



UNIVERSIDADE FEDERAL DE SANTA CATARINA
CAMPUS FLORIANÓPOLIS
PROGRAMA DE PÓS-GRADUAÇÃO EM MÉTODOS E GESTÃO EM AVALIAÇÃO

Adrián Silveira Aberastury

**FATORES ASSOCIADOS AO DESEMPENHO ESCOLAR NO URUGUAI:
ANÁLISE VIA MODELOS LINEARES HIERÁRQUICOS COM VALORES
PLAUSÍVEIS**

FLORIANÓPOLIS

2019

Adrián Silveira Aberastury

**FATORES ASSOCIADOS AO DESEMPENHO ESCOLAR NO URUGUAI:
ANÁLISE VIA MODELOS LINEARES HIERÁRQUICOS COM VALORES
PLAUSÍVEIS**

Dissertação submetida ao Programa de Pós-Graduação em Métodos e Gestão em Avaliação - PPGMGA da Universidade Federal de Santa Catarina para a obtenção do título de Mestre em Métodos e Gestão em Avaliação.
Orientador: Prof. Dr. Pedro Alberto Barbeta.

Florianópolis
2019

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Silveira Aberastury, Adrián

Fatores associados ao desempenho escolar no Uruguai:
Análise via modelos lineares hierárquicos com valores
plausíveis / Adrián Silveira Aberastury ; orientador, Pedro
Alberto Barbetta, 2019.

151 p.

Dissertação (mestrado profissional) - Universidade
Federal de Santa Catarina, Centro Tecnológico, Programa de
Pós-Graduação em Métodos e Gestão em Avaliação, Florianópolis,
2019.

Inclui referências.

1. Métodos e Gestão em Avaliação. 2. Modelos Lineares
Hierárquicos. 3. Valores plausíveis. 4. Fatores associados à
aprendizagem. 5. Avaliação Internacional em Larga Escala.
I. Barbetta, Pedro Alberto . II. Universidade Federal de
Santa Catarina. Programa de Pós-Graduação em Métodos e
Gestão em Avaliação. III. Título.

Adrián Silveira Aberastury

**FATORES ASSOCIADOS AO DESEMPENHO ESCOLAR NO URUGUAI:
ANÁLISE VIA MODELOS LINEARES HIERÁRQUICOS COM VALORES
PLAUSÍVEIS**

O presente trabalho em nível de mestrado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Prof. Pedro Barbeta, Dr.

Universidade Federal de Santa Catarina

Prof. Luiz Ricardo Nakamura, Dr.

Universidade Federal de Santa Catarina

Prof. Andrés Peri Hada, Dr.

División de Investigación, Evaluación y Estadística, ANEP-CODICEN

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi julgado adequado para obtenção do título de mestre em Métodos e Gestão em Avaliação.

Prof. Dr. Marcelo Menezes Reis

Coordenador do Programa

Prof. Dr. Pedro Alberto Barbeta

Orientador

Florianópolis, 13 de setembro de 2019.

Dedicado a Mica e Fausti.

AGRADECIMENTOS

Reuni momentos e experiências inesquecível desde que esta ideia se materializou até hoje quando chega seu fim. Neste tempo tem havido muitas pessoas a quem me sinto profundamente grato e certamente é injusto mencionar apenas alguns.

Agradeço em primeiro lugar a Micaela, minha companheira de tantos sonhos. Sem sua paciência e carinho incondicional, este projeto nunca teria se tornado realidade.

Foi igualmente importante para mim o apoio da minha família, meu pai, Leo, meus amigos e colegas de trabalho, onde sempre encontrei o impulso para avançar mesmo nos momentos mais difíceis.

Cada palavra que escrevi foi exaustivamente revisada por meu orientador, o Professor Dr. Pedro Barbeta, que sempre tentou, com interminável paciência e generosidade, que meu trabalho tivesse a melhor qualidade possível. Sem seu constante acompanhamento do progresso do meu trabalho, eu não teria chegado a este lugar. Obrigado Pedro.

Andrés Peri merece um agradecimento especial, pois além de aceitar o convite para integrar a banca examinadora, foi uma das primeiras pessoas que me incentivou a continuar minha formação e quem gerou todas as condições necessárias para cumprir com este programa acadêmico.

Sou muito grato aos meus amigos Thomas Evans e Santiago Cardozo por sua dedicação para ler e comentar versões preliminares do manuscrito desta dissertação.

Aos colegas de turma no Mestrado, que generosamente me ajudaram a tornar as coisas muito mais simples para um estudante estrangeiro.

Quero agradecer o apoio da Universidad de la República que financiou uma parte importante desta formação através do “Programa de apoio à mobilidade de professores que estudam cursos de pós-graduação na região”.

Finalmente, à Universidade Federal de Santa Catarina, ao corpo docente, à Coordenação e Secretaria do PPGMGA, meus sinceros agradecimentos pela oportunidade de formação humana e acadêmica nesta instituição.

—Uma última pergunta— disse o jornalista e continuou. Suponha que esta entrevista tenha sido lida por nossos descendentes, como em algumas centenas de anos. O que você acha que valeria a pena contar àquela geração sobre a vida que você viveu e as lições que aprendeu nela?

—Gostaria de me referir a duas questões: uma intelectual e outra moral. O intelectual é que quando você está estudando algum assunto ou considerando qualquer filosofia, pergunte a si mesmo: Quais são os fatos? Qual é a verdade que os fatos revelam? Nunca se deixe desviar, seja por causa do que você quer acreditar ou por causa do que você acha que lhe traria benefícios se fosse acreditado. Olhe apenas, sem dúvida, quais são os fatos.

—E a moral? — respondeu o jornalista.

—A moral é ainda mais simples. O amor é sábio, o ódio é tolo. Nesse mundo que está cada vez mais interconectado, nós temos que aprender a tolerar uns aos outros. Temos que aprender aceitar o fato de que algumas pessoas dizem coisas que não gostamos. Nós precisamos aprender um novo tipo de caridade e tolerância que é absolutamente vital para a continuação da vida humana neste planeta.

RESUMO

O presente estudo tem como objetivo principal identificar os fatores que influenciam o aprendizado dos estudantes ao final do ensino fundamental no Uruguai. Foram utilizados os dados do Terceiro Estudo Regional Comparativo e Explicativo (TERCE) da UNESCO para identificar os fatores associados ao desempenho escolar em Leitura, Matemática e Ciências Naturais. O TERCE é o primeiro estudo em larga escala na América Latina, no qual as distribuições da proficiência da população foram estimadas com a metodologia dos valores plausíveis. Foram analisadas as relações dos indicadores contextuais com a proficiência dos estudantes por meio de modelos lineares hierárquicos (MLH) que consideram os efeitos das características das escolas e seus entornos, dos professores, dos estudantes e de suas famílias sobre as proficiências medidas em valores plausíveis. O trabalho mostra empiricamente a importância do uso adequado do set de valores plausíveis para estimativas não viesadas das médias e dos erros padrões da proficiência dos estudantes nos modelos MLH. A análise confirma a importância de alguns fatores individuais associados à trajetória escolar dos alunos e suas características como sexo, origem familiar, as expectativas educacionais dos pais e a defasagem idade/série. Em relação aos fatores de nível escolar, o resultado mais notável é que, em comparação com as características do ambiente e contexto, os fatores organizacionais representaram uma parte menor da variância total explicada. Verificou-se que, em termos gerais, a capacidade das escolas de influenciar a aprendizagem é consideravelmente limitada em comparação com o peso das condições socioculturais dos alunos e do ambiente da própria escola.

Palavras-chave: Modelos Lineares Hierárquicos. Valores plausíveis. Fatores associados à aprendizagem. Avaliação Internacional em Larga Escala.

ABSTRACT

The main objective of this study is to identify the factors that influence students' achievement at the end of elementary school in Uruguay. Data from UNESCO's Third Regional Comparative and Explanatory Study (TERCE) were used to identify the factors associated with school performance in Reading, Mathematics and Natural Sciences. TERCE is the first International Large-Scale Assessments (ILSA) focus on student's achievement in Latin America, in which population proficiency distributions were estimated using the plausible values methodology. The relationships between contextual indicators and student proficiency were analyzed through hierarchical linear models (MLH) that consider the effects of the characteristics of schools, teachers, students and their families on proficiency measured in plausible values. The dissertation shows empirically the importance of the appropriate use of the set of plausible values for non-biased estimates of the means and standard errors of student proficiency when MLH is used. The analysis confirms the importance of student level factors associated to characteristics such as gender, family background, parents' educational expectations and grade retention. Regarding school-level effects, the institutional factors represented a smaller part of the total variance explained. Overall, the ability of schools to influence learning is considerably limited even after taking account of the school composition and its contextual effects.

Keywords: Hierarchical Linear Model. Plausible values. Factors associated with academic achievement. International Large-Scale Assessments.

LISTA DE QUADROS

Quadro 4.1. Questionários de fatores associados.....	59
Quadro 4.2. Proporção de variância explicada em relação ao Modelo Nulo.	69
Quadro 5.1. Variáveis e questões selecionadas no cálculo do NSE.	79
Quadro 5.2. Variáveis e questões selecionadas no cálculo do IVB.....	82
Quadro 5.3. Variáveis e questões selecionadas no cálculo do ISAE.	84
Quadro 5.4. Variáveis e questões selecionadas no cálculo do IPPE.	86
Quadro 5.5. Variáveis e questões selecionadas no cálculo do IDGA.....	88
Quadro 5.6. Variáveis e questões selecionadas no cálculo do IBUL.	89
Quadro 5.7. Variáveis e questões selecionadas no cálculo do INFRA.....	90
Quadro 5.8. Variáveis e questões selecionadas no cálculo do LID.....	92
Quadro 5.9. Variáveis e questões selecionadas no cálculo do CLI.....	94
Quadro 6.1. Descrição dos indicadores, níveis e efeito esperado na proficiência (sem distinção de disciplina) no modelo empírico. Nível hierárquico ESTUDANTE.	99
Quadro 6.2. Descrição dos indicadores, níveis e efeito esperado na proficiência (sem distinção de disciplina) no modelo empírico. Nível hierárquico ESCOLA.....	100
Quadro 6.3. Modelos e variáveis composicionais e contextuais para o estudo da de decomposição da variância.	104
Quadro 6.4. Ajuste dos Modelos Finais (M6) para as diferentes áreas avaliadas. Resumo conceitual da intensidade e sentido da associação.	121

LISTA DE TABELAS

Tabela 4.1. Tamanho da amostra final de estudantes por prova e questionário na 6 ^o série de TERCE.	60
Tabela 5.1. Estimativas dos parâmetros dos itens e erros padrões associados ao NSE.	81
Tabela 5.2. Estimativas dos parâmetros dos itens e erros padrões associados ao IVB.	83
Tabela 5.3. Estimativas dos parâmetros dos itens e erros padrões associados ao ISAE.	85
Tabela 5.4. Estimativas dos parâmetros dos itens e erros padrões associados ao IPPE.	87
Tabela 5.5. Estimativas dos parâmetros dos itens e erros padrões associados ao IDGA.	89
Tabela 5.6. Estimativas dos parâmetros dos itens e erros padrões associados ao IBUL.	90
Tabela 5.7. Estimativas dos parâmetros dos itens e erros padrões associados ao INFRA.	91
Tabela 5.8. Estimativas dos parâmetros dos itens e erros padrões associados ao LID.	93
Tabela 5.9. Estimativas dos parâmetros dos itens e erros padrões associados ao CLI.	95
Tabela 6.1. Modelo incondicional (Modelo Nulo) e partição da variância.	102
Tabela 6.2. Decomposição da variância, controladas pela composição socioeconômica e cultural da escola e fatores contextuais.	106
Tabela 6.3. Ajuste dos modelos hierárquicos em Leitura. Estimativas para os coeficientes, erros-padrão e estatística T.	112
Tabela 6.4. Ajuste dos modelos hierárquicos em Matemática. Estimativas para os coeficientes, erros-padrão e estatístico T.	115
Tabela 6.5. Ajuste dos modelos hierárquicos em Ciências Naturais. Estimativas para os coeficientes, erros-padrão e estatístico T.	119
Tabela 6.6. Estimativas dos modelos MLH (coeficientes, erros padrões e variâncias) usando diferentes estratégias dos valores plausíveis para as três áreas avaliadas	130

LISTA DE ABREVIATURAS E SIGLAS

ANEP	Administración Nacional de Educación Pública
A.PR.EN.D.E.R.	Atención Prioritaria en Entornos con Dificultades Estructurales Relativas
CCI	Curva Característica do Item
CCI	Coeficiente de Correlação Intraclasse
CEIBAL	Programa de Conectividad Educativa de Informática Básica para el Aprendizaje en Línea
CEIP	Consejo de Educación Inicial y Primaria
CEPAL	Comisión Económica para América Latina y el Caribe
CES	Consejo de Educación Secundaria
CETP	Consejo de Educación Técnico- Profesional
CFE	Consejo de Formación en Educación
CII	Curva de Informação do Item
CLI	Índice de Clima Escolar
CODICEN	Consejo Directivo Central
DIEE	División de Investigación, Evaluación y Estadística
EAP	Estimação pela média a posteriori
EMS	Ensino Médio Superior
ENEM	Exame Nacional do Ensino Médio
ERCE	Estudo Regional Comparativo e Explicativo
IBUL	Índice de bullying entre colegas na escola
IDGA	Índice de Disciplina e Gestão da sala de Aula
IEA	International Association for the Evaluation of Educational Achievement
INEED	Instituto Nacional de Evaluación Educativa
INFRA	Índice de Infraestrutura do local escolar
IPPE	Índice de Práticas Pedagógicas de Ensino
ISAE	Índice de Supervisão de Estudos e Atividades Educativas na Casa
IVB	Índice de Violência no Bairro
LLECE	Laboratório Latino-americano de Avaliação da Qualidade da Educação

LID	Índice de Liderança do Diretor
MEC	Ministério da Educação e Cultura
MLH	Modelo Linear Hierárquico
MML	Máxima Verossimilhança Marginal
NAEP	National Assessment of Education Progress
NSE	Índice do Nível Socioeconômico e Cultural do aluno e suas famílias
OCDE	Organização para Cooperação e Desenvolvimento Econômico
ONE	Operativo Nacional de Evaluación de la Calidad
PERCE	Primeiro Estudo Regional Comparativo e Explicativo
PIB	Produto Interno Bruto
PIRLS	Progress in International Reading Literacy Study
PISA	Programme for International Student Assessment
SAEB	Sistema de Avaliação da Educação Básica
SEA	Sistema de Evaluación de Aprendizajes
SERCE	Segundo Estudo Regional Comparativo e Explicativo
TC	Escuelas de Tiempo Completo
TERCE	Terceiro Estudo Regional Comparativo e Explicativo
TIMSS	Trends in International Mathematics and Science Survey
TRI	Teoria de Resposta ao Item
UNESCO	United Nations Educational, Scientific and Cultural Organization
VP	Valores Plausíveis
WLE	Warm's Mean Weighted Likelihood Estimates

SUMÁRIO

1.	INTRODUÇÃO	15
1.1	PROBLEMA DE PESQUISA.....	16
1.2	OBJETIVOS.....	19
1.3	JUSTIFICATIVA E RELEVÂNCIA.....	20
1.4	DELIMITAÇÃO.....	21
1.5	ESTRUTURA DO TRABALHO.....	22
2.	PANORAMA DA EDUCAÇÃO NO URUGUAI	24
2.1	EVOLUÇÃO HISTÓRICA.....	24
2.2	ESTRUTURA INSTITUCIONAL DO SISTEMA EDUCACIONAL	26
2.3	TENDÊNCIAS EM ACESSO E COBERTURA	28
2.4	POLÍTICAS EDUCACIONAIS RECENTES.....	29
2.5	O SISTEMA DE AVALIAÇÃO DA EDUCAÇÃO NACIONAL.....	32
3.	FUNDAMENTAÇÃO TEÓRICA	34
3.1	. FATORES INDIVIDUAIS E FAMILIARES	35
3.1.1.	Fatores Socioeconômicos e Culturais	36
3.1.2.	O contexto sociocultural das escolas	38
3.1.3.	Gênero	39
3.1.4.	Trabalho infantil	40
3.1.5.	Trajetória acadêmica dos estudantes	41
3.1.5.1.	Repetência: A distorção idade/série.....	41
3.1.5.2.	Educação infantil e Pré-escola.....	44
3.1.6.	Tempo de exposição ao ensino	45
3.1.6.1.	Absenteísmo escolar.....	45
3.1.6.2.	Dedicação ao estudo fora da escola	46
3.1.7.	Expectativas familiares e aspirações educacionais	47

3.1.8.	Envolvimento dos pais na vida escolar	48
3.2.	FATORES ESCOLARES.....	49
3.2.1.	Práticas e atitudes dos professores no processo de aprendizagem	52
3.2.2.	Violência No Entorno Escolar	53
3.2.3.	Clima Escolar	54
4.	MATERIAIS E MÉTODOS	55
4.1.	CONTEXTO DO ESTUDO: TERCE	55
4.1.1.	O Laboratório Latino-Americano da Avaliação da Qualidade da Educação	55
4.1.2.	Os Estudos da UNESCO: PERCE, SERCE e TERCE	56
4.1.3.	Objetivos e Características do TERCE	57
4.1.3.1.	Provas e questionários do TERCE	58
4.1.3.2.	Dados e variáveis do estudo TERCE	59
4.2.	TEORIA DE RESPOSTA AO ITEM.....	61
4.3.	MODELO LINEAR HIERÁRQUICO	62
4.3.1.	Modelo Nulo	67
4.4.	USO DOS VALORES PLAUSÍVEIS EM AVALIAÇÕES EM LARGA ESCALA ⁶⁹	
4.4.1.	Definição dos valores plausíveis	70
4.4.2.	O Método de Imputação Via Valores Plausíveis	72
4.4.3.	A Metodologia dos Valores Plausíveis no TERCE	74
4.5.	O PACOTE 'BIFIESURVEY' DO R.....	76
5.	CONSTRUÇÃO DE INDICADORES	78
5.1.	ÍNDICE DO NÍVEL SOCIOECONÔMICO E CULTURAL DO ALUNO E SUA FAMÍLIA.....	78
5.2.	ÍNDICE DE VIOLÊNCIA NO BAIRRO	82
5.3.	ÍNDICE DE SUPERVISÃO DE ESTUDOS E ATIVIDADES EDUCATIVAS NA CASA	83
5.4.	ÍNDICE DE PRÁTICAS PEDAGÓGICAS DE ENSINO	85

5.5.	ÍNDICE DE DISCIPLINA E GESTÃO DA SALA DE AULA.....	88
5.6.	ÍNDICE DE BULLYING ENTRE COLEGAS.....	89
5.7.	ÍNDICE DE INFRAESTRUTURA DO LOCAL ESCOLAR.....	90
5.8.	ÍNDICE DE LIDERANÇA DO DIRETOR	91
5.9.	ÍNDICE DE CLIMA ESCOLAR	94
6.	RESULTADOS	97
6.1.	DECOMPOSIÇÃO DA VARIÂNCIA NOS NÍVEIS HIERÁRQUICOS: EFEITO ESCOLA.....	101
6.2.	AJUSTE DE MODELOS LINEARES HIERÁRQUICOS COM VARIÁVEIS EXPLICATIVAS NOS NÍVEIS INDIVIDUAIS E ESCOLARES	107
6.2.1.	Etapas do ajuste do modelo hierárquico linear	107
6.2.2.	Ajuste de modelos em Leitura	108
6.2.3.	Ajuste de modelos em Matemática.....	113
6.2.4.	Ajuste de modelos em Ciências Naturais.....	116
6.3.	DISCUSSÃO DOS RESULTADOS DOS MODELOS FINAIS.....	120
6.4.	TESTE EMPÍRICO SOBRE O USO DE VALORES PLAUSÍVEIS EM MODELOS LINEARES HIERÁRQUICOS.....	126
6.4.1.	Componentes do erro padrão das estimativas de proficiência nas avaliações em larga escala	126
6.4.2.	Comparação de estimativas via modelos lineares hierárquicos com valores plausíveis.....	128
7.	CONSIDERAÇÕES FINAIS	132
	REFERÊNCIAS	138

1. INTRODUÇÃO

Embora a desigualdade tenha diminuído na América Latina desde 2002, ainda é uma das regiões mais desiguais do mundo. Metade dos 20 países mais desiguais do mundo estão na América Latina (DURYEA e ROBLES, 2016). A desigualdade na distribuição de recursos na população está relacionada ao acesso desigual das famílias a bens e serviços de qualidade, como saúde e educação.

Neste contexto, um dos principais desafios que qualquer sistema educacional enfrenta, em particular na América Latina, é melhorar a aprendizagem dos estudantes e, simultaneamente, enfrentar o problema da desigualdade, em termos de reduzir a distância entre os estudantes com alto e baixo desempenho acadêmico.

A qualidade da aprendizagem é sempre o resultado de uma combinação de fatores escolares e sociais relacionados. Além das condições sociais e familiares, existem condições de aprendizagem fundamentais, como interesse ou motivação para aprender, que são, em suma, produto da boa prática escolar. Somam-se a isso outros fatores sistêmicos, como equipamentos, condições de trabalho para professores, projetos curriculares atuais e relevantes, professores bem treinados, estratégias e métodos didáticos adequados às características da população estudantil, etc.

As políticas educacionais e, em especial, a melhoria da qualidade da aprendizagem, são hoje uma questão de agenda pública no Uruguai. Apesar dos importantes avanços em relação ao acesso, a qualidade da educação ainda é bastante precária, o que pode ser constatado a partir dos resultados alcançados pelos estudantes uruguaios em avaliações em larga escala internacionais como o PISA¹. Embora o país esteja entre os de melhores desempenhos na América Latina, em estudos comparativos internacionais, ainda fica a uma distância considerável da média dos países mais desenvolvidos.

¹ Programme for International Student Assessment.

Além disso, esse desempenho educativo inclui, infelizmente, uma das características mais notórias dos sistemas educacionais latino-americanos: o Uruguai é um dos países com maiores desigualdades de aprendizado, ou seja, apresenta uma alta dispersão de resultados do aprendizado em avaliações em larga escala e, o é que ainda pior, com alta inequidade, produto da forte incidência do nível socioeconômico no desempenho educacional.

Este trabalho procura mostrar o efeito deste e de outros fatores no desempenho educacional dos estudantes. Assim, a dissertação contribui para a área de pesquisa no âmbito da avaliação educacional em larga escala. O trabalho pretende identificar os fatores que influenciam o aprendizado dos estudantes ao final do ensino fundamental no Uruguai.

O foco está no estudo da desigualdade no desempenho escolar através dos dados do Terceiro Estudo Regional Comparativo e Explicativo (TERCE). O TERCE é coordenado pelo Escritório Regional de Educação da UNESCO, e participaram 15 países da América Latina. Pela primeira vez, em um estudo comparativo regional em larga escala, para se estimar as distribuições de proficiência da população e subpopulações, foi usada a metodologia de Valores Plausíveis, descrita por Mislevy, Johnson e Muraki (1992).

O foco de interesse é analisar os fatores associados às proficiências dos estudantes, em três áreas cognitivas: Leitura, Matemática e Ciências Naturais, por meio de modelos lineares hierárquicos (MLH) considerando os efeitos das características das escolas e seus entornos, dos professores, dos estudantes e de suas famílias.

1.1 PROBLEMA DE PESQUISA

Até pouco tempo atrás, a reflexão acadêmica sobre as dimensões que afetavam a educação considerava o resultado do processo educativo através da expansão da matrícula escolar ou da noção de desempenho escolar, expressada como sucesso na aprovação das séries escolares ou, o seu fracasso, como repetência e evasão. Essas discussões foram, portanto, focadas em um problema de acesso e fluxo dentro do sistema.

Mais recentemente, os trabalhos foram reorientados para o estudo dos efeitos dos fatores familiares e escolares sobre o desempenho dos estudantes em avaliações padronizadas, possibilitando uma discussão mais ampla sobre a qualidade da educação (SOARES et al., 2001; BARBOSA; FERNANDES, 2001; ALBERNAZ et al., 2002).

