gráficos 47A e 47B.

Gráfico 47 - Valor agregado na proficiência final em Leitura ligada à infraestrutura de Laboratório de Informática e acesso à Internet



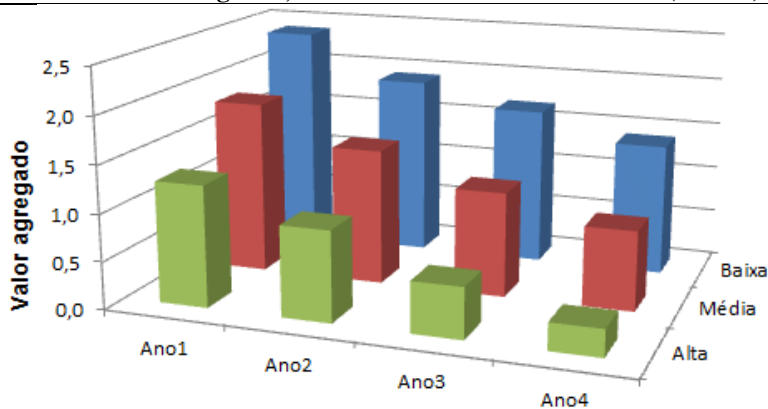
Ligada à gestão da escola, a variável FED25, que controla o número de anos que o diretor trabalha especificamente na Educação, mostra que essa variável é relevante, como mostra o Gráfico 48.

Gráfico 48 - Valor agregado na proficiência final em Leitura devido à experiência do Diretor em Educação (FED25)



Outro resultado importante, que deve ser ressaltado, é o encontrado com a variável FEP49, que controla o tempo usado em disciplinas complementares, não diretamente ligadas às que cujas proficiências são medidas. Esta variável é composta por dois itens do questionário respondido pelos professores em 2006: “A porcentagem do tempo de aula de história e geografia oferecida para os alunos desta turma” e “A porcentagem do tempo de aula de ciências naturais oferecida para os alunos desta turma”. As alternativas de respostas para esses itens do questionário são: “0%” (0), “15%” (1), “30%” (2), “45%” (3), “60%” (4), e “75%” (5). Como a média nesta variável é de 3,7 com desvio-padrão de 0,8, significa que aproximadamente 68% das escolas assumem valores entre 3 e 5, o que destaca a relevância dessa variável na proficiência final, uma vez que o valor agregado no Gráfico 49 refere-se ao valor agregado de uma unidade, nessa variável.

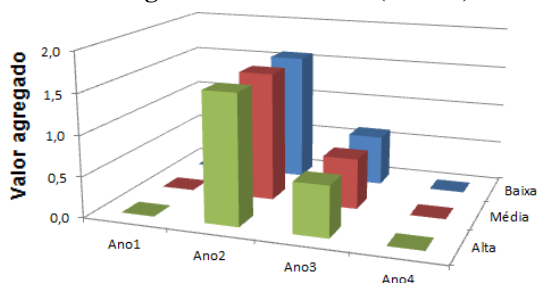
Gráfico 49 - Valor agregado na proficiência final em Leitura devido ao Ensino de Geografia, História e Ciências da Natureza (FEP49)



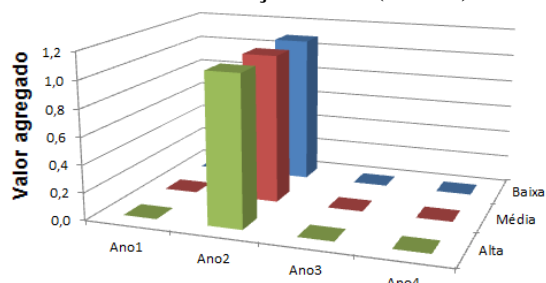
Essa variável foi mais relevante que outras duas que controlam o tempo usado nas disciplinas relacionadas com as proficiências medidas (FEP47) e o tempo usado nas disciplinas não relacionadas com as proficiências, Artes e Educação Física (FEP48). Os efeitos dessas duas variáveis são apresentados nos Gráficos 50A e 50B.

Gráfico 50 - Valor agregado na proficiência final em Leitura ligado ao tempo usado em outras disciplinas

A – Relacionadas com as proficiências – Língua Portuguesa e Matemática (FEP47)



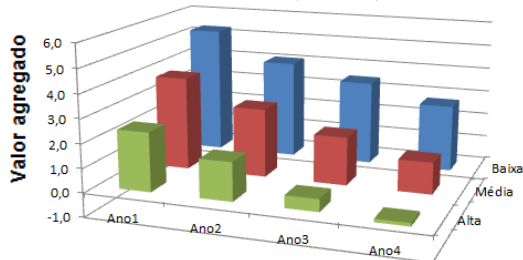
B – Não relacionadas com as proficiências – Artes e Educação Física (FEP48)



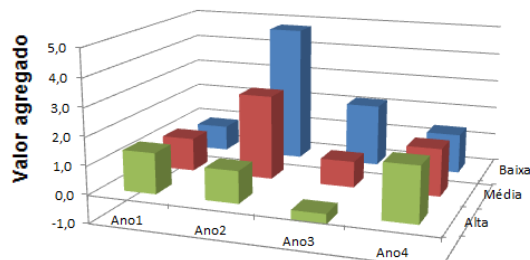
Na perspectiva da metodologia de ensino, duas variáveis apareceram como relevantes como mostram os Gráficos 51A e 51B.

Gráfico 51 - Valor agregado na proficiência final em Leitura ligada à metodologia de ensino

A – Leitura individual e produção de texto individual (FEP42)



B – Ensino com ênfase em memorização (FEP44)



A variável FEP42 resulta da AFE de três itens do questionário do professor, respondidos na quarta Onda de aplicação dos testes, em 2007, que são:

- No ensino de língua portuguesa dou prioridade para: Alunos leem um texto e sublinham com cores diferentes substantivos e adjetivos. A atividade é individual;
- No ensino de língua portuguesa dou prioridade para: Após explicação da professora, alunos leem texto procurando identificar e sublinhar palavras com ditongos, e;

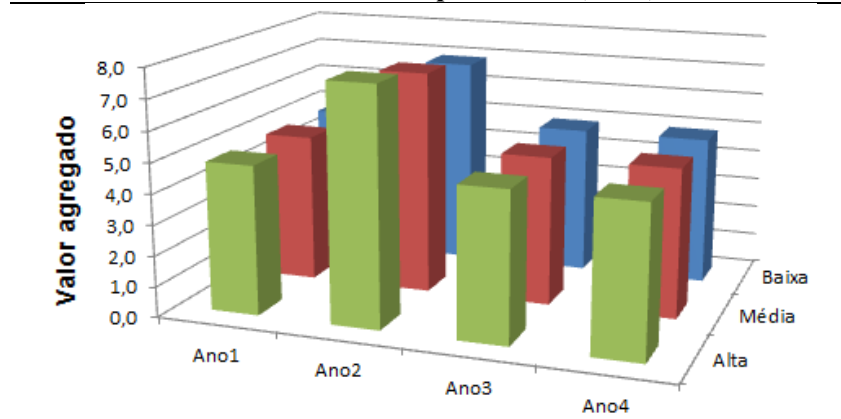
- No ensino de língua portuguesa dou prioridade para: Após explicação da professora, alunos leem texto procurando identificar e diferenciar palavras com ditongos crescentes e decrescentes.

Já a variável FEP44, refere-se a metodologia mais ligada à matemática e que refletiu nos resultados em leitura. Certamente é uma variável que reflete um estilo específico de conduzir as aulas do professor. Essa variável é resultado da AFE dos seguintes itens de questionário do professor respondido em 2006 e 2007:

- lidaram com situações para memorizar conceitos e regras?
- decoraram regras e aplicaram-nas em situações-problema?

Todas as variáveis apresentadas até agora estão associadas às características da escola e a única que apareceu sendo ligada ao perfil do aluno é a variável FA07, cujos efeitos são apresentados no Gráfico 52.

Gráfico 52 - Valor agregado na proficiência final em Leitura devido ao Estímulo extrínseco para estudo (FA07)



Essa variável é composta pelos seguintes itens de questionário respondido pelo aluno, e que mostram o quanto os estímulos são extrínsecos a ele:

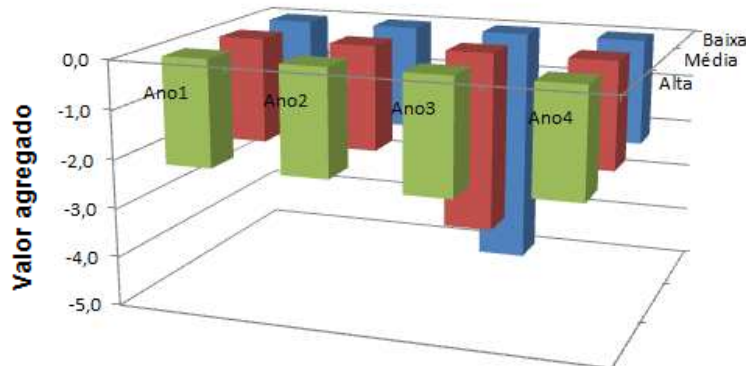
- Eu estudo porque meus pais mandam;
- Eu faço o dever de casa por obrigação;
- Eu estudo porque meus pais prometem me dar presentes, se as minhas notas forem boas;
- Eu estudo porque minha professora acha importante;
- Eu só estudo para não me sair mal na escola;

- Eu estudo para que meus pais me deixem brincar com os meus amigos;
- ou fazer as coisas que eu gosto;
- Eu só estudo para agradar meus professores;
- Eu só faço meu dever de casa (lição) porque meus pais acham importante.

Com efeitos relevantes e que influenciam negativamente na proficiência final em leitura aparecem duas variáveis: a FEP34, que controla o grau de expectativas negativa do aprendizado e a FEP51, que controla a Interrupção das aulas por necessidade administrativa.

O Gráfico 53 mostra o efeito negativo que tem quando o professor tem expectativas negativas em relação à aprendizagem de seus alunos.

Gráfico 53 - Valor agregado na proficiência final em Leitura devido às expectativas negativas quanto à aprendizagem do aluno (FEP34)



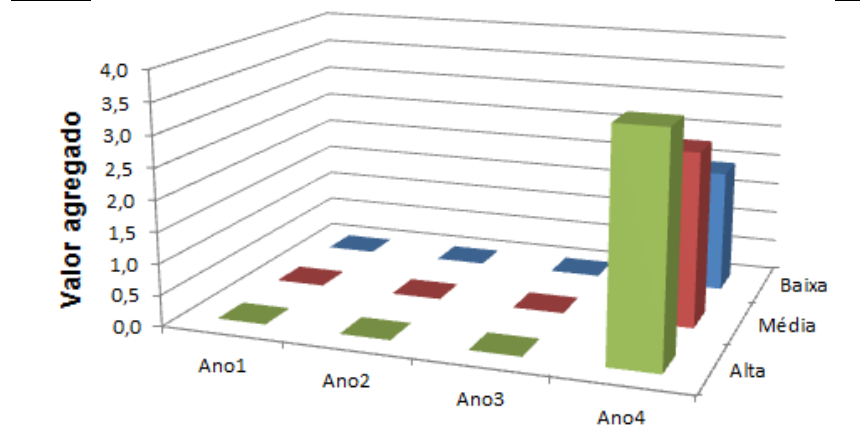
Essa variável é medida a partir de quatro itens de questionário que buscam captar a percepção que o professor tem acerca de suas possibilidades de o aluno aprender. Esses itens são:

- Com todos os atrativos que as crianças podem ter acesso hoje em dia, é muito difícil para a escola fazer seu trabalho;
- Diante das dificuldades desta escola, um pequeno aprendizado dos alunos já é um bom resultado;
- Para que os alunos desta escola pudessem realmente aprender seria necessário que a educação fosse levada mais a sério neste país;
- Com as famílias que os alunos dessa escola têm, o aprendizado fica muito comprometido.

Essa variável mostra que, quanto mais o professor tem a percepção de que os condicionantes externos às suas possibilidades de trabalho são dificultadores da aprendizagem do aluno, menor é a proficiência inicial. Mas essa variável pode também refletir as condições inadequadas de trabalho desse professor, o que também interfere negativamente na proficiência final, afinal, só um professor que se encontra em condições precárias de trabalho tem essa percepção.

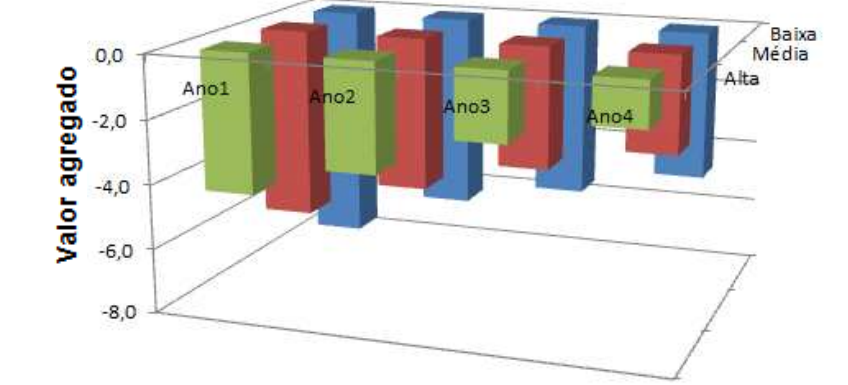
A constatação do efeito dessa variável é confirmada, mesmo de forma pontual, pelo resultado apresentado em Matemática, pela variável FEP35, que controla em um sentido contrário, o grau de expectativas positiva do aprendizado, como mostra o Gráfico 54. Essa variável é composta por um único item do questionário que verificou a opinião se “A maioria dos professores mantém altas expectativas sobre o aprendizado dos alunos”.

Gráfico 54 - Valor agregado na proficiência final em Matemática devido às expectativas positivas quanto à aprendizagem do aluno (FEP35)



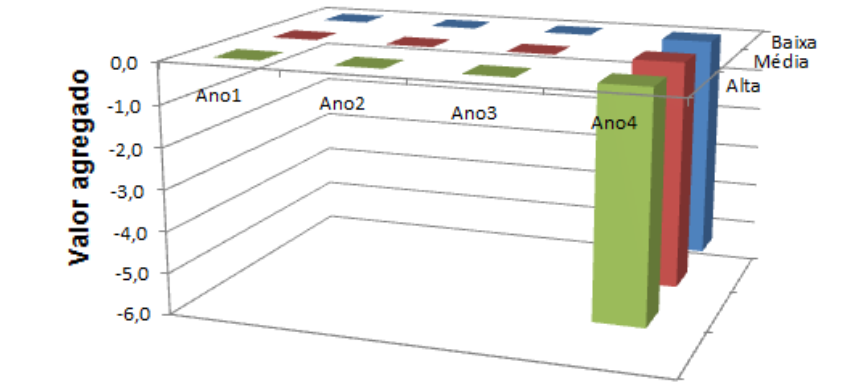
A outra variável que influencia negativamente a proficiência final em leitura aparecem novamente as interrupções das aulas, segundo a percepção dos professores, controlada pela variável FEP51, porém agora provocadas por necessidades administrativas, tais como anúncios ou comunicações da direção, coordenação e/ou secretaria. Esses efeitos negativos são apresentados no Gráfico 55.

Gráfico 55 - Valor agregado na proficiência final em Leitura devido às interrupções das aulas por necessidade administrativa (FEP51)



A constatação do efeito dessa variável é confirmada, mesmo de forma pontual, pelo resultado apresentado em Matemática, como mostra o Gráfico 56.

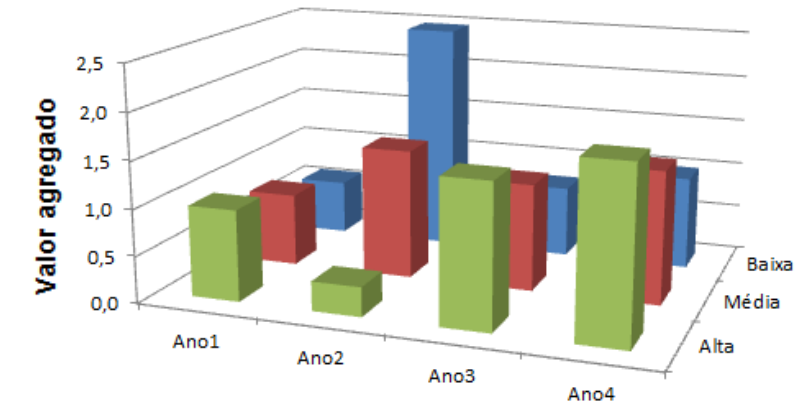
Gráfico 56 - Valor agregado na proficiência final em Matemática devido às interrupções das aulas por necessidade administrativa (FEP51)



Na Matemática, três variáveis tiveram seus efeitos agregando positivamente na proficiência final.

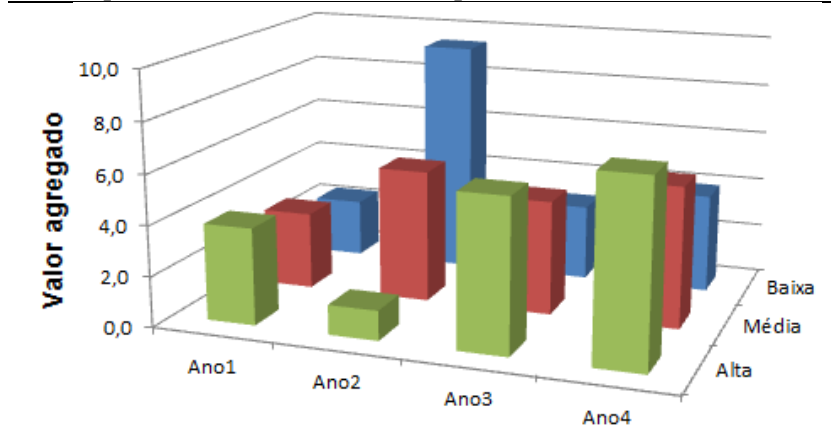
A primeira delas, ligada à infraestrutura, controla a existência da Biblioteca e, que a princípio, parece não ter ligação direta com a aprendizagem matemática, mas seus efeitos são relevantes como mostra o Gráfico 57.

Gráfico 57 - Valor agregado na proficiência final em Matemática devido às condições da Biblioteca (FEE05)



Assim como ocorreu em Leitura, uma das características da gestão, teve resultado positivo sobre a proficiência em matemática, porém com a variável FED22, que controla a proporção de diretores que trabalham em uma mesma escola. Essa variável foi retirada do questionário de diretores nos diferentes anos, e verificou em quantas escolas o diretor trabalhava. Os efeitos quando se constata que diretor trabalha em apenas uma escola está apresentado no Gráfico 58.

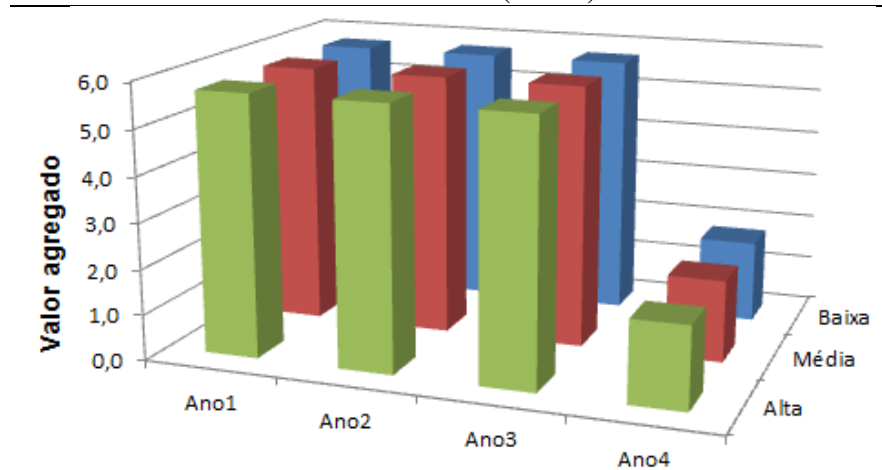
Gráfico 58 - Valor agregado na proficiência final em Matemática quando o diretor trabalha em apenas uma escola (FED22)



No aspecto metodológico, a variável FEP33, composto por dois itens o questionário que verificava a frequência com o dever de casa era passado pelos alunos e também pela proporção dos alunos que faziam o dever ("Com que frequência o professor passa dever de

casa?” e “Quantos alunos do professor fazem dever de casa?”), mostrou ter efeitos positivos na proficiência final de matemática, como mostra o Gráfico 59.

Gráfico 59 - Valor agregado na proficiência final em Matemática devido ao dever de casa (FEP33)



Com a descrição dessas variáveis alguns dos fatores que interferem nas proficiências finais dos alunos em Leitura e Matemática estão identificados, que, em certa medida, leva o cumprimento dos objetivos desta pesquisa.

No entanto, é também importante a análise não só das variáveis cujos efeitos foram estatisticamente significantes, mas com efeito contrário ao esperado e também daquelas que, apesar da literatura apontar sua relevância nos estudos de eficácia escolar, neste estudo com os Modelos de Referência não foram estatisticamente significantes.

Esses dois agrupamentos de variáveis são analisados a seguir.

6.3.4.2) Comentários acerca das variáveis estatisticamente significantes e com efeito contrário ao esperado.

No Quadro 98 a seguir são apresentados os comentários de algumas das variáveis⁵⁵.

⁵⁵ Caso haja o interesse de analisar outras variáveis basta a consulta aos Anexos 03 e 04.

Quadro 98 - Variáveis com efeitos consistentes em todos os anos, mas contrários ao esperado

Efeito em:	Descrição
Leitura e Matemática	Espaço sala de leitura (FEE04)
	Estímulo interno para lição de casa (FA11)
	Meus professores pensam que sou um aluno lento (FA12)
Leitura	Conservação e limpeza área de alimentação (FEE03)
	Existência de terrário e aquário (FEE11)
	Boa interação entre a equipe (FEP28)
	Percepção do aluno quanto ao incentivo do professor (FA09)
Matemática	Tempo que o(a) diretor (a) trabalha em anos na educação (FED25)
	Pouco trabalho coletivo (FEP29)
	Uso de recursos audiovisuais (FEP52)
	Algumas vezes eu sinto que não sou bom o bastante (FA14)

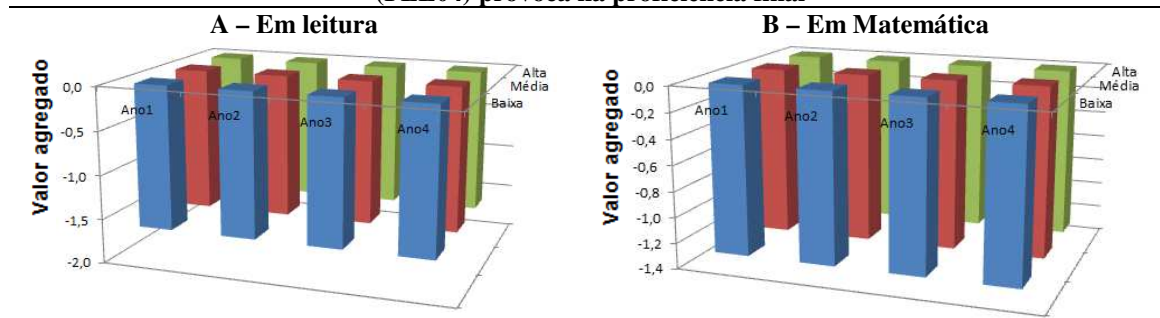
É recorrente esperar que qualquer espaço disponível e organizado disponível nas escolas colabore para a melhor no desempenho. Em recente estudo, Bonamino e Lima (2013), não fazendo distinção entre biblioteca e espaço de leitura, mostram que concordam com essa expectativa e citam que existem “várias pesquisas brasileiras alinhadas ao campo de estudo sobre escolas eficazes apontam que os espaços escolares, os equipamentos da escola e espaços como a biblioteca e/ou sala de leitura contribuem para o melhor desempenho dos alunos” (p. 111) e fazem referência a um estudo realizado pelo INEP (2010), com base no Censo Escolar, aponta que “de acordo com os resultados, biblioteca/sala de leitura, quadra de esportes e laboratório de informática são os que mais contribuem para o bom desempenho dos alunos no Índice de desenvolvimento da Educação Básica” (BONAMINO e LIMA, 2013, p. 11).

No entanto no estudo realizado pelas autoras, utilizando os mesmos dados do Projeto GERES, afirmam que

a variável *biblioteca* não apresentou significância estatística para a onda 2 em nenhuma das duas disciplinas e por isso não foi incluída nos modelos. Na Onda 5, essa variável exerce efeito positivo nas duas disciplinas, sendo ligeiramente maior em leitura. (BONAMINO e LIMA, 2013, p. 110, grifo das autoras)

Como pode ser percebido, no estudo das autoras, no qual são usados modelos lineares hierárquicos com dados longitudinais do Projeto GERES e também com a proficiência inicial como variável explicativa, porém com intervalos bianuais, a variável tem impactos distintos nos diferentes anos, e difere dos resultados encontrados nessa pesquisa como mostra o Gráfico 60.

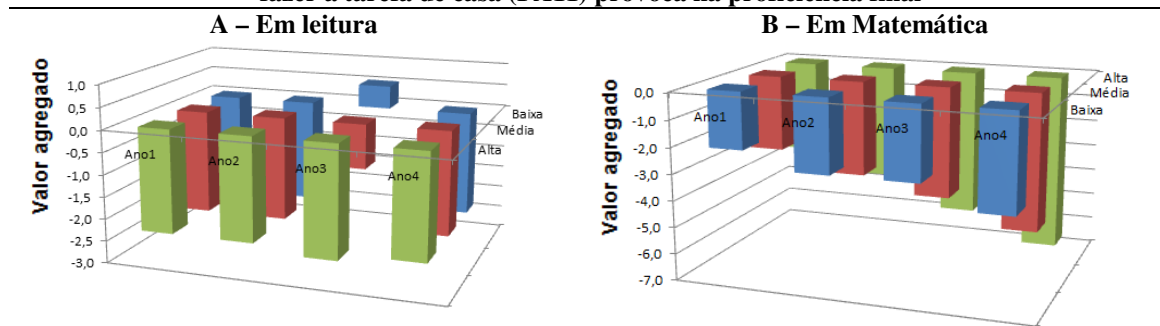
Gráfico 60 - Valor agregado negativo que a variável que controla a existência da Sala de Leitura (FEE04) provoca na proficiência final



Essa variável poderia ser descartada uma vez que são inúmeros os argumentos relacionados a falhas metodológicas desse estudo que ser usados para justificar esse resultado, no entanto, da mesma maneira, inúmeros argumentos podem ser usados para mostrar que esse resultado, apesar de contraditório, pode ser real. Um desses argumentos refere-se ao possível mau uso desse espaço, onde os alunos podem ser deslocados de suas salas de aula e usam o tempo no espaço para atividades que não levam à aprendizagem.

Outro resultado negativo encontrado no estudo foi encontrado na variável FA11, que controla o estímulo interno para tarefa de casa, como mostra o Gráfico 61.

Gráfico 61 - Valor agregado negativo que a variável que controla o estímulo interno do aluno para fazer a tarefa de casa (FA11) provoca na proficiência final



Essa variável é resultante da AFE a partir de três itens de questionário:

- Eu estudo porque estudar é importante para mim
- Eu gosto de estudar
- Eu gosto de estudar porque me dá prazer e alegria

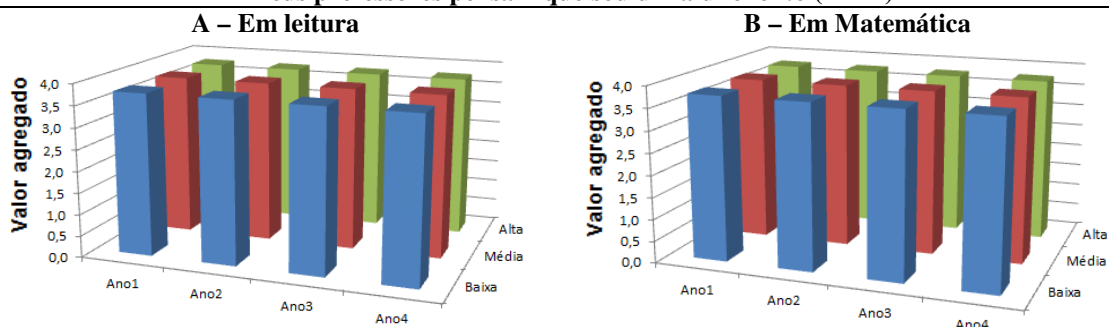
Cada um desses itens teve como opções de respostas as seguintes alternativas: “nunca” (1), “às vezes” (2) e “sempre” (3).

Diferentemente da variável anterior, argumentos que justificam tais resultados não são fáceis de serem encontrados, restando apenas a possibilidade de alguma falha metodológica.

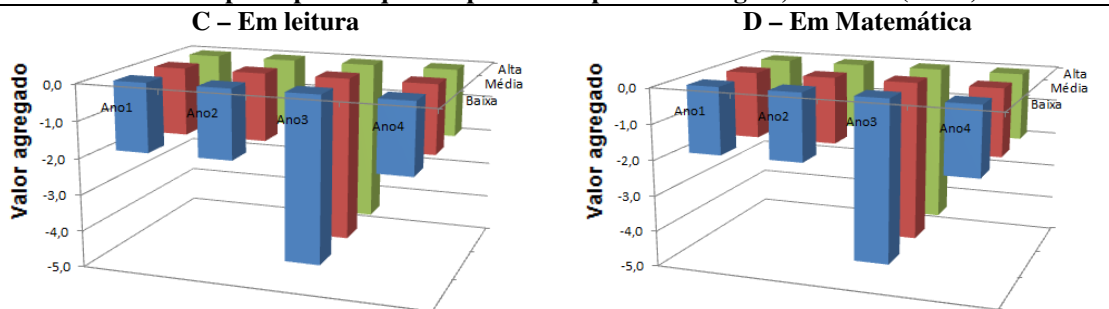
Na busca dessa falha, todos os procedimentos metodológicos foram refeitos para essa variável, assim como para as demais com resultados semelhantes. A primeira hipótese de falha levantada foi a inversão dos sinais, que transformaria esse resultado em algo adequado com a literatura. No entanto, ao final do processo de depuração não foram encontradas falhas nos procedimentos. Restando apenas como possibilidade a falha na coleta dos dados, nos procedimentos usados⁵⁶ ou no processo de depuração.

Os mesmos resultados contraditórios, apresentados nos Gráficos 63A, 63B, 63C e 62D são encontrados nas variáveis que controlam a percepção do aluno acerca autoestima do aluno.

Gráfico 62 - Valor agregado negativo que as variáveis associadas à autoestima (FA12) e (FA13)
Meus professores pensam que sou um aluno lento (FA12)



Sou tipo de pessoa que sempre será capaz de conseguir, se tentar (FA13)

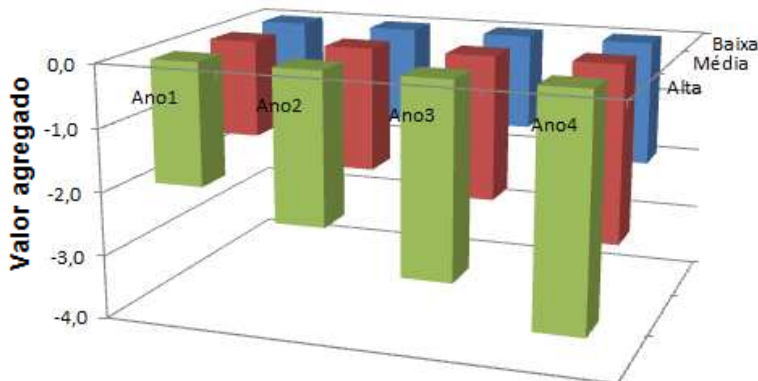


56 Em especial, no processo da Análise Fatorial Exploratória, cujos constructos usados nesta pesquisa forma diferentes ao proposto, *a priori*, pelo Projeto GERES. Isso ocorreu por conta de que a divulgação só ocorreu após a conclusão dessa etapa do trabalho nesta pesquisa. Uma das justificativas dessas diferenças deve-se às diferentes possibilidades de respostas ao itens, sendo que em alguns deles, eram ordinal, de escala Likert e dicotômicos.

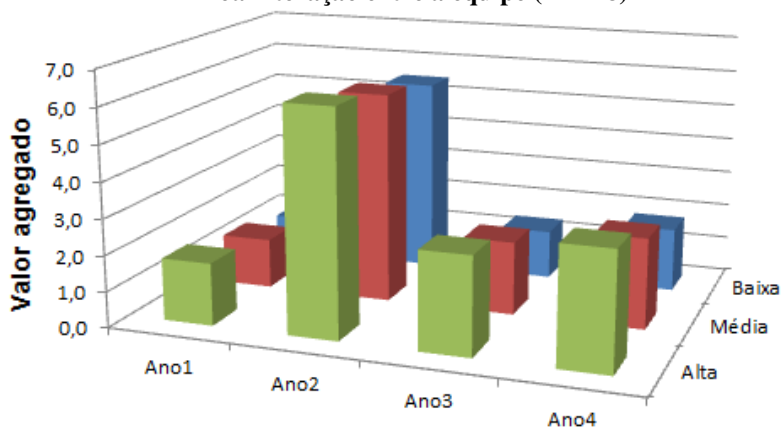
Especificamente em Leitura, os resultados com efeitos contrários ao esperado, em todos os anos de escolarização, foram observados nas três variáveis apresentadas nos Gráficos 63A, 63B e 63C.

Gráfico 63 - Variáveis com agregado negativo em Leitura (FEE03- FEP28 - FA09)

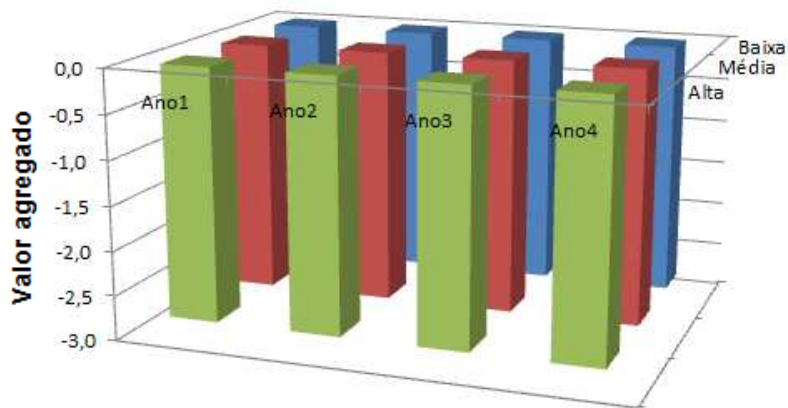
A - Conservação e limpeza área de alimentação (FEE03)



B - Boa interação entre a equipe (FEP28)



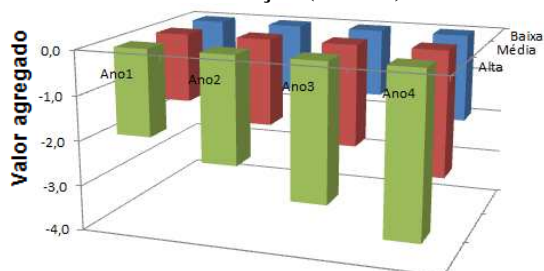
C - Percepção do aluno quanto ao incentivo do professor (FA09)



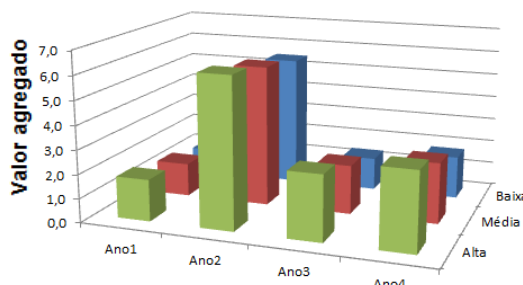
Em matemática, quatro variáveis tiveram efeitos contrários ao esperado, como mostram os Gráficos 64A, 64B, 64C e 64D.

Gráfico 64 - Variáveis com agregado negativo em Matemática (FED25 - FEP29 - FEP52 - FA14)

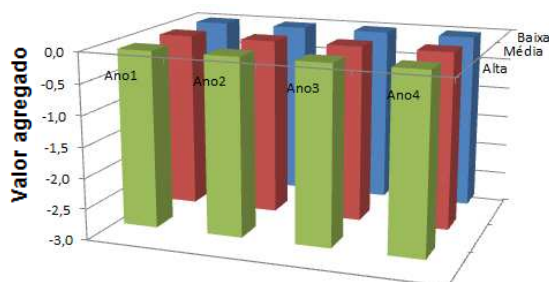
A - Tempo que o(a) diretor (a) trabalha em anos na educação (FED25)



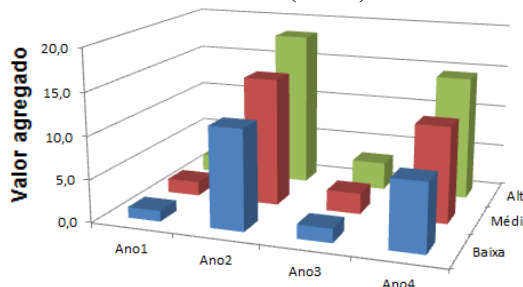
B - Pouco trabalho coletivo (FEP29)



C - Uso de recursos audiovisuais (FEP52)



D - Algumas vezes eu sinto que não sou bom o bastante (FA14)



Curiosamente, a variável FED25, que controla o tempo que o(a) diretor (a) trabalha em anos na educação e que em Matemática foi contrário ao esperado, em Leitura aparece com efeito de acordo com o esperado.

6.3.4.3) Comentários acerca das variáveis sem significância estatística.

Neste capítulo, com uso do Modelo de Referência, é esperado identificar a influência no valor agregado de cada uma das variáveis isoladamente, eliminando as diferenças entre os alunos devido às influências anteriores a sua entrada na escola (com a inclusão da proficiência inicial como variável explicativa) e também daquelas provocadas pelos nível socioeconômico do aluno e da escola, variáveis notadamente reconhecidas como influentes, após sua entrada na escola.

O número de variáveis constatadas como não significantes foi maior que o esperado, uma vez que não considera a interação com as outras variáveis. Nove variáveis não foram

estatisticamente significativa em Leitura e Matemática, outras dezesseis só em Leitura e dezenove só em Matemática, todas relacionadas no Quadro 99.

Quadro 99 - Variáveis sem significância estatística

Efeito em:	Descrição
Leitura e Matemática	Existência de equipamentos para fotocópia e escaneamento (FEE09);
	Interferência do perfil do professor (FED13);
	Interferência da direção na orientação pedagógica (FED15);
	Organização de eventos (FED17);
	Carga horária de trabalho do(a) diretor(a) na escola (FED23);
	Proposição de leitura (FEP37);
	Proposição de redação (FEP38);
	Grau de importância inerente à indicação externa do livro (FEP46);
	Proporção de professores do sexo feminino (FEP55).
Leitura	Sala de artes e laboratório de ciências (FEE07)
	Interferência por intimidação, violência, discriminação e drogas (FED12)
	Interferência do perfil do aluno, família (FED14)
	Presença física do diretor (FED16)
	Frequência de encontros com a comunidade (FED18)
	Proporção de diretores que possuem Ens. Sup. em Pedagogia (FED20)
	Proporção de diretores que trabalham em apenas uma escola (FED22)
	Tempo que o diretor trabalha na escola ininterruptamente na função de direção (FED24)
	Idade do(a) diretor(a) em anos (FED26)
	Violência no ambiente escolar (FEP32)
	Frequência do dever de casa (FEP33)
	Grau de expectativas positiva do aprendizado (FEP35)
	Proposição de cópia ou caligrafia (FEP39)
	Leitura em prática coletiva (FEP41)
	Uso de recursos audiovisuais (FEP52)
Idade média dos professores (FEP56)	
Matemática	Disponibilidade de computador e internet para professor (FEE08)
	Existência de material concreto para matemática (FEE10)
	Relação com o diretor (FEP27)
	Boa interação entre a equipe (FEP28)
	Boa relação entre professores (FEP30)
	Unidade de propósito entre os professores (FEP31)
	Acesso a atividades culturais (FEP36)
	Leitura individual e produção de texto individual (FEP42)
	Adequação do livro ao ensino (FEP45)
	Ensino com ênfase em memorização (FEP44)
	Tempo usado nas disciplinas relacionadas com as proficiências (FEP47)
	Tempo usado em disciplinas de formação complementar (FEP49)
	Uso de materiais concretos de matemática e geografia (FEP54)
	Renda Familiar média dos professores, em salários mínimos (FEP57)
	Tempo de exercício médio do professor na ocupação (FEP58)
	Número médio de pessoas da família que mora com o professor (FEP59)
	Estímulo extrínseco para estudo (FA07)
	Estímulo intrínseco para o estudo (FA08)
	Percepção do aluno quanto ao incentivo do professor (FA09)

Uma vez que no Modelo de Referência em Matemática ajustado com as variável de Nível 1 teve o *p-value* do coeficiente aleatório $\tau_{\pi00}$ maior que 0,5, era esperada uma tendência de que as variáveis de Nível 2 fossem não significantes estatisticamente.

Bonamino e Lima (2013) afirmam que “intervalos de tempos curtos (por exemplo, entre um ano escolar e outro) não são adequados, uma vez que a variação da aprendizagem não é grande” (p. 107) e esta pode ser uma das justificativas para esse número elevado de variáveis sem significância estatística.

No entanto, uma justificativa mais provável seja a coleta de um dado característico em um determinado momento seja considerado válido para todos os anos. Por exemplo, um dado de gestão válido em um momento, talvez não o seja nos momentos anteriores ou nem mesmo para momentos posteriores. Ou ainda, a disponibilização de um espaço disponível, como recurso didático, pode ser usado adequadamente em um momento e não em outro.

Essa especificidade de um momento sendo generalizada para os demais momentos é, certamente, um dos principais dificultadores metodológicos da pesquisa longitudinal.

Em contrapartida, diante dessa adversidade, as variáveis que se apresentam consistente nos diversos anos ganham confiabilidade maior, requerendo atenção especial aquelas, cujos resultados contrariam ao que é esperado.

CONCLUSÕES

Esta pesquisa, nos mesmos moldes do que ocorre em outras partes do mundo, desenvolve-se no Brasil em um período em que as avaliações externas deixam de ser apenas fontes de dados para orientação das políticas educacionais (*low stake*) e passam cada vez mais a, também, interferir diretamente na vida de alunos, professores e escolas (*high stake*), desconsiderando uma série de preocupações devido a algumas limitações técnicas das metodologias empregadas.

No contexto brasileiro, a principal limitação técnica deve-se ao fato de tais avaliações coletarem dados seccionais, que consistem na mensuração do desempenho de alunos em um determinado momento, e, a partir deles, produzirem uma única medida, favorecendo a produção de ranqueamentos, sejam eles individuais de alunos, sejam de escolas, redes de ensino, municípios e estados. Esses ranqueamentos são produzidos por meio da média do desempenho dos alunos, porém, em todos os casos, desconsiderando que parte dessas medidas não depende somente dos alunos ou da escola, mas também de outros fatores, como as relações educacionais que eles estabelecem entre si, as relações que estabelecem com o entorno e as famílias e também a proficiência com que iniciam cada período de escolarização.

Esta pesquisa insere-se nessa preocupação e coloca como objetivo identificar tais fatores por meio da análise com modelos lineares hierárquicos, buscando a superação dessa limitação básica, descrita nos Capítulos 1 e 2 e sintetizada por Brooke e Bonamino (2011, p. 12), que dizem:

Empreendimentos dessa natureza possuem uma limitação básica: a medida de desempenho em leitura e matemática é um agregado do aprendizado dos alunos ao longo de muitos anos. As informações sobre professores e as escolas, por sua vez, referem-se às condições escolares no ano da coleta de dados. Essa falta de sintonia temporal entre a medida do desempenho e as medidas das condições escolares fragiliza as análises e inviabiliza a formulação de políticas de qualidade e equidade baseada em evidências sólidas.

Conforme detalhado no Capítulo 2 deste trabalho, para uma mensuração mais precisa de como o perfil dos alunos, seus familiares, professores e gestores, as características das escolas e de seu entorno e as práticas escolares influenciam a proficiência dos alunos, é necessário um acompanhamento contínuo do desempenho do aluno e, simultaneamente, das condições em que o processo de escolarização se desenvolve. Tal acompanhamento requer,

então, dados longitudinais, e não dados seccionais, como os que têm sido coletados no Brasil desde o início da década de 1990.

Visando à construção de conhecimentos acerca de avaliações de larga escala com dados longitudinais, foi desenvolvido o Projeto GERES, que mediu a proficiência de uma amostra de alunos ao longo de quatro anos de escolarização, em Leitura e em Matemática, assim como das práticas educativas e das condições escolares.

Esta pesquisa usou esses dados coletados pelo Projeto GERES com o objetivo de identificar quais desses inúmeros fatores influenciam a proficiência dos alunos. Para tal, foram desenvolvidos modelos estatísticos que consideraram a complexidade do fenômeno escolar, isto é, modelos que levam em conta a influência simultânea de inúmeros fatores na proficiência final em cada ano de escolarização.

Ao final da pesquisa, ficou claro que esses modelos que identificam os fatores que agregam valor às proficiências, os chamados Modelos de Valor Agregado (VAM – *Value Added Models*), não eliminam todas as fragilidades dos resultados das avaliações educacionais, representando apenas a superação de uma de suas limitações.

Durante o desenvolvimento da pesquisa, inúmeras decisões metodológicas delinearam um caminho, com muitas idas e vindas, que culminou nos resultados e análise apresentados neste texto. Este, em certa medida, colabora com a finalidade do Projeto GERES de produzir conhecimentos sobre as avaliações educacionais longitudinais, em que se incluem os Modelos de Valor Agregado, ainda incipientes no Brasil, mas comuns em outras partes do mundo, especialmente nos Estados Unidos, onde

aparece[m] em todos os lugares hoje em dia, dos jornais à televisão para campanhas políticas. Os modelos de valor agregado são fortemente promovidos com entusiasmo desenfreado e acrítico pela imprensa, por políticos, e até mesmo por (alguns) especialistas em educação, e são apontados como o caminho moderno, "científico", para medir o sucesso educativo em tudo (EWING, 2011, p. 667, tradução do autor).

Ao término desta pesquisa, talvez a principal constatação seja a de que o caminho percorrido é um dentre inúmeros outros que poderiam ser tomados, cujos modelos levariam à identificação de outras variáveis. Esta é a justificativa para a afirmação anterior de que os Modelos de Valor Agregado não eliminam todas as fragilidades dos resultados das avaliações educacionais, ou seja, os resultados encontrados em qualquer avaliação educacional, assim

como em qualquer outro resultado na área da sociologia, precisam ser apresentados como verdadeiros, porém verdadeiros segundo um modelo selecionado.

Essa constatação pode ser considerada óbvia para alguns, no entanto, precisa ser questionada e divulgada para muitos outros, pois rompe com a herança positivista, que temos em nossa cultura, que nos leva à ideia da precisão e objetividade dos números, além do poder determinístico que a estatística nunca professou ter em seus modelos. Esse traço cultural pode ser percebido desde o âmbito da sala de aula, na entrega da correção de uma prova escrita, até o âmbito do sistema educacional, quando é divulgado o Índice de Desenvolvimento da Educação Básica (IDEB). Em certa medida, esse traço também pode ser percebido em situações como quando qualquer resultado educacional é publicado em relatórios produzidos pelos órgãos governamentais ou privados, como, por exemplo, o objetivo geral, que é “produzir informações sobre os fatores do contexto socioeconômico, cultural e escolar que influenciam o desempenho dos alunos” (INEP, 2011).

O caminho percorrido durante a pesquisa começou a ser escolhido com as decisões tomadas diante das seguintes perguntas:

a) Deverão ser consideradas apenas as proficiências dos alunos que permanecem na mesma escola ao longo do período estudado?

b) O valor agregado a ser estudado deverá ser anual, bianual ou considerado no período completo de cinco anos?

A opção por não excluir os alunos que trocaram de escola deve-se ao fato de que, como visto no Capítulo 3, praticamente 49,76% dos alunos da amostra se transferiram de ou para uma escola pertencente ao Projeto GERES durante o período de coleta de dados, que ocorreu entre o início de 2005 e o final de 2008. Assim, manter na amostra apenas os alunos que permaneceram na mesma escola faria estimativas de turmas que não existiram, pois as interferências dos demais alunos seriam eliminadas; conseqüentemente, a modelagem representaria um pouco menos a realidade do sistema educacional brasileiro. Além disso, a exclusão desses alunos diminuiria substancialmente o número de elementos da amostra, tendo como conseqüência um aumento do erro-padrão das estimativas.

A decisão de manter todos os alunos com pelo menos duas medidas consecutivas de proficiências, juntamente com a hipótese de que diversos fatores influenciam de forma diferente a cada ano a proficiência dos alunos, colaborou para a decisão de que o estudo fosse

feito ano a ano, uma vez que, para cada aluno que mudou de escola, foram consideradas as características da escola na qual ele estudou no respectivo ano de escolarização.

Com essa segunda decisão, que consistiu em fazer o estudo com base no valor agregado anualmente, veio à tona a preocupação de alguns pesquisadores, dentre eles, Bonamino e Lima (2013), de que o intervalo de um ano letivo não é adequado por apresentar variabilidade muito pequena na proficiência.

Essa preocupação, apesar de ser procedente, foi questionada, uma vez que, teoricamente, em períodos maiores que um ano de escolarização, uma variável só consolida seu efeito quando tem um efeito acumulativo nesses distintos períodos; caso contrário, por exemplo, um efeito positivo em um ano pode ser anulado em outro. Neste estudo, constata-se que, empiricamente, essa preocupação não se sustenta por completo. Como visto no Quadro 94, do Capítulo 5, onde estão apresentados os modelos ajustados para Leitura e para Matemática com a influência simultânea de todas as variáveis explicativas, verificou-se que, das 59 variáveis do Nível 3, apenas cinco não foram significantes estatisticamente em Leitura e 17 em Matemática. Dentre as 11 variáveis de Nível 2, todas foram significantes em Leitura e uma não o foi em Matemática.

A mesma constatação ocorre quando as variáveis são analisadas individualmente no Modelo de Referência, cujos resultados estão descritos no Capítulo 4. Nessa outra modelagem, apenas nove variáveis explicativas do Nível 3 não foram estatisticamente significantes em Leitura e em Matemática, 16 foram só em Leitura e outras 16 só em Matemática. Dentre as variáveis de Nível 2, três não foram significantes em pelo menos um ano somente em Matemática. Com essas constatações, fica evidente que a modelagem proposta nesta pesquisa permite captar a influência das variáveis explicativas em um ano de escolarização.

Por um lado, era previsto que o número de variáveis sem significância estatística fosse menor, uma vez que, como apresentado no Capítulo 2, todas as variáveis que fazem parte do banco de dados foram apontadas como sendo relevantes por diversos estudos que precederam este, ou seja, era esperado que todas elas tivessem, minimamente, alguma influência no valor agregado.

Por outro lado, com a inserção de 67 variáveis⁵⁷ influenciando o valor agregado, que, em média, como mostra a Tabela 28, varia de 12 a 19 em Leitura e de 27 a 44 em Matemática, é esperado que, em média, a influência de cada uma delas seja muito próxima de zero.

Tabela 28 - Valor Agregado Médio

	Leitura	Matemática
Ano 1	19	30
Ano 2	17	27
Ano 3	16	44
Ano 4	12	44

Era também esperado que, em algumas delas, esse valor se aproximasse ainda mais de zero, já que, nesse conjunto de variáveis, há algumas com efeitos maiores.

Nessa perspectiva, é importante destacar que, no teste de Wald, usado para estudar a significância estatística de uma variável explicativa, é verificada a probabilidade de uma estimativa ser igual a zero. Quando essa probabilidade, expressa pelo *p-value*, tem valor menor que 0,05 (e, em alguns casos, é admissível um valor até 0,1), isso significa que a estimativa da variável deve ser considerada diferente de zero. Dadas as características desse teste, quando o número de elementos da amostra é elevado, há uma tendência para que o *p-value* seja menor que 0,05 para quaisquer estimativas cujo valor seja realmente diferente de zero. Como nesta pesquisa o número de elementos da amostra é elevado, essa característica do teste de Wald é importante, porque atribui grande confiabilidade para as estimativas das variáveis cujo *p-value* foram maiores que 0,1, indicando que realmente não são estatisticamente significantes. A relevância dessa observação deve-se ao fato de que, na presente pesquisa, como o número de variáveis também é alto, a razão entre o valor agregado e o número de variáveis leva a uma tendência de existir um número elevado de variáveis não-significantes, sobretudo quando as variáveis estão associadas ao incremento por unidade de proficiência inicial, em que as estimativas estão na ordem de centésimos e até décimos de milésimos.

Com essas considerações, é possível afirmar que o objetivo da pesquisa de elaborar um modelo que considere a complexidade do fenômeno educacional foi cumprido.

⁵⁷ Nessa perspectiva, não se pode deixar de considerar que o nível socioeconômico individual do aluno e o nível socioeconômico médio da escola, além do controle sobre a retenção do aluno, também são variáveis que compõem o valor agregado.

Como apresentado no Capítulo 2, a modelagem proposta nesta pesquisa, feita em três níveis hierárquicos e usando dados longitudinais, permite:

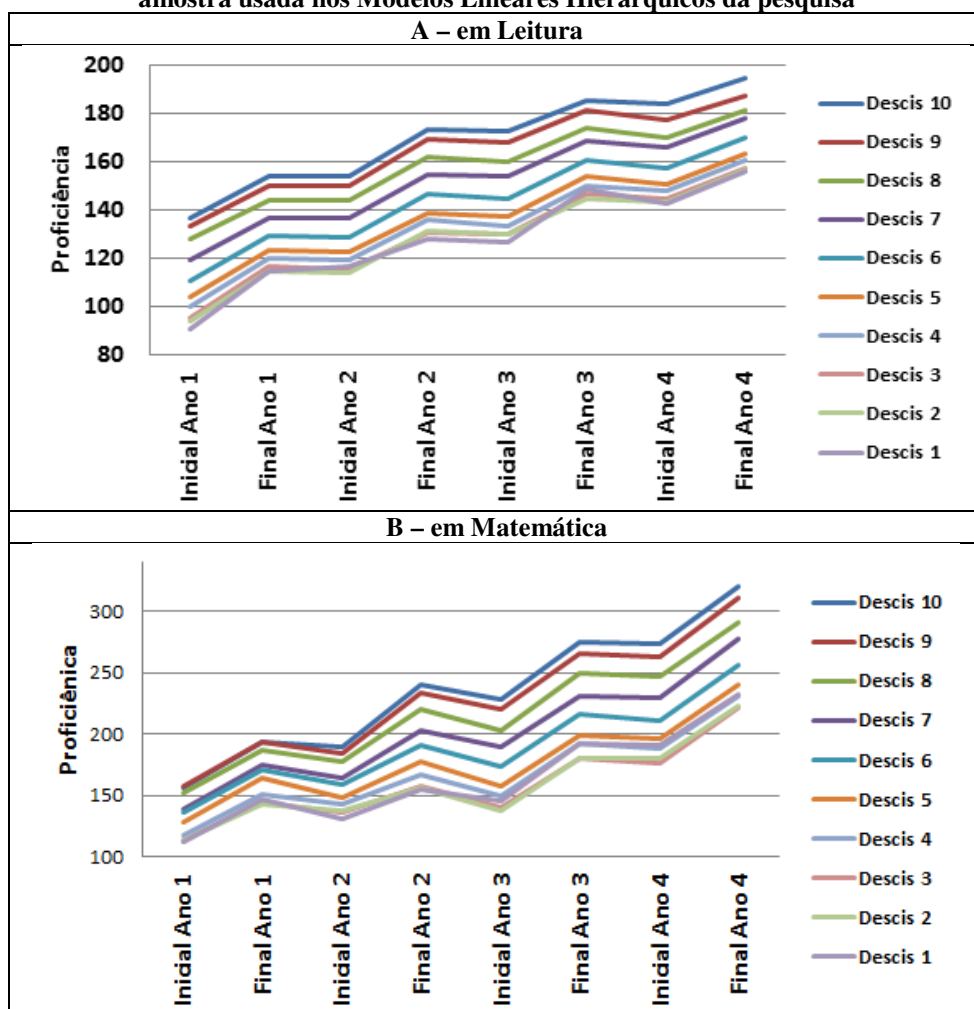
- a) identificar as variáveis significativas em cada um dos anos de escolarização;
- b) estimar o comportamento do efeito que cada uma dessas variáveis tem sobre a proficiência final de alunos com diferentes proficiências iniciais.

Essas características fizeram com que os modelos fossem sensíveis a qualquer variação das inúmeras variáveis que estavam sendo analisadas, gerando resultados instáveis que dificultam as análises.

No entanto, mesmo diante dessas instabilidades, a análise dos efeitos das variáveis feita nos Capítulos 5 e 6 mostra a tendência de que, durante o período de escolarização estudado nesta pesquisa, em Matemática, os alunos com proficiência inicial mais baixa têm um valor agregado médio menor do que aqueles que entram com proficiência inicial mais alta. Isso evidencia que a diferença de proficiência média entre os alunos com níveis socioeconômicos aumenta no decorrer dos anos de escolarização.

Apesar de ocorrer o contrário em Leitura, ou seja, os alunos com proficiência inicial mais baixa têm um valor agregado médio maior do que aqueles que entram com proficiência inicial mais alta, diminuindo a diferença das proficiências médias entre os alunos com níveis socioeconômicos diferentes, mas não ao ponto de maneira suficiente para eliminar tal diferença. Esse comportamento do valor agregado nas duas áreas pode ser verificado nos Gráficos 1A e 1B do Capítulo 2, que representam os desempenhos médios dos alunos que permaneceram na mesma escola ao longo dos quatro anos de escolarização, assim como pode também ser observado nos Gráficos 65A e 65B, que representam as proficiências médias iniciais e finais de cada ano de escolarização, em Leitura e em Matemática.

Gráfico 65 - Proficiências médias iniciais e finais de cada ano de escolarização da amostra usada nos Modelos Lineares Hierárquicos da pesquisa



Nesse gráfico⁵⁸, cada linha corresponde à proficiência média de um dos 10 agrupamentos de alunos feitos em função do seu nível socioeconômico, em ambas as áreas de conhecimento.

Ainda nos Gráficos 65A e 65B, constata-se que a proficiência média final de um ano é maior que a proficiência média inicial do ano subsequente. Essa diferença é provocada pela mobilidade de alunos, indicando que aqueles que saíram das escolas GERES têm proficiências maiores do que aqueles que entraram.