O problema dos fatores associados à aprendizagem escolar é complexo devido à multiplicidade de conceitos e variáveis que devem ser levados em conta, bem como aos diferentes níveis de análise envolvidos.

Na medida que cresce o interesse nos processos educacionais e seu impacto sobre o futuro da vida dos indivíduos, as avaliações em larga escala do sistema escolar expandiram-se consideravelmente (BLOSSFELD, SCHNEIDER e DOLL, 2009; VON DAVIER, GONZALEZ, KIRSCH e YAMAMOTO, 2013). Essas avaliações fornecem informações para uma variedade de partes interessadas, tais como acadêmicos e investigadores, decisores de políticas educacionais, professores e comunidade educativa (HANUSHEK, 2011, 2013; RITZEN, 2013).

Além disso, como indica Soares (2013), as recomendações de políticas públicas devem ser apoiadas por evidências sólidas. Assim, em matéria de políticas educacionais, podem ser derivadas dos resultados obtidos por avaliações educacionais.

A avaliação de estudantes em larga escala tem sido muito popular, especialmente após os anos 70 (HOEPFL, 2000), e a quantidade de testes tem aumentado desde então (SHAVELSON, SCHNEIDER e SHULMAN, 2007).

Para Klein e Fontanive,

“Os objetivos da avaliação em larga escala do sistema escolar são os de informar o que populações e subpopulações de estudantes em diferentes séries sabem e são capazes de fazer em um determinado momento, e acompanhar sua evolução ao longo dos anos. Não é seu objetivo fornecer informações sobre estudantes ou escolas individuais” (KLEIN e FONTANIVE, 1995)

Os estudos em larga escala, geralmente, avaliam uma variedade de domínios cognitivos, e variáveis contextuais (o que é chamado de background) das

escolas, professores, estudantes e suas famílias, permitindo monitorar o desempenho dos estudantes e os fatores que estão associados a seu desempenho educacional. Portanto, uma ampla gama de questões pode ser abordada, relacionadas as proficiências dos estudantes e os fatores que as influenciam.

Atualmente, os dados das avaliações de larga escala têm sido amplamente utilizados em estudos educacionais, sociais e econômicos; e o desempenho dos estudantes é considerado como um dos principais indicadores para a qualidade do sistema educacional (ALVES e CANDIDO, 2017)

Assumindo que tais avaliações representam bons indicadores da qualidade do ensino ofertado, considera-se importante identificar quais são os determinantes do rendimento escolar. Soares (2004) afirma que os fatores que determinam o desempenho cognitivo pertencem a três grandes categorias: aqueles associados à estrutura escolar, os relacionados à família e os relacionados ao próprio estudante. Nesse estudo é apresentado um modelo conceitual que mostra como os fatores intraescolares e extraescolares estão associados ao desempenho cognitivo dos estudantes.

Os modelos multiníveis, também conhecidos como modelos hierárquicos ou modelos de efeitos aleatórios, são indicados para analisar dados que apresentem uma estrutura hierárquica. Na maioria das avaliações educacionais em larga escala essa hierarquização é facilmente observada, uma vez que os estudantes são agrupados em turmas e as turmas, por sua vez, são organizadas em escolas.

Os modelos multiníveis têm se destacado como metodologia empírica para detecção de fatores ligados ao desempenho (FERRÃO et al., 2001; SOARES, CÉSAR e MAMBRINI, 1997; ALBERNAZ et al., 2002). As vantagens de considerar uma estrutura hierárquica, presente nos dados de avaliações educacionais, é permitir a avaliação dos efeitos do indivíduo e do grupo ao qual pertence, possibilitando a identificação do efeito das escolas sobre o desempenho dos estudantes.

Nesta dissertação, serão aplicados modelos multiníveis para o estudo de fatores associados, utilizando como variável de desempenho o set de valores plausíveis de cada estudante, disponíveis para cada área avaliada.

Para Laukaityte e Wiberg (2017) reportar as proficiências dos estudantes através de valores plausíveis é a maneira mais eficiente de estimar parâmetros,

como a média, a variância e a correlação. Através dos valores plausíveis, a estrutura hierárquica dos dados e a complexidade das amostras são considerados, portanto, atualmente, tornou-se a norma em avaliações internacionais em larga escala. Com esse impulso, gradualmente, na região, as avaliações educacionais começaram a incorporar essa metodologia. Nesta dissertação, vamos aproveitar esta oportunidade para usar intensivamente dessa metodologia, que em geral, não é tão popular na pesquisa educacional em América Latina.

1.2 OBJETIVOS

A presente dissertação tem dois objetivos fundamentais e complementares. Primeiro, mostrar evidência sobre os principais fatores que estão associados ao desempenho dos estudantes uruguaios nas proficiências em Leitura, Matemática e Ciências Naturais. E em segundo lugar, realizar essa análise com as proficiências estimadas através da metodologia dos valores plausíveis.

O Uruguai participa do PISA (que se aplica aos 15 anos de idade), mas nunca participou do TIMSS ou do PIRLS². Para Uruguai, TERCE é a primeira avaliação em larga escala no ensino fundamental onde a proficiência de cada estudante é expressa em termos de uma amostra de cinco valores plausíveis. Portanto, o TERCE, significa a primeira oportunidade de fazer essas análises com valores plausíveis para o ensino fundamental.

Assim, baseados na estrutura hierárquica desses dados, combina-se a análise de fatores associados via MLH e usar os valores plausíveis das proficiências como variáveis dependentes.

Os objetivos específicos são:

- a) Criar medidas e indicadores que podem estar associados ao desempenho cognitivo dos estudantes, com base em itens dos questionários de contextualização do TERCE.

² *Trends in International Mathematics and Science Study* (TIMSS) e o *Progress in International Reading Literacy Study* (PIRLS) foram desenvolvidos pela Associação Internacional para Avaliação do Desempenho Educacional (IEA).

- b) Identificar possíveis fatores que influenciam a aprendizagem.
- c) Aplicar modelos hierárquicos lineares com valores plausíveis da proficiência dos estudantes.
- d) Demonstrar a importância de utilizar adequadamente o set de valores plausíveis nas estimativas das médias e dos erros padrões da proficiência dos estudantes.
- e) Discutir sobre os fatores associados ao desempenho escolar.

1.3 JUSTIFICATIVA E RELEVÂNCIA

No Uruguai, assim como no Brasil e no resto da América Latina, a pesquisa educacional que utiliza métodos quantitativos foi historicamente escassa (MORDUCHOWICZ, 2004). Esse processo se expressa na pouca receptividade nas faculdades de educação (BROOKE, 2005) e na baixa apropriação pelos tomadores de decisão da política educacional (IOSCHPE, 2004).

No entanto, particularmente no Brasil, há um número crescente de estudos que usam métodos quantitativos para identificar fatores associados ao desempenho dos estudantes (WAISELFISZ, 2000; BARROS et al., 2001; ALBERNAZ et al., 2002; FELICIO e FERNANDES, 2005; MENEZES-FILHO, 2006; MENEZES-FILHO e PAZELLO, 2007; FELICIO, 2007; ALVES E SOARES, 2007; MACHADO et al., 2008; CASTRO, 2009; OLIVEIRA, 2010).

A identificação de fatores associados à aprendizagem e a determinação de seu grau de influência nas habilidades dos estudantes tem sido objeto de estudo, tanto para fins acadêmicos, como para desenvolvimento de políticas e ações de gestão no sistema educacional. Assim, considera-se que o aprofundamento da identificação dos fatores associados à aprendizagem seja uma das formas de promover a melhoria da qualidade do ensino e a otimização dos recursos utilizados na educação.

Em segundo lugar, no caso do Uruguai, uma análise aprofundada dos fatores associados à aprendizagem baseada nos dados do TERCE ainda não foi realizada com exaustividade. O Relatório Regional do TERCE analisa a associação de cada fator (ou variável) ao desempenho acadêmico, mas não faz a análise global e em profundidade de todos os fatores intervenientes.

“Modelos de regressão multinível foram elaborados nos níveis de escola e estudante em cada país, série e disciplina avaliada. A robustez e magnitude da associação de cada fator ou variável ao desempenho acadêmico foi avaliada antes e depois de considerar a influência no resultado de cada país do índice socioeconômico dos estudantes, tanto no nível individual quanto na média das escolas”. (UNESCO, 2016: 17)

Desta forma, este estudo tem como objetivo completar a visão geral da descrição dos fatores associados à aprendizagem.

Em terceiro lugar, como mostram Laukaityte e Wiberg (2017), reportar as habilidades dos estudantes através de valores plausíveis é a maneira mais eficiente de estimar parâmetros, como a média, a variância e a correlação, para analisar as habilidades dos estudantes. Através dos valores plausíveis, a estrutura hierárquica dos dados e a complexidade dos desenhos de amostragem são considerados. E por esse motivo, são usados atualmente em avaliações internacionais de larga escala, como TIMSS, PIRLS ou PISA. Os estudos da UNESCO, como o TERCE, também incorporaram essa metodologia recentemente e esta é a base para a análise desta dissertação.

Por fim, esta dissertação propõe a construção de um conjunto de índices que sintetizam fatores teoricamente relevantes relacionados à aprendizagem. Nesse sentido, é frequente que a observação de fatores associados à aprendizagem seja realizada por meio da análise isolada de itens em questionários contextuais. A construção de índices baseados em conjuntos de itens é importante, porque eles proporcionam melhor diferenciação entre indivíduos e entre escolas, podendo medir com maior precisão a influência dos construtos de interesse.

1.4 DELIMITAÇÃO

Embora no Uruguai tenha havido avaliações nacionais de desempenho no ensino fundamental desde 1996, nenhuma estimou as variáveis de proficiência por meio da metodologia de valores plausíveis. O TERCE é o primeiro estudo em educação básica que permite fazê-lo.

O estudo regional foi realizado pelo Laboratório Latino-americano de Avaliação da Qualidade da Educação (LLECE em espanhol) da UNESCO. O TERCE fornece informações de alta qualidade que permitem a comparação entre os países da América Latina (UNESCO, 2016).

A dissertação limita-se na análise da proficiência em Leitura, Matemática e Ciências Naturais de uma amostra representativa nacional de estudantes da 6ª série do ensino fundamental em escolas da rede pública e escolas particulares do Uruguai, que participaram em 2013 da avaliação TERCE.

Apesar do fato de que o TERCE recolhe dados sobre aprendizagens na 3ª e 6ª séries, foi decidido somente trabalhar com a 6ª série devido a três razões substantivas. Em primeiro lugar, supõe-se que o processo educacional é, por definição, cumulativo. Isto é, devido à sua natureza cumulativa, o que um estudante demonstra ser capaz de fazer em um teste padronizado é o resultado (entre outros fatores) do que ele aprendeu ao longo de sua vida escolar. É por isso que é considerado adequado focalizar no nível de ensino mais avançado. Em segundo lugar, a 6ª série significa um momento importante na escolarização, já que, segundo o sistema educacional uruguaio, a escola primária atinge até a 6ª série. A 7ª série corresponde ao Ensino Médio e implica mudanças importantes, pois os estudantes devem, necessariamente, mudar de escola (estabelecimentos de ensino), os conteúdos curriculares são diversificados em 11 disciplinas com pouca coordenação entre professores, etc. Tudo isso faz da 6ª série o momento final de um ciclo de aprendizagem.

No terceiro termo, há uma condição instrumental. O questionário de fatores associados que completam os estudantes em ambas as séries mostra, para a 3ª série, deficiências importantes. Alguns estudantes da 3ª série ainda não sabem ler corretamente ou manter a atenção por longo período, o que ameaça diretamente a confiabilidade desse tipo de instrumento.

1.5 ESTRUTURA DO TRABALHO

O presente trabalho está dividido em sete capítulos. O próximo capítulo descreve as principais tendências do sistema educacional uruguaio nos últimos anos. O capítulo três revisa a literatura que fundamenta o argumento principal da

dissertação, que é, essencialmente, trabalhos que discorrem sobre os possíveis fatores associados ao desempenho acadêmico dos estudantes, em particular, sobre as características familiares, individuais e contextuais das escolas e seus professores.

A descrição do TERCE, a amostragem nacional, o tratamento dos dados, as características dos questionários contextuais, a teoria da resposta ao item aplicada para a geração dos índices e medidas contextuais e, finalmente, a descrição e exposição das vantagens da metodologia de utilização de valores plausíveis são apresentados no capítulo quatro.

No capítulo cinco são detalhados os procedimentos para a geração dos índices que sintetizam os possíveis fatores que, segundo a literatura, influenciam a aprendizagem. E no capítulo seis, são apresentados e discutidos os resultados obtidos a partir da modelagem MLH dos principais fatores associados ao aprendizado. Na seção final do capítulo apresenta-se um teste empírico sobre a importância de utilizar adequadamente o set de valores plausíveis para estimativas não viesadas das médias e dos erros padrões da proficiência dos estudantes. No último capítulo são abordadas as considerações finais e recomendações para trabalhos futuros.

2. PANORAMA DA EDUCAÇÃO NO URUGUAI

Este capítulo descreve as principais tendências do sistema educacional uruguaio nos últimos anos. O capítulo está organizado em cinco grandes seções. No primeiro, são brevemente revisados os principais aspectos da evolução histórica do sistema educacional. A segunda seção descreve a estrutura institucional do sistema educacional, em particular, analisando as características da governança do sistema e seus mecanismos de funcionamento. A terceira apresenta uma breve descrição das principais tendências na oferta educacional, com ênfase nos indicadores de acesso e cobertura, e destaca o papel do Estado na prestação desse serviço público. A quarta seção, enfoca as principais políticas educacionais dos últimos anos. Em quinto, e último lugar, descrevesse o sistema de avaliação educacional e o papel dos diferentes atores na avaliação da aprendizagem.

2.1 EVOLUÇÃO HISTÓRICA

Desde o final do século XIX, o sistema educacional uruguaio, e em particular a escola pública, caracterizou-se como uma ferramenta fundamental para a institucionalização da ordem social que exigia o que foi conhecido como “processo disciplinar”. Esse esforço possibilitou atingir altos níveis de alfabetização desde o início, o que colocou a escola como uma instituição privilegiada na incipiente formação cívica da cidadania. Neste contexto, a educação pública uruguaia foi caracterizada por cumprir múltiplos papéis que iam desde a integração das quotas migrantes e a construção de um patrimônio cultural comum imaginário e nacional, para ser o mecanismo fundamental de mobilidade entre os diferentes segmentos sociais. A educação também foi constituída no mais amplo e universal sistema de benefício social do Estado no país. Esses elementos distinguiram o país como uma das melhores qualidades educacionais da região.

No entanto, esse projeto técnico-político permaneceu em vigor até meados do século XX, quando o modelo de acumulação econômica mostrou sua exaustão, obscurecendo o papel da educação como instrumento efetivo de mobilidade social e

inserção laboral. Quase imperceptivelmente, esses problemas estavam ganhando em magnitude e aumentando a complexidade na agitada sociedade uruguaia entre 1960 e 1980 (ANEP, 2005).

A educação expandiu-se fortemente desde a década de 1950, que até então tinha acesso quase universal ao ensino primário. Mesmo quando apresentou taxas de evasão muito elevadas, sua matrícula cresceu 30% entre 1950 e 1970. No entanto, o aumento mais notável ocorreu no ensino médio, passando de 34.226 estudantes em 1950 para 132.145 em 1970 (CEPAL, 1992).

A constituição de 1967 (ainda vigente) elevou a obrigatoriedade da educação até três anos de ensino médio, ou seja, até completar o que se chama de Ensino Médio Básico (EMB). No entanto, a estrutura educacional mostrou diferenças institucionais radicais entre os níveis primário e secundário para enfrentar as crescentes ondas de massa. Enquanto o desenvolvimento inicial do ensino fundamental determinou que, em meados do século XX, suas bases institucionais já estavam consolidadas e com ampla expansão territorial, o ensino médio ainda era um nível educacional reservado a uma pequena elite. Isso significa que o ensino médio estava ancorado em um paradigma de tipo generalista e propedêutico para os estudos universitários e de alto nível acadêmico, o que correspondia às necessidades, expectativas e interesses de um grupo seletivo da população. (ANEP, 2005).

Após a restauração democrática (1985), o Estado reafirmou seu objetivo de ampliar a obrigatoriedade do ensino médio, até que, finalmente, a Lei Geral de Educação de 2008 estendeu sua obrigatoriedade ao Ensino Médio Superior (EMS). Nesse período, as matrículas no nível primário permaneceram essencialmente estáveis, enquanto as correspondentes ao ciclo médio de educação aumentaram significativamente. Aqui registra-se o acesso de um grupo de pessoas anteriormente excluídas desses níveis educacionais. Essa nova expansão do ensino, portanto, deveu-se principalmente à "extensão vertical", incorporando estudantes de estratos de baixa renda familiar, e "extensão horizontal", incorporando populações anteriormente não usuárias, mulheres e a população do interior do país (CARDOZO, 2010).

Esse aumento no ciclo médio foi acompanhado apenas parcialmente pelo aumento concomitante de recursos didáticos e materiais. Isso agravou a ineficiência do sistema, uma vez que o aumento de matrículas não teve contrapartida semelhante em termos de retenção e graduação desses estudantes. Ou seja, houve progresso nas "áreas mais flexíveis" do desenvolvimento educacional (matrícula), mas muito menos em suas "áreas difíceis" (retenção, fluxo e egresso). É nesse sentido que De Armas e Retamoso indicam que:

"Embora a obrigatoriedade e a universalização da educação não sejam conceitos equivalentes, pode-se supor que, ao definir uma seção da educação como obrigatória, os países estão assumindo um forte compromisso de atingir a meta de todas as pessoas que acessam esse nível e o atingem, ao mesmo tempo, recebem uma educação de qualidade. Nesse sentido, pode-se dizer que o país ainda não alcançou os objetivos propostos há quase quarenta anos em termos de acesso e graduação do ensino médio. Mais preocupante é o fato de que a brechas que existia no início dos anos setenta do século passado entre a realidade e os objetivos declarados no marco legal não foi reduzida significativamente nas últimas quatro décadas, ou que, em qualquer caso, a diminuição dessa brecha foi modesta." (DE ARMAS e RETAMOSO, 2010: 24).

Mais recentemente, no âmbito das reformas educacionais que a América Latina experimentou na década de 90, o Uruguai desenvolveu, entre 1995 e 1999, uma Reforma Educacional orientada para quatro objetivos fundamentais: i) a melhoria da qualidade da educação, ii) progresso em equidade, iii) profissionalização docente e iv) modernização da gestão. Essa reforma foi caracterizada como bem-sucedida, na medida em que alcançou a implementação de grande parte dos objetivos inicialmente propostos (MANCEBO e TORELLO, 2008).

2.2 ESTRUTURA INSTITUCIONAL DO SISTEMA EDUCACIONAL

O Sistema Nacional de Educação é definido como um conjunto de propostas educacionais constituídas pela educação formal e não-formal, assim como na primeira infância. O sistema educacional uruguaio apresenta um desenho institucional diferente dos padrões internacionais.

O poder de formular, implementar e gerenciar políticas de educação no ensino obrigatório é um organismo autônomo, a Administração Nacional de

Educação Pública (ANEP) e não do Ministério da Educação e Cultura (MEC) como é normalmente o padrão internacional. Ao mesmo tempo, é integrado por um conjunto de organizações que operam com certa independência umas das outras e com baixos níveis de articulação. Nesse contexto, o sistema educacional é caracterizado por um alto grau de centralização funcional e geográfica. Isso implica que a maioria das decisões administrativas e pedagógicas são tomadas no nível central (INEED, 2015).

Na estrutura do sistema educacional, a ANEP regula e administra parte da educação na primeira infância e pré-escolar; todo o ensino fundamental e médio; formação de professores e educação técnica e profissional. A ANEP é estruturada em o Conselho Executivo Central (CODICEN), como órgão máximo e controla quatro Conselhos de Educação: o Conselho de Educação Infantil e Primária (CEIP), o Conselho do Ensino Secundário (CES), o Conselho de Educação Técnica e Profissional (CETP) e o Conselho de Formação em Educação (CFE).

O CODICEN está encarregado de definir as diretrizes gerais para todos os níveis e tipos de educação. Além disso, ele é responsável pela elaboração do orçamento educacional, tem autoridade sobre o calendário escolar e decide sobre a criação de novas escolas.

O CODICEN coordena o trabalho dos quatro conselhos e é hierarquicamente superior a eles, mas, por sua vez, os conselhos são considerados autônomos em suas áreas de responsabilidade. Isso faz com que os Conselhos operem com relativa independência um do outro, o que na prática dificulta a coordenação das políticas educacionais e impede uma eficiente articulação entre as políticas educacionais (INEED, 2015).

O Ministério da Educação e Cultura (MEC), por sua vez, estabelece diretrizes gerais para a política educacional. Os poderes do MEC limitam-se à regulação e articulação do sistema educacional e não à formulação e implementação de políticas.

Como pode ser visto, há desafios importantes em termos do redesenho institucional do sistema educacional. Parece necessário abordar uma mudança nos modelos de gestão e governança. Isso implica a necessidade de articular históricas

inércias com reconfigurações de atores e papéis atuais. É essencial para a eficiência institucional as oportunidades de rearticular funções e órgãos de liderança técnica e política do modelo educacional atual, mas também envolve desafios a nível local, no âmbito da promoção de uma maior autonomia das escolas e fortalecer as comunidades educativas.

2.3 TENDÊNCIAS EM ACESSO E COBERTURA

A educação constitui um elemento central para a construção da cidadania, a integração social e a conquista de objetivos individuais e coletivos. Em termos comparativos, o Uruguai tem-se destacado na região por ter alcançado desde o início a quase universalização do ensino fundamental. Esses números não registraram variações significantes nos últimos 25 anos. Além disso, nas últimas décadas, também se estendeu ao nível inicial de cinco anos e quase todas as crianças de quatro anos.

A universalização da cobertura no nível primário resulta da combinação de dois fatores: uma quase completa inscrição das coortes na idade de início e taxas de abandono muito baixas (MANCEBO e TORELLO, 2008).

Por outro lado, no ensino médio, apesar da melhoria nas taxas de matrícula escolar, seus números estão longe de ser satisfatórios, não só em comparação com o ensino primário, mas também mostram discrepância no contexto regional.

O papel do Estado na provisão da educação é, e tem sido, um elemento distintivo do Uruguai em todos os níveis educacionais. Esse padrão tem sido constante nos últimos anos, além de pequenas oscilações e da transferência, mais recentemente, de uma proporção de estudantes para as escolas particulares³. Em 2015, 83,4% dos estudantes nos níveis obrigatórios (educação inicial de 4 e 5 anos, primária e ensino médio) são cobertos pelo setor público. Sua participação é um pouco menor nas idades correspondentes à educação inicial, especialmente em crianças de três anos (nível não obrigatório), que em qualquer caso, cobre mais de

³ Em geral, as escolas particulares não são financiadas diretamente por o nível público, embora que estão isentas de alguns impostos. Tem algumas exceções nos níveis de educação na primeira infância, creche pré-escola. As escolas particulares estão concentradas em Montevideu e nos departamentos vizinhos de Canelones e Maldonado. Eles seguem principalmente o currículo nacional e usam sua autonomia para o fornecimento de atividades extracurriculares.

60% dos estudantes. Por sua vez, e como tem sido o padrão histórico, a presença do estado é predominante na população de baixa renda: o setor público cobre quase todos os estudantes do primeiro quintil de renda e mais de 85% dos que estão no quintil 2.

Pela natureza da sua função, a ANEP é a rede de serviços do estado com maior presença no território, com 2.752 escolas e aproximadamente 675.000 estudantes, incluindo escolas primárias e estabelecimentos de pré-escolar, escolas de ensino médio e institutos de formação de professores (ANEP, 2015b).

A evolução dos últimos 30 anos marca dois períodos diferentes: um de forte aumento na cobertura, com pequenas flutuações, de meados da década de 1980 para os primeiros anos da década de 2000, e outro marcado por uma redução de inscrição, um pouco mais pronunciada até 2007 e mais leve, embora sustentada, a partir de então.

É necessário ressaltar que a redução no número absoluto de estudantes não tem sido o resultado de uma diminuição da cobertura educacional, mas tem três efeitos complementares: i) a redução no tamanho das coortes de estudantes que atingem o sistema de ensino, derivado da diminuição do número de nascimentos. Este efeito tem operado nos últimos anos no nível de educação inicial e primária e começará a ser observado a partir de agora no ensino médio básico; ii) melhorias no fluxo das trajetórias escolares, como consequência da redução das taxas de repetência; iii) transferência leve, mas sustentada, de estudantes para escolas particulares.

Em suma, a matrícula pública em 2015 é distribuída da seguinte forma nos subsistemas da ANEP: o ensino fundamental e inicial representa 50% do total, o ensino médio 33%, o ensino técnico profissional 14% e Formação educacional em 3% (ANEP, 2015b).

2.4 POLÍTICAS EDUCACIONAIS RECENTES

Desde 2005, promoveu-se uma série de mudanças no quadro de referência que sustenta as políticas públicas de educação do país. Assim, podem ser

identificados, numa primeira fase, políticas de fortalecimento da educação pública em três áreas: i) através da perspectiva dos direitos humanos e desenvolvimento econômico e social sustentável e inclusivo, ii) como uma estratégia para a integração social e, iii) como espaço para o exercício ativo da cidadania.

Numa segunda fase, as políticas são aprofundadas em ações de continuidade e fortalecimento das trajetórias escolares com desenvolvimento locais para promover a continuidade educacional e melhoria da qualidade da aprendizagem.

Durante este período, há um claro aumento nos gastos públicos em educação. Antes de 2005, esse gasto não alcançou 3,2% e atingiu 4,5% do PIB até 2015. Isso resultou em melhorias significantes nos salários dos professores e no investimento em infraestrutura (ANEP, 2015b).

A diversificação de formatos escolares no âmbito de um programa curricular comum tem sido uma das estratégias da política educacional orientado a melhorar a qualidade da aprendizagem e reduzir as inequidades. As escolas A.PR.EN.D.E.R.⁴ e as modalidades de extensão do tempo pedagógico: Escolas de Tempo Integral⁵ e, mais recentemente, a criação das Escolas de Tempo Estendido constituem as principais políticas empreendidas neste sentido. Nos últimos dez anos, o ensino primário público duplicou o número de escolas e estudantes com tempo pedagógico estendido.

⁴ O Programa A.PR.EN.D.E.R. (Atenção Prioritária em Ambientes com Dificuldades Estruturais Relativas) é um programa de inclusão educacional que busca garantir o acesso e a permanência de todas as crianças no sistema educacional como a obtenção de uma aprendizagem de qualidade. Responde a reconhecer a necessidade de integrar e articular ações voltadas para os setores mais vulneráveis. Procura servir as escolas com base nas características socioculturais da comunidade a que os seus estudantes pertencem, reflete a necessidade de compreender o contexto para trabalhar a partir do respeito pela diversidade e gestão participativa de todos os atores.

⁵ O Programa de *Escuelas de Tiempo Completo* (em espanhol) constitui uma das principais políticas educacionais promovidas pelo Uruguai ao nível do ensino primário. O desenho do programa foi consolidado no final da década de 1990, alguns anos após a criação das primeiras escolas em período integral. Desde então, um número significativo de escolas foi criado ou convertido em tempo integral (Cardozo et al., 2017).

O desenho do Programa define-o como um modelo pedagógico integral baseado em três eixos fundamentais: i) Extensão do dia escolar com reconfiguração do modelo de escola de tempo parcial vai de quatro a sete horas e meia por dia; ii) Formação em serviço para professores e fortalecimento institucional com aumentar o componente de alimentação e a incorporação de diversas disciplinas de artes visuais, música, educação física e inglês; iii) A política também significa um investimento importante na construção de infra-estrutura e equipamentos e materiais de ensino.

Assim, o Programa significou um novo modelo de escola a partir do ponto de vista pedagógico e organizacional. Inicialmente estas escolas estão enfocadas no cuidado das crianças em contextos socioculturais desfavoráveis, mas também se espalhou para algumas escolas de contextos mais favoráveis.

Conjuntamente, as modalidades que supõem extensão do tempo de aula (Tempo Integral e Tempo Estendido) e a modalidade de escolas A.PR.EN.D.E.R. deram cobertura em 2014 a quase metade (47%) dos estudantes do 1º a 6º série em escolas da rede pública (ANEP, 2015b).

Tais transformações responderam a duas estratégias específicas: a) a forte extensão das escolas A.PR.EN.D.E.R. durante a segunda metade da década de 2000, que passou a cobrir 32% das matrículas do nível primário em 2008 e manteve essa proporção nos anos seguintes; b) o aumento gradual das Escolas de Tempo Integral, mais acelerada no último período, juntamente com a subsequente incorporação de Escolas Tempo Estendido, que proporcionam uma cobertura em 2014 de 16% dos estudantes (em comparação com 7% em 2004).

Atualmente, as escolas A.PR.EN.D.E.R. oferecem cobertura para 82,5% dos estudantes de 1ª a 6ª série em escolas urbanas do primeiro quintil socioeconômico e 74,0% no quintil 2. Os restantes estudantes desses dois quintis atendem Escolas de Tempo Integral ou Tempo Estendido (18% e 22%, respectivamente). No total, estas escolas cobrem, em conjunto, 16% das matrículas em educação urbana e as escolas A.PR.EN.D.E.R. 31% (ANEP, 2015b).

Paralelamente à criação e extensão de formatos escolares, as políticas de educação básica envolveram a progressiva incorporação de novos atores, com perfis complementares aos do professor em sala de aula. Foram criados cargos associados a programas específicos, como o Programa de Professores comunitários (PMC), que em 2014 trabalhou em 318 escolas, o programa “Trânsito entre ciclos educativos” (Programa de coordenação entre escolas de níveis primário y médio), que cobriu 254 escolas nesse mesmo ano. No conjunto, essas transformações refletem transformações importantes na oferta oferecida atualmente pela escola pública, em particular, na estrutura de apoio à tarefa central do professor tradicional em sala de aula.

Finalmente, como parte deste ciclo de transformações, deve-se destacar a incorporação de um amplo plano de desenvolvimento computacional (internet, software e hardware) conhecido como Plano CEIBAL, que envolveu, entre muitos

processos, o fornecimento de um próprio computador pessoal e de uso individual para cada estudante, primeiro no ensino primário e depois no ensino médio.

O Plano CEIBAL tornou-se uma ferramenta fundamental para introduzir um modelo educacional focado no século XXI, pois consegue provocar mudanças e transformações que hoje ocorrem em ritmo acelerado na sociedade do conhecimento e na revolução digital.

2.5 O SISTEMA DE AVALIAÇÃO DA EDUCAÇÃO NACIONAL

No Uruguai, desde 1996, com a implementação de um censo nacional de aprendizagem em Leitura e Matemática, há uma cultura de avaliação educacional no nível nacional desenvolvida na órbita da ANEP. Esta avaliação foi subsequentemente reeditada em base de amostragem nos anos de 1999, 2002, 2005, 2009 e 2013.

Seus resultados assinalam uma associação persistente entre diferentes contextos e níveis de proficiência ou teste de desempenho socioculturais, em todos os casos marcados por uma forte brecha de aprendizagem. Por exemplo, a avaliação de 2002 concluiu que:

"Na prova de matemática de 2002 as escolas de desempenho de "suficiência" de contexto muito favorável duplicavam a muito as de contexto desfavorável (35,7% e 72,3% respectivamente). Também é verdade que houve uma melhora nos níveis de proficiência que se concentra, em termos proporcionais, em contextos desfavoráveis" (CLAVIJO, FRANCIA e RETAMOSO, 2005: 11)

Além das avaliações nacionais, Uruguai participou em 2006 do Segundo Estudo Regional Comparativo e Explicativo (SERCE) e, em 2013, do TERCE, liderado pelo LLECE no Escritório Regional de Educação para a América Latina e o Caribe (OREALC / UNESCO Santiago). Os dados do estudo TERCE são a principal fonte de informação desta dissertação.

No ensino médio, o país participa continuamente do teste PISA desde 2003. No entanto, não existem estudos sistemáticos de aprendizagem nesse nível.

Institucionalmente, as avaliações de aprendizagem são de responsabilidade da Divisão de Pesquisa, Avaliação e Estatísticas (DIEE), na Secretaria de

Planejamento da Educação do ANEP, e tem a responsabilidade de desenvolver e implementar avaliações da aprendizagem a fim de fornecer elementos suficientes para o monitoramento e a avaliação de políticas educacionais da ANEP.

Por definição, as avaliações da DIEE são avaliações formativas, ou seja, são destinadas a melhoria da aprendizagem, enquanto procuram contribuir para a tomada de uma decisão sobre o processo educacional. O DIEE foi impulsionado pelo Sistema de Avaliação de Aprendizagem (SEA), que através da avaliação on-line e aplicação por computador, permitindo um conjunto diversificado de ferramentas de avaliação.

O DIEE também é responsável pela implementação nacional dos estudos PISA (nos ciclos 2003, 2006, 2009, 2012, 2015 e 2018) e ERCE-UNESCO (SERCE-2006 e TERCE em 2013).

Finalmente, deve-se notar entre as suas inovações, a Lei de Educação 2008 que criou o Instituto Nacional de Avaliação Educacional (INEED), que começou a funcionar em 2012. O INEED é uma instituição autônoma, com a responsabilidade de avaliar a qualidade da educação infantil, primária e média. Os objetivos do INEED incluem oferecer informações sobre a aprendizagem dos estudantes no Uruguai, desenvolver e disseminar conhecimento sobre procedimentos de avaliação e fazer recomendações para melhorar a educação.

3. FUNDAMENTAÇÃO TEÓRICA

Os estudos de avaliação educacional em larga escala, tanto a nível nacional como internacional, caracterizam-se por coletar, além das provas, um conjunto de informações sobre as características pessoais dos estudantes e suas famílias, professores e escolas, a fim de identificar as principais variáveis relacionadas aos desempenhos escolares (RAVELA, 2006). O estudo dessas condições relacionadas com o desempenho de estudantes e escolas é geralmente chamado de fatores associados.

O objetivo do estudo dos fatores associados é identificar as condições que influenciam a aprendizagem e medir a magnitude dos efeitos dessas variáveis no desempenho. O estudo de fatores associados à aprendizagem educacional é complexo, em parte, devido à variedade de conceitos e variáveis que devem ser avaliados, tanto no nível individual quanto no nível institucional. Deve-se ter em mente que, embora a aprendizagem seja um processo de resultado individual, ela faz parte de um campo de interação social. Assim, como indica Fernández (2004), a aprendizagem implica um processo de acumulação de experiências grupais, de forma que os níveis de aprendizado de um estudante é, uma função do aprendizado médio do grupo.

Nesta secção, apresentamos uma revisão e discussão sobre as principais correntes de pesquisa sobre as variáveis explicativas, ou seja, os fatores associados ao rendimento escolar. São definidas as relações entre um conjunto de fatores teoricamente relevantes e a qualidade da aprendizagem. Considerando que muitos desses fatores sejam definidos e observados de diferentes maneiras, é importante especificar como ele é compreendido e como cada conceito será observado neste trabalho.

Na análise dos fatores que determinam o desempenho cognitivo, a literatura, em geral, distingue os fatores em três grandes categorias: aqueles associados à estrutura escolar, os relacionados à família e os relacionados ao próprio estudante (SOARES, 2004; 2005; ANDRADE, 2007; MENEZES-FILHO, 2007; BLANCO, 2011; FERNÁNDEZ, 2004).

Se as diferenças sociais entre os estudantes que as escolas recebem são imutáveis a partir das capacidades do sistema educacional, há um conjunto de variáveis que a instituição escolar está em posição de manipular. A forma como o currículo é efetivamente implementado e organizado, o compromisso dos pais e da comunidade, a criação de um clima favorável para o desenvolvimento profissional dos professores e um agradável ambiente de aprendizagem são aspectos puramente institucionais que afetam o desempenho escolar.

Além disso, dentro da estrutura escolar pode-se distinguir os fatores que são parte da sala de aula, que se referem a características dos professores e do ensino, frente aqueles que são institucionais, ou seja, relacionado as formas da administração, recursos, infraestrutura e gestão escolar.

Nessa perspectiva, a escola, e especificamente a aula, é considerada o núcleo fundamental dos processos de interação, através dos quais a aprendizagem escolar ocorre. Neste núcleo distingue dois elementos fundamentais: os técnicos (pedagogia, didática, cobertura curricular, tempo de ensino) e os sociais (relações entre sujeitos baseados nos seus papéis como estudantes/professores e a definição estrutural/situacional desses papéis). Segundo Blanco (2011), os primeiros são sintetizados no conceito de "oportunidades de aprendizagem" e os segundos no conceito de "clima de aula".

Desta forma, cada seção deste capítulo é dedicada à apresentação de um fator (ou dimensão) teoricamente vinculados ao desempenho educacional.

3.1. FATORES INDIVIDUAIS E FAMILIARES

Esta primeira seção foca nos fatores pertencentes ao nível de análise individual, ou seja, as características dos estudantes e das suas famílias.

A revisão da literatura internacional indica que o histórico escolar (repetência e evasão), as práticas educacionais em casa e os atributos socioeconômicos, demográficos e culturais das famílias estão associados ao desempenho acadêmico (CRESPO, SOARES e SOUZA, 2000; SOARES, CÉSAR e MAMBRINI, 1997; SOARES, 2003; RAUDENBUSH e WILLMS, 1995; WILLMS, 2002).

Na trajetória educacional, a repetência e a ausência das aulas mantêm uma associação negativa robusta com o desempenho escolar. Pelo contrário, ter frequentado a educação pré-escolar é um fator que tem uma relação positiva com a aprendizagem.

O ambiente familiar, como as expectativas educacionais, o acompanhamento do desempenho acadêmico dos estudantes, os hábitos de leitura e o estudo diário em casa são fatores associados positivamente à aprendizagem.

É importante considerar que o estudante, além das suas próprias características pessoais, como gênero, cor/raça ou idade, também carrega toda a influência de seu contexto socioeconômico, incluindo aspectos que podem ser representados pela renda familiar, escolaridade e ocupação de seus pais (DIAZ, 2007; GONTIJO et al., 2011, SILVA; VENDRAMINI e LOPES, 2010).

Embora as características dos estudantes sejam exógenas, isso não significa que a alocação de estudantes em diferentes escolas seja aleatória. Ou seja, na medida que a escola não tenha o poder de decidir a escolha de seus estudantes, ela não deixa de exercer alguma influência no perfil dos estudantes que recebe. Critério similar se aplica aos professores e sua escolha de escolas. Dessa forma, assim como nem todos os fatores são totalmente aleatórios, nem todos são puramente exógenos.

3.1.1. Fatores Socioeconômicos e Culturais

O desempenho dos estudantes em cada etapa da vida escolar é o resultado de um conjunto de fatores que muitas vezes escapam aos limites das escolas. Dentre a magnitude dos fatores intervenientes (pessoal, familiar, social e institucional), há maior consenso acadêmico em identificar os principais determinantes do desempenho escolar que se referem às características socioeconômicas das famílias dos estudantes.

No Brasil, vários estudos buscam identificar esses determinantes através da avaliação de larga escala (FERRÃO et al., 2001; SOARES, CÉSAR e MAMBRINI, 1997; ALBERNAZ et al., 2002; SOARES e COLLARES, 2006). Um ponto comum entre esses estudos é a constatação da grande importância do *background* familiar sobre o desempenho do estudante. Estudos internacionais também destacam as

características do estudante e do seu grupo familiar como principal determinante do sucesso escolar.

A partir da publicação de “Equality of Educational Opportunity”, também conhecido como “Relatório Coleman” (COLEMAN et al., 1966), foi colocada a questão de que as características das escolas exerciam menor poder explicativo nos rendimentos escolares do que os aspectos referentes aos fatores familiares.

Assim, abriu-se o debate sobre até que ponto o sistema educacional reproduz as desigualdades sociais. O relatório concluiu que as diferenças nos resultados escolares foram explicadas quase inteiramente pelas diferenças familiares (raciais e socioeconômicas), enquanto as escolas contribuíram muito pouco para a diferenciação dos resultados dos estudantes. Verificou-se também que os recursos das escolas estavam altamente correlacionados com o status social de seus estudantes.

A pesquisa britânica conhecida como “Plowden Report”, realizada também na década de 1960, buscou avaliar as relações entre diversos fatores familiares, escolares e sociais e o desenvolvimento dos estudantes nas escolas. Os resultados e conclusões dessa pesquisa são semelhantes aos resultados do relatório Coleman, porém com um diferencial: o desempenho escolar tem uma maior relação com fatores psicológicos (atitudes e valores dos pais em relação aos estudos e o futuro escolar) da família do que fatores materiais e objetivos da mesma família (renda, escolaridade, condições materiais). Porém ambos os fatores, objetivos e subjetivos da família, possuem maior importância do que as condições materiais e os processos de escolarização das instituições escolares (FORQUIN, 1995: 32).

Estes resultados, devido ao seu interesse e controvérsia, exerceram uma grande influência na pesquisa acadêmica subsequente, por exemplo, as referidas pesquisas que relacionam as pertinências entre origem social e desempenho escolar foram elementos constitutivos para a produção teórica de Pierre Bourdieu. Assim, por mais de meio século, a pesquisa sistematicamente mostrou que o nível educacional dos pais dos estudantes, as condições culturais da casa e a situação econômica da família são os aspectos que mais afetam a aprendizagem dos estudantes.

3.1.2. O contexto sociocultural das escolas

Se ademais da origem socioeconômico individual, se agrega a concentração nas escolas de famílias de um determinado nível socioeconômico, assunto explicado pela segregação territorial e pelos processos de seleção e custos de matrícula nas escolas, a medida do Nível Socioeconômico dos Estudantes (NSE) de uma mesma escola também afeta a aprendizagem. Isto é de grande relevância devido à baixa inclusão social nos sistemas educacionais na América Latina (UNESCO, 2015).

Assim, segundo a literatura, o principal fator determinante não é tanto a origem social individual de cada estudante, mas a composição social da escola. Assim, a pesquisa de Hattie (1999) em base a meta-análise de 180.000 estudos que representam 50 milhões de estudantes mostra que a influência do nível socioeconômico é maior quando considerada no nível da escola do que no nível individual (HATTIE, 1999).

Assim, o contexto sociocultural das escolas e a heterogeneidade do *status socioeconômico e cultural* de seus estudantes são aspectos que têm grande impacto no desempenho.

Cervini (2002) destaca a importância de NSE médio nas escolas na aprendizagem. Utilizando dados da Argentina no *Operativo Nacional de Evaluación de la Calidad* (ONE), onde foram aplicadas provas padronizadas de Matemática e linguagem em estudantes da sétima série. O autor descobriu que grande parte da diferença no desempenho em Matemática e linguagem está ligada à variação do capital econômico e cultural dos estudantes em nível inter-escolar⁶.

No TIMSS 2015, em quase todos os países, os estudantes da 4ª série que frequentavam escolas com maior proporção de estudantes com vantagem econômica obtiveram um melhor desempenho em Matemática (MULLIS et al., 2016). Esse mesmo padrão de resultados ocorreu na 8ª série.

A mesma associação entre o nível socioeconômico escolar e a proficiência em Leitura foi observada no teste PIRLS 2011 (MULLIS et al., 2012).

⁶ Medido com base na disponibilidade de bens, medida de habitação superlotada, educação parental e disponibilidade de livros e material escolar.

Há abundante evidência local que coloca o Uruguai como um dos países com mais inequidade da América Latina (ANEP, 2014, 2015; INEED, 2014, 2015a, 2017; RIVAS, 2015). Essa forte segregação escolar deve-se às famílias enviarem seus filhos para escolas próximas de suas residências, em um contexto de forte segregação residencial. Como consequência, as escolas, mesmo entre as públicas, tendem a ser claramente diferenciadas de acordo com o nível socioeconômico dos estudantes que as compõem (INEED, 2015).

3.1.3. Gênero

Estreitar as diferenças educacionais de gênero constitui um dos Objetivos de Desenvolvimento Sustentável das Nações Unidas por sua contribuição direta para a redução da desigualdade. Pesquisas empíricas evidenciaram diferenças no acesso à educação por estudantes do sexo feminino, o que não seria explicado exclusivamente pelas condições socioeconômicas de suas famílias. Em particular, um estudo de Post (2001) conclui que a taxa de evasão de estudantes mexicanas é maior que a dos homens daquele país. O autor ressalta que esta persistência menor é devido aos papéis tradicionalmente atribuídos às mulheres, tais como uma maior colaboração em casa, e a importância relativa que as famílias atribuem ao acesso de suas filhas às escolas. As brechas de gênero na educação escolar também são observadas na aprendizagem. A equidade de gênero nessa área é desejável, pois as diferenças entre homens e mulheres minam o princípio do direito à educação (UNESCO, 2016).

No contexto da América Latina, a tendência em outros estudos internacionais com foco no ensino fundamental como o TIMSS e PIRLS (MARTIN, MULLIS, FOY e HOOPER, 2016) verificam que os homens obtêm melhores resultados em Matemática e, as mulheres, em Leitura. Não há diferenças significantes nas Ciências Naturais.

3.1.4. Trabalho infantil

A proibição do trabalho infantil pago se deve, em parte, à crença de que o período da infância deve ser dedicado ao desenvolvimento humano, especialmente ao aprendizado escolar.

Ainda não está claro que se uma família que decide que seu filho deve trabalhar, além de estudar, prejudica suas oportunidades futuras. Em contextos específicos do mercado de trabalho, dedicar tempo ao trabalho remunerado pode ser um investimento positivo para a família. Em outras palavras, a alternativa de estudar pode ter vantagens marginais. Apesar desses possíveis benefícios econômicos, o que explicaria as decisões das famílias, a evidência é conclusiva em demonstrar que, realizar o trabalho remunerado prejudica a aprendizagem dos estudantes (POST e PONG, 2000).

Vários estudos sugerem que as experiências de trabalho infantil em idade escolar, estão associados a resultados de aprendizagem mais baixos e com maior probabilidade de evasão e repetência e da pobreza na idade adulta (SEDLACEK, DURYEYEA, ILAHI e SASAKI, 2009).

Essa informação é preocupante, em parte porque cerca de 8% dos estudantes entre os 4 e 15 anos trabalham na América Latina, conforme dados de 2012 (CEPAL, 2012). Este valor sem considerar o trabalho não remunerado, isto é, o que os estudantes fazem em suas casas que é generalizado e subestimado pelos dados oficiais.

No caso da América Latina, desde a aplicação do estudo PERCE (1997), confirmou-se uma associação negativa entre o trabalho infantil remunerado e a proficiência dos estudantes.

Charlin e Torres (2017), com base nos resultados do TERCE, encontrou maior probabilidade de afastamento da escola entre os estudantes trabalhadores. Outros estudos empíricos indicam que essa associação afetaria em maior grau a aprendizagem de estudantes do sexo masculino, sendo mais pronunciada em Matemática e Ciências Naturais.

3.1.5. Trajetória acadêmica dos estudantes

No breve período que decorre desde o início da escolaridade até a 6ª série do ensino fundamental, é muito provável que, para a maioria dos estudantes, o seu trânsito pelo sistema educacional não tenha tido grandes transtornos. No entanto, para um número significativo de estudantes, também é provável que eles tenham experimentado experiências como repetição, absenteísmo ou evasão do sistema educacional.

Devido à relevância desses fatores, neste trabalho a trajetória educacional é resumida em dois indicadores: a experiência da repetição e os antecedentes da educação pré-escolar.

3.1.5.1. Repetência: A distorção idade/série

A repetência é uma prática comum nos sistemas educacionais de muitos países. Em outro, no entanto, é totalmente ausente ou usado para situações muito específicas. Assim, a repetência não é um instrumento uniforme em todos os sistemas educacionais que a utilizam. Alguns sistemas educacionais projetaram formas de certificação (ou creditação) que permitem decidir a passagem de série. Nestes casos, os estudantes avançam para a próxima série se, por exemplo, obtiverem a pontuação necessária nos diferentes testes nacionais, caso contrário, repetem automaticamente. A França e a Alemanha são bons exemplos onde a repetência é um fenômeno com alta prevalência.