58 As escalas usadas nos gráficos são diferentes porque as escalas de proficiência em Leitura e em Matemática também são distintas, o que não permite a comparação entre elas, isto é, cada unidade em cada uma delas tem interpretações pedagógicas evidentemente distintas.

Além disso, a decomposição da variância observada nos Modelos Nulo, Modelo de Nível 1 e de Referência em Matemática, todos apresentados no Capítulo 4, mostra três diferentes estágios:

- No Modelo Nulo, quando não há a inserção de qualquer variável, a decomposição das variâncias mostra que 62,4% se devem às diferenças entre as proficiências finais dos quatro anos de escolarização, 15,1% se devem às diferenças entre os alunos e 22,5% se devem às diferenças entre as escolas. No entanto, comparando apenas as variâncias referentes aos alunos e às escolas, a partição é de 40,2% e 59,8%, respectivamente;

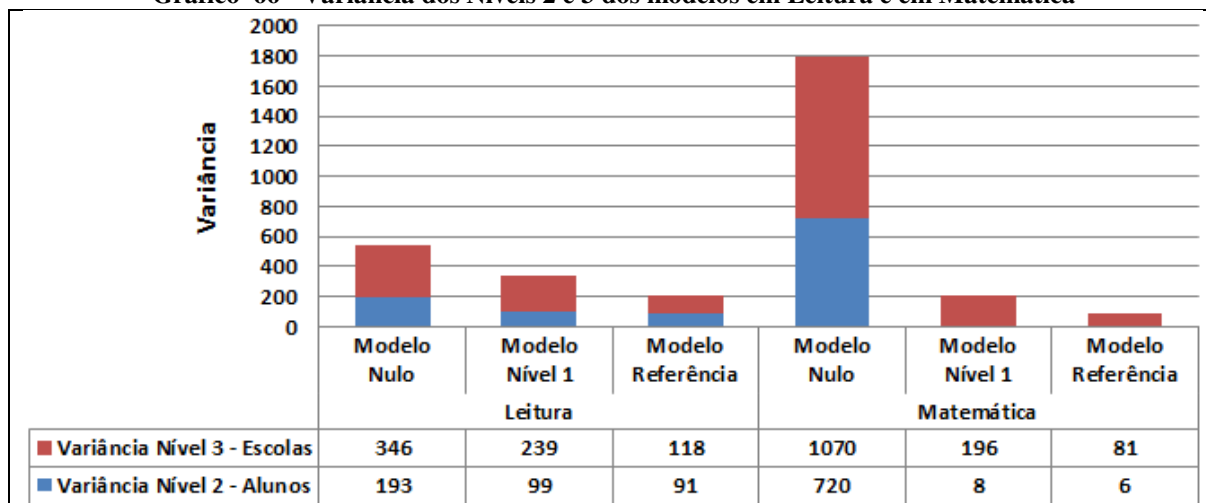
- No Modelo de Nível 1, quando foram inseridas as proficiências iniciais como variável explicativa e também a que controla a retenção do aluno no processo, a partição da variância entre os alunos e as escolas passa a ser de 3,7% e 96,3%, respectivamente. Com isso, pode-se notar que as diferenças entre os alunos de uma mesma escola é praticamente a mesma, em todas as escolas, enfatizando que as diferenças são predominantes entre as escolas. Isso ocorre porque a diferença entre os alunos passa a ser controlada, indiretamente, pela proficiência inicial;

- No Modelo de Referência, quando são inseridas as variáveis que controlam o nível socioeconômico dos alunos e das escolas, a partição da variância entre os alunos e as escolas passa a ser de 6,9% e 93,1%.

A redução da variância estimada nos modelos lineares hierárquicos, em Leitura e Matemática, assim como o quanto resta dessa variância em cada um dos níveis com a inserção das variáveis explicativas, é explicitada no Gráfico 66.

Com esse gráfico, fica evidente que, após o controle da proficiência inicial, da retenção do aluno no processo, do nível socioeconômico do aluno e da escola, a variabilidade da proficiência final em Matemática, diferentemente do que ocorre em Leitura, pode ser explicada apenas com variáveis do Nível 3, ou seja, variáveis que controlam as características da escola.

Gráfico 66 - Variância dos Níveis 2 e 3 dos modelos em Leitura e em Matemática



Esses gráficos, da mesma maneira que os gráficos 1A e 1B, já discutidos no Capítulo 2, mostram que as proficiências iniciais em Leitura e em Matemática são determinantes nas proficiências finais. Por isso, para a análise dos fatores que interferem na proficiência final, é essencial que a proficiência inicial seja colocada como uma variável explicativa, como foi feito no modelo usado nesta pesquisa, remetendo o estudo à eficácia diferencial (*differential school effects*), ou seja, a estudos que se caracterizam por comparar as escolas com alunos com a mesma característica socioeconômica.

Essa opção também justifica o resultado encontrado nesta pesquisa, em que o nível socioeconômico médio da escola aparece como um fator mais influente do que o nível socioeconômico individual do aluno, mas de fato não o é. Esse resultado aparece porque o nível socioeconômico do aluno é controlado indiretamente por meio do controle da proficiência inicial, dada a alta correlação existente essas duas variáveis. Esse resultado corrobora o que é afirmado por Ferrão e Couto (2013, p. 137):

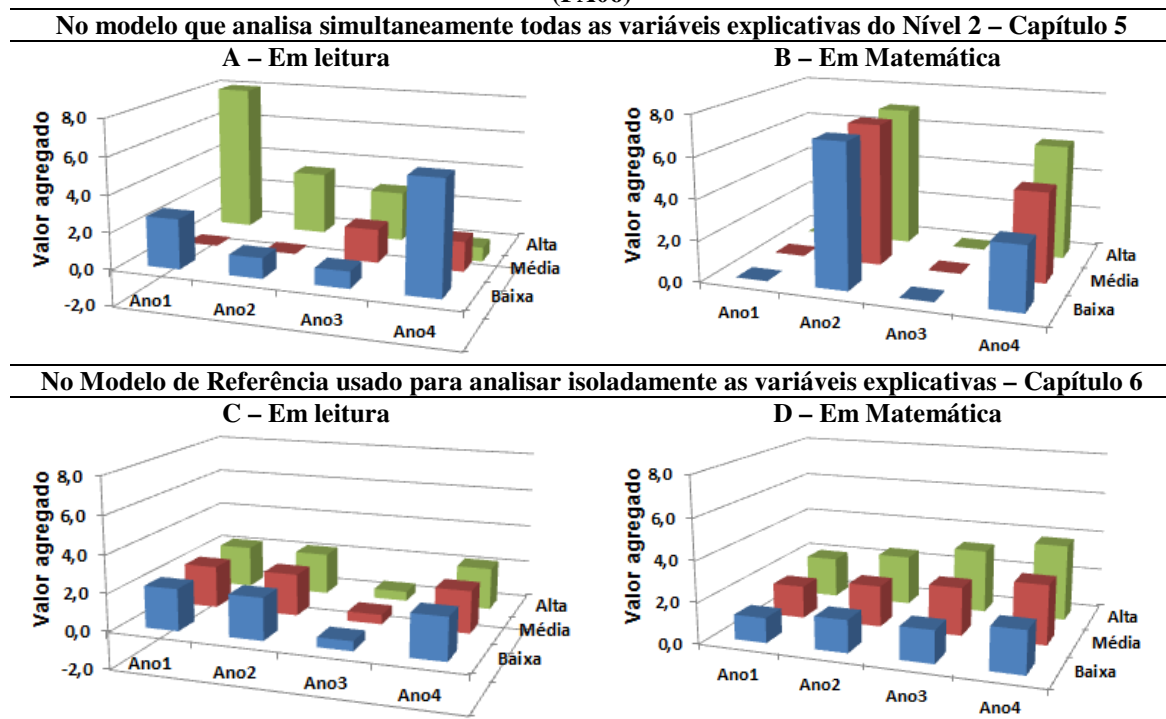
Alguns dos estudos que mostram evidência da ocorrência de eficácia diferencial (*differential school effects*), isto é o efeito-escola podem ser influenciados pelo perfil do aluno à entrada na escola (*students' intake*), revelam ainda que a composição da população discente (*school composition effects*) também influencia os resultados atingidos. Estes colocam em evidência o facto de o desempenho individual dos alunos de uma escola se encontrar associado ao perfil dominante dos seus alunos (*aggregated students' intake*).

Esse contexto em que o nível socioeconômico médio da escola mostra ser mais influente do que o nível socioeconômico individual do aluno na composição de sua

proficiência final é mais um exemplo do cuidado estritamente necessário para a adequada análise dos resultados de uma avaliação educacional e, mais ainda, para a divulgação desses resultados.

Os resultados encontrados nos Capítulos 5 e 6 mostram que o Nível socioeconômico individual do aluno continua a interferir em sua proficiência final, mesmo quando comparado a outros alunos com igual nível socioeconômico, depois de sua entrada na escola, como mostram os Gráficos 67A, 67B, 67C e 67D.

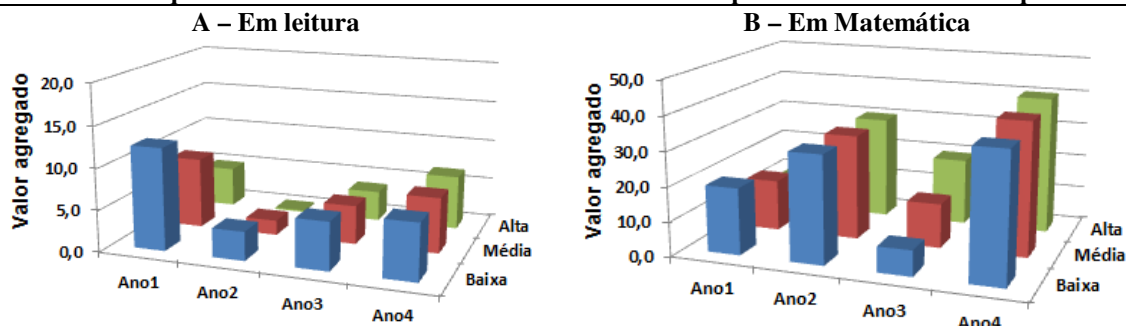
Gráfico 67 - Valor agregado correspondente a uma unidade no nível socioeconômico do aluno (FA06)



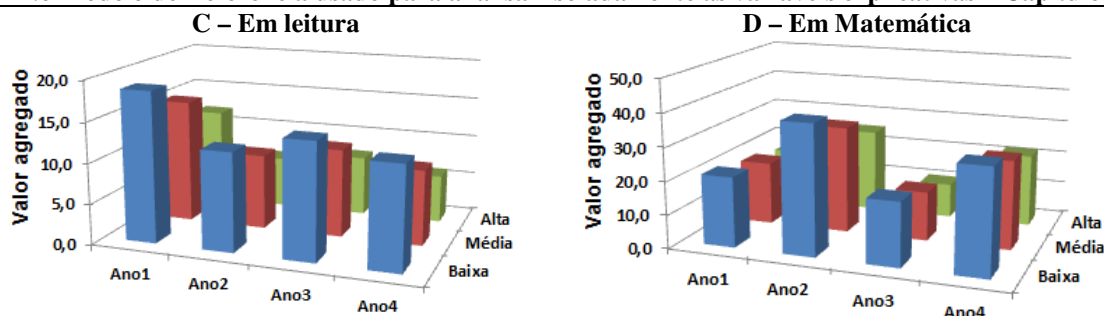
Com esse gráfico, fica evidente que o efeito da variável nível socioeconômico individual do aluno é diferente em Leitura e em Matemática, assim como também é nos distintos modelos, restando apenas a certeza de que, após entrada do aluno na escola, é um fator que certamente continua a influenciar e que a melhoria do contexto de vida do aluno colabora para o aumento da proficiência.

Gráfico 68 - Valor agregado correspondente a uma unidade no nível socioeconômico médio da escola (FEA01)

No modelo que analisa simultaneamente todas as variáveis explicativas do Nível 3 – Capítulo 5



No Modelo de Referência usado para analisar isoladamente as variáveis explicativas – Capítulo 6



Os valores que resultaram nessa variável parecem discrepantes em uma primeira análise, no entanto, deve-se lembrar que representa o valor agregado referente ao aumento de uma unidade no nível socioeconômico médio da escola. Porém, essa variável oscila entre - 0,46 e +1,28, sendo que a média é 0,18, com desvio-padrão de 0,44, o que significa que as escolas com nível socioeconômico correspondente ao valor médio da amostra têm atribuído nessa variável o valor 0,18 e que aproximadamente 68% das escolas estão entre o valor de - 0,26 e 0,62. Ou seja, a variação de uma unidade nessa variável da escola faz com que seus alunos, por exemplo, migrem agrupamento do primeiro para o décimo estrato, como representado nos Gráficos 1A, 1B, 65A e 65B.

Dessa forma, os modelos comprovam a importância que o nível socioeconômico dos alunos, que, agregado, se torna o da escola, tem na proficiência final, antes e depois da entrada do aluno na escola.

Esse resultado dá indícios fortes para a inadequação do uso de avaliações com dados seccionais e, ao mesmo tempo, destaca a importância do uso de modelos de valor agregado, uma vez que as escolas que, com ou sem o apoio de sua comunidade, garantem a permanência dos alunos no decorrer dos anos de escolarização, mesmo com um bom trabalho, estão por um

período fadadas a ter uma proficiência média baixa e a ser classificadas como más, mesmo estando cumprindo plenamente com o seu papel social.

Em um sentido contrário, esses resultados também explicam as estratégias usadas pelas escolas, de maneira branda ou não, de selecionar seus alunos para que tenham melhores desempenhos nas avaliações externas seccionais, como registra Fraga (2013, p.4) em sua entrevista ao professor José Francisco Soares, quando critica as políticas das escolas que buscam atrair apenas bons alunos.

O que a criança leva da casa é muito mais determinante do que a escola consegue acrescentar. Quem não tem boa escola não pode avançar, as escolas não podem riscar a contribuição das famílias [...] Parte do sucesso das escolas que ficam muito bem nos rankings não é o sucesso pedagógico. É o sucesso na atração do aluno com grande potencial de aprendizagem.

Por sua vez, essa influência do contexto de vida do aluno a importância de estudos de eficácia escolar apoiarem-se nas medidas de valor agregado, e não na média das proficiências obtidas pelos seus alunos em tais avaliações, sobretudo para minimizar, mas não resolver os problemas do inadequado uso meritocrático dos resultados da avaliação, como descrevem Ferrão e Couto (2013, p. 139),⁵⁹ comparando os modelos que usam o valor agregado com aqueles que usam dados seccionais

Estes refletem o desempenho num determinado momento do tempo e a sua eventual utilização como critério de um sistema de incentivos tende a premiar escolas com reduzido valor agregado e a desvalorizar a prestação das escolas inseridas em contextos adversos. [...] Algumas das escolas onde os alunos atingem níveis de desempenho elevado pouco contribuem para esse desempenho: tendencialmente são escolas frequentadas por alunos com elevado nível de desempenho à entrada (escolas seletivas) e provenientes de grupos sociais favorecidos. Em contrapartida, as escolas que recebem alunos com baixo desempenho e onde, no final do período sob avaliação, o desempenho se mantém aquém da meta educativa estabelecida, podem, apesar disso, ter contribuído fortemente para alavancar a aprendizagem desses alunos e, portanto, apresentam elevado VA.

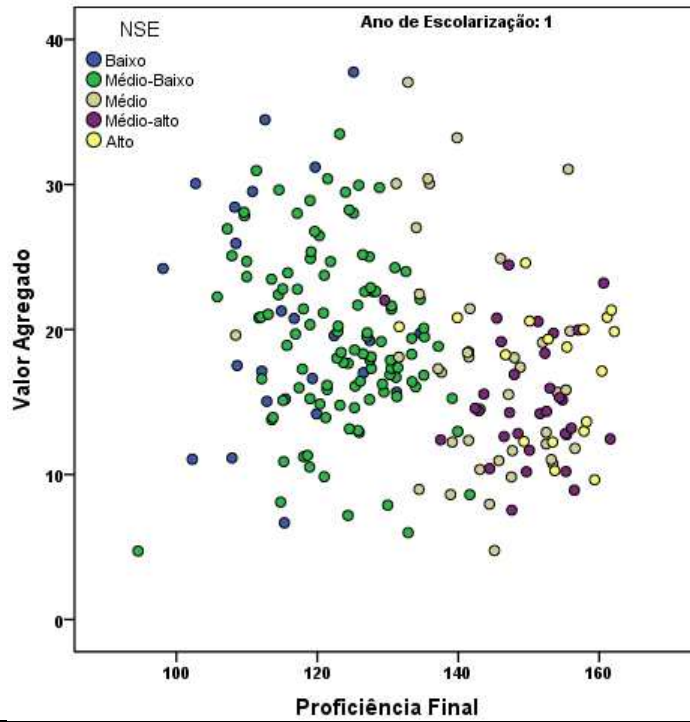
Os Gráficos 69A e 69B representam a dispersão das escolas com base no valor agregado médio e na proficiência final média no primeiro ano de escolarização em Leitura e

59 Os autores citam estes trabalhos para fundamentar suas afirmações: BRAUN et al., 2010; OECD, 2008; GOLDSTEIN, 1997; SAMMONS; THOMAS; MORTIMORE; BOSKER, 1997; SCHEERENS; BOSKER, 1997; RAUDENBUSH; WILLMS, 1995; MORTIMORE et al., 1988.

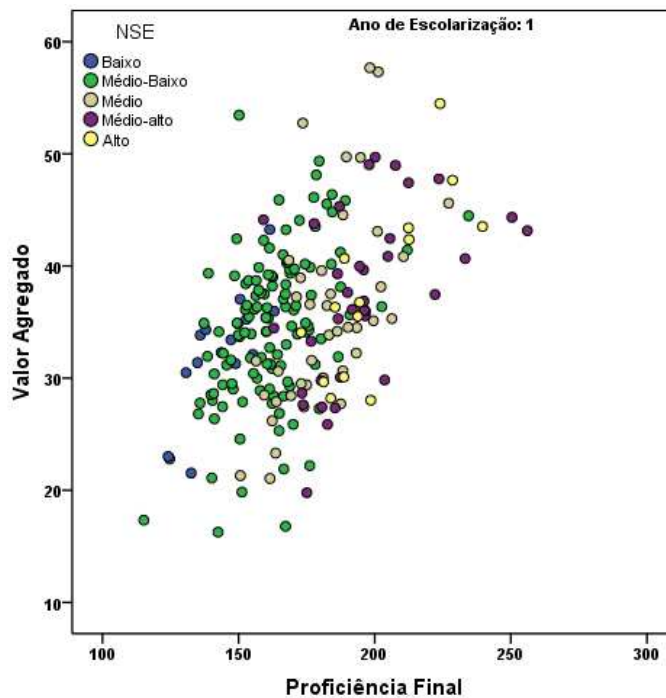
Matemática, de acordo com a proposta feita pelo National Research Council and National Academy of Education (2010).

Gráfico 69 - Dispersão das escolas em função de seu valor agregado médio e da proficiência final média

69 A – Em Leitura



69 B – Em Matemática



Essas escolas foram agrupadas em cinco níveis socioeconômicos, identificados com cores distintas com o objetivo de evidenciar que, nessa dispersão, há um agrupamento também em função do nível socioeconômico médio da escola, ou seja, o desempenho nessa relação entre o valor agregado e a proficiência tem alguma correlação com o nível socioeconômico. O mesmo é observado nos demais anos de escolarização e também em Leitura.

Com as constatações apresentadas até o momento, dos quatro benefícios que os modelos de valor agregado proporcionam aos estudos sobre a eficácia escolar e ao uso dos resultados das avaliações educacionais para sua melhoria, propostos por Ferrão e Couto, os três primeiros⁶⁰ sintetizam bem algumas das conclusões desta pesquisa:

Primeiro, permitem a identificação das escolas que mais contribuem para a aprendizagem dos alunos. Segundo, possibilitam ainda que escolas inseridas em contextos sociais, econômicos e culturais adversos possam ser reconhecidas pelo seu desempenho. Terceiro, tornam possível a identificação de situações críticas e de sucesso que possam apoiar a adoção de políticas de melhoria da eficácia da escola e de redução das desigualdades na educação. Finalmente, permitem que a avaliação educacional possa estabelecer uma relação mais rigorosa e justa entre desempenho e incentivos do que a viabilizada apenas pelos indicadores de status (FERRÃO e COUTO, 2013, p. 138).

No que se refere ao primeiro item apontado pelos autores, no Capítulo 6, foi feita uma análise individual das variáveis explicativas por meio de um Modelo de Referência composto apenas pelas variáveis que controlam o nível socioeconômico do aluno e da escola, fatores que não dependem da escola, com o objetivo de identificar quais os fatores que dependem da ação escolar para a melhora da proficiência.

Os resultados dessa análise apontam para duas ações que mostraram-se importantes em Leitura e em Matemática e que cabe ao sistema educativo implementá-las. A primeira refere-se à ampliação do Ensino Fundamental para nove anos, tornando obrigatória a matrícula das crianças com seis anos de idade. Foi uma medida acertada, uma vez que os alunos que cursaram a então denominada pré-escola tiveram um ganho significativo nas proficiências em Leitura e em Matemática. A segunda ação está associada à exclusividade do trabalho do diretor, pois as escolas onde os diretores tinham outra atividade que complementa sua renda mensal tiveram perdas significativas na proficiência nas duas áreas. Ainda associado

60 O quarto benefício é questionado mais adiante.

ao perfil do diretor, o tempo que o diretor trabalha em anos na educação mostrou influenciar especificamente a proficiência em Leitura, e o fato de o diretor trabalhar em apenas uma escola influencia a proficiência em Matemática.

Esta segunda variável já se mostrava importante, pois aparece nos constructos dos sete trabalhos (WILLMS, 1992; LEE, BRYK e SMITH, 1993; SAMMONS, HILLMAN e MORTIMORE, 1995; MURILLO, 2003; RACZINSKI e MUÑOZ, 2004; MELLO, 2004 e MEC, 1998) usados pelo Projeto GERES para fundamentar a elaboração de seus instrumentos de coleta de dados. A importância do diretor escolar também é destacada por Betini (2009) quando, em um estudo qualitativo, se inseriu como pesquisador-apoiador para a implementação de processo que levasse a Avaliação Institucional Participativa a escolas pertencentes às redes municipais de Campinas participantes do Projeto GERES. Segundo ele,

Embora não tenha sido um pressuposto inicial, destacou-se durante o exercício de intervenção do apoiador na escola, no caminhar da coleta de dados e na análise dos mesmos, revelando-se ao pesquisador clara e simplesmente: a importância do papel político do dirigente escolar. Constata-se o quanto a direção das escolas mergulha no cotidiano, num trefismo que lhes tira a energia para o objetivo maior que a traz e mantém na escola. Exatamente num momento em que impera na discussão nacional a gestão escolar, o papel político do dirigente escolar se apresenta nesse trabalho muito fortemente como um motor possível para tirar a escola do marasmo do cotidiano e das tarefas burocráticas (BETINI, 2009, p.324).

Outro trabalho ligado ao Projeto GERES com destaque para a importância do diretor foi feito por Polon (2009, p. 303), que afirma que a gestão escolar, “quando praticada com ênfase nos aspectos pedagógicos, parece contribuir para que as escolas obtenham melhores resultados, o que faz dessa tipologia um possível apoio à elaboração de hipóteses sobre indicadores de eficácia”.

Na dimensão pedagógica, outras duas variáveis foram significativas. A primeira delas colabora para o aumento da proficiência em Leitura e Matemática e está associada à metodologia de resolução de problemas. Essa variável é resultante de um conjunto de itens de questionário que por si só esclarecem o porquê de essa variável ter sido tão significativa estatisticamente; são eles:

- contaram para os colegas sobre como eles tentam resolver um problema?
- lidaram com problemas que exigem raciocínios diferentes ou mais complexos?
- discutiram as possibilidades de resolução de um problema após resolvê-lo?

- lidaram com problemas que envolvam a representação de situações em linguagem matemática?

- lidaram com situações relacionadas ao cotidiano?

- os alunos trabalharam com temas que aparecem em jornais ou revistas, relacionando-os com a Matemática?

- apresentaram para os colegas as estratégias que eles utilizaram para resolver as questões do dever de casa?

Apesar de ser uma variável aparentemente ligada ao ensino da Matemática, a dinâmica de trabalho que atende positivamente a essas questões beneficia também as demais áreas, afinal,

o fato de que seja preciso ensinar os alunos a resolver problemas escolares próprios da área não deve significar que em cada área se enfrente o ensino da solução de problemas de um modo diferente ou desvinculado do que ocorre em outras áreas. Embora os conhecimentos conceituais e algumas das estratégias necessárias para resolver um problema de matemática e para realizar um jogo de simulação em geografia sejam diferentes, [...] existem muitas dificuldades comuns ao ensino e à aprendizagem da resolução de problemas nesses diversos domínios (POZO e ANGÓN, 1998, p. 140).

Em um sentido contrário, os resultados dessa pesquisa mostram que as interrupções das aulas por bagunça dos alunos, pela entrada de alunos que chegam atrasados ou por barulho nos corredores prejudicam significativamente o aumento da proficiência, tanto em Leitura quanto em Matemática. Especificamente em Leitura, as interrupções das aulas por necessidades administrativas, como anúncios ou comunicações da direção, coordenação ou secretaria, também interferem negativamente na proficiência.

Ainda em termos de ação pedagógica, em Matemática, a única variável que impactou positivamente foi a que está relacionada à frequência com que o dever de casa (FEP33) é passado para os alunos, enquanto que, em Leitura, as práticas com leitura individual e produção de texto individual (FEP42) e as com ênfase em memorização (FEP44) impactaram positivamente na proficiência. Nesta segunda, o resultado gera certo estranhamento, pois pertence ao constructo ligado à área de Matemática, composto por itens que questionam o professor, verificando se sua prática leva os alunos a lidarem com situações para memorizar conceitos e regras e também se suas propostas de atividade levam os alunos a decorarem regras e a aplicá-las em situações-problema. De qualquer forma, essas práticas definem um

perfil de ação do professor que vai de encontro às da metodologia de resolução de problemas consideradas atualmente mais adequadas, no entanto leva a uma maior proficiência em Leitura.

Nesse mesmo sentido, resultado interessante foi encontrado com a variável que controlou o tempo usado em disciplinas de formação complementar às áreas de Leitura e Matemática. Essa variável (FEP49) resultante da Análise Fatorial Exploratória, composta por itens do questionário que controlaram a proporção do tempo usado para o ensino de Ciências, Geografia e História, mostram que um tempo maior destinado a essas disciplinas aumenta a proficiência em Leitura. Esse é um resultado importante, pois mostra que a priorização do tempo nas disciplinas consideradas nas avaliações em larga escala em detrimento das demais é uma estratégia equivocada até mesmo para essa finalidade. Muitos estudos mostram a influência que as avaliações externas, sobretudo com a ênfase nos processos de *accountability*, têm tido sobre os currículos escolares, pois as matrizes dessas avaliações se tornam o currículo da escola.

Quanto à infraestrutura disponível nas escolas, o uso do laboratório de informática com acesso à Internet (FEE06) e a disponibilidade de computador e Internet para o professor (FEE08) influenciaram positivamente a proficiência em Leitura. As condições de uso da biblioteca influenciaram positivamente a proficiência em Matemática, apesar de essa influência ser esperada em Leitura, o que não aconteceu⁶¹.

Essas variáveis mostram que, mesmo enquanto os problemas do nível socioeconômico não são resolvidos, a escola e seus profissionais têm a potencialidade de melhorar a proficiência dos alunos, sobretudo em Matemática, em que as diferenças entre os alunos com distintos níveis socioeconômicos são aumentadas no decorrer dos anos de escolarização.

No entanto, nem todos os resultados encontrados se manifestaram de maneira que traduz aquilo que é esperado. Por exemplo, é bastante improvável que o uso do espaço destinado à sala de leitura (FEE04), o estímulo intrínseco do aluno para realização de sua tarefa de casa (FA11) ou a percepção de que o aluno tem de que o professor o acha um aluno

61 É importante destacar que, dadas as características dos modelos que usam de alguma forma a regressão linear, dentre eles, o modelo linear hierárquico usado nesta pesquisa, a ausência da biblioteca, assim como outros fatores que não foram significantes estatisticamente, deve ser interpretada como uma variável que não conseguiu explicar a variabilidade da proficiência final e não como não sendo importante.

lento (FA12) sejam fatores que colaboram para o aumento da proficiência em Leitura e em Matemática e que o pouco trabalho coletivo na escola provoque um aumento na proficiência nessas duas áreas.

Evidentemente, algumas hipóteses podem ser levantadas para justificar os resultados – em princípio, incoerentes – dessas quatro variáveis. Algumas delas poderiam ser associadas às falhas metodológicas da pesquisa ou a erro em algum dos procedimentos realizados em seu desenvolvimento. Outras, assumindo-se todos os procedimentos como corretos, poderiam buscar argumentos que requereriam outras pesquisas qualitativas para a sua confirmação.

Os Capítulos 5 e 6 estão repletos de exemplos de variáveis que apresentam resultados que estão de acordo com o esperado. Da mesma maneira, o contrário é encontrado, além de haver variáveis consideradas instáveis ou não-significantes.

Comparando os resultados das análises com os Modelos de Referência, apenas a que controla as condições de uso da biblioteca (FEE05) e a que controla a quantidade de anos em que o diretor trabalha na área educacional (FED25) apresentaram resultados contraditórios entre si. As análises mostram que o uso da biblioteca colabora para a proficiência em Matemática ao mesmo tempo em que prejudica a Leitura, enquanto que, quanto mais anos o diretor trabalhar na área educacional, maior será a proficiência em leitura, mas o contrário em Matemática.

Apesar de os modelos mostrarem-se suficientemente robustos, a instabilidade inerente aos estudos dessa natureza é a justificativa para relativizar a afirmação apresentada anteriormente por Ferrão e Couto (2013, p. 138) acerca dos benefícios do uso de modelos de valor agregado para que “a avaliação educacional possa estabelecer uma relação mais rigorosa e justa entre desempenho e incentivos do que a viabilizada apenas pelos indicadores de status”.

É bastante coerente afirmar que os modelos de valor agregado são mais rigorosos e justos com relação aos modelos que usam dados seccionais; no entanto, não são suficientemente justos para tornar as avaliações externas “*high stake tests*”.

O Quadro 100 faz uma síntese quantitativa para explicitar a instabilidade das variáveis, tal como já foi evidenciado no Capítulo 5, em que as variáveis foram analisadas simultaneamente.

Quadro 100 - Comparativo entre os resultados em Leitura e Matemática com o Modelo de Referência

Com o mesmo resultado	de acordo com o esperado	5
	contrário ao esperado	4
	não-significante	9
	instável	3
Subtotal		21
Com resultados diferentes	instável em uma das áreas e não-significante em outra	17
	instável em uma das áreas e de acordo com o esperado em outra	5
	instável em uma das áreas e contrário ao esperado em outra	5
	não-significante em uma das áreas e de acordo com o esperado	11
	não-significante em uma das áreas e contrário ao esperado em outra	7
	de acordo com o esperado em uma das áreas e contrário em outra	2
Subtotal		47
Total		68

Diante desse resultado, o aprimoramento na coleta dos dados, assim como o seu tratamento até chegar à análise, diminuindo-se o número de variáveis não-significantes estatisticamente e com a instabilidade percebida, além da seleção daquelas que são pertinentes à natureza desta pesquisa, mostram-se necessários para que os resultados reflitam um pouco mais a realidade, na busca da pretensa justiça, mas com a certeza de que sempre estarão longe de representar a realidade.

A seleção é necessária porque algumas variáveis, como, por exemplo, aquelas ligadas ao estilo de gestão, quando sofrem alteração em um determinado instante, seus efeitos são percebidos em outros, assim como as marcas do estilo anterior permanecem por um tempo considerável. Outras variáveis usadas para pesquisa acadêmica nos mais diversos objetos de pesquisa, tais como gestão escolar, didática e organização curricular, devem ser deixadas para estudos com essa finalidade. Uma proposta seria manter apenas as variáveis importantes para o monitoramento das escolas, seja por elas mesmas, seja para a orientação de políticas públicas. Sempre, é claro, sem qualquer finalidade meritocrática, para evitar os mais diversos tipos de fraudes ou, minimamente, distorção de resultados. Nessas condições, as avaliações educacionais poderiam cumprir mais plenamente sua finalidade, que é fornecer dados para a melhora educacional para os decisores de políticas públicas e, principalmente, para a comunidade interna da escola em processos de avaliação institucional participativa.

Os resultados encontrados no Capítulo 5 mostram que a tecnologia disponível nos dias de hoje permite processar um grande número de variáveis, mas um limite está colocado, quer por condições técnicas, de tempo ou financeiras para a coleta de um número grande de dados, quer por limitações dos programas computacionais ou de *hardware*.

A proposta feita por Ferrão e Couto (2013, p.134), dentre outros autores, é “adotar plano de recolha de dados longitudinal que envolva no mínimo três coortes de alunos” para verificar a consistência das estimativas, que consiste no grau de concordância entre elas nas diferentes coortes, tendo como referência diferentes conjuntos de variáveis e modelos. Essa proposição também é sustentada por Raudenbush (2004, p.127

Raudenbush, Bryk e Ponisciak (2003) analisam dados coletados em cinco coortes de alunos ao longo de cinco anos, em Washington, D.C.. Mesmo com mais de 50 mil estudantes, a precisão na estimativa de efeitos de professores foi modesta. Usar várias coortes parece essencial para obter a precisão adequada. Além disso, o efeito das escolas só podem ser determinados ao longo de vários grupos, a fim de obter um efeito de média estável. Finalmente, as tendências de aproveitamento (ganhos de valor agregado) não podem ser calculadas sem múltiplas coortes.

Mesmo assim, os modelos de valor agregado estão sendo criticados nos Estados Unidos, onde seu uso é recorrente para avaliar o desempenho de alunos, professores e escolas. Dentre muitos outros pesquisadores (CHETTY, FRIEDMAN, ROCKOFF, 2011; LAVIGNE, GOOD, 2013), Baker *et al* sintetizam essas observações:

Avanços estatísticos recentes tornaram possível olhar para ganhos de desempenho dos alunos após o ajuste para algumas características do aluno e da escola. Estas abordagens que medem o crescimento usando modelos de valor agregado são comparações mais justas de professores do que julgamentos com base em resultados de testes de seus alunos em um único ponto no tempo ou comparações de coortes de estudantes que envolvem diferentes alunos em dois pontos no tempo. Os modelos de valor agregado também contribuem para análises mais robustas do progresso escolar, das influências de programas governamentais, bem como da validade dos métodos de avaliação do que anteriormente. No entanto, há um amplo consenso entre os estatísticos, psicometristas e economistas de que as notas dos alunos só são indicadores, não são suficientemente confiáveis e válidas de eficácia do professor para serem usadas em decisões pessoais, de alto impacto, mesmo quando são usados aplicativos estatísticos mais sofisticados para a modelagem com valor agregado.

Assim, constata-se que as decisões que definem o modelo, o tratamento dos dados, a coleta dos dados e os instrumentos de coleta (ALMEIDA, 2002; GUNTHER, 2003; MADAUS, RUSSELL e HIGGINS, 2009) definem diretamente os resultados encontrados nos

estudos. Portanto, a certeza que decorre deste estudo é que um modelo não retrata a realidade, mas sim uma realidade vista e apropriada nos limites do recorte feito diante das inúmeras decisões metodológicas tomadas e desconsiderando um conjunto de outras variáveis não controladas no processo.

O Brasil deveria caminhar para a implantação desses modelos, pois em muito contribuiria para a melhoria da qualidade educacional oferecida a sua população. Evidentemente desvinculado de qualquer processo meritocrático, cujas restrições foram observadas nesse estudo, confirmando o que já vem sendo apontado há algum tempo em pesquisa internacionais, afinal o sucesso ou o fracasso não é da escola ou do professor, tampouco do aluno ou sua família. O sucesso ou o fracasso na educação é do sistema educativo e da própria sociedade, que ocorre por meio da escola, do professor, da família e do aluno.

Para tal, bastaria organizar os dados de maneira longitudinal, uma vez que são feitas medidas censitárias e em períodos regulares de maneira já estão consolidadas promovidas nos entes federativos. Essa adoção, não só produziria dados longitudinais, mas também, com o passar dos anos, cortes que poderiam ajudar na análise de consistência e estabilidade.

Evidentemente, esses estudos devem manter distantes das influências governamentais como é proposto por Goldstein e Woodhouse (2008, p. 423)

é claro que muitas das descobertas da eficácia escolar são relevantes para as agendas de governo, e um modo de ver o papel dos pesquisadores da eficácia escolar como se fossem amigos críticos, que podem contribuir para o debate público das políticas, enquanto mantêm sua integridade, especialmente quando ignoradas ou ameaçadas pelo governo. Isso pode ser difícil quando o próprio governo tende a considerar a comunidade de pesquisa formada por amigos comprometidos ou inimigos subversivos.

A complexidade inerente aos modelos não favorece tais debates, sejam eles metodológicos ou simplesmente de seus resultados, e, por isso, em meio à idolatria de alguns e à fobia de outros com relação a tais modelos, os estudos de eficácia escolar podem se tornar um campo fértil para o uso até nefasto desses resultados, tendo como consequência a ineficácia e a inequidade do sistema educativo.

Nesse sentido, é válido o alerta dado por Freitas (2003, p. 35) quando afirma que

a escola é um pouco mais complexa do que um conjunto de variáveis a serem manipuladas, e os valores dessas variáveis são de difícil estabilização e transferência para outras situações. [...] Isso não significa dizer que nada há a ser aprendido com os estudos quantitativos e com as avaliações de larga escala. Significa que devemos colocá-los em seu devido lugar.

A devida identificação dos alunos nos diferentes anos de aplicação das avaliações censitárias já existentes no Brasil, promovidas pelos diferentes entes federativos, poderia fornecer dados longitudinais, viabilizando o uso de modelos de valor agregado, e, diante de diferentes coortes, permitir o avanço aqui proposto. Além disso, o desenvolvimento de estudos de eficácia escolar colocando outros indicadores como variável resposta poderia ajudar as políticas públicas e as escolas a olharem para outras finalidades das escolas.

Com as análises finais deste estudo, juntamente com as dos demais citados neste trabalho, é possível afirmar que os modelos estatísticos já atingiram um patamar de sofisticação tecnológica de tal forma que a melhora de seus resultados estão limitados à qualidade dos dados coletados, à forma com que são modelados e à maneira que os resultados são publicizados. Com isso, o desafio está em conseguir com que estes resultados sejam usados para a melhoria da qualidade oferecida pelo sistema educativo.

Neste sentido, na perspectiva técnica, as limitações metodológicas inerentes a esses modelos são impeditivas para o uso meritocrático de seus resultados, ao mesmo tempo em que são potentes na produção de informações para orientar reflexões no interior do sistema educativo na busca da garantia da aprendizagem, que ocorre por meio da escola e do professor.

REFERÊNCIAS BIBLIOGRÁFICAS

AFONSO, Almerindo J. **Avaliação Educacional: regulação e emancipação**. São Paulo: Cortez, 2000.

AITKIN, Murray; LONGFORD, Nicholas. Statistical modelling issues in school effectiveness studies. **Journal of the Royal Statistical Society**. Series A (General), p. 1-43, 1986.

ALEXANDRE, João W. C. *et al.* Uma proposta de análise de um constructo para fatores da gestão pela qualidade por intermédio da Teoria da resposta ao Item. In: **Gestão & Produção**, Fortaleza, v.9, n.2, p.129-141, 2002.

ALMEIDA, Alberto. C. **O efeito do contexto e posição da pergunta no questionário sobre o resultado da medição**. Opinião Pública, vol.8, n.2, 2002.

ALMEIDA, Ivanete B. **Análise do desempenho de escolas públicas cicladas e não cicladas pertencentes ao ensino fundamental**. Tese (Doutorado). 250 p. Campinas: Faculdade de Educação-Unicamp, Campinas, 2009.

ALMEIDA, Luana C. **Relação entre o desempenho e o entorno social em escolas municipais de Campinas: a voz dos sujeitos**. Tese (Doutorado). 279 p. Campinas: Faculdade de Educação-Unicamp, 2014.

ALVES, Maria T. G.; SOARES, José F. Medidas de nível socioeconômico em pesquisas sociais: uma aplicação aos dados de uma pesquisa educacional. **Opinião Pública**, v. 15, n.1, p.1-30, 2009.

ALVES, Maria T. G.; SOARES, José F. O efeito das escolas no aprendizado dos alunos: um estudo com dados longitudinais no Ensino Fundamental. **Educação e Pesquisa**, v.34, n.3, p. 527-544, 2008.

ANDRADE, Dalton F.; TAVARES, Heliton R.; VALLE, Raquel C. **Teoria da Resposta ao Item: Conceitos e Aplicações**. São Paulo: ABE – Associação Brasileira de Estatística, 2000.

BABBIE, Earl. **Métodos de Pesquisas de Survey**. Tradução de Guilherme Cezarino. Belo Horizonte: Editora da UFMG, 2005.

BACKER, Eva. L. *et al.* Problems with the Use of Student test scores to evaluate teachers. **EPI Briefing Paper**. Economic Policy Intitute. n. 278, Aug. 2010. Disponível em: http://epi.3cdn.net/b9667271ee6c154195_t9m6iij8k.pdf. Acesso em: 28 jul. 2013.

BARROSO, João. O Estado, a Educação e a Regulação das Políticas Públicas. **Educação & Sociedade**, vol.26, n. 92, p. 725-751, 2005.

BETINI, Geraldo A. **Avaliação Institucional em Escolas Públicas de Ensino Fundamental de Campinas**. Tese (Doutorado). 394p. Campinas: Faculdade de Educação-Unicamp, 2009.

BIRD, Sheila M. *et al.* Performance indicators: good, bad, and ugly. **Journal of the Royal Statistical Society** (Series A), v. 168, n. 1, p. 1-27, 2005.

BONAMINO, Alicia; FRANCO, Creso. Avaliação e política educacional: o processo de institucionalização do SAEB. **Cadernos de Pesquisa**, v. 108, p. 101-132, 1999.

BONAMINO, Alicia; LIMA, Naira C, M.. Aspectos da gestão escolar e seus efeitos no desempenho dos alunos nos anos iniciais do ensino fundamental. In: MARTINS, Ana M. *et al.* **Políticas e Gestão da Educação: desafios em tempos de mudança.** Campinas: Autores associados, 2013.

BRAUN, Henry I. **Using student progress to evaluate teachers: A primer on value-added models.** Princeton, NJ: Educational Testing Service, 2005.

BROOKE, Nigel; BONAMINO, Alicia. **GERES 2005: Razões e Resultados de uma pesquisa longitudinal sobre eficácia escolar.** Rio de Janeiro: Walprint, 2011.

BROOKE, Nigel; SOARES, José F. **Pesquisa em eficácia escolar: origem e trajetórias.** Belo Horizonte: Editora UFMG, 2008.

BRYK, Anthony. S.; RAUDENBUSH, Stephen W. e PONISCIAK, Steven. **A value-added model for assessing improvements in school productivity: Results from the Washington, DC public schools and an analysis of their statistical conclusion validity annual meeting of American Educational Research Association,** San Diego, CA. 2003.

BUCHMANN, Claudia. Measuring Family Background in International Studies of Education: Conceptual Issues and Methodological Challenges. In: PORTE, Andrew; GAMORAN, Adam. **Methodological Advances in Cross-National Surveys of Educational Achievement.** Washington, D.C.: National Academy Press, 2002.

CANGUSSU, Maria A. R. **Características escolares associadas ao desempenho dos estudantes na Pesquisa GERES: a escola pode fazer diferença?** Tese (Doutorado). 247p. Campinas: Faculdade de Educação-Unicamp, 2010.

CASTRO, Guiomar. N. M. Sistemas de Avaliação da Educação no Brasil: avanços e novos desafios. **Perspectiva,** São Paulo, v. 23, n. 1, p. 5-18, 2009.

CÉSAR, Cibele C.; SOARES, José F. Desigualdades acadêmicas induzidas pelo contexto escolar. **Revista Brasileira de Estudos Populacionais,** Rio de Janeiro, v. 18, n. 1/2, p. 97-110, 2001.

CHETTY, Raj; FRIEDMAN, John N.; ROCKOFF, Jonah E. The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood. **NBER Working Paper.** Cambridge. n. 17699, 2011.

COLEMAN, James S. *et al.* **Equality of educational opportunity.** Washington, D.C: U.S. Dept. of Health, Education, and Welfare, Office of Education; U.S. Government Printing Office, 1966.

COSTA, Maria S. F. P. M. **Avaliação institucional no ensino fundamental: a participação dos estudantes.** Tese (Doutorado). 402p. Campinas: Faculdade de Educação-Unicamp, 2012.

CRAMER, Duncan. **Advanced quantitative data analysis.** Berkshire: McGraw-Hill Education, 2003.

CRUZ, Cláudia C. M. S. **Modelos Multinível: Fundamentos e Aplicações.** Dissertação (Mestrado).200 p. Lisboa: Universidade Aberta, 2010.

DALBEN, Adilson. **Avaliação Institucional Participativa na Educação Básica: possibilidades, limitações e potencialidades.** Dissertação (Mestrado). 253 p. Campinas: Faculdade de Educação-Unicamp, 2008.

DIAS SOBRINHO, José. Campo e Caminhos da Avaliação: a avaliação da educação superior no Brasil. In: FREITAS, Luiz C. (org.) **Avaliação: construindo o campo e a crítica**. Florianópolis: Editora Insular, 2002.

DRAPER, Norman R.; SMITH, Harry. **Applied regression analysis**. New York: John Wiley, 1981.

ELLIOTT, John. School effectiveness research and its critics: Alternative visions of schooling. **Cambridge journal of education**, v. 26, n. 2, p. 199-224, 1996.

EWING, John. Mathematical Intimidation: Driven by Data. **Notices of the AMS (American Mathematical Society)**. Vol. 58. n. 5, p. 667 - 673, Mai. 2011. Disponível em <http://www.ams.org/notices/201105/rtx110500667p.pdf>. Acesso em: 15 nov. 2013.

FACCENDA, Odival; DALBEN, Adilson; FREITAS, Luiz C. Capacidade explicativa de questionários de contexto: aspectos metodológicos. In: **Revista Brasileira de Estudos Pedagógicos**. Vol. 92, n. 231, p. 246-267, mai.-ago.2011.

FERRÃO, Maria E. **Introdução aos modelos de regressão multinível em educação**. Campinas: Komedi, 2003.

FERRÃO, Maria E. Avaliação educacional e modelos de valor acrescentado: tópicos para reflexão. **Educação & Sociedade**. Campinas: Cedes, v. 33, p. 455-470, n. 119, abr.-jun. 2012.

FERRÃO, Maria E.; COUTO, Alcino. Indicador de Valor Acrescentado e Tópicos sobre Consistência e Estabilidade: uma aplicação ao Brasil. In: **Ensaio: avaliação e políticas públicas em educação**, Rio de Janeiro, v. 21, n. 78, p. 131-164, jan./mar. 2013. Disponível em: http://www.scielo.br/scielo.php?script=sci_issuetoc&pid=0104-403620130001&lng=en&nrm=iso. Acesso em 10 out. 2013.

FERRÃO, Maria E.; FERNANDES, Cristiano. 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, 2001.

FERRÃO, Maria E.; KLEIN, Rubem. Editorial da Revista 78. In: **Ensaio: avaliação e políticas públicas em educação**, Rio de Janeiro, v. 21, n. 78, p. 7-9, 2013. Disponível em: http://www.scielo.br/scielo.php?script=sci_issuetoc&pid=0104-403620130001&lng=en&nrm=iso. Acesso em 22 dez. 2013.

FERRÃO, Maria. E. *et al.* O SAEB – Sistema Nacional de Avaliação da Educação Básica: objetivos, características e contribuições na investigação da escola eficaz. **Revista Brasileira de Estudos Populacionais**, v.18, n.1 e 2, p.111-130, 2001.

FLETCHER, Philip R.. A Teoria da Resposta ao Item: medidas invariantes do desempenho escolar. **Ensaio: aval. pol. públ. educ.**, Rio de Janeiro, v. 01, n. 02, 1994. Disponível em <http://educa.fcc.org.br/scielo.php?script=sci_arttext&pid=S0104-40361994000100004&lng=pt&nrm=iso>. Acesso em 13 jun. 2013

FLETCHER, Philip. R. **À procura do ensino eficaz**. Relatório técnico. Brasília: MEC-DAEB, 1998.

FORQUIN, Jean C. (org.) **Sociologia da educação: dez anos de pesquisa**. Petrópolis/RJ: Vozes, 1995.

FRAGA, Érica. Pais não devem escolher escolas só com base no ENEM, diz pesquisador. **Folha de São Paulo**, São Paulo, 29 nov. 2013. Cotidiano, p. 4.

FRANCO, C; BROOKE, N.; ALVES, F. Estudo longitudinal sobre qualidade e equidade no ensino fundamental brasileiro: GERES 2005. **Ensaio: Avaliação e Políticas públicas em Educação**. Rio de Janeiro, v. 16, n. 61, p. 625-638, 2008.

FRANCO, Creso. *et al.* O referencial teórico na construção dos questionários contextuais do SAEB 2001. **Estudos em Avaliação Educacional**, n.28, p.39-74, 2007.

FRANCO, Creso. O SAEB - Sistema de Avaliação da Educação Básica: potencialidades, problemas e desafios. **Revista Brasileira de Educação**, Rio de Janeiro, n. 17, 2001. Disponível em <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-24782001000200010&lng=pt&nrm=iso>. Acesso em: 13 dez. 2013.

FREITAS, Luiz C. (org.) **Avaliação: construindo o campo e a crítica**. (Coleção Avaliação: construindo o campo e a crítica). Florianópolis: Editora Insular, 2002.

FREITAS, Luiz C. **Ciclos, Seriação e Avaliação: confronto de lógicas**. São Paulo: Moderna, 2003.

FREITAS, Luiz C. Eliminação adiada: o ocaso das classes populares no interior da escola e a ocultação da (má) qualidade do ensino. **Educação & Sociedade**, Campinas, v. 28, n. 100 - Especial, p. 965-987, out. 2007.

FREITAS, Luiz C. *et al.* (org.). **Avaliação e políticas públicas educacionais: Ensaios contrarregulatórios em debate**. (Coleção Avaliação: construindo o campo e a crítica). Campinas-SP: Leitura Crítica, 2012.

FREITAS, Luiz C.; BELLONI, Isaura; SOARES, José F. (Org.). **Avaliação de escolas e universidades**. (Coleção Avaliação: construindo o campo e a crítica). Campinas: Komedi, 2003.

FREITAS, Luiz C.; GATTI, Bernadete A.; SOUZA, Sandra M. Z. L. (Org.). **Questões de avaliação educacional**. (Coleção Avaliação: construindo o campo e a crítica). Campinas: Komedi, 2003.

FREITAS, Luiz C.; SORDI, Mara R. L.; MALAVASI, Maria M. S.; FREITAS, Helena C.L. **Avaliação educacional: caminhando pela contramão**. Petrópolis: Vozes, 2009.

GHUNTHER, Hartmut. **Como elaborar um questionário**. Brasília: Laboratório de Psicologia Ambiental, 2003.

GIBSON, Alex; ASTHANA, Sheena. Schools, pupils and exam results: Contextualizing school performance. **British Educational Research Journal**, v. 24, n. 3, p. 195-210, 1998.

GOLDSTEIN, Harvey. **Multilevel statistical models**. London: Edward Arnould, 1995.

GOLDSTEIN, Harvey; WOODHOUSE, Goeffrey. Pesquisa sobre eficácia escolar e políticas educacionais. In: BROOKE, Nigel e SOARES, José F., **Pesquisa em eficácia escolar: origem e trajetórias**. Belo Horizonte: Editora UFMG, 2008.

GOMES, Cândido A. **Fundamentos de uma lei de responsabilidade educacional**. Brasília, DF: Unesco, 2008. (Debates, 10). Disponível em: <unesdoc.unesco.br>. Acesso em: 22 fev. 2012.

GRAY, J. Desenvolvendo métodos de Valor Agregado para a Avaliação da Escola. In: BROOKE, Nigel e SOARES, José F. **Pesquisa em eficácia escolar: origem e trajetórias**. Belo Horizonte: Editora UFMG, 2008.

HAIR, Joseph F. *et al.* **Análise multivariada de dados**. (6a. Ed.). Editora: ARTMED; 2009.

HAMILTON, David. Peddling feel good fictions. **Forum** 38, v. 2, p. 54-56, 1996.

HOX, Joop. **Multilevel analysis: Techniques and applications**. Cambridge: Routledge, 2010.

INEP. Instituto Nacional de Estudos e Pesquisas Educacionais. **Prova Brasil: Avaliação do rendimento escolar**. Disponível em: <http://provabrasil.inep.gov.br>. Acesso em 12 de abril de 2011.

JENKS, Christopher. Desigualdade no aproveitamento Educacional. In: BROOKE, Nigel e SOARES, José F., **Pesquisa em eficácia escolar: origem e trajetórias**. Belo Horizonte: Editora UFMG, 2008.

JENKS, Christopher. **Inequality: a Reassessment of the Effect of Family in Schooling in America**. New York: Basic Books, 1972.

KREFT, Ita e LEEWN, Jan. **Introducing multilevel modeling**. Londres: Sage, 1998.

LAHIRE, Bernard. Reprodução ou prolongamento críticos? **Educação e Sociedade**. Campinas, SP, CEDES, ano XXIII, n. 78, p. 37-55, abr. 2002.

LAROS, Jacob A.; MARCIANO, João L. P. Análise multinível aplicada aos dados do NELS:88 In: **Estudos em Avaliação Educacional**, v. 19, n. 40, mai. – ago., 2008.

LAVIGNE, Alyson; GOOD, Leah. **Thomas L. Teacher and student evaluation: moving beyond the failure of school reform**. New York: Routledge, 2013.

LEE, Valerie E. A necessidade dos dados longitudinais na identificação do efeito-escola. In: **Revista Brasileira de Estudos Pedagógicos**, Brasília, v. 91, n. 229, p. 471-480, set. -dez., 2010a.

LEE, Valerie E. Dados longitudinais em educação: um componente essencial da abordagem de valor agregado no que se refere à avaliação de desempenho escolar. In: **Estudos em Avaliação Educacional**, São Paulo, v. 21, n. 47, p. 531-542, set.- dez., 2010b.

LEE, Valerie E. Utilização de modelos hierárquicos lineares para estudar contextos sociais: o caso dos efeitos escola. In: BROOKE, Nigel e SOARES, José F., **Pesquisa em eficácia escolar: origem e trajetórias**. Belo Horizonte: Editora UFMG, 2008.

LEE, Valerie E. What are multilevel questions, and how might we explore them with quantitative methods? In: **Estudos em Avaliação Educacional**, São Paulo, v. 24, p. 31-45, 2001.

LEE, Valerie E.; BRYK, Anthony S.; SMITH, Julia. The Organization of Effective Secondary Schools. In: DARLING-HAMMOND, L. **Review of Research in Education**. Washington, D. C.: American Educational Research Association, 1993.

MADAUS, George F.; AIRASIAN, Peter W.; KELLAGHAN, Thomas. Insumos Escolares, processos e recursos. In: BROOKE, Nigel e SOARES, José F., **Pesquisa em eficácia escolar: origem e trajetórias**. Belo Horizonte: Editora UFMG, 2008.

MADAUS, George; RUSSEL, Michael e HIGGINS, Jennifer. **The Paradoxes of High Stakes Testing: how they affect students, their parents, teachers, principals, schools, and society**. Charlotte: IAP, 2009.

MALAVAZI, Maria M. S.; BERTAGNA, Regiane H.; FREITAS, Luiz C. (Org.). **Avaliação: desafio dos novos tempos**. (Coleção Avaliação: construindo o campo e a crítica). Campinas: Komedi, 2006.

MANLY, Bryan F. J. **Métodos estatísticos multivariados: uma introdução**. Porto Alegre: Bookman, 2008.

MEC – MINISTÉRIO DA EDUCAÇÃO E DO ESPORTO, Secretaria de Educação Fundamental. **Referências para Formação de professores**. Documento técnico, Brasília, 1998.

MELLO, Guiomar N. Escolas eficazes: em tema revisitado. In: XAVIER, Antonio C. *et al.* **Gestão Escolar: desafios e tendências**. Brasília: IPEA, 2004.

MIRANDA, Antonio C. **O desafio da construção de referências de qualidade para os sistemas de ensino: uma avaliação com o uso de análise envoltória de dados – DEA**. Tese (Doutorado). 351p. Campinas: Faculdade de Educação-Unicamp, 2008.

MIRANDA, Eliana C. M. **O SAEB-2003 no estado de São Paulo: um estudo multinível**. Campinas: Dissertação (Mestrado). Campinas: Faculdade de Educação-Unicamp, 2006.

MIRANDA, Eliana C. M.; ANDRADE, Dalton F. **GERES: Relatório Técnico do Plano Amostral**, 2005.

MIRANDA, Eliana. C. M. **O Saeb-2003 no Estado de São Paulo: um estudo multinível (HLM)**. Dissertação (Mestrado em Educação) – Universidade Estadual de Campinas, Campinas, 2006.

MORAES, Mônica C. M. **Avaliação da escola pelos pais**. Tese (Doutorado). 151p. Campinas: Faculdade de Educação-Unicamp, 2010.

MORTIMORE, Peter. *et al.* **School Matters: the junior year**. Wells: Open Books, 1998, 1998.

MORTIMORE, Peter. **The Road to Improvement: Reflections on School Effectiveness**. Netherlands: Swets&Zeitlinger Publishers, 1998.

MORTIMORE, Peter. **The use of performance indicators**. Paris: OCDE, 1991.

MOSTELLER, Frederick e MOYNIHAN, Daniel P. Um Relatório Inovador. In: BROOKE, Nigel e SOARES, José F., **Pesquisa em eficácia escolar: origem e trajetórias**. Belo Horizonte: Editora UFMG, 2008.

MURILLO, Francisco J. ¿Importa la escuela? Una estimación de los efectos escolares en España. In: **Tendencias Pedagógicas**, 10, p. 29-45, 2005.

MURILLO, Francisco J. **La investigación sobre eficacia escolar en Iberoamérica: Revisión Internacional sobre el estado del arte**. Bogotá: Convenio Andrés Bello, 2003.

MURILLO, Francisco J. Mejora De La Eficacia Escolar en Iberoamérica. In: **Revista Iberoamericana de Educación**, n. 55, p. 49-83, 2011. Disponível em: <http://www.rieoei.org/rie55a02.pdf>. Acesso em: 17 jun 2013.