Em contraste, o que é conhecido como *promoção social* é a prática de passar os estudantes para a próxima série, independentemente de seu desempenho educacional. Isso parece prevalecer em sociedades mais igualitárias, como Dinamarca, Finlândia, Noruega e Reino Unido.

Ao mesmo tempo, a repetição representa um instrumento de reparação na educação; em alguns países é o principal, e as vezes, o único instrumento para esse fim. No entanto, seu uso implica custos significantes. A repetição consome recursos significantes, uma vez que aumenta necessariamente o *stock* de estudantes

matriculados no sistema educacional. Em termos econômicos, os repetentes são indivíduos que, se concluírem sua educação formal, mais tarde se voltam para o mercado de trabalho ou podem ter tido uma vida laboral mais longa. Além disso, existem custos substantivos a longo prazo, no sentido de que estudos mostram que repetentes tendem a piorar a renda em sua vida laboral (BRODATY, GARY-BOBO e PRIETO, 2012).

Além disso, alguns estudos têm mostrado que a mera presença de repetição como uma ferramenta funciona como um incentivo e pode aumentar o esforço do estudante (DE FRAJA, OLIVEIRA e ZANCHI, 2010).

Muitos aspectos importantes de uma análise de custo-benefício ainda não são claramente compreendidos. Consequentemente, dado o seu uso generalizado, é difícil afirmar se a repetição é mais benéfica do que o sistema de promoção social, ou qual dos dois é mais valioso para a política educacional. Esta é uma questão de grande debate e a comparação internacional mostra evidências em favor de ambas as posições.

Tendências recentes na política educacional mostram um reposicionamento. Enquanto na França há uma redução na repetição, nos Estados Unidos ela foi reincorporada de uma perspectiva da rendição de contas (*accountability*) nas escolas.

Os resultados do PISA 2012 (OECD, 2013), indicam que nos sistemas educacionais em que os estudantes têm maior probabilidade de repetir uma série, o impacto do NSE dos estudantes no seu desempenho é maior do que nos sistemas em que esse mecanismo é usado com menos frequência. Em 35 dos 61 países e economias avaliadas, ao comparar dois estudantes com desempenho semelhante em Matemática, o estudante com menor NSE é mais provável que tenha repetido um ano.

Da mesma forma, nos resultados do PISA 2015 (OECD, 2016) esta relação entre distorção idade-série e desempenho é confirmada. A distorção idade-série é mais prevalente em sistemas educacionais nos quais os estudantes são menos proficientes no teste de ciências do PISA. Afirma-se também que quanto menos comum é a incidência da distorção idade-série, mais equitativo é o sistema educacional, ou seja, a associação entre o NSE do estudante e o desempenho em ciência é menor. Em geral, nos países da OCDE, observa-se que os estudantes que

repetem uma série geralmente são meninos (versus meninas), estudantes com NSE baixo (versus aqueles com NSE alto) e estudantes imigrantes (versus não-imigrantes).

O fenômeno da distorção idade-série apresenta uma alta prevalência na região da América Latina. No Brasil, a sua influência é bastante forte. Apresenta um alto índice de atraso escolar. Segundo o relatório internacional do Pisa 2012 menos da metade dos estudantes que participaram da avaliação cursava a série adequada para essa idade.

As consequências da distorção idade-série não são fáceis de estimar. No caso do Brasil, Bacchetto (2016) mostrou que é um mecanismo de alto custo. Segundo seus cálculos, a repetição total gerada pelo sistema educacional brasileiro ultrapassa o valor estimado de R\$ 8,8 bilhões e representa 9,1% do total do FUNDEB, em 2012.

Alves, Ortigão e Franco (2007) mostram que a distorção idade-série incide diferentemente em diversos grupos sociais de acordo com certas características, tais como o gênero, a etnia e o nível socioeconômico. Nesta perspectiva, entende-se que a repetência consiste em um mecanismo que não necessariamente atua somente por motivação pedagógica, supostamente baseada na aprendizagem dos estudantes, mas que reflete as desigualdades sociais brasileiras e de acesso à educação de qualidade.

Por outra parte, Riani, Silva e Soares (2012) mostram que os estudantes repetentes, em média, apresentam resultados em testes piores do que seus pares não repetentes, o que questiona a eficácia pedagógica deste mecanismo.

O Uruguai também apresenta uma situação preocupante em atraso escolar. O uso da repetição está permanentemente em debate. As autoridades nacionais manifestaram a sua vontade de reduzir as taxas de repetência e, assim, contribuir para melhorar as taxas de conclusão dos diferentes níveis de escolaridade obrigatória.

Segundo estatísticas oficiais, aproximadamente três em cada dez estudantes de escolas da rede pública concluem a 6ª série com, pelo menos, um ano de distorção idade-série (ANEP, 2015b).

Os estudos longitudinais do PISA-L sobre trajetórias educacionais no ensino médio no Uruguai, mostram sistematicamente o condicionamento quase "determinante" da defasagem escolar nas trajetórias educacionais⁷:

"Apenas 6% dos estudantes que o PISA 2009 avaliou com um ano ou mais de atraso para a idade (estudantes que estavam no 9º ano no máximo aos 15 anos de idade) conseguiram credenciar o ensino médio seis anos depois" (CARDOZO, 2016)

Por outro lado, todas as avaliações de aprendizagem indicam fortes brechas de desempenho produzido pela distorção idade-série no final do ensino fundamental, mas também em séries anteriores (ANEP, 1997; 1999; 2002; 2007; 2009).

3.1.5.2. Educação infantil e Pré-escola

O processo de consolidação do acesso à educação infantil (pré-escola, entre 4 a 6 anos de idade) é um dos fenômenos de importância crescente na política de educação em todo o mundo e, especialmente, na América Latina (ANDERS, et al., 2016).

A inclusão desta etapa tem suas maiores consequências para os estudantes mais desfavorecidos, tornando-se uma medida altamente custo-efetiva (BARNETT e MASSE, 2007). Os mecanismos de influência na aprendizagem dos estudantes são explicados de duas maneiras: por um lado, uma maior plasticidade das habilidades cognitivas durante os primeiros anos de vida, e em segundo lugar, a associação positiva entre as atividades pré-escolares (como a leitura em voz alta ou contar) e a futura disposição dos estudantes para habilidades mais avançadas, como compreensão de leitura.

⁷ Entre 2001 e 2003, o PISA Governing Bureau (PGB) do PISA estudou a viabilidade de implementar um levantamento longitudinal sobre as questões de transição da *escolar ao trabalho*. Esta proposta, chamada de PISA-L, procurou testar se as habilidades cognitivas desenvolvidas por estudantes após o ensino obrigatório tiveram um impacto específico sobre as transições para o mercado de trabalho e do ensino superior. Embora o projeto original não prosperar como um programa internacional, vários países têm implementado com sucesso estudos longitudinais com base no monitoramento de amostras nacionais do PISA: Canadá, Suíça, Dinamarca, Austrália, República Checa e Uruguai (BOADO e FERNÁNDEZ, 2010; CARDOZO e FERNÁNDEZ, 2013; CARDOZO, 2016).

Esses mecanismos têm sido observados ao associar a frequência pré-escolar dos estudantes aos seus resultados em avaliações em larga escala (ANDERS et al., 2016).

A associação entre o acesso a pré-escola e o desempenho também se manifesta nos resultados do PISA 2015. Em praticamente todos os países da OCDE com dados disponíveis, os estudantes que estavam pelo menos um ano no pré-escolar tiveram um desempenho melhor do que os outros (sem controle do NSE do estudante e da escola). Além disso, a associação é maior em sistemas educacionais que oferecem educação pré-escolar por um período maior de tempo, que têm uma proporção menor de estudantes para professores de pré-escola e que investem mais em crianças no nível pré-escolar (OECD, 2017). Esta mesma relação foi observada no TIMSS 2015 (MULLIS et al., 2016), na quarta série de estudantes de Matemática.

3.1.6. Tempo de exposição ao ensino

O tempo dedicado à aprendizagem tem se mostrado um aspecto amplamente relacionado à aprendizagem dos estudantes, sendo fundamental nas oportunidades de aprendizado dos estudantes durante os anos escolares (KUGER, 2016). Mesmo assim, as evidências são inconclusivas sobre a magnitude da associação entre o tempo em aprendizado e o aproveitamento escolar. Alguns autores enfatizam que essa associação é maior para os estudantes que inicialmente possuem maior desempenho de aprendizagem (HUEBENER et al., 2016); outros apontam que o investimento de tempo é mais benéfico na tenra idade (MARCOTTE, 2007). Especificamente há evidências de que o absenteísmo escolar é mais prejudicial para os estudantes com baixo desempenho (AUCEJO e ROMANO, 2014).

3.1.6.1. Absenteísmo escolar

Esta dimensão tem sido estudada em diferentes níveis, por exemplo, o tempo obrigatório que vários sistemas estabelecem para a educação, a dedicação à

aprendizagem nas escolas, nas salas de aula e ao nível do estudante, com o indicador da frequência o estudante faltou à escola. Em todos eles, existem fatores que limitam a dedicação efetiva do ensino, um dos quais é a falta dos estudantes. Através desta informação, se procura capturar parte do tempo dedicado à aprendizagem acadêmica (BERLINER, 1990).

Devido às características das avaliações padronizadas, como o TERCE, é difícil observar outras distinções, como o tempo das aulas em que os estudantes estão efetivamente comprometidos com a aprendizagem.

De acordo com os resultados do PISA 215, observa-se uma relação negativa na maioria dos países do estudo entre a ausência (se falta duas ou mais vezes por mês) e a proficiência, em qualquer uma das disciplinas avaliadas (OECD, 2016).

Em 44 países e economias avaliadas em PISA, o não comparecimento às aulas é mais comum em escolas socioeconomicamente desfavorecidas do que em escolas com *status* elevado. Em média, nos países da OCDE, os estudantes que faltam às aulas pelo menos um dia durante as duas semanas anteriores à avaliação do PISA obtiveram 45 pontos a menos em Ciências do que os estudantes que não faltaram à escola, explicando 2,7% da variância. Os oito países da América Latina com dados para essa variável mostram um comportamento similar (OECD, 2016).

Os resultados do TIMSS 2015 (MULLIS et al., 2016) confirmam a relação negativa entre o não comparecimento às aulas e o desempenho nas aprendizagens.

3.1.6.2. Dedicação ao estudo fora da escola

Como o tempo na escola não é a única variável que contribui para a aprendizagem, em um estudo de fatores associados é necessário incluir a dedicação fora da escola.

As oportunidades de aprendizagem fora da escola são variadas, desde a realização de tarefas no lar até atividades mais complexas. Estudos internacionais apontam para grandes diferenças entre os países e dentro deles em termos dessa dedicação adicional (BRAY et al., 2013).

A pesquisa de Kruger (2016) sugere que algumas dessas diferenças são devido às diferenças socioeconômicas das famílias e, também, às características individuais, como o sexo dos estudantes (as meninas seriam mais propensas a

gastar mais tempo estudando fora das escolas). Especificamente, as evidências indicam que a magnitude da associação com a aprendizagem é positiva. Alguns autores encontraram que a associação só existe quando se considera esta variável no nível de classe (ou seja, como a média dos estudantes na mesma aula dedicada à aprendizagem fora da escola). (TRAUTWEIN, KÖLLER, SCHMITZ, e BAUMERT, 2002).

Embora os resultados não são conclusivos. O PISA 2015 indica que os sistemas educacionais nos quais os estudantes passam mais tempo estudando depois da escola tendem a apresentar um desempenho pior em Ciência. Em Matemática, tem-se o mesmo padrão (OCDE, 0216).

Os resultados do TIMSS 2015 (MULLIS et al., 2016) são menos claros, porque, em geral, estudantes da 8ª série que declararam gastar menos de uma hora por semana para estudar fora da escola obtiveram resultados mais elevados, quando comparando com aqueles que passaram mais tempo.

3.1.7. Expectativas familiares e aspirações educacionais

As condições econômicas, culturais e familiares em que vive o estudante não são os únicos determinantes do desempenho educacional. A chamada socialização acadêmica dos estudantes tem um grande impacto em sua aprendizagem e é altamente dependente de suas famílias. A pesquisa é consistente em apontar que as famílias são os agentes primários desse processo, delineando essa socialização por meio de suas crenças, práticas e experiências pessoais (TAYLOR, CLAYTON e ROWLEY, 2004).

Do exposto, existe uma extensa literatura sobre o que os pais fazem e o que os pais são e pensam. Suas expectativas sobre o nível de escolaridade alcançado por seus filhos estão incluídas no segundo tipo; isto é associado à aprendizagem educacional, mas também a outros resultados, como por exemplo, o salário em sua vida adulta (HILL, 1997).

Em estudos como PIRLS 2011 (MULLIS et al., 2012) se indica uma correlação positiva entre as expectativas dos pais quanto ao alcance da educação

superior por seus filhos e sua proficiência em Leitura. No entanto, deve-se atender que esta associação tende a diminuir quando se considera o NSE do estudante e da escola.

3.1.8. Envolvimento dos pais na vida escolar

O envolvimento dos pais na aprendizagem dos estudantes é um fator que tem sido amplamente desenvolvido como um elemento que facilitaria o aprendizado. Em particular, destacam-se três áreas em que sua intervenção afetaria as conquistas de aprendizado dos estudantes: as atividades que realizam com os estudantes em casa, o grau de envolvimento nas atividades escolares (seja colaborando com algumas atividades específicas ou mantendo a comunicação com os professores) e suas motivações ou expectativas sobre sua participação na aprendizagem (HERTEL e JUDE, 2016).

Em geral, tanto o envolvimento nas escolas quanto em casa favoreceria o desenvolvimento cognitivo e socioemocional das crianças. Para a primeira área, há ampla evidência da associação positiva do envolvimento na escola e dos resultados dos estudantes. Mesmo assim, o grau de participação dos pais nas escolas (seja em entrevistas com professores ou outras instâncias) depende da disposição dessas e das características individuais das famílias. Assim, os antecedentes constataam que os benefícios desse envolvimento nas escolas são amplificados para os estudantes de maior NSE, e mais modestos para aqueles que vêm de famílias mais desfavorecidas. Esses resultados apontam para a importância de incentivar o envolvimento dos pais entre os estudantes com condições socioeconômicas mais desfavorecidas (HILL et al., 2004).

As condições de aprendizagem nas casas afetam o desenvolvimento de competências e do aprendizagem dos estudantes. Alguns deles referem-se a oportunidades de leitura precoce, desenvolvimento de linguagem e habilidades socioemocionais e apoio emocional dos pais (HERTEL e JUDE, 2016).

Outro aspecto positivamente ligado à aprendizagem do estudante refere-se à estimulação cognitiva que os estudantes recebem em suas casas. As evidências mostram que haveria uma relação entre as condições socioeconômicas das famílias e a presença de atitudes ou ações no lar que favorecem o aprendizado.

Especificamente, os pais mais favorecidos apresentariam, em maior medida, ações que estimulassem a aprendizagem e ambientes favoráveis ao aprendizado e, com menor probabilidade, transferissem essa responsabilidade para as escolas (SON e MORRISON, 2010).

3.2. FATORES ESCOLARES

Os fatores sociais estão fora do controle das escolas e do sistema educacional. De acordo com Soares (2004), os fatores externos, ou fatores *extraescolares*, estão associados às condições socioeconômicas dos estudantes, ou seja, levam em consideração as características da família, do próprio estudante e do contexto social no qual eles estão inseridos. Nesse campo, a política educacional tem pouco a fazer, exceto evitar contribuir para a segregação social da matrícula ou estimular a motivação dos pais para suporte de seus filhos na escola.

Existe, ao contrário, um conjunto de fatores que estão diretamente na esfera de influência dos sistemas educacionais. Os fatores internos à escola, ou *intraescolares*, estão relacionados à gestão, aos professores, ao ambiente e infraestrutura escolar. Esses fatores, principalmente ligados aos recursos e práticas dentro das escolas, os níveis de autonomia com que atuam e os apoios que recebem dos níveis centrais, são geralmente referidos como *modificáveis*. Ou seja, são aqueles que podem ser objeto de política educacional, uma vez que é possível intervir neles através da tomada de decisão dentro do sistema educacional.

Do trabalho de Klitgaard e Hall (1979), Edmonds (1982) e Brookover (1979) emerge uma nova abordagem dentro da sociologia da educação, que redefine o "efeito escola". Miller (1985) refere-se assim ao grupo de estudos que não mais tentam demonstrar o óbvio (por exemplo, que existem escolas extraordinariamente eficazes, uma vez que os fatores socioeconômicos são controlados), mas passaram para questões mais substantivas, como melhoria da qualidade da pesquisa empírica ou análise de processos de mudança educacional e organizacional.

Essa perspectiva decorre da observação de que estudantes pertencentes a um mesmo contexto apresentam diferenças de desempenho que não podem ser

plenamente explicadas por variáveis de estratificação familiar, e a identificação de que essas diferenças estão sistematicamente relacionadas às características da escola.

Dentro deste último, fatores como ênfase acadêmica, práticas de ensino, disponibilidade e tipos de incentivos são observados, e o grau em que os próprios estudantes assumem a responsabilidade da aprendizagem. A sociologia da educação experimentou uma mudança para uma visão mais otimista do papel da escola, enfocando "a eficácia da escola" e "o desenho da organização escolar" como um nível de análise (Fernández, 1999).

Como indica Soares (2004), existem ao interior dois claras linhas de pesquisa,

“Na literatura americana e inglesa usam-se duas nomenclaturas para identificar essa linha de pesquisa. Os autores que descrevem seu objeto como School Effectiveness Research enfatizam modelos teóricos oriundos da Economia. Outros autores, como Lee (2000), preferem descrever sua área de interesse como pesquisa sobre Schools Effects Research. Aqui a abordagem é mais próxima da Sociologia da Educação e da Pedagogia. As complexidades presentes em uma escola real, associadas à interação das pessoas que fazem o seu dia-a-dia, são centrais ao raciocínio. Nessa abordagem a questão da equidade de resultados merece tanto ou mais atenção do que a da eficácia escolar. Nessa vertente os trabalhos são mais teóricos e os modelos, naturalmente, mais complexos. Embora seja útil reconhecer as duas diferentes abordagens existentes na área, é preciso destacar a impossibilidade de se classificar, seja a maioria dos autores, seja a dos artigos, como pertencentes somente a uma ou outra categoria” (SOARES, 2004: 84)

Em relação aos insumos escolares, Soares e Candian (2007), utilizando dados do SAEB e do PISA 2003, avaliam o efeito das escolas brasileiras no desempenho dos estudantes. Os autores ressaltam que as escolas brasileiras são muito distintas e que isso se deve à segmentação socioeconômica, além de enfatizarem a importância da gestão escolar na promoção de atividades que utilizem os recursos pedagógicos existentes, de forma adequada a garantir uma melhoria do desempenho cognitivo dos estudantes. Assim, em relação ao papel dos insumos escolares, não há um consenso sobre quais são os principais fatores determinantes do sucesso escolar. Hanushek, Rivkin e Taylor (1996), a partir de uma vasta revisão de estudos anteriores e de novas análises empíricas, concluíram que artigos que

indicam uma relação positiva e significativa entre os insumos educacionais e o desempenho dos estudantes utilizavam dados com tal grau de agregação que acabavam por inflar os coeficientes estimados. Já para os estudos nos quais os dados empregados possuem um baixo nível de agregação, as evidências de quaisquer relações positivas entre as características da escola e o desempenho se mostram bem menores.

Hanushek e Luque (2003) discordam de um resultado encontrado por Krueger e Whitmore (1999), afirmando que turmas menores são benéficas apenas nas séries iniciais e não em todas as séries.

Observando especificamente o impacto dos professores, Hanushek e Luque (2003) dizem que apenas 9% dos estudos avaliados por eles encontraram uma relação significativa e positiva entre a escolaridade dos professores e o desempenho dos estudantes, e 49% encontram uma relação positiva e significativa entre experiência dos professores e desempenho.

Albernaz et al. (2002) encontraram, no Brasil, um impacto positivo da escolaridade dos professores sobre as notas, mesmo controlando diversos fatores. O salário dos professores também é significativo quando colocado no modelo como substituto da escolaridade. Além disso, o impacto dessas variáveis é maior para os estudantes de melhores condições socioeconômicas. Rios-Neto et al. (2002) também encontram evidência do efeito da escolaridade do professor sobre o desempenho dos estudantes. Por sua vez, Felício (2004) obteve resultados inconclusivos tanto com relação à experiência do professor como com relação à escolaridade dos mesmos.

Em definitivo, nesta seção, vamos aprofundar nos fatores que são frequentemente chamados de *intraescolares*, já que dependem da dinâmica do espaço institucional.

3.2.1. Práticas e atitudes dos professores no processo de aprendizagem

As evidências mostram que o desempenho do professor é o mais relevante para explicar a contribuição do fator escolar na aprendizagem dos estudantes (ROCKOFF, 2004; RIVKIN, HANUSHEK e KAIN, 2005; ARAUJO, CARNEIRO, CROSS-AGUAYO e SCHADY, 2014). Esta situação é ainda mais relevante em contextos com alta desigualdade, tais como os países da América Latina, onde a associação dos resultados dos estudantes com suas condições socioeconômicas de suas famílias é o fator que mais afeta a aprendizagem (BRUNS e LUQUE, 2014).

Para estudar práticas de ensino, crenças e atitudes, suas diferentes dimensões devem ser diferenciadas. Para tanto, optou-se pela identificação dos domínios do Marco de Ensino de Charlotte Danielson (2011), principalmente porque as evidências sustentam que estes são domínios associados à aprendizagem do estudante. Em sua versão resumida⁸, o modelo distingue as seguintes dimensões: i) A gestão da sala de aula, ou seja, a capacidade de o professor criar um ambiente apropriado para a aprendizagem; ii) Interação pedagógica, que inclui suas intervenções e organização de experiências de ensino e aprendizagem para promovê-lo e; iii) Responsabilidades profissionais, que se referem à capacidade reflexiva de melhoria contínua e ao exercício responsável da profissão.

É importante notar que o modelo de Danielson enfoca a prática de ensino em sala de aula com base no fato de que a melhoria na aprendizagem se deve principalmente às práticas de apoio observáveis através do relacionamento professor-aluno. Essa perspectiva se opõe à crença de que o aprendizado melhora através de práticas de pressão.

Finalmente, este tipo de indicador tem alta confiabilidade e validade quando se trata de identificar práticas de ensino que proporcionam maiores incentivos para a aprendizagem do estudante⁹ (HO e KANE, 2012; KANE, McCAFFREY e STAIGER, 2010; 2012).

⁸ O modelo completo inclui uma dimensão muito relevante associada à qualidade do planejamento e preparação das aulas, mas devido à complexidade de medir alguns desses indicadores por meio de questionários, decidiu-se excluí-lo como um fator associado (DANIELSON, 2011).

⁹ É medido através da opinião dos estudantes sobre: i) o que eles fazem, pensam ou sentem durante as aulas e, ii) como eles percebem seus professores. Assim, os questionários reúnem opiniões, tanto

Por práticas pedagógicas refere-se à atuação do professor em sala de aula, considerando as intervenções pedagógicas e como estruturar as experiências da aprendizagem. Segundo Danielson (2011), os aspectos relevantes dessa dimensão são: comunicação com os estudantes, técnicas de questionamento e discussão utilizadas, motivação para a aprendizagem e uso formativo das avaliações.

De acordo com as evidências, essas perguntas são válidas e confiáveis quando se consultam estudantes do ensino fundamental (FAUTH, DECRISTAN, RIESER, KLIEME e BÜTTNER, 2014; DOWNER et al, 2015). Portanto, as informações sobre esta área serão coletadas principalmente através dos estudantes e sua visão sobre o que eles fazem ou pensam durante as aulas, e como eles percebem seus professores e o tipo de relacionamentos e interações que eles promovem.

Também serão incluídas perguntas aos professores sobre suas práticas de sala de aula. Especificamente, os professores são consultados sobre a frequência das atividades em grupo dos estudantes, discussões, trabalhos em projetos que envolvem outros assuntos e a participação dos estudantes nas aulas.

3.2.2. Violência No Entorno Escolar

A exposição à violência pode gerar transtornos emocionais e psicológicos e problemas de saúde física (BOWEN e CHAPMAN, 1996). Espera-se que, se puder gerar esses efeitos, também afete o desempenho escolar por meio do impacto negativo no desenvolvimento cognitivo e emocional do estudante. A violência no ambiente também pode influenciar as famílias a estabelecer rotinas mais rigorosas e proteger seus filhos, mantendo-os em casa, o que está negativamente relacionado ao desempenho acadêmico (SHONKOFF e PHILLIPS, 2000). Finalmente, a violência também pode fragmentar o tecido social da comunidade, isolando as pessoas, aumentando a criminalidade do setor e influenciando negativamente o comportamento dos estudantes (HARDING, 2009).

sobre a percepção de esforço e motivação que os estudantes vivenciam nas aulas, como sobre o manejo da disciplina, a didática e o tipo de relacionamento que seus professores promovem.

Para fins dessas análises, definiremos a violência no ambiente escolar como baseada na percepção que as famílias dos estudantes têm de situações de agressão ou comportamento ilegal no ambiente escolar (por exemplo, uso e venda de drogas, delinquência, agressões entre vizinhos).

Bowen e Bowen (1999) descobriram que quanto mais os estudantes foram expostos à violência no bairro e na escola, mais eles perderam a escola, pior eles se comportaram e piores foram suas notas. Da mesma forma, Burdick-Will (2016) encontrou um efeito a longo prazo da exposição à violência na vizinhança em idade precoce sobre o desempenho em testes padronizados ao longo do tempo.

3.2.3.Clima Escolar

A existência de um ambiente de convivência saudável no espaço educacional é um elemento fundamental para uma educação de qualidade. Isso se relaciona, em um sentido amplo, com o desenvolvimento educacional, com a complexidade da vida escolar e envolve aspectos fundamentais da dinâmica escolar.

Como Carrasco (2016) aponta, em geral, concorda-se que um clima escolar favorável é importante para o bom funcionamento de uma escola, no entanto, o conceito de clima escolar não é unívoco e, portanto, está associado a diferentes dimensões e maneiras de medi-lo. Portanto, no estudo da relação entre o clima escolar e os resultados educacionais, é necessário, em primeiro lugar, esclarecer a natureza desse conceito.

Nos estudos educacionais, o clima escolar pode ser definido como o conjunto de percepções e expectativas por parte dos membros da comunidade escolar. Corresponde a uma avaliação subjetiva em relação a dimensões inter-relacionadas, como normas, objetivos, valores, relações humanas, organização institucional e estruturas físicas, pedagógicas e administrativas presentes em uma instituição de ensino (MORO, 2018).

As escolas são ambientes privilegiados para desenvolver as capacidades de alteridade, de identificação com o outro, respeito e um conjunto de valores tendentes a uma coexistência harmoniosa (GOERGEN, 2005). Nesse sentido, a aprendizagem é um processo com implicações cognitivas e afetivas, onde os aspectos emocionais exercem papel fundamental.

4. MATERIAIS E MÉTODOS

Nesse capítulo serão apresentados os procedimentos utilizados para análise dos dados. Uma vez que se baseia nos dados do TERCE, é feita uma breve descrição dessa avaliação internacional.

Utilizando os questionários de contexto do TERCE, será construído um conjunto de escalas. Para a construção das escalas será utilizada a Teoria da Resposta ao Item – TRI, modelo de resposta gradual de Samejima (1969).

Os indicadores serão submetidos, juntamente com as demais variáveis individuais, ao Modelo Linear Hierárquico (MLH). Na teoria o MLH possibilita relacionar as variáveis explicativas considerando a hierarquia dos dados, que na estrutura escolar refere-se ao aluno, a sala de aula e a escola.

Dentro desse contexto são também expressas considerações sobre a metodologia dos Valores Plausíveis como estimativa da proficiência em Leitura, Matemática e Ciências Naturais dos estudantes da 6ª série no TERCE de 2013. As seções seguintes farão a descrição passo a passo das metodologias.

4.1. CONTEXTO DO ESTUDO: TERCE

4.1.1. O Laboratório Latino-Americano da Avaliação da Qualidade da Educação

O Laboratório Latino-Americano da Avaliação da Qualidade da Educação (LLECE, em espanhol), da Organização das Nações Unidas para a Educação (UNESCO), foi criado em 1994 como uma rede de unidades para medir e avaliar a qualidade dos sistemas educacionais dos países latino-americanos. A Coordenação Técnica é feita por uma equipe da Oficina Regional de Educação da UNESCO para América Latina e Caribe (OREALC / UNESCO Santiago).

Desde a sua criação, o LLECE tornou-se uma referência regional para a cooperação entre países no campo da avaliação da educação e, também, como suporte técnico para o treinamento de equipes responsáveis por sistemas nacionais de medição e avaliação.

A abordagem de avaliação do LLECE considera a qualidade da educação como um conceito multidimensional, em que a aprendizagem é um elemento essencial, mas não o único. De acordo com a UNESCO (2007), a educação de qualidade também se refere a variáveis como a eficiência (bom uso dos recursos), equidade (distribuição de benefícios educacionais), relevância (que atenda às necessidades da sociedade).

No âmbito desta conceituação, alcançar a melhoria da aprendizagem ocupa um lugar importante como indicador de qualidade. O monitoramento dos resultados obtidos pelos diferentes grupos de estudantes, tem sido crucial para identificar as lacunas na equidade, particularmente, quando se considera a variável de pertencer a grupos indígenas e sexo do estudante.

Durante os últimos vinte anos, a maioria dos países da região experimentou importantes avanços em áreas-chave como o desenvolvimento global o crescimento econômico e, em menor medida, a redução da pobreza. Tudo isso gerou um contexto favorável para o avanço da educação.

No entanto, a persistência de altos níveis de desigualdade e pobreza e a alta proporção da população que vive em áreas rurais, continuam a oferecer dificuldades adicionais à expansão da educação de qualidade na maior parte da região.

4.1.2. Os Estudos da UNESCO: PERCE, SERCE e TERCE

Dentro deste quadro de ação, o LLECE desenvolveu três versões do seu Estudo Regional Comparativo e Explicativo (ERCE). Em 1997, o primeiro estudo, chamado PERCE, foi aplicado em 13 países. Esta pesquisa avaliou Matemática e Leitura no terceiro e no quarto ano do ensino fundamental. Além disso, outras dimensões associadas a aspectos do espaço escolar foram avaliadas. É possível verificar, nos resultados do relatório, que, entre os efeitos associados à qualidade da educação, o clima em sala de aula se mostrou particularmente expressivo, uma vez que assinala que nas instituições nas quais os estudantes se importam uns com os

outros, reconhecem a importância do outro, em que a incidência de brigas é pequena, prevalecendo o respeito, os resultados foram superiores (UNESCO, 2001).

O segundo estudo (SERCE) foi aplicado nove anos depois, em 2006, e avaliou as áreas de Matemática e Leitura na terceira e na sexta série de ensino fundamental; e Ciências Naturais na sexta série. O terceiro estudo, o TERCE, foi aplicado em 2013, avaliando as mesmas áreas do SERCE e adiciona a avaliação na área da Escrita.

4.1.3. Objetivos e Características do TERCE

O principal objetivo do Terceiro Estudo Regional Comparativo e Explicativo, TERCE, foi avaliar a qualidade da educação em termos de aprendizagem nos países participantes da América Latina e Caribe e identificar os fatores associados a ela (UNESCO, 2016).

Nesse sentido, o TERCE não apenas fornece um diagnóstico da situação com base no alcance da aprendizagem, mas também procura contribuir para a melhoria educacional, sob a premissa de que conhecer os fatores associados auxilia na formulação de políticas públicas e na revisão do processo de aprendizagem.

O TERCE, fundamentalmente, tenta responder a duas questões de pesquisa:

- i) Qual é o nível de desempenho escolar geral dos estudantes do ensino fundamental, na 3ª e 6ª séries, nas áreas de linguagem (Leitura e Escrita), Matemática e Ciências Naturais nos países participantes?
- ii) Qual é a relação entre o desempenho escolar e outras variáveis (fatores associados) vinculadas aos estudantes e suas famílias; aos professores e salas de aula; e às escolas e sua gestão?

Correspondendo a esses dois eixos principais de pesquisa, o TERCE utilizou dois tipos de instrumentos de coleta de informações para alcançar seu objetivo: testes (provas) de avaliação de aprendizagem e questionários contextuais.

A aprendizagem foi medida através de testes de Leitura e Matemática na 3ª e na 6ª série do ensino fundamental, além de Ciências Naturais somente na 6ª série.

Os questionários contextuais foram desenvolvidos levando em consideração o referencial teórico do estudo, a partir de uma revisão da literatura sobre os fatores que podem afetar o alcance da aprendizagem no nível fundamental. O TERCE tem questionários para estudantes, famílias, professores e diretores. As informações consultadas por meio desses instrumentos permitiram analisar fatores associados às principais características dos sistemas educacionais participantes.

O desenho amostral do TERCE possibilita que os resultados do estudo sejam representativos de duas populações dentro do sistema educacional: estudantes na 3ª e 6ª séries dos países participantes.

Cabe ressaltar que, ao contrário do SERCE, no TERCE todos os países participaram da avaliação das Ciências Naturais, refletindo a importância dessa disciplina no contexto da avaliação e na formação recebida por crianças e jovens nos sistemas de ensino da região.

4.1.3.1. Provas e questionários do TERCE

Além do interesse em gerar informações sobre aprendizagem em diferentes países da América Latina, o TERCE investigou fatores sociais, econômicos, culturais, organizacionais e políticas educacionais que impactam o desempenho acadêmico dos estudantes da região. Como indica a Relatório Técnico do TERCE:

“Por um lado, a realização de um diagnóstico comparativo e periódico de aprendizagens é fundamental para gerar avanços educacionais. Por outro lado, uma parte importante do significado e escopo deste diagnóstico é enfraquecida se não for combinada com a intenção de entender as condições sob as quais os processos educacionais dentro da região são enquadrados, os fatores que explicariam as diferenças de desempenho acadêmico e intervenções que poderiam melhorar os processos de ensino-aprendizagem no contexto latino-americano”. (UNESCO, 2016a: 164)

Assim, questionários contextuais acompanham as provas de desempenho acadêmico em Matemática, Leitura e Ciências Naturais. Em primeiro lugar, espera-se que os dados dos questionários permitam estudar as hipóteses do estudo sobre os fatores ou variáveis que seriam associados aos resultados de aprendizagem dos estudantes do ensino fundamental. Portanto, na seleção das dimensões a serem

incluídas nesses instrumentos, foi dada prioridade àquelas que, segundo estudos anteriores, apresentam associação com a aprendizagem.

A coleta de informações sobre essas questões fundamentais é justamente a finalidade dos Questionários de Contexto, que - através das percepções dos diferentes atores do processo educacional - permitem caracterizar o contexto micro e macro no qual a experiência educativa está inserida. Para este fim, os questionários apresentados no Quadro 4.1, coletam informações através de consulta com estudantes, famílias, professores e diretores das escolas.

Quadro 4.1. Questionários de fatores associados.

Questionários	Objetivo
Estudante	Conhecer as características pessoais, assim como seus lares e escolas, e investigue o relacionamento com seus professores e a atitude que eles percebem deles. Além disso, procura identificar suas atividades fora da escola, tanto acadêmicas quanto recreativas.
Família	Informe-se sobre as características socioeconômicas, demográficas e culturais do lar, bem como sobre a disponibilidade de recursos educacionais (livros), atitude em relação à leitura na família e comportamentos da criança (frequência às aulas, frequência e horários de estudo em casa etc. Coleta também informações sobre o envolvimento da família no processo de aprendizagem e sobre as expectativas de desenvolvimento acadêmico relacionadas aos alunos. Por fim, coleta informações sobre as características do bairro em que a escola está inserida.
Professor	Conhecer as características pessoais e histórico de trabalho dos professores e investigar o seu trabalho, além das expectativas sobre seus alunos, clima, liderança e gestão escolar.165
Diretor	Reúne informações sobre suas características pessoais, o ambiente e a infraestrutura da escola e sobre o gerenciamento da escola.

Fonte: Extraído do Relatório TERCE (UNESCO, 2016).

4.1.3.2. Dados e variáveis do estudo TERCE

Os dados utilizados nesta análise são provenientes da avaliação regional TERCE e correspondem à 6ª série do ensino fundamental para o ano de 2013.

O desenho amostral foi de tipo estratificado em dois estágios de conglomeração, representativo em nível nacional. No primeiro estágio, as escolas são selecionadas (em cada estrato explícito) com probabilidade de seleção proporcional ao tamanho¹⁰. No segundo estágio, uma sala de aula é selecionada aleatoriamente e aplicada a todos os alunos que pertencem a essa sala de aula.

Após a depuração da base de dados, o número de casos disponíveis para a análise é apresentado na Tabela 4.1 para cada uma das áreas avaliadas.

Tabela 4.1. Tamanho da amostra final de estudantes por prova e questionário na 6° série de TERCE.

		Instrumento	Amostra
Provas		Leitura	2.799
		Matemática	2.799
		Ciências Naturais	2.803
Questionários		Estudante	2.939
		Família	2.645
		Professor	165
		Diretor	167

Fonte: Elaborado pelo autor em base de dados do Uruguai no estudo TERCE (UNESCO 2013).

As variáveis dependentes são as três escalas de habilidades (Leitura, Matemática e Ciências Naturais) com escores padronizados, construídas de acordo com o modelo de Rasch da TRI.

A escala de pontuação do TERCE é construída de tal forma que sua média, ponderada pelos pesos amostrais e considerando cada país com um universo de alunos de igual número, é de 700 pontos e seu desvio padrão é de 100. Além disso, pontuações a uma distância maior que 4,5 vezes o desvio padrão acima da média são truncadas. Na prática, isso implica que o escore máximo teórico de cada teste corresponde a 1150 pontos e o mínimo teórico a 250. Esta distância sobre a média global para realizar o truncamento é escolhida com cuidado para não causar um

¹⁰ A definição de escola que é aplicada no TERCE, refere-se às unidades educacionais onde pelo menos o terceiro e / ou o sexto ano são ensinados, e que têm as seguintes características:

i) Tem uma localização física única; ii) Tem um número definido e identificável de alunos e professores; iii) Possuem um sistema único de administração / gestão educacional; iv) Tem um espaço social identificável (UNESCO, 2016: 206)

efeito de truncamento maciço em países com pontuações muito altas ou muito baixas (o que é conhecido como efeito "teto" ou "chão") (UNESCO, 2016).

Os índices e variáveis independentes foram construídos a partir dos questionários contextuais aplicados a estudantes, famílias, professores e diretores.

4.2. TEORIA DE RESPOSTA AO ITEM

Para a construção das escalas dos indicadores foi utilizada a Teoria de Resposta ao Item (TRI). A utilização da TRI é comum em várias áreas de conhecimento, em destaque em avaliações na área educacional. A TRI possui vários modelos que permitem medir um traço latente. Um traço latente pode ser entendido como a característica que não pode ser observada diretamente no indivíduo, mas que pode ser medida utilizando variáveis secundárias relacionadas a ela.

De acordo com Alves e Soares (2007), as vantagens da TRI em relação a outras metodologias de construção de escalas são que além da forma natural que tem de abordar os dados ausentes, o resultado é uma escala hierárquica que mede cada indicador a partir da combinação ponderada de diferentes variáveis. Além disso, a TRI permite a comparação entre populações ao longo do tempo, desde que sejam submetidas a alguns itens comuns. Permite ainda, a comparação de indivíduos que dentro da mesma população tenha se submetido a testes completamente diferentes (ANDRADE, TAVARES e VALLE, 2000).

O modelo de TRI utilizado neste trabalho é o de resposta gradual de Samejima (1969). Esse modelo possibilita medir o traço latente, ou seja, a característica do indivíduo, agrupando diversos tipos de itens. Por exemplo, os questionários contextuais do TERCE apresentam itens de diversas formas, uns são dicotômicos, ou seja, com duas alternativas de respostas, e outros politômicos. Ressalta-se que as alternativas dos questionários contextuais do TERCE não são consideradas certas ou erradas, são características relacionadas ao indivíduo.

No modelo de Samejima, os itens do questionário do estudante no TERCE serão recodificados de forma gradual, considerando que, quanto maior for o valor atribuído à alternativa da questão, maior será o valor do traço latente.

O modelo de Samejima pode ser descrito pela seguinte equação:

$$P_{ik}^+(\theta_j) = \frac{1}{1 + e^{-\alpha_i(\theta_j - b_{i,k})}}$$

Com $i = 1, 2, \dots, l$; $j = 1, 2, \dots, n$; e $k = 2, 3, \dots, n$, em que:

b_{ik} representa o parâmetro que indica a inflexão da curva de probabilidade da resposta da k -ésima categoria do item i .

θ_j representa a habilidade ou nível de concordância (traço latente) do j -ésimo indivíduo.

P_{ik}^+ é a probabilidade de um indivíduo j com habilidade ou nível de concordância θ_j responder o item i na categoria k ou superior.

α_i é o parâmetro de discriminação (ou de inclinação) do item i , com valor proporcional à inclinação da Curva Característica do Item – CCI no ponto b_{ik} .

O valor de θ é estimado com base nas respostas dos questionários e é tão maior quanto maiores forem os valores atribuídos às respostas dos itens. No modelo são considerados apenas os itens para os quais os indivíduos forneceram resposta, desconsiderando aqueles em branco. Para a análise dos itens que compõe cada indicador foi levado em consideração a CCI e sua curva de informação (CII). A CCI permite analisar o nível de informação do item e estimar a habilidade do indivíduo, permite ainda, relacionar a função de informação com os parâmetros do item. Já a CII compreende a soma de informações fornecidas por cada item do teste.

4.3. MODELO LINEAR HIERÁRQUICO

O Modelo Linear Hierárquico (RAUDENBUSH e BRYK, 1986; RAUDENBUSH, 1993) é um modelo estatístico que relaciona variáveis independentes (de diferentes níveis) com o objetivo de explicar a variação de uma variável dependente. Estes são, em essência, modelos de regressão, mas apresenta algumas particularidades.

A análise através do ajuste desses modelos também é conhecida como 'random coefficient model'(DE LEEUW e KREFT, 1986) ou 'variance component model' (LONGFORD, 1993). Também a literatura de estatística refere-se a estes como modelos mistos (LITTELL, MILLIKEN, STROUP, WOLFINGER e SCHABENBERGER, 1996).

Apesar de suas pequenas diferenças, em todos os casos esses nomes referem-se, genericamente, a modelos de regressão multinível, pois assume que o modelo apresenta níveis diferentes de agrupamento dos dados.

Para Barbeta, Andrade e Tavares (2018) os modelos lineares hierárquicos podem ser considerados uma extensão da regressão linear múltipla para acomodar uma estrutura hierárquica, como é o caso dos estudantes que estão agrupados em escolas. Nesses modelos, procura-se avaliar a parcela do desempenho do estudante devida a cada fator associado ao estudante (nível socioeconômico familiar, percepção do estudante quanto ao clima escolar, necessidade de o estudante fazer trabalhos domésticos etc.) e cada fator associado à escola (infraestrutura, localização, dependência administrativa etc.).

Para cumprir o objetivo deste trabalho, modelos de regressão multinível são aplicados. Cervini (1999, 2002) sugere que esta é uma técnica adequada para estudar as variações nas características escolares dos indivíduos que fazem parte de grupos (por exemplo, neste caso, escolas e turmas). Portanto, quando os dados são estruturados dessa forma, esses modelos são mais apropriados, pois incorporam essas informações em relação ao aninhamento, ou seja, consideram que as fontes de variabilidade se encontram aninhadas.

Essas estruturas hierárquicas de dados são caracterizadas por estruturas observacionais agrupadas que, por sua vez, podem estar agrupadas em outras estruturas ainda maiores e assim sucessivamente por vários níveis, reproduzindo formato de dados em "árvore". Esses agrupamentos tomam a forma de níveis no modelo; o primeiro nível descreve as características individuais do objeto de estudo; tem, além disso, outros níveis (aninhados dentro dos grupos) que agrupam informações dos coletivos em questão.

O MLH permite que cada um dos níveis seja especificado separadamente e, depois, reunidos em um único modelo. Como as amostras também são selecionadas com uma estrutura hierárquica, a necessidade de uma análise que reconheça essa estrutura aninhada, pois não há independência das observações dentro de cada nível (HOX, 1995).

Para Macedo Alves (2004) é importante ter em mente que na análise dos dados de questionários existem advertências metodológicas que devem ser consideradas.

“Algumas das mais relevantes críticas ao debate sobre a magnitude do impacto dos fatores associados ao rendimento escolar se deram no âmbito metodológico. Essas passavam pelo questionamento da qualidade das informações socioeconômicas utilizadas pelos estudos. Em pesquisas como a “Avaliação de desempenho: fatores associados” e o Sistema Nacional de Avaliação da Educação Básica, o SAEB, essas informações são reportadas pelos estudantes que fazem os testes de rendimento, crianças de 7 a 14 anos, que normalmente não são os melhores respondentes para a mensuração deste tipo de variável. Também foram bastante discutidas as abordagens dos modelos, considerando principalmente a análise do coeficiente de explicação do modelo, o R^2 , ressaltado pela literatura como ultrapassado. Já a análise da “escola eficaz” se pautou na crítica metodológica enfatizando a importância do uso dos modelos hierárquicos (ou multinível) para contemplar a estrutura dos dados de rendimento. Essa nova ótica dos estudos, utilizando modelos hierárquicos, retomou o debate sobre o impacto dos fatores escolares e familiares na determinação do rendimento escolar” (MACEDO ALVES, 2004: 4)

Para Soares, César e Mambrini (1997) a utilização de um modelo que incorpora a estrutura hierárquica presente nos dados apresenta as seguintes vantagens: (1) obtenção de melhores estimativas para os parâmetros relativos a unidades específicas, pois por meio do modelo hierárquico é possível obter uma equação para cada escola, fazendo assim uso de toda a informação presente na amostra de forma eficiente; (2) possibilidade de formular e testar hipóteses relativas a efeitos entre níveis; (3) partição da variância em componentes, ou seja, para um modelo com três níveis, tem-se a variância do nível 1 (estudantes), variância do nível 2 (turmas) e variância do nível 3 (escolas). Para este estudo são considerados apenas dois níveis, estudantes e escolas, já que uma vez selecionada a escola se aplica a prova somente em uma turma.

Outra vantagem do uso da análise multinível, neste caso, é que ela possibilita estudar, ao mesmo tempo, os efeitos de variáveis individuais e de grupo nos resultados individuais e, além disso, incorpora a consideração da dependência das observações dentro de cada grupo (DEZ ROUX, 2002).

Desta forma, a utilização de modelos multiníveis permite obter melhores estimadores dos coeficientes de regressão e sua variação, em comparação com os modelos tradicionais (DE LA CRUZ, 2008). Além disso, os testes estatísticos padrões dependem fortemente da suposição de independência das observações e, se isso for violado (como nas estruturas multiníveis), as estimativas de erros-padrão dos testes estatísticos convencionais são muito menores e seus resultados espúrios (HOX, 1995).

Os sistemas escolares são um exemplo típico de estrutura hierárquica, pois, os estudantes estão agrupados em turmas, as turmas agrupadas em escolas (Soares e Mendonça, 2003). Portanto, esse tipo de análise é conveniente quando se estudam os determinantes do desempenho educacional (HOX, 1995; BRYK e RAUDENBUSCH, 1988). Quando informações sobre escolas e estudantes estão disponíveis, um modelo hierárquico de dois níveis é proposto. No TERCE há dados de dois níveis: ao primeiro nível (Nível 1: Estudantes), e o segundo nível (Nível 2: Escola), considerando ainda, que a escola, potencialmente, exerce influência sobre os estudantes.