MURILLO, Francisco J. Um panorama da pesquisa Ibero-americana. In: BROOKE, Nigel e SOARES, José F., **Pesquisa em eficácia escolar: origem e trajetórias**. Belo Horizonte: Editora UFMG, 2008.

MURILLO, Francisco. J. Mejora de la eficacia escolar. In: **Cuadernos de Pedagogia**, n. 300, p. 47-74, 2001.

NATIONAL RESEARCH COUNCIL AND NATIONAL ACADEMY OF EDUCATION. **Getting Value Out of Value-Added: Report of a Workshop**. Committee on Value-Added Methodology for Instructional Improvement, program Evaluation, and Educational Accountability, Henry Braun, Naomi Chudowsky, and Judith Loenig, Editors. Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press, 2010.

OECD. **Literacy skills for the world of tomorrow: further results from PISA 2000**. Paris, 2004.

OLIVEIRA, Lina K. M.; FRANCO, Creso; SOARES, Tufi M. C. Projeto GERES / 2005: Novos indicadores para a construção e interpretação da escala de proficiência. In.: **REICE - Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación**. v. 5, p. 154-182, n. 2, 2007

PLEWIS, Ian. **Statistics in Education**. London: Edward Arnold, 1997.

PLOWDEN COMMITTEE. **Children and their primary school**. London: HMSO, 1967.

POLI, Edneia. C. **Estudo longitudinal em Matemática: possibilidades e leitura de uma realidade do ensino fundamental**. Tese (Doutorado). 274p. Campinas: Faculdade de Educação- Unicamp, 2007.

POLON, Thelma L. P. **Identificação dos perfis de liderança e características relacionadas à gestão pedagógica eficaz nas escolas participantes do Projeto Geres: Estudo Longitudinal - Geração Escolar 2005 - Pólo Rio de Janeiro**. Tese (Doutorado). 323p. Rio de Janeiro: Departamento de Educação da Pontifícia Universidade Católica do Rio de Janeiro, 2009.

POZO, Juan I.; ANGÓN, Yolanda P. A solução de problemas como conteúdo procedimental da educação Básica. In: POZO, Juan I. *et al.* **A solução de problemas: aprender a resolver, resolver para aprender**. Porto Alegre: Artmed, 1998.

PRING, Richard. Educating Persons: Putting Education Back into Educational Research. (The 1995 SERA lecture). **Scottish Educational Review**, v. 27, p. 101-112, 1995.

RAUDENBUSH, Stephen W. **Schooling, Statistics, and Poverty: Can We Measure School Improvement?**. Educational Testing Service, 2004.

RAUDENBUSH, Stephen, RACZINSKI, D.; MUÑOZ, G. **Factores que desafían los buenos resultados educativos de escuelas en sectores de pobreza**. Santiago de Chile: Preal, 2004.

RAUDENBUSH, Stephen W.; BRYK, Anthony S. **Hierarchical linear models: Applications and data analysis methods**. California: Sage, 2002.

ROBINSON, John P.; SHAVER, Phillip R.; WRIGHTSMAN, Lawrence S. *Criteria for scale and evaluation*. In.: ROBINSON, John P.; SHAVER, Phillip R.; WRIGHTSMAN, Lawrence

S. **Measures of Personality and social Psychological Attitudes**. San Diego, CA: Academic Press, 1991.

RODRIGUES, Sueli C. **Construção de uma metodologia alternativa para a avaliação das escolas públicas de ensino fundamental através do uso da análise por Envoltória de Dados (DEA)**. 488p. Campinas: Tese de Doutorado, UNICAMP, 2005.

RUTTER, Michael. **Fifteen Thousand hours**. London: Open Books, 1979

SAMMONS, Pam; HILLMAN, Josh; MORTIMORE, Peter. **Key Characteristics of Effective Schools: a review of school effectiveness research**. London: Office for Standards in Education, 1995.

SNIJEDERS, Tom A. B.; BOSKER, Roel J. **Multilevel analysis: an introduction to basic and advanced multilevel modeling**. London: Sage, 1999.

SOARES, José F. O efeito da escola no desempenho cognitivo de seus alunos. **Reice: Revista Eletrônica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación**, v. 2, n. 2, p. 83-104. 2004.

SOARES, José F.; ANDRADE, R.J. Nível socioeconômico, qualidade e equidade das escolas de Belo Horizonte. **Ensaio: Avaliação e Políticas Públicas Educacionais**, Rio de Janeiro, v. 14, n. 50, p. 107-126, jan./mar. 2006.

SOARES, José F.; CANDIAN, J.F. O efeito da escola básica brasileira: as evidências do PISA e do SAEB. **Revista Contemporânea de Educação**, v. 2, n. 4, p. 45-64, 2007.

SOARES, José F., CÉSAR, C. C.; MAMBRINI, J. Determinantes de Desempenho dos Alunos do Ensino Básico Brasileiro: Evidências do SAEB de 1997. In: FRANCO, C. (Org.) **Promoção, ciclos e avaliação educacional**. Porto Alegre: Artmed, 2001.

SOARES, José F.; COLLARES, Ana C. M. Recursos Familiares e o Desempenho Cognitivo dos Alunos do Ensino Básico Brasileiro. **DADOS - Revista de Ciências Sociais**, v. 49, n. 3, p. 615-481, 2006.

SOARES, Tufi M. Utilização da Teoria de Resposta ao Item na produção de indicadores socioeconômicos. In.: **Pesquisa Operacional**, v.25, n.1, p. 83-112, 2005.

SOARES, Tufi M. Modelo de três níveis hierárquicos para a proficiência dos alunos de 4a série avaliados no teste de língua portuguesa do SIMAVE/ PROEB-2002. **Revista Brasileira de Educação**, n. 29, p. 73-87, 2005.

SORDI, Mara R. L. Entendendo as lógicas da avaliação Institucional para dar sentido ao contexto interpretativo. In: VILLAS BOAS, B. M. F. **Avaliação: políticas e práticas**. Campinas: Papirus, 2002.

SOUZA, Sandra M. Z. L. Avaliação do rendimento escolar como instrumento de gestão educacional. In: OLIVEIRA, Dalila A. (org.). **Gestão democrática da educação: desafios contemporâneos**. Petrópolis: Vozes, 1997.

STEVÃO, Christiane B. G. A. **Teoria da resposta do item : um estudo inicial dos dados GERES**. Dissertação (Mestrado em Educação) – Universidade Estadual de Campinas, Campinas, 2008.

STRAUSS, Valerie. Leading mathematician debunks 'value-added'. **The Washington Post**, Washigton D.C., 05 set 2011. The Answer Sheet. Disponível em: http://www.washingtonpost.com/blogs/answer-sheet/post/leading-mathematician-debunks-value-added/2011/05/08/AFb999UG_blog.html. Acesso em: 10 out. 2013.

TABACHNICK, Bárbara G.; FIDELL, Linda S. **Using multivariate statistics**. Northridge, CA: Harper Collins College, 1996.

WILLMOTT, Robert. School effectiveness research: an ideological commitment?. **Journal of philosophy of education**, v. 33, n. 2, p. 253-268, 1999.

WILLMS, Jon D. A Estimação do Efeito Escola. In: BROOKE, Nigel e SOARES, José F., J. F. **Pesquisa em eficácia escolar: origem e trajetórias**. Belo Horizonte: Editora UFMG, 2008.

WILLMS, Jon D. **Monitoring School Performance**. Washington, D. C.: The Falmer Press, 1992.

WORTHINGTON, Roger; WHITTAKER, Tiffany. Scale development research. A content analysis and recommendations for best practices. **The Counseling Psychologist**, v.34, n.6, p. 806-838, Nov,2006.

XIMENES, Salomão B. Responsabilidade educacional: concepções diferentes e riscos eminentes ao direito à educação. **Educação & Sociedade**. Campinas: Cedes, v. 33, p. 353-378, n. 119, 2012.

APÊNDICES

Apêndice 1 - Arquivos resultantes da AFE e itens que compõem cada constructo

1) Arquivo Alunos

A Tabela 29 mostra as variáveis que compõe o arquivo que contém informações sobre os alunos e que resultou do processo de análise para redução de variáveis. Neste arquivo são mantidos os 35538 registros, porém, como o instrumento de coletou as informações sobre os alunos foi aplicado apenas na quinta onda, a quantidade de dados faltantes diminui em algumas das variáveis é diminuído em relação ao arquivo original, uma vez que, as variáveis que resultam da análise fatorial são compostas pela média aritmética dos valores dos itens que o compõe.

Tabela 29 - Dicionário de dados do Arquivo Alunos

Variável	Descrição
FA01	Sexo
FA02	Autodenominação da cor da pele do aluno
FA03	Anos de defasagem idade série
FA04	Indica se o aluno fez a pré-escola (variável Indicadora)
FA05	Indica a retenção do aluno pelo menos uma vez no processo
FA06	Nível socioeconômico dos alunos GERES
FA07	Estímulo externo para estudo
FA08	Estímulo interno para o estudo
FA09	Incentivo do Professor
FA10	Gosto pela escola
FA11	Estímulo interno para lição de casa (variável Indicadora)
FA12	Meus professores pensam que sou um aluno lento
FA13	Sou tipo de pessoa que sempre será capaz de conseguir, se tentar
FA14	Algumas vezes eu sinto que não sou bom o bastante

Obs: Variável FA01: 0 = Masculino e 1 = Feminino
Demais variáveis: 0 = Não e 1 = Sim

Essas 14 variáveis foram trabalhadas em quatro diferentes constructos, descritos a seguir.

1.01 - Construto 1: Características individuais do aluno

Este constructo é composto por cinco fatores, dimensionados diretamente do arquivo original, oferecido pelo Projeto GERES, que continha informações sobre os alunos. Esses fatores estão na Tabela 30.

Tabela 30 - Variáveis do Constructo 01 do arquivo Alunos obtidos diretamente do questionário

Variável	Descrição da variável	Item do questionário
FA01	Sexo	q001
FA02	Autodenominação da cor da pele do aluno	q002
FA03	Anos de defasagem idade série	q003
FA04	Indica se o aluno fez a pré-escola	q004
FA05	Indica a retenção do aluno pelo menos uma vez no processo	q009

1.02 - Construto 2: Nível socioeconômico

É composto pela variável FA06, que foi fornecida pelo Projeto GERES⁶². Neste trabalho é considerado como sendo um constructo, uma vez que foi determinada pela redução de diversos itens de questionário que tratavam sobre escolaridade, ocupação e a renda dos pais dos alunos. Esta última determinada pela identificação de itens de conforto.

Esta é uma importante variável, pois representa a condição socioeconômica dos alunos e sua determinação foi feita através da aplicação da Teoria de Resposta ao Item – TRI. Sua escolha deve-se ao fato de oferecer vantagens em relação a outras abordagens, como análise fatorial, pois permite a utilização de indicadores categóricos e apresenta uma tolerância maior com dados ausentes.

1.03 - Construto 3: Motivação para o estudo

Na Tabela 31 são apresentados os resultados da matriz de análise fatorial com as respectivas cargas fatoriais, bem como, os indicadores que, segundo critério descrito na metodologia, possibilitou a redução dos vinte e três itens em um conjunto de cinco fatores.

62 Para mais informações sobre a elaboração desse índice que quantifica o nível socioeconômico consultar Alves e Soares (2009)

Tabela 31 - Arquivo Alunos - Construto 3 - Motivação para o estudo

Variáveis	Cargas rotacionadas por Varimax					Comunalidades	Correlação Item-fator
	Fator						
	FA07	FA08	FA09	FA10	FA11		
o5_q36	0,739					0,568	0,723
o5_q35	0,715					0,546	0,713
o5_q34	0,708					0,515	0,683
o5_q31	0,681					0,482	0,671
o5_q33	0,660					0,445	0,647
o5_q29	0,653					0,439	0,648
o5_q32	0,625					0,487	0,634
o5_q30	0,518					0,434	0,548
o5_q26		0,799				0,683	0,841
o5_q28		0,761				0,643	0,841
o5_q23		0,684				0,501	0,650
o5_q42			0,832			0,709	0,838
o5_q43			0,804			0,683	0,853
o5_q41				0,847		0,735	0,819
o5_q40				0,742		0,643	0,851
o5_q25					0,746	0,607	0,805
o5_q24					0,706	0,559	0,809
KMO (p < 0,001)	0,868					Total	
Auto valor inicial	3,562	1,966	1,441	1,409	1,298	17	
% do Traço	20,955	11,562	8,478	8,29	7,637	56,92%	
Alfa de Cronbach	0,82	0,68	0,59	0,55	0,45		

Estas cinco fatores latentes juntas explicam aproximadamente 57% da variância total dos dados. O valor estimado de Kaiser-Meyer-Olkin = 0,87 e juntamente com outros indicadores apresentados na Tabela 31, indicando uma boa adequação dos fatores que representam o construto. Com relação ao coeficiente de confiabilidade Alpha de Cronbach os três últimos fatores ficaram abaixo do mínimo preconizado de 0,6, onde os resultados das escalas tendem a ser diferentes se aplicados novamente. Por outro lado se observarmos o resultado da correlação item-fator verifica-se que os mesmos ficam bem acima do mínimo preconizado de 0,5. Assim em futuros estudos recomenda-se melhorar a composição dos itens que compõe este construto, no sentido de manter no mínimo três em cada fator.

A correlação entre os cinco fatores situou-se entre 0,00 e 0,35, consideradas baixas no sentido dos fatores não estarem sobreestimados e de não apresentarem multicolinearidade quando utilizados na análise de rejeição como preditoras do desempenho acadêmico. Portanto, os fatores representam conceitos separados e independentes, características desejáveis para a finalidade que se propõe.

As Tabelas 32, 33, 34, 35 e 36 apresentam os itens do questionário que compõem cada uma das variáveis deste constructo.

Tabela 32 - Itens que compõem a variável FP07 – Estímulo externo para estudar

Item	Descrição dos itens do questionário
o5_q29	Eu estudo porque meus pais mandam
o5_q30	Eu faço o dever de casa por obrigação
o5_q31	Eu estudo porque meus pais prometem me dar presentes, se as minhas notas forem boas
o5_q32	Eu estudo porque minha professora acha importante
o5_q33	Eu só estudo para não me sair mal na escola
o5_q34	Eu estudo para que meus pais me deixem brincar com os meus amigos ou fazer as coisas que eu gosto
o5_q35	Eu só estudo para agradar meus professores
o5_q36	Eu só faço meu dever de casa (lição) porque meus pais acham importante

Tabela 33 - Itens que compõem a variável FP08 – Estímulo interno para o estudo

Item	Descrição dos itens do questionário
o5_q23	Eu estudo porque estudar é importante para mim
o5_q26	Eu gosto de estudar
o5_q28	Eu gosto de estudar porque me dá prazer e alegria

Tabela 34 - Itens que compõem a variável FP09 – Incentivo do professor

Item	Descrição dos itens do questionário
o5_q42	Minha professora incentiva os alunos a melhorar
o5_q43	Minha professora mostra interesse no aprendizado de todos os alunos

Tabela 35 - Itens que compõem a variável FP10 – Gosto pela escola

Item	Descrição dos itens do questionário
o5_q40	Eu me sinto bem na minha escola
o5_q41	Eu me divirto na escola

Tabela 36 - Itens que compõem a variável FP11 – Estímulo interno para lição de casa

Item	Descrição dos itens do questionário
o5_q24	Eu me esforço bastante nos trabalhos de casa, mesmo sabendo que não vão valer como nota
o5_q25	Eu faço minhas lições de casa, mesmo que meus pais não me peçam

Neste constructo, alguns dos itens do questionário tiveram que ser eliminados pois apresentaram carga cruzada. Estes itens estão descritos na Tabela 37.

Tabela 37 - Itens que foram excluídos do constructo 03 do arquivo de alunos

Item	Descrição dos itens do questionário
o5_q27	Participo das decisões relacionadas ao meu trabalho
o5_q37	Participo das decisões relacionadas ao meu trabalho
o5_q38	Participo das decisões relacionadas ao meu trabalho
o5_q39	Poucos professores assumem a responsabilidade de melhorar a escola

Os itens q27 e q39 apresentaram comunalidade 0,31 e 0,38 indicando uma fraca correlação com os cinco fatores. Os itens q37, q38 e q39 apresentaram cargas cruzadas, menores que 0,15, em dois fatores. O item q39 apresentou carga fatorial inferior a 0,40. Desta forma, por haver mais evidências contrárias, optou-se por eliminar os itens q27 e q39, segundo orientação dos autores, (WORTHINGTON e WHITTAKER, 2006), em que recomendam a retirada do item quando sua comunalidade for inferior a 0,4. Os itens q37 e q38 apresentaram carga fatorial relevante em dois fatores em que a diferença entre as cargas fatoriais foi inferior a 0,15. Mesmo que os autores Worthington e Whittaker (2006), recomendem que para esses casos a eliminação dos itens deve ser equacionada. No presente estudo, os autores entenderam que as questões não estão bem redigidas possibilitando dupla interpretação ao aluno. Assim optou-se por não incluí-las em nenhuma dimensão, para evitar dificuldades posteriores na sua interpretação como variáveis preditoras de desempenho acadêmico.

1.04 - Construto 4: Auto Avaliação

Este constructo é composto por três fatores dimensionados diretamente do arquivo original, oferecido pelo Projeto GERES, que continha informações sobre os alunos. Esses fatores estão na Tabela 38.

Tabela 38 - Variáveis do Constructo 01 do arquivo alunos obtidos diretamente do questionário

Variável	Descrição da variável	Item do questionário
FA12	Meus professores pensam que sou um aluno lento	O5_q44
FA13	Autoconfiança	O5_q45
FA14	Algumas vezes eu sinto que não sou bom o bastante	O5_q46

1) Arquivo Professores

A Tabela 39 traz as variáveis que resultaram do processo de redução de variáveis.

Tabela 39 - Dicionário de dados do arquivo Professores

Variável	Descrição
FP01	Relação com o diretor
FP02	Boa Interação entre a equipe
FP03	Pouco Trabalho Coletivo
FP04	Boa Relação entre professores
FP05	Unidade de propósito entre os professores
FP06	Violência no ambiente escolar
FP07	Frequência do dever de casa
FP08	Expectativas negativas em relação à educação
FP09	Expectativas positivas em relação à educação
FP10	Acesso à atividades culturais
FP11	Prática pedagógica de leitura
FP12	Proposição de escrita
FP13	Proposição de cópia ou caligrafia
FP14	Proposição atividade leitura individual é precedida de explicação
FP15	Proposição de leitura em prática coletiva
FP16	Proposição de atividade individual de leitura e produção de texto
FP17	Ensino de matemática com ênfase na resolução de problemas
FP18	Ensino de matemática com ênfase na reprodução
FP19	Usa a premiação para estimular o estudo
FP20	Usa a diversificação de atividades para estimular o estudo
FP21	Valoriza publicamente o desempenho
FP22	Estimula pelo sucesso
FP23	Adequação do livro ao ensino
FP24	Importância da indicação externa do livro
FP25	Tempo usado nas disciplinas relacionadas com as proficiências
FP26	Tempo usado nas disciplinas não relacionadas com as proficiências
FP27	Tempo usado em disciplinas de formação complementar
FP28	Interrupção das aulas por indisciplina
FP29	Interrupção das aulas por necessidade administrativa
FP30	Uso de recursos audiovisuais
FP31	Uso de laboratório e multimídia
FP32	Uso de materiais concretos de matemática e geografia
FP33	Trabalho coletivo
FP34	Grau de inclusão do professor com o trabalho escolar
FP35*	Sexo dos professores (variável Indicadora)
FP36*	Idade media dos professores
FP37*	Renda Familiar média dos professores
FP38*	Tempo de exercício médio na ocupação
FP39*	Numero de familiares que moram com o professor

*Variáveis com conteúdo obtido diretamente dos questionários de contexto.

Os estudos que usarão a regressão linear para analisar esses fatores deverão excluir as variáveis FP33 e FP34, pois o coeficiente de correlação entre elas é maior que 0,95 e, portanto FP01 explica a mesma coisa que FP33 e FP03 explica a mesma coisa que FP34. Além disso, atenções especiais deverão ser colocadas nos fatores FP11 e FP12 que são fortemente correlacionados (0,72), sugerindo que as duas práticas poderiam ser juntadas, No entanto, constata-se uma correlação alta e inversa (-0,744) com o gênero. Outra sugestão é a exclusão dos fatores FP19, FP20, FP21 e FP22 por terem mais de 6% de valores faltantes.

Essas 39 variáveis resultam do trabalho com 17 diferentes constructos descritos a seguir:

2.01 - Construto 01: Visão da atuação do diretor pelo professor

Na Tabela 40 são apresentados os principais resultados da matriz de análise fatorial com as respectivas cargas fatoriais, bem como, os indicadores que, segundo critério descrito na metodologia, possibilitou a redução dos nove itens em um único fator.

Tabela 40 - Arquivo Professores- Construto 01 - Visão da atuação do diretor pelo professor

Variáveis	Cargas rotacionadas por Varimax		Comunidades	Correlação Item-fator
	Fator			
	FP01			
q001	0,879		0,773	0,885
q002	0,907		0,823	0,911
q003	0,847		0,718	0,852
q004	0,887		0,786	0,887
q005	0,876		0,767	0,879
q006	0,756		0,572	0,758
q007	0,776		0,603	0,774
q008	0,833		0,693	0,826
q009	0,676		0,457	0,651
KMO (p < 0,001)	0,939		Total	
Auto valor inicial	6,19		9	
% do Traço	68,79%		68,79%	
Alfa de Cronbach	0,94			

O construto “Visão da atuação do diretor pelo professor” ficou representado por um único fator e mede o nível de legitimidade / reconhecimento da direção. Este fator latente, explica aproximadamente 69% da variância total dos dados. O valor estimado de Kaiser-Meyer-Olkin = 0,94 juntamente com outros indicadores, apresentados na Tabela 40, indicam uma adequação muito satisfatória do fator que representa o construto. O Alpha de Cronbach e as correlações item-fator indicam uma excelente confiabilidade da escala devendo apresentar resultados muito semelhantes em caso de novas aplicações.

A Tabela 41 apresenta os itens de questionários que compõe o único fator deste constructo.

Tabela 41 - Itens que compõem a variável FP01 – Visão da atuação do diretor pelo professor

Item	Descrição dos itens do questionário
q001	O diretor me anima e motiva para o trabalho
q002	Tenho plena confiança profissional no(a) diretor(a)
q003	O diretor consegue que os professores se comprometam com a escola
q004	O diretor estimula as atividades inovadoras
q005	O diretor dá atenção especial aos aspectos relacionados à aprendizagem dos alunos
q006	O diretor dá atenção especial aos aspectos relacionados às normas administrativas
q007	O diretor dá atenção especial aos aspectos relacionados à manutenção da escola
q008	Sinto-me respeitado pelo diretor
q009	Respeito o diretor

2.02 - Construto 02: Visão de metas compartilhadas

Na Tabela 42 são apresentados os principais resultados da matriz de análise fatorial com as respectivas cargas fatoriais, bem como, os indicadores que, segundo critério descrito na metodologia, possibilitou a redução dos 21 itens em um conjunto de quatro fatores.

Estes quatro fatores latentes juntos explicam aproximadamente 61% da variância total dos dados. O valor estimado de Kaiser-Meyer-Olkin = 0,87 e juntamente com outros indicadores apresentados na Tabela 42 indicam uma adequação muito satisfatória dos fatores que representam o construto. Para os três primeiros fatores o Alpha de Cronbach e as correlações item-fator indicam uma excelente confiabilidade das escalas tendendo a apresentar os mesmos resultados se aplicados novamente. O fator FP05 apresenta o Alpha de Cronbach

baixo e a correlação item-fator alta, neste caso, em estudos futuros recomenda-se aumentar para no mínimo três o número dos itens que compõe este fator.

Tabela 42 - Arquivo Professores - Construto 03 - Visão de metas compartilhadas

Variáveis	Cargas rotacionadas por Varimax				Comunalidades	Correlação Item-fator
	Fator					
	FP02	FP03	FP04	FP05		
q025	0,791				0,723	0,784
q015	0,730				0,644	0,759
q014	0,713				0,601	0,779
q024	0,710				0,600	0,750
q027	0,647				0,548	0,672
q032	0,560				0,455	0,718
q031		0,792			0,663	0,785
q034		0,751			0,628	0,791
q030		0,714			0,602	0,736
q033		0,665			0,502	0,721
q026		0,626			0,505	0,708
q035		0,531			0,329	0,632
q012			0,856		0,763	0,882
q029_05			0,813		0,769	0,839
q011			0,796		0,746	0,901
q028_05			0,759		0,717	0,870
q019_05				0,834	0,711	0,771
q018_05				0,647	0,502	0,837
KMO (p < 0,001)	0,871				Total	
Auto valor inicial	6,48	1,97	1,56	0,99	18	
% do Traço	19,17	17,73	16,66	7,58	61,16%	
Alfa de Cronbach	0,83	0,82	0,87	0,43		

A correlação entre os quatro fatores ($r_{(2;3)} = -0,53$; $r_{(2;4)} = 0,53$; $r_{(2;5)} = 0,38$; $r_{(3;4)} = -0,35$; $r_{(3;5)} = -0,17$ e $r_{(4;5)} = 0,25$) ficou relativamente alta. Desta forma, segundo os autores Cramer, (2003); Hair *et al.*, (2006), o resultado da matriz de estrutura estão sobreestimados. Não é difícil de ver que boa interação entre equipes e pouco trabalho coletivo medem bastante da mesma coisa, porém em sentidos opostos, da mesma forma com boa interação entre equipes e boa relação entre professores, neste caso estão no mesmo sentido, mas os fatores num certo sentido estão sobrepostos.

O agrupamento dos itens do construto visão de metas compartilhadas, na concepção do professor, indica uma estrutura de quatro fatores o fator FP02 representa a boa interação entre

a equipe, o fator FP03 representa o pouco trabalho coletivo, o fator FP04 mede a boa relação entre professores e o fator FP05 representa à unidade de propósito entre os professores.

As Tabelas 43, 44, 45 e 46 apresentam os itens do questionário que compõem cada uma das variáveis deste constructo.

Tabela 43 - Itens que compõem a variável FP02 – Boa Interação entre a equipe

Item	Descrição dos itens do questionário
q014	Os professores desta escola se esforçam para coordenar o conteúdo das matérias entre as séries
q015	Os diretores, professores e demais membros da equipe da escola colaboram para a escola funcionar bem
q024	A maioria dos professores é receptiva à implementação de novas idéias
q025	A maior parte dos professores está empenhada em melhorar suas aulas
q027	A maioria dos professores sente-se responsável pelo desempenho dos alunos
q032	O projeto educacional desta escola é consequência da troca de idéias entre os professores

Tabela 44 - Itens que compõem a variável FP03 – Pouco Trabalho Coletivo

Item	Descrição dos itens do questionário
q026	Nesta escola, poucos professores trocam idéias e experiências de modo a viabilizar que todos alunos aprendam
q030	Nesta escola, tenho dificuldade em compartilhar minhas preocupações e frustrações profissionais com outros professores
q031	Nesta escola, tenho poucas oportunidades em discutir idéias sobre ensino-aprendizagem
q033	O conteúdo programático entre diferentes séries não é planejado em equipe
q034	Nesta escola, tenho poucas oportunidades em discutir o conteúdo programático da minha turma com a equipe da escola
q035	Existem muitos projetos nesta escola e eu não consigo ter uma visão geral deles

Tabela 45 - Itens que compõem a variável FP04 – Boa relação entre os professores

Item	Descrição dos itens do questionário
q11	A equipe de professores leva em consideração minhas idéias
q12	Eu levo em consideração as idéias de outros colegas
q28_05	A equipe de professores leva em consideração as minhas idéias
q29_05	Eu levo em consideração as idéias dos meus colegas

Tabela 46 - Itens que compõem a variável FP05 – Unidade de propósitos entre os professores

Item	Descrição dos itens do questionário
Q18_05	No início do EF a tarefa mais importante desta escola é a socialização dos alunos
Q19_05	Nesta escola, no início do EF, a maior parte do tempo de sala de aula é dedicada ao aprendizado da leitura e escrita

Neste constructo, alguns dos itens do questionário tiveram que ser eliminados por apresentarem carga cruzada. Estes itens estão descritos na Tabela 47.

Tabela 47 - Itens que foram excluídos do constructo 01 do arquivo de professores

Item	Descrição dos itens do questionário
q010	Participo das decisões relacionadas ao meu trabalho O ensino que a escola oferece aos alunos é muito
q013	influenciado pela troca de idéias entre professores
q021	Poucos professores assumem a responsabilidade de melhorar a escola

O item q010 apresentou comunalidade igual a 0,301 e carga cruzada. Os itens q013 e q021 apresentaram carga cruzada. Desta forma, procedeu-se a análise novamente sem incluir o item q010, os itens q013 e q021 tornaram a apresentar carga cruzada, desta foram eliminados e a análise refeita. Assim optou-se por não incluí-las em nenhuma dimensão, para evitar dificuldades posteriores na sua interpretação como variáveis preditoras de desempenho acadêmico. Mesmo que a estrutura fatorial apresentada por nós tenha passado pelo teste da estabilidade em quatro subamostras, é interessante reestudar os itens apresentados neste construto no sentido de verificar até que ponto eles não estão, de certa forma, contribuindo para duas ou mais dimensões. Pois além de ter excluídos estes três itens outros apresentaram cargas cruzadas, q026, q027, q028_05 e q032, porém com diferenças maiores que 0,15.

2.03 - Construto 03: Violência no ambiente escolar

Na Tabela 48 são apresentados os principais resultados da matriz de análise fatorial com as respectivas cargas fatoriais, bem como, os indicadores que, segundo critério descrito na metodologia, possibilitou a redução dos 13 itens em apenas um fator.

Tabela 48 - Arquivo Professores- Construto 03 - Violência no ambiente escolar

Variáveis	Cargas rotacionadas por Varimax		Comuna- lidades	Correlação Item-fator
	Fator FP06			
q047	0,863		0,744	0,840
q048	0,858		0,737	0,839
q046	0,808		0,653	0,815
q039	0,803		0,645	0,792
q045	0,800		0,640	0,810
q041	0,732		0,536	0,740
q037	0,690		0,476	0,706
q038	0,686		0,470	0,691
KMO ($p < 0,001$)	0,918		Total	
Auto valor inicial	4,9		9	
% do Traço	61,26%		61,26%	
Alfa de Cronbach	0,905			

O agrupamento dos itens do construto clima organizacional, na concepção do professor, indica uma estrutura de um fator o fator FP06 violência e depredação no ambiente escolar.

Este fator latente explica aproximadamente 61% da variância total dos dados. O valor estimado de Kaiser-Meyer-Olkin = 0,92 e juntamente com outros indicadores apresentados na Tabela 48 apresentam uma adequação muito satisfatória dos fatores que representam o construto.

Os itens deste constructo estão relacionados na Tabela 49.

Tabela 49 - Itens que compõem a variável FP06 – Violência no ambiente escolar

Item	Descrição dos itens do questionário
q047	Em que medida o consumo de drogas dentro da escola impede o melhoramento dela?
q048	Em que medida a interferência do tráfico de drogas dentro da escola impede o melhoramento dela?
q046	Em que medida a interferência do tráfico de drogas impede o melhoramento da escola?
q039	Em que medida a violência física contra professores e funcionários impede o melhoramento desta escola?
q045	Em que medida o consumo de drogas nas proximidades da escola impede o melhoramento dela?
q041	Em que medida furto ou roubos de equipamentos da escola impede o melhoramento dela?
q037	Em que medida a violência física contra alunos impede o melhoramento desta escola?
q038	Em que medida a intimidação a professores e funcionários impede o melhoramento desta escola?

Neste constructo, algumas variáveis precisaram ser excluídas devido às cargas cruzadas. Estas variáveis estão relacionadas na Tabela 50.

Itens que foram excluídos do fator q036, q043, q044, q042 e q040. Pois, na hipótese de se extrair dois fatores, existiam itens com cargas fatoriais relevantes ($\geq 0,30$) nos dois fatores e com diferença entre elas não significativa (cargas cruzadas $< 0,15$).

Tabela 50 - Itens que foram excluídos do constructo 03 do arquivo de professores

Item	Descrição dos itens do questionário
q036	Em que medida a intimidação a alunos impede o melhoramento desta escola?
q040	Em que medida a depredação de equipamentos da escola impede o melhoramento dela?
q042	Em que medida pichações na escola impede o melhoramento dela?
q043	Em que medida depredação de banheiros da escola impede o melhoramento dela?
q044	Em que medida depredação das dependências da escola impede o melhoramento dela?

2.04 - Construto 04: Frequência do dever de casa

O constructo “Frequência do dever de casa” ficou representada pelo fator FP08 agrupando a frequência em que passa o dever de casa e a proporção de alunos que o faz.

Este fator latente, explica aproximadamente 72% da variância total dos dados, apresentou um valor estimado de Kaiser-Meyer-Olkin = 0,50, indicando uma qualidade ruim na extração dos fatores.

Tabela 51 - Arquivo Professores- Construto 04 - Dever de casa

Variáveis	Cargas rotacionadas por Varimax		Comuna- lidades	Correlação Item-fator
	Fator			
		FP07		
q059ir	0,847		0,718	0,925
q060i	0,847		0,718	0,745
KMO ($p < 0,001$)	0,5		Total	
Auto valor inicial	1,436		2	
% do Traço	71,78%		71,78%	
Alfa de Cronbach	0,544			

Este constructo é composto pelos itens de questionário relacionados na Tabela 52.

Tabela 52 - Itens que compõem a variável FP07 – Reforço Positivo

Item	Descrição dos itens do questionário
q059ir	Com que frequência o professor passa dever de casa?
q060i	Quantos alunos do professor fazem dever de casa?

2.05 - Construto 05: Expectativas em relação à educação

Na Tabela 53 são apresentados os principais resultados da matriz de análise fatorial com as respectivas cargas fatoriais, bem como, os indicadores que, segundo critério descrito na metodologia, possibilitou a redução dos cinco itens em um conjunto de dois fatores.

O arranjo dos itens do construto expectativas, na concepção do professor, ficou representado por dois fatores o FP08 que representa a expectativa positiva e o grau de comprometimento do mesmo com o aprendizado do aluno e a qualidade da aula e o fator FP09 que representa o grau de expectativa negativa em relação ao aprendizado do aluno, na visão dos professores.

Tabela 53 - Arquivo Professores - Construto 05 - Expectativas em relação à educação

Variáveis	Cargas rotacionadas por Varimax		Comuna- lidades	Correlação Item-fator
	Fator			
	FP08	FP09		
q022		0,980	0,961	1,000
q025_06	0,743		0,564	0,731
q020	0,692		0,503	0,779
q032_06	0,690		0,501	0,681
q017_06	0,565		0,302	0,599
KMO (p < 0,001)	0,683		Total	
Auto valor inicial	1,834	1,014	5	
% do Traço	36,52	20,44	56,96%	
Alfa de Cronbach	0,6			

Estes dois fatores latentes juntas explicam aproximadamente 57% da variância total dos dados. O valor estimado de Kaiser-Meyer-Olkin = 0,68 e juntamente com os outros indicadores apresentados na Tabela 53, indicam uma adequação mediana dos fatores que

representam o construto. O primeiro fator apresenta Alpha de Cronbach baixo e correlação item-fator alta, o segundo fator apresenta um item somente, neste caso, em estudos futuros recomenda-se melhorar a distribuição dos itens entre os fatores, cada fator deverá ter no mínimo três itens, (TABACHNICK e FIDELL, 2007). A comunalidade do item q017 e o Alpha de Cronbach baixos devem ter ocorrido pelo fato de que alguns itens foram respondidos somente em um ano da coleta e outros durante 4 anos, desta forma a amostra não é igual para todos os itens.

Este constructo é composto por duas variáveis que resultaram do processo de análise fatorial e explicam aproximadamente 57% da variância total dos dados e a segunda variável é composta por um único item de questionário como mostram as Tabelas 54 e 55.

Tabela 54 - Itens que compõem a variável FP08 – Expectativas negativas em relação à educação

Item	Descrição dos itens do questionário
q017_06	Com todos os atrativos que as crianças podem ter acesso hoje em dia, é muito difícil para a escola fazer seu trabalho
q020	Diante das dificuldades desta escola, um pequeno aprendizado dos alunos já é um bom resultado
q025_06	Para que os alunos desta escola pudessem realmente aprender seria necessário que a educação fosse levada mais a sério neste país
q032_06	Com as famílias que os alunos dessa escola têm, o aprendizado fica muito comprometido

Tabela 55 - Itens que compõem a variável FP09 – Expectativas positivas em relação à educação

Item	Descrição dos itens do questionário
q022	A maioria dos professores mantém altas expectativas sobre o aprendizado dos alunos

Neste construto, por ser formado por poucos itens, optou-se por não excluir nenhum deles.

2.06 - Construto 06: Acesso às atividades culturais

Os principais resultados da matriz de análise fatorial são apresentados na Tabela 56, bem como, os principais indicadores que possibilitaram a redução dos seis itens em um fator somente.

No construto perfil do professor, na própria concepção, ficou representado por um fator somente, FP10, e representa acesso a atividades culturais.

Este fator latente explica aproximadamente 46% da variância total dos dados. O valor estimado de Kaiser-Meyer-Olkin = 0,81e juntamente com outros indicadores apresentados na Tabela 56 indicam uma adequação boa do fator que representa o construto. O Alpha de Cronbach e as correlações item-fator indicam uma boa confiabilidade das escalas tendendo a apresentar resultados semelhantes em caso de reaplicá-las.

Tabela 56 - Arquivo Professores- Construto 06 - Acesso às atividades culturais

Variáveis	Cargas rotacionadas por Varimax		Comunidades	Correlação Item-fator
	Fator			
	FP10			
q119_05	0,695		0,483	0,74
q120_05	0,759		0,577	0,739
q121_05	0,635		0,403	0,57
q122_05	0,64		0,409	0,606
q123_05	0,697		0,485	0,687
q124_05	0,615		0,379	0,651
KMO (p < 0,001)	0,812		Total	
Auto valor inicial	2,736		6	
% do Traço	45,59		45,59%	
Alfa de Cronbach	0,749			

Este constructo, composto por uma variável que explica aproximadamente 46% da variância total dos dados, foi reduzido a partir de seis itens de questionário que estão relacionados na Tabela 57.

Tabela 57 - Itens que compõem a variável FP10 – Acesso à atividades culturais

Item	Descrição dos itens do questionário
q119_05	Nos últimos 12 meses, com que frequência o professor foi ao cinema?
q120_05	Nos últimos 12 meses, com que frequência o professor foi ao teatro?
q121_05	Nos últimos 12 meses, com que frequência o professor foi a uma ópera ou a um concerto de música clássica?
q122_05	Nos últimos 12 meses, com que frequência o professor foi a um balé ou espetáculo de dança?
q123_05	Nos últimos 12 meses, com que frequência o professor visitou museus?
q124_05	Nos últimos 12 meses, com que frequência o professor foi à livraria?

Neste construto deve ser observado que as condições em que os itens estão contemplando não são iguais de pólo para pólo, pois em alguns estas atividades são minimamente disponível enquanto que em outros são de uso comum. Este fato contribui fortemente para que os indicadores da matriz da análise fatorial não sejam bons.

2.07 - Construto 07: Prática pedagógica de leitura

Os principais resultados da matriz de análise fatorial do construto prática de leitura são apresentados na Tabela 58, bem como, os principais indicadores que possibilitaram a redução dos sete itens em um fator somente.

Tabela 58 - Arquivo Professores- Construto 07 - Prática pedagógica de leitura

Variáveis	Cargas rotacionadas por Varimax		Comuna- lidades	Correlação Item-fator
	Fator	FP11		
q067	0,926		0,858	0,927
q068	0,873		0,762	0,868
q071	0,861		0,742	0,858
q072	0,859		0,738	0,855
q073	0,852		0,726	0,856
q069	0,842		0,71	0,844
q070	0,79		0,624	0,786
KMO ($p < 0,001$)	0,924		Total	
Auto valor inicial	5,16		7	
% do Traço	73,71		73,71	
Alfa de Cronbach	0,94			

Este construto, na concepção do professor, ficou representado por um fator FP11 e representa a proposição de leitura.

Este fator latente explica aproximadamente 74% da variância total dos dados. O valor estimado de Kaiser-Meyer-Olkin = 0,92 e juntamente com outros indicadores apresentados na Tabela 58 indicam uma adequação muito satisfatória do fator que representa o construto.

Este constructo, composto por uma variável que explica aproximadamente de 74% da variância total dos dados, foi reduzido a partir de sete itens de questionário que estão relacionados na Tabela 59.

Tabela 59 - Itens que compõem a variável FP11 – Prática pedagógica de leitura

Item	Descrição dos itens do questionário
q069	Com que frequência houve leitura silenciosa pelos alunos do livro didático, neste ano?
q070	Com que frequência houve leitura silenciosa pelos alunos de textos escolhidos por eles, neste ano?
q071	Com que frequência houve leitura silenciosa pelos alunos de textos escolhidos pelo professor, neste ano?
q072	Com que frequência houve leitura oral, individual e alternada, neste ano?
q073	Com que frequência houve leitura coletiva em voz alta pelos alunos, neste ano?
q067	Com que frequência houve leitura em voz alta de história ou outros textos para os alunos desta turma, neste ano?
q068	Com que frequência contou-se uma história para os alunos desta turma, neste ano?

2.08 - Construto 08: Prática pedagógica de escrita

Os resultados da matriz de análise fatorial do construto prática de escrita são apresentados na Tabela 60, bem como, os principais indicadores que possibilitaram a redução dos seis itens em dois fatores.

O construto, na concepção do professor, ficou representado por dois fatores que representam a proposição de redação, FP12, e proposição de cópia ou caligrafia, FP13.

Tabela 60 - Arquivo Professores - Construto 08 - Prática pedagógica de escrita

Variáveis	Cargas rotacionadas por Varimax		Comuna- lidades	Correlação Item-fator
	Fator			
	FP12	FP13		
q074	0,724		0,615	0,771
q077	0,742		0,565	0,741
q078	0,887		0,788	0,848
q079	0,857		0,736	0,85
q075		0,714	0,658	0,851
q076_05		0,881	0,782	0,836
KMO ($p < 0,001$)	0,761		Total	
Auto valor inicial	2,956	1,188	6	
% do Traço	45,83	23,23	69,06%	
Alfa de Cronbach	0,809	0,514		

Estes fatores latentes explicam em torno de 69% da variância total dos dados e apresentou um valor estimado de Kaiser-Meyer-Olkin = 0,76 e juntamente com outros indicadores apresentados na Tabela 60, indicando uma adequação mediana do modelo dos fatores que representam o construto. O Alpha de Cronbach e as correlações item-total indicam uma confiabilidade muito boa da escala para o primeiro fator e tendendo a apresentar os mesmos resultados em caso de replicação. O fator dois apresenta o mesmo problema de ser constituído por um número muito pequeno de itens. A correlação entre os dois fatores ($r_{(12;13)} = 0,32$) ficou dentro de uma faixa aceitável.

Este constructo, composto por duas variáveis que explicam aproximadamente 69% da variância total dos dados, foi reduzido a partir de seis itens de questionário que estão relacionados nas Tabelas 61 e 62.

Tabela 61 - Itens que compõem a variável FP12 – Proposição de redação

Item	Descrição dos itens do questionário
q074	Com que frequência houve ditados, neste ano?
q077	Com que frequência houve redação de um texto sobre o tema escolhido pelos alunos, neste ano?
q078	Com que frequência houve redação de um texto sobre temas que o professor escolheu, neste ano?
q079	Com que frequência estudantes responderam por escrito perguntas feitas ao final da história lida, neste ano?

Tabela 62 - Itens que compõem a variável FP13 – Proposição de cópia e caligrafia

Item	Descrição dos itens do questionário
q075	Com que frequência houve cópia de textos, neste ano?
q076_05	Com que frequência houve exercício de caligrafia, neste ano?

Devido ao baixo número de itens decidiu-se manter todos, ainda que os itens q079 e q075 apresentassem cargas relevantes nos dois fatores e com diferença maior que 0,15. Desta forma recomenda-se que o construto ou os fatores do construto sejam reestudados para sua melhor utilização em futuras pesquisas.

2.09 - Construto 9: Ensino de Língua Portuguesa

Os resultados da matriz de análise fatorial do construto ensino da língua portuguesa são apresentados na Tabela 63, bem como, os principais indicadores que possibilitaram a redução dos doze itens em três fatores.

Este construto, na concepção do professor, ficou representado por três fatores o fator FP14 é aquela que a leitura individual é precedida da explicação do professor. O fator FP15 representa leitura com prática coletiva. O fator FP16 representa atividade de leitura e produção de texto individual.

Tabela 63 - Arquivo Professores - Construto 09 - Ensino da língua portuguesa

Variáveis	Cargas rotacionadas por Varimax			Comunalidades	Correlação Item-fator
	Fator				
	FP14	FP15	FP16		
q092_07	0,872			0,781	0,882
q093_07	0,864			0,756	0,875
q086_07	0,747			0,582	0,791
q096_07		0,697		0,517	0,681
q091_07		0,667		0,582	0,749
q089_07		0,662		0,518	0,687
q097_07		0,608		0,432	0,707
q088_07			0,774	0,641	0,802
q094_07			0,733	0,561	0,8
KMO (p < 0,001)	0,704			Total	
Auto valor inicial	2,547	1,635	1,188	9	
% do Traço	25,41	19,59	14,67	59,66	
Alfa de Cronbach	0,804	0,592	0,43		

Estes três fatores latentes juntas explicam aproximadamente 60% da variância total dos dados. O valor estimado de Kaiser-Meyer-Olkin = 0,70 e juntamente com outros indicadores apresentados na Tabela 63 indicam uma adequação mediana dos fatores que representam o construto. O Alpha de Cronbach e as correlações item-total indicam para o primeiro fator uma excelente confiabilidade da escala tendendo a apresentar os mesmos resultados em caso de replicação. O fator dois está muito próximo do limite inferior, porém as correlações item-fator estão bem acima do mínimo recomendado, 0,5, indicando ser o construto confiável. O fator três apresenta o mesmo problema de ser constituído por um número muito pequeno de itens, desta forma o alfa é baixo, mas a correlação é alta. Neste caso, ainda que como estudo exploratório o tenha sido mantido, é recomendável que este construto fosse repensado no sentido de melhorar a redação dos itens para que cada conjunto

de itens represente fatores independentes. A correlação entre os três fatores ficou na faixa de 0,15 a 0,30 e, portanto, numa faixa relativamente baixa, considerada aceitável.

Este constructo, composto por três variáveis que explicam aproximadamente 60% da variância total dos dados, foi reduzido a partir de catorze itens de questionário que estão relacionados nas Tabelas 64, 65 e 66.

Tabela 64 - Itens que compõem a variável FP14 – Proposição de leitura individual precedida de explicação

Item	Descrição dos itens do questionário
q086_07	No ensino de língua portuguesa dou prioridade para: Alunos lêem um texto e sublinham com cores diferentes substantivos e adjetivos. A atividade é individual
q092_07	No ensino de língua portuguesa dou prioridade para: Após explicação da professora, alunos lêem texto procurando identificar e sublinhar palavras com ditongos
q093_07	No ensino de língua portuguesa dou prioridade para: Após explicação da professora, alunos lêem texto procurando identificar e diferenciar palavras com ditongos crescentes e decrescentes

Tabela 65 - Itens que compõem a variável FP15 – Proposição de leitura em prática coletiva

Item	Descrição dos itens do questionário
q089_07	No ensino de língua portuguesa dou prioridade para: Alunos lêem um texto e, por escrito, respondem a perguntas sobre o texto. A atividade é em dupla
q091_07	No ensino de língua portuguesa dou prioridade para: Em grupos alunos participam de um jogo onde devem falar uma palavra com o mesmo som da sílaba inicial da palavra dita pela professora
q096_07	No ensino de língua portuguesa dou prioridade para: Os alunos lêem dois textos de gêneros diferentes sobre o mesmo assunto, procurando estabelecer comparações
q097_07	No ensino de língua portuguesa dou prioridade para: A cada dificuldade ortográfica trabalhada, os alunos procuram palavras correspondentes em jornais e revistas

Tabela 66 - Itens que compõem a variável FP16 – Proposição de atividade individual de leitura e produção de texto

Item	Descrição dos itens do questionário
q088_07	No ensino de língua portuguesa dou prioridade para: Alunos lêem um texto e, por escrito, respondem a perguntas sobre o texto. A atividade é individual
q094_07	No ensino de língua portuguesa dou prioridade para: A professora pede aos alunos que escrevam uma redação sobre um determinado tema

Neste constructo, algumas variáveis precisaram ser excluídas devido às cargas cruzadas. Estas variáveis estão relacionadas na Tabela 67.

Tabela 67 - Itens que foram excluídos do constructo 09 do arquivo de professores

Item	Descrição dos itens do questionário
q087_07	No ensino de língua portuguesa dou prioridade para: Um aluno lê em voz alta um texto. Os demais acompanham. A professora pede que outro aluno continue a leitura
q090_07	No ensino de língua portuguesa dou prioridade para: Alunos usam revistas em sala de aula. Com a tesoura recortam exemplos de substantivos e adjetivos. Cada grupo faz cartazes, um com os exemplos de substantivos e outro com exemplos de adjetivos
q095_07	No ensino de língua portuguesa dou prioridade para: Alunos copiam, no caderno, texto do livro didático
q098_07	No ensino de língua portuguesa dou prioridade para: Após a explicação da professora, alunos copiam no caderno as definições de substantivo e adjetivo
q099_07	No ensino de língua portuguesa dou prioridade para: O professor seleciona palavras com questões ortográficas mais complexas e pede aos alunos que listem palavras da mesma família, para que observem a coincidência ortográfica

Itens que foram excluídos dos fatores q087, q090, q095, q098 e q099. Os itens q090, q095, q098 e q099 foram excluídos por apresentar cargas fatoriais relevantes ($\geq 0,30$) em dois fatores e com diferença entre elas não significativa (cargas cruzadas $< 0,15$). O item q087 foi extraído pelo fato da estrutura fatorial não passar no teste de estabilidade aplicado em subamostras e também por apresentar comunalidade 0,37 abaixo do mínimo recomendado.

2.10 - Construto 10: Ensino de Matemática

Na Tabela 68 são apresentados os principais resultados da matriz de análise fatorial com as respectivas cargas fatoriais, bem como, os indicadores que, segundo critério descrito na metodologia, possibilitou a redução dos 14 itens em um conjunto de dois fatores.

O agrupamento dos itens do construto práticas utilizadas no ensino da matemática, na concepção do professor, indica uma estrutura de dois fatores o fator FP17 representa uso de metodologia de resolução de problemas e o fator FP18 representa práticas de ensino com ênfase em memorização.

Tabela 68 - Arquivo Professores - Construto 10 - Ensino de Matemática

Variáveis	Cargas rotacionadas por Varimax		Comuna- lidades	Correlação Item-fator
	Fator			
	FP17	FP18		
q103_07_06	0,777		0,619	0,755
q102_07_06	0,72		0,541	0,727
q101_07_06	0,717		0,515	0,714
q113_07_06	0,683		0,493	0,727
q106_07_06	0,681		0,486	0,662
q105_07_06	0,638		0,541	0,682
q108_07_06	0,592		0,367	0,648
q111_07_06		0,837	0,702	0,829
q104_07_06		0,781	0,649	0,808
q110_07_06		0,632	0,501	0,696
KMO (p < 0,001)	0,864		Total	
Auto valor inicial	4,03	1,381	10	
% do Traço	34,66	19,48	54,13	
Alfa de Cronbach	0,823	0,672		

Estes dois fatores latentes juntos explicam aproximadamente 54% da variância total dos dados. O valor estimado de Kaiser-Meyer-Olkin = 0,86 e juntamente com outros indicadores apresentados na Tabela 68, como carga fatorial, indicam que os fatores apresentam uma adequação muito satisfatória do construto. O Alpha de Cronbach e as correlações item-fator indicam para os dois fatores uma confiabilidade da escala boa tendendo a apresentar os mesmos resultados em caso de replicação. Esta estrutura passou pelo teste de estabilidade em subamostras e, portanto, representa com boa precisão o conceito de interesse. A correlação entre os dois fatores ($r_{(17;18)} = 0,401$) num limite minimamente aceitável.

Este constructo, composto por duas variáveis que explicam aproximadamente 54% da variância total dos dados, foi reduzido a partir de catorze itens de questionário que estão relacionados nas Tabelas 69 e 70.

Tabela 69 - Itens que compõem a variável FP17 – Ensino de Matemática com ênfase na resolução de problemas

Item	Descrição dos itens do questionário
q101_06_07	contaram para os colegas sobre como eles tentam resolver um problema?
q102_06_07	lidaram com problemas que exigem raciocínios diferentes ou mais complexos?
q103_06_07	discutiram as possibilidades de resolução de um problema após resolvê-lo?
q105_06_07	lidaram com problemas que envolvam a representação de situações em linguagem matemática?
q106_06_07	lidaram com situações relacionadas ao cotidiano?
q108_06_07	os alunos trabalharam com temas que aparecem em jornais ou revistas, relacionando-os com a Matemática?
q113_06_07	apresentaram para os colegas as estratégias que eles utilizaram para resolver as questões do dever de casa?

Tabela 70 - Itens que compõem a variável FP18 – Ensino de Matemática com ênfase na reprodução

Item	Descrição dos itens do questionário
q104_07	lidaram com situações para memorizar conceitos e regras?
q110_06	lidaram com situações para memorizar conceitos e regras?
q111_07	decoraram regras e aplicaram-nas em situações-problema?

Neste constructo, algumas variáveis precisaram ser excluídas devido às cargas cruzadas. Estas variáveis estão relacionadas na Tabela 71.

Tabela 71 - Itens que foram excluídos do constructo 10 do arquivo de professores

Item	Descrição dos itens do questionário
Q100_07_06	trabalharam com jogos e quebra-cabeça?
Q107_07_06	lidaram com problemas que envolvam definições ou reconhecimento de fatos?
Q109_07_06	lidaram com situações para fixar procedimentos?
Q112_07_06	lidaram com problemas que envolvem cálculos ou algoritmos?

Os itens q100 e q112 baixa comunalidade. Os itens q107, q109 e q112 foram excluídos por apresentar cargas fatoriais relevantes ($\geq 0,30$) em dois fatores. Itens que violaram este critério, mas foram mantidos pelos autores, q105 e q110. Para isso levou-se em consideração outros aspectos, semântica e estabilidade em subamostras. Destaca-se que este construto foi aplicado em duas ondas e, portanto, não foi respondido por todos os professores, de qualquer

forma recomenda-se que seja repensado para futuras aplicações, de forma que o conjunto de itens em cada fator possa representar conceitos separados.

2.11 - Construto 11: Prática Pedagógica

Os resultados da matriz de análise fatorial do construto práticas pedagógicas são apresentados na Tabela 72, bem como, os principais indicadores que possibilitaram a redução dos 16 itens em quatro fatores.

Este construto, na concepção do professor, ficou representado por quatro fatores o fator FP19 onde o professor dá prêmio para estimular o estudo. O fator FP20 é aquela que o professor estimula as atividades práticas com inovação. O fator FP21 é aquele que o professor valoriza publicamente o desempenho. O fator FP22 é aquele que o professor estimula pelo sucesso.

Tabela 72 - Arquivo Professores - Construto 11 - Prática Pedagógica

Variáveis	Cargas rotacionadas por Varimax				Comunidades	Correlação Item-fator
	Fator					
	FP19	FP20	FP21	FP22		
q058_08	0,905				0,833	0,904
q059_08	0,897				0,825	0,898
q063_08	0,792				0,672	0,808
q060_08		0,755			0,600	0,723
q061_08		0,685			0,480	0,710
q062_08		0,669			0,459	0,677
q057_08		0,659			0,462	0,592
q065_08			0,826		0,702	0,823
q064_08			0,737		0,563	0,708
q066_08			0,634		0,491	0,693
q071_08				0,900	0,836	0,931
q070_08				0,879	0,816	0,913
KMO ($p < 0,001$)	0,690				Total	
Auto valor inicial	3,11	2,02	1,41	1,19	12	
% do Traço	19,38	16,99	14,36	13,75	64,49%	
Alfa de Cronbach	0,84	0,6	0,57	0,81		

Estes quatro fatores latentes juntas explicam aproximadamente de 65% da variância total dos dados. O valor estimado de Kaiser-Meyer-Olkin = 0,69 e juntamente com outros

indicadores apresentados na Tabela 72 indicam uma boa adequação dos fatores que representam o construto. O Alpha de Cronbach e as correlações item-fator indicam uma boa confiabilidade da escala, tendendo a apresentar os mesmos resultados em caso de replicação. Esta estrutura passou pelo teste de estabilidade em subamostras e, portanto, representa com boa precisão o conceito de interesse. A única ressalva é que os itens foram respondidos apenas pelos professores que trabalharam no quarto ano de escolarização e o último fator ficou apenas com dois itens quando se recomenda no mínimo três. A correlação entre os quatro fatores ficou na faixa de 0,05 a 0,30 e, portanto, numa faixa relativamente baixa, indicativo de que os fatores indicam conceitos independentes.

Este constructo, composto por quatro variáveis que explicam aproximadamente 65% da variância total dos dados, foi reduzido a partir de dezoito itens de questionário que estão relacionados nas Tabelas 73, 74, 75 e 76.

Tabela 73 - Itens que compõem a variável FP19 – Usa a premiação para estimular o estudo

Item	Descrição dos itens do questionário
q058_08	prêmio por trabalhos de aula para estimular o estudo
q059_08	prêmio por trabalhos de casa para estimular o estudo
q063_08	prêmio por resultado de provas para estimular o estudo

Tabela 74 - Itens que compõem a variável FP20 – Usa a diversificação de atividades para estimular o estudo

Item	Descrição dos itens do questionário
q057_08	atividades que despertem curiosidade
q060_08	atividades que lidem com novidades
q061_08	atividades de livros paradidáticos
q062_08	atividades práticas ou experiências

Tabela 75 - Itens que compõem a variável FP21 – Valoriza publicamente o desempenho

Item	Descrição dos itens do questionário
q064_08	valoriza o desempenho do estudante perante a turma
q065_08	valoriza o desempenho do estudante perante a escola
q066_08	valoriza o desempenho do estudante perante a família

Tabela 76 - Itens que compõem a variável FP22 – Usa a mobilidade social para estimular o estudo

Item	Descrição dos itens do questionário
q057_08	atividades que despertem curiosidade
q060_08	atividades que lidem com novidades
q070_08	argumento da ascensão social para estimular o estudo
q071_08	argumento do sucesso financeiro para estimular o estudo

Neste constructo, algumas variáveis precisaram ser excluídas devido às cargas cruzadas. Estas variáveis estão relacionadas na Tabela 77.

Tabela 77 - Itens que foram excluídos do constructo 11 do arquivo de professores

Item	Descrição dos itens do questionário
q067_08	Utiliza atividades que estimulem a socialização
q068_08	Utiliza atividade de jogos
q069_08	Utiliza atividades de competição entre os estudantes
q072_08	Valoriza a autoestima para estimular o estudo

Todos os itens foram excluídos por apresentar cargas fatoriais relevantes ($\geq 0,30$) em dois ou mais fatores.

2.12 - Construto 12: Escolha de Livro Didático

Na Tabela 78 são apresentados os principais resultados da matriz de análise fatorial com as respectivas cargas fatoriais, bem como, os indicadores que, segundo critério descrito na metodologia, possibilitou a redução dos nove itens em um conjunto de dois fatores.

O agrupamento dos itens do construto à escolha do livro didático, na concepção do professor, ficou representado por dois fatores FP23 adequação do livro ao ensino e FP24 grau de importância inerente a indicação externa do livro.

Estes dois fatores latentes juntas explicam aproximadamente 80% da variância total dos dados. O valor estimado de Kaiser-Meyer-Olkin = 0,94 e juntamente com outros indicadores apresentados na Tabela 78 indicam uma adequação excelente dos fatores que representam o construto. O Alpha de Cronbach e as correlações item-fator indicam para o fator FP23 uma excelente confiabilidade da escala, no fator FP24 o alfa se encontra no limite

mínimo de aceitabilidade devido provavelmente de ser composto por apenas dois itens. Desta forma recomenda-se incluir mais itens neste fator em futuros estudos.

Tabela 78 - Arquivo Professores - Construto 12 - Escolha de Livro Didático

Variáveis	Cargas rotacionadas por Varimax		Comuna- lidades	Correlação Item-fator
	Fator			
	FP23	FP24		
q092_05	0,918		0,884	0,938
q091_05	0,916		0,886	0,941
q089_05	0,894		0,838	0,913
q087_05	0,887		0,862	0,928
q088_05	0,864		0,81	0,907
q090_05	0,861		0,77	0,882
q093_05	0,838		0,756	0,872
q085_05		0,839	0,733	0,856
q086_05		0,792	0,683	0,849
KMO (p < 0,001)	0,941		Total	
Auto valor inicial	6,18	1,04	9	
% do Traço	61,56	18,69	80,26	
Alfa de Cronbach	0,966	0,591		

A correlação entre os dois fatores ($r_{(23;24)} = 0,49$) ficou relativamente alta, ainda que não tivéssemos encontrado cargas importantes do mesmo item em dois fatores. Portanto estes fatores guardam entre si uma relação de dependência.

Este constructo, composto por duas variáveis que explicam aproximadamente 80% da variância total dos dados, foi reduzido a partir de nove itens de questionário que estão relacionados nas Tabelas 79 e 80.

Tabela 79 - Itens que compõem a variável FP23 – Adequação do livro ao ensino

Item	Descrição dos itens do questionário
q087_05	os temas abordados no livro são importantes?
q088_05	a adequação do livro ao projeto pedagógico da escola é importante?
q089_05	a adequação do livro ao estilo de ensino adotado pelo professor é importante?
q090_05	a adequação do livro ao nível médio dos alunos é importante?
q091_05	o enfoque pedagógico e/ou metodológico do livro é importante?
q092_05	as atividades e exercícios propostos no livro são importantes?
q093_05	a orientação oferecida para o trabalho que o professor desenvolve é importante?

Tabela 80 - Itens que compõem a variável FP24 – Importância da Indicação externa do livro

Item	Descrição dos itens do questionário
q085_05	a indicação da obra em Guias de Livros Didáticos do MEC é importante?
q086_05	a indicação da obra por coordenadores, especialista ou técnica é importante?

Neste construto foram utilizados todos os itens.

2.13 - Construto 13: Planejamento do tempo para as disciplinas

Os resultados da matriz de análise fatorial do construto planejamento do tempo para as disciplinas são apresentados na Tabela 81, bem como, os principais indicadores que possibilitaram a redução dos 6 itens em três fatores.

Este construto, na concepção do professor, ficou representado por três fatores o fator FP25 representa o tempo nas disciplinas relacionadas com a proficiência. O fator FP26 representa o tempo usado em disciplinas não relacionadas com as proficiências. O fator FP27 representa o tempo usado em disciplinas de formação complementar.

Tabela 81 - Arquivo Professores - Construto 13 - Planejamento do tempo para as disciplinas

Variáveis	Cargas rotacionadas por Varimax			Comunalidades	Correlação Item-fator
	Fator				
	FP25	FP26	FP27		
q069_06	0,884			0,803	0,661
q070_06	0,874			0,811	0,649
q067_06		0,919		0,78	0,59
q068_06		0,857		0,837	0,675
q072_06			0,929	0,816	0,928
q071_06			0,822	0,834	0,920
KMO (p < 0,001)	0,751			Total	
Auto valor inicial	3,77	1,11	0,59	6	
% do Traço	31,63	30,92	28,67	91,23	
Alfa de Cronbach	0,964	0,893	0,825		

Estes três fatores latentes juntas explicam em torno de 91% da variância total dos dados. O valor estimado de Kaiser-Meyer-Olkin = 0,75 e juntamente com outros indicadores apresentados na Tabela 81, indicando uma adequação muito satisfatória dos fatores que

representam o construto. O Alpha de Cronbach e as correlações item-fator indicam para os fatores uma confiabilidade muito boa da escala, tendendo a apresentar os mesmos resultados em caso de replicação.

A correlação entre os fatores é alta ($r_{(25;26)} = 0,67$; $r_{(25;27)} = 0,38$ e $r_{(26;27)} = 0,56$). Este fato ocorre devido a cargas fatoriais importantes em dois fatores e que não foram eliminadas. Entendeu-se que cada fator representa conceitos (não sei o que) diferentes, ainda que aja sobreposição dos mesmos.

As três variáveis desse constructo que explicam aproximadamente 91% da variância total dos dados e foram reduzidas a partir de seis itens de questionário que estão relacionados nas Tabelas 82, 83 e 84.

Tabela 82 - Itens que compõem a variável FP25 – Tempo usado nas disciplinas não relacionadas com as proficiências

Item	Descrição dos itens do questionário
q069_06	A porcentagem do tempo de aula de história e geografia oferecida para os alunos desta turma
q070_06	A porcentagem do tempo de aula de ciências naturais oferecida para os alunos desta turma

Tabela 83 - Itens que compõem a variável FP26 – Tempo usado nas disciplinas relacionadas com as proficiências

Item	Descrição dos itens do questionário
q067_06	A porcentagem do tempo de aula de língua portuguesa oferecida para os alunos desta turma
q068_06	A porcentagem do tempo de aula de matemática oferecida para os alunos desta turma

Tabela 84 - Itens que compõem a variável FP27 – Tempo usado nas disciplinas de formação complementar

Item	Descrição dos itens do questionário
q071_06	A porcentagem do tempo de aula de artes oferecida para os alunos desta turma
q072_06	A porcentagem do tempo de aula de educação física oferecida para os alunos desta turma

Não foi eliminado nenhum item devido ao baixo número deles em cada fator.

2.14 - Construto 14: Interrupção de aulas

Na Tabela 85 são apresentados os principais resultados da matriz de análise fatorial com as respectivas cargas fatoriais, bem como, os indicadores que, segundo critério descrito na metodologia, possibilitou a redução dos cinco itens em um conjunto de dois fatores.

O agrupamento dos itens do construto, na concepção do professor, ficou representado por dois fatores que representam o fator FP28, interrupção das aulas por indisciplina, e a FP29, interrupção das aulas por necessidade administrativa.

Tabela 85 - Arquivo Professores - Construto 14 - Interrupção de aulas

Variáveis	Cargas rotacionadas por Varimax		Comunalidades	Correlação Item-fator
	Fator			
	FP28	FP29		
q065	0,818		0,673	0,756
q063	0,766		0,587	0,765
q066	0,651		0,522	0,774
q064		0,974	0,959	1
KMO ($p < 0,001$)	0,672		Total	
Auto valor inicial	1,865	0,876	4	
% do Traço	46,62	22,91	68,53	
Alfa de Cronbach	0,623			

Estes dois fatores latentes juntos explicam aproximadamente 58% da variância total dos dados. O valor estimado de Kaiser-Meyer-Olkin = 0,63 e juntamente com outros indicadores apresentados na Tabela 85, indicando uma adequação medíocre dos fatores que representam o construto. O primeiro fator apresenta Alpha de Cronbach minimamente aceitável e alta correlação item-fator indicando para este fator uma confiabilidade boa da escala o segundo fator apresenta um item somente, neste caso, em estudos futuros recomenda-se melhorar a distribuição dos itens entre os fatores. A correlação entre os dois fatores ($r_{(28;29)} = 0,24$) é relativamente baixa, o que nos possibilita afirmar que os fatores representam conceitos distintos.

As duas variáveis deste constructo explicam aproximadamente 58% da variância total dos dados e foram reduzidas a partir de quatro itens de questionário que estão relacionados nas Tabelas 86 e 87.

Tabela 86 - Itens que compõem a variável FP28 – Interrupção de aulas por indisciplina dos alunos

Item	Descrição dos itens do questionário
q063	bagunça dos alunos
q065	estudantes atrasados (na entrada, na volta do recreio)
q066	barulho no corredor

Tabela 87 - Itens que compõem a variável FP29 – Tempo usado nas disciplinas não relacionadas com as proficiências

Item	Descrição dos itens do questionário
q064	anúncios ou comunicações da direção, coordenação e/ou secretaria

Não houve eliminação de itens neste construto.

2.15 - Construto 15: Recursos técnico-pedagógicos

Os resultados da matriz de análise fatorial do construto planejamento do tempo para as disciplinas são apresentados na Tabela 88, bem como, os principais indicadores que possibilitaram a redução dos dez itens em três fatores.

Este construto, na concepção do professor, ficou representado por três fatores o fator FP30 é aquela que representa a intensidade de uso de recursos audiovisuais. O fator FP31 representa a intensidade de uso de laboratório multimídia. O fator FP32 representa o uso de materiais concretos de matemática e geografia.

Tabela 88 - Arquivo Professores - Construto 15 - Recursos técnico-pedagógicos

Variáveis	Cargas rotacionadas por Varimax			Comunalidades	Correlação Item-fator
	Fator				
	FP30	FP31	FP32		
q054	0,861			0,749	0,845
q056	0,852			0,745	0,845
q053	0,767			0,719	0,787
q055	0,708	0,312		0,609	0,766
q057		0,857		0,752	0,831
q051		0,789		0,628	0,757
q058		0,590		0,406	0,711
q049			0,827	0,704	0,867
q050			0,774	0,646	0,798
KMO (p < 0,001)	0,739			Total	
Auto valor inicial	2,963	1,748	1,246	9	
% do Traço	28,83	20,59	16,77	66,19%	
Alfa de Cronbach	0,816	0,623	0,543		

Estes três fatores latentes juntas explicam aproximadamente 66% da variância total dos dados. O valor estimado de Kaiser-Meyer-Olkin = 0,74 e juntamente com outros indicadores apresentados na Tabela 88 indicam uma adequação boa dos fatores que representam o construto. O Alpha de Cronbach e as correlações item-fator indicam uma boa confiabilidade da escala para o fator FP30, para os fatores FP31 e FP32 é complicado, pois o número de itens é muito pequeno, nesse caso o alfa é influenciado pela escala e portanto recomenda-se aumentar o número de itens neste construto. A correlação entre os fatores é relativamente baixa ($r_{(30;31)} = 0,194$; $r_{(30;32)} = 0,259$ e $r_{(31;32)} = 0,063$). Indicando valor um pouco mais elevado onde existem itens com carga cruzada.

Este constructo, composto por três variáveis que explicam aproximadamente 66% da variância total dos dados, foi reduzido a partir de dez itens de questionário que estão relacionados nas Tabelas 89, 90 e 91.

Tabela 89 - Itens que compõem a variável FP30 – Uso de recursos audiovisuais

Item	Descrição dos itens do questionário
q053	fitas de vídeo/DVD (educativas)
q054	fitas de vídeo/DVD (lazer)
q055	televisão
q056	vídeo cassete/DVD

Tabela 90 - Itens que compõem a variável FP31 – Uso de laboratório e multimídia

Item	Descrição dos itens do questionário
q051	terrário e aquário
q057	canhão multimídia e Datashow
q058	computador

Tabela 91 - Itens que compõem a variável FP32 – Uso de materiais concretos de matemática e geografia complementar

Item	Descrição dos itens do questionário
q049	material concreto de matemática
q050	mapas e globos

Neste constructo, apenas uma variável foi excluídas devido às cargas cruzadas e está descrita na Tabela 92.

Tabela 92 - Itens que foram excluídos do construto 11 do arquivo de professores

Item	Descrição dos itens do questionário
q052	Com que frequência o professor utiliza diagramas do corpo humano e modelos anatômicos?

2.16 - Construto 16: Trabalho Coletivo

Na Tabela 93 são apresentados os principais resultados da matriz de análise fatorial com as respectivas cargas fatoriais, bem como, os indicadores que, segundo critério descrito na metodologia, possibilitou a redução dos doze itens em um conjunto de dois fatores.

O agrupamento dos itens do construto, na concepção do professor, ficou representado por dois fatores o fator FP33 representa o clima existente em relação ao trabalho coletivo. O fator FP34 representa o grau de inclusão do professor em relação ao trabalho escolar.

Tabela 93 - Arquivo Professores - Construto 16 – Trabalho Coletivo

Variáveis	Cargas rotacionadas por		Comuna- lidades	Correlação Item-fator
	Varimax			
	Fator			
	FP33	FP34		
q002	0,894		0,81	0,9
q008	0,862		0,749	0,847
q001	0,859		0,754	0,876
q003	0,815		0,71	0,851
q009	0,738		0,546	0,696
q015	0,679		0,579	0,737
q010	0,609		0,393	0,637
q031		0,796	0,66	0,808
q030		0,761	0,597	0,768
q026		0,743	0,561	0,746
q034		0,692	0,532	0,739
q021		0,630	0,398	0,659
KMO (p < 0,001)	0,885		Total	
Auto valor inicial	5,147	2,142	12	
% do Traço	42,896	17,847	60,74	
Alfa de Cronbach	0,901	0,795		

Estes dois fatores latentes juntas explicam em torno de de 61% da variância total dos dados. O valor estimado de Kaiser-Meyer-Olkin = 0,89 juntamente com outros indicadores apresentados na Tabela 93, indicando uma adequação muito satisfatória dos fatores que representam o construto. O Alpha de Cronbach e as correlações item-fator indicam uma confiabilidade muito boa da escala, tendendo a apresentar os mesmos resultados em caso de

replicação. Existe uma correlação negativa apreciável entre os dois fatores de -0,35. O que não dá para dizer que são independentes.

As duas variáveis desse constructo explicam aproximadamente 61% da variância total dos dados dos 12 itens de questionário, que estão nas Tabelas 94 e 95.

Tabela 94 - Itens que compõem a variável FP33 – Trabalho coletivo

Item	Descrição dos itens do questionário
q001	O diretor me anima e motiva para o trabalho
q002	Tenho plena confiança profissional no(a) diretor(a)
q003	O diretor consegue que os professores se comprometam com a escola
q008	Sinto-me respeitado pelo diretor
q009	Respeito o diretor
q010	Participo das decisões relacionadas ao meu trabalho
q015	Os diretores, professores e demais membros da equipe da escola colaboram para a escola funcionar bem

Tabela 95 - Itens que compõem a variável FP34 – Grau de inclusão do professor com o trabalho escolar

Item	Descrição dos itens do questionário
q021	Poucos professores assumem a responsabilidade de melhorar a escola
q026	Nesta escola, poucos professores trocam idéias e experiências de modo a viabilizar que todos alunos aprendam
q030	Nesta escola, tenho dificuldade em compartilhar minhas preocupações e frustrações profissionais com outros professores
q031	Nesta escola, tenho poucas oportunidades em discutir idéias sobre ensino-aprendizagem
q034	Nesta escola, tenho poucas oportunidades em discutir o conteúdo programático da minha turma com a equipe da escola

Não houve eliminação de itens neste construto.

2.17 - Construto 17: Características individuais do professor

Este constructo é composto por cinco fatores dimensionados diretamente do arquivo original com as informações sobre os professores. Esses fatores estão na Tabela 96.

Tabela 96 - Variáveis do Construto 17 do arquivo professores obtidos diretamente do questionário

Variável	Descrição da variável	Item do questionário	Descrição do item do questionário
FP35	Sexo	111	Sexo
FP36	Idade	q110	Idade
FP37	Renda familiar	q114	Renda familiar bruta
FP38	Tempo de trabalho	q108	Há quantos anos o professor trabalha nesta ocupação?
FP39	Numero de familiares que moram com o professor	q115	Quantas pessoas da família moram com o professor?

2) Arquivo Diretores

Com o objetivo de se construir o perfil dos diretores que atuaram nas escolas do Projeto GERES foram aplicados questionários de contexto com questões que envolviam desde a coleta de dados pessoais como sexo, idade, estado civil, trabalho, etc., como também foram elaboradas questões visando obter informações amplas referentes a aspectos que direta ou indiretamente dizem respeito ao processo de escolarização dos alunos.

A Tabela 97 traz as variáveis que resultaram do processo de redução de variáveis do arquivo diretores.

Tabela 97 - Dicionário de dados do arquivo Professores

Variável	Descrição	Domínio
FD01	Interferência por intimidação, violência, discriminação e drogas	$1 \leq FD19 \leq 3$
FD02	Interferência do perfil do professor	$1 \leq FD20 \leq 3$
FD03	Interferência do perfil do aluno, família	$1 \leq FD21 \leq 3$
FD04	Interferência da orientação pedagógica	$1 \leq FD23 \leq 4$
FD05	Presença física do diretor	$1 \leq FD23 \leq 4$
FD06	Organização de eventos	$1 \leq FD23 \leq 4$
FD07	Frequência de encontros com a comunidade	$1 \leq FD24 \leq 5$
FD08	Frequência da presença na comunidade	$1 \leq FD25 \leq 5$
FD09*	Formação (Ensino superior em pedagogia)	0 ou 1
FD10*	Exercícios de outras atividades remuneradas	0 ou 1
FD11*	Dedicação a uma só escola	0 ou 1
FD12*	Carga horária de trabalho	$1 \leq FD32 \leq 4$
FD13*	Tempo de exercício ininterrupto na função de direção	$1 \leq FD33 \leq 5$
FD14*	Tempo de trabalho em anos na educação	$1 \leq FD34 \leq 5$
FD15*	Idade	$1 \leq FD35 \leq 8$

*Variáveis com observação direta dos questionários de contexto.

Essas 15 variáveis foram trabalhadas em quatro diferentes constructos, descritos a seguir.

3.01 - Construto 01: Perfil da Comunidade escolar

Na Tabela 98 são apresentados os principais resultados da matriz de análise fatorial com as respectivas cargas fatoriais, bem como, os indicadores que, segundo critério descrito na metodologia, possibilitou a redução dos 26 itens em um conjunto de três fatores.

O agrupamento dos itens do construto que caracteriza os obstáculos que impedem o melhoramento da escola, na concepção do diretor, ficou representado por três fatores o fator

FD01 é aquele que representa a intensidade que a interferência com intimidação, violência, discriminação e drogas influencia o melhoramento da escola. O fator FD02 representa a intensidade que a interferência no perfil do professor impede o melhoramento da escola. O fator FD03 representa a intensidade que o perfil do aluno e família interferem no melhoramento da escola.

Tabela 98 - Arquivo Diretores - Construto 01 - Perfil da Comunidade Escolar

Variáveis	Cargas rotacionadas por Varimax			Comunidades	Correlação Item-fator
	Fator				
	FP30	FP31	FP32		
q088	0,762			0,624	0,782
q089	0,761			0,597	0,717
q086	0,743			0,654	0,815
q087	0,74		0,328	0,666	0,817
q098	0,68			0,483	0,687
q096	0,665			0,531	0,765
q085	0,647			0,486	0,693
q078		0,779		0,625	0,775
q076		0,767		0,635	0,788
q077		0,765		0,623	0,773
q079		0,741		0,583	0,798
q080			0,838	0,772	0,891
q081			0,815	0,74	0,885
KMO ($p < 0,001$)	0,866			Total	
Auto valor inicial	5,041	1,843	1,135	13	
% do Traço	28,79	19,39	13,5	61,68	
Alfa de Cronbach	0,86	0,787	0,729		

Estes três fatores latentes juntas explicam aproximadamente 62% da variância total dos dados. O valor estimado de Kaiser-Meyer-Olkin = 0,87 e juntamente com outros indicadores apresentados na Tabela 98 indicam uma adequação muito boa dos fatores que representam o construto. O Alpha de Cronbach e as correlações item-total indicam uma boa confiabilidade da escala, tendendo a apresentar os mesmos resultados em caso de replicação. A correlação entre os três fatores, $r_{(1;2)} = 0,36$; $r_{(1;3)} = 0,46$; $r_{(1;4)} = 0,36$ se manteve abaixo de 0,5, limite estabelecido neste trabalho como máximo tolerável.

Este constructo, composto por três variáveis que explicam aproximadamente 62% da variância total dos dados, foi reduzido a partir de vinte e quatro itens de questionário que estão relacionados nas Tabelas 99, 100 e 101.

Tabela 99 - Itens que compõem a variável FD01 – Interferência por intimidação, violência, discriminação e drogas

Item	Descrição dos itens do questionário
q088	intimidação a professores e funcionários é um fator que impede o melhoramento da escola?
q089	violência física contra professores e funcionários é um fator que impede o melhoramento da escola?
q086	intimidação a alunos é um fator que impede o melhoramento da escola?
q087	violência física contra alunos é um fator que impede o melhoramento da escola?
q098	a interferência do tráfico de drogas dentro da escola é um fator que impede o melhoramento dela?
q096	a interferência do tráfico de drogas é um fator que impede o melhoramento da escola?
q085	a discriminação racial na comunidade é um fator que impede o melhoramento da escola?

Tabela 100 - Itens que compõem a variável FD02 – Interferência do perfil do professor

Item	Descrição dos itens do questionário
q078	a falta de uma avaliação dos professores é um fator que impede o melhoramento da escola?
q077	a falta de habilidades e conhecimento por parte de alguns professores é um fator que impede o melhoramento da escola?
q076	a apatia dos profissionais da escola é um fator que impede o melhoramento da escola
q079	a dificuldade de demitir professores pouco envolvidos é um fator que impede o melhoramento da escola?

Tabela 101 - Itens que compõem a variável FD03 – Interferência do perfil do aluno, família

Item	Descrição dos itens do questionário
q081	pais apáticos com relação à escolarização dos filhos é um fator que impede o melhoramento da escola?
q080	estudantes problemáticos é um fator que impede o melhoramento da escola?

Neste constructo, algumas variáveis precisaram ser excluídas devido às cargas cruzadas. Estas variáveis estão relacionadas na Tabela 102.

Tabela 102 - Itens que foram excluídos do constructo 1 do arquivo de diretores

Item	Descrição dos itens do questionário
q074	a rotatividade dos professores é um fator que impede o melhoramento da escola?
q075	a falta de tempo para o trabalho em equipe dos professores é um fator que impede o melhoramento da escola?
q082	a desconfiança entre professores e pais é um fator que impede o melhoramento da escola?
q083	a falta de envolvimento da comunidade é um fator que impede o melhoramento da escola?
q084	problemas sociais na comunidade é um fator que impede o melhoramento da escola?
q090	a depredação de equipamentos da escola é um fator que impede o melhoramento dela?
q091	furto ou roubos de equipamentos da escola é um fator que impede o melhoramento dela?
q092	pichações na escola é um fator que impede o melhoramento dela?
q093	depredação de banheiros da escola é um fator que impede o melhoramento dela?
q094	depredação das dependências da escola é um fator que impede o melhoramento dela?
q097	o consumo de drogas dentro da escola é um fator que impede o melhoramento dela?

3.02 - Construto 2: Responsabilidade da direção

Na Tabela 103 são apresentados os principais resultados da matriz de análise fatorial com as respectivas cargas fatoriais, bem como, os indicadores que, segundo critério descrito na metodologia, possibilitou a redução dos dez itens em um conjunto de três fatores.

O agrupamento dos itens do construto, na concepção do diretor, ficou representado por três fatores o FD06 é aquele que representa a interferência da direção na orientação pedagógica, o FD07 representa a intensidade que ocorre a presença física do diretor e o fator FD08 representa a intensidade que a direção organiza e promove eventos na escola.

Tabela 103 - Arquivo Diretores - Construto 01 - Perfil da Comunidade Escolar

Variáveis	Cargas rotacionadas por Varimax			Comuna- lidades	Correlação Item-fator
	Fator				
	FD04	FD05	FD06		
q010_06	0,808			0,704	0,829
q006_06	0,767			0,626	0,781
q011_06	0,765			0,682	0,813
q004_06	0,724			0,547	0,711
q002_06		0,747		0,57	0,741
q008_06		0,696		0,512	0,699
q003_06		0,681		0,467	0,663
q009_06			0,832	0,721	0,800
q007_06			0,761	0,647	0,856
KMO (p < 0,001)	0,767			Total	
Auto valor inicial	2,964	1,495	1,017	16	
% do Traço	27,117	17,177	16,547	60,84	
Alfa de Cronbach	0,790	0,486	0,538		

Estes três fatores latentes juntos explicam aproximadamente 61% da variância total dos dados. O valor estimado de Kaiser-Meyer-Olkin = 0,77 e juntamente com outros indicadores apresentados na Tabela 103 indicam uma adequação boa dos fatores que representam o construto. O Alpha de Cronbach e as correlações item-total indicam uma confiabilidade boa para o primeiro fator e medíocre para os outros dois, desta forma em estudos futuros recomenda-se rever a composição dos itens deste construto nos fatores FD05 e FD06. A correlação entre os fatores é relativamente baixa ($r_{(4;5)} = 0,147$; $r_{(4;6)} = 0,39$ e $r_{(5;6)} = 0,12$).

Este constructo, composto por três variáveis que explicam aproximadamente 61% da variância total dos dados, foi reduzido a partir de nove itens de questionário que estão relacionados nas Tabelas 104, 105 e 106.

Tabela 104 - Itens que compõem a variável FD04 – Interferência da orientação pedagógica

Item	Descrição dos itens do questionário
q010_06	orienta os professores na elaboração de deveres escolares e outras produções acadêmicas?
q006_06	orienta a produção do planejamento escolar: conteúdos a priorizar, estratégias de ensino etc.?
q011_06	orienta os professores na elaboração de projetos didáticos diferenciados?
q004_06	assiste as aulas e orienta pedagogicamente os professores a partir delas?

Tabela 105 - Itens que compõem a variável FD05 – Presença física do diretor

Item	Descrição dos itens do questionário
q002_06	acompanha a entrada e saída dos alunos?
q008_06	atende os pais?
q003_06	atende aos pedidos dos professores (reprodução de material etc.)?

Tabela 106 - Itens que compõem a variável FD06 – Organização de eventos

Item	Descrição dos itens do questionário
q002_06	acompanha a entrada e saída dos alunos?
q003_06	atende aos pedidos dos professores (reprodução de material etc.)?

Nenhum item foi excluído neste constructo.

3.03 - Construto 3: Sociabilidade do Diretor

Na Tabela 107 são apresentados os principais resultados da matriz de análise fatorial com as respectivas cargas fatoriais, bem como, os indicadores que, segundo critério descrito na metodologia, possibilitou a redução dos seis itens em um conjunto de dois fatores.

O agrupamento dos itens do construto sociabilidade do diretor, na concepção dele mesmo, ficou definido por dois fatores o fator FD07 que representa a frequência de encontros do diretor com representantes e líderes comunitários e o fator FD08 representa a frequência da presença com que o diretor vai a eventos.

Tabela 107 - Arquivo Diretores - Construto 03 - Sociabilidade do Diretor

Variáveis	Cargas rotacionadas por Varimax		Comuna- lidades	Correlação Item-fator
	Fator			
	FD07	FD08		
q058	0,906		0,825	0,919
q057	0,895		0,824	0,932
q055		0,747	0,595	0,796
q053		0,705	0,511	0,63
q054		0,702	0,597	0,74
KMO (p < 0,001)	0,621		Total	
Auto valor inicial	2,173	1,18	5	
% do Traço	35,574	31,487	67,06	
Alfa de Cronbach	0,817	0,542		

Estes dois fatores latentes juntas explicam aproximadamente 67% da variância total dos dados. O valor estimado de Kaiser-Meyer-Olkin = 0,62 e juntamente com outros indicadores apresentados na Tabela 107 indicam uma adequação razoável dos fatores que representam o construto. Para o primeiro fator o Alpha de Cronbach e as correlações item-total indicam uma boa confiabilidade da escala, no segundo fator essa escala apresenta uma instabilidade razoável oferecendo poucas condições de replicar os resultados auferidos, se replicado o estudo..

Este constructo, composto por duas variáveis que explicam aproximadamente 67% da variância total dos dados, foi reduzido a partir de catorze itens de questionário que estão relacionados nas Tabelas 108 e 109.

Tabela 108 - Itens que compõem a variável FD07 – Frequência de encontros com a comunidade

Item	Descrição dos itens do questionário
q058	encontra com representantes de organizações e associações de bairro que assistem crianças e jovens?
q057	encontra com líderes comunitários?

Tabela 109 - Itens que compõem a variável FD08 – Frequência da presença na comunidade

Item	Descrição dos itens do questionário
q055	vai a eventos religiosos que os alunos também freqüentam?
q053	vai a eventos na comunidade onde a escola está localizada?
q054	faz compras na comunidade onde a escola está localizada?

Neste constructo, algumas variáveis precisaram ser excluídas devido às cargas cruzadas. Estas variáveis estão relacionadas na Tabela 110.

Tabela 110 - Item excluído do constructo 3 do arquivo de diretores

Item	Descrição dos itens do questionário
q056	3) visita a casa de alunos da escola?

3.04 - Construto 4: Interação escola, pais e comunidade

Decidiu-se por excluí-lo desta análise, por apresentar respostas dicotômicas e ter muitos itens correlacionados. Este constructo não teve qualquer variável inserida no arquivo, pois as variáveis são dicotômicas, característica essa que as impedem de ser submetidas ao processo de análise fatorial. A Tabela 111 apresenta os itens do questionário que seriam incluídos nestes constructos caso não tivesse a característica anteriormente citada.

Tabela 111 - Itens que foram desconsiderados nestes constructo

Item	Descrição dos itens do questionário
q052	a comunidade participou de mutirões para a manutenção da estrutura física da escola
q051	a comunidade participou de mutirões para a limpeza da escola
q050	a comunidade colaborou na manutenção de hortas, pomar e jardins na escola
q046	houve eventos de terceiros realizados na escola e abertos para a comunidade
q045	houve eventos da comunidade utilizando as instalações, equipamentos ou recursos da escola
q047	houve eventos da escola destinados à comunidade externa
q051	houve campanhas de solidariedade promovidas pela escola
q050	houve campanhas de solidariedade propostas pela comunidade, envolvendo a escola

3.05 - Construto 05: Características individuais do diretor

Este constructo é composto por cinco fatores dimensionados diretamente do arquivo original com as informações sobre os diretores. Esses fatores estão na Tabela 112.

Tabela 112 - Variáveis do Constructo 5 do arquivo diretores obtidos diretamente do questionário

Variável	Descrição da variável	Item do questionário	Descrição do item do questionário
FD09*	Formação (Ensino superior em pedagogia)	q006	O diretor (a) possui formação Ens_Sup_Pedag
FD10*	Exercícios de outras atividades remuneradas	q016	O(a) diretor(a) possui outra atividade que contribui para a renda pessoal
FD11*	Dedicação a uma só escola	017r	o(a) diretor(a) trabalha em apenas uma escola
FD12*	Carga horária de trabalho	q018	Carga horária de trabalho do(a) diretor(a) na escola
FD13*	Tempo de exercício ininterrupto na função de direção	q019	Tempo em anos que o(a) diretor(a) trabalha na escola ininterruptamente (na função de diretoria)
FD14*	Tempo de trabalho em anos na educação	q021	Tempo que o(a) diretor (a) trabalha em anos na educação
FD15*	Idade	q101	Idade do(a) diretor(a) em anos

4) Arquivo Escola

Com o objetivo de descrever as escolas que participaram do projeto GERES foram aplicados questionários de contexto com questões que envolviam desde a coleta de dados pessoais como sexo, idade, estado civil, trabalho, etc., como também foram elaboradas questões visando obter informações amplas referentes a aspectos que direta ou indiretamente dizem respeito ao processo de escolarização dos alunos.

A Tabela 113 traz as variáveis que resultaram do processo de redução de variáveis do arquivo diretores.

Tabela 113 - Dicionário de dados do arquivo Professores

Variável	Descrição	Domínio
FA01	NSE médio da escola a partir do nse dos alunos	$-0,71 \leq FA01 \leq 1,28$
FE02	Conservação e limpeza supraestrutura	$1 \leq FE02 \leq 4$
FE03	Conservação e limpeza área de alimentação	$1 \leq FE03 \leq 4$
FE04	Espaço sala de leitura	$1 \leq FE07 \leq 3$
FE05	Condições da Biblioteca	$1 \leq FE07 \leq 3$
FE06	Laboratório de Informática	$1 \leq FE07 \leq 3$
FE07	Sala de Artes e Laboratório de Ciências	$1 \leq FE07 \leq 3$
FE08	Disponibilidade de computador e Internet para Professor	$1 \leq FE15 \leq 3$
FE09	Existência de Equipamentos para fotocópia e escaneamento	$1 \leq FE16 \leq 3$
FE10	Existência de material concreto para matemática	$1 \leq FE17 \leq 3$
FE11	Existência de Terrário e aquário	$1 \leq FE18 \leq 3$

Neste arquivo há registro de 311 escolas dos cinco Pólos participantes do Projeto GERES, das quais 303 participaram nas quatro primeiras ondas e apenas 237 participaram da quinta onda, isso deve-se ao fato de que o Pólo de Salvador aplicou a primeira onda nos alunos da segunda série, e por conseguinte, a quinta onda não foi possível de ser aplicada porque seus alunos já não estavam na primeira fase do Ensino Fundamental.

De maneira geral, os dados revelam que 34,4% são da rede privada de ensino e 65,6 da rede pública, sendo que, das que pertencem a rede pública 40,2% são da rede municipal, 22,5% estadual e 2,9% especial.

Revelam também que, do total de escolas, 19,3% pertencem ao pólo de Belo Horizonte, 19,6% ao de Campinas, 19% ao de Campo Grande, 24,4% ao do Rio de Janeiro e 17,7% ao de Salvador.

Essas 11 variáveis foram trabalhadas em quatro diferentes constructos, descritos a seguir.

4.01 - Construto 1: NSE médio da escola a partir do NSE dos alunos

Este constructo, composto por uma única variável, indica o nível socioeconômico médio da escola e foi obtido através da média aritmética simples do nível socioeconômico de todos os alunos vinculados a essa escola, obtido do arquivo alunos.

4.02 - Construto 2: Conservação e limpeza da escola

Na Tabela 114 são apresentados os principais resultados da matriz de análise fatorial com as respectivas cargas fatoriais, bem como, os indicadores que, segundo critério descrito na metodologia, possibilitou a redução dos 22 itens em um conjunto de dois fatores.

O agrupamento dos itens do construto relacionado com a conservação da escola, na concepção dos observadores externos, ficou representado por dois fatores o fator FE02 que avalia a conservação da estrutura superior da escola e o fator FE03 a conservação da área de alimentação da escola.

Tabela 114 - Arquivo Escola - Construto 02 - Conservação e limpeza da escola

Variáveis	Cargas rotacionadas por Varimax		Comunidades	Correlação Item-fator
	Fator			
	FE02	FE03		
e004	0,782		0,616	0,756
e002	0,782		0,617	0,751
e015	0,776		0,606	0,760
e014	0,756		0,573	0,736
e005	0,721		0,521	0,746
e003	0,708		0,501	0,695
e019	0,653		0,427	0,677
e012	0,646		0,420	0,642
e018	0,631		0,419	0,647
e007		0,841	0,708	0,882
e022		0,829	0,689	0,872
e020		0,803	0,645	0,765
e006		0,774	0,618	0,741
KMO (p < 0,001)	0,770		Total	
Auto valor inicial	4,732	2,630	13,000	
% do Traço	36,02	20,6	56,6	
Alfa de Cronbach	0,88	0,83		

Estes dois fatores latentes juntas explicam aproximadamente 60% da variância total dos dados. O valor estimado de Kaiser-Meyer-Olkin = 0,79 e juntamente com outros indicadores apresentados na Tabela 114 indicam uma boa adequação dos fatores que representam o construto. O Alpha de Cronbach e as correlações item-fator indicam para o fator FE02 uma boa confiabilidade da escala, tendendo a apresentar os mesmos resultados em caso de replicação. A presença de apenas dois itens no fator FE03 deve ser o motivo de

apresentar um valor baixo para o Alpha, no entanto, nota-se que a correlação item-fator é alta no caso. A correlação entre os fatores $r_{(2;3)} = 0,091$ é baixa.

Para estudos que envolverão estes dados com regressão linear, é importante observar que existe alta correlação ($r = 0,70$) entre NSE e FE02, podendo gerar problemas oriundos da multicolinearidade entre essas variáveis.

Este constructo, composto por duas variáveis que explicam aproximadamente 60% da variância total dos dados, foi reduzido a partir de vinte e dois itens de questionário que estão relacionados nas Tabelas 115 e 116.

**Tabela 115 - Itens que compõem a variável FE02 –
Conservação e limpeza da escola**

Item	Descrição dos itens do questionário
e004	Estado de conservação de portas e janelas
e002	Estado de conservação das paredes
e015	Limpeza das portas e janelas
e014	Limpeza das paredes
e005	Estado de conservação dos banheiros
e003	Estado de conservação do piso
e019	Limpeza dos banheiros
e012	Estado de conservação das instalações elétricas
e018	Limpeza das salas de aula

**Tabela 116 - Itens que compõem a variável FE03
– Conservação e limpeza da área de alimentação**

Item	Descrição dos itens do questionário
e007	Estado de conservação do refeitório
e022	Limpeza da refeitório
e020	Limpeza da cozinha
e006	Estado de conservação da cozinha

Neste constructo, algumas variáveis precisaram ser excluídas devido às cargas cruzadas. Estas variáveis estão relacionadas na Tabela 90.

Tabela 117 - Itens que foram excluídos do constructo 2 do arquivo de escolas

Item	Descrição dos itens do questionário
e001	Estado de conservação do telhado
e008	Estado de conservação da cantina
e009	Estado de conservação da área de recreação
e010	Estado de conservação do auditório
e011	Estado de conservação das instalações hidráulicas
e013	Limpeza da entrada do prédio
e016	Limpeza da área de recreação e pátio
e017	Limpeza dos corredores
e021	Limpeza da cantina

4.03 - Construto 3: Disponibilidade de espaços físicos

Os resultados da matriz de análise fatorial do construto planejamento do tempo para as disciplinas são apresentados na Tabela 118, bem como, os principais indicadores que possibilitaram a redução dos itens em quatro fatores.

O construto que caracteriza a disponibilidade de espaços físicos e afins, na concepção dos observadores externos, ficou representado por quatro fatores o FE04 é aquele que caracteriza a disponibilidade de espaço para sala de sala de leitura, o FE05 caracteriza disponibilidade de espaço e acervo para a biblioteca, o FE06 disponibilidade de espaço para laboratório de informática e FE07 Disponibilidade de espaço para sala de artes e laboratório de ciências.

Tabela 118 - Arquivo Escolas - Construto 03 - Disponibilidade de espaços físicos

Variáveis	Cargas rotacionadas por Varimax				Comuna- lidades	Correlação Item-fator
	Fator					
	FE04	FE05	FE06	FE07		
e032	0,954				0,916	0,939
e031	0,952				0,919	0,977
e030		0,937			0,92	0,956
e029		0,933			0,918	0,963
e028			0,924		0,891	0,954
e025			0,913		0,886	0,915
e024				0,835	0,75	0,854
e023				0,830	0,745	0,874
KMO (p < 0,001)	0,607				Total	
Auto valor inicial	2,81	1,97	1,19	0,98	8	
% do Traço	23,24	22,94	22,16	18,47	86,81	
Alfa de Cronbach	0,86	0,91	0,83	0,66		

Estes quatro fatores latentes juntas explicam aproximadamente de 87% da variância total dos dados. O valor estimado de Kaiser-Meyer-Olkin = 0,60 e juntamente com outros indicadores apresentados na Tabela 118 indicam uma adequação aceitável dos fatores que representam o construto. O Alpha de Cronbach e as correlações item-fator indicam confiabilidade muito boa da escala tendendo a apresentar os mesmos resultados em caso de

replicação. A correlação entre os fatores é baixa, situando-se entre 0,04 a 0,32, o que significa que os fatores estão representando conceitos separados.

Este constructo, composto por quatro variáveis que explicam aproximadamente 87% da variância total dos dados, foi reduzido a partir de dez itens de questionário que estão relacionados nas Tabelas 119, 120, 121 e 122.

Tabela 119 - Itens que compõem a variável FE04 – Sala de Leitura

Item	Descrição dos itens do questionário
e032	Há professores com carga horária dedicada à Sala de Leitura?
e031	Esta escola tem Sala de Leitura?

Tabela 120 - Itens que compõem a variável FE05 – Condições da Biblioteca

Item	Descrição dos itens do questionário
e030	Como é considerado o acervo da biblioteca?
e029	Esta escola tem Biblioteca?

Tabela 121 - Itens que compõem a variável FE06 – Laboratório de Informática

Item	Descrição dos itens do questionário
e025	Existência de laboratório de informática
e028	Computadores do Lab. de Informática ligados à Internet?

Tabela 122 - Itens que compõem a variável FE07 – Sala de Artes e Laboratório de Ciências

Item	Descrição dos itens do questionário
e023	Existência de sala para atividade de artes
e024	Existência de laboratório de ciências

Neste constructo, algumas variáveis precisaram ser excluídas devido às cargas cruzadas. Estas variáveis estão relacionadas na Tabela 123.

Tabela 123 - Itens que foram excluídos do constructo 3 do arquivo de escolas

Item	Descrição dos itens do questionário
e026	Existência de horta
e027	Existência de quadra de esportes

4.04 - Construto 4: Recursos materiais disponíveis

Na Tabela 124 são apresentados os principais resultados da matriz de análise fatorial com as respectivas cargas fatoriais, bem como, os indicadores que, segundo critério descrito na metodologia, possibilitou a redução dos 21 itens em um conjunto de quatro fatores.

O agrupamento dos itens do construto que caracteriza os recursos materiais disponíveis, na concepção dos observadores externos, ficou representado por quatro fatores o fator FE08 avalia a existência de computador e internet para professor, o fator FE09 avalia a disponibilidade equipamento para cópia eletrônica, o FE10 de material concreto de matemática e o FE11 existência de terrário / aquário.

Tabela 124 - Arquivo Escolas - Construto 04 - Recursos materiais disponíveis

Variáveis	Cargas rotacionadas por Varimax				Comunalidades	Correlação Item-fator
	Fator					
	FE08	FE09	FE10	FE11		
e042	0,906				0,863	0,912
e043	0,876				0,847	0,914
e051		0,842			0,740	0,801
e053		0,759			0,656	0,843
e033			0,963		0,968	1,000
e035				0,987	0,996	1,000
KMO (p < 0,001)	0,665				Total	
Auto valor inicial	2,37	0,983	0,883	0,834	6	
% do Traço	27,654	22,992	17,041	16,82	84,51	
Alfa de Cronbach	0,8	0,516				

Estes quatro fatores latentes juntas explicam aproximadamente 84% da variância total dos dados. O valor estimado de Kaiser-Meyer-Olkin = 0,66 e juntamente com outros indicadores apresentados na Tabela 124 indicam uma adequação muito precária dos fatores que representam o construto. Na verdade não se conseguiu extrair fatores separados com um número mínimo aceitável de três itens. Desta forma, estes poucos itens que restaram possam de alguma forma representar os recursos materiais mais disponíveis e utilizados nas escolas. A correlação entre os fatores é baixa, situando-se entre 0,10 a 0,40, o que significa que os fatores estão representando conceitos com grau razoável de independência.

Este constructo, composto por quatro variáveis que explicam aproximadamente 84% da variância total dos dados, foi reduzido a partir de 21 itens de questionário que estão relacionados nas Tabelas 125, 126, 127 e 128.

Tabela 125 - Itens que compõem a variável FE08 – Existência de computador e internet para professor

Item	Descrição dos itens do questionário
e042	Existência de Computador(es) para uso dos professores
e043	Existência de Acesso à Internet para uso dos

Tabela 126 - Itens que compõem a variável FE09 - Existência de equipamento para cópia eletrônica

Item	Descrição dos itens do questionário
e051	Existência de Máquina fotocopadora
e053	Existência de Scanner

Tabela 127 - Itens que compõem a variável FE10 – Material concreto de matemática

Item	Descrição dos itens do questionário
e033	Existência de Material concreto de matemática

Tabela 128 - Itens que compõem a variável FE11 – Existência de terrário ou aquário

Item	Descrição dos itens do questionário
e035	Existência de Terrário / Aquário

Neste constructo, houve dois tipos de exclusão. O primeiro, assim como nos constructos anteriores, devido às cargas cruzadas. As variáveis excluídas por esse motivo estão relacionadas na Tabela 129.

Tabela 129 - Itens que foram excluídos do constructo 3 do arquivo de escolas devido ao critério de cargas fatoriais cruzadas

Item	Descrição dos itens do questionário
e036	Existência de Diagramas do corpo humano
e038	Existência de Fitas de vídeo/DVD(educativas)
e039	Existência de Fitas de vídeo/DVD(lazer)
e041	Existência de Livros de consulta para professores
e046	Existência de Retroprojeter
e049	Existência de Canhão Multimídia
e050	Existência de Mimeógrafo

Um segundo grupo de variáveis foi excluído devido à falta de variabilidade nas respostas. A exclusão foi feita quando uma das respostas possíveis teve um predomínio maior que 90% em relação às demais. As variáveis excluídas por esse motivo estão relacionadas na Tabela 130.

Tabela 130 - Itens que foram excluídos do constructo 3 do arquivo de escolas devido ao predomínio maior que 90% de uma das respostas

Item	Descrição dos itens do questionário
e034	Existência de Mapas Geográficos/Globos
e037	Existência de Material para ativ.lúdicas
e040	Existência de Livros didáticos do professor
e044	Existência de Computador(es) para uso administrativo
e045	Existência de Aparelho de som
e047	Existência de Televisão
e048	Existência de Vídeo cassete/DVD
e052	Existência de Impressora

Apêndice 2 - Análise descritiva das variáveis usadas nos modelos

A análise descritiva apresentada a seguir foi feita a partir dos três arquivos que resultaram de todo o processo de preparação feito a partir dos arquivos fornecidos pelo Projeto GERES.

1 - Apresentação da estatística descritiva das variáveis do arquivo do Nível 1

Os arquivos de Nível 1 contém as medidas de proficiência final e inicial de cada um dos anos de escolarização. Como já apontado anteriormente, a partir do segundo ano de escolarização, a proficiência inicial de um ano é a final do ano anterior.

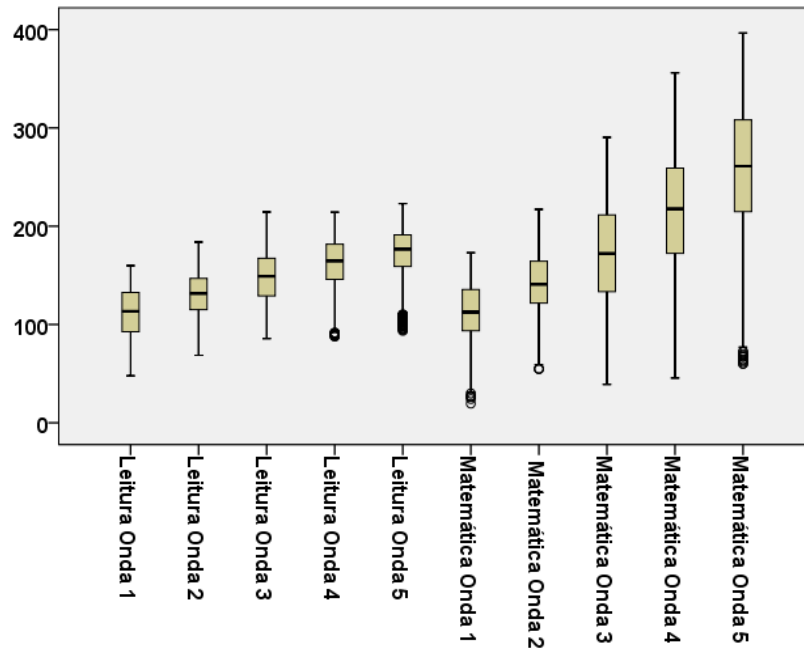
Na Tabela 131 são apresentadas as descrições das proficiências em leitura e em Matemática medidas nas cinco ondas.

Tabela 131- Descrição das proficiência nas cinco ondas

	Leitura					Matemática				
	Onda 1	Onda 2	Onda 3	Onda 4	Onda 5	Onda 1	Onda 2	Onda 3	Onda 4	Onda 5
Média	112,4	130,9	147,8	162,6	174,0	113,5	143,1	171,3	214,4	259,4
Desvio Padrão	25,4	22,0	26,3	25,1	22,8	27,6	31,1	54,6	59,0	62,4
Mínimo	47,8	68,7	85,7	87,8	93,5	19,5	54,7	39,1	45,5	58,6
Máximo	159,9	183,9	214,5	214,2	223,0	173,0	217,1	290,4	356,0	396,5
Amplitude	112,1	115,2	128,8	126,5	129,4	153,5	162,4	251,3	310,5	338,0

O Gráfico 70 representa esses valores e permite a comparação entre os intervalos de proficiências de Leitura e Matemática medidas nas diferentes ondas de aplicação dos testes.

Gráfico 70 - Proficiência medidas em Leitura e matemática dos arquivos do Nível 1



Desse gráfico, duas observações podem ser feitas:

- a) A variabilidade dos dados em matemática é maior que em leitura;
- b) Tanto em leitura, quanto em matemática há uma significativa sobreposição no intervalo de proficiência nos diferentes anos.

2 - Apresentação da estatística descritiva das variáveis do arquivo do Nível 2

Os arquivos de Nível 2 trazem informações que definem o perfil do aluno. Mesmo com a constatação de que ambos os arquivos têm aproximadamente 97% dos registros em comum, indicando uma distribuição relativamente próxima entre as distintas categorias, optou-se por descrever as variáveis de ambos os arquivos para a confirmação dessa hipótese inicial. Sempre que possível, quando a variável dos arquivos de Nível 2 foram extraída diretamente do arquivo original do Projeto GERES, não passando pela AFE, é feita a comparação das amostras usadas no modelo linear hierárquico com as do Projeto GERES.

A variável FA01, que controla o sexo tem distribuição como mostra a Tabela 132.

Tabela 132- Descrição da Variável FA01 - Sexo

	Leitura		Matemática	
	Frequência	%	Frequência	%
Masculino	6181	50,7	6163	50,7
Feminino	5999	49,3	5990	49,3
Total	12180	100,0	12153	100,0

A distribuição constatada confirma que em ambos os arquivos a distribuição relativa é a mesma e que o número de alunos do sexo masculino é um pouco maior que o do sexo feminino, resultado que também se observa no arquivo completo do Projeto GERES onde os meninos correspondem a 52,6% da amostra e as meninas 47,4%, considerando os alunos que responderam a esse item.

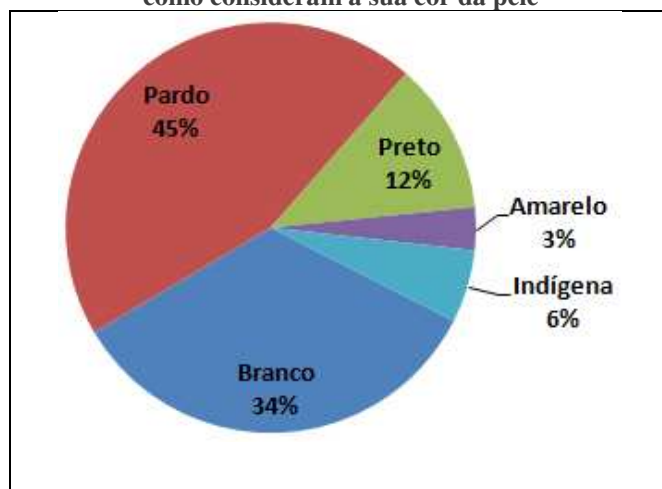
A variável FA02, que controla o com o alunos considera sua cor de pele, tem a seguinte descrição:

Tabela 133- Descrição da Variável FA02_Cor

	Leitura		Matemática		GERES	
	Frequência	%	Frequência	%	Frequência	%
Branco	4138	34,0%	4101	33,7%	5651	32,8%
Pardo	5480	45,0%	5485	45,1%	7667	44,5%
Preto	1452	11,9%	1458	12,0%	2308	13,4%
Amarelo	413	3,4%	409	3,4%	612	3,6%
Indígena	697	5,7%	700	5,8%	978	5,7%
Total	12180		12153		17216	

O Gráfico 71 representa a partição da amostra nas diferentes categorias de como os alunos se consideram em relação à cor da pele, no qual se destaca a predominância da cor parda.

Gráfico 71 - Distribuição dos alunos em função de como consideram a sua cor da pele



Percebe-se que a distribuição relativa entre as amostras de Leitura e Matemática coincidem entre si e também com aquela encontrada no arquivo completo do Projeto GERES.

Outra constatação importante é que o número de alunos que permanecem na amostra é aproximadamente 70% do número de alunos do arquivo completo do Projeto GERES que responderam a esse item.

Sobre o item que controla se o aluno frequentou a Educação Infantil antes de seu ingresso no Ensino Fundamental, a variável FE04, foi encontrada a distribuição de frequência, expressa na Tabela 134, na qual se constata o grande número de alunos que não passaram pela Educação Infantil.

Tabela 134- Frequência de alunos que fizeram a Educação Infantil antes do ingresso ao Ensino Fundamental (FA04)

	Leitura		Matemática		GERES	
	Frequência	%	Frequência	%	Frequência	%
Sim	1713	14,1%	1715	14,1%	2647	15,4%
Não	10467	85,9%	10438	85,9%	14527	84,6%
Total	12180		12153		17174	

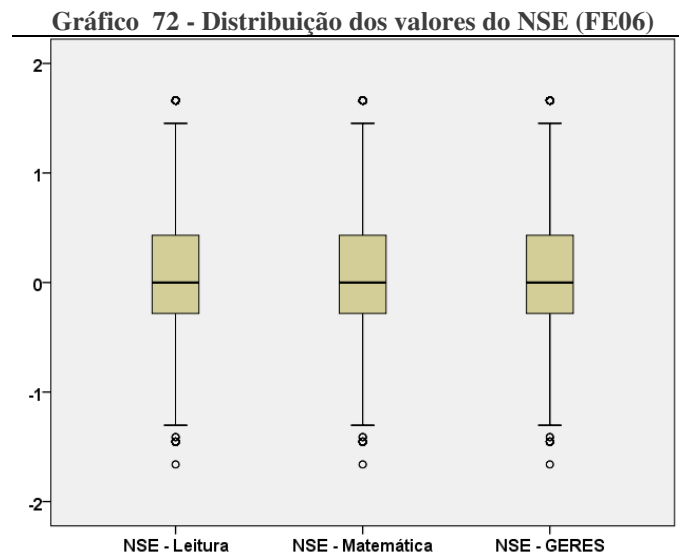
A variável FA06, que controla o nível socioeconômico dos alunos, tem sua descrição feita na Tabela 135.

Tabela 135- Descrição da variável nível socioeconômico (FA06)

	N	Mínimo	Máximo	Média	Desvio Padrão
NSE - Leitura	12180	-1,661	1,661	,10532	,577
NSE - Matemática	12153	-1,661	1,661	,10510	,578
NSE - GERES	29013	-1,661	1,661	-,02307	,544

Como os valores dessa variável são padronizados, percebe-se um deslocamento da média, a exclusão dos registros durante a preparação das amostras de Leitura e Matemática, provocou um deslocamento na ordem de 20% do desvio padrão estimado.

Constata-se também que as amostras de Leitura e Matemática mantiveram os registros com os valores mínimos e máximos existentes no arquivo completo fornecido pelo Projeto GERES, que certamente podem ser considerados *outliers*, como pode ser observado no Gráfico 72.



As variáveis FA07, FA08, FA09, FA10 e FA11 foram produzidas através da AFE e envolveram itens do questionário, cujas respostas têm valores ordinais que representam desde o “nunca” (1), “às vezes” (2) e “sempre” (3). Como são produzidas a partir da média aritmética desses diversos itens, o valor armazenado passou a ser considerado como escalar.

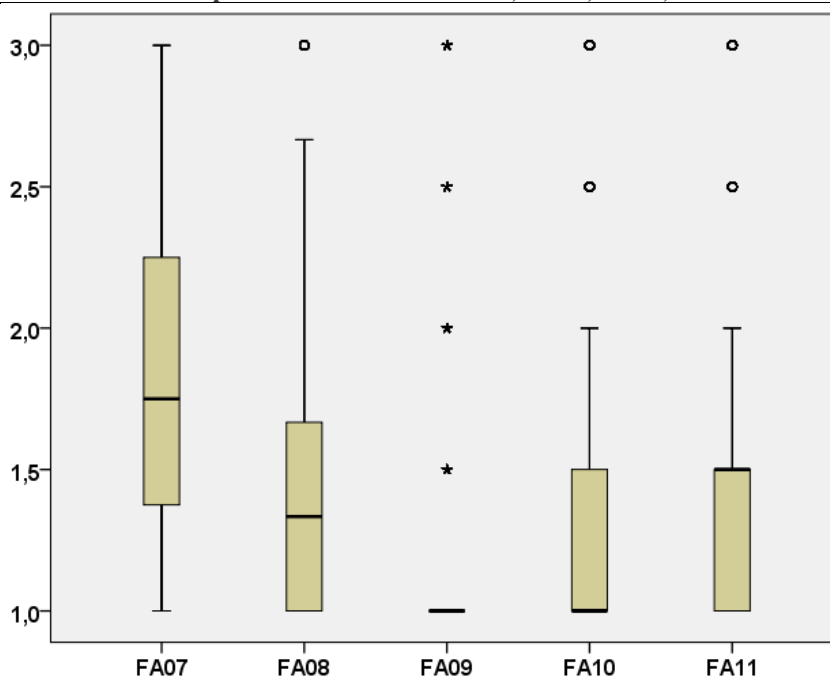
Como pode ser percebido na Tabela 136, não há diferenças significativas entre as amostras de Leitura e de Matemática, afinal os alunos são praticamente os mesmos.