Finalmente, deve-se notar que os MLH podem incorporar efeitos fixos e aleatórios. A forma mais simples é que a variabilidade entre os grupos (efeitos aleatórios) é evidenciada apenas no intercepto, mas formas mais complexas aparecem quando esta variabilidade é incorporada nos coeficientes das variáveis explicativas.

Dado que os MLH têm níveis diferentes, existem variáveis explicativas em cada uma delas. Por exemplo, se considerarmos o caso do desempenho escolar em que existem dois níveis: escolas e estudantes, haverá algumas variáveis que caracterizarão as escolas (nível 2) e outras os estudantes (nível 1).

As variáveis do nível 2 serão idênticas para todos os estudantes de uma determinada escola, portanto, essas variáveis só influenciarão os interceptos das

escolas. Variáveis de nível 1 podem ser incorporadas com efeitos fixos ou aleatórios, dependendo da teoria referente ao objeto de estudo e aos objetivos da pesquisa (OECD, 2003). Se uma variável de nível 1 é incorporada com efeitos fixos, assume-se que não há diferença entre as escolas quanto ao efeito da dita variável na variável dependente. Por outro lado, se uma variável de nível 1 é incluída atribuindo efeito aleatório, está sendo considerado que seu efeito na variável dependente difere entre as escolas.

No caso do desempenho escolar, se os efeitos são fixos, os coeficientes de regressão fornecem informações sobre o que acontece com o resultado educacional — em uma escola — quando o valor de uma variável explicativa é alterado (*ceteris paribus*). Como os efeitos são fixos, esse coeficiente é o mesmo para todas as escolas (OECD, 2003).

Da mesma forma, se os efeitos forem aleatórios, os coeficientes serão interpretados da mesma maneira, mas terão um valor diferente em cada escola. Esses coeficientes podem ser divididos em duas partes, uma fixa, compartilhada por todas as escolas, e uma aleatória, que representa a distância residual entre o coeficiente de cada escola e a parte comum (OECD, 2003).

Formalmente, assumindo que existe um modelo de desempenho educacional composto por dois níveis e que incorpora três variáveis explicativas, uma no nível 2 (P) e duas no nível 1 (X representando a variável de efeito fixo e Z a de efeito aleatório), o modelo pode ser escrito da seguinte forma:

- Nível 1:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + \beta_{2j}Z_{ij} + r_{ij} \quad (1)$$

- Nível 2:

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

$$\beta_{1j} = \gamma_{10} \quad (3)$$

$$\beta_{2j} = \gamma_{20} + v_{2j} \quad (4)$$

- Modelo Completo:

$$Y_{ij} = \gamma_{00} + \gamma_{01}P_j + \gamma_{10}X_{ij} + \gamma_{20}Z_{ij} + v_{0j} + v_{2j}Z_{ij} + r_{ij} \quad (5)$$

em que:

- Y_{ij} é o resultado educacional (esperado) do estudante "i" na escola "j";
- β_{0j} é o intercepto da regressão para a escola "j";
- β_{1j} é o coeficiente que acompanha a variável explicativa X, que é incorporada no modelo com efeito fixo; portanto, β_{1j} , é o mesmo para todos as escolas é representado por γ_{10} ;
- β_{2j} é o coeficiente que acompanha a variável explicativa Z, que é incorporada no modelo com efeito aleatório; portanto, varia entre as escolas. Consiste em uma parte fixa (γ_{20}) e uma parte aleatória (v_{2j}). Esta última representa a distância residual do coeficiente da escola j em relação ao coeficiente médio de regressão (γ_{20});
- γ_{01} é o coeficiente que acompanha a variável explicativa P (na situação colocada, o primeiro e único do nível 2. Como pertence ao nível 2, existe um valor de P para todas as escolas j. Como pode ser visto na equação, esta variável influencia o valor do intercepto β_{0j} ;
- r_{ij} é a variância residual dentro de cada escola;
- v_{0j} é a variância residual entre escolas.

4.3.1. Modelo Nulo

Um conceito importante usado em modelos com estrutura hierárquica aplicados a dados educacionais é o chamado *efeito escola*. Como dito anteriormente, uma das grandes vantagens desta metodologia é permitir a avaliação do efeito individual e do efeito do grupo. Desta forma, o efeito escola é usado para descrever o impacto da instituição de ensino sobre o desempenho dos seus

estudantes e pode ser calculado a partir das componentes de variação devida a cada um dos níveis do modelo.

Na análise multinível é útil estimar um modelo sem incluir variáveis explicativas, o que é chamado de Modelo Nulo. Isso fornece informações sobre que proporção da variabilidade dos resultados do desempenho é devida às diferenças entre escolas (*between schools*) e que proporção está ligada a desigualdades dentro dela (*within schools*). Formalmente, o modelo é expresso da seguinte forma:

$$Y_{ij} = \gamma_{00} + v_{0j} + r_{ij} \quad (6)$$

onde:

γ_{00} representa o efeito fixo ou determinístico (intercepção global).

r_{ij} e v_{0j} representam efeitos aleatórios ou estocásticos

Neste caso, os interceptos de cada escola (β_{0j}) são iguais às médias das escolas ou muito próximos deles. β_{0j} é formado por uma parte fixa — comum a todas as escolas (γ_{00}) — e uma parte aleatória (v_{0j}), que representa o desvio da escola j de γ_{00} . Deve-se notar que γ_{00} é o resultado do conjunto de escolas (média) e é chamado de *intercepto*. Então, v_{0j} é o desvio da escola j da média do conjunto de escolas e representa a variância entre elas.

Por outro lado, r_{ij} é o desvio do resultado do indivíduo i em relação à média da escola j à qual pertence. Como nesse modelo cada estudante recebe a média de sua escola como uma pontuação prevista, r_{ij} é igual à variação dentro de cada escola.

Então, quando não há variáveis explicativas no modelo, as variâncias residuais *intraescolar* e *interescolar* são iguais às estimativas de variâncias *intraescola* e *interescolar* (OECD, 2003).

Normalmente existem dois tipos de índices relevantes na análise multinível: os coeficientes de regressão e a decomposição da variância entre os diferentes níveis. Um indicador comumente utilizado chamado *Coefficiente de Correlação*

Intraclasse (ρ) que representa a proporção da variância residual explicada pelas diferenças entre as escolas ($\rho = v_{0j}/(v_{0j} + r_{ij})$).

No entanto, é interessante conhecer a variância explicada, que pode ser obtida comparando o modelo chamado de final (ajustado com variáveis de níveis 1 e 2) e o modelo nulo. A maneira de calcular a variância é explicada no Quadro 4.2.

Quadro 4.2. Proporção de variância explicada em relação ao Modelo Nulo.

Total	$1 - (v_{0j} + r_{ij}) \text{ Modelo final} / (v_{0j} + r_{ij}) \text{ Modelo nulo}$
Nível 1 (estudantes)	$1 - (r_{ij}) \text{ Modelo final} / (r_{ij}) \text{ Modelo nulo}$
Nível 2 (escolas)	$1 - (v_{0j}) \text{ Modelo final} / (v_{0j}) \text{ Modelo nulo}$

Fonte: elaborado com base em informações da OCDE (2003).

Para Soares e Candian (2007), a denominação mais adequada do *Coefficiente de Correlação Intraclasse* seria de “Índice de dissimilaridade entre escolas”, pois valores próximos de 1 indicam que as escolas são muito diferentes em termos do desempenho de seus estudantes, ou seja, o efeito escola é elevado, enquanto valores próximos de zero indicam que as escolas são similares entre si.

4.4. USO DOS VALORES PLAUSÍVEIS EM AVALIAÇÕES EM LARGA ESCALA

A avaliação é uma atividade necessária e fundamental durante o processo educacional. Constitui uma atividade altamente complexa, que pode ser orientada para diferentes objetivos e executada em diferentes níveis, desde uma sala de aula até o sistema educacional como um todo.

Nesse sentido, o uso de dados de avaliação em larga escala é cada vez mais comum para descrever as habilidades e conhecimentos dos estudantes. Assim, com a crescente influência das avaliações em larga escala sobre as políticas públicas educacionais, a discussão sobre a precisão das inferências para a população de referência torna-se altamente relevante.

Para medir o desempenho escolar, deve-se ter em mente que o que se deseja é obter estimativas mais precisas de características no nível do estudante ou da população. Quando o objetivo é o segundo, então as estimativas pontuais usuais para a habilidade dos estudantes (EAP ou WLE) podem dar resultados enviesados em alguns grupos de indivíduos (Laukaityte e Wiberg, 2017).

Reis Costa (2017) diz que diferentemente de exames como o Exame Nacional do Ensino Médio (ENEM), cujo foco são resultados individuais, as avaliações educacionais em larga escala, como o PISA, TIMSS, PIRLS ou TERCE, não fornecem resultados individuais aos avaliados, pois são planejadas para se ter um diagnóstico mais preciso ao nível nacional ou de subgrupos específicos da população.

Nesses estudos, o cálculo da proficiência dos estudantes é baseada em valores plausíveis, que representam o intervalo de habilidades em que se localiza a verdadeira proficiência do estudante. Ao invés de calcular diretamente a proficiência, são gerados diferentes valores plausíveis para a habilidade do estudante em cada escala de proficiência. Os resultados são agregados para se obter as estimativas finais e seus respectivos erros padrão.

4.4.1. Definição dos valores plausíveis

Em termos gerais, pode-se dizer que o procedimento de valores plausíveis consiste em uma abordagem de múltipla imputação. A base teórica para a imputação de valores plausíveis foi desenvolvida por Rubin (1987) e aplicada em uma avaliação em larga escala por Mislevy e outros (MISLEVY, 1991; MISLEVY, JOHNSON e MURAKI, 1992; MISLEVY, BEATON, KAPLAN e SHEEHAN, 1992; BEATON e JOHNSON, 1992; MISLEVY, 1993; VON DAVIER, GONZALEZ e MISLEVY, 2009).

“The simplest way to describe plausible values is to say that plausible values are a representation of the range of abilities that a student might reasonably have. (...). Instead of directly estimating a student’s ability θ , a probability distribution for a student’s θ , is estimated. That is, instead of obtaining a point estimate for θ , (like a WLE), a range of possible values for a student’s θ , with an associated probability for each of these values is estimated. Plausible values are random draws

from this (estimated) distribution for a student's θ ." (WU e ADAMS, 2002)

É importante destacar que os valores plausíveis são valores aleatórios desenhados a partir da distribuição de escores que podem ser razoavelmente atribuídos para cada indivíduo. Não são notas e, portanto, não devem ser tratados como escores individuais. Os valores plausíveis não são estatísticas adequadas para estimar parâmetros individuais.

A metodologia dos valores plausíveis foi introduzida pela primeira vez para a *National Assessment of Education Progress Data* (NAEP) dos Estados Unidos 1983-1984. Em avaliações internacionais de larga escala, desde a aplicação do terceiro ciclo do TIMSS, conduzidas pela IEA em 1995, a estimativa do desempenho dos estudantes é feita através de valores plausíveis.

É comum a utilização de diversos escores em avaliações como PISA (10 valores plausíveis) ou NAEP (20 valores plausíveis). Por exemplo, o Relatório Técnico do TERCE (UNESCO, 2016) indica que estimativas com esses valores devem ser realizadas cinco vezes, uma para cada valor plausível relevante para a análise. Para todos os estudantes das escolas amostradas, a base de dados do TERCE continha cinco valores plausíveis para a avaliação de cada área.

O método dos valores plausíveis permite estimativas melhores da variabilidade dos dados e, conseqüentemente, inferências melhores sobre os parâmetros populacionais de interesse:

"...the sets of plausible values give a more realistic representation of the expected values in subgroups as well as of the variances within actual results for these subgroups. Because the true values are in most cases unknown, and because we do not have accurate estimates of individual performance on short tests, plausible values are a very useful tool for generating values that have more accurate statistical properties than do observed scores for subgroup comparisons."(von Davier, Gonzalez e Mislevy, 2009, p. 22)

Variabilidades subestimadas podem, por exemplo, levar a conclusões que médias populacionais são diferentes quando, na verdade, elas são iguais, portanto,

o método apresentado é mais adequado nas comparações entre subpopulações avaliadas em testes em larga escala.

4.4.2. O Método de Imputação Via Valores Plausíveis

Uma grande preocupação de qualquer pesquisador na coleta dos dados de pesquisas de campo refere-se ao número de casos faltantes (missing data) na amostra. A retirada total desses casos pode comprometer ou invalidar os resultados do estudo. É comum, no entanto, a utilização de técnicas de imputação de dados para atribuir valores de substituição para aqueles em falta, inválidos ou inconsistentes.

A imputação se classifica em: simples (ou única) e múltipla (Rubin, 1987). A imputação simples (ou única) ocorre quando os dados faltantes são substituídos por um único valor (por exemplo, a média ou a mediana dos dados), enquanto na múltipla cada dado faltante é substituído por dois ou mais valores imputados de forma a representar a incerteza sobre o valor atribuído. Os valores plausíveis são imputações do último tipo.

Segundo Reis Costa (2017), no caso das avaliações em larga escala, a técnica de imputação deve ser utilizada, uma vez que na montagem dos diferentes cadernos de provas, a estrutura de dados faltantes é especificamente planejada. Os valores plausíveis têm por objetivo medir com maior precisão características populacionais, considerando que os estudantes foram avaliados por apenas um subconjunto de itens de uma gama maior de itens analisados.

Na maioria das avaliações em larga escala, não se pede a cada estudante que responda ao conjunto completo de todos os itens que foram projetados para esse teste, mas apenas uma parte. Caso contrário, o tempo necessário para completar o teste seria extremamente longo. Portanto, grupos diferentes de estudantes recebem cadernos de teste diferentes, vinculados por alguns itens. Isso permite avaliar uma ampla gama de domínios de habilidades.

“The relatively small number of items per block and the relatively small number of blocks per test booklet mean that the accuracy of measurement at the individual level of these assessments is considerably lower than is the level of accuracy common for individual tests used for diagnosis, tracking, and/or admission purposes. In tests

for individual reporting, the number of items administered is considerably more than the number contained in a typical booklet in a large-scale survey assessment.” (von Davier et al., 2009, p. 11)

Contudo, as abordagens tradicionais para estimar a proficiência individual, como estimativas de *máxima verossimilhança marginal* (MVM) e a *estimação pela média a posteriori* (EAP), são estimativas pontuais adequadas para estudantes individualmente, mas não para estimativas em nível de grupo. Ou seja, essas abordagens resultam, conseqüentemente, em estimativas que podem apresentar viés, quando a estimativa é para um grupo da população (Laukaityt, 2016).

No TERCE, assim como no PISA ou no TIMSS, a geração de valores plausíveis envolve, para a imputação, uma regressão linear com outras variáveis contextuais dos estudantes. Os resultados dos estudantes avaliados são combinados com fatores intervenientes no desempenho dos estudantes: variáveis contextuais, atitudes e sociodemográficas, para proporcionar estimativas do desempenho geral e subpopulações.

Este processo é feito após a calibração conjunta dos itens pela TRI. Os valores plausíveis são extraídos da distribuição preditiva dos valores dos escores calibrados para cada estudante (esse processo é chamado de condicional).

Para melhorar a estabilidade numérica das escalas, as variáveis de contexto originais são padronizadas e transformadas em um conjunto de variáveis linearmente independentes pela extração dos componentes principais das variáveis originais. Esse modelo de regressão baseia-se em um modelo populacional que utiliza respostas de teste e respostas ao questionário de fatores associados dos estudantes como covariáveis. A habilidade não é observada, mas presume-se que depende das respostas do item de teste e das respostas do questionário de fatores associados. Essa variável é tratada como uma variável ausente, seguindo a abordagem de Rubin (1987) e é usada para derivar valores plausíveis.

Assim, em vez de se utilizar as variáveis originais, são usados os componentes principais para se ter informações a priori dos avaliados. A abordagem de valores plausíveis usa as respostas dos estudantes aos itens, juntamente com

todos os dados do estudante, para estimar diretamente as características das populações e subpopulações dos estudantes (YAMAMOTO e KULICK, 2000).

Os valores plausíveis são gerados por múltiplas imputações, usando a distribuição de habilidade e explicando o erro ou a incerteza no nível individual.

O conjunto de dados de valores plausíveis é mais adequado para descrever o desempenho da população do que para estimar a pontuação individual do estudante. Ou seja, os valores plausíveis não devem ser tratados como escores de teste para indivíduos, mas como medidas do desempenho da população.

4.4.3. A Metodologia dos Valores Plausíveis no TERCE

Em termos gerais, pode-se afirmar que a utilidade dos valores plausíveis está no fato de ser uma metodologia que permite estimar as distribuições de proficiência da população e subpopulações para construir estimadores populacionais não viesados.

Nesta metodologia, considera-se a informação contextual dos estudantes e a variância do grupo ao qual pertencem (neste caso, a escola) para tornar mais precisa a estimativa de proficiência baseado em um número limitado de respostas (UNESCO, 2016a).

Como o TERCE é um estudo em larga escala com o objetivo de estudar o desempenho de várias subpopulações, os resultados foram gerados a partir da metodologia de valores plausíveis. Ter conjuntos de valores plausíveis permite obter uma representação mais realista dos valores esperados em diferentes grupos, bem como a variância destes (Chmielewski e Dhuey, 2017).

Dado que a precisão da estimativa da habilidade individual é limitada se for condicionada por apenas um conjunto específico de respostas (neste caso, um único caderno de prova), os valores plausíveis são uma alternativa mais precisa para conhecer as propriedades estatísticas dos grupos de interesse (von Davier et al., 2009).

O Relatório Técnico do TERCE (UNESCO, 2016a) indica que os valores plausíveis podem ser vistos como outra maneira de estimar a capacidade de um estudante, com mais informações do que a proficiência da TRI obtida pela prova.

Os valores plausíveis correspondem a valores imputados da distribuição empírica da proficiência construída com base nas respostas do estudante aos itens que lhe são administrados, além de outras variáveis de contexto, com o objetivo de estudar diretamente as características de uma população ou subpopulação de estudantes.

Matematicamente, os valores plausíveis são amostras aleatórias da distribuição de probabilidade da habilidade de um estudante, condicionadas ao seu padrão de respostas e a algumas variáveis condicionantes.

Ou seja, corresponde a uma amostra aleatória de:

$$h(\theta | x) = \frac{f(x|\theta) g(\theta)}{\int f(x|\theta) g(\theta) d\theta}$$

Em que $g(\theta)$ é a função de densidade da habilidade do indivíduo, que pode depender de algumas variáveis de contexto, geralmente referidas como variáveis de condicionamento, de modo que a distribuição de habilidades é a seguinte:

$$\theta \sim N(X\beta, \sigma^2)$$

É importante notar que ao realizar análises que relacionam o desempenho dos estudantes a outra variável de interesse, usando valores plausíveis, o modelo que gera os valores plausíveis deve conter as referidas variáveis. Neste sentido, o Relatório Técnico do TERCE (UNESCO, 2016) indica que a estimação dos valores plausíveis é feita separadamente para cada país, permitindo assim que os parâmetros de dificuldade do teste, comuns a nível regional, sejam complementados com o efeito específico dessas covariáveis em cada país.

Como foi apontado por Reis Costa (2017), a metodologia dos valores plausíveis contribui para o debate do aperfeiçoamento das técnicas de análise utilizadas nas avaliações educacionais. As medidas de proficiências calculadas na Prova Brasil, por exemplo, são obtidas via EAP, que podem ser vistas com o valor

médio de um conjunto de valores plausíveis para cada estudante, assumindo que essas estimativas foram computadas de uma distribuição *a posteriori* com mesma modelagem da técnica de imputação múltipla.

Como destaca Laukaitytė (2016), se variáveis de contexto não forem consideradas na criação dessas medidas, as estimativas via EAP tendem a se encolher em torno da média populacional, subestimando a variância dos dados, componente essencial para se fazer corretas inferências sobre determinados subgrupos da população.

4.5. O PACOTE 'BIFIESURVEY' DO R

Apesar da importância das técnicas de estimação de variância para correção de erros padrão, um grande número de artigos não usa o método de estimativa de variância corretamente ou não fornecem detalhes sobre seus procedimentos.

Isso pode ser devido em grande parte à dificuldade de usar tais métodos no contexto de projetos de pesquisa complexos, como avaliações em larga escala.

Para facilitar a análise neste tipo de estudo, o *Centro de Pesquisa e Processamento de Dados* da IEA lançou um programa para o pacote de software SPSS chamado "*IDB Analyzer*". O programa incorpora pesos, procedimentos de estimativa de variância baseados em replicação e valores plausíveis, tanto para estudos da IEA como da OCDE. No entanto, o programa só executa análises muito simples, como médias, frequências e regressão OLS, e não pode ser usado para muitas análises mais complexas, como os modelos MLH.

No caso dos valores plausíveis, muitas vezes será possível usar comandos integrados projetados para dados de imputação múltiplos. Caso contrário, é relativamente simples executar cada valor plausível separadamente, calcular as médias e calcular os erros padrão usando fórmulas padrão para a variação da imputação. Procedimentos de estimação de variância com base na replicação podem ser implementados usando comandos de dados de levantamento integrados (por exemplo, "*svy*" do Stata) (Kreuter e Valliant, 2007).

Recentemente, a OCDE lançou o comando "*repest*" para Stata (Avvisati & Keslair, 2014), que calcula pesos replicados e valores plausíveis e é muito mais flexível para programar modelos complexos. No entanto, o "*repest*" pode ser usado

apenas com bancos de dados da OCDE, não com o IEA. (Chmielewski e Dhuey, 2017).

No campo do software livre, o pacote "*BIFIEsurvey*" é a ferramenta mais recente para a análise estatística de avaliações educacionais. São incluídas estatística descritiva, regressão linear e logística, modelos de *path* para variáveis latentes com correção de erro de medição e regressões hierárquicas de dois níveis para amostras ponderadas. A inferência estatística pode ser conduzida para conjuntos de dados com imputação múltipla é particularmente adequada para a análise de valores plausíveis.

Por esses motivos, as estimações e modelos foram gerados com o pacote *BIFIEsurvey* que possibilita a estimação de Modelos Lineares Hierárquicos com valores plausíveis e pesos amostrais para estudantes e escolas.

5. CONSTRUÇÃO DE INDICADORES

A construção das escalas de medida as quais serão indicadores para mensuração dos fatores associados à aprendizagem, são feitas com a exploração das observações disponíveis nas bases de dados do TERCE.

As escalas foram construídas com o emprego da TRI, estabelecidas com média zero e desvio padrão um. Em geral, foi utilizado o modelo de resposta gradual de Samejima (1969).

O processo de estimar os parâmetros dos itens é conhecido como calibração. Neste processo foram feitas as estimativas dos seguintes parâmetros: o parâmetro de discriminação de itens (a) e os parâmetros de dificuldade (b_k), com seus respectivos erros padrões (EP). Esses parâmetros foram analisados para reter apenas os itens que apresentaram valores coerentes com referência na literatura.

Para a construção destes indicadores, foi realizada uma análise preliminar onde foram identificados os itens que, segundo a literatura, são relevantes para a melhor mensuração de cada indicador. Em todos os casos, os itens politômicos foram considerados em todas as suas categorias. Quando algumas opções de resposta apresentaram poucos respondentes, as categorias foram agrupadas. Mesmo em alguns casos específicos, foi feito o agrupamento de categorias desses itens, passando de itens politômicos para itens dicotômicos.

5.1. ÍNDICE DO NÍVEL SOCIOECONÔMICO E CULTURAL DO ALUNO E SUA FAMÍLIA

Uma boa medida do status socioeconômico e cultural da família do estudante, genericamente conhecida em inglês como *family background* é essencial ao analisar dados de avaliações educacionais em larga escala.

A maior parte da literatura mostra a alta associação entre o status socioeconômico e cultural e o desempenho dos estudantes, inclusive tem forte relação com um conjunto de atitudes e fatores relevantes no processo de aprendizagem.