Tabela 136- Descrição das variáveis FA07, FA08, FA09, FA10 e FA11

	Leitura				Matemática			
	Média	Desvio Padrão	Mínimo	Máximo	Média	Desvio Padrão	Mínimo	Máximo
FA07 - Estímulo externo para estudo	1,84	,56	1,00	3,00	1,84	,55	1,00	3,00
FA08 - Estímulo interno para o estudo	1,36	,42	1,00	3,00	1,36	,42	1,00	3,00
FA09 - Percepção do alunos quanto ao incentivo do professor	1,12	,31	1,00	3,00	1,12	,31	1,00	3,00
FA10 - Gosto pela escola	1,32	,44	1,00	3,00	1,32	,44	1,00	3,00
FA11 - Estímulo interno para lição de casa	1,39	,45	1,00	3,00	1,39	,45	1,00	3,00

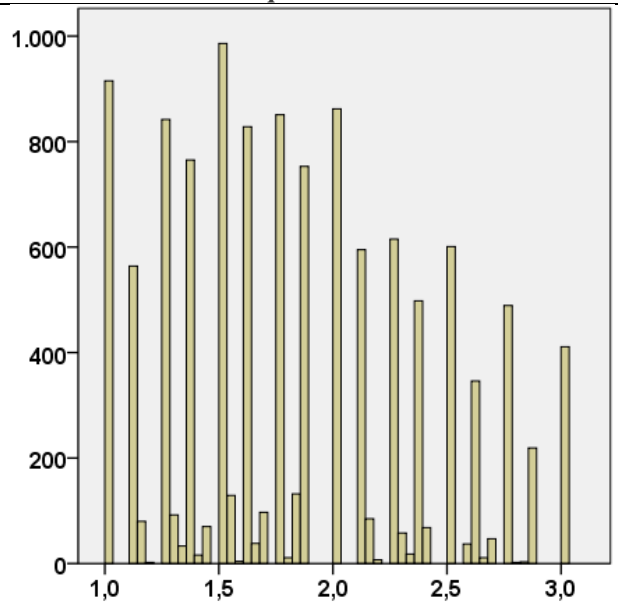
O Gráfico 73 representa a dispersão dos valores encontrados nessas variáveis na amostra de Leitura, já que ela é a mesma na amostra de matemática.

Gráfico 73 - Dispersão das variáveis FA07, FA08, FA09, FA10 e FA11



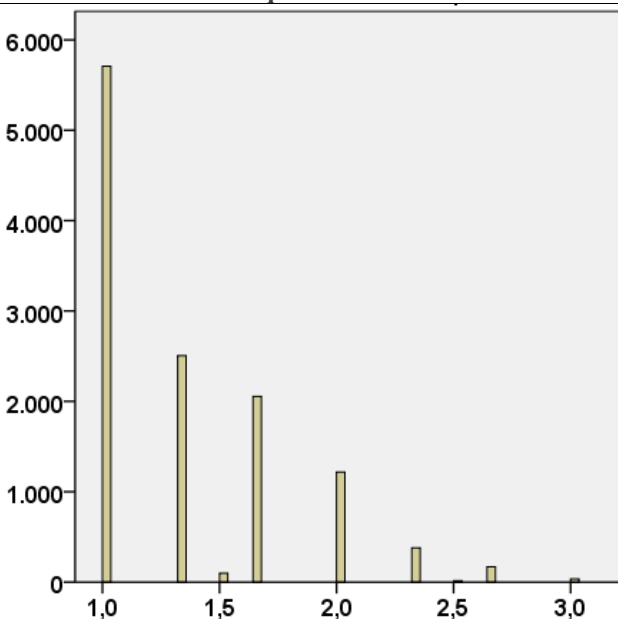
Como a variável FA07, que controla o estímulo extrínseco para estudo, é composta por 8 itens de questionário, há uma diversidade maior de valores como mostra o Gráfico 74.

Gráfico 74 - Frequência da variável FA07

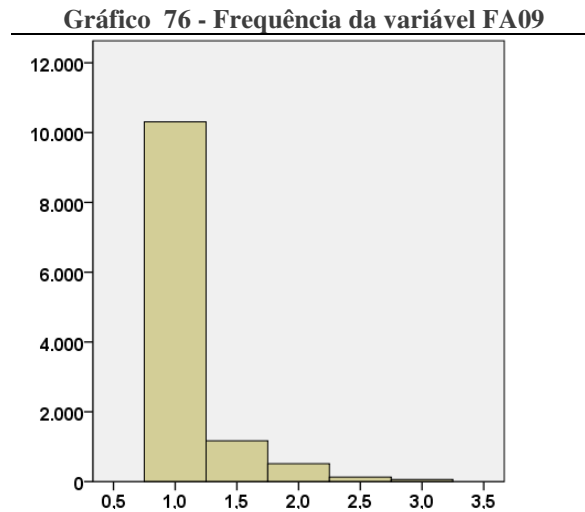


Uma diversidade menor de resultados é encontrada na variável FA08, Estímulo intrínseco para estudo, pois é composto por apenas 3 itens de questionário. Esses valores estão representados no Gráfico 75.

Gráfico 75 - Frequência da variável FA08



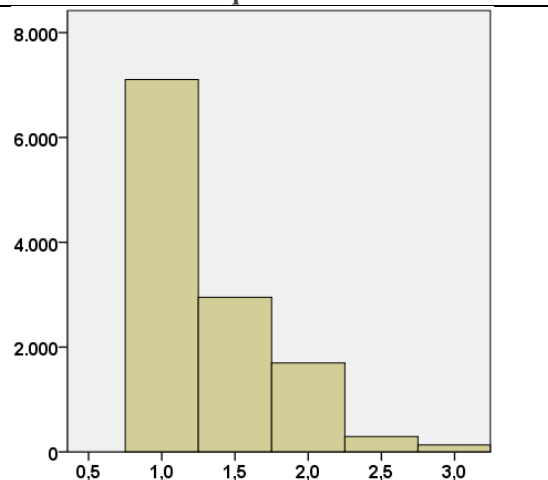
A variável FA09, que controla Percepção do alunos quanto ao incentivo do professor, é composta por 2 itens do questionário e sua concentração no valor 1, indica inexistir tal incentivo. Uma interpretação mais detalhada é possível ser feita com o Gráfico 76.



Esse é um resultado preocupante, pois em torno de 85% dos alunos indicam que não recebe incentivo do professor.

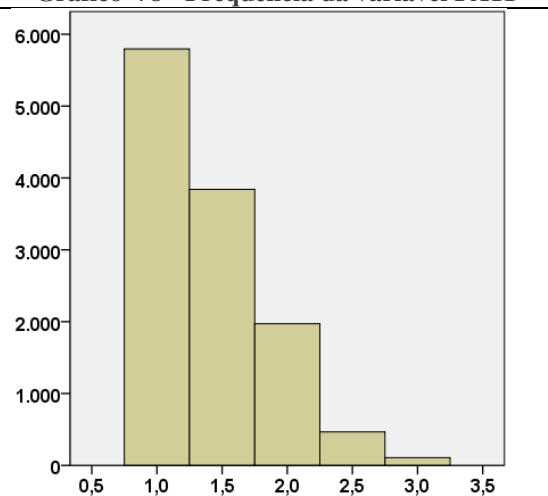
Resultado igualmente surpreendente pode ser observado na variável FA10, que controla a periodicidade com que o aluno gosta de estar na escola e composto por 2 itens de questionário. Nessa variável pode ser observado que 58,3% dos alunos indicam resposta 1, ou seja, indica que nunca se diverte na escola e se sente bem nela (estes são os 2 itens do questionário que compõe essa variável). Apenas 1,1% dos alunos assinalam essas alternativas como sempre. A frequência dos valores encontrados nessa variável está apresentada no Gráfico 77.

Gráfico 77 - Frequência da variável FA10



A variável FA11, composta por dois itens na AFE e que controla o estímulo interno para a tarefa, tem seus valores apresentados no Gráfico 78.

Gráfico 78 - Frequência da variável FA11



Essa variável indica que 47,6% dos alunos nunca fazem a tarefa de casa se ela não valer nota ou se os pais não os obrigarem. Do lado oposto, apenas 0,9% dos alunos fazem as tarefas de casa independentemente da nota e da obrigatoriedade colocada pelos pais.

As variáveis FA12, FA13, e FA14 controlam, de alguma forma, a auto-estima do aluno e são retiradas diretamente do arquivo do Projeto GERES e portanto, não foram elaboradas pela AFE.

A FA12 controla a percepção que o aluno tem sobre a imagem que o professor faz dele mesmo (aluno). Os dados dessa variável mostram que 79,9% dos alunos indicam que seus professores não o acham lento, como mostra a Tabela 137.

Tabela 137- Meus professores pensam que sou um aluno lento (FA12)

	Leitura		Matemática		GERES	
	Frequência	%	Frequência	%	Frequência	%
Sim	2474	20,3%	2476	20,4%	3620	22,3%
Não	9706	79,7%	9677	79,6%	12624	77,7%
Total	12180		12153		16244	

A variável FA13 verificou a autoconfiança do próprio aluno, e constata que 91,4% dos alunos dão resposta positiva, como mostra a Tabela 138.

Tabela 138- Sou tipo de pessoa que sempre será capaz de conseguir, se tentar (FA13)

	Leitura		Matemática		GERES	
	Frequência	%	Frequência	%	Frequência	%
Sim	11128	91,4%	11103	91,4%	14871	90,9%
Não	1052	8,6%	1050	8,6%	1483	9,1%
Total	12180		12153		16354	

Resposta contrária é encontrada na variável FA14, que verificou se algumas vezes o aluno não se sente que é suficientemente bom, obteve o resultado apresentado na Tabela 139.

Tabela 139- Algumas vezes eu sinto que não sou bom o bastante (FA14)

	Leitura		Matemática		GERES	
	Frequência	%	Frequência	%	Frequência	%
Sim	7859	64,5%	7837	64,5%	10547	64,9%
Não	4321	35,5%	4316	35,5%	5701	35,1%
Total	12180		12153		16248	

De maneira geral as amostras de Leitura e Matemática têm resultados relativos muito próximos entre si e o mesmo ocorre quando comparado ao arquivo completo do Projeto GERES.

3 - Apresentação da estatística descritiva das variáveis do arquivo do Nível 3

Nesta apresentação, como a amostra de leitura possui apenas um registro a mais que a amostra de matemática, todas as estatísticas descritivas serão feitas com base no arquivo de leitura, já que o resultado não será significativamente diferente.

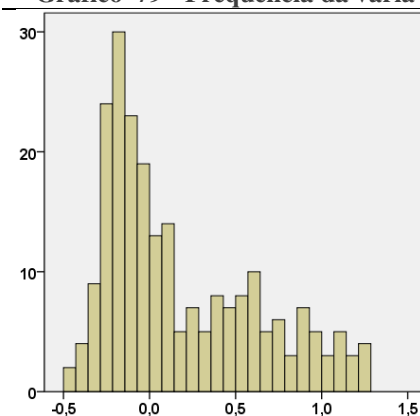
A variável FEA01 controla o nível socioeconômico médio da escola e é determinado a partir da média do nível socioeconômico dos alunos da escola antes da exclusão dos registros dos alunos que não possuem duas médias de proficiências consecutivas. Sua estatística descritiva está apresentada na Tabela 140.

Tabela 140- Descritiva da Variável FEA01

Média	0,183
Desvio Padrão	0,442
Mínimo	-0,457
Máximo	1,280

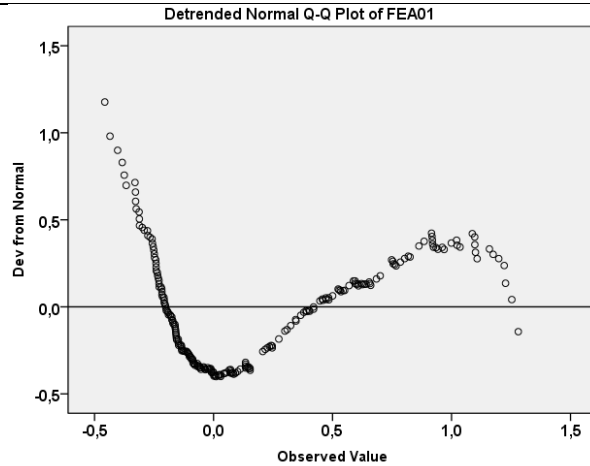
Como os valores do nível socioeconômico dos alunos são normalizados, no arquivo completo do Projeto GERES a média é igual a zero. Da Tabela 140 pode-se perceber que a média das escolas na amostra dos arquivos do Nível 3 teve um deslocamento significativo 0,183. A distribuição da frequência dos valores encontrados nessa variável está no Gráfico 79.

Gráfico 79 - Frequência da variável FEA01



Pode-se notar visualmente que a maior parte da frequência das escolas tem um valor médio inferior a zero e que os dados fogem significativamente da normalidade o que é comprovado pelo Gráfico 80 praticamente em todos os valores observados.

Gráfico 80 – Resíduo da variável FEA01

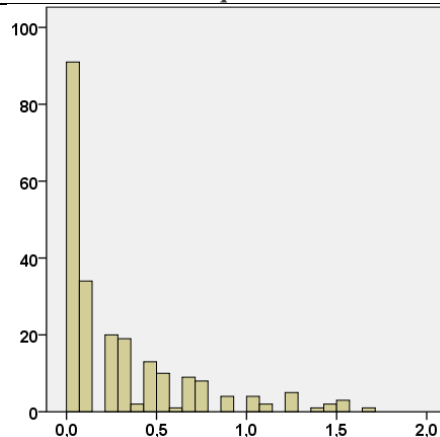


As variáveis FEE02 a FEE11 foram retiradas diretamente dos arquivos produzidos na AFE com dados sobre a escola.

As variáveis FEE02 e FEE03 são obtidas através da AFE e concatenou resultados de perguntas que oscilam de “inexistente” (0) até “adequado” (3)

O Gráfico 81 apresenta os valores encontrados nessa variável FEE02, que controla a conservação e a limpeza do espaço escolar.

Gráfico 81 - Frequência da variável FEE02

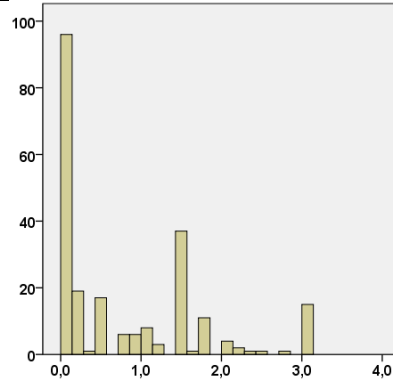


Fica evidente que em geral a limpeza é precária, sendo que o valor médio encontrado foi de 0,29 com desvio padrão de 0,38.

A variável FEE03, que controla a conservação e a limpeza do espaço de alimentação na escola teve um resultado semelhante, com média de 0,77 e desvio padrão de 0,92,

mostrando igualmente a precariedade da limpeza, também nesse espaço Os resultados estão apresentados no Gráfico 82.

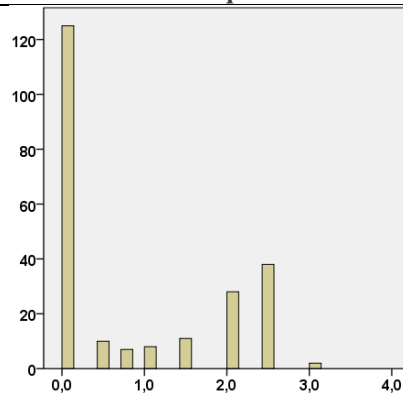
Gráfico 82 - Frequência da variável FEE03



As variáveis FEE04 a FEE06 são obtidas através da AFE e resultam da concatenação dos resultados de dois itens do questionário. As respostas possíveis para essas variáveis, nos itens do questionário, variam de “não tem” (0) até “uso regular e orientado” (3).

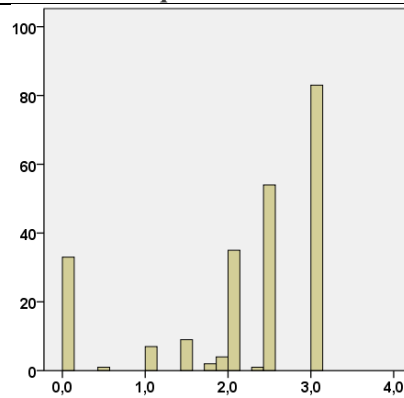
O Gráfico 83 apresenta os valores encontrados na variável FEE04, que controla a existência de sala de leitura, e evidencia a inexistência desse espaço na maioria das escolas. As respostas possíveis para essas variáveis, nos itens do questionário, variam de “não tem” (0) até “uso regular e orientado” (3).

Gráfico 83 - Frequência da variável FEE04



A variável FEE05 tem seus valores representados no Gráfico 84 e evidência que a maioria das escolas possui biblioteca em uso.

Gráfico 84 - Frequência da variável FEE05



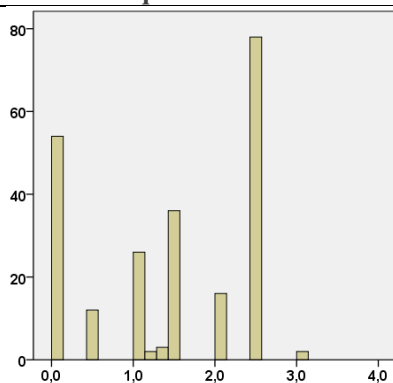
Quanto aos laboratórios de informática, os dados contidos na variável FEE06 e que estão apresentados na Tabela 141 indicam que aproximadamente 40% das escolas possuem laboratório em uso, 30% não possuem o laboratório e as demais têm laboratório, mas não estão em uso.

Tabela 141 - Frequência da Variável FEE06

Valor	Freq	%
0,0	54	23,6
0,5	12	5,2
1,0	26	11,4
1,2	1	0,4
1,3	1	0,4
1,3	1	0,4
1,4	1	0,4
1,4	1	0,4
1,5	35	15,3
1,6	1	0,4
2,0	16	7,0
2,5	78	34,1
3,0	2	0,9
Total	229	100,0

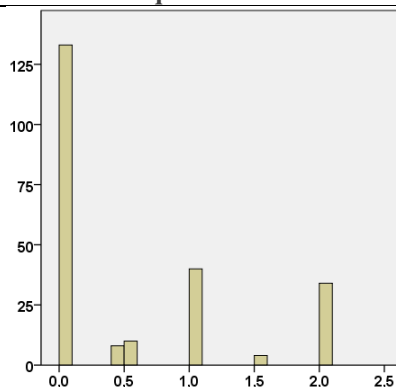
Esses valores estão representados no Gráfico 85.

Gráfico 85 - Frequência da variável FEE06



A inexistência do laboratório de ciência é percebida em 58,1% das escolas que permaneceram na amostra em apenas 15% delas, eles estão disponíveis, como mostra o Gráfico 86 que representa os valores encontrados na variável FEE07. As respostas possíveis para essa variável, nos itens do questionário, variam de “não tem” (0) até “uso regular e orientado” (2).

Gráfico 86 - Frequência da variável FEE07

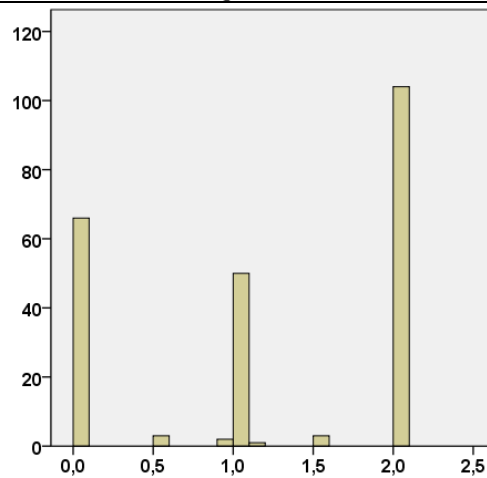


As variáveis FEE08 e FEE09 resultam da concatenação dos resultados de dois itens do questionário, enquanto que as variáveis FEE10 e FEE11 resultam de apenas um item, no entanto, todas passaram pelo processo de AFE em um mesmo constructo ligado à disponibilidade de recursos na escola. As respostas possíveis para essas variáveis, nos itens do questionário, variam de “não tem” (0) até “tem e é usado” (2), passando por “tem, mas não é usado” (1).

A disponibilidade de computadores e internet para o professor, controlado pela variável FEE08, tem uma característica semelhante à disponibilidade dos laboratórios de

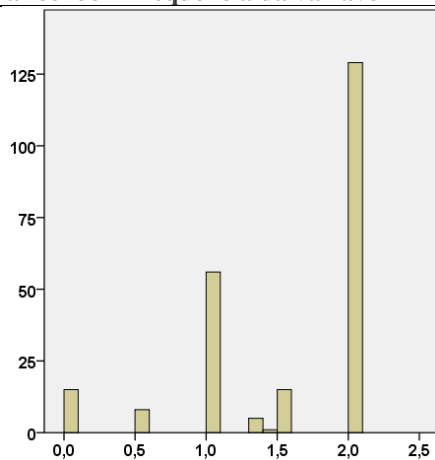
informática (FEE06), de forma que em aproximadamente 45% das escolas os professores usam a Internet e o computador, 30% não disponibilizam esse recurso e o restante 25% tem o recurso, mas não o disponibilizam. As distribuição de frequências da variável FEE08 está representada no Gráfico 87.

Gráfico 87 - Frequência da variável FEE08

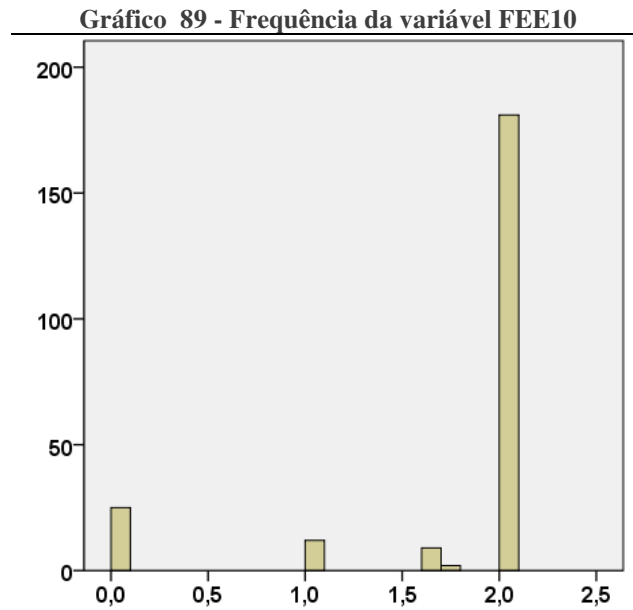


A variável FEE09 controla a existência de equipamentos para fotocópia e escaneamento, indica que aproximadamente 56,3% das escolas usam esses equipamentos e apenas 6,6% delas não os têm. Para as demais escolas, os resultados encontrados nessa variável se aproximam da indicação de que elas possuem esses equipamentos, mas não os usam. Esses resultados são representados no Gráfico 88.

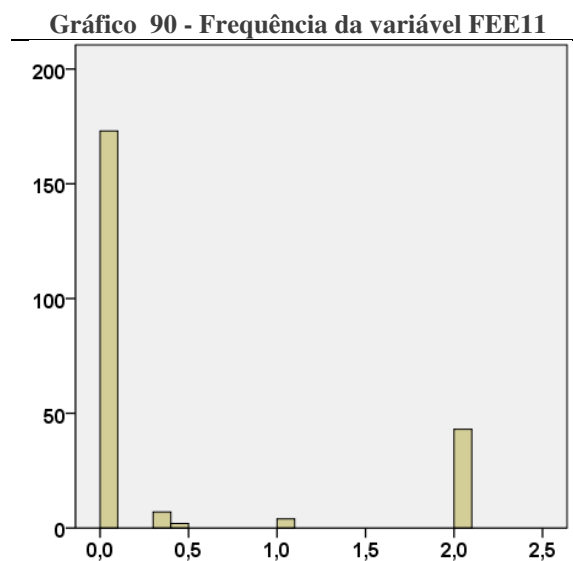
Gráfico 88 - Frequência da variável FEE09



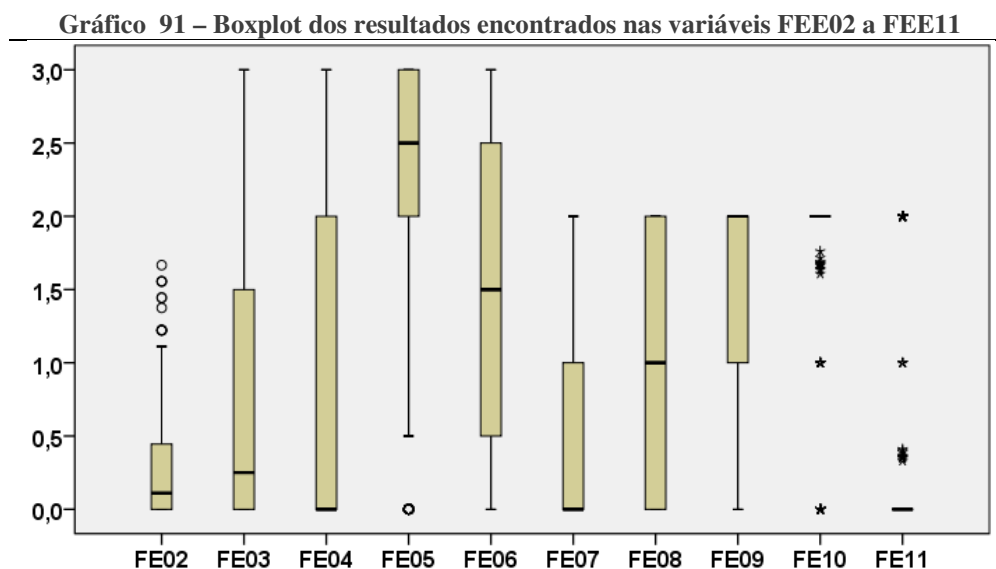
O Gráfico 89 representa os resultados encontrados na variável FEE10, que controla a existência de material concreto para o ensino da matemática. Dele podemos constatar que aproximadamente 80% das escolas usam esse recurso.



O terrário e aquário são recursos inexistentes em 75,5% das escolas da amostra e são considerados em uso em 18,8% delas. Esse recurso é controlado pela variável FEE11, cuja distribuição de frequência está apresentada no Gráfico 90.



O Gráfico 91 apresenta uma síntese agrupando os intervalos dos resultados encontrados nas variáveis associadas ao arquivo que armazena os dados de contexto da escola.

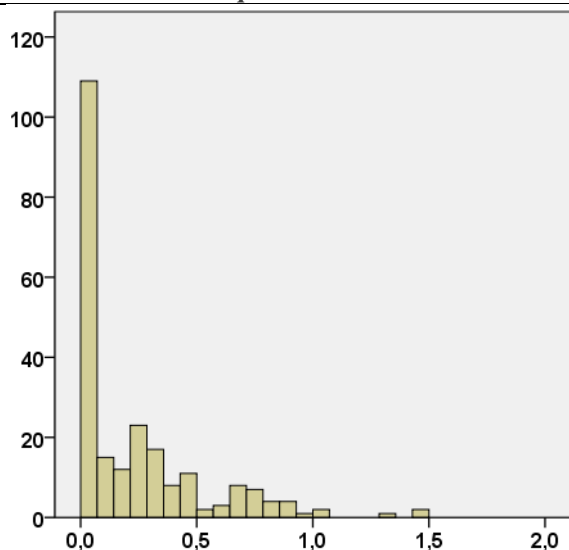


As próximas variáveis estão associadas ao questionário de contexto respondido pelos diretores de escola. Os valores a elas atribuídos foram obtidos a partir da concatenação de vários itens do questionário agrupados na AFE e em seguida, no caso das escolas que tiveram mais que um diretor no período da pesquisa, foi atribuído a média das respostas desses vários diretores.

As variáveis FED12, FED13 e FED14 fazem parte de um único constructo da AFE por verificarem a opinião sobre a influência do perfil do entorno, dos professores e das famílias no melhoramento da escola. As respostas possíveis para os itens do questionário, varia em uma escala de 0 a 2, onde o extremo inferior indica que “Não é um fator” (0) e o outro indica “É um sério fator” (2). O valor intermediário é “Em alguma medida é um fator”(1).

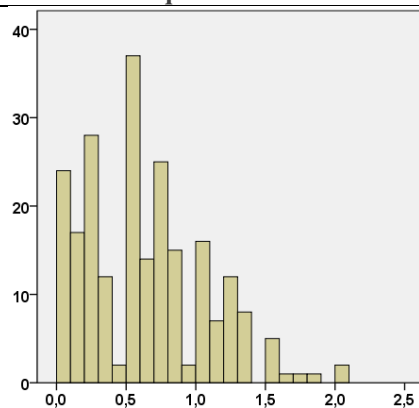
A variável FED12 controla a opinião dos diretores sobre a intimidação, violência, discriminação e drogas são fatores que impedem o melhoramento da escola. O resultado indica que 41,5% dos respondentes entendem que esses aspectos não são fatores a serem considerados. O Gráfico 92 explicita os resultados encontrados nessa variável.

Gráfico 92 - Frequência da variável FEE12



A opinião do diretor sobre o impedimento que representa o perfil do professor está controlada pela variável FEE 13, cuja distribuição de frequências está representada no Gráfico 93. Da análise descritiva constata-se que 10% dos diretores indicam que esse não é um fator que impede o melhoramento da escola e, como a média encontrada nessa variável foi de 0,63 e o desvio padrão de 0,44, os resultados indicam que no mínimo outro 70% dos diretores tendem a ter a mesma opinião. Apenas 0,9% indicam que é um sério fator.

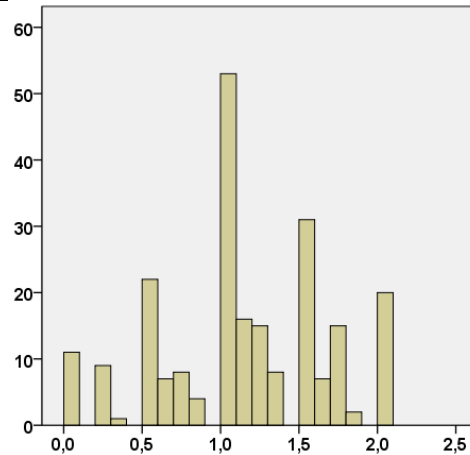
Gráfico 93 - Frequência da variável FEE13



No Gráfico 94 encontra-se a distribuição de frequência observada na variável FEE14 que controla a opinião do diretor sobre a interferência do perfil do aluno e família no melhoramento da escola. A média encontrada de 1,12 e o desvio padrão de 0,52 sugerem que

aproximadamente 70% dos diretores acreditam que esse é um fator em certa medida. Dentre os diretores 4,8% indicam que esse não é um fator e 8,7% indicam é um forte fator.

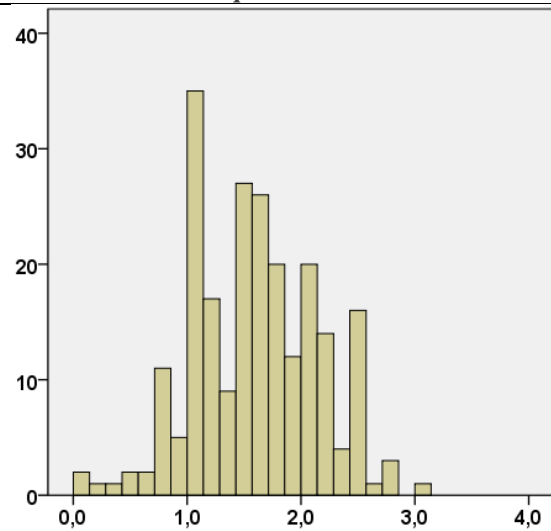
Gráfico 94 - Frequência da variável FEE14



As próximas quatro variáveis também são oriundas dos arquivos que contém informações sobre o contexto de diretores e fazem parte de um constructo denominado responsabilidade do diretor composto por três fatores que são as variáveis explicativas FED15, FED16 e FED17. As respostas possíveis oscilam em uma escala ordinal que varia de 0 a 3, cujos significados correspondem a “Nunca”, “Algumas vezes por mês”, “Algumas vezes por semana” e “Todos os dias da semana”, respectivamente.

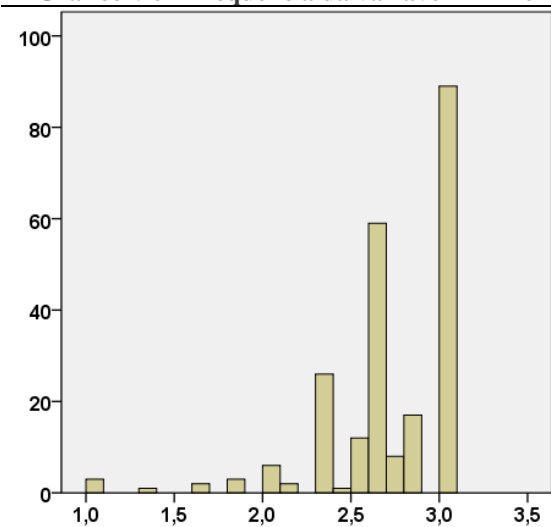
A variável FED15 contém os dados que indicam a periodicidade das interferências da direção na orientação pedagógica e possuem uma média de 1,58 e desvio padrão de 0,56 o que indica que aproximadamente 70% dos diretores interferem pedagogicamente algumas vezes na semana ou no mês. Nesse intervalo, 11,8% deles interferem algumas vezes por mês e 6,1% algumas vezes na semana. Nas extremidades encontram-se 0,9% dos diretores que nunca interferem e 0,4% interferem todos os dias. As frequências dos resultados encontrados nessa variável estão representadas no Gráfico 95.

Gráfico 95 - Frequência da variável FED15



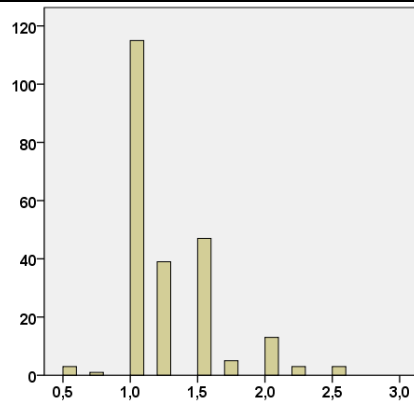
A periodicidade da presença física do diretor, controlada pela variável FED16 evidencia que 39,8% dos diretores estão presentes todos os dias da semana. Apenas 1,3% apontam que estão presentes na escola alguns dias do mês. Com o valor médio de 2,7 e desvio padrão de 0,38, pode-se constatar que aproximadamente 95% dos diretores estão fisicamente presentes na escola todos os dias ou alguns dias da semana. Essa distribuição pode ser observada no Gráfico 96.

Gráfico 96 - Frequência da variável FED16



Em 70% das escolas a direção organiza e promove eventos na escola com a periodicidade de algumas vezes por mês. Essa é a indicação dada pela média de 1,25 e desvio padrão de 0,35 encontrados na variável FED17, cujos valores encontrados estão representados no Gráfico 97.

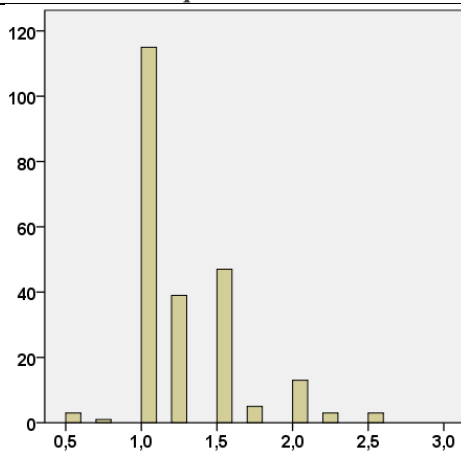
Gráfico 97 - Frequência da variável FED17



As variáveis FED18 e FED19 fazem parte de um constructo sociabilidade do diretor, segundo sua própria percepção e concepção. As respostas possíveis oscilam em uma escala ordinal que varia de 0 a 4, cujos significados correspondem, respectivamente, a “Nunca”, “Uma vez por bimestre”, “Uma vez por mês”, “Algumas vezes por mês” e “semanalmente”.

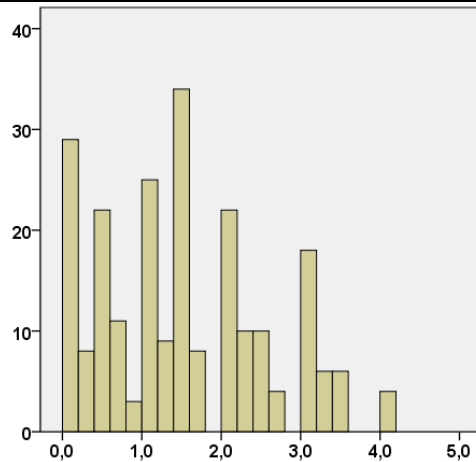
A frequência de encontros do diretor com a comunidade ou seus representantes, controlada pela variável FED18, evidencia que em 50,2% das escolas esse encontro é bimestral. Com o valor médio de 1,25 e desvio padrão de 0,35, pode-se constatar que em aproximadamente 70% das escolas esse encontro é bimestral ou mensal, como também mostra o Gráfico 98.

Gráfico 98 - Frequência da variável FED18



Já a presença da comunidade na escola, controlada pela variável FED19, mostra um resultado semelhante nas diversas alternativas, como pode ser observado no Gráfico 99 e também indicado pela média de 1,48 e desvio padrão de 1,05.

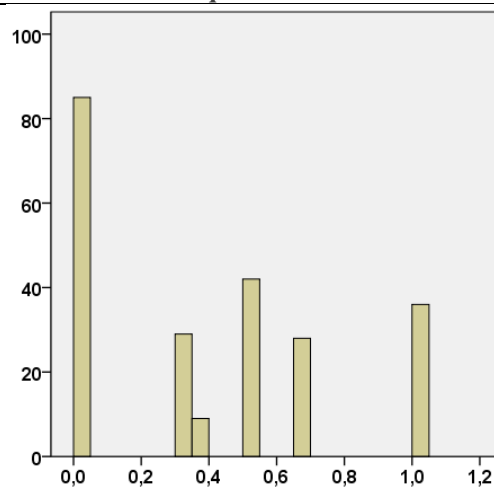
Gráfico 99 - Frequência da variável FED19



Ainda associadas ao perfil do diretor, as próximas variáveis não passaram pela AFE e foram, portanto, retiradas diretamente dos itens do questionário de contexto do diretor.

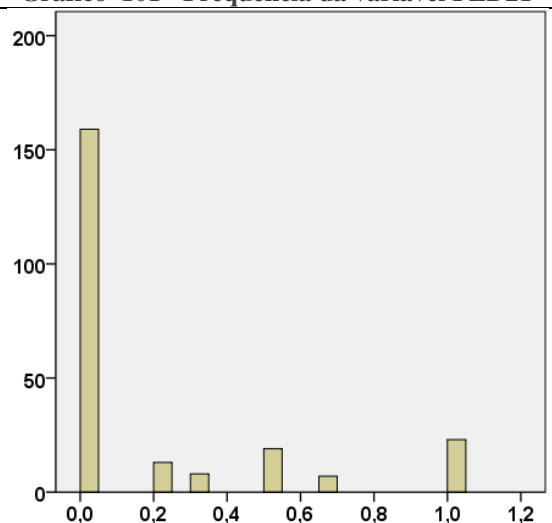
A variável FED20 indica a proporção dos diretores que passaram pela escola e possuem Ensino Superior em Pedagogia. Os resultados indicam que em 15,7% das escolas o diretor ou todos os diretores tiveram curso superior em pedagogia e em 37,1% os diretores ou o diretor não têm o curso. Os resultado estão apresentados no Gráfico 100.

Gráfico 100 - Frequência da variável FED20



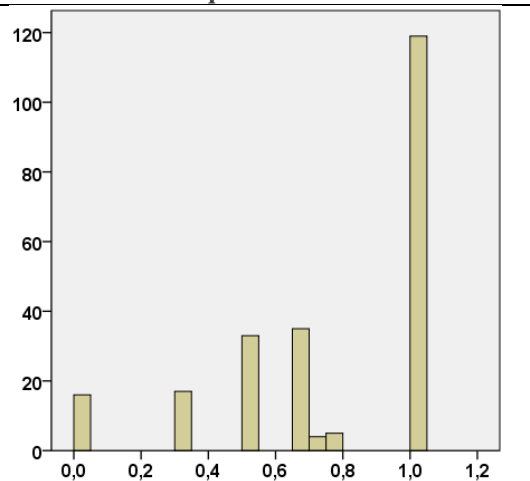
O Gráfico 101 mostra os resultados da variável FED21 que controla a proporção de diretores que possuem outra atividade que contribui para a renda pessoal, portanto seus valores podem variar de 0 a 1. Analisando os resultados, constata-se que em 69,4% das escolas, os diretores ou o diretor não têm outra atividade e em 10% os diretores exercem outras atividades.

Gráfico 101 - Frequência da variável FED21



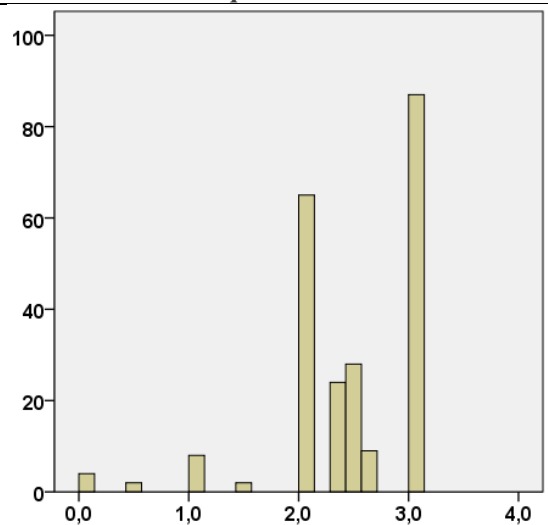
A variável FED22 controla a proporção de diretores que trabalham em apenas uma escola e, portanto tem valores que podem variar de 0 a 1. Nela constata-se que em 52% das escolas, seus diretores trabalham em apenas uma escola e em 7% delas os diretores trabalham em outras escolas. Os valores encontrados estão representados no Gráfico 102.

Gráfico 102 - Frequência da variável FED22



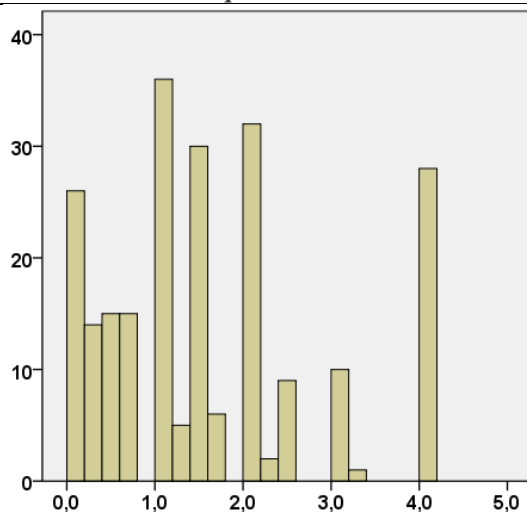
A variável FED23 onde está registrada a carga horária de trabalho do(a) diretor(a) na escola, tem seus valores representados no Gráfico 103. As respostas possíveis oscilam em uma escala ordinal que varia de 0 a 3, cujos significados correspondem, respectivamente, a “Até 20 horas semanais”, “De 21 a 30 horas”, “De 31 a 40 horas semanais” e “mais de 41 horas semanais”. A análise desta variável mostra que em 38% das escolas, seus diretores se posicionam exatamente na categoria dos que trabalham mais de 41 horas semanais e em 28,4% eles se posicionam na categoria dos que trabalham de 31 a 40 horas semanais. A média de 2,41 e o desvio padrão de 0,63 indicam que em 70% das escolas, os diretores se colocam nessas duas faixas. A distribuição de frequência encontradas nessa variável está representada no Gráfico 103.

Gráfico 103 - Frequência da variável FED23



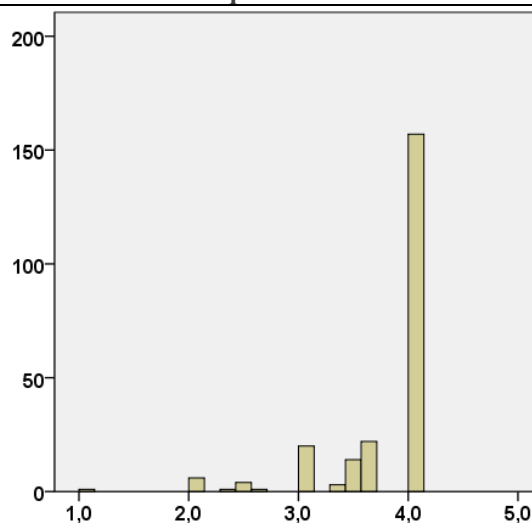
O Gráfico 104 representa a distribuição de frequência encontrada na variável FED24 que controla o tempo em anos que o(a) diretor(a) trabalha na escola ininterruptamente na função de diretoria. As respostas possíveis oscilam em uma escala ordinal que varia de 0 a 4, cujos significados correspondem, respectivamente, a “Há menos de 2 anos”, “de 2 a 4 anos”, “de 5 a 10 anos”, “de 11 a 15 anos” e “Há mais de 15 anos”. Essa distribuição indica que 47,2% das escolas, seus diretores trabalham na escola entre 2 e 10 anos. Em 11,4% delas, eles trabalham a menos de 2 anos, e em 12,% eles trabalham a a mais de 15 anos.

Gráfico 104 - Frequência da variável FED24



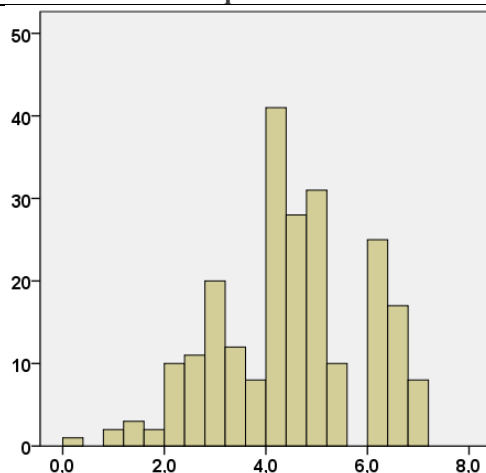
O tempo que o diretor trabalha na educação é controlado pela variável FED25, cuja distribuição de frequência está representada no Gráfico 105. As respostas possíveis oscilam em uma escala ordinal que varia de 0 a 4, cujos significados correspondem, respectivamente, a “Há menos de 2 anos”, “de 2 a 4 anos”, “de 5 a 10 anos”, “de 11 a 15 anos” e “Há mais de 15 anos”. Em 68,6% das escolas os diretores trabalham na educação há mais de 15 anos. A média de 3,73 e o desvio padrão de 0,50 indicam que em aproximadamente 70% das escolas, em média, seus diretores se posicionam entre essas duas categorias.

Gráfico 105 - Frequência da variável FED25



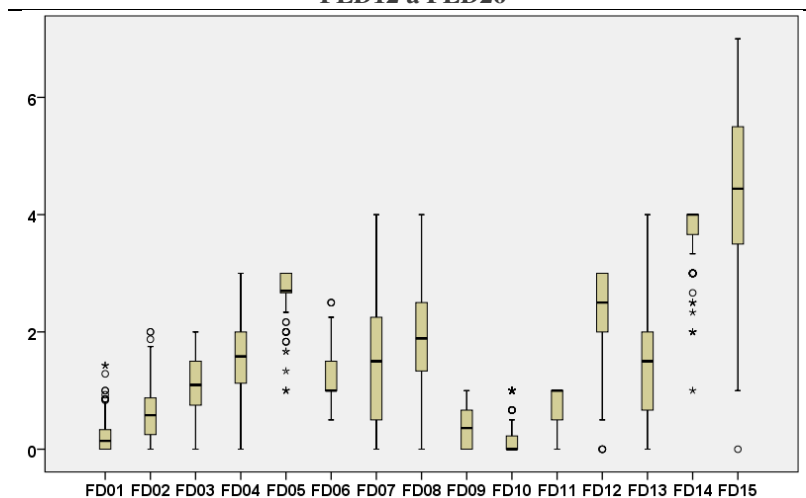
A variável FED26 que controla Idade do(a) diretor(a) em anos. As respostas possíveis oscilam em uma escala ordinal que varia de 0 a 7, cujos significados correspondem a “Até 24 anos”, “de 25 a 29”, “de 30 a 34”, “de 35 a 39”, “de 40 a 44”, “de 45 a 49”, “de 50 a 54” e “mais de 55 anos”, respectivamente. A média de 4,39 e o desvio padrão de 1,40 indicam 70% das escolas, seus diretores, em média a idade que oscila entre 30 e 54 anos. Em 17,5% os diretores têm, em média, ente 40 e 44 anos. Essa distribuição de frequências pode ser observada no Gráfico 106.

Gráfico 106 - Frequência da variável FED26



O Gráfico 107 apresenta uma síntese agrupando os intervalos dos resultados encontrados nas variáveis associadas ao arquivo que armazena os dados de contexto do diretor.

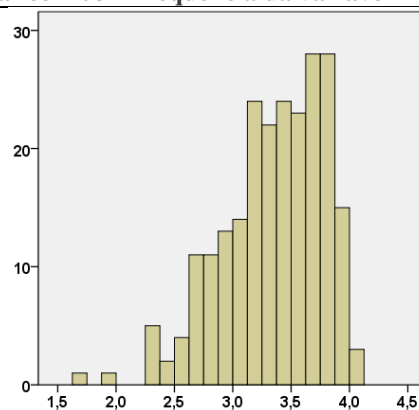
Gráfico 107 - Boxplot dos resultados encontrados nas variáveis FED12 a FED26



As próximas variáveis estão associadas ao questionário de contexto respondido pelos professores da escola. Da mesma maneira como ocorreu com as informações oriundas no arquivo de diretores, os valores atribuídos a essas variáveis foram definidos a partir da média das respostas dadas pelos diversos professores aos distintos itens concatenados do questionário agrupados na AFE.

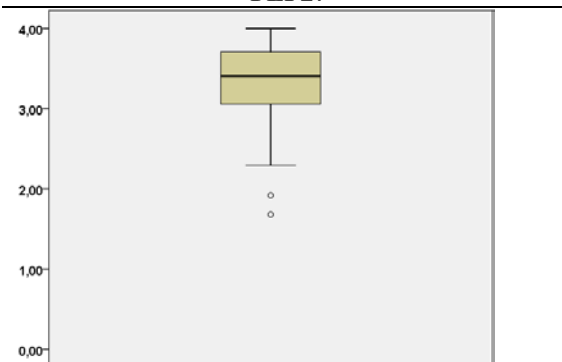
A variável FEP27 que armazena as informações sobre a percepção do professor acerca de sua relação com o diretor. É o único fator que resultou da AFE em um constructo chamado Visão da atuação do diretor pelo professor, composto por 9 itens do questionário. Nesses itens as respostas podem variar de 0 a 4, seguindo a escala Likert, onde o valor do 0 significa “discordo plenamente” e o 4 significa concordo plenamente”. O Gráfico 108 representa a distribuição das frequência dessa variável.

Gráfico 108 - Frequência da variável FEP27



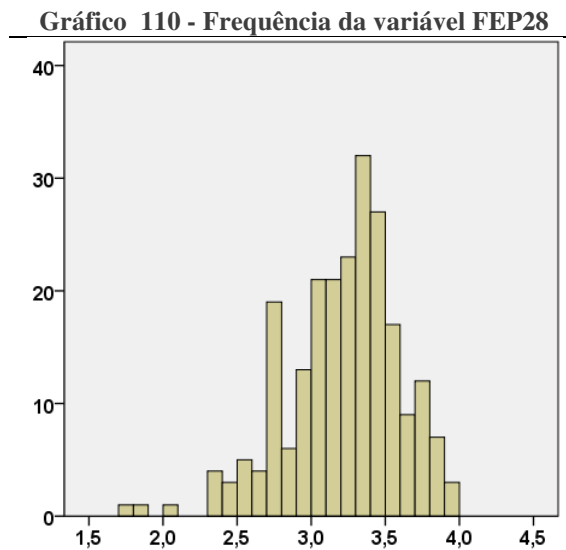
A média de 3,35 e o desvio padrão de 0,43 indicam um resultado positivo nessa variável para todas as escolas como mostra o Gráfico 109.

Gráfico 109 - Intervalo de valores da variável FEP27

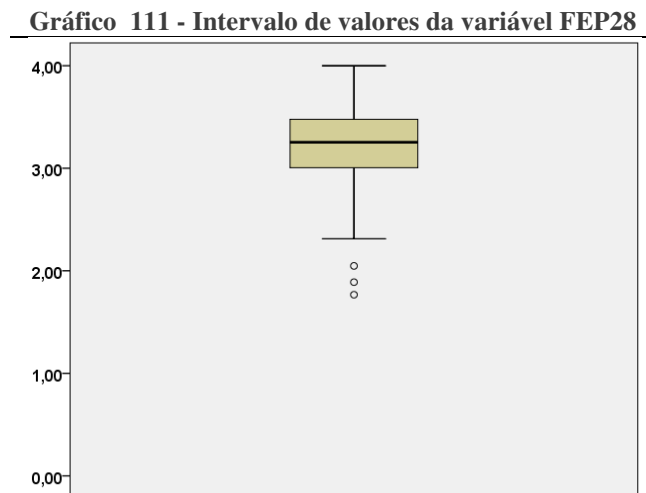


As variáveis FEP28, FEP29, FEP30 e FEP31 fazem parte de um constructo denominado metas compartilhadas e também tem como respostas valores que variam de 0 a 4, seguindo a escala Likert, onde o valor do 0 significa “discordo plenamente” e o 4 significa concordo plenamente”.

A variável FEP28 que armazena as informações sobre a percepção do professor da boa interação indica a aprovação neste aspecto como mostra o Gráfico 110.

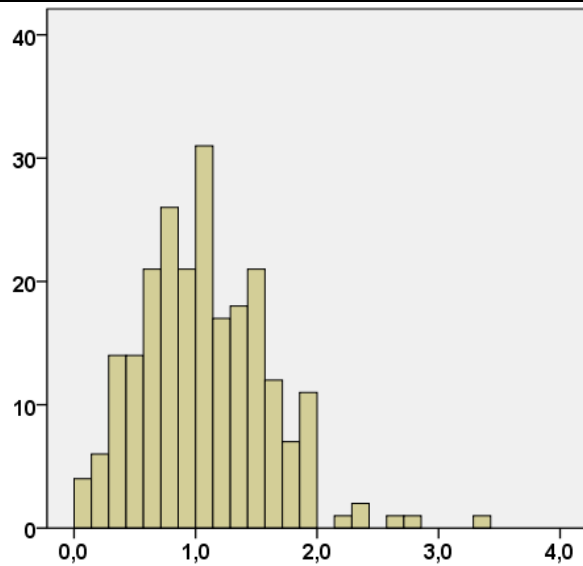


Essa avaliação é confirmada pela média de 3,21 e o desvio padrão de 0,39 observado na variável e pelo Gráfico 111.



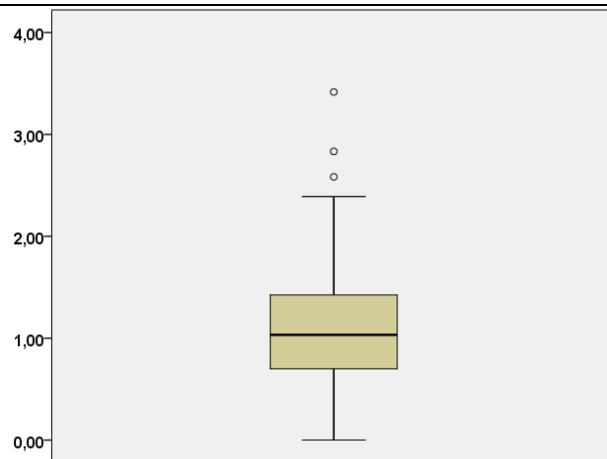
A variável FEP29 que armazena as informações sobre a percepção do professor acerca do pouco trabalho coletivo mostra uma aprovação nesse aspecto. Em quase todas as escolas, os professores indicam uma fragilidade neste aspecto como mostra o Gráfico 112.

Gráfico 112 - Frequência da variável FEP29

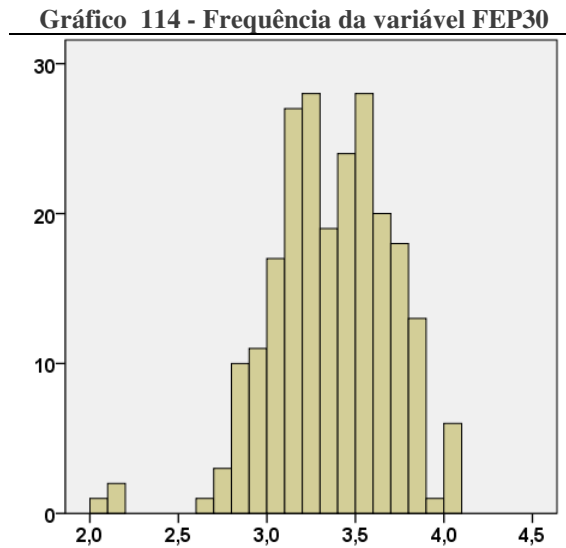


A média de 1,07 e o desvio padrão de 0,53 confirmam que mais de 75% das escolas discordam do pouco trabalho coletivo e apenas 2,3% concordam, pelo menos em alguma medida. O intervalo de valores encontrados nessa variável é explícito no Gráfico 113.

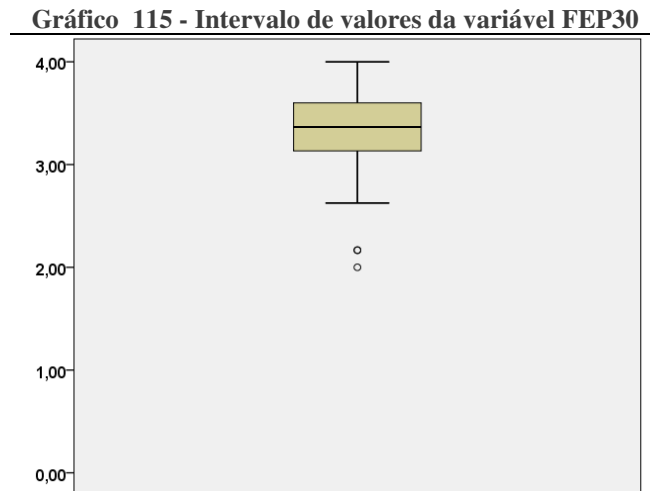
Gráfico 113 - Intervalo de valores da variável FEP29



A variável FEP30 que armazena as informações sobre a percepção do professor da boa relação entre professores indica também uma aprovação neste aspecto. O Gráfico 114 representa os valores encontrados nesta variável.

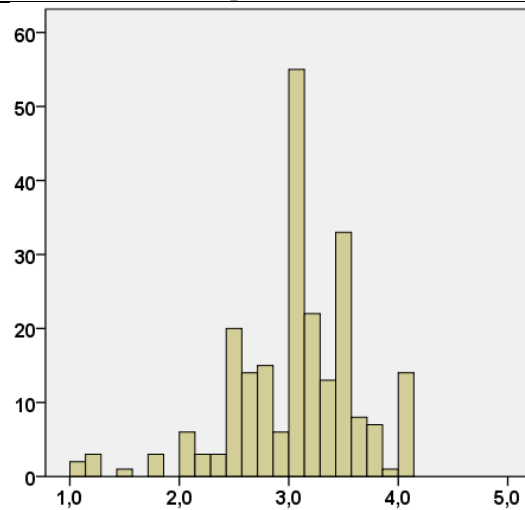


A média de 3,35 e o desvio padrão de 0,33 também indicam um resultado positivo nessa variável como mostra o Gráfico 115.



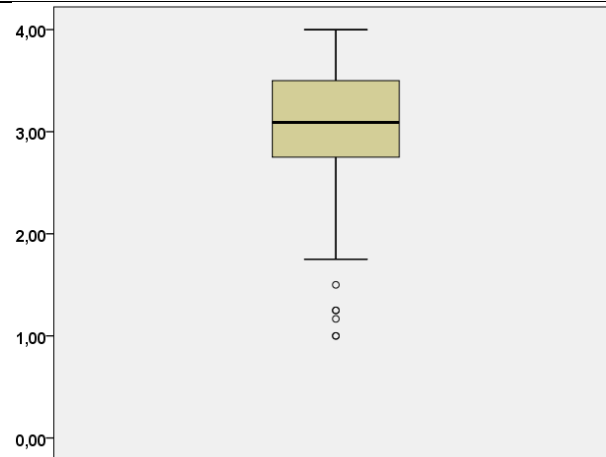
O Gráfico 116 representa a distribuição de frequências encontradas na variável FEP31, indica a percepção que o professor tem acerca da unidade de propósito entre os professores na escola.

Gráfico 116 - Frequência da variável FEP31



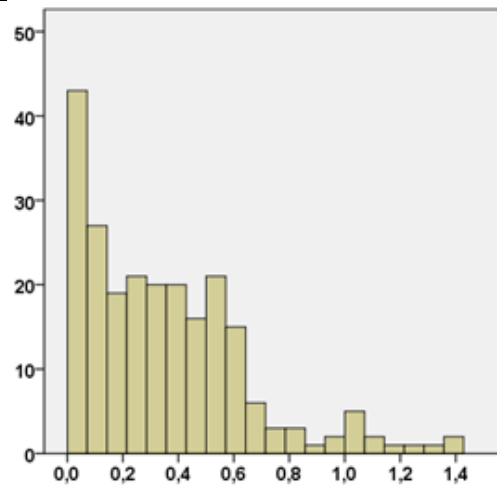
A média de 3,04 e o desvio padrão de 0,57 encontrado na variável FEP31 indicam, mais uma vez, um resultado positivo nessa variável como evidencia o Gráfico 117.

Gráfico 117 - Intervalo de valores da variável FEP31



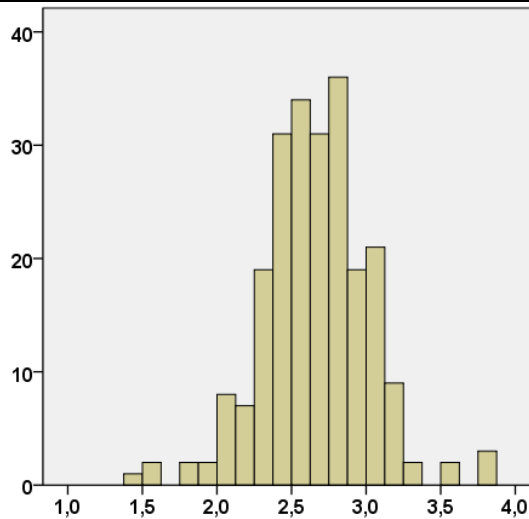
A variável FEP32, única neste constructo, armazena as informações sobre a percepção do professor acerca do impedimento que a violência tem no melhoramento escolar. As respostas nesta variável pode variar entre 0, 1 ou 2, sendo que significam, respectivamente “não impede”, “em alguma medida” e “impede muito”. A média de 0,35 e o desvio padrão de 0,30 indicam que, na opinião dos professores, a violência não é um impedimento significativo para o melhoramento da escola. O gráfico 118 representa os valores encontrados nesta variável.

Gráfico 118 - Frequência da variável FEP32



A variável FEP33, única neste constructo relacionado às tarefas de casa e composto por dois itens do questionário que quantifica a produção de tarefa de casa tem valor que pode variar entre 0 e 4. A média de 2,64 e o desvio padrão de 0,35 indicam que a produção quantitativa de tarefa produzida é mediana, como mostram os resultados do Gráfico 119.

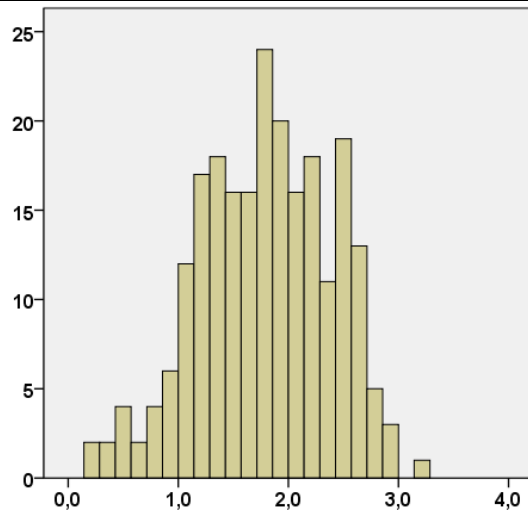
Gráfico 119 - Frequência da variável FEP33



As variáveis FEP34 e FEP35 fazem parte de um constructo denominado expectativas em relação à aprendizagem e tem como respostas valores que variam de 0 a 4, seguindo a escala Likert, onde o valor do 0 significa “discordo plenamente” e o 4 significa concordo plenamente”.

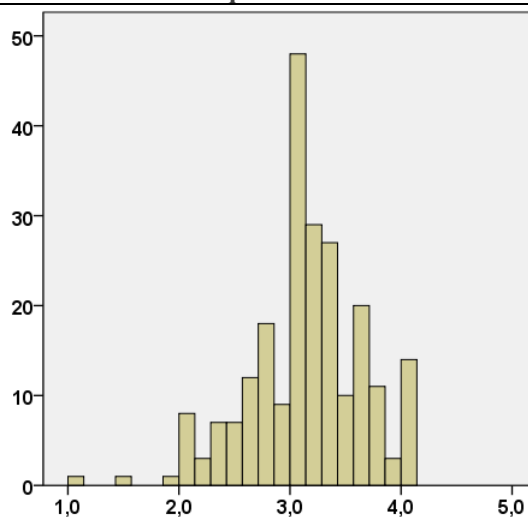
A variável FEP34 controla as expectativas positivas que o professor tem relação à Educação e pelo Gráfico 120 pode-se perceber que tais expectativas são medianas com uma tendência para ser negativa. Essa avaliação é confirmada pela média de 1,78 e o desvio padrão de 0,60.

Gráfico 120 - Frequência da variável FEP34



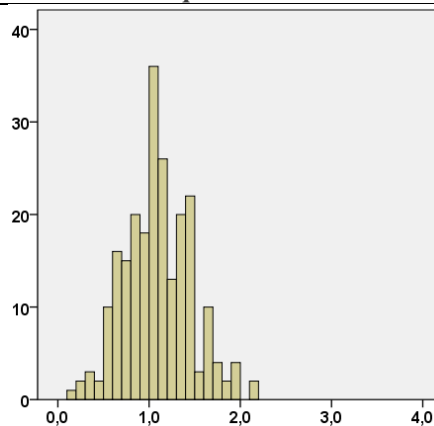
Finalidade análoga tem a variável FEP35, porém controla as expectativas positivas que o professor tem em relação à aprendizagem, como mostra o Gráfico 121. A média de 3,12 e o desvio padrão de 0,51, confirmam o apresentado na variável FEP34, porém com maior intensidade.

Gráfico 121 - Frequência da variável FEP35



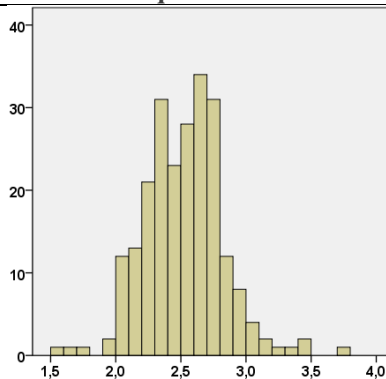
No Gráfico 122 está a distribuição de frequência da variável FEP36 que controla o acesso a atividades culturais que o professor teve em um ano. As respostas nesta variável podem variar no arquivo entre 0 e 3, sendo que significam, respectivamente “nenhuma”, “uma ou duas vezes”, “três ou quatro vezes” e “mais que quatro vezes”. A média de 1,1 e o desvio padrão de 0,37 indicam que 70% dos professores tiveram acesso atividades culturais uma ou duas vezes no ano.

Gráfico 122 - Frequência da variável FEP36



A variável FEP37 controla a proposição de leitura aos alunos e tem como possíveis respostas a escala de 0 e 4 que significam, respectivamente “nunca”, “raramente”, “algumas vezes no bimestre”, “cerca de uma vez por semana” e “várias vezes por semana”. A média de 2,52 e o desvio padrão de 0,34 indicam que 70% dos professores propõem leitura para os alunos em torno de uma vez por semana e algumas vezes no bimestre. O Gráfico 123 apresenta a distribuição de frequências nessa variável.

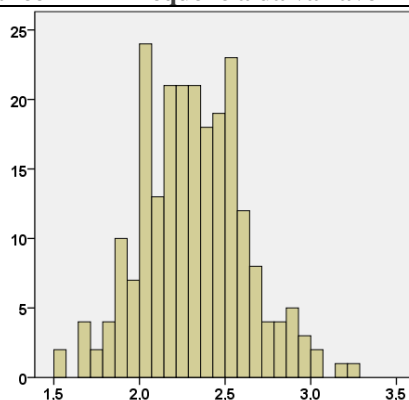
Gráfico 123 - Frequência da variável FEP37



As variáveis FEP38 e FEP39 resultam de um mesmo constructo na AFE denominado prática de escrita e são compostos por itens do questionário cujas alternativas e valores correspondentes nos arquivos variam em uma escala de 0 e 4 que significam, respectivamente “nunca”, “raramente”, “algumas vezes no bimestre”, “cerca de uma vez por semana” e “várias vezes por semana”.

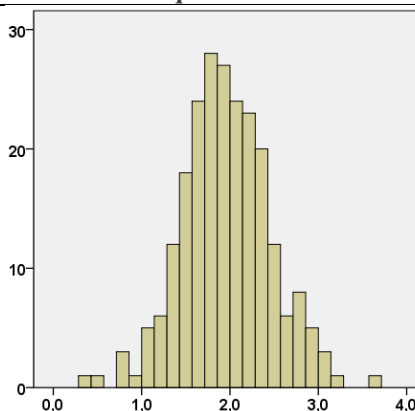
A variável FEP38 controla a frequência da proposição da escrita através da produção de textos. O Gráfico 124 e a média de 2,3 e o desvio padrão de 0,3 indicam que 70% dos professores fazem essa proposta algumas vezes por bimestre.

Gráfico 124 - Frequência da variável FEP38



No Gráfico 125 está a distribuição de frequência da variável FEP39 que controla a periodicidade com que são propostas cópias e exercícios de caligrafia. A média de 1,9 e o desvio padrão de 0,5 sugerem que 70% dos professores fazem essa proposta algumas vezes por bimestre.

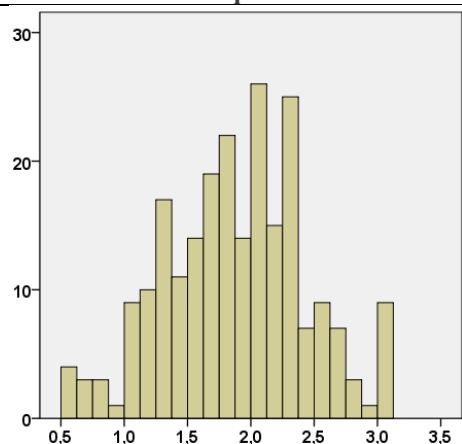
Gráfico 125 - Frequência da variável FEP39



As variáveis FEP40, FEP41 e FEP42 resultam de um mesmo constructo na AFE ligado a metodologia da Língua Portuguesa nas escolas. As alternativas de respostas nos itens do questionário e os valores correspondentes nos arquivos são: “nenhuma” (0), “baixa” (1), “média” (2) e “alta” (3).

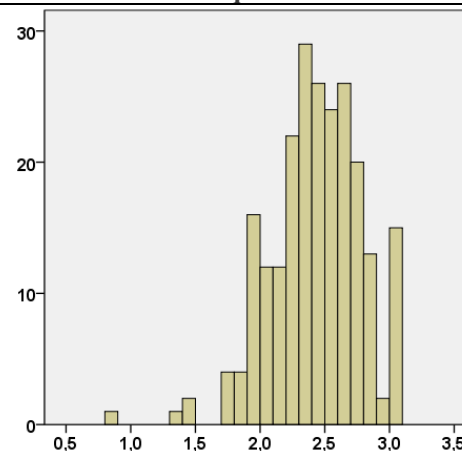
A variável FEP40 quantifica a proposição de leitura individual precedida da explicação do professor que é feita para o aluno. No Gráfico 126, a média de 1,9 e o desvio padrão de 0,6 indicam que 70% das escolas os professores fazem essa proposta com intensidade média.

Gráfico 126 - Frequência da variável FEP40



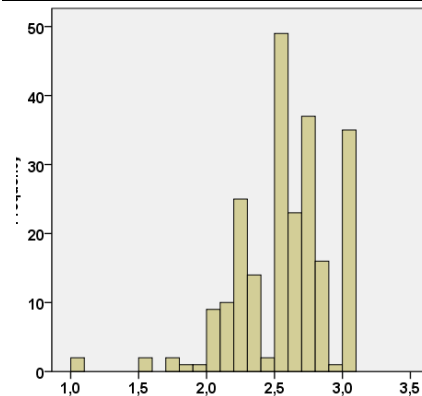
No Gráfico 127 está a distribuição de frequência da variável FEP15 que quantifica a intensidade com que a leitura como prática coletiva é proposta aos alunos. A média de 2,4 e o desvio padrão de 0,3 indicam que em 70% das escolas os professores fazem essa proposta com intensidade média para alta.

Gráfico 127 - Frequência da variável FEP41



A frequência com que é proposta a leitura individual seguida da produção de texto individual é controlada pela variável FEP42 e os resultados observados estão descritos no Gráfico 128. A média de 2,5 e o desvio padrão de 0,3 indicam que em 70% das escolas, os professores fazem essa proposta com intensidade média.

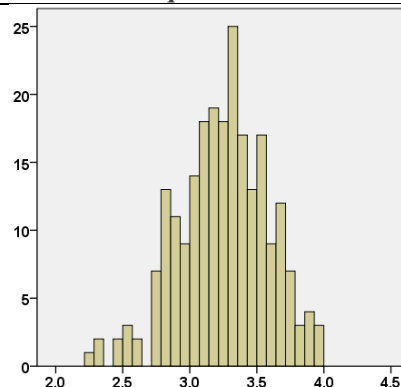
Gráfico 128 - Frequência da variável FEP42



As variáveis FEP43 e FEP44 resultam de um mesmo constructo na AFE ligado a metodologia do ensino da matemática. As alternativas de respostas nos itens do questionário e os valores correspondentes nos arquivos são: “nunca” (0), “raramente” (1), “algumas vezes no bimestre” (2), “cerca de uma vez por semana” (3) e “várias vezes por semana” (4) para a periodicidade com que os professores fazem propostas dentro desse enfoque aos seus alunos.

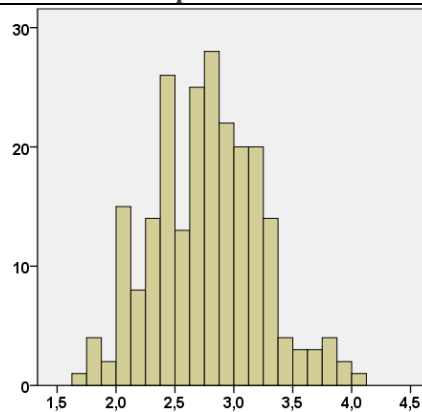
A variável FEP43 controla a periodicidade de atividades com enfoque na resolução de problemas e os resultados observados estão no Gráfico 129. A média de 3,2 e o desvio padrão de 0,3 sugerem que em 70% das escolas os professores fazem propostas com essas características em torno de uma vez por semana.

Gráfico 129 - Frequência da variável FEP43



No Gráfico 130 está a distribuição de frequência da variável FEP44, que controla a periodicidade com atividades matemáticas com enfoque na memorização são propostas aos alunos. A média de 2,8 e o desvio padrão de 0,5 indicam que em 70% das escolas, os professores fazem propostas com essas características em torno de uma vez por semana.

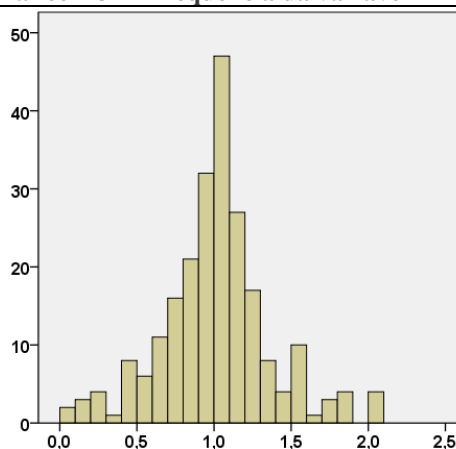
Gráfico 130 - Frequência da variável FEP44



As variáveis FEP45 e FEP46 resultam de um mesmo constructo na AFE ligado ao uso do livro didático. As alternativas de respostas para os itens do questionário e os valores correspondentes nos arquivos são: “nenhuma” (0), “pouca” (1) e “muita” (2).

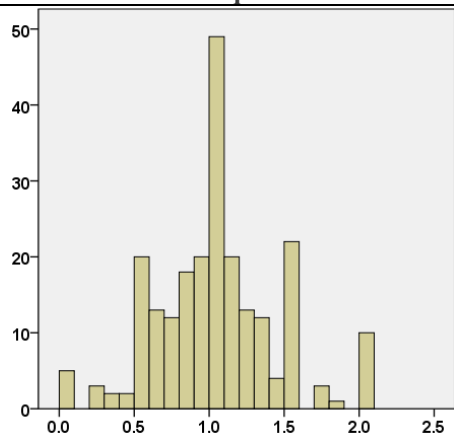
A variável FEP45 controla a opinião do professor acerca da adequação do livro didático ao ensino. No Gráfico 131 está representada a distribuição de frequência observada nessa atividade. A média de 1,0 e o desvio padrão de 0,3 sugerem que em 70% das escolas os professores consideram os livros didáticos pouco adequados.

Gráfico 131 - Frequência da variável FEP45



No Gráfico 132 está representada a distribuição de frequência da variável FEP46, que controla a opinião acerca do grau de importância inerente a indicação externa do livro. A média de 1,0 e o desvio padrão de 0,4 indicam que os professores de 70% das escolas consideram pouco importante as orientações externas à escola sobre a adoção do livro didático.

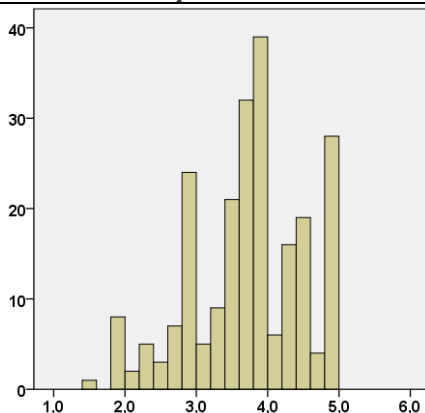
Gráfico 132 - Frequência da variável FEP46



As variáveis FEP47, FEP48 e FEP49 resultam de um mesmo constructo na AFE denominado planejamento do tempo para as disciplinas. As alternativas de respostas nos itens do questionário e os valores correspondentes nos arquivos são: “0%” (0), “15%” (1), “30%” (2), “45%” (3), “60%” (4), e “75%” (5).

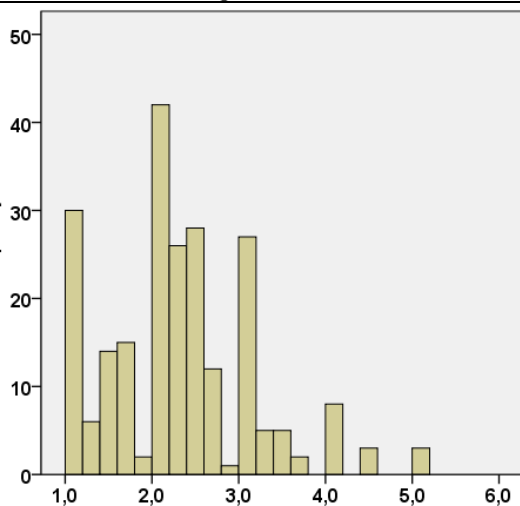
O Gráfico 133 apresenta os resultados encontrados na variável FEP47 onde estão as informações sobre o tempo usado nas disciplinas relacionadas com as proficiências avaliadas. A média de 3,7 e o desvio padrão de 0,8 mostram que em 70% das escolas esse tempo é de 60%.

Gráfico 133 - Frequência da variável FEP47



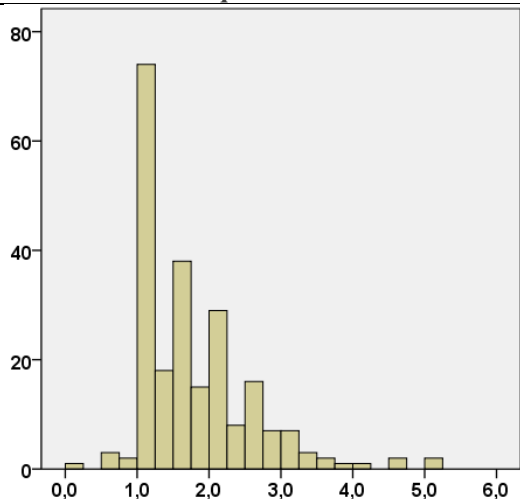
A variável FEP48 traz informações sobre o tempo usado nas disciplinas não relacionadas com as proficiências medidas e os seus resultados estão apresentados no Gráfico 134. A média de 2,3 e o desvio padrão de 0,9 indicam que em 70% das escolas o tempo destinado a essas disciplinas são de aproximadamente 30%.

Gráfico 134 - Frequência da variável FEP48



O tempo usado em disciplinas de formação complementar é controlado pela variável FEP49, cuja distribuição de frequências observadas está apresentada no Gráfico 135. A média de 1,7 e o desvio padrão de 0,87 sugerem que indicam que em 70% das escolas, o tempo destinado a disciplinas com formação complementar é de aproximadamente 30%.

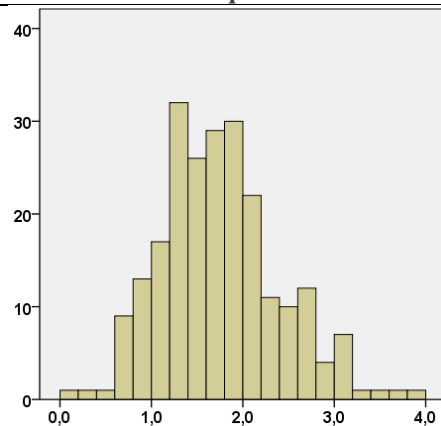
Gráfico 135 - Frequência da variável FEP49



As variáveis FEP50 e FEP51 resultam de um mesmo constructo na AFE denominado planejamento do tempo para as disciplinas. As alternativas de respostas para um dia normal de aulas nos itens do questionário e os valores correspondentes nos arquivos são: “Nenhuma” (0), “Uma vez” (1), “Duas vezes” (2), “três a quatro vezes” (3), “cinco a seis vezes” (4), e “sete ou mais vezes” (5).

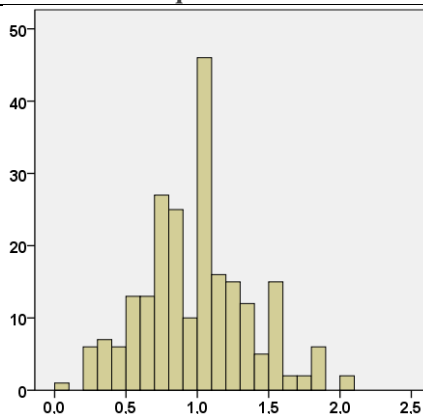
A interrupção das aulas por indisciplina é controlada pela variável FEP50. A média de 1,7 e o desvio padrão de 0,7 observados indicam que em 70% das escolas essas interrupções ocorrem em torno de duas vezes ao dia. O Gráfico 136 representa os valores observados nessa variável.

Gráfico 136 - Frequência da variável FEP50



No Gráfico 137 está representada a distribuição de frequência da variável FEP51 que controla a quantidade de interrupção das aulas por necessidade. A média de 1,0 e o desvio padrão de 0,4 sugerem que em 70% das escolas interrupções dessa natureza ocorrem duas vezes ao dia.

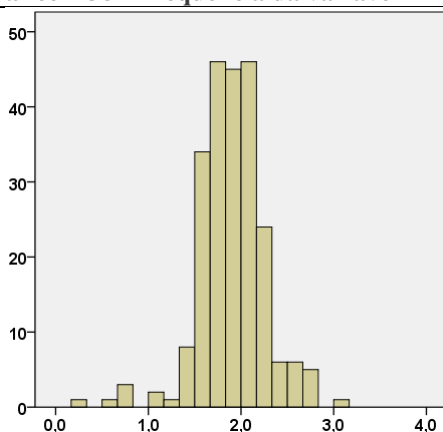
Gráfico 137 - Frequência da variável FEP51



As variáveis FEP52, FEP53 e FEP54 resultam de um mesmo constructo na AFE denominado “uso de recursos técnico pedagógicos”. Os valores correspondentes no arquivo e os itens que os questionários tinham como alternativas de respostas que variam de 0 a 4, sendo que cada uma delas significa “Nunca”, “Raramente”, “Algumas vezes no bimestre”, “Cerca de uma vez por semana” e “Várias vezes por semana”, respectivamente.

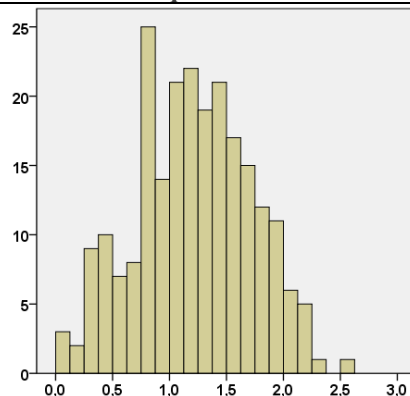
A distribuição da frequência dos valores observados na variável FEP52 que controla o uso de recursos audiovisuais está representada no Gráfico 138. A média de 1,9 e o desvio padrão de 0,4 mostra que os recursos como videocassete e DVD são usados por aproximadamente 70% das escolas algumas vezes no bimestre.

Gráfico 138 - Frequência da variável FEP52



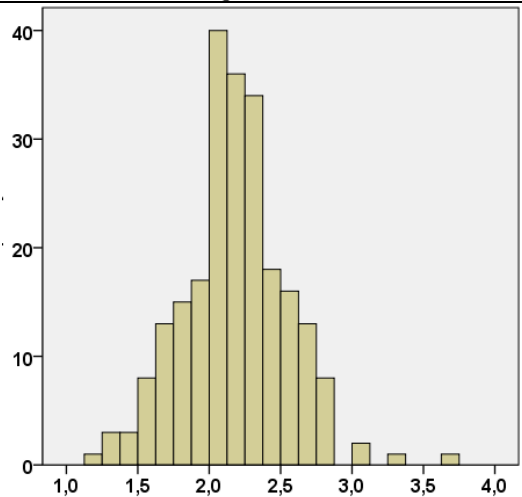
Os recursos como terrários, aquários, projetores e computadores são usados também algumas vezes no bimestre. Essa constatação é obtida pela média de 2,1 e o desvio padrão de 0,4 observados na variável FEP53, cujos resultados estão apresentados no Gráfico 139.

Gráfico 139 - Frequência da variável FEP53



No Gráfico 140 está representada a distribuição de frequência da variável FEP54, que controla o uso de materiais concretos de matemática e geografia. A média de 0,7 e o desvio padrão de 0,1 indicam que em 70% das escolas esses materiais são usados raramente.

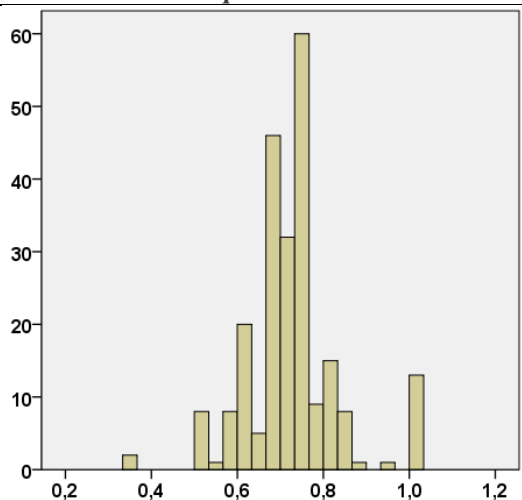
Gráfico 140 - Frequência da variável FEP54



As próximas variáveis caracterizam as características médias dos professores.

A variável FEP55 controla a proporção de professores do sexo feminino na escola, cujos valores podem variar de 0 a 1, onde se constata que em 70% das escolas o corpo docente é composto por 70% de mulheres. Os dados mostram que em 5,7% das escolas todos os docentes são mulheres, como pode ser observado no Gráfico 141.

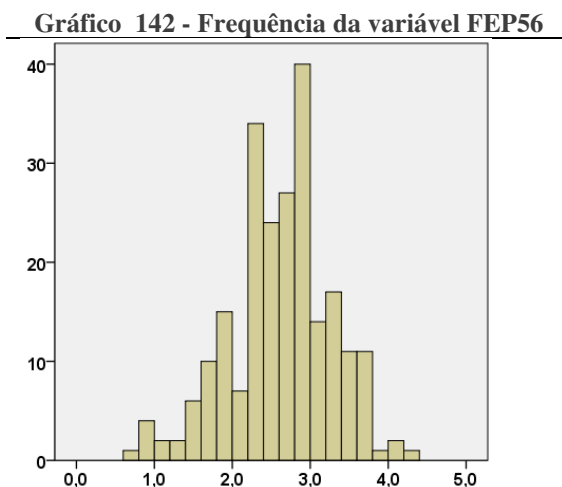
Gráfico 141 - Frequência da variável FEP55



A variável FEP36 controla a faixa etária dos professores da escola. No item do questionário as alternativas eram: (0) Até 24 anos; (1) De 25 a 29 anos; (2) De 30 a 39 anos; (3) De 40 a 49 anos; (4) De 50 a 54 anos e (5) Mais de 55 anos.

Em média as escolas têm docentes com idade que oscilam entre 25 e 39 anos, dada a média de 2,6 e o desvio padrão de 0,7 observados nessa variável.

A distribuição das frequência observadas está apresentada no Gráfico 142.



O Gráfico 143 mostra a distribuição de frequência da variável FEP57 controla a renda familiar média dos professores. No item do questionário as alternativas e os valores correspondentes no arquivo eram:

- 0) Até R\$ 300,00
- 1) De R\$ 301,00 a R\$ 500,00
- 2) De R\$ 501,00 a R\$ 700,00
- 3) De R\$ 701,00 a R\$ 900,00
- 4) De R\$ 901,00 a R\$ 1.100,00
- 5) De R\$ 1.101,00 a R\$ 1.300,00
- 6) De R\$ 1.301,00 a R\$ 1.500,00
- 7) De R\$ 1.501,00 a R\$ 1.700,00
- 8) De R\$ 1.701,00 a R\$ 1.900,00
- 9) De R\$ 1.901,00 a R\$ 2.300,00

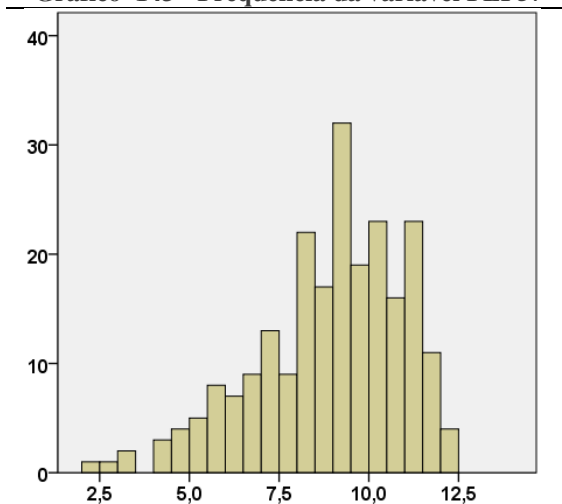
10) De R\$ 2.301,00 a R\$ 2.700,00

11) De R\$ 2.701,00 a R\$ 3.100,00

12) Mais de R\$ 3.100,00

A média de 8,8 e o desvio padrão de 2,0 observados na variável mostra que em 70% das escolas, a renda varia de R\$ 1501,00 a R\$ 3100,00.

Gráfico 143 - Frequência da variável FEP57

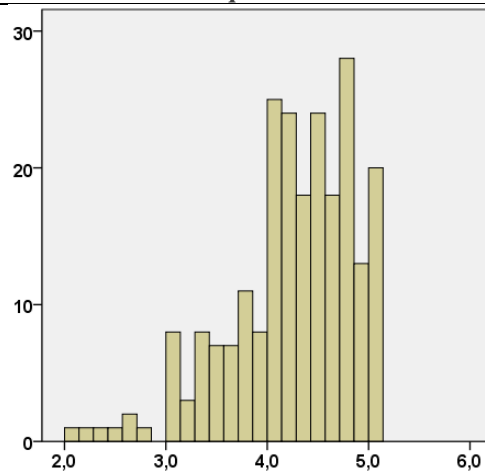


A variável FEP58 controla o tempo de exercício médio do professor na ocupação acesso a atividades culturais que o professore teve em um ano. As alternativas propostas no item do questionário que coletou essa informação e os valores correspondentes no arquivo são:

- 0) Há menos de 1 ano
- 1) De 1 a 2 anos
- 2) De 3 a 4 anos
- 3) De 5 a 10 anos
- 4) De 11 a 15 anos
- 5) Há mais de 15 anos

A média de 4,2 e o desvio padrão de 06 indicam que em 70% dos professores exercem a profissão de 11 a 15 anos. O Gráfico 144 apresenta os dados encontrados nessa variável.

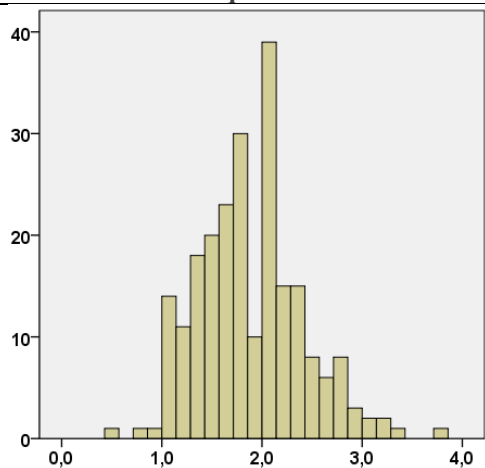
Gráfico 144 - Frequência da variável FEP58



No Gráfico 145 está representada a distribuição de frequência da variável FEP59, que controla o número médio de pessoas da família moram com o professor. As respostas oferecidas no item do questionário que coletou essa informação e os valores correspondentes no arquivo são: (0) Moro sozinho(a), (1) 2 pessoas, (2) 3 pessoas; (3) 4 pessoas; (4) 5 pessoas; (5) 6 pessoas e (6) 7 pessoas ou mais

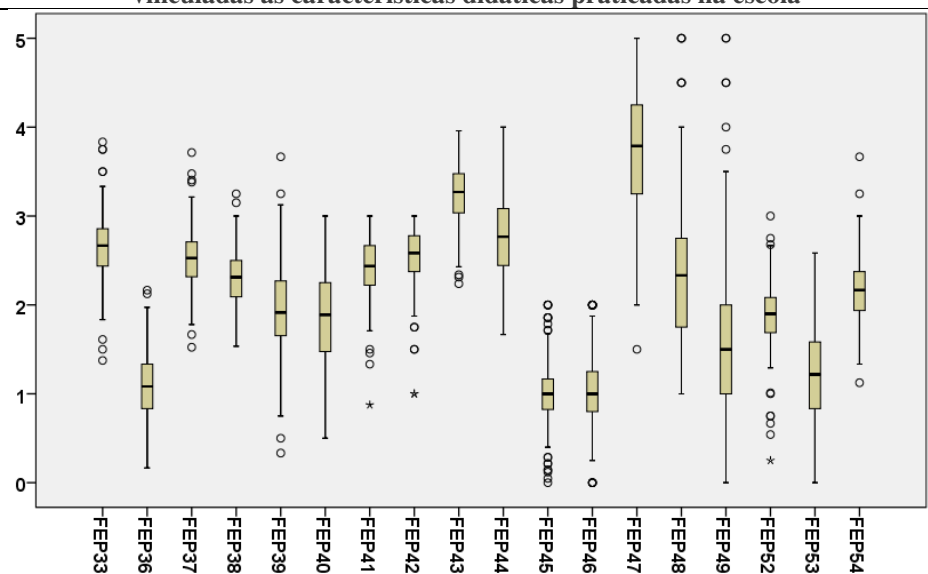
A média de 1,9 e o desvio padrão de 0,5 sugere que em 70% das escolas, o professor mora com mais três pessoas.

Gráfico 145 - Frequência da variável FEP59



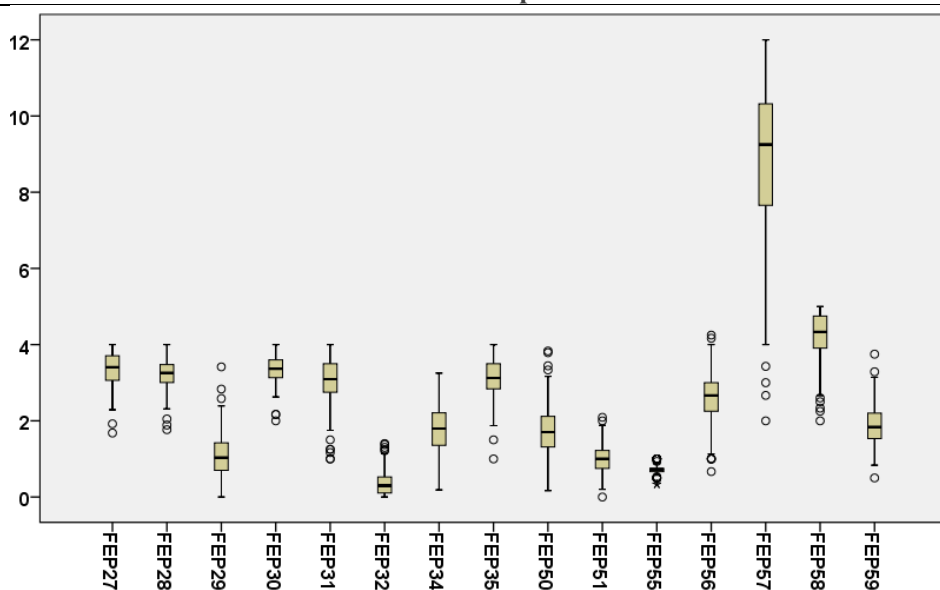
O Gráfico 146 apresenta uma síntese das variáveis de Nível 3 vinculadas às práticas didáticas da escola oriundas do arquivo que armazena as informações de contexto dos professores, que agrupados representam uma média da escola.

Gráfico 146 - Boxplot dos resultados encontrados nas variáveis de Nível 3 vinculadas às características didáticas praticadas na escola



O Gráfico 147, de forma análoga, apresenta uma síntese do intervalo de valores encontrados das variáveis de Nível 3 vinculadas às características organizacionais da escola, cujas informações também se originam do arquivo de professor.

Gráfico 147 - Boxplot resultados encontrados nas variáveis de Nível 3 vinculadas às características didáticas praticadas na escola



O Quadro 101 apresenta a síntese de todas as variáveis do Nível 3.

Quadro 101 - Síntese das variáveis do arquivo de Nível 3

(continua)

Identificação		Limites possíveis		Valores Observados		Média	Desvio Padrão
Nível 3	AFE	Inf	Sup	Min	Max		
FEA01	FA01	-3	3	-0,46	1,28	0,18	0,44
FEE02	FE02	0	3	0,00	1,70	0,29	0,38
FEE03	FE03	0	3	0,00	3,00	0,77	0,91
FEE04	FE04	0	3	0,00	3,00	0,84	1,04
FEE05	FE05	0	3	0,00	3,00	2,13	1,02
FEE06	FE06	0	3	0,00	3,00	1,42	0,99
FEE07	FE07	0	2	0,00	2,00	0,54	0,74
FEE08	FE08	0	2	0,00	2,00	1,17	0,85
FEE09	FE09	0	2	0,00	2,00	1,52	0,62
FEE10	FE10	0	2	0,00	2,00	1,71	0,64
FEE11	FE11	0	2	0,00	2,00	0,41	0,78
FED12	FD01	0	2	0,00	1,43	0,22	0,29
FED13	FD02	0	2	0,00	2,00	0,63	0,44
FED14	FD03	0	2	0,00	2,00	1,12	0,53
FED15	FD04	0	3	0,00	3,00	1,58	0,56
FED16	FD05	0	3	1,00	3,00	2,69	0,37
FED17	FD06	0	3	0,50	2,50	1,25	0,35
FED18	FD07	0	4	0,00	4,00	1,48	1,05
FED19	FD08	0	4	0,00	4,00	1,87	0,84
FED20*	FD09	0	1	0,00	1,00	0,39	0,36
FED21*	FD10	0	1	0,00	1,00	0,19	0,33
FED22*	FD11	0	1	0,00	1,00	0,75	0,31
FED23*	FD12	0	3	0,00	3,00	2,42	0,63
FED24*	FD13	0	4	0,00	4,00	1,56	1,20
FED25*	FD14	0	4	0,00	4,00	3,74	0,50
FED26*	FD15	0	7	0,00	7,00	4,40	1,41
FEP27	FP01	0	4	1,68	4,00	3,35	0,43
FEP28	FP02	0	4	1,77	4,00	3,21	0,39
FEP29	FP03	0	4	0,00	3,42	1,07	0,53
FEP30	FP04	0	4	2,00	4,00	3,36	0,33
FEP31	FP05	0	4	1,00	4,00	3,04	0,57
FEP32	FP06	0	2	0,00	1,41	0,35	0,30
FEP33	FP07	0	4	1,38	3,83	2,64	0,35
FEP34	FP08	0	4	0,19	3,25	1,78	0,60
FEP35	FP09	0	4	1,00	4,00	3,12	0,51
FEP36	FP10	0	3	0,17	2,17	1,10	0,37
FEP37	FP11	0	4	1,52	3,71	2,52	0,30
FEP38	FP12	0	4	1,54	3,25	2,31	0,30
FEP39	FP13	0	4	0,33	3,67	1,95	0,50
FEP40	FP14	0	3	0,50	3,00	1,86	0,55
FEP41	FP15	0	3	0,88	3,00	2,43	0,34
FEP42	FP16	0	3	1,00	3,00	2,55	0,34
FEP43	FP17	0	4	2,24	3,96	3,24	0,33
FEP44	FP18	0	4	1,67	4,00	2,76	0,46
FEP45	FP23	0	2	0,00	2,00	1,01	0,36

(conclusão)

Identificação		Limites possíveis		Valores Observados		Média	Desvio Padrão
Nível 3	AFE	Inf	Sup	Min	Max		
FEP46	FP24	0	2	0,00	2,00	1,03	0,40
FEP47	FP25	0	5	1,50	5,00	3,76	0,79
FEP48	FP26	0	5	1,00	5,00	2,28	0,85
FEP49	FP27	0	5	0,00	5,00	1,70	0,80
FEP50	FP28	0	5	0,17	3,83	1,75	0,65
FEP51	FP29	0	5	0,00	2,09	0,98	0,37
FEP52	FP30	0	4	0,25	3,00	1,88	0,36
FEP53	FP31	0	4	0,00	2,58	1,21	0,51
FEP54	FP32	0	4	1,13	3,67	2,16	0,36
FEP55*	FP35*	0	1	0,33	1,00	0,72	0,11
FEP56*	FP36*	0	5	0,67	4,25	2,63	0,65
FEP57*	FP37*	0	12	2,00	12,00	8,83	2,05
FEP58*	FP38*	0	5	2,00	5,00	4,23	0,61
FEP59*	FP39*	0	6	0,50	3,75	1,87	0,52

ANEXOS

ANEXO 01 - Interpretação pedagógica dos níveis de proficiência em Leitura

Nível 1: Neste nível os alunos reconhecem a letra do alfabeto e a direção da escrita. São capazes de ler frases tanto com estrutura canônica (sujeito + verbo + objeto) quanto frases que fogem a esse padrão associando-as a uma imagem. Além disso, já têm capacidade de localizar uma informação em rótulo. Estas são consideradas habilidade básicas, e sua consolidação é fundamental para que os alunos progridam em seu desenvolvimento como leitores, visto que se colocam como condições essenciais para que leiam com compreensão e de forma autônoma.

Nível 2: Neste nível os alunos demonstram reconhecer sílabas de uma palavra. No Domínio Estratégia de Leitura, os alunos são capazes de localizar informações num texto curto (com até dez linhas) com poucas informações e linguagem familiar em um texto de extensão mediana (com até quinze linhas). Identificam o assunto, indicando pelo título, de um texto informativo e, também, de uma reportagem a partir das pistas verbais e não verbais da capa de uma revista. Além disso, são capazes de inferir uma informação em uma tirinha e de identificar a finalidade de um bilhete.

Nível 3: Nesse nível, os alunos localizam informações em um texto de gêneros mais variados (conto de extensão mediana, textos informativos curtos e de extensão mediana e mapas de tempo). Identificam o assunto indicado pelo título do texto informativo curto e de um folheto de divulgação de uma campanha, por meio da associação da linguagem verbal e não verbal. São capazes de inferir uma informação, em texto de gênero menos familiar, apoiando-se em pistas não verbais, e também o sentido de uma palavra em um poema com linguagem simples. Ainda, identificam a finalidade de um texto instrucional. Eles revelam uma expansão de suas habilidades de leitura, agregando habilidades relacionadas ao Domínio Processamento do Texto. Nesse domínio, os alunos estabelecem relações lógico-discursivas entre partes de um texto por meio de formas verbais, identificando um referente distante, bem como relação de causa-consequência, sem marcação explícita, em um conto com temática familiar. Além disso, criam relações entre convites, reconhecendo-lhes diferenças.

Nível 4: Nesse nível os alunos ampliam as habilidades relacionadas à relação de inferência, mostrando-se capaz de inferir uma informação em uma tirinha (com linguagem não verbal e com linguagem mista) e uma ficha de identificação de espécie da fauna. Inferem o sentido de uma palavra em uma receita. Reconhecem, em uma tirinha, efeitos de humor, o

efeito de sentido do uso de reticência e de onomatopeia. Identificam a finalidade de texto de gênero pouco familiar, uma ficha de identificação de espécie da fauna. Estabelecem relações entre textos de mesmo gênero, reconhecendo diferenças quanto à informação. São capazes também de criar relações entre partes de um texto por meio de substituição lexical e de causa-consequência, em um poema curto e em texto de extensão mediana. Além disso, identificam interlocutores em uma fábula e as marcas explícitas que os evidenciam em um discurso. Ainda com relação aos elementos de uma narrativa, revelam ser capazes de identificar o narrador de um poema com temática e vocabulário familiar, bem como a situação problemática em um fábula extensa.

Nível 5: Neste nível, na competência Realiza Inferência, os alunos inferem informação em texto com linguagem mista em torno de extensão mediana (até quinze linhas) com linguagem simples e em texto informativo curto que utiliza elementos gráficos (códigos de barras), além de reconhecerem o efeito de humor em uma anedota. Identificam a finalidade de uma fábula e de uma anedota. Estabelecem relações lógico-discursivas de tempo, em um poema, marcadas por uma locução adverbias e de causa-consequência em uma lenda. Reconhecem a opinião de um personagem em discurso direto, e do narrador. Identificam, ainda, a ordem em que os fatos ocorrem em uma narrativa literária, a caracterização de uma personagem em um conto conhecido, os interlocutores, em um discurso direto, por meio das marcas explícitas que os caracterizam, bem como o narrador em um conto mais extenso (até trinta linhas). Observa-se, neste nível, um crescimento tanto no que se diz respeito ao aumento das habilidades quanto no que concerne à complexidade dessa habilidade, as quais exigem maior refinamento dos processos cognitivos envolvidos no desenvolvimento da literatura, pressupondo-se assim, um aluno-leitor que ampliar gradativamente seu repertório de leituras.

Nível 6: Nesse nível, os alunos localizam informações em um texto informativo curto com vocabulário mais sofisticado em texto expositivo-argumentativo. Identificam assunto de um texto informativo e de uma letra de música. Inferem, em um poema, o sentido de uma palavra, bem como reconhecem marcas de ironia em texto com linguagem simples e temática familiar e efeito de sentido do emprego de um diminutivo. Além disso, estabelecem relações lógico-discursivas de causa-consequência em texto informativo com vocabulário mais sofisticado e em um conto. São capazes, também, de reconhecer o ponto de vista do

enunciador por meio do emprego de um adjetivo. Acrescente-se, ainda, que reconhecem marcas linguísticas características de situações informais.

ANEXO 02 - Interpretação pedagógica dos níveis de proficiência em Matemática

Nível 1: Nesse nível, os alunos comparam pequenas quantidades com o uso de recursos gráficos. Eles destacam o que possui a maior quantidade e ainda a altura de objetos, indicando o mais baixo e o mais alto. Identificam o símbolo numérico (números com 2 algarismos) e comparam números naturais de dois algarismos, com e sem apoio gráfico. Realizam a contagem seletiva, conseguindo associar quantidades aos números. Além disso, são capazes de coordenar as ações de contar e juntar quantidades para resolver situações – problema simples a fim de determinar o total até 20 e resolvem problemas envolvendo as ideias de contar e retirar uma quantidade de outra (mimando até 10), a partir de apoio gráfico. A contagem é uma atividade que envolve aspectos cognitivos e motores. É cognitiva, por que repousa sobre um conhecimento abstrato que diz respeito à ordem e à cardinalidade. Estas habilidades muito elementares, na maioria das vezes, são construídas por meio de conhecimento adquiridos socialmente.

Nível 2: Nesse nível, os alunos localizam um objetivo entre dois outros, indicam seus tamanhos, apontando qual deles é o menor, o maior ou o médio. Identificam a ordem crescente de grupos de objetos dispostos aleatoriamente. Agrupam pequenas quantidades em unidades e dezenas com apoio gráfico ou utilizando o sistema monetário brasileiro. Resolvem problemas de adição (ação juntar) e subtração (ação de retirar e completar), sem apoio gráfico. Neste nível, eles já identificam a operação de subtração como a solução de uma dada situação, além de ler e interpretar informações em quadros ou tabelas simples.

Nível 3: Neste nível, os alunos evidenciam maior expansão do campo numérico que atinge o grupo de números representado por três ou quatro algarismos. Além de identificar esses números associando a escrita por extenso ao símbolo numérico, eles demonstram ser capazes de identificar o antecessor de um número e realizar a sua decomposição. Também resolvem problemas mais complexos envolvendo a subtração com a ideia de complementação, comparação e equalização, de multiplicação envolvendo o princípio multiplicativo e de divisão com o significado de repartir, que se encontram em processos mais avançado de construção e são resolvidos quando inseridos em contextos, ou que indica que a criança tem a compreensão da ação operatória.

Nível 4: Neste nível os alunos resolvem problemas envolvendo a composição e a decomposição de valores monetários, habilidade essa decorrente de outras sedimentadas

anteriormente. Afinal, compor e decompor quantias em reais tem suporte na composição e decomposição de números naturais, bem como na troca de valores monetários. As habilidades importantes pertinentes ao campo de Grandezas e Medidas, apresentadas nesse nível, também evidenciam o trabalho com a matemática escolar. Os alunos comparam unidades de medidas aplicadas a grandezas mensuráveis presentes no contexto diário, tais como massa, capacidade de medir de tempo em situações-problema envolvendo mês e ano. Neste nível constata-se que os alunos compreendem o Sistema de Numeração Decimal e o significado das operações fundamentais na resolução e problemas, apontando para uma expansão do conhecimento matemático necessário para as séries iniciais do Ensino Fundamental. Nota-se que os alunos situados nesse nível resolvem problemas que envolvem subtração de números naturais em situações mais complexas, com estado inicial desconhecido, inserindo números com ideias de complementação e comparação, bem como problemas que contêm multiplicação de números naturais e divisão exata de um número natural por outro de dois algarismos com resto. Há crescimento não apenas no sentido de ampliação do leque de habilidades, mas também na profundidade e complexidade do item, que exige do aluno mais discernimento, perspicácia e desempenho.

Nível 5: Nesse nível, os alunos desenvolveram as habilidades sobre o trabalho realizado em sala de aula. Nota-se que há progressão na habilidade de resolver problemas numéricos envolvendo os diferentes significados das operações, como de equalização com a mudança de uma quantidade, para que as duas se tornem equivalentes, de comparar na subtração de racionais na forma de decimal, de proporcionalidade, assim como os que contêm a configuração retangular da multiplicação, presentes nos contextos dos problemas. Enfim, as evidências expressas nos resultados dos alunos que alcançaram uma proficiência média situada nesse nível demonstram que eles já possuem noções matemáticas mais elaboradas, o que deve ser feito do trabalho realizado na escola. Algumas habilidades, como a resolução de problemas que envolvem as operações fundamentais, ampliam-se e caminham para a consolidação. Os itens solucionados pelos alunos, alocados a partir desse nível contêm problemas que inserem operações com números naturais, racionais, sob a forma decimal e porcentagem, medidas de comprimento, massa e capacidade e suas aplicações, como o cálculo do perímetro e da área, assim como as medidas de tempo e valor.

ANEXO 03 - Relatório HLM – Leitura – Modelo Nulo

Program: HLM 6 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: HLM3.EXE (6.08.29257.1)
Date: 7 September 2011, Wednesday
Time: 7:36: 2

SPECIFICATIONS FOR THIS HLM3 RUN

Problem Title: no title

The data source for this run = Estudo_c_Ret.mdm

The command file for this run = C:\Users\Adilson Dalben\Documents\Meus Documentos
Adilson\DOCTORADO\00 - Tese_Banco-Geres\Estudo com retido\Modelo00-Nulo_c_P_R.hlm

Output file name = C:\Users\Adilson Dalben\Documents\Meus Documentos
Adilson\DOCTORADO\00 - Tese_Banco-Geres\Estudo com retido\hlm3.txt

The maximum number of level-1 units = 34657

The maximum number of level-2 units = 12180

The maximum number of level-3 units = 229

The maximum number of iterations = 100

Method of estimation: full maximum likelihood

Weighting Specification

Weight
Variable
Weighting? Name Normalized?
Level 1 no
Level 2 yes PESOAL yes
Level 3 no
Precision no

The outcome variable is LEITF

The model specified for the fixed effects was:

Level-1 Level-2 Level-3
Coefficients Predictors Predictors

INTRCPT1, P0 INTRCPT2, B00 INTRCPT3, G000

Summary of the model specified (in equation format)

Level-1 Model
Y = P0 + E

Level-2 Model

$$P0 = B00 + R0$$

Level-3 Model

$$B00 = G000 + U00$$

For starting values, data from 34657 level-1 and 12180 level-2 records were used

Iterations stopped due to small change in likelihood function

***** ITERATION 15 *****

Standard errors for sigma_squared, Tau(pi), Tau(beta) are not computable.