Dada a complexidade desse constructo, ele foi explorado por meio de um indicador que integra: as origens educacionais, econômicas e de trabalho dos pais, bens materiais e o conforto, os livros e as características da construção da casa da família. Hipotetiza-se uma relação direta e positiva entre o nível socioeconômico e a aprendizagem.

Quadro 5.1. Variáveis e questões selecionadas no cálculo do NSE.

Variável	Questões	Categorias
dqfit09_01	Qual o maior nível de escolaridade que o pai do estudante têm?	1: Não tem estudos. 2: Ensino fundamental completo. 3: Ensino médio completo. 4: Tecnólogo ou licenciatura de curta duração. 5: Ensino superior completo. 6: Mestrado ou Doutorado.
dqfit09_02	Qual o maior nível de escolaridade que a mãe do estudante têm?	1: Não tem estudos. 2: Ensino fundamental completo. 3: Ensino médio completo. 4: Tecnólogo ou licenciatura de curta duração. 5: Ensino superior completo. 6: Mestrado ou Doutorado.
dqfit10_01	Em que trabalha regularmente o pai do estudante?	1: Tem um trabalho pago estável ou permanente. 2: Tem trabalho por temporadas. 3: Trabalha ocasionalmente. 4: Estuda, ou não trabalha, ou tem um trabalho não remunerado.
dqfit10_02	Em que trabalha regularmente a mãe do estudante?	1: Tem um trabalho pago estável ou permanente. 2: Tem trabalho por temporadas. 3: Trabalha ocasionalmente. 4: Estuda, ou não trabalha, ou tem um trabalho não remunerado.
dqfit14	De que material é feito a maior parte dos pisos de sua casa?	1: Terra 2: Cimento 3: Tábuas de madeira sem polir 4: Lajota, cerâmica ou similar. 5: Taco, madeira polida ou piso acarpetado.
dqfit15_01	Luz elétrica	1: Sim; 0: Não
dqfit15_02	Água potável	1: Sim; 0: Não
dqfit15_03	Fossa ou esgoto	1: Sim; 0: Não
dqfit15_04	Coleta de lixo	1: Sim; 0: Não
dqfit15_05	Telefone fixo	1: Sim; 0: Não
dqfit15_06	Televisão a cabo	1: Sim; 0: Não
dqfit15_07	Conexão à internet	1: Sim; 0: Não
dqfit16_01	Televisor	1: 0; 2: 1; 3: 2; 4: 3 ou mais

dqfit16_02	Rádio ou aparelho de som	1:0; 2: 1; 3:2; 4: 3 ou mais
dqfit16_03	Computador	1:0; 2: 1; 3:2; 4: 3 ou mais
dqfit16_04	Geladeira	1:0; 2: 1; 3:2; 4: 3 ou mais
dqfit16_05	Lavadora de roupa	1:0; 2: 1; 3:2; 4: 3 ou mais
dqfit16_06	Celular sem conexão à internet	1:0; 2: 1; 3:2; 4: 3 ou mais
dqfit16_07	Celular com conexão à internet	1:0; 2: 1; 3:2; 4: 3 ou mais
dqfit16_08	Veículo motorizado (carro, moto, camionete, etc.)	1:0; 2: 1; 3:2; 4: 3 ou mais
dqfit21	Quantos livros há na casa do estudante? Considere todos os tipos de livro: poesia, romance, dicionários, livros didáticos, etc.	1: Não há livros. 2: Há 10 livros ou menos. 3: De 11 a 20 livros. 4: De 21 a 30 livros. 5: De 31 a 50 livros. 6: Há mais de 50 livros.

Fonte: Elaborado pelo autor com base nos questionários do TERCE (UNESCO 2013).

Partiu-se para a estimação dos parâmetros dos itens (etapa de calibração). Para viabilizar tal procedimento decidiu-se utilizar todas as observações disponíveis. Como critério de exclusão para os itens, foram considerados dois elementos: (i) Parâmetro de discriminação (a) não razoável, que para fins desta análise foi considerado em $< 0,60$ (MENEGON, VICENZI, DE ANDRADE e VINK, 2017), apesar de que é comum encontrar critérios mais exigentes que o estabeleçam em valores mais altos ($a < 0,70$) (PEREIRA, MAGALHÃES e SCHMITT, 2017; BARBETTA, TREVISAN, TAVARES e AZEVEDO, 2014); (ii) Valores de Erros Padrão (EP) maiores que 1. Valores grandes de EP indicam que o item não contém informações suficientes, por exemplo, casos em que há poucas respostas.

Na primeira calibração, observou-se um conjunto de itens que não ajustaram parâmetros aceitáveis de acordo com os critérios estabelecidos e, portanto, foram retirados da calibração final. Por exemplo, numa tentativa inicial constatou-se que o item dqfit10_01 e dqfit10_02 (trabalho dos pais) apresentou problemas de discriminação (no parâmetro a negativo). Tentou-se a recodificação em apenas duas categorias, sendo 1 para “Tem um trabalho pago estável ou permanente” e 0 para “Trabalhos ocasionais” ou similares e também não deu certo.

No entanto, o item dqfit16_06 – Celular com internet – ficou abaixo do critério adotado ($a = 0,45$), indicando pouca discriminação para o NSE. Diante disso, eles foram desconsiderados na construção final do índice NSE.

Tabela 5.1. Estimativas dos parâmetros dos itens e erros padrões associados ao NSE.

Variável	Questões	Parâmetros						
		a (EP)	b1 (EP)	b2 (EP)	b3 (EP)	b4 (EP)	b5 (EP)	b6 (EP)
dqfit09_01	Qual o maior nível de escolaridade que o pai do estudante têm?	1.234 (0.035)	-3.54 (0.107)	-0.132 (0.026)	1.94 (0.050)	2.482 (0.063)	3.333 (0.089)	3.764 (0.107)
dqfit09_02	Qual o maior nível de escolaridade que a mãe do estudante têm?	1.487 (0.037)	-3.368 (0.089)	-0.35 (0.023)	1.524 (0.036)	2.18 (0.049)	3.383 (0.085)	4.092 (0.124)
dqfit14	De que material é feito a maior parte dos pisos de sua casa?	1.368 (0.038)	-3.385 (0.092)	-0.919 (0.029)	-0.874 (0.028)	2.562 (0.062)		
dqfit15_01	Luz elétrica	1.616 (0.147)	-3.521 (0.210)					
dqfit15_02	Água potável	0.987 (0.063)	-3.302 (0.171)					
dqfit15_03	Fossa ou esgoto	0.942 (0.034)	-0.586 (0.033)					
dqfit15_04	Coleta de lixo	0.947 (0.043)	-2.13 (0.083)					
dqfit15_05	Telefonefixo	2.451 (0.074)	0.164 (0.018)					
dqfit15_06	Televisão a cabo	1.221 (0.041)	-0.853 (0.031)					
dqfit15_07	Conexão à internet	3.46 (0.124)	-0.077 (0.016)					
dqfit16_01	Televisor	1.682 (0.040)	-3.054 (0.072)	-0.476 (0.022)	1.054 (0.027)			
dqfit16_02	Rádio ou aparelho de som	1.29 (0.035)	-1.692 (0.043)	1.41 (0.038)	3.078 (0.079)			
dqfit16_03	Computador	1.43 (0.034)	-0.828 (0.028)	0.814 (0.026)	1.766 (0.040)			
dqfit16_04	Geladeira	1.404 (0.051)	-2.671 (0.076)	2.406 (0.066)	4.341 (0.161)			
dqfit16_05	Lavadora de roupa	1.561 (0.049)	-1.263 (0.033)	3.142 (0.083)	4.708 (0.198)			
dqfit16_07	Celular com conexão à internet	1.434 (0.037)	0.253 (0.022)	1.223 (0.032)	2.32 (0.054)			
dqfit16_08	Veículo com motor(carro, moto, camionete, etc.)	0.885 (0.028)	-0.747 (0.037)	1.499 (0.051)	3.646 (0.114)			
dqfit21	Quantos livros há na casa do estudante?	1.366 (0.032)	-2.508 (0.058)	-0.552 (0.025)	0.202 (0.023)			

Fonte: Elaborado pelo autor em base de dados do Uruguai no estudo TERCE (UNESCO 2013).

Os parâmetros dos itens da calibração final são apresentados na tabela 5.1 e no Anexo apresenta se a curva de informação do índice NSE, mostrando em que região de Theta o índice é mais informativo.

5.2. ÍNDICE DE VIOLÊNCIA NO BAIRRO

As características da comunidade ou do bairro onde fica a escola devem estar ligadas à aprendizagem dos estudantes, pois costumam marcar o tipo de relacionamento estabelecido na comunidade e o acesso a oportunidades no nível local. Enquanto o acesso a serviços e capital cultural estaria positivamente associado à aprendizagem, o oposto ocorreria em torno dos níveis de violência e crime na comunidade.

Com esse objetivo, foi criado a partir de itens do questionário familiar, o Índice de Violência no Bairro na percepção da família do estudante (IVB). O procedimento de construção desse índice é o mesmo que o realizado para o índice NSE.

Quadro 5.2. Variáveis e questões selecionadas no cálculo do IVB.

Variável	Questões [No bairro...qual e a chance de acontecerem...]	Categorias	Escala invertida
dqfit19_01	Venda explícita ou consumo de drogas	1: Muito pouco provável. 2: Pouco provável. 3: Provável. 4: Muito provável	Não
dqfit19_02	Convitesde vizinhos para reuniões ou festas	Idem	Não
dqfit19_03	Atos de vandalismo	Idem	Não
dqfit19_04	Oportunidades para compartilhar entre os vizinhos os problemas do bairro	Idem	Sim
dqfit19_05	Brigas entre vizinhos	Idem	Não
dqfit19_06	Desenvolvimento de ações para ajudar algum vizinho com problemas	Idem	Sim
dqfit19_07	Brigas com armas	Idem	Não
dqfit19_08	Organização de festas, jantar ou rifas em que participam muitos vizinhos	Idem	Sim
dqfit19_09	Intercâmbio ou empréstimos de coisas entre vizinhos	Idem	Sim
dqfit19_10	Agressões em que alguém fique gravemente ferido ou morto	Idem	Não
dqfit19_11	Roubos	Idem	Não

Fonte: Elaborado pelo autor com base nos questionários do TERCE (UNESCO 2013).

Depois da primeira calibração foram retirados os itens com capacidade de discriminação muito baixa ($\alpha < 0,6$): dqfit19_02; dqfit19_04; dqfit19_06; dqfit19_08 e dqfit19_09.

Os parâmetros dos itens da calibração final são apresentados na tabela 5.2.

Tabela 5.2. Estimativas dos parâmetros dos itens e erros padrões associados ao IVB.

Variável	Questões	Parâmetros			
		a (EP)	b1 (EP)	b2 (EP)	b3 (EP)
dqfit19_01	Venda explícita ou consumo de drogas	1.832 (0.049)	-0.242 (0.022)	0.472 (0.022)	1.564 (0.037)
dqfit19_03	Atos de vandalismo	2.459 (0.066)	-0.378 (0.020)	0.449 (0.019)	1.453 (0.031)
dqfit19_05	Brigas entre vizinhos	1.664 (0.046)	0.086 (0.022)	1.268 (0.033)	2.394 (0.057)
dqfit19_07	Brigas com armas	3.147 (0.109)	0.672 (0.019)	1.430 (0.028)	2.153 (0.044)
dqfit19_10	Agressões em que alguém fique gravemente ferido ou morto	2.862 (0.095)	0.746 (0.020)	1.511 (0.031)	2.180 (0.046)
dqfit19_11	Roubos	2.173 (0.056)	-0.877 (0.024)	-0.06 (0.019)	1.12 (0.027)

Fonte: Elaborado pelo autor em base de dados do Uruguai no estudo TERCE (UNESCO 2013).

É importante considerar que, do jeito que foi construído, valores positivos no Índice IVB indicam níveis mais altos de violência no bairro.

5.3. ÍNDICE DE SUPERVISÃO DE ESTUDOS E ATIVIDADES EDUCATIVAS NA CASA

As oportunidades de aprendizagem fora da escola são variadas. Este tipo de atividades, em geral, pode ser estudar ou fazer tarefas em casa. Estudos internacionais (BRAY et al., 2013) apresentam grandes diferenças entre os países e dentro deles, em termos dessa dedicação adicional. O índice ISAE tenta sintetizar um conjunto de atividades realizadas pela família, que indica a frequência com que esta participa da supervisão das tarefas e das atividades educativas com o aluno. Os ítemes deste índice são baseados na percepção dos pais no questionário de família.

Quadro 5.3. Variáveis e questões selecionadas no cálculo do ISAE.

Variável	Questões	Categorias
dqfit27	Em geral, por quanto tempo o aluno se dedica para estudar as matérias escolares ou fazer as tarefas em casa?	1: Não estuda depois das aulas; 2: 30 minutos ou menos por dia 3: Mais de 30 minutos, porém menos de uma hora por dia; 4: De 1 a 2 horas por dia; 5: Mais de 2 horas por dia.
dqfit28_01	[Com que frequência...] Lemos juntos um livro ou uma revista	1: Nunca ou quase nunca; 2: Às vezes; 3: Sempre o quase sempre
dqfit28_04	[Com que frequência...] Fazemos as tarefas do colégio juntos	1: Nunca ou quase nunca; 2: Às vezes; 3: Sempre o quase sempre
dqfit28_05	[Com que frequência...] Vamos a museus ou bibliotecas	1: Nunca ou quase nunca; 2: Às vezes; 3: Sempre o quase sempre
dqfit32_01	[Com que frequência...] Asseguro-me de que o estudante fez todas as tarefas escolares	1: Nunca ou quase nunca; 2: Às vezes; 3: Sempre o quase sempre
dqfit32_02	[Com que frequência...] Pergunto ao estudante o que fez na escola	1: Nunca ou quase nunca; 2: Às vezes; 3: Sempre o quase sempre
dqfit32_03	[Com que frequência...] Pergunto ao estudante que notas obteve na escola	1: Nunca ou quase nunca; 2: Às vezes; 3: Sempre o quase sempre

Fonte: Elaborado pelo autor com base nos questionários do TERCE (UNESCO 2013).

Depois da primeira calibração não foi necessário retirar nenhum item pois não apresentaram problemas na estimação dos parâmetros. Os resultados são apresentados na Tabela 5.3.

Pela forma em que foi construído, valores positivos no Índice ISAE indicam níveis mais altos de supervisão e acompanhamento em atividades educativas dos estudantes por parte das famílias.

Tabela 5.3. Estimativas dos parâmetros dos itens e erros padrões associados ao ISAE.

Variável	Questões	Parâmetros				
		a (EP)	b1 (EP)	b2 (EP)	b3 (EP)	b4 (EP)
dqfit27	Em geral, por quanto tempo o aluno se dedica para estudar as matérias escolares ou fazer as tarefas em casa?	0.817 (0.036)	-3.258 (0.134)	-0.896 (0.048)	1.126 (0.054)	4.471 (0.190)
dqfit28_01	Com que frequência...] Lemos juntos um livro ou uma revista	1.29 (0.053)	-1.261 (0.044)	1.969 (0.064)		
dqfit28_04	[Com que frequência...] Fazemos as tarefas do colégio juntos	1.447 (0.058)	-2.407 (0.073)	-0.106 (0.024)		
dqfit28_05	[Com que frequência...] Vamos a museus ou bibliotecas	0.791 (0.051)	1.832 (0.104)	5.189 (0.312)		
dqfit32_01	[Com que frequência...] Asseguro-me de que o estudante fez todas as tarefas escolares	2.106 (0.106)	-3.118 (0.103)	-1.406 (0.039)		
dqfit32_02	[Com que frequência...] Pergunto ao estudante o que fez na escola	1.884 (0.100)	-3.798 (0.157)	-1.704 (0.052)		
dqfit32_03	[Com que frequência...] Pergunto ao estudante que notas obteve na escola	1.048 (0.051)	-3.954 (0.172)	-1.433 (0.059)		

Fonte: Elaborado pelo autor em base de dados do Uruguai no estudo TERCE (UNESCO 2013).

5.4. ÍNDICE DE PRÁTICAS PEDAGÓGICAS DE ENSINO

As evidências mostram que as práticas de ensino, crenças e atitudes dos professores são aspectos fundamentais no processo de aprendizagem (ROCKOFF, 2004; RIVKIN, HANUSHEK e KAIN, 2005; ARAUJO, CARNEIRO, CROSS-AGUAYO e SCHADY, 2014).

O índice desenvolvido, chamado Índice de Práticas Pedagógicas de Ensino (IPPE), procura sintetizar as condições da interação pedagógica, sua atuação na sala de aula e como estrutura as experiências de aprendizagem. Inclui suas intervenções e organização de ensino e aprendizagem.

Os aspectos relevantes e dimensões do IPPE são: comunicação com os estudantes, técnicas de questionamento e discussão utilizadas, motivação para a aprendizagem e uso formativo das avaliações.

Especificamente, os estudantes são consultados sobre a frequência com que algumas atividades acontecem na aula. E um índice de nível 1 (estudante).

Quadro 5.4. Variáveis e questões selecionadas no cálculo do IPPE.

Variável	Questões [Com que frequência destas coisas acontecem em suas aulas?]	Categorias
dqa6it17_02	Prestamos atenção quando os professores falam	1: Nunca ou quase nunca; 2: Az vezes; 3: Sempre ou quase sempre
dqa6it17_07	Os professores estão contentes em dar aula	Idem
dqa6it17_08	Os professores nos cumprimentam quando fazemos algo correto	Idem
dqa6it17_09	Os professores nos motivam para que continuemos estudando	Idem
dqa6it17_10	Os professores me animam quando encontro dificuldade na matéria	Idem
dqa6it17_11	Os professores são simpáticos comigo	Idem
dqa6it17_12	Os professores me escutam com atenção quando faço algum comentário	Idem
dqa6it17_13	Os professores nos explicam com paciência	Idem
dqa6it17_14	Os professores chegam com as aulas bem preparadas	Idem
dqa6it17_15	Os professores têm pronto os materiais que usaremos na sua aula	Idem
dqa6it17_16	Os professores se preocupam em que aproveitemos o tempo ao máximo	Idem
dqa6it17_18	Ao passar para um novo conteúdo, os professores perguntam que coisas já sabemos	Idem
dqa6it17_20	Os professores me perguntam o que entendi e o que não entendi	Idem
dqa6it17_21	Se não entendemos algo, os professores buscam outras formas de explicar	Idem
dqa6it17_22	Se me equivoquei, os professores me ajudam a ver meus erros	Idem
dqa6it17_23	Os professores nos ensinam e nós escutamos e tomamos notas	Idem
dqa6it17_25	Os professores permitem que os alunos expliquem ao resto como se resolve um problema	Idem
dqa6it17_26	Os professores me pedem que explique minhas respostas	Idem
dqa6it17_27	Os professores usam nossos próprios exercícios ou trabalhos para nos explicar como fazê-lo melhor	Idem

Fonte: Elaborado pelo autor com base nos questionários do TERCE (UNESCO 2013).

Depois da primeira calibração não foi necessário retirar nenhum item, porque não houve problema na estimação dos parâmetros. As estimativas dos parâmetros dos itens são apresentados na tabela 5.4.

Tabela 5.4. Estimativas dos parâmetros dos itens e erros padrões associados ao IPPE.

Variável	Questões	Parâmetros		
		a (EP)	b1 (EP)	b2 (EP)
dqa6it17_02	Prestamos atenção quando os professores falam	0.676 (0.050)	-6.106 (0.461)	-0.133 (0.063)
dqa6it17_07	Os professores estão contentes em dar aula	1.211 (0.066)	-3.001 (0.146)	-0.750 (0.049)
dqa6it17_08	Os professores nos cumprimentam quando fazemos algo correto	1.44 (0.073)	-3.016 (0.136)	-0.562 (0.041)
dqa6it17_09	Os professores nos motivam para que continuemos estudando	1.588 (0.090)	-3.019 (0.143)	-1.359 (0.059)
dqa6it17_10	Os professores me animam quando encontro dificuldade na matéria	1.901 (0.089)	-2.306 (0.086)	-0.46 (0.034)
dqa6it17_11	Os professores são simpáticos comigo	1.699 (0.083)	-2.72 (0.112)	-0.557 (0.037)
dqa6it17_12	Os professores me escutam com atenção quando faço algum comentário	1.515 (0.077)	-2.861 (0.126)	-0.718 (0.042)
dqa6it17_13	Os professores nos explicam com paciência	1.776 (0.089)	-2.741 (0.114)	-0.796 (0.040)
dqa6it17_14	Os professores chegam com as aulas bem preparadas	1.687 (0.101)	-3.378 (0.171)	-1.508 (0.064)
dqa6it17_15	Os professores têm pronto os materiais que usaremos na sua aula	1.347 (0.081)	-3.637 (0.200)	-1.385 (0.068)
dqa6it17_16	Os professores se preocupam em que aproveitemos o tempo ao máximo	1.313 (0.076)	-3.21 (0.164)	-1.308 (0.065)
dqa6it17_18	Ao passar para um novo conteúdo, os professores perguntam que coisas já sabemos	1.018 (0.055)	-3.091 (0.158)	0.064 (0.046)
dqa6it17_20	Os professores me perguntam o que entendi e o que não entendi	1.438 (0.076)	-2.884 (0.132)	-0.919 (0.048)
dqa6it17_21	Se não entendemos algo, os professores buscam outras formas de explicar	2.031 (0.110)	-2.903 (0.125)	-1.098 (0.044)
dqa6it17_22	Se me equivoco, os professores me ajudam a ver meus erros	1.862 (0.099)	-2.937 (0.129)	-1.038 (0.044)
dqa6it17_23	Os professores nos ensinam e nós escutamos e tomamos notas	0.902 (0.051)	-2.524 (0.137)	0.723 (0.061)
dqa6it17_25	Os professores permitem que os alunos expliquem ao resto como se resolve um problema	0.759 (0.049)	-3.524 (0.221)	0.79 (0.073)
dqa6it17_26	Os professores me pedem que explique minhas respostas	0.753 (0.049)	-3.691 (0.235)	0.414 (0.063)
dqa6it17_27	Os professores usam nossos próprios exercícios ou trabalhos para nos explicar como fazê-lo melhor	1.08 (0.056)	-2.609 (0.127)	0.227 (0.045)

Fonte: Elaborado pelo autor em base de dados do Uruguai no estudo TERCE (UNESCO 2013).

5.5. ÍNDICE DE DISCIPLINA E GESTÃO DA SALA DE AULA

Em este índice IDGA, os estudantes são consultados sobre a frequência com que algumas atividades acontecem na aula vinculadas a disciplina e a forma como os professores controlam a sala de aula. Os alunos são os que respondem a essas perguntas.

Quadro 5.5. Variáveis e questões selecionadas no cálculo do IDGA.

Variável	Questões <i>[Com que frequência destas coisas acontecem em suas aulas?]</i>	Categorias	Escala invertida
dqa6it17_01	Há barulho e bagunça	1: Nunca ou quase nunca; 2: Às vezes; 3: Sempre ou quase sempre	Sim
dqa6it17_03	Há zombarias entre colegas	1: Nunca ou quase nunca; 2: Às vezes; 3: Sempre ou quase sempre	Sim
dqa6it17_04	Os professores faltam às aulas	1: Nunca ou quase nunca; 2: Às vezes; 3: Sempre ou quase sempre	Sim
dqa6it17_05	Os professores chegam tarde às aulas	1: Nunca ou quase nunca; 2: Às vezes; 3: Sempre ou quase sempre	Sim
dqa6it17_06	Os professores saem mais cedo	1: Nunca ou quase nunca; 2: Às vezes; 3: Sempre ou quase sempre	Sim
dqa6it17_17	Quando passamos de uma atividade para outra, fazemos desordem na sala	1: Nunca ou quase nunca; 2: Às vezes; 3: Sempre ou quase sempre	Sim

Fonte: Elaborado pelo autor com base nos questionários do TERCE (UNESCO 2013).