Sigma_squared = 422.65182

Tau(pi)

INTRCPT1,P0 191.54959

Tau(pi) (as correlations)

INTRCPT1,P0 1.000

Random level-1 coefficient Reliability estimate

INTRCPT1, P0 0.538

Tau(beta)

INTRCPT1
INTRCPT2,B00
346.43496

Tau(beta) (as correlations)

INTRCPT1/INTRCPT2,B00 1.000

Random level-2 coefficient Reliability estimate

INTRCPT1/INTRCPT2, B00 0.978

The value of the likelihood function at iteration 15 = -1.593720E+005

The outcome variable is LEITF

Final estimation of fixed effects

(with robust standard errors)

Fixed Effect Standard Approx.
 Coefficient Error T-ratio d.f. P-value

For INTRCPT1, P0
For INTRCPT2, B00
 INTRCPT3, G000 152.989297 3.967249 38.563 228 0.000

Final estimation of level-1 and level-2 variance components:

Random Effect		Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, level-1,	R0 E	13.84014 20.55850	191.54959 422.65182	11951	27828.50965	0.000

Final estimation of level-3 variance components:

Random Effect		Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1/INTRCPT2, U00		18.61276	346.43496	228	6398.93652	0.000

Statistics for current covariance components model

Deviance = 318744.036708
Number of estimated parameters = 4

ANEXO 04 - Relatório HLM – Leitura – Modelo de Referência

Program: HLM 6 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: HLM3.EXE (6.08.29257.1)
Date: 7 September 2011, Wednesday
Time: 7:43:33

SPECIFICATIONS FOR THIS HLM3 RUN

Problem Title: no title

The data source for this run = Estudo.mdm
The command file for this run = C:\Users\Adilson Dalben\Documents\DOCTORADO\Modelo-Ref.hlm
Output file name = C:\Users\Adilson Dalben\Documents\DOCTORADO\hlm3.txt
The maximum number of level-1 units = 34657
The maximum number of level-2 units = 12180
The maximum number of level-3 units = 229
The maximum number of iterations = 100
Method of estimation: full maximum likelihood

Weighting Specification

Weight
Variable
Weighting? Name Normalized?
Level 1 no
Level 2 yes PESOAL yes
Level 3 no
Precision no

The outcome variable is LEITF

The model specified for the fixed effects was:

Level-1 Level-2 Level-3
Coefficients Predictors Predictors

INTRCPT1, P0 INTRCPT2, B00 INTRCPT3, G000
T2 slope, P1 # INTRCPT2, B10 INTRCPT3, G100
T3 slope, P2 # INTRCPT2, B20 INTRCPT3, G200
T4 slope, P3 # INTRCPT2, B30 INTRCPT3, G300
LEITI slope, P4 INTRCPT2, B40 INTRCPT3, G400
LEITIXT2 slope, P5 # INTRCPT2, B50 INTRCPT3, G500
LEITIXT3 slope, P6 # INTRCPT2, B60 INTRCPT3, G600
LEITIXT4 slope, P7 # INTRCPT2, B70 INTRCPT3, G700
RET_AC slope, P8 # INTRCPT2, B80 INTRCPT3, G800

'#' - The residual parameter variance for the parameter has been set to zero

Summary of the model specified (in equation format)

Level-1 Model

$$Y = P0 + P1*(T2) + P2*(T3) + P3*(T4) + P4*(LEITI) + P5*(LEITIXT2) + P6*(LEITIXT3) + P7*(LEITIXT4) + P8*(RET_AC) + E$$

Level-2 Model

$$\begin{aligned} P0 &= B00 + R0 \\ P1 &= B10 \\ P2 &= B20 \\ P3 &= B30 \\ P4 &= B40 + R4 \\ P5 &= B50 \\ P6 &= B60 \\ P7 &= B70 \\ P8 &= B80 \end{aligned}$$

Level-3 Model

$$\begin{aligned} B00 &= G000 + U00 \\ B10 &= G100 \\ B20 &= G200 \\ B30 &= G300 \\ B40 &= G400 + U40 \\ B50 &= G500 \\ B60 &= G600 \\ B70 &= G700 \\ B80 &= G800 \end{aligned}$$

For starting values, data from 30012 level-1 and 8655 level-2 records were used

Iterations stopped due to small change in likelihood function

***** ITERATION 52 *****

Standard errors for sigma_squared, Tau(pi), Tau(beta) are not computable.

Sigma_squared = 184.94396

Tau(pi)

INTRCPT1,P0	110.17466	-0.60527
LEITI,P1	-0.60527	0.00343

Tau(pi) (as correlations)

INTRCPT1,P0	1.000	-0.984
LEITI,P4	-0.984	1.000

Random level-1 coefficient Reliability estimate

INTRCPT1, P0	0.048
LEITI, P4	0.028

Note: The reliability estimates reported above are based on only 8655 of 12180 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Tau(beta)
 INTRCPT1 LEITI
 INTRCPT2,B00 INTRCPT2,B40
 249.75333 -1.08385
 -1.08385 0.00509

Tau(beta) (as correlations)
 INTRCPT1/INTRCPT2,B00 1.000 -0.961
 LEITI/INTRCPT2,B40 -0.961 1.000

Random level-2 coefficient	Reliability estimate
INTRCPT1/INTRCPT2, B00	0.836
LEITI/INTRCPT2, B40	0.722

The value of the likelihood function at iteration 52 = -1.411130E+005
 The outcome variable is LEITF

Final estimation of fixed effects
 (with robust standard errors)

Fixed Effect	Standard Coefficient	Standard Error	Approx. T-ratio	d.f.	P-value
For INTRCPT1, P0					
For INTRCPT2, B00					
INTRCPT3, G000	63.956412	4.679876	13.666	228	0.000
For T2 slope, P1					
For INTRCPT2, B10					
INTRCPT3, G100	-26.676918	3.322691	-8.029	34648	0.000
For T3 slope, P2					
For INTRCPT2, B20					
INTRCPT3, G200	-8.259770	2.803781	-2.946	34648	0.004
For T4 slope, P3					
For INTRCPT2, B30					
INTRCPT3, G300	-1.023668	3.638341	-0.281	34648	0.778
For LEITI slope, P4					
For INTRCPT2, B40					
INTRCPT3, G400	0.534836	0.025440	21.023	228	0.000
For LEITIXT2 slope, P5					
For INTRCPT2, B50					
INTRCPT3, G500	0.260323	0.025740	10.114	34648	0.000
For LEITIXT3 slope, P6					
For INTRCPT2, B60					
INTRCPT3, G600	0.148268	0.021979	6.746	34648	0.000
For LEITIXT4 slope, P7					
For INTRCPT2, B70					

INTRCPT3, G700	0.111320	0.025914	4.296	34648	0.000
For RET_AC slope, P8					
For INTRCPT2, B80					
INTRCPT3, G800	5.627392	1.543986	3.645	34648	0.000

Final estimation of level-1 and level-2 variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, R0	10.49641	110.17466	8426	8742.71411	0.008
LEITI slope, R4	0.05858	0.00343	8426	8751.99848	0.007
level-1, E	13.59941	184.94396			

Note: The chi-square statistics reported above are based on only 8655 of 12180 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Final estimation of level-3 variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1/INTRCPT2, U00	15.80359	249.75333	228	1036.77550	0.000
LEITI/INTRCPT2, U40	0.07133	0.00509	228	873.62912	0.000

Statistics for current covariance components model

Deviance = 282225.968490
Number of estimated parameters = 16

Model comparison test

Chi-square statistic = 36518.06822
Number of degrees of freedom = 12
P-value = 0.000

ANEXO 05 - Relatório HLM – Leitura – Modelo Nível 2

Program: HLM 6 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: HLM3.EXE (6.08.29257.1)
Date: 12 September 2011, Monday
Time: 15:40: 5

SPECIFICATIONS FOR THIS HLM3 RUN

Problem Title: no title

The data source for this run = Estudo.mdm
The command file for this run = C:\Users\Adilson Dalben\Documents\DOCTORADO\Modelo-Nivel2.hlm
Output file name = C:\Users\Adilson Dalben\Documents\DOCTORADO\hlm3.txt
The maximum number of level-1 units = 34657
The maximum number of level-2 units = 12180
The maximum number of level-3 units = 229
The maximum number of iterations = 100
Method of estimation: full maximum likelihood

Weighting Specification

	Weight	Variable	
	Weighting?	Name	Normalized?
Level 1	no		
Level 2	yes	PESOAL	yes
Level 3	no		
Precision	no		

The outcome variable is LEITF

The model specified for the fixed effects was:

Level-1	Level-2	Level-3
Coefficients	Predictors	Predictors
	INTRCPT1, P0	INTRCPT2, B00 INTRCPT3, G000
	# FA04_PRE, B01	INTRCPT3, G010
	# FA07, B02	INTRCPT3, G020
	# FA09, B03	INTRCPT3, G030
	# FA12, B04	INTRCPT3, G040
	# FA14, B05	INTRCPT3, G050
# T2 slope, P1	# INTRCPT2, B10	INTRCPT3, G100
	# FA04_PRE, B11	INTRCPT3, G110
	# FA07, B12	INTRCPT3, G120
	# FA09, B13	INTRCPT3, G130
	# FA10, B14	INTRCPT3, G140
# T3 slope, P2	# INTRCPT2, B20	INTRCPT3, G200

```

# FA11, B21 INTRCPT3, G210
# FA13, B22 INTRCPT3, G220
# T4 slope, P3 # INTRCPT2, B30 INTRCPT3, G300
#%FA06_NSE, B31 INTRCPT3, G310
# FA07, B32 INTRCPT3, G320
# FA09, B33 INTRCPT3, G330
LEITI slope, P4 INTRCPT2, B40 INTRCPT3, G400
#%FA06_NSE, B41 INTRCPT3, G410
# FA08, B42 INTRCPT3, G420
# FA12, B43 INTRCPT3, G430
# FA13, B44 INTRCPT3, G440
# FA14, B45 INTRCPT3, G450
# LEITIXT2 slope, P5 # INTRCPT2, B50 INTRCPT3, G500
# FA04_PRE, B51 INTRCPT3, G510
#%FA06_NSE, B52 INTRCPT3, G520
# FA08, B53 INTRCPT3, G530
# FA09, B54 INTRCPT3, G540
# FA10, B55 INTRCPT3, G550
# LEITIXT3 slope, P6 # INTRCPT2, B60 INTRCPT3, G600
#%FA06_NSE, B61 INTRCPT3, G610
# FA08, B62 INTRCPT3, G620
# LEITIXT4 slope, P7 # INTRCPT2, B70 INTRCPT3, G700
#%FA06_NSE, B71 INTRCPT3, G710
# FA07, B72 INTRCPT3, G720
# FA08, B73 INTRCPT3, G730
# FA09, B74 INTRCPT3, G740
# FA14, B75 INTRCPT3, G750
# RET_AC slope, P8 # INTRCPT2, B80 INTRCPT3, G800
# FA01_SEX, B81 INTRCPT3, G810
# FA11, B82 INTRCPT3, G820

```

'#' - The residual parameter variance for the parameter has been set to zero

'%' - This variable has been centered around its grand mean

Summary of the model specified (in equation format)

Level-1 Model

$$Y = P0 + P1*(T2) + P2*(T3) + P3*(T4) + P4*(LEITI) + P5*(LEITIXT2) + P6*(LEITIXT3) + P7*(LEITIXT4) + P8*(RET_AC) + E$$

Level-2 Model

$$\begin{aligned}
P0 &= B00 + B01*(FA04_PRE) + B02*(FA07) + B03*(FA09) + B04*(FA12) \\
&\quad + B05*(FA14) + R0 \\
P1 &= B10 + B11*(FA04_PRE) + B12*(FA07) + B13*(FA09) + B14*(FA10) \\
P2 &= B20 + B21*(FA11) + B22*(FA13) \\
P3 &= B30 + B31*(FA06_NSE) + B32*(FA07) + B33*(FA09) \\
P4 &= B40 + B41*(FA06_NSE) + B42*(FA08) + B43*(FA12) + B44*(FA13) \\
&\quad + B45*(FA14) + R4 \\
P5 &= B50 + B51*(FA04_PRE) + B52*(FA06_NSE) + B53*(FA08) + B54*(FA09) \\
&\quad + B55*(FA10) \\
P6 &= B60 + B61*(FA06_NSE) + B62*(FA08) \\
P7 &= B70 + B71*(FA06_NSE) + B72*(FA07) + B73*(FA08) + B74*(FA09)
\end{aligned}$$

$$+ B75*(FA14)$$
$$P8 = B80 + B81*(FA01_SEX) + B82*(FA11)$$

Level-3 Model

B00 = G000 + U00
B01 = G010
B02 = G020
B03 = G030
B04 = G040
B05 = G050
B10 = G100
B11 = G110
B12 = G120
B13 = G130
B14 = G140
B20 = G200
B21 = G210
B22 = G220
B30 = G300
B31 = G310
B32 = G320
B33 = G330
B40 = G400 + U40
B41 = G410
B42 = G420
B43 = G430
B44 = G440
B45 = G450
B50 = G500
B51 = G510
B52 = G520
B53 = G530
B54 = G540
B55 = G550
B60 = G600
B61 = G610
B62 = G620
B70 = G700
B71 = G710
B72 = G720
B73 = G730
B74 = G740
B75 = G750
B80 = G800
B81 = G810
B82 = G820

For starting values, data from 30012 level-1 and 8655 level-2 records were used

Iterations stopped due to small change in likelihood function

***** ITERATION 57 *****

Standard errors for sigma_squared, Tau(pi), Tau(beta) are not computable.

Sigma_squared = 170.97253

Tau(pi)

INTRCPT1,P0 86.99567 -0.44829
LEITI,P1 -0.44829 0.00248

Tau(pi) (as correlations)

INTRCPT1,P0 1.000 -0.966
LEITI,P4 -0.966 1.000

Random level-1 coefficient	Reliability estimate
INTRCPT1, P0	0.041
LEITI, P4	0.022

Note: The reliability estimates reported above are based on only 8655 of 12180 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Tau(beta)

INTRCPT1 LEITI
INTRCPT2,B00 INTRCPT2,B40
175.34289 -0.82167
-0.82167 0.00410

Tau(beta) (as correlations)

INTRCPT1/INTRCPT2,B00 1.000 -0.969
LEITI/INTRCPT2,B40 -0.969 1.000

Random level-2 coefficient	Reliability estimate
INTRCPT1/INTRCPT2, B00	0.802
LEITI/INTRCPT2, B40	0.698

The value of the likelihood function at iteration 57 = -1.398592E+005

The outcome variable is LEITF

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Standard Coefficient	Standard Error	Approx. T-ratio	d.f.	P-value
For INTRCPT1, P0					
For INTRCPT2, B00					
INTRCPT3, G000	60.358064	3.927039	15.370	228	0.000
For FA04_PRE, B01					
INTRCPT3, G010	2.005712	0.459187	4.368	12174	0.000

For FA07, B02						
INTRCPT3, G020	4.503866	0.394784	11.408	12174	0.000	
For FA09, B03						
INTRCPT3, G030	-2.289278	0.598164	-3.827	12174	0.000	
For FA12, B04						
INTRCPT3, G040	6.004041	1.527733	3.930	12174	0.000	
For FA14, B05						
INTRCPT3, G050	-3.924416	1.300164	-3.018	12174	0.003	
For T2 slope, P1						
For INTRCPT2, B10						
INTRCPT3, G100	-16.520948	8.747607	-1.889	34615	0.058	
For FA04_PRE, B11						
INTRCPT3, G110	-17.654839	6.215450	-2.840	34615	0.005	
For FA07, B12						
INTRCPT3, G120	2.450565	0.529578	4.627	34615	0.000	
For FA09, B13						
INTRCPT3, G130	-10.027613	4.966512	-2.019	34615	0.043	
For FA10, B14						
INTRCPT3, G140	11.644016	3.863217	3.014	34615	0.003	
For T3 slope, P2						
For INTRCPT2, B20						
INTRCPT3, G200	-10.186626	3.053597	-3.336	34615	0.001	
For FA11, B21						
INTRCPT3, G210	1.989575	0.860097	2.313	34615	0.021	
For FA13, B22						
INTRCPT3, G220	-3.117672	0.996583	-3.128	34615	0.002	
For T4 slope, P3						
For INTRCPT2, B30						
INTRCPT3, G300	5.534380	6.679316	0.829	34615	0.408	
For FA06_NSE, B31						
INTRCPT3, G310	6.353627	2.475192	2.567	34615	0.011	
For FA07, B32						
INTRCPT3, G320	9.081303	3.135097	2.897	34615	0.004	
For FA09, B33						
INTRCPT3, G330	-15.211943	3.585303	-4.243	34615	0.000	
For LEITI slope, P4						
For INTRCPT2, B40						
INTRCPT3, G400	0.441680	0.024545	17.995	228	0.000	
For FA06_NSE, B41						
INTRCPT3, G410	0.026758	0.004178	6.404	12174	0.000	
For FA08, B42						
INTRCPT3, G420	0.024382	0.004571	5.334	12174	0.000	
For FA12, B43						
INTRCPT3, G430	-0.021803	0.011310	-1.928	12174	0.053	
For FA13, B44						
INTRCPT3, G440	-0.008208	0.004349	-1.887	12174	0.059	
For FA14, B45						
INTRCPT3, G450	0.026445	0.010097	2.619	12174	0.009	
For LEITIXT2 slope, P5						
For INTRCPT2, B50						
INTRCPT3, G500	0.172961	0.073060	2.367	34615	0.018	
For FA04_PRE, B51						
INTRCPT3, G510	0.145958	0.052791	2.765	34615	0.006	
For FA06_NSE, B52						
INTRCPT3, G520	-0.015518	0.005617	-2.763	34615	0.006	
For FA08, B53						

INTRCPT3, G530	-0.015455	0.006519	-2.371	34615	0.018
For FA09, B54					
INTRCPT3, G540	0.067148	0.038593	1.740	34615	0.081
For FA10, B55					
INTRCPT3, G550	-0.085319	0.029493	-2.893	34615	0.004
For LEITIXT3 slope, P6					
For INTRCPT2, B60					
INTRCPT3, G600	0.185426	0.023626	7.849	34615	0.000
For FA06_NSE, B61					
INTRCPT3, G610	-0.017814	0.004323	-4.120	34615	0.000
For FA08, B62					
INTRCPT3, G620	-0.018133	0.006301	-2.878	34615	0.004
For LEITIXT4 slope, P7					
For INTRCPT2, B70					
INTRCPT3, G700	0.114087	0.046602	2.448	34615	0.015
For FA06_NSE, B71					
INTRCPT3, G710	-0.050586	0.015168	-3.335	34615	0.001
For FA07, B72					
INTRCPT3, G720	-0.052126	0.018628	-2.798	34615	0.006
For FA08, B73					
INTRCPT3, G730	-0.012100	0.005632	-2.148	34615	0.031
For FA09, B74					
INTRCPT3, G740	0.087121	0.024769	3.517	34615	0.001
For FA14, B75					
INTRCPT3, G750	-0.006328	0.003781	-1.674	34615	0.094
For RET_AC slope, P8					
For INTRCPT2, B80					
INTRCPT3, G800	8.267556	1.478621	5.591	34615	0.000
For FA01_SEX, B81					
INTRCPT3, G810	1.209754	0.274355	4.409	34615	0.000
For FA11, B82					
INTRCPT3, G820	-2.369643	0.496097	-4.777	34615	0.000

Final estimation of level-1 and level-2 variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, R0	9.32715	86.99567	8421	8467.89587	0.357
LEITI slope, R4	0.04975	0.00248	8421	8475.72000	0.335
level-1, E	13.07565	170.97253			

Note: The chi-square statistics reported above are based on only 8655 of 12180 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Final estimation of level-3 variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1/INTRCPT2, U00	13.24171	175.34289	228	787.27743	0.000

LEITI/INTRCPT2, U40 0.06403 0.00410 228 694.09467 0.000

Statistics for current covariance components model

Deviance = 279718.494750

Number of estimated parameters = 49

Model comparison test

Chi-square statistic = 2507.47374

Number of degrees of freedom = 33

P-value = 0.000

ANEXO 06 - Relatório HLM – Leitura – Modelo Nível 3

Program: HLM 6 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: HLM3.EXE (6.08.29257.1)
Date: 5 October 2011, Wednesday
Time: 11:14:16

SPECIFICATIONS FOR THIS HLM3 RUN

Problem Title: no title

The data source for this run = Estudo.mdm
The command file for this run = C:\Users\Adilson Dalben\Documents\DOUTORADO\Modelo-Nivel3.hlm
Output file name = C:\Users\Adilson Dalben\Documents\DOUTORADO\hlm3.txt
The maximum number of level-1 units = 34657
The maximum number of level-2 units = 12180
The maximum number of level-3 units = 229
The maximum number of iterations = 100
Method of estimation: full maximum likelihood

Weighting Specification

Weight
Variable
Weighting? Name Normalized?
Level 1 no
Level 2 yes PESOAL yes
Level 3 no
Precision no

The outcome variable is LEITF

The model specified for the fixed effects was:

Level-1 Level-2 Level-3
Coefficients Predictors Predictors

INTRCPT1, P0 INTRCPT2, B00 INTRCPT3, G000
 % FEA01, G001
 FEE02, G002
 FEE04, G003
 FEE05, G004
 FEE08, G005
 FED12, G006
 FED13, G007
 FED14, G008
 FED22, G009
 FED23, G0010

FED25, G0011
 FEP29, G0012
 FEP31, G0013
 FEP32, G0014
 FEP33, G0015
 FEP34, G0016
 FEP35, G0017
 FEP36, G0018
 FEP37, G0019
 FEP38, G0020
 FEP39, G0021
 FEP40, G0022
 FEP41, G0023
 FEP42, G0024
 FEP43, G0025
 FEP46, G0026
 FEP47, G0027
 FEP48, G0028
 FEP51, G0029
 FEP54, G0030
 FEP58, G0031
 # FA04_PRE, B01 INTRCPT3, G010
 # FA07, B02 INTRCPT3, G020
 # FA09, B03 INTRCPT3, G030
 # FA12, B04 INTRCPT3, G040
 # FA14, B05 INTRCPT3, G050
 # T2 slope, P1 # INTRCPT2, B10 INTRCPT3, G100
 % FEA01, G101
 FEE02, G102
 FEE03, G103
 FEE06, G104
 FEE07, G105
 FEE10, G106
 FED14, G107
 FED19, G108
 FED20, G109
 FED21, G1010
 FED24, G1011
 FED25, G1012
 FED26, G1013
 FEP29, G1014
 FEP31, G1015
 FEP32, G1016
 FEP33, G1017
 FEP34, G1018
 FEP35, G1019
 FEP38, G1020
 FEP39, G1021
 FEP40, G1022
 FEP45, G1023
 FEP48, G1024
 FEP50, G1025
 FEP52, G1026
 FEP53, G1027
 FEP54, G1028
 FEP55, G1029

```

FEP58, G1030
# FA04_PRE, B11 INTRCPT3, G110
# FA07, B12 INTRCPT3, G120
# FA09, B13 INTRCPT3, G130
# FA10, B14 INTRCPT3, G140
# T3 slope, P2 # INTRCPT2, B20 INTRCPT3, G200
    % FEA01, G201
    FEE02, G202
    FEE03, G203
    FED12, G204
    FED14, G205
    FED15, G206
    FED19, G207
    FED22, G208
    FED24, G209
    FED26, G2010
    FEP31, G2011
    FEP32, G2012
    FEP33, G2013
    FEP34, G2014
    FEP39, G2015
    FEP42, G2016
    FEP43, G2017
    FEP44, G2018
    FEP49, G2019
    FEP51, G2020
    FEP57, G2021
    FEP58, G2022
    # FA11, B21 INTRCPT3, G210
    # FA13, B22 INTRCPT3, G220
# T4 slope, P3 # INTRCPT2, B30 INTRCPT3, G300
    % FEA01, G301
    FEE02, G302
    FEE04, G303
    FEE07, G304
    FEE11, G305
    FED13, G306
    FED14, G307
    FED17, G308
    FED18, G309
    FED21, G3010
    FED22, G3011
    FED23, G3012
    FED25, G3013
    FEP28, G3014
    FEP29, G3015
    FEP31, G3016
    FEP32, G3017
    FEP34, G3018
    FEP35, G3019
    FEP36, G3020
    FEP37, G3021
    FEP38, G3022
    FEP39, G3023
    FEP40, G3024
    FEP41, G3025

```

FEP44, G3026
 FEP47, G3027
 FEP48, G3028
 FEP49, G3029
 FEP51, G3030
 FEP52, G3031
 FEP53, G3032
 FEP55, G3033
 FEP58, G3034
 #%FA06_NSE, B31 INTRCPT3, G310
 # FA07, B32 INTRCPT3, G320
 # FA09, B33 INTRCPT3, G330
 LEITI slope, P4 INTRCPT2, B40 INTRCPT3, G400
 % FEA01, G401
 FEE02, G402
 FEE04, G403
 FEE08, G404
 FED12, G405
 FED13, G406
 FED14, G407
 FED22, G408
 FED23, G409
 FED25, G4010
 FEP29, G4011
 FEP31, G4012
 FEP33, G4013
 FEP34, G4014
 FEP35, G4015
 FEP36, G4016
 FEP37, G4017
 FEP38, G4018
 FEP39, G4019
 FEP40, G4020
 FEP42, G4021
 FEP45, G4022
 FEP46, G4023
 FEP48, G4024
 FEP54, G4025
 FEP58, G4026
 #%FA06_NSE, B41 INTRCPT3, G410
 # FA08, B42 INTRCPT3, G420
 # FA12, B43 INTRCPT3, G430
 # FA13, B44 INTRCPT3, G440
 # FA14, B45 INTRCPT3, G450
 # LEITIXT2 slope, P5 # INTRCPT2, B50 INTRCPT3, G500
 % FEA01, G501
 FEE02, G502
 FEE03, G503
 FEE10, G504
 FED14, G505
 FED19, G506
 FED20, G507
 FED23, G508
 FED24, G509
 FED25, G5010
 FED26, G5011

FEP29, G5012
 FEP30, G5013
 FEP31, G5014
 FEP32, G5015
 FEP33, G5016
 FEP35, G5017
 FEP38, G5018
 FEP39, G5019
 FEP40, G5020
 FEP45, G5021
 FEP48, G5022
 FEP49, G5023
 FEP52, G5024
 FEP53, G5025
 FEP54, G5026
 FEP57, G5027
 FEP58, G5028
 # FA04_PRE, B51 INTRCPT3, G510
 # %FA06_NSE, B52 INTRCPT3, G520
 # FA08, B53 INTRCPT3, G530
 # FA09, B54 INTRCPT3, G540
 # FA10, B55 INTRCPT3, G550
 # LEITIXT3 slope, P6 # INTRCPT2, B60 INTRCPT3, G600
 % FEA01, G601
 FEE02, G602
 FEE10, G603
 FEE11, G604
 FED12, G605
 FED14, G606
 FED15, G607
 FED18, G608
 FED19, G609
 FED22, G6010
 FED25, G6011
 FEP31, G6012
 FEP32, G6013
 FEP34, G6014
 FEP39, G6015
 FEP40, G6016
 FEP43, G6017
 FEP44, G6018
 FEP49, G6019
 FEP57, G6020
 FEP58, G6021
 # %FA06_NSE, B61 INTRCPT3, G610
 # FA08, B62 INTRCPT3, G620
 # LEITIXT4 slope, P7 # INTRCPT2, B70 INTRCPT3, G700
 % FEA01, G701
 FEE04, G702
 FEE06, G703
 FEE07, G704
 FEE09, G705
 FEE11, G706
 FED12, G707
 FED13, G708
 FED14, G709

FED17, G7010
 FED19, G7011
 FED22, G7012
 FED25, G7013
 FEP28, G7014
 FEP29, G7015
 FEP31, G7016
 FEP32, G7017
 FEP33, G7018
 FEP34, G7019
 FEP36, G7020
 FEP37, G7021
 FEP38, G7022
 FEP39, G7023
 FEP41, G7024
 FEP44, G7025
 FEP45, G7026
 FEP46, G7027
 FEP47, G7028
 FEP48, G7029
 FEP49, G7030
 FEP50, G7031
 FEP51, G7032
 FEP53, G7033
 FEP55, G7034
 FEP58, G7035

##FA06_NSE, B71 INTRCPT3, G710

FA07, B72 INTRCPT3, G720

FA08, B73 INTRCPT3, G730

FA09, B74 INTRCPT3, G740

FA14, B75 INTRCPT3, G750

RET_AC slope, P8 # INTRCPT2, B80 INTRCPT3, G800

FA01_SEX, B81 INTRCPT3, G810

FA11, B82 INTRCPT3, G820

'#' - The residual parameter variance for the parameter has been set to zero

'%' - This variable has been centered around its grand mean

Summary of the model specified (in equation format)

 Level-1 Model

$$Y = P0 + P1*(T2) + P2*(T3) + P3*(T4) + P4*(LEITI) + P5*(LEITIXT2) + P6*(LEITIXT3) + P7*(LEITIXT4) + P8*(RET_AC) + E$$

Level-2 Model

$$P0 = B00 + B01*(FA04_PRE) + B02*(FA07) + B03*(FA09) + B04*(FA12) + B05*(FA14) + R0$$

$$P1 = B10 + B11*(FA04_PRE) + B12*(FA07) + B13*(FA09) + B14*(FA10)$$

$$P2 = B20 + B21*(FA11) + B22*(FA13)$$

$$P3 = B30 + B31*(FA06_NSE) + B32*(FA07) + B33*(FA09)$$

$$P4 = B40 + B41*(FA06_NSE) + B42*(FA08) + B43*(FA12) + B44*(FA13) + B45*(FA14) + R4$$

$$\begin{aligned}
P5 &= B50 + B51*(FA04_PRE) + B52*(FA06_NSE) + B53*(FA08) + B54*(FA09) \\
&\quad + B55*(FA10) \\
P6 &= B60 + B61*(FA06_NSE) + B62*(FA08) \\
P7 &= B70 + B71*(FA06_NSE) + B72*(FA07) + B73*(FA08) + B74*(FA09) \\
&\quad + B75*(FA14) \\
P8 &= B80 + B81*(FA01_SEX) + B82*(FA11)
\end{aligned}$$

Level-3 Model

$$\begin{aligned}
B00 &= G000 + G001(FA01) + G002(FEE02) + G003(FEE04) + G004(FEE05) \\
&\quad + G005(FEE08) + G006(FED12) + G007(FED13) + G008(FED14) \\
&\quad + G009(FED22) + G0010(FED23) + G0011(FED25) + G0012(FEP29) \\
&\quad + G0013(FEP31) + G0014(FEP32) + G0015(FEP33) + G0016(FEP34) + G0017(FEP35) + \\
&G0018(FEP36) + G0019(FEP37) + G0020(FEP38) + G0021(FEP39) + G0022(FEP40) + G0023(FEP41) + \\
&G0024(FEP42) + G0025(FEP43) + G0026(FEP46) + G0027(FEP47) + G0028(FEP48) + G0029(FEP51) + \\
&G0030(FEP54) + G0031(FEP58) + U00 \\
B01 &= G010 \\
B02 &= G020 \\
B03 &= G030 \\
B04 &= G040 \\
B05 &= G050 \\
B10 &= G100 + G101(FA01) + G102(FEE02) + G103(FEE03) + G104(FEE06) \\
&\quad + G105(FEE07) + G106(FEE10) + G107(FED14) + G108(FED19) \\
&\quad + G109(FED20) + G1010(FED21) + G1011(FED24) + G1012(FED25) \\
&\quad + G1013(FED26) + G1014(FEP29) + G1015(FEP31) + G1016(FEP32) + G1017(FEP33) + \\
&G1018(FEP34) + G1019(FEP35) + G1020(FEP38) + G1021(FEP39) + G1022(FEP40) + G1023(FEP45) + \\
&G1024(FEP48) + G1025(FEP50) + G1026(FEP52) + G1027(FEP53) + G1028(FEP54) + G1029(FEP55) + \\
&G1030(FEP58) \\
B11 &= G110 \\
B12 &= G120 \\
B13 &= G130 \\
B14 &= G140 \\
B20 &= G200 + G201(FA01) + G202(FEE02) + G203(FEE03) + G204(FED12) \\
&\quad + G205(FED14) + G206(FED15) + G207(FED19) + G208(FED22) \\
&\quad + G209(FED24) + G2010(FED26) + G2011(FEP31) + G2012(FEP32) \\
&\quad + G2013(FEP33) + G2014(FEP34) + G2015(FEP39) + G2016(FEP42) + G2017(FEP43) + \\
&G2018(FEP44) + G2019(FEP49) + G2020(FEP51) + G2021(FEP57) + G2022(FEP58) \\
B21 &= G210 \\
B22 &= G220 \\
B30 &= G300 + G301(FA01) + G302(FEE02) + G303(FEE04) + G304(FEE07) \\
&\quad + G305(FEE11) + G306(FED13) + G307(FED14) + G308(FED17) \\
&\quad + G309(FED18) + G3010(FED21) + G3011(FED22) + G3012(FED23) \\
&\quad + G3013(FED25) + G3014(FEP28) + G3015(FEP29) + G3016(FEP31) + G3017(FEP32) + \\
&G3018(FEP34) + G3019(FEP35) + G3020(FEP36) + G3021(FEP37) + G3022(FEP38) + G3023(FEP39) + \\
&G3024(FEP40) + G3025(FEP41) + G3026(FEP44) + G3027(FEP47) + G3028(FEP48) + G3029(FEP49) + \\
&G3030(FEP51) + G3031(FEP52) + G3032(FEP53) + G3033(FEP55) + G3034(FEP58) \\
B31 &= G310 \\
B32 &= G320 \\
B33 &= G330 \\
B40 &= G400 + G401(FA01) + G402(FEE02) + G403(FEE04) + G404(FEE08) \\
&\quad + G405(FED12) + G406(FED13) + G407(FED14) + G408(FED22) \\
&\quad + G409(FED23) + G4010(FED25) + G4011(FEP29) + G4012(FEP31) \\
&\quad + G4013(FEP33) + G4014(FEP34) + G4015(FEP35) + G4016(FEP36) + G4017(FEP37) + \\
&G4018(FEP38) + G4019(FEP39) + G4020(FEP40) + G4021(FEP42) + G4022(FEP45) + G4023(FEP46) + \\
&G4024(FEP48) + G4025(FEP54) + G4026(FEP58) + U40
\end{aligned}$$

B41 = G410
 B42 = G420
 B43 = G430
 B44 = G440
 B45 = G450
 B50 = G500 + G501(FEA01) + G502(FEE02) + G503(FEE03) + G504(FEE10)
 + G505(FED14) + G506(FED19) + G507(FED20) + G508(FED23)
 + G509(FED24) + G5010(FED25) + G5011(FED26) + G5012(FEP29)
 + G5013(FEP30) + G5014(FEP31) + G5015(FEP32) + G5016(FEP33) + G5017(FEP35) +
 G5018(FEP38) + G5019(FEP39) + G5020(FEP40) + G5021(FEP45) + G5022(FEP48) + G5023(FEP49) +
 G5024(FEP52) + G5025(FEP53) + G5026(FEP54) + G5027(FEP57) + G5028(FEP58)
 B51 = G510
 B52 = G520
 B53 = G530
 B54 = G540
 B55 = G550
 B60 = G600 + G601(FEA01) + G602(FEE02) + G603(FEE10) + G604(FEE11)
 + G605(FED12) + G606(FED14) + G607(FED15) + G608(FED18)
 + G609(FED19) + G6010(FED22) + G6011(FED25) + G6012(FEP31)
 + G6013(FEP32) + G6014(FEP34) + G6015(FEP39) + G6016(FEP40) + G6017(FEP43) +
 G6018(FEP44) + G6019(FEP49) + G6020(FEP57) + G6021(FEP58)
 B61 = G610
 B62 = G620
 B70 = G700 + G701(FEA01) + G702(FEE04) + G703(FEE06) + G704(FEE07)
 + G705(FEE09) + G706(FEE11) + G707(FED12) + G708(FED13)
 + G709(FED14) + G7010(FED17) + G7011(FED19) + G7012(FED22)
 + G7013(FED25) + G7014(FEP28) + G7015(FEP29) + G7016(FEP31) + G7017(FEP32) +
 G7018(FEP33) + G7019(FEP34) + G7020(FEP36) + G7021(FEP37) + G7022(FEP38) + G7023(FEP39) +
 G7024(FEP41) + G7025(FEP44) + G7026(FEP45) + G7027(FEP46) + G7028(FEP47) + G7029(FEP48) +
 G7030(FEP49) + G7031(FEP50) + G7032(FEP51) + G7033(FEP53) + G7034(FEP55) + G7035(FEP58)
 B71 = G710
 B72 = G720
 B73 = G730
 B74 = G740
 B75 = G750
 B80 = G800
 B81 = G810
 B82 = G820

For starting values, data from 30012 level-1 and 8655 level-2 records were used

Iterations stopped due to small change in likelihood function

***** ITERATION 174 *****

Standard errors for sigma_squared, Tau(pi), Tau(beta) are not computable.

Sigma_squared = 167.62694

Tau(pi)

INTRCPT1,P0 60.37075 -0.29426
 LEITI,P1 -0.29426 0.00150

Tau(pi) (as correlations)

INTRCPT1,P0 1.000 -0.979
 LEITI,P4 -0.979 1.000

Random level-1 coefficient Reliability estimate

INTRCPT1, P0	0.030
LEITI, P4	0.014

Note: The reliability estimates reported above are based on only 8655 of 12180 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Tau(beta)

INTRCPT1	LEITI
INTRCPT2,B00	INTRCPT2,B40
10.59862	-0.06205
-0.06205	0.00044

Tau(beta) (as correlations)

INTRCPT1/INTRCPT2,B00	1.000	-0.910
LEITI/INTRCPT2,B40	-0.910	1.000

Random level-2 coefficient Reliability estimate

INTRCPT1/INTRCPT2, B00	0.298
LEITI/INTRCPT2, B40	0.269

The value of the likelihood function at iteration 174 = -1.389420E+005

The robust standard errors cannot be computed for this model.

The outcome variable is LEITF

Final estimation of fixed effects:

Fixed Effect	Standard Coefficient	Standard Error	Approx. T-ratio	d.f.	P-value
For INTRCPT1, P0					
For INTRCPT2, B00					
INTRCPT3, G000	18.868885	17.560205	1.075	197	0.284
FEA01, G001	25.076443	3.957045	6.337	197	0.000
FEE02, G002	8.924717	1.669669	5.345	197	0.000
FEE04, G003	-3.292213	0.643699	-5.115	197	0.000
FEE05, G004	-0.535085	0.160057	-3.343	197	0.001
FEE08, G005	4.078116	0.646538	6.308	197	0.000
FED12, G006	-6.590487	2.794509	-2.358	197	0.019
FED13, G007	5.937845	1.704131	3.484	197	0.001
FED14, G008	10.175413	2.509554	4.055	197	0.000
FED22, G009	11.279911	2.473743	4.560	197	0.000
FED23, G0010	-5.239586	1.083577	-4.835	197	0.000
FED25, G0011	4.170230	1.536516	2.714	197	0.008
FEP29, G0012	-3.298442	1.560354	-2.114	197	0.036
FEP31, G0013	6.970771	1.689877	4.125	197	0.000
FEP32, G0014	-4.225370	0.998501	-4.232	197	0.000
FEP33, G0015	-7.608408	2.106763	-3.611	197	0.001
FEP34, G0016	7.880698	1.760707	4.476	197	0.000

FEP35, G0017	-5.875418	1.394812	-4.212	197	0.000
FEP36, G0018	4.319456	1.915607	2.255	197	0.025
FEP37, G0019	-17.517024	3.227746	-5.427	197	0.000
FEP38, G0020	13.447055	3.524707	3.815	197	0.000
FEP39, G0021	-10.124262	2.077662	-4.873	197	0.000
FEP40, G0022	-5.454846	1.294856	-4.213	197	0.000
FEP41, G0023	-1.159059	0.526919	-2.200	197	0.029
FEP42, G0024	5.809007	1.730559	3.357	197	0.001
FEP43, G0025	2.402434	0.584800	4.108	197	0.000
FEP46, G0026	-8.959368	1.533335	-5.843	197	0.000
FEP47, G0027	1.206162	0.322695	3.738	197	0.000
FEP48, G0028	-3.535668	1.033404	-3.421	197	0.001
FEP51, G0029	-1.643687	0.483731	-3.398	197	0.001
FEP54, G0030	6.238716	1.708372	3.652	197	0.001
FEP58, G0031	9.470685	1.979672	4.784	197	0.000
For FA04_PRE, B01					
INTRCPT3, G010	1.997967	0.264161	7.563	12174	0.000
For FA07, B02					
INTRCPT3, G020	4.308195	0.212259	20.297	12174	0.000
For FA09, B03					
INTRCPT3, G030	-2.273596	0.352730	-6.446	12174	0.000
For FA12, B04					
INTRCPT3, G040	5.884671	0.933988	6.301	12174	0.000
For FA14, B05					
INTRCPT3, G050	-4.425480	0.828124	-5.344	12174	0.000
For T2 slope, P1					
For INTRCPT2, B10					
INTRCPT3, G100	33.940917	26.097339	1.301	34388	0.194
FEA01, G101	-13.744214	6.334205	-2.170	34388	0.030
FEE02, G102	-20.495366	3.449747	-5.941	34388	0.000
FEE03, G103	4.160145	1.700973	2.446	34388	0.015
FEE06, G104	-0.923775	0.247256	-3.736	34388	0.000
FEE07, G105	1.973766	0.466147	4.234	34388	0.000
FEE10, G106	-3.421246	1.840593	-1.859	34388	0.063
FED14, G107	-18.575601	3.562975	-5.214	34388	0.000
FED19, G108	-3.542728	1.449887	-2.443	34388	0.015
FED20, G109	-7.176973	3.261343	-2.201	34388	0.028
FED21, G1010	-2.507220	0.705855	-3.552	34388	0.001
FED24, G1011	3.664373	1.250683	2.930	34388	0.004
FED25, G1012	-6.635119	2.531944	-2.621	34388	0.009
FED26, G1013	-2.544130	0.935995	-2.718	34388	0.007
FEP29, G1014	7.225762	3.012456	2.399	34388	0.017
FEP31, G1015	-13.346000	2.576397	-5.180	34388	0.000
FEP32, G1016	17.815751	5.649127	3.154	34388	0.002
FEP33, G1017	16.662062	4.253342	3.917	34388	0.000
FEP34, G1018	-1.675732	0.617941	-2.712	34388	0.007
FEP35, G1019	7.795435	2.935566	2.656	34388	0.008
FEP38, G1020	-14.609046	5.148414	-2.838	34388	0.005
FEP39, G1021	8.205856	3.067097	2.675	34388	0.008
FEP40, G1022	6.703959	2.449871	2.736	34388	0.007
FEP45, G1023	18.581963	3.401494	5.463	34388	0.000
FEP48, G1024	4.820414	1.836567	2.625	34388	0.009
FEP50, G1025	-1.059772	0.447949	-2.366	34388	0.018
FEP52, G1026	10.243337	4.453825	2.300	34388	0.021
FEP53, G1027	-10.550274	3.065157	-3.442	34388	0.001
FEP54, G1028	8.025950	3.828342	2.096	34388	0.036

FEP55, G1029	9.881197	2.593648	3.810	34388	0.000
FEP58, G1030	-17.553171	2.867018	-6.122	34388	0.000
For FA04_PRE, B11					
INTRCPT3, G110	-19.036251	3.067590	-6.206	34388	0.000
For FA07, B12					
INTRCPT3, G120	2.339973	0.357359	6.548	34388	0.000
For FA09, B13					
INTRCPT3, G130	-5.431695	2.867215	-1.894	34388	0.058
For FA10, B14					
INTRCPT3, G140	10.320955	2.291537	4.504	34388	0.000
For T3 slope, P2					
For INTRCPT2, B20					
INTRCPT3, G200	75.246327	19.794583	3.801	34388	0.000
FEA01, G201	-14.570065	5.062544	-2.878	34388	0.004
FEE02, G202	-5.622300	2.834692	-1.983	34388	0.047
FEE03, G203	-0.533949	0.248010	-2.153	34388	0.031
FED12, G204	16.814603	4.410820	3.812	34388	0.000
FED14, G205	-8.790715	3.277509	-2.682	34388	0.008
FED15, G206	-5.242108	1.735376	-3.021	34388	0.003
FED19, G207	6.078526	1.219467	4.985	34388	0.000
FED22, G208	-7.231337	3.286711	-2.200	34388	0.028
FED24, G209	0.420298	0.191043	2.200	34388	0.028
FED26, G2010	-0.329394	0.154138	-2.137	34388	0.032
FEP31, G2011	-9.291439	2.253419	-4.123	34388	0.000
FEP32, G2012	-23.167641	4.389968	-5.277	34388	0.000
FEP33, G2013	-3.199399	0.892836	-3.583	34388	0.001
FEP34, G2014	-15.459180	2.503875	-6.174	34388	0.000
FEP39, G2015	7.892376	2.595686	3.041	34388	0.003
FEP42, G2016	-1.097295	0.588369	-1.865	34388	0.062
FEP43, G2017	-7.711106	3.554562	-2.169	34388	0.030
FEP44, G2018	10.485566	2.687628	3.901	34388	0.000
FEP49, G2019	7.041952	1.395163	5.047	34388	0.000
FEP51, G2020	1.904206	0.577335	3.298	34388	0.001
FEP57, G2021	1.697509	0.623429	2.723	34388	0.007
FEP58, G2022	-10.955519	2.679548	-4.089	34388	0.000
For FA11, B21					
INTRCPT3, G210	1.697226	0.387774	4.377	34388	0.000
For FA13, B22					
INTRCPT3, G220	-2.853929	0.602494	-4.737	34388	0.000
For T4 slope, P3					
For INTRCPT2, B30					
INTRCPT3, G300	35.022748	27.491375	1.274	34388	0.203
FEA01, G301	-17.646327	6.134367	-2.877	34388	0.004
FEE02, G302	4.015503	0.845252	4.751	34388	0.000
FEE04, G303	3.181989	1.020779	3.117	34388	0.002
FEE07, G304	14.642165	2.224927	6.581	34388	0.000
FEE11, G305	5.986768	1.810781	3.306	34388	0.001
FED13, G306	-5.032136	2.771564	-1.816	34388	0.069
FED14, G307	-13.606613	3.198043	-4.255	34388	0.000
FED17, G308	13.783495	3.115007	4.425	34388	0.000
FED18, G309	-0.468311	0.209881	-2.231	34388	0.026
FED21, G3010	-1.547025	0.652928	-2.369	34388	0.018
FED22, G3011	-8.677447	3.350223	-2.590	34388	0.010
FED23, G3012	-1.661227	0.403918	-4.113	34388	0.000
FED25, G3013	-7.192533	2.257315	-3.186	34388	0.002
FEP28, G3014	9.049056	3.552034	2.548	34388	0.011

FEP29, G3015	17.019455	3.175844	5.359	34388	0.000
FEP31, G3016	-6.719512	2.288738	-2.936	34388	0.004
FEP32, G3017	-12.078703	4.398708	-2.746	34388	0.006
FEP34, G3018	-7.598346	2.494348	-3.046	34388	0.003
FEP35, G3019	-1.391194	0.562026	-2.475	34388	0.014
FEP36, G3020	-19.253890	3.174933	-6.064	34388	0.000
FEP37, G3021	22.951629	5.404689	4.247	34388	0.000
FEP38, G3022	-19.423649	5.460492	-3.557	34388	0.001
FEP39, G3023	6.278891	2.703238	2.323	34388	0.020
FEP40, G3024	-1.688584	0.636270	-2.654	34388	0.008
FEP41, G3025	8.498390	3.067128	2.771	34388	0.006
FEP44, G3026	8.456585	2.528656	3.344	34388	0.001
FEP47, G3027	-7.562158	2.044028	-3.700	34388	0.000
FEP48, G3028	8.122642	2.372429	3.424	34388	0.001
FEP49, G3029	3.814763	1.555781	2.452	34388	0.014
FEP51, G3030	-11.446946	2.515828	-4.550	34388	0.000
FEP52, G3031	1.301475	0.616232	2.112	34388	0.034
FEP53, G3032	-6.191346	2.186384	-2.832	34388	0.005
FEP55, G3033	47.804948	11.799338	4.051	34388	0.000
FEP58, G3034	-15.681019	2.568885	-6.104	34388	0.000
For FA06_NSE, B31					
INTRCPT3, G310	3.771545	1.828103	2.063	34388	0.039
For FA07, B32					
INTRCPT3, G320	8.615349	1.513257	5.693	34388	0.000
For FA09, B33					
INTRCPT3, G330	-14.396644	1.980629	-7.269	34388	0.000
For LEITI slope, P4					
For INTRCPT2, B40					
INTRCPT3, G400	0.783081	0.140116	5.589	202	0.000
FEA01, G401	-0.147625	0.030179	-4.892	202	0.000
FEE02, G402	-0.083167	0.013757	-6.046	202	0.000
FEE04, G403	0.017386	0.004702	3.697	202	0.000
FEE08, G404	-0.021317	0.004474	-4.765	202	0.000
FED12, G405	0.049547	0.022122	2.240	202	0.026
FED13, G406	-0.042188	0.012431	-3.394	202	0.001
FED14, G407	-0.090978	0.020169	-4.511	202	0.000
FED22, G408	-0.076280	0.019177	-3.978	202	0.000
FED23, G409	0.038585	0.007501	5.144	202	0.000
FED25, G4010	-0.035542	0.012463	-2.852	202	0.005
FEP29, G4011	0.033669	0.011287	2.983	202	0.004
FEP31, G4012	-0.057376	0.013926	-4.120	202	0.000
FEP33, G4013	0.066117	0.016663	3.968	202	0.000
FEP34, G4014	-0.061852	0.014294	-4.327	202	0.000
FEP35, G4015	0.040912	0.009905	4.130	202	0.000
FEP36, G4016	-0.034124	0.013846	-2.465	202	0.015
FEP37, G4017	0.112721	0.023561	4.784	202	0.000
FEP38, G4018	-0.099808	0.025237	-3.955	202	0.000
FEP39, G4019	0.078255	0.016438	4.760	202	0.000
FEP40, G4020	0.036342	0.009942	3.655	202	0.001
FEP42, G4021	-0.028846	0.011552	-2.497	202	0.014
FEP45, G4022	-0.008339	0.004152	-2.009	202	0.046
FEP46, G4023	0.057264	0.010975	5.218	202	0.000
FEP48, G4024	0.018201	0.007403	2.459	202	0.015
FEP54, G4025	-0.033350	0.011616	-2.871	202	0.005
FEP58, G4026	-0.074770	0.015916	-4.698	202	0.000
For FA06_NSE, B41					

INTRCPT3, G410	0.016717	0.003315	5.043	12174	0.000
For FA08, B42					
INTRCPT3, G420	0.022636	0.003487	6.491	12174	0.000
For FA12, B43					
INTRCPT3, G430	-0.021917	0.006867	-3.192	12174	0.002
For FA13, B44					
INTRCPT3, G440	-0.008374	0.002341	-3.576	12174	0.001
For FA14, B45					
INTRCPT3, G450	0.030239	0.006181	4.892	12174	0.000
For LEITIXT2 slope, P5					
For INTRCPT2, B50					
INTRCPT3, G500	-0.200783	0.203084	-0.989	34388	0.323
FEA01, G501	0.074551	0.046198	1.614	34388	0.106
FEE02, G502	0.175446	0.028003	6.265	34388	0.000
FEE03, G503	-0.028385	0.012062	-2.353	34388	0.019
FEE10, G504	0.024706	0.014352	1.721	34388	0.085
FED14, G505	0.147808	0.027290	5.416	34388	0.000
FED19, G506	0.032511	0.011047	2.943	34388	0.004
FED20, G507	0.052597	0.024717	2.128	34388	0.033
FED23, G508	-0.007409	0.002933	-2.526	34388	0.012
FED24, G509	-0.020800	0.009191	-2.263	34388	0.024
FED25, G5010	0.048872	0.019738	2.476	34388	0.014
FED26, G5011	0.018394	0.007074	2.600	34388	0.010
FEP29, G5012	-0.042976	0.022604	-1.901	34388	0.057
FEP30, G5013	-0.022043	0.006006	-3.670	34388	0.000
FEP31, G5014	0.108731	0.020169	5.391	34388	0.000
FEP32, G5015	-0.111402	0.043381	-2.568	34388	0.011
FEP33, G5016	-0.143132	0.032792	-4.365	34388	0.000
FEP35, G5017	-0.057847	0.021688	-2.667	34388	0.008
FEP38, G5018	0.106324	0.037937	2.803	34388	0.006
FEP39, G5019	-0.066865	0.023472	-2.849	34388	0.005
FEP40, G5020	-0.039060	0.018412	-2.121	34388	0.034
FEP45, G5021	-0.126829	0.026846	-4.724	34388	0.000
FEP48, G5022	-0.031302	0.013589	-2.303	34388	0.021
FEP49, G5023	-0.005118	0.002575	-1.988	34388	0.046
FEP52, G5024	-0.067987	0.033952	-2.002	34388	0.045
FEP53, G5025	0.063838	0.023372	2.731	34388	0.007
FEP54, G5026	-0.061356	0.028870	-2.125	34388	0.033
FEP57, G5027	0.003701	0.001067	3.470	34388	0.001
FEP58, G5028	0.131385	0.022125	5.938	34388	0.000
For FA04_PRE, B51					
INTRCPT3, G510	0.156949	0.025151	6.240	34388	0.000
For FA06_NSE, B52					
INTRCPT3, G520	-0.003672	0.004330	-0.848	34388	0.397
For FA08, B53					
INTRCPT3, G530	-0.014901	0.004537	-3.284	34388	0.001
For FA09, B54					
INTRCPT3, G540	0.032720	0.022733	1.439	34388	0.150
For FA10, B55					
INTRCPT3, G550	-0.076467	0.017584	-4.349	34388	0.000
For LEITIXT3 slope, P6					
For INTRCPT2, B60					
INTRCPT3, G600	-0.377355	0.147864	-2.552	34388	0.011
FEA01, G601	0.107839	0.035954	2.999	34388	0.003
FEE02, G602	0.058264	0.021359	2.728	34388	0.007
FEE10, G603	-0.005955	0.001976	-3.014	34388	0.003

FEE11, G604	-0.009541	0.002222	-4.293	34388	0.000
FED12, G605	-0.126939	0.032538	-3.901	34388	0.000
FED14, G606	0.075658	0.024215	3.124	34388	0.002
FED15, G607	0.036310	0.011856	3.062	34388	0.003
FED18, G608	-0.003822	0.001550	-2.466	34388	0.014
FED19, G609	-0.038770	0.008457	-4.584	34388	0.000
FED22, G6010	0.049948	0.024392	2.048	34388	0.040
FED25, G6011	0.008925	0.004275	2.088	34388	0.037
FEP31, G6012	0.067907	0.016960	4.004	34388	0.000
FEP32, G6013	0.191914	0.030910	6.209	34388	0.000
FEP34, G6014	0.094523	0.018469	5.118	34388	0.000
FEP39, G6015	-0.046740	0.019241	-2.429	34388	0.015
FEP40, G6016	-0.010148	0.003850	-2.636	34388	0.009
FEP43, G6017	0.045778	0.024344	1.880	34388	0.060
FEP44, G6018	-0.067800	0.018126	-3.740	34388	0.000
FEP49, G6019	-0.048198	0.009748	-4.944	34388	0.000
FEP57, G6020	-0.010458	0.004271	-2.449	34388	0.015
FEP58, G6021	0.077477	0.019870	3.899	34388	0.000
For FA06_NSE, B61					
INTRCPT3, G610	-0.008709	0.004007	-2.174	34388	0.030
For FA08, B62					
INTRCPT3, G620	-0.015761	0.004259	-3.701	34388	0.000
For LEITIXT4 slope, P7					
For INTRCPT2, B70					
INTRCPT3, G700	-0.161195	0.190986	-0.844	34388	0.399
FEA01, G701	0.142884	0.041152	3.472	34388	0.001
FEE04, G702	-0.018956	0.006774	-2.798	34388	0.006
FEE06, G703	-0.004427	0.001405	-3.150	34388	0.002
FEE07, G704	-0.089644	0.013741	-6.524	34388	0.000
FEE09, G705	0.004521	0.002122	2.131	34388	0.033
FEE11, G706	-0.039096	0.010939	-3.574	34388	0.001
FED12, G707	-0.014439	0.007746	-1.864	34388	0.062
FED13, G708	0.045146	0.018364	2.458	34388	0.014
FED14, G709	0.105647	0.023159	4.562	34388	0.000
FED17, G7010	-0.081132	0.019501	-4.160	34388	0.000
FED19, G7011	0.004537	0.001563	2.903	34388	0.004
FED22, G7012	0.066156	0.023494	2.816	34388	0.005
FED25, G7013	0.051409	0.015967	3.220	34388	0.002
FEP28, G7014	-0.049330	0.022073	-2.235	34388	0.025
FEP29, G7015	-0.108015	0.020414	-5.291	34388	0.000
FEP31, G7016	0.055071	0.016675	3.303	34388	0.001
FEP32, G7017	0.097640	0.028244	3.457	34388	0.001
FEP33, G7018	-0.031258	0.006737	-4.640	34388	0.000
FEP34, G7019	0.068474	0.017907	3.824	34388	0.000
FEP36, G7020	0.128296	0.020873	6.147	34388	0.000
FEP37, G7021	-0.148794	0.035834	-4.152	34388	0.000
FEP38, G7022	0.119779	0.035922	3.334	34388	0.001
FEP39, G7023	-0.050523	0.019480	-2.594	34388	0.010
FEP41, G7024	-0.044803	0.019414	-2.308	34388	0.021
FEP44, G7025	-0.050072	0.015753	-3.179	34388	0.002
FEP45, G7026	0.008857	0.004335	2.043	34388	0.041
FEP46, G7027	-0.014958	0.003692	-4.051	34388	0.000
FEP47, G7028	0.046310	0.012567	3.685	34388	0.000
FEP48, G7029	-0.054999	0.015111	-3.640	34388	0.001
FEP49, G7030	-0.026932	0.009882	-2.725	34388	0.007
FEP50, G7031	-0.014150	0.002512	-5.633	34388	0.000

FEP51, G7032	0.074530	0.016148	4.616	34388	0.000
FEP53, G7033	0.036284	0.013946	2.602	34388	0.010
FEP55, G7034	-0.252117	0.075197	-3.353	34388	0.001
FEP58, G7035	0.112512	0.018664	6.028	34388	0.000
For FA06_NSE, B71					
INTRCPT3, G710	-0.029563	0.012016	-2.460	34388	0.014
For FA07, B72					
INTRCPT3, G720	-0.049761	0.009368	-5.312	34388	0.000
For FA08, B73					
INTRCPT3, G730	-0.012032	0.003921	-3.068	34388	0.003
For FA09, B74					
INTRCPT3, G740	0.082939	0.013349	6.213	34388	0.000
For FA14, B75					
INTRCPT3, G750	-0.006946	0.002301	-3.018	34388	0.003
For RET_AC slope, P8					
For INTRCPT2, B80					
INTRCPT3, G800	8.489026	0.492891	17.223	34388	0.000
For FA01_SEX, B81					
INTRCPT3, G810	1.133238	0.163252	6.942	34388	0.000
For FA11, B82					
INTRCPT3, G820	-2.174559	0.213145	-10.202	34388	0.000

Given the small number of level 3 units in your model, the standard errors above may underestimate the uncertainty about the fixed effects.

Final estimation of level-1 and level-2 variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, R0	7.76986	60.37075	8421	8458.77617	0.384
LEITI slope, R4	0.03869	0.00150	8421	8487.24937	0.303
level-1, E	12.94708	167.62694			

Note: The chi-square statistics reported above are based on only 8655 of 12180 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Final estimation of level-3 variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1/INTRCPT2, U00	3.25555	10.59862	197	437.74373	0.000
LEITI/INTRCPT2, U40	0.02094	0.00044	202	426.35693	0.000

Statistics for current covariance components model

Deviance = 277884.002617
Number of estimated parameters = 276

ANEXO 07 - Relatório HLM – Matemática – Modelo Nulo

Program: HLM 6 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: HLM3.EXE (6.08.29257.1)
Date: 2 June 2012, Saturday
Time: 11:45: 4

SPECIFICATIONS FOR THIS HLM3 RUN

Problem Title: no title
The data source for this run = Modelo_matemática.mdm
The command file for this run = C:\Users\Adilson Dalben\Documents\Meus Documentos Adilson\Doutorado\00 - Tese_Banco-Geres\00-Interpretação_HLM_Matematica\Modelo N1 - 00 - NULO.hlm
Output file name = C:\Users\Adilson Dalben\Documents\Meus Documentos Adilson\Doutorado\00 - Tese_Banco-Geres\00-Interpretação_HLM_Matematica\hlm3.txt
The maximum number of level-1 units = 34561
The maximum number of level-2 units = 12153
The maximum number of level-3 units = 230
The maximum number of iterations = 100
Method of estimation: full maximum likelihood

Weighting Specification

Weight
Variable
Weighting? Name Normalized?
Level 1 no
Level 2 yes PESO_ALU yes
Level 3 no
Precision no

The outcome variable is MAT_F
The model specified for the fixed effects was:

Level-1 Level-2 Level-3
Coefficients Predictors Predictors

INTRCPT1, P0 INTRCPT2, B00 INTRCPT3, G000

Summary of the model specified (in equation format)

Level-1 Model
 $Y = P0 + E$

Level-2 Model
 $P0 = B00 + R0$

Level-3 Model
 $B00 = G000 + U00$

For starting values, data from 34561 level-1 and 12153 level-2 records were used
Iterations stopped due to small change in likelihood function

***** ITERATION 26 *****

Standard errors for sigma_squared, Tau(pi), Tau(beta) are not computable.

Sigma_squared = 2975.65079

Tau(pi)

INTRCPT1,P0 719.83685

Tau(pi) (as correlations)

INTRCPT1,P0 1.000

 Random level-1 coefficient Reliability estimate

INTRCPT1, P0 0.393

Tau(beta)

INTRCPT1

INTRCPT2,B00

1070.17241

Tau(beta) (as correlations)

INTRCPT1/INTRCPT2,B00 1.000

 Random level-2 coefficient Reliability estimate

INTRCPT1/INTRCPT2, B00 0.962

The value of the likelihood function at iteration 26 = -1.908140E+005

The outcome variable is MAT_F

Final estimation of fixed effects

(with robust standard errors)

Fixed Effect	Standard Coefficient	Standard Error	Approx. T-ratio	d.f.	P-value
For INTRCPT1, P0					
For INTRCPT2, B00					
INTRCPT3, G000	200.723047	4.149735	48.370	229	0.000

Final estimation of level-1 and level-2 variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, level-1, R0	26.82978	719.83685	11923	20451.04703	0.000
E	54.54953	2975.65079			

Final estimation of level-3 variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1/INTRCPT2, U00	32.71349	1070.17241	229	5706.51630	0.000

Statistics for current covariance components model

Deviance = 381628.059542
Number of estimated parameters = 4

ANEXO 08 - Relatório HLM – Matemática – Modelo Referência

Program: HLM 6 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: HLM3.EXE (6.08.29257.1)
Date: 2 June 2012, Saturday
Time: 11:46: 3

SPECIFICATIONS FOR THIS HLM3 RUN

Problem Title: no title

The data source for this run = Modelo_matemática.mdm

The command file for this run = C:\Users\Adilson Dalben\Documents\Meus Documentos Adilson\Doutorado\00 - Tese_Banco-Geres\00-Interpretação_HLM_ Matematica\Modelo N1 - 01 - REFERENCIA.hlm

Output file name = C:\Users\Adilson Dalben\Documents\Meus Documentos Adilson\Doutorado\00 - Tese_Banco-Geres\00-Interpretação_HLM_ Matematica\hlm3.txt

The maximum number of level-1 units = 34561

The maximum number of level-2 units = 12153

The maximum number of level-3 units = 230

The maximum number of iterations = 100

Method of estimation: full maximum likelihood

Weighting Specification

Weight
Variable
Weighting? Name Normalized?
Level 1 no
Level 2 yes PESO_ALU yes
Level 3 no
Precision no

The outcome variable is MAT_F

The model specified for the fixed effects was:

Level-1 Level-2 Level-3
Coefficients Predictors Predictors

INTRCPT1, P0 INTRCPT2, B00 INTRCPT3, G000
T2 slope, P1 # INTRCPT2, B10 INTRCPT3, G100
T3 slope, P2 # INTRCPT2, B20 INTRCPT3, G200
T4 slope, P3 # INTRCPT2, B30 INTRCPT3, G300
MAT_I slope, P4 INTRCPT2, B40 INTRCPT3, G400
T2XMAT_I slope, P5 # INTRCPT2, B50 INTRCPT3, G500
T3XMAT_I slope, P6 # INTRCPT2, B60 INTRCPT3, G600
T4XMAT_I slope, P7 # INTRCPT2, B70 INTRCPT3, G700
DEFAS slope, P8 # INTRCPT2, B80 INTRCPT3, G800

'#' - The residual parameter variance for the parameter has been set to zero

Summary of the model specified (in equation format)

Level-1 Model

$$Y = P0 + P1*(T2) + P2*(T3) + P3*(T4) + P4*(MAT_I) + P5*(T2XMAT_I) + P6*(T3XMAT_I) + P7*(T4XMAT_I) + P8*(DEFAS) + E$$

Level-2 Model

$$\begin{aligned} P0 &= B00 + R0 \\ P1 &= B10 \\ P2 &= B20 \\ P3 &= B30 \\ P4 &= B40 + R4 \\ P5 &= B50 \\ P6 &= B60 \\ P7 &= B70 \\ P8 &= B80 \end{aligned}$$

Level-3 Model

$$\begin{aligned} B00 &= G000 + U00 \\ B10 &= G100 \\ B20 &= G200 \\ B30 &= G300 \\ B40 &= G400 + U40 \\ B50 &= G500 \\ B60 &= G600 \\ B70 &= G700 \\ B80 &= G800 \end{aligned}$$

For starting values, data from 31207 level-1 and 9207 level-2 records were used

Iterations stopped due to small change in likelihood function

***** ITERATION 2463 *****

Standard errors for sigma_squared, Tau(pi), Tau(beta) are not computable.

Sigma_squared = 971.35803

Tau(pi)

INTRCPT1,P0	6.26968	-0.03163
MAT_I,P1	-0.03163	0.00020

Tau(pi) (as correlations)

INTRCPT1,P0	1.000	-0.902
-------------	-------	--------

MAT_I,P4 -0.902 1.000

```

-----
Random level-1 coefficient  Reliability estimate
-----
INTRCPT1, P0                0.001
  MAT_I, P4                  0.001
-----

```

Note: The reliability estimates reported above are based on only 9207 of 12153 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Tau(beta)

INTRCPT1	MAT_I
INTRCPT2,B00	INTRCPT2,B40
191.22998	-0.58144
-0.58144	0.00434

Tau(beta) (as correlations)

INTRCPT1/INTRCPT2,B00	1.000	-0.638
MAT_I/INTRCPT2,B40	-0.638	1.000

```

-----
Random level-2 coefficient  Reliability estimate
-----
INTRCPT1/INTRCPT2, B00      0.731
  MAT_I/INTRCPT2, B40      0.662
-----

```

The value of the likelihood function at iteration 2463 = -1.684360E+005
The outcome variable is MAT_F

Final estimation of fixed effects
(with robust standard errors)

```

-----

```

Fixed Effect	Standard Coefficient	Standard Error	Approx. T-ratio	d.f.	P-value
For INTRCPT1, P0					
For INTRCPT2, B00					
INTRCPT3, G000	59.692826	4.924900	12.121	229	0.000
For T2 slope, P1					
For INTRCPT2, B10					
INTRCPT3, G100	-81.617714	5.009702	-16.292	34552	0.000
For T3 slope, P2					
For INTRCPT2, B20					
INTRCPT3, G200	4.153126	4.283248	0.970	34552	0.333
For T4 slope, P3					
For INTRCPT2, B30					
INTRCPT3, G300	16.636345	3.922345	4.241	34552	0.000
For MAT_I slope, P4					
For INTRCPT2, B40					
INTRCPT3, G400	0.588629	0.033772	17.429	229	0.000
For T2XMAT_I slope, P5					

```

-----

```

For INTRCPT2, B50						
INTRCPT3, G500	0.627854	0.037168	16.893	34552	0.000	
For T3XMAT_I slope, P6						
For INTRCPT2, B60						
INTRCPT3, G600	0.186425	0.033485	5.567	34552	0.000	
For T4XMAT_I slope, P7						
For INTRCPT2, B70						
INTRCPT3, G700	0.174227	0.030955	5.628	34552	0.000	
For DEFAS slope, P8						
For INTRCPT2, B80						
INTRCPT3, G800	16.962911	3.104164	5.465	34552	0.000	

Final estimation of level-1 and level-2 variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, R0	2.50393	6.26968	8977	8267.45761	>.500
MAT_I slope, R4	0.01401	0.00020	8977	9421.11452	0.001
level-1, E	31.16662	971.35803			

Note: The chi-square statistics reported above are based on only 9207 of 12153 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Final estimation of level-3 variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1/INTRCPT2, U00	13.82859	191.22998	229	920.42314	0.000
MAT_I/INTRCPT2, U40	0.06590	0.00434	229	778.64564	0.000

Statistics for current covariance components model

Deviance = 336871.962245
Number of estimated parameters = 16

Model comparison test

Chi-square statistic = 46409.53170
Number of degrees of freedom = 12
P-value = 0.000

ANEXO 09 - Relatório HLM – Matemática – Modelo Nível 2

Program: HLM 6 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: HLM3.EXE (6.08.29257.1)
Date: 19 July 2012, Thursday
Time: 16:45:32

SPECIFICATIONS FOR THIS HLM3 RUN

Problem Title: no title

The data source for this run = Modelo_matemática.mdm

The command file for this run = C:\Users\Adilson Dalben\Documents\Meus Documentos Adilson\Doutorado\00 - Tese_Banco-Geres\00-Interpretação_HLM_Matematica\Modelo N2 - 18 - PM005.hlm

Output file name = C:\Users\Adilson Dalben\Documents\Meus Documentos Adilson\Doutorado\00 - Tese_Banco-Geres\00-Interpretação_HLM_Matematica\hlm3.txt

The maximum number of level-1 units = 34561

The maximum number of level-2 units = 12153

The maximum number of level-3 units = 230

The maximum number of iterations = 100

Method of estimation: full maximum likelihood

Weighting Specification

Weight
Variable
Weighting? Name Normalized?
Level 1 no
Level 2 yes PESO_ALU yes
Level 3 no
Precision no

The outcome variable is MAT_F

The model specified for the fixed effects was:

Level-1 Level-2 Level-3
Coefficients Predictors Predictors

INTRCPT1, P0 INTRCPT2, B00 INTRCPT3, G000
 # FA04_PRE, B01 INTRCPT3, G010
 # FA12, B02 INTRCPT3, G020
 # FA13, B03 INTRCPT3, G030
 # FA14, B04 INTRCPT3, G040
T2 slope, P1 # INTRCPT2, B10 INTRCPT3, G100
 # FA04_PRE, B11 INTRCPT3, G110
 #FA06_NSE, B12 INTRCPT3, G120

```

# FA07, B13 INTRCPT3, G130
# FA09, B14 INTRCPT3, G140
# T3 slope, P2 # INTRCPT2, B20 INTRCPT3, G200
# FA07, B21 INTRCPT3, G210
# T4 slope, P3 # INTRCPT2, B30 INTRCPT3, G300
# FA01_SEX, B31 INTRCPT3, G310
# FA09, B32 INTRCPT3, G320
# FA14, B33 INTRCPT3, G330
MAT_I slope, P4 INTRCPT2, B40 INTRCPT3, G400
# FA01_SEX, B41 INTRCPT3, G410
# FA04_PRE, B42 INTRCPT3, G420
# FA07, B43 INTRCPT3, G430
# FA08, B44 INTRCPT3, G440
# T2XMAT_I slope, P5 # INTRCPT2, B50 INTRCPT3, G500
# FA04_PRE, B51 INTRCPT3, G510
# FA07, B52 INTRCPT3, G520
# FA09, B53 INTRCPT3, G530
# FA12, B54 INTRCPT3, G540
# T3XMAT_I slope, P6 # INTRCPT2, B60 INTRCPT3, G600
# FA07, B61 INTRCPT3, G610
# FA12, B62 INTRCPT3, G620
# T4XMAT_I slope, P7 # INTRCPT2, B70 INTRCPT3, G700
#%FA06_NSE, B71 INTRCPT3, G710
# FA07, B72 INTRCPT3, G720
# FA12, B73 INTRCPT3, G730
# FA14, B74 INTRCPT3, G740
# DEFAS slope, P8 # INTRCPT2, B80 INTRCPT3, G800
# FA09, B81 INTRCPT3, G810
# FA11, B82 INTRCPT3, G820
# FA12, B83 INTRCPT3, G830
# FA14, B84 INTRCPT3, G840

```

'#' - The residual parameter variance for the parameter has been set to zero

'%' - This variable has been centered around its grand mean

Summary of the model specified (in equation format)

Level-1 Model

$$Y = P0 + P1*(T2) + P2*(T3) + P3*(T4) + P4*(MAT_I) + P5*(T2XMAT_I) + P6*(T3XMAT_I) + P7*(T4XMAT_I) + P8*(DEFAS) + E$$

Level-2 Model

$$\begin{aligned}
P0 &= B00 + B01*(FA04_PRE) + B02*(FA12) + B03*(FA13) + B04*(FA14) + R0 \\
P1 &= B10 + B11*(FA04_PRE) + B12*(FA06_NSE) + B13*(FA07) + B14*(FA09) \\
P2 &= B20 + B21*(FA07) \\
P3 &= B30 + B31*(FA01_SEX) + B32*(FA09) + B33*(FA14) \\
P4 &= B40 + B41*(FA01_SEX) + B42*(FA04_PRE) + B43*(FA07) + B44*(FA08) + R4 \\
P5 &= B50 + B51*(FA04_PRE) + B52*(FA07) + B53*(FA09) + B54*(FA12) \\
P6 &= B60 + B61*(FA07) + B62*(FA12) \\
P7 &= B70 + B71*(FA06_NSE) + B72*(FA07) + B73*(FA12) + B74*(FA14) \\
P8 &= B80 + B81*(FA09) + B82*(FA11) + B83*(FA12) + B84*(FA14)
\end{aligned}$$

Level-3 Model

B00 = G000 + U00
B01 = G010
B02 = G020
B03 = G030
B04 = G040
B10 = G100
B11 = G110
B12 = G120
B13 = G130
B14 = G140
B20 = G200
B21 = G210
B30 = G300
B31 = G310
B32 = G320
B33 = G330
B40 = G400 + U40
B41 = G410
B42 = G420
B43 = G430
B44 = G440
B50 = G500
B51 = G510
B52 = G520
B53 = G530
B54 = G540
B60 = G600
B61 = G610
B62 = G620
B70 = G700
B71 = G710
B72 = G720
B73 = G730
B74 = G740
B80 = G800
B81 = G810
B82 = G820
B83 = G830
B84 = G840

For starting values, data from 31207 level-1 and 9207 level-2 records were used

Iterations stopped due to small change in likelihood function

***** ITERATION 2324 *****

Standard errors for sigma_squared, Tau(pi), Tau(beta) are not computable.

Sigma_squared = 921.17879

Tau(pi)

INTRCPT1,P0	7.31846	-0.03324
MAT_I,P1	-0.03324	0.00020

Tau(pi) (as correlations)
 INTRCPT1,P0 1.000 -0.862
 MAT_I,P4 -0.862 1.000

Random level-1 coefficient	Reliability estimate
INTRCPT1, P0	0.002
MAT_I, P4	0.001

Note: The reliability estimates reported above are based on only 9207 of 12153 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Tau(beta)
 INTRCPT1 MAT_I
 INTRCPT2,B00 INTRCPT2,B40
 132.89709 -0.48128
 -0.48128 0.00375

Tau(beta) (as correlations)
 INTRCPT1/INTRCPT2,B00 1.000 -0.682
 MAT_I/INTRCPT2,B40 -0.682 1.000

Random level-2 coefficient	Reliability estimate
INTRCPT1/INTRCPT2, B00	0.676
MAT_I/INTRCPT2, B40	0.643

The value of the likelihood function at iteration 2324 = -1.674852E+005
 The outcome variable is MAT_F

Final estimation of fixed effects
 (with robust standard errors)

Fixed Effect	Standard Coefficient	Standard Error	Approx. T-ratio	d.f.	P-value
For INTRCPT1, P0					
For INTRCPT2, B00					
INTRCPT3, G000	48.770027	5.222531	9.338	229	0.000
For FA04_PRE, B01					
INTRCPT3, G010	9.276324	2.759337	3.362	12148	0.001
For FA12, B02					
INTRCPT3, G020	7.391590	2.801224	2.639	12148	0.009
For FA13, B03					
INTRCPT3, G030	-4.708028	0.989605	-4.757	12148	0.000
For FA14, B04					
INTRCPT3, G040	7.975560	2.231573	3.574	12148	0.001
For T2 slope, P1					
For INTRCPT2, B10					

INTRCPT3, G100	-61.559830	17.521280	-3.513	34522	0.001
For FA04_PRE, B11					
INTRCPT3, G110	-21.895423	8.843496	-2.476	34522	0.014
For FA06_NSE, B12					
INTRCPT3, G120	7.005423	1.284022	5.456	34522	0.000
For FA07, B13					
INTRCPT3, G130	19.427428	5.723818	3.394	34522	0.001
For FA09, B14					
INTRCPT3, G140	-18.579223	6.006047	-3.093	34522	0.002
For T3 slope, P2					
For INTRCPT2, B20					
INTRCPT3, G200	-16.490463	7.034952	-2.344	34522	0.019
For FA07, B21					
INTRCPT3, G210	11.955186	3.721553	3.212	34522	0.002
For T4 slope, P3					
For INTRCPT2, B30					
INTRCPT3, G300	32.633458	4.846822	6.733	34522	0.000
For FA01_SEX, B31					
INTRCPT3, G310	4.241767	1.495949	2.836	34522	0.005
For FA09, B32					
INTRCPT3, G320	-6.667184	1.796276	-3.712	34522	0.000
For FA14, B33					
INTRCPT3, G330	-9.630809	3.781674	-2.547	34522	0.011
For MAT_I slope, P4					
For INTRCPT2, B40					
INTRCPT3, G400	0.436983	0.041483	10.534	229	0.000
For FA01_SEX, B41					
INTRCPT3, G410	-0.023060	0.004959	-4.650	12148	0.000
For FA04_PRE, B42					
INTRCPT3, G420	-0.040359	0.017019	-2.371	12148	0.018
For FA07, B43					
INTRCPT3, G430	0.073028	0.007600	9.609	12148	0.000
For FA08, B44					
INTRCPT3, G440	0.009481	0.003893	2.436	12148	0.015
For T2XMAT_I slope, P5					
For INTRCPT2, B50					
INTRCPT3, G500	0.453704	0.137418	3.302	34522	0.001
For FA04_PRE, B51					
INTRCPT3, G510	0.176573	0.067438	2.618	34522	0.009
For FA07, B52					
INTRCPT3, G520	-0.130838	0.036820	-3.553	34522	0.001
For FA09, B53					
INTRCPT3, G530	0.107854	0.044452	2.426	34522	0.015
For FA12, B54					
INTRCPT3, G540	0.047502	0.021987	2.160	34522	0.031
For T3XMAT_I slope, P6					
For INTRCPT2, B60					
INTRCPT3, G600	0.334916	0.051360	6.521	34522	0.000
For FA07, B61					
INTRCPT3, G610	-0.085513	0.020132	-4.248	34522	0.000
For FA12, B62					
INTRCPT3, G620	0.022787	0.010550	2.160	34522	0.031
For T4XMAT_I slope, P7					
For INTRCPT2, B70					
INTRCPT3, G700	0.176998	0.042881	4.128	34522	0.000
For FA06_NSE, B71					

INTRCPT3, G710	0.021805	0.005181	4.208	34522	0.000
For FA07, B72					
INTRCPT3, G720	-0.031480	0.008345	-3.772	34522	0.000
For FA12, B73					
INTRCPT3, G730	0.026684	0.008045	3.317	34522	0.001
For FA14, B74					
INTRCPT3, G740	0.050687	0.018878	2.685	34522	0.008
For DEFAS slope, P8					
For INTRCPT2, B80					
INTRCPT3, G800	32.701877	3.395342	9.631	34522	0.000
For FA09, B81					
INTRCPT3, G810	-4.982747	1.304549	-3.820	34522	0.000
For FA11, B82					
INTRCPT3, G820	-2.826747	0.645762	-4.377	34522	0.000
For FA12, B83					
INTRCPT3, G830	-4.861191	2.580137	-1.884	34522	0.059
For FA14, B84					
INTRCPT3, G840	-8.363138	2.220337	-3.767	34522	0.000

Final estimation of level-1 and level-2 variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, R0	2.70527	7.31846	8973	8081.35876	>.500
MAT_I slope, R4	0.01425	0.00020	8973	9165.47287	0.076
level-1, E	30.35093	921.17879			

Note: The chi-square statistics reported above are based on only 9207 of 12153 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Final estimation of level-3 variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1/INTRCPT2, U00	11.52810	132.89709	229	822.24242	0.000
MAT_I/INTRCPT2, U40	0.06120	0.00375	229	763.91348	0.000

Statistics for current covariance components model

Deviance = 334970.432292
Number of estimated parameters = 46

Model comparison test

Chi-square statistic = 1901.52995

Number of degrees of freedom = 30
P-value = 0.000

ANEXO 10 - Relatório HLM - Matemática- Modelo Nível 3

Program: HLM 6 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: HLM3.EXE (6.08.29257.1)
 Date: 22 July 2012, Sunday
 Time: 8:57: 0

SPECIFICATIONS FOR THIS HLM3 RUN

Problem Title: no title

The data source for this run = Modelo_matemática.mdm

The command file for this run = C:\Users\Adilson Dalben\Documents\Meus Documentos Adilson\Doutorado\00 - Tese_Banco-Geres\00-Interpretação_HLM_ Matematica\Modelo N3 - 38 - PM007 - PCSI e-5.hlm

Output file name = C:\Users\Adilson Dalben\Documents\Meus Documentos Adilson\Doutorado\00 - Tese_Banco-Geres\00-Interpretação_HLM_ Matematica\hlm3.txt

The maximum number of level-1 units = 34561

The maximum number of level-2 units = 12153

The maximum number of level-3 units = 230

The maximum number of iterations = 100

Method of estimation: full maximum likelihood

Weighting Specification

	Weight	Variable	Normalized?
	Weighting?	Name	
Level 1	no		
Level 2	yes	PESO_ALU	yes
Level 3	no		
Precision	no		

The outcome variable is MAT_F

The model specified for the fixed effects was:

Level-1	Level-2	Level-3
Coefficients	Predictors	Predictors
INTRCPT1, P0	INTRCPT2, B00	INTRCPT3, G000
	% FEA01, G001	
	FEE02, G002	
	FEE07, G003	
	FED12, G004	
	FED18, G005	

```

        FED25, G006
        FEP32, G007
        FEP34, G008
        FEP39, G009
        FEP41, G0010
        FEP48, G0011
        FEP49, G0012
        FEP54, G0013
# FA04_PRE, B01 INTRCPT3, G010
# FA12, B02 INTRCPT3, G020
# FA13, B03 INTRCPT3, G030
# FA14, B04 INTRCPT3, G040
# T2 slope, P1 # INTRCPT2, B10 INTRCPT3, G100
        FEE02, G101
        FEE03, G102
        FEE04, G103
        FEE11, G104
        FED22, G105
        FED23, G106
        FED25, G107
        FEP32, G108
        FEP34, G109
        FEP38, G1010
        FEP41, G1011
        FEP45, G1012
        FEP47, G1013
        FEP49, G1014
        FEP53, G1015
        FEP56, G1016
        FEP58, G1017
# FA04_PRE, B11 INTRCPT3, G110
#%FA06_NSE, B12 INTRCPT3, G120
# FA07, B13 INTRCPT3, G130
# FA09, B14 INTRCPT3, G140
# T3 slope, P2 # INTRCPT2, B20 INTRCPT3, G200
        % FEA01, G201
        FEE04, G202
        FEE08, G203
        FEP34, G204
        FEP40, G205
        FEP49, G206
        FEP50, G207
# FA07, B21 INTRCPT3, G210
# T4 slope, P3 # INTRCPT2, B30 INTRCPT3, G300
        FEE02, G301
        FEE09, G302
        FEE10, G303
        FED16, G304
        FED17, G305
        FED19, G306
        FED20, G307
        FEP33, G308
        FEP40, G309
        FEP41, G3010
        FEP49, G3011
        FEP50, G3012

```

```

          FEP51, G3013
# FA01_SEX, B31 INTRCPT3, G310
# FA09, B32 INTRCPT3, G320
# FA14, B33 INTRCPT3, G330
MAT_I slope, P4 INTRCPT2, B40 INTRCPT3, G400
          % FEA01, G401
          FEE02, G402
          FEE03, G403
          FEE07, G404
          FEE11, G405
          FED13, G406
          FED18, G407
          FED21, G408
          FED25, G409
          FEP29, G4010
          FEP32, G4011
          FEP34, G4012
          FEP35, G4013
          FEP37, G4014
          FEP39, G4015
          FEP47, G4016
          FEP51, G4017
# FA01_SEX, B41 INTRCPT3, G410
# FA04_PRE, B42 INTRCPT3, G420
# FA07, B43 INTRCPT3, G430
# FA08, B44 INTRCPT3, G440
# T2XMAT_I slope, P5 # INTRCPT2, B50 INTRCPT3, G500
          % FEA01, G501
          FEE02, G502
          FEE11, G503
          FED21, G504
          FED22, G505
          FEP27, G506
          FEP32, G507
          FEP34, G508
          FEP41, G509
          FEP45, G5010
          FEP47, G5011
          FEP49, G5012
          FEP51, G5013
          FEP53, G5014
# FA04_PRE, B51 INTRCPT3, G510
# FA07, B52 INTRCPT3, G520
# FA09, B53 INTRCPT3, G530
# FA12, B54 INTRCPT3, G540
# T3XMAT_I slope, P6 # INTRCPT2, B60 INTRCPT3, G600
          % FEA01, G601
          FEE02, G602
          FEE03, G603
          FEE04, G604
          FEE08, G605
          FED18, G606
          FED21, G607
          FEP34, G608
          FEP49, G609
          FEP50, G6010

```

```

                FEP51, G6011
                FEP55, G6012
#   FA07, B61  INTRCPT3, G610
#   FA12, B62  INTRCPT3, G620
# T4XMAT_I slope, P7 # INTRCPT2, B70  INTRCPT3, G700
                % FEA01, G701
                FEE02, G702
                FEE09, G703
                FED13, G704
                FED16, G705
                FED18, G706
                FED19, G707
                FED20, G708
                FED21, G709
                FEP33, G7010
                FEP34, G7011
                FEP48, G7012
                FEP53, G7013
                FEP55, G7014
#%FA06_NSE, B71  INTRCPT3, G710
#   FA07, B72  INTRCPT3, G720
#   FA12, B73  INTRCPT3, G730
#   FA14, B74  INTRCPT3, G740
#  DEFAS slope, P8 # INTRCPT2, B80  INTRCPT3, G800
#   FA09, B81  INTRCPT3, G810
#   FA11, B82  INTRCPT3, G820
#   FA12, B83  INTRCPT3, G830
#   FA14, B84  INTRCPT3, G840

```

'#' - The residual parameter variance for the parameter has been set to zero

'%' - This variable has been centered around its grand mean

Summary of the model specified (in equation format)

Level-1 Model

$$Y = P0 + P1*(T2) + P2*(T3) + P3*(T4) + P4*(MAT_I) + P5*(T2XMAT_I) + P6*(T3XMAT_I) + P7*(T4XMAT_I) + P8*(DEFAS) + E$$

Level-2 Model

$$\begin{aligned}
 P0 &= B00 + B01*(FA04_PRE) + B02*(FA12) + B03*(FA13) + B04*(FA14) + R0 \\
 P1 &= B10 + B11*(FA04_PRE) + B12*(FA06_NSE) + B13*(FA07) + B14*(FA09) \\
 P2 &= B20 + B21*(FA07) \\
 P3 &= B30 + B31*(FA01_SEX) + B32*(FA09) + B33*(FA14) \\
 P4 &= B40 + B41*(FA01_SEX) + B42*(FA04_PRE) + B43*(FA07) + B44*(FA08) + R4 \\
 P5 &= B50 + B51*(FA04_PRE) + B52*(FA07) + B53*(FA09) + B54*(FA12) \\
 P6 &= B60 + B61*(FA07) + B62*(FA12) \\
 P7 &= B70 + B71*(FA06_NSE) + B72*(FA07) + B73*(FA12) + B74*(FA14) \\
 P8 &= B80 + B81*(FA09) + B82*(FA11) + B83*(FA12) + B84*(FA14)
 \end{aligned}$$

Level-3 Model

B00 = G000 + G001(FEA01) + G002(FEE02) + G003(FEE07) + G004(FED12)
 + G005(FED18) + G006(FED25) + G007(FEP32) + G008(FEP34)
 + G009(FEP39) + G0010(FEP41) + G0011(FEP48) + G0012(FEP49)
 + G0013(FEP54) + U00
 B01 = G010
 B02 = G020
 B03 = G030
 B04 = G040
 B10 = G100 + G101(FEE02) + G102(FEE03) + G103(FEE04) + G104(FEE11)
 + G105(FED22) + G106(FED23) + G107(FED25) + G108(FEP32)
 + G109(FEP34) + G1010(FEP38) + G1011(FEP41) + G1012(FEP45)
 + G1013(FEP47) + G1014(FEP49) + G1015(FEP53) + G1016(FEP56) + G1017(FEP58)
 B11 = G110
 B12 = G120
 B13 = G130
 B14 = G140
 B20 = G200 + G201(FEA01) + G202(FEE04) + G203(FEE08) + G204(FEP34)
 + G205(FEP40) + G206(FEP49) + G207(FEP50)
 B21 = G210
 B30 = G300 + G301(FEE02) + G302(FEE09) + G303(FEE10) + G304(FED16)
 + G305(FED17) + G306(FED19) + G307(FED20) + G308(FEP33)
 + G309(FEP40) + G3010(FEP41) + G3011(FEP49) + G3012(FEP50)
 + G3013(FEP51)
 B31 = G310
 B32 = G320
 B33 = G330
 B40 = G400 + G401(FEA01) + G402(FEE02) + G403(FEE03) + G404(FEE07)
 + G405(FEE11) + G406(FED13) + G407(FED18) + G408(FED21)
 + G409(FED25) + G4010(FEP29) + G4011(FEP32) + G4012(FEP34)
 + G4013(FEP35) + G4014(FEP37) + G4015(FEP39) + G4016(FEP47) + G4017(FEP51) + U40
 B41 = G410
 B42 = G420
 B43 = G430
 B44 = G440
 B50 = G500 + G501(FEA01) + G502(FEE02) + G503(FEE11) + G504(FED21)
 + G505(FED22) + G506(FEP27) + G507(FEP32) + G508(FEP34)
 + G509(FEP41) + G5010(FEP45) + G5011(FEP47) + G5012(FEP49)
 + G5013(FEP51) + G5014(FEP53)
 B51 = G510
 B52 = G520
 B53 = G530
 B54 = G540
 B60 = G600 + G601(FEA01) + G602(FEE02) + G603(FEE03) + G604(FEE04)
 + G605(FEE08) + G606(FED18) + G607(FED21) + G608(FEP34)
 + G609(FEP49) + G6010(FEP50) + G6011(FEP51) + G6012(FEP55)
 B61 = G610
 B62 = G620
 B70 = G700 + G701(FEA01) + G702(FEE02) + G703(FEE09) + G704(FED13)
 + G705(FED16) + G706(FED18) + G707(FED19) + G708(FED20)
 + G709(FED21) + G7010(FEP33) + G7011(FEP34) + G7012(FEP48)
 + G7013(FEP53) + G7014(FEP55)
 B71 = G710
 B72 = G720
 B73 = G730
 B74 = G740
 B80 = G800

B81 = G810
B82 = G820
B83 = G830
B84 = G840

For starting values, data from 31207 level-1 and 9207 level-2 records were used

Iterations stopped due to small change in likelihood function

***** ITERATION 661 *****

Standard errors for sigma_squared, Tau(pi), Tau(beta) are not computable.

Sigma_squared = 892.93850

Tau(pi)

INTRCPT1,P0 18.50033 -0.08247
MAT_I,P1 -0.08247 0.00059

Tau(pi) (as correlations)

INTRCPT1,P0 1.000 -0.787
MAT_I,P4 -0.787 1.000

Random level-1 coefficient Reliability estimate

INTRCPT1, P0 0.004
MAT_I, P4 0.004

Note: The reliability estimates reported above are based on only 9207 of 12153 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Tau(beta)

INTRCPT1 MAT_I
INTRCPT2,B00 INTRCPT2,B40
24.03127 -0.14195
-0.14195 0.00112

Tau(beta) (as correlations)

INTRCPT1/INTRCPT2,B00 1.000 -0.864
MAT_I/INTRCPT2,B40 -0.864 1.000

Random level-2 coefficient Reliability estimate

INTRCPT1/INTRCPT2, B00 0.333
MAT_I/INTRCPT2, B40 0.395

The value of the likelihood function at iteration 661 = -1.667968E+005

The outcome variable is MAT_F

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Standard Coefficient	Standard Error	Approx. T-ratio	d.f.	P-value

For INTRCPT1, P0					
For INTRCPT2, B00					
INTRCPT3, G000	41.025168	10.558704	3.885	216	0.000
FEA01, G001	33.165886	6.994829	4.741	216	0.000
FEE02, G002	14.246174	3.652748	3.900	216	0.000
FEE07, G003	5.128520	1.652795	3.103	216	0.003
FED12, G004	-4.333201	1.215579	-3.565	216	0.001
FED18, G005	-4.077523	1.350889	-3.018	216	0.003
FED25, G006	5.295337	1.744492	3.035	216	0.003
FEP32, G007	-10.962443	3.714114	-2.952	216	0.004
FEP34, G008	9.942335	3.102397	3.205	216	0.002
FEP39, G009	-6.045304	1.604683	-3.767	216	0.000
FEP41, G0010	-2.482307	1.134406	-2.188	216	0.030
FEP48, G0011	-1.632464	0.675077	-2.418	216	0.017
FEP49, G0012	-2.317526	0.705829	-3.283	216	0.002
FEP54, G0013	1.809653	0.942890	1.919	216	0.056
For FA04_PRE, B01					
INTRCPT3, G010	8.422401	2.782104	3.027	12148	0.003
For FA12, B02					
INTRCPT3, G020	7.889119	2.590890	3.045	12148	0.003
For FA13, B03					
INTRCPT3, G030	-4.692511	0.985473	-4.762	12148	0.000
For FA14, B04					
INTRCPT3, G040	7.945600	2.124939	3.739	12148	0.000
For T2 slope, P1					
For INTRCPT2, B10					
INTRCPT3, G100	-133.381240	39.783373	-3.353	34415	0.001
FEE02, G101	-25.404787	8.492398	-2.991	34415	0.003
FEE03, G102	1.928560	0.972730	1.983	34415	0.047
FEE04, G103	-1.295460	0.605933	-2.138	34415	0.032
FEE11, G104	-9.394964	4.313821	-2.178	34415	0.029
FED22, G105	27.824024	7.850924	3.544	34415	0.001
FED23, G106	2.339467	1.149680	2.035	34415	0.042
FED25, G107	-3.694980	1.144222	-3.229	34415	0.002
FEP32, G108	28.843693	9.980327	2.890	34415	0.004
FEP34, G109	-16.604962	5.317021	-3.123	34415	0.002
FEP38, G1010	-5.654391	2.631362	-2.149	34415	0.031
FEP41, G1011	22.218076	8.357057	2.659	34415	0.008
FEP45, G1012	19.058976	7.048664	2.704	34415	0.007
FEP47, G1013	9.794128	4.750098	2.062	34415	0.039
FEP49, G1014	-9.966265	3.442593	-2.895	34415	0.004
FEP53, G1015	17.542389	5.700962	3.077	34415	0.003
FEP56, G1016	5.227831	1.529594	3.418	34415	0.001
FEP58, G1017	-3.195044	1.465393	-2.180	34415	0.029
For FA04_PRE, B11					
INTRCPT3, G110	-24.237324	10.366160	-2.338	34415	0.019
For FA06_NSE, B12					
INTRCPT3, G120	3.556424	1.831337	1.942	34415	0.052
For FA07, B13					

INTRCPT3, G130	16.777776	5.289025	3.172	34415	0.002
For FA09, B14					
INTRCPT3, G140	-17.096710	5.742085	-2.977	34415	0.003
For T3 slope, P2					
For INTRCPT2, B20					
INTRCPT3, G200	19.470153	11.360403	1.714	34415	0.086
FEA01, G201	-36.854036	9.081615	-4.058	34415	0.000
FEE04, G202	-4.009030	1.579751	-2.538	34415	0.011
FEE08, G203	5.754431	1.808163	3.182	34415	0.002
FEP34, G204	-18.207177	4.411625	-4.127	34415	0.000
FEP40, G205	-3.736032	1.239056	-3.015	34415	0.003
FEP49, G206	8.092382	2.178232	3.715	34415	0.000
FEP50, G207	-9.286714	2.858766	-3.249	34415	0.002
For FA07, B21					
INTRCPT3, G210	11.305804	3.574029	3.163	34415	0.002
For T4 slope, P3					
For INTRCPT2, B30					
INTRCPT3, G300	116.297766	20.310576	5.726	34415	0.000
FEE02, G301	-16.056832	5.984990	-2.683	34415	0.008
FEE09, G302	9.479314	2.795191	3.391	34415	0.001
FEE10, G303	2.154169	0.848873	2.538	34415	0.011
FED16, G304	-17.366232	5.286444	-3.285	34415	0.001
FED17, G305	-4.204757	1.845811	-2.278	34415	0.023
FED19, G306	6.284429	1.950142	3.223	34415	0.002
FED20, G307	24.292398	4.537584	5.354	34415	0.000
FEP33, G308	-21.027977	5.110172	-4.115	34415	0.000
FEP40, G309	-4.731587	1.248673	-3.789	34415	0.000
FEP41, G3010	7.660355	2.362053	3.243	34415	0.002
FEP49, G3011	2.331982	0.919210	2.537	34415	0.011
FEP50, G3012	-5.184145	1.013740	-5.114	34415	0.000
FEP51, G3013	-10.388826	2.587892	-4.014	34415	0.000
For FA01_SEX, B31					
INTRCPT3, G310	4.165475	1.445105	2.882	34415	0.004
For FA09, B32					
INTRCPT3, G320	-6.678862	1.757395	-3.800	34415	0.000
For FA14, B33					
INTRCPT3, G330	-10.788647	3.729577	-2.893	34415	0.004
For MAT_I slope, P4					
For INTRCPT2, B40					
INTRCPT3, G400	0.417061	0.073114	5.704	212	0.000
FEA01, G401	-0.162401	0.049983	-3.249	212	0.002
FEE02, G402	-0.155649	0.036105	-4.311	212	0.000
FEE03, G403	-0.011986	0.003590	-3.339	212	0.001
FEE07, G404	-0.030060	0.010788	-2.786	212	0.006
FEE11, G405	-0.012386	0.003478	-3.561	212	0.001
FED13, G406	-0.019424	0.007859	-2.472	212	0.014
FED18, G407	0.036984	0.011111	3.328	212	0.001
FED21, G408	0.026417	0.013626	1.939	212	0.053
FED25, G409	-0.033855	0.009738	-3.477	212	0.001
FEP29, G4010	0.026756	0.006522	4.102	212	0.000
FEP32, G4011	0.083147	0.022413	3.710	212	0.000
FEP34, G4012	-0.097177	0.026460	-3.673	212	0.001
FEP35, G4013	0.010902	0.004604	2.368	212	0.019
FEP37, G4014	0.027563	0.010130	2.721	212	0.007
FEP39, G4015	0.039728	0.009678	4.105	212	0.000
FEP47, G4016	0.013661	0.003629	3.765	212	0.000

FEP51, G4017	0.035425	0.013127	2.699	212	0.008
For FA01_SEX, B41					
INTRCPT3, G410	-0.024005	0.004849	-4.951	12148	0.000
For FA04_PRE, B42					
INTRCPT3, G420	-0.033418	0.016800	-1.989	12148	0.046
For FA07, B43					
INTRCPT3, G430	0.074275	0.006027	12.324	12148	0.000
For FA08, B44					
INTRCPT3, G440	0.008464	0.003911	2.164	12148	0.030
For T2XMAT_I slope, P5					
For INTRCPT2, B50					
INTRCPT3, G500	1.177037	0.259405	4.537	34415	0.000
FEA01, G501	0.144345	0.028276	5.105	34415	0.000
FEE02, G502	0.201391	0.062497	3.222	34415	0.002
FEE11, G503	0.064028	0.026988	2.372	34415	0.018
FED21, G504	-0.047469	0.018969	-2.502	34415	0.013
FED22, G505	-0.187684	0.060439	-3.105	34415	0.002
FEP27, G506	-0.034687	0.012196	-2.844	34415	0.005
FEP32, G507	-0.179297	0.068673	-2.611	34415	0.009
FEP34, G508	0.123841	0.039156	3.163	34415	0.002
FEP41, G509	-0.128073	0.051981	-2.464	34415	0.014
FEP45, G5010	-0.108641	0.049729	-2.185	34415	0.029
FEP47, G5011	-0.067864	0.029561	-2.296	34415	0.022
FEP49, G5012	0.085002	0.026603	3.195	34415	0.002
FEP51, G5013	-0.049210	0.018803	-2.617	34415	0.009
FEP53, G5014	-0.144487	0.042195	-3.424	34415	0.001
For FA04_PRE, B51					
INTRCPT3, G510	0.197177	0.075119	2.625	34415	0.009
For FA07, B52					
INTRCPT3, G520	-0.118347	0.033991	-3.482	34415	0.001
For FA09, B53					
INTRCPT3, G530	0.099682	0.041531	2.400	34415	0.017
For FA12, B54					
INTRCPT3, G540	0.044347	0.018407	2.409	34415	0.016
For T3XMAT_I slope, P6					
For INTRCPT2, B60					
INTRCPT3, G600	0.205723	0.077167	2.666	34415	0.008
FEA01, G601	0.263477	0.051271	5.139	34415	0.000
FEE02, G602	0.055356	0.020151	2.747	34415	0.006
FEE03, G603	0.015904	0.006488	2.451	34415	0.014
FEE04, G604	0.022890	0.009257	2.473	34415	0.014
FEE08, G605	-0.034105	0.010092	-3.380	34415	0.001
FED18, G606	-0.016589	0.005972	-2.778	34415	0.006
FED21, G607	-0.051772	0.015965	-3.243	34415	0.002
FEP34, G608	0.135340	0.027835	4.862	34415	0.000
FEP49, G609	-0.044259	0.014303	-3.094	34415	0.002
FEP50, G6010	0.055931	0.016412	3.408	34415	0.001
FEP51, G6011	-0.030540	0.015576	-1.961	34415	0.050
FEP55, G6012	-0.070829	0.035890	-1.974	34415	0.048
For FA07, B61					
INTRCPT3, G610	-0.083191	0.020297	-4.099	34415	0.000
For FA12, B62					
INTRCPT3, G620	0.020872	0.009357	2.231	34415	0.026
For T4XMAT_I slope, P7					
For INTRCPT2, B70					
INTRCPT3, G700	-0.148363	0.100344	-1.479	34415	0.139

FEA01, G701	0.188926	0.026010	7.264	34415	0.000
FEE02, G702	0.151100	0.043579	3.467	34415	0.001
FEE09, G703	-0.039571	0.014055	-2.815	34415	0.005
FED13, G704	0.026644	0.009973	2.671	34415	0.008
FED16, G705	0.096213	0.025164	3.823	34415	0.000
FED18, G706	-0.014866	0.006200	-2.398	34415	0.017
FED19, G707	-0.032817	0.010485	-3.130	34415	0.002
FED20, G708	-0.129004	0.023326	-5.530	34415	0.000
FED21, G709	-0.037802	0.014760	-2.561	34415	0.011
FEP33, G7010	0.084414	0.024746	3.411	34415	0.001
FEP34, G7011	0.053725	0.014229	3.776	34415	0.000
FEP48, G7012	-0.015592	0.004827	-3.230	34415	0.002
FEP53, G7013	-0.027665	0.007803	-3.546	34415	0.001
FEP55, G7014	-0.067260	0.027175	-2.475	34415	0.014
For FA06_NSE, B71					
INTRCPT3, G710	0.009485	0.005439	1.744	34415	0.081
For FA07, B72					
INTRCPT3, G720	-0.032253	0.006400	-5.040	34415	0.000
For FA12, B73					
INTRCPT3, G730	0.024670	0.007567	3.260	34415	0.001
For FA14, B74					
INTRCPT3, G740	0.057328	0.017987	3.187	34415	0.002
For DEFAS slope, P8					
For INTRCPT2, B80					
INTRCPT3, G800	31.895261	2.719875	11.727	34415	0.000
For FA09, B81					
INTRCPT3, G810	-4.671535	1.240502	-3.766	34415	0.000
For FA11, B82					
INTRCPT3, G820	-2.887286	0.620645	-4.652	34415	0.000
For FA12, B83					
INTRCPT3, G830	-5.043325	2.453753	-2.055	34415	0.040
For FA14, B84					
INTRCPT3, G840	-8.192341	2.117026	-3.870	34415	0.000

Final estimation of level-1 and level-2 variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, R0	4.30120	18.50033	8973	8100.30737	>.500
MAT_I slope, R4	0.02438	0.00059	8973	9220.94317	0.033
level-1, E	29.88208	892.93850			

Note: The chi-square statistics reported above are based on only 9207 of 12153 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Final estimation of level-3 variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
---------------	--------------------	--------------------	----	------------	---------

INTRCPT1/INTRCPT2, U00 4.90217 24.03127 216 435.33106 0.000
MAT_I/INTRCPT2, U40 0.03351 0.00112 212 465.82884 0.000

Statistics for current covariance components model

Deviance = 333593.526300
Number of estimated parameters = 153

Model comparison test

Chi-square statistic = 1376.90599
Number of degrees of freedom = 107
P-value = 0.000

ANEXO 11 - Relatório HLM – Modelo de Referência em Leitura

Program: HLM 6 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: HLM3.EXE (6.08.29257.1)
Date: 23 November 2013, Saturday
Time: 17: 1: 7

SPECIFICATIONS FOR THIS HLM3 RUN

Problem Title: no title

The data source for this run = Modelo_Leitura.mdm

The command file for this run = C:\Users\Adilson_Dalben\Documents\Meus Documentos
Adilson\Doutorado\00 - DEFESA\01 - Defesa_HLM_Leitura\Modelo_Leitura-02-REFERENCIA-09.hlm

Output file name = C:\Users\Adilson_Dalben\Documents\Meus Documentos
Adilson\DOUTORADO\00 - DEFESA\01 - Defesa_HLM_Leitura\hlm3.txt

The maximum number of level-1 units = 34657

The maximum number of level-2 units = 12180

The maximum number of level-3 units = 229

The maximum number of iterations = 100

Method of estimation: full maximum likelihood

Weighting Specification

Weight
Variable
Weighting? Name Normalized?
Level 1 no
Level 2 yes PESOAL yes
Level 3 no
Precision no

The outcome variable is LEITF

The model specified for the fixed effects was:

Level-1 Level-2 Level-3
Coefficients Predictors Predictors

INTRCPT1, P0 INTRCPT2, B00 INTRCPT3, G000
% FEA01, G001
#%FA06_NSE, B01 INTRCPT3, G010
T2 slope, P1 # INTRCPT2, B10 INTRCPT3, G100
% FEA01, G101
T3 slope, P2 # INTRCPT2, B20 INTRCPT3, G200
#%FA06_NSE, B21 INTRCPT3, G210
LEITI slope, P3 INTRCPT2, B30 INTRCPT3, G300
% FEA01, G301
LEITIXT2 slope, P4 # INTRCPT2, B40 INTRCPT3, G400

LEITIXT3 slope, P5 # INTRCPT2, B50 INTRCPT3, G500
 # LEITIXT4 slope, P6 # INTRCPT2, B60 INTRCPT3, G600
 # RET_AC slope, P7 # INTRCPT2, B70 INTRCPT3, G700

'#' - The residual parameter variance for the parameter has been set to zero
 '%' - This variable has been centered around its grand mean

Summary of the model specified (in equation format)

Level-1 Model

$$Y = P0 + P1*(T2) + P2*(T3) + P3*(LEITI) + P4*(LEITIXT2) + P5*(LEITIXT3) + P6*(LEITIXT4) + P7*(RET_AC) + E$$

Level-2 Model

$$\begin{aligned} P0 &= B00 + B01*(FA06_NSE) + R0 \\ P1 &= B10 \\ P2 &= B20 + B21*(FA06_NSE) \\ P3 &= B30 + R3 \\ P4 &= B40 \\ P5 &= B50 \\ P6 &= B60 \\ P7 &= B70 \end{aligned}$$

Level-3 Model

$$\begin{aligned} B00 &= G000 + G001(FAA01) + U00 \\ B01 &= G010 \\ B10 &= G100 + G101(FAA01) \\ B20 &= G200 \\ B21 &= G210 \\ B30 &= G300 + G301(FAA01) + U30 \\ B40 &= G400 \\ B50 &= G500 \\ B60 &= G600 \\ B70 &= G700 \end{aligned}$$

For starting values, data from 30012 level-1 and 8655 level-2 records were used

Iterations stopped due to small change in likelihood function

***** ITERATION 1877 *****

Standard errors for sigma_squared, Tau(pi), Tau(beta) are not computable.

Sigma_squared = 185.88017

Tau(pi)

INTRCPT1,P0 91.12224 -0.48154
 LEITI,P1 -0.48154 0.00256

Tau(pi) (as correlations)
 INTRCPT1,P0 1.000 -0.997
 LEITI,P3 -0.997 1.000

Random level-1 coefficient	Reliability estimate
INTRCPT1, P0	0.040
LEITI, P3	0.021

Note: The reliability estimates reported above are based on only 8655 of 12180 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Tau(beta)
 INTRCPT1 LEITI
 INTRCPT2,B00 INTRCPT2,B30
 117.95789 -0.58737
 -0.58737 0.00320

Tau(beta) (as correlations)
 INTRCPT1/INTRCPT2,B00 1.000 -0.956
 LEITI/INTRCPT2,B30 -0.956 1.000

Random level-2 coefficient	Reliability estimate
INTRCPT1/INTRCPT2, B00	0.736
LEITI/INTRCPT2, B30	0.642

The value of the likelihood function at iteration 1877 = -1.409177E+005
 The outcome variable is LEITF

Final estimation of fixed effects
 (with robust standard errors)

Fixed Effect	Standard Coefficient	Error	Approx. T-ratio	d.f.	P-value
For INTRCPT1, P0					
For INTRCPT2, B00					
INTRCPT3, G000	65.873322	2.799586	23.530	227	0.000
FEA01, G001	29.955380	5.844429	5.125	227	0.000
For FA06_NSE, B01					
INTRCPT3, G010	2.213134	0.327540	6.757	12178	0.000
For T2 slope, P1					
For INTRCPT2, B10					
INTRCPT3, G100	-31.077944	3.034150	-10.243	34644	0.000
FEA01, G101	-3.603676	1.216110	-2.963	34644	0.004
For T3 slope, P2					
For INTRCPT2, B20					
INTRCPT3, G200	-10.538071	2.709232	-3.890	34644	0.000
For FA06_NSE, B21					

INTRCPT3, G210	-1.708899	0.536781	-3.184	34644	0.002
For LEITI slope, P3					
For INTRCPT2, B30					
INTRCPT3, G300	0.518063	0.017726	29.225	227	0.000
FEA01, G301	-0.131387	0.026949	-4.875	227	0.000
For LEITIXT2 slope, P4					
For INTRCPT2, B40					
INTRCPT3, G400	0.293439	0.022620	12.972	34644	0.000
For LEITIXT3 slope, P5					
For INTRCPT2, B50					
INTRCPT3, G500	0.167611	0.020592	8.140	34644	0.000
For LEITIXT4 slope, P6					
For INTRCPT2, B60					
INTRCPT3, G600	0.109310	0.006084	17.967	34644	0.000
For RET_AC slope, P7					
For INTRCPT2, B70					
INTRCPT3, G700	6.030269	1.590242	3.792	34644	0.000

Final estimation of level-1 and level-2 variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, R0	9.54580	91.12224	8425	8692.61533	0.020
LEITI slope, R3	0.05058	0.00256	8426	8716.02437	0.013
level-1, E	13.63379	185.88017			

Note: The chi-square statistics reported above are based on only 8655 of 12180 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Final estimation of level-3 variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1/INTRCPT2, U00	10.86084	117.95789	227	790.77491	0.000
LEITI/INTRCPT2, U30	0.05656	0.00320	227	762.20683	0.000

Statistics for current covariance components model

Deviance = 281835.377705
Number of estimated parameters = 20


```
# T2XMAT_I slope, P4 # INTRCPT2, B40 INTRCPT3, G400
% FEA01, G401
# T3XMAT_I slope, P5 # INTRCPT2, B50 INTRCPT3, G500
# T4XMAT_I slope, P6 # INTRCPT2, B60 INTRCPT3, G600
# DEFAS slope, P7 # INTRCPT2, B70 INTRCPT3, G700
```

'#' - The residual parameter variance for the parameter has been set to zero
 '%' - This variable has been centered around its grand mean

Summary of the model specified (in equation format)

 Level-1 Model

$$Y = P0 + P1*(T2) + P2*(T4) + P3*(MAT_I) + P4*(T2XMAT_I) + P5*(T3XMAT_I) + P6*(T4XMAT_I) + P7*(DEFAS) + E$$

Level-2 Model

$$\begin{aligned} P0 &= B00 + R0 \\ P1 &= B10 \\ P2 &= B20 \\ P3 &= B30 + B31*(FA06_NSE) + R3 \\ P4 &= B40 \\ P5 &= B50 \\ P6 &= B60 \\ P7 &= B70 \end{aligned}$$

Level-3 Model

$$\begin{aligned} B00 &= G000 + G001(FA01) + U00 \\ B10 &= G100 + G101(FA01) \\ B20 &= G200 + G201(FA01) \\ B30 &= G300 + G301(FA01) + U30 \\ B31 &= G310 \\ B40 &= G400 + G401(FA01) \\ B50 &= G500 \\ B60 &= G600 \\ B70 &= G700 \end{aligned}$$

For starting values, data from 31207 level-1 and 9207 level-2 records were used

Iterations stopped due to small change in likelihood function

***** ITERATION 2516 *****

Standard errors for sigma_squared, Tau(pi), Tau(beta) are not computable.

Sigma_squared = 965.27893

Tau(pi)

```
INTRCPT1,P0 6.02049 -0.03037
MAT_I,P1 -0.03037 0.00020
```

Tau(pi) (as correlations)
 INTRCPT1,P0 1.000 -0.878
 MAT_I,P3 -0.878 1.000

Random level-1 coefficient	Reliability estimate
INTRCPT1, P0	0.001
MAT_I, P3	0.001

Note: The reliability estimates reported above are based on only 9207 of 12153 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Tau(beta)
 INTRCPT1 MAT_I
 INTRCPT2,B00 INTRCPT2,B30
 80.83350 -0.47340
 -0.47340 0.00381

Tau(beta) (as correlations)
 INTRCPT1/INTRCPT2,B00 1.000 -0.853
 MAT_I/INTRCPT2,B30 -0.853 1.000

Random level-2 coefficient	Reliability estimate
INTRCPT1/INTRCPT2, B00	0.569
MAT_I/INTRCPT2, B30	0.637

The value of the likelihood function at iteration 2516 = -1.681677E+005
 The outcome variable is MAT_F

Final estimation of fixed effects
 (with robust standard errors)

Fixed Effect	Standard Coefficient	Error	Approx. T-ratio	d.f.	P-value
For INTRCPT1, P0					
For INTRCPT2, B00					
INTRCPT3, G000	61.854773	3.944732	15.680	228	0.000
FEA01, G001	27.923909	3.790146	7.368	228	0.000
For T2 slope, P1					
For INTRCPT2, B10					
INTRCPT3, G100	-67.898859	4.341013	-15.641	34547	0.000
FEA01, G101	35.302276	9.014035	3.916	34547	0.000
For T4 slope, P2					
For INTRCPT2, B20					
INTRCPT3, G200	22.771229	3.312672	6.874	34547	0.000
FEA01, G201	14.861215	3.015723	4.928	34547	0.000
For MAT_I slope, P3					
For INTRCPT2, B30					

INTRCPT3, G300	0.578427	0.020938	27.626	228	0.000
FEA01, G301	-0.079935	0.023608	-3.386	228	0.001
For FA06_NSE, B31					
INTRCPT3, G310	0.014401	0.005035	2.860	12151	0.005
For T2XMAT_I slope, P4					
For INTRCPT2, B40					
INTRCPT3, G400	0.547525	0.031142	17.582	34547	0.000
FEA01, G401	-0.140216	0.052315	-2.680	34547	0.008
For T3XMAT_I slope, P5					
For INTRCPT2, B50					
INTRCPT3, G500	0.213092	0.009862	21.607	34547	0.000
For T4XMAT_I slope, P6					
For INTRCPT2, B60					
INTRCPT3, G600	0.153125	0.021874	7.000	34547	0.000
For DEFAS slope, P7					
For INTRCPT2, B70					
INTRCPT3, G700	16.415656	3.063738	5.358	34547	0.000

Final estimation of level-1 and level-2 variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, R0	2.45367	6.02049	8977	8251.73473	>.500
MAT_I slope, R3	0.01409	0.00020	8976	9402.33337	0.001
level-1, E	31.06894	965.27893			

Note: The chi-square statistics reported above are based on only 9207 of 12153 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Final estimation of level-3 variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1/INTRCPT2, U00	8.99075	80.83350	228	637.04185	0.000
MAT_I/INTRCPT2, U30	0.06171	0.00381	228	768.52313	0.000

Statistics for current covariance components model

Deviance = 336335.420552
Number of estimated parameters = 21

Model comparison test

Chi-square statistic = 19.26858
Number of degrees of freedom = 8
P-value = 0.014