Depois da primeira calibração os itens mostraram-se consistentes na construção da medida uma vez que apresentaram valores de discriminação bastante superiores ao valor de referência da literatura. Somente foi retirado o item com capacidade de discriminação muito baixa ($a > 0.6$): dqa6it17_06. Os parâmetros dos itens da calibração final são apresentados na tabela 5.5.

Tabela 5.5. Estimativas dos parâmetros dos itens e erros padrões associados ao IDGA.

Variável	Questões	Parâmetros		
		a (EP)	b1 (EP)	b2 (EP)
dqa6it17_01	Há barulho e bagunça	1.318 (0.095)	-0.794 (0.054)	2.494 (0.136)
dqa6it17_03	Há zombarias entre colegas	0.903 (0.069)	-1.147 (0.087)	2.481 (0.166)
dqa6it17_04	Os professores faltam às aulas	0.675 (0.071)	-6.188 (0.630)	-0.564 (0.080)
dqa6it17_05	Os professores chegam tarde às aulas	1.034 (0.098)	-4.221 (0.346)	-1.586 (0.119)
dqa6it17_17	Quando passamos de uma atividade para outra, fazemos desordem na sala	1.607 (0.127)	-0.797 (0.050)	1.862 (0.095)

Fonte: Elaborado pelo autor em base de dados do Uruguai no estudo TERCE (UNESCO 2013).

5.6. ÍNDICE DE BULLYING ENTRE COLEGAS

Em este índice IBUL, os estudantes são consultados sobre algumas situações que acontecem na escola vinculadas ao relacionamento entre seus próprios colegas. Os alunos são os que respondem a essas perguntas.

Quadro 5.6. Variáveis e questões selecionadas no cálculo do IBUL.

Variável	Questões <i>[Com que frequência destas coisas acontecem em suas aulas?]</i>	Categorias
dqa6it19_01	Tenho medo de algum dos meus colegas	1: Sim; 0: Não
dqa6it19_02	Sinto-meameaçado por algum colega	1: Sim; 0: Não
dqa6it19_03	Temo que um dos meus colegas me bata ou me faça mal	1: Sim; 0: Não
dqa6it19_04	Meus colegas fazem zombarias comigo	1: Sim; 0: Não
dqa6it19_05	Meus colegas me deixam sozinho	1: Sim; 0: Não
dqa6it19_06	Meus colegas me forçam a fazer coisas que eu não quero fazer	1: Sim; 0: Não

Fonte: Elaborado pelo autor com base nos questionários do TERCE (UNESCO 2013).

Depois da primeira calibração os itens mostraram-se consistentes na construção da medida uma vez que apresentaram valores de discriminação (parâmetro a) bastante superiores ao valor de referência da literatura. Não foi retirado nenhum item. Os parâmetros dos itens da calibração final são apresentados na Tabela 5.6.

Tabela 5.6. Estimativas dos parâmetros dos itens e erros padrões associados ao IBUL.

Variável	Questões	Parâmetros	
		a (EP)	b1 (EP)
dqa6it17_01	Há barulho e bagunça	3.36 (0.316)	1.494 (0.051)
dqa6it17_03	Há zombarias entre colegas	3.444 (0.333)	1.596 (0.055)
dqa6it17_04	Os professores faltam às aulas	4.032 (0.439)	1.408 (0.046)
dqa6it17_05	Os professores chegam tarde às aulas	1.763 (0.126)	0.984 (0.050)
dqa6it17_17	Quando passamos de uma atividade para outra, fazemos desordem na sala	1.335 (0.109)	1.994 (0.117)

Fonte: Elaborado pelo autor em base de dados do Uruguai no estudo TERCE (UNESCO 2013).

5.7. ÍNDICE DE INFRAESTRUTURA DO LOCAL ESCOLAR

Para o índice INFRA, são consultados os Diretores das escolas. As questões são sobre a infraestrutura física, recursos materiais e didático-pedagógico.

Quadro 5.7. Variáveis e questões selecionadas no cálculo do INFRA.

Variável	Questões <i>[Que instalações tem a escola?]</i>	Categorias
dqdit14_01	Sala para o diretor	1: Sim; 0: Não
dqdit14_02	Salas adicionais (secretaria, administração, etc.)	1: Sim; 0: Não
dqdit14_03	Sala de reunião para os professores	1: Sim; 0: Não
dqdit14_04	Campo ou quadra de esportes	1: Sim; 0: Não
dqdit14_05	Ginásio	1: Sim; 0: Não
dqdit14_06	Sala de computação	1: Sim; 0: Não
dqdit14_07	Auditório	1: Sim; 0: Não
dqdit14_08	Sala de artes e/ou música	1: Sim; 0: Não
dqdit14_09	Enfermaria	1: Sim; 0: Não
dqdit14_10	Laboratório(s) de ciências	1: Sim; 0: Não
dqdit14_11	Biblioteca	1: Sim; 0: Não

Fonte: Elaborado pelo autor com base nos questionários do TERCE (UNESCO 2013).

Depois da primeira calibração os itens mostraram-se consistentes na construção da medida uma vez que apresentaram valores de discriminação (parâmetro a) bastante superiores ao valor de referência da literatura. Somente

foram retirados os itens *dqdit14_04* e *dqdit14_11*. Os parâmetros dos itens da calibração final são apresentados na Tabela 5.7.

Tabela 5.7. Estimativas dos parâmetros dos itens e erros padrões associados ao INFRA.

Variável	Questões	Parâmetros	
		a (EP)	b1 (EP)
dqdit14_01	Sala para o diretor	1.103 (0.103)	-2.363 (0.169)
dqdit14_02	Salas adicionais (secretaria, administração, etc.)	1.595 (0.089)	0.295 (0.036)
dqdit14_03	Sala de reunião para os professores	0.944 (0.061)	0.164 (0.048)
dqdit14_05	Ginásio	6.033 (0.694)	1.148 (0.034)
dqdit14_06	Sala de computação	4.386 (0.374)	0.599 (0.028)
dqdit14_07	Auditório	1.842 (0.102)	0.933 (0.044)
dqdit14_08	Sala de artes e/ou música	2.39 (0.138)	0.832 (0.037)
dqdit14_09	Enfermaria	1.626 (0.136)	2.517 (0.135)
dqdit14_10	Laboratório(s) de ciências	2.571 (0.155)	1.207 (0.043)

Fonte: Elaborado pelo autor em base de dados do Uruguai no estudo TERCE (UNESCO 2013).

5.8. ÍNDICE DE LIDERANÇA DO DIRETOR

Em este índice LID, os professores são consultados sobre algumas questões que tem relação com a avaliação da gestão da equipe de direção da escola e o relacionamento entre os próprios docentes.

Quadro 5.8. Variáveis e questões selecionadas no cálculo do LID.

Variável	Questões <i>[Com que frequência destas coisas acontecem em suas aulas?]</i>	Categorias
dqpit26_01	Quando um professor apresenta um problema da sala de aula, resolve-o em conjunto com o diretor	1: Muito em desacordo 2: Em desacordo 3: De acordo 4: Muito de acordo
dqpit26_02	A equipe da direção promove o intercâmbio de material de ensino entre professores	Idem
dqpit26_03	A equipe da direção incentiva que nós, os docentes, planejem nossas aulas em conjunto	Idem
dqpit26_04	As reuniões de professores são uma instância para nos apoiarmos no nosso trabalho	Idem
dqpit26_05	As reuniões de professores muitas vezes são uma perda de tempo	Idem
dqpit26_06	Nas reuniões de professores se fala de estratégias de ensino e/ou avaliação	Idem
dqpit26_07	A equipe da direção se interessa pelo aperfeiçoamento dos docentes	Idem
dqpit26_08	A equipe da direção fornece todas as facilidades a quem deseja participar de cursos ou aperfeiçoamentos	Idem
dqpit26_09	Os objetivos que esta escola procura alcançar, são conhecidos por todos os professores	Idem
dqpit26_10	Desenvolvo atividades na sala de aula que estão de acordo com os objetivos desta escola	Idem
dqpit27_01	Utiliza métodos eficazes para se manter informada sobre os rendimentos dos estudantes	Idem
dqpit27_02	Lembra-nos permanentemente de nossas metas	Idem
dqpit27_03	Coleta informação que permite conhecer em profundidade o que se passa nesta escola	Idem
dqpit27_04	Coleta informação para fazer melhoras na escola	Idem
dqpit27_05	Escuta as sugestões dos docentes	Idem
dqpit27_06	Explica-nos as razões de suas decisões	Idem
dqpit27_07	Trata-nos como iguais	Idem
dqpit28_01	Visita nossas salas e observa como damos as aulas	1: Nunca; 2: Uma vez por semestre; 3: Mais de uma vez por semestre; 4: Uma vez por mês.
dqpit28_02	Conversa conosco sobre nossa forma de dar aulas	Idem
dqpit28_03	Conversamos sobre a forma como avaliamos os estudantes	Idem
dqpit28_04	Conversamos sobre nossos planejamentos	Idem
dqpit28_05	Conversamos sobre nossa forma de controlar a turma	Idem

Fonte: Elaborado pelo autor com base nos questionários do TERCE (UNESCO 2013).

Depois da primeira calibração os itens mostraram-se consistentes na construção da medida uma vez que apresentaram valores de discriminação

(parâmetro a) bastante superiores ao valor de referência da literatura. Não foi retirado nenhum item.

Tabela 5.8. Estimativas dos parâmetros dos itens e erros padrões associados ao LID.

Variável	Questões	Parâmetros			
		a (EP)	b1 (EP)	b2 (EP)	b3 (EP)
dqpit26_01	Quando um professor apresenta um problema da sala de aula, resolve-o em conjunto com o diretor	3.008 (0.108)	-3.892 (0.162)	-1.374 (0.041)	0.4 (0.027)
dqpit26_02	A equipe da direção promove o intercâmbio de material de ensino entre professores	3.582 (0.125)	-2.339 (0.071)	-1.232 (0.036)	0.307 (0.026)
dqpit26_03	A equipe da direção incentiva que nós, os docentes, planejem nossas aulas em conjunto	1.957 (0.067)	-2.525 (0.082)	-1.411 (0.047)	0.944 (0.039)
dqpit26_04	As reuniões de professores são uma instância para nos apoiarmos no nosso trabalho	2.213 (0.077)	-2.484 (0.082)	-1.635 (0.051)	0.299 (0.030)
dqpit26_05	As reuniões de professores muitas vezes são uma perda de tempo	1.137 (0.048)	-2.871 (0.120)	-1.548 (0.070)	1.189 (0.059)
dqpit26_06	Nas reuniões de professores se fala de estratégias de ensino e/ou avaliação	2.41 (0.084)	-2.627 (0.087)	-1.539 (0.048)	0.94 (0.035)
dqpit26_07	A equipe da direção se interessa pelo aperfeiçoamento dos docentes	3.568 (0.125)	-2.24 (0.064)	-1.26 (0.036)	0.428 (0.026)
dqpit26_08	A equipe da direção fornece todas as facilidades a quem deseja participar de cursos ou aperfeiçoamentos	3.773 (0.128)	-2.134 (0.060)	-1.058 (0.033)	0.448 (0.026)
dqpit26_09	Os objetivos que esta escola procura alcançar, são conhecidos por todos os professores	2.669 (0.095)	-2.642 (0.083)	-1.891 (0.054)	0.275 (0.027)
dqpit26_10	Desenvolvo atividades na sala de aula que estão de acordo com os objetivos desta escola	1.478 (0.065)	-3.987 (0.176)	0.258 (0.036)	
dqpit27_01	Utiliza métodos eficazes para se manter informada sobre os rendimentos dos estudantes	4.465 (0.169)	-2.095 (0.056)	-1.163 (0.033)	0.679 (0.027)
dqpit27_02	Lembra-nos permanentemente de nossas metas	2.257 (0.073)	-2.204 (0.066)	-0.741 (0.033)	1.151 (0.039)
dqpit27_03	Coleta informação que permite conhecer em profundidade o que se passa nesta escola	5.977 (0.252)	-1.891 (0.047)	-0.924 (0.029)	0.652 (0.026)
dqpit27_04	Coleta informação para fazer melhoras na escola	5.153 (0.208)	-2.24 (0.062)	-1.12 (0.032)	0.568 (0.026)
dqpit27_05	Escuta as sugestões dos docentes	3.788 (0.135)	-2.641 (0.079)	-1.204 (0.035)	0.423 (0.026)
dqpit27_06	Explica-nos as razões de suas decisões	2.747 (0.092)	-2.302 (0.068)	-1.352 (0.040)	0.575 (0.030)
dqpit27_07	Trata-nos como iguais	3.301 (0.110)	-2.264 (0.065)	-1.05 (0.034)	0.293 (0.026)
dqpit28_01	Visita nossas salas e observa como damos as aulas	1.537 (0.054)	-1.427 (0.055)	0.023 (0.034)	1.136 (0.047)
dqpit28_02	Conversa conosco sobre nossa forma de dar aulas	1.525 (0.055)	-1.097 (0.048)	0.007 (0.034)	1.057 (0.046)
dqpit28_03	Conversamos sobre a forma como avaliamos os estudantes	1.556 (0.056)	-0.918 (0.044)	-0.105 (0.034)	1.289 (0.051)
dqpit28_04	Conversamos sobre nossos planejamentos	1.523	-1.502	-0.216	1.238

		(0.053)	(0.057)	(0.035)	(0.050)
dqpit28_05	Conversamos sobre nossa forma de controlar a turma	1.540	-1.208	-0.366	0.941
		(0.056)	(0.051)	(0.035)	(0.044)

Fonte: Elaborado pelo autor em base de dados do Uruguai no estudo TERCE (UNESCO 2013).

5.9. ÍNDICE DE CLIMA ESCOLAR

No índice de clima escolar (CLI), os professores são consultados sobre algumas questões que tem relação com as relações entre os professores, diretor, estudantes e família, ou seja, as principais pessoas na escola.

Quadro 5.9. Variáveis e questões selecionadas no cálculo do CLI.

Variável	Questões <i>[Em geral, como você acha que são as relações dentro da escola?]</i>	Categorias	Escala invertida
dqpit22_01	As relações entre os professores	1: Muito ruins; 2: Ruim; 3: Regulares; 4: Boas; 5: Muito boas.	Não
dqpit22_02	As relações entre os professores e o diretor	Idem	Não
dqpit22_03	As relações entre os professores e os pais	Idem	Não
dqpit22_04	As relações entre os professores e os estudantes	Idem	Não
dqpit22_05	As relações entre os estudantes	Idem	Não
dqpit23_01	Contamos com o apoio do diretor a todo o momento	1: Muito em desacordo 2: Em desacordo 3: De acordo 4: Muito de acordo	Não
dqpit23_02	O diretor se preocupa pelos estudantes muito além do seu rendimento escolar	Idem	Não
dqpit23_03	Há um bom relacionamento entre os professores	Idem	Não
dqpit23_04	Sei que posso contar com meus colegas no caso de ter dúvidas ou precisar de orientação	Idem	Não
dqpit23_05	O diretor nos trata com respeito inclusive quando alguma coisa não lhe parece bem	Idem	Não
dqpit23_06	Os estudantes mostram respeito por seus colegas	Idem	Não
dqpit24_01	Um estudante insultou ou ameaçou outro estudante	1: Não aconteceu; 2: Aconteceu algumas vezes; 3: Aconteceu muitas vezes	Sim
dqpit24_02	Um estudante bateu ou machucou outro estudante	Idem	Sim
dqpit24_03	Um estudante foi excluído de um jogo ou atividade realizada pelos colegas	Idem	Sim
dqpit24_04	Um estudante insultou ou ameaçou algum professor	Idem	Sim

dqpit25_01	Posso realizar minhas aulas sem interrupções	1: Muito em desacordo 2: Em desacordo 3: De acordo 4: Muito de acordo	Não
dqpit25_02	Quando estou explicando algo, os estudantes prestam atenção	Idem	Não
dqpit25_03	Gosto muito de dar aulas nesta turma	Idem	Não
dqpit25_04	Os estudantes são agressivos entre si	Idem	Sim
dqpit25_05	Os estudantes frequentemente ajudam àqueles que têm mais dificuldade	Idem	Não
dqpit25_06	Os estudantes mostram respeito por seus colegas	Idem	Não

Fonte: Elaborado pelo autor com base nos questionários do TERCE (UNESCO 2013).

Depois da primeira calibração os itens mostraram-se consistentes na construção da medida uma vez que apresentaram valores de discriminação (parâmetro a) bastante superiores ao valor de referência da literatura. Não foi retirado nenhum item. Os parâmetros dos itens da calibração final são apresentados na Tabela 5.9.

Tabela 5.9. Estimativas dos parâmetros dos itens e erros padrões associados ao CLI.

Variável	Questões	Parâmetros				
		a (EP)	b1 (EP)	b2 (EP)	b3 (EP)	b4 (EP)
dqpit22_01	As relações entre os professores	1.118 (0.053)	-4.45 (0.229)	-2.619 (0.116)	0.147 (0.042)	
dqpit22_02	As relações entre os professores e o diretor	0.958 (0.048)	-4.977 (0.275)	-2.408 (0.113)	0.193 (0.048)	
dqpit22_03	As relações entre os professores e os pais	1.512 (0.063)	-4.588 (0.296)	-3.615 (0.170)	-1.543 (0.058)	1.307 (0.052)
dqpit22_04	As relações entre os professores e os estudantes	1.781 (0.074)	-2.366 (0.085)	0.144 (0.032)		
dqpit22_05	As relações entre os estudantes	2.766 (0.107)	-2.841 (0.107)	-2.047 (0.058)	-1.056 (0.034)	1.154 (0.036)
dqpit23_01	Contamos com o apoio do diretor a todo o momento	0.798 (0.046)	-5.049 (0.298)	-2.767 (0.153)	0.056 (0.053)	
dqpit23_02	O diretor se preocupa pelos estudantes muito além do seu rendimento escolar	0.927 (0.048)	-4.423 (0.238)	-2.858 (0.140)	0.202 (0.049)	
dqpit23_03	Há um bom relacionamento entre os professores	1.011 (0.049)	-5.526 (0.330)	-1.895 (0.090)	1.276 (0.069)	
dqpit23_04	Sei que posso contar com meus colegas no caso de ter dúvidas ou precisar de orientação	0.944 (0.051)	-5.783 (0.373)	-3.741 (0.199)	0.297 (0.049)	
dqpit23_05	O diretor nos trata com respeito inclusive quando alguma coisa não lhe parece bem	0.630 (0.044)	-6.195 (0.448)	-5.026 (0.351)	-0.212 (0.065)	

dqpit23_06	Os estudantes mostram respeito por seus colegas	3.196 (0.124)	-2.367 (0.073)	-0.819 (0.030)	1.281 (0.036)
dqpit24_01	Um estudante insultou ou ameaçou outro estudante	1.744 (0.069)	-1.577 (0.054)	0.781 (0.038)	
dqpit24_02	Um estudante bateu ou machucou outro estudante	1.605 (0.065)	-2.343 (0.083)	-0.018 (0.034)	
dqpit24_03	Um estudante foi excluído de um jogo ou atividade realizada pelos colegas	1.478 (0.062)	-2.406 (0.090)	0.461 (0.038)	
dqpit24_04	Um estudante insultou ou ameaçou algum professor	1.603 (0.073)	-3.289 (0.145)	-0.742 (0.039)	
dqpit25_01	Posso realizar minhas aulas sem interrupções	1.086 (0.047)	-3.052 (0.131)	-0.85 (0.052)	1.482 (0.069)
dqpit25_02	Quando estou explicando algo, os estudantes prestam atenção	1.539 (0.065)	-3.017 (0.121)	-1.492 (0.056)	1.397 (0.054)
dqpit25_03	Gosto muito de dar aulas nesta turma	1.405 (0.059)	-3.918 (0.185)	-2.248 (0.086)	0.255 (0.037)
dqpit25_04	Os estudantes são agressivos entre si	1.856 (0.066)	-0.719 (0.036)	0.683 (0.035)	3.524 (0.147)
dqpit25_05	Os estudantes frequentemente ajudam àqueles que têm mais dificuldade	1.332 (0.060)	-4.493 (0.250)	-2.021 (0.080)	1.294 (0.057)
dqpit25_06	Os estudantes mostram respeito por seus colegas	2.348 (0.087)	-2.57 (0.086)	-0.803 (0.033)	1.23 (0.040)

Fonte: Elaborado pelo autor em base de dados do Uruguai no estudo TERCE (UNESCO 2013).

6. RESULTADOS

O capítulo 6 apresenta os resultados da relação dos fatores associados ao desempenho cognitivo dos estudantes na prova TERCE em Leitura, Matemática e Ciências Naturais.

No trabalho de Barbeta, Andrade e Tavares (2018) destaca-se que, em termos práticos, nos estudos de fatores associados ao desempenho escolar, a metodologia de modelos hierárquicos se destaca, já que existem fatores associados ao estudante e fatores relacionados à escola onde ele estuda, ou seja, há um nível de hierarquia primário – os estudantes – e um nível agregado – as escolas.

Assim, com base nos dados disponíveis do TERCE, serão utilizados modelos MLH de dois níveis: O “nível 1” refere a fatores associados ao estudante, que podem ser imutáveis, como o sexo, a trajetória educativa e raça/cor; ou mutáveis, como nível socioeconômico; contudo essas características estão fora do controle da gestão escolar.

O “nível 2” refere a fatores ligados à escola, como, a rede de ensino, o nível socioeconômico da escola, as práticas pedagógicas e o clima escolar da escola. No entanto, é necessário aqui ser mais específico na definição dos componentes do nível 2. Para Willms (2010) é preciso, porém, diferenciar dentro do nível 2 os fatores composicionais daqueles que são próprios do contexto escolar. Por isso, na análise, essas variáveis serão incorporadas sucessivamente ao modelo.

Como foi introduzido no Capítulo 4, os MLH permitem estimar as variâncias em dois níveis. As variáveis a serem controladas referem-se ao escopo do aluno e sua família (nível 1) e a sala de aula e a escola (nível 2). Deve ser lembrado que apenas uma sala de aula por escola foi tomada, portanto a variabilidade dentro de uma escola não pode ser estimada.

A análise está organizada da seguinte forma: em primeiro lugar, é apresentado o modelo nulo, sem variáveis explicativas, com o objetivo de analisar a influência dos níveis tanto para a proficiência em Leitura quanto em Matemática e Ciências Naturais.

Na sequência foram considerados modelos para a análise de fatores associados com o desempenho escolar, onde foram incorporadas as variáveis explicativas, teoricamente relevantes, relacionadas aos atributos dos estudantes e suas famílias (nível 1) e, em seguida, aquelas relacionadas às características das escolas, professores e diretores (nível 2).

Na última seção, tentaremos demonstrar, com os dados do TERCE, a importância de utilizar o set de valores plausíveis em MLH. Para isso, os resultados dos modelos são comparados de quatro maneiras: i) considerando cada valor plausível individualmente, ii) usando a média dos valores plausíveis, iii) usando o set de valores plausíveis disponíveis e iv) usando o valor único de proficiência estimado por WLE¹¹.

As principais dimensões e indicadores considerados para análise multinível são apresentados nos Quadros a seguir. Para separá-los em níveis o Quadro 6.1 apresenta os fatores do nível 1, isto é, do tipo individual, e o Quadro 6.2 os fatores escolares (nível 2).

¹¹ Warm's Mean Weighted Likelihood Estimates.

Quadro 6.1. Descrição dos indicadores, níveis e efeito esperado na proficiência (sem distinção de disciplina) no modelo empírico. Nível hierárquico ESTUDANTE.

Dimensão	Indicador	Descrição	Efeito esperado	Tipo de variável
I. Características sociodemográficas, econômicas e condições familiares	nina	Gênero feminino	Positivo	Dummy
	nse	NSE: Índice do Nível Socioeconômico e cultural do estudante.	Positivo	Quantitativa
	maypa	Integração familiar: Mora com os pais.	Positivo	Dummy
	traba2	Trabalha (inclui declaração de trabalho remunerado e não remunerado dentro e fora de casa).	Negativo	Dummy
	ivb_def	IVB: Índice de Violência no Bairro na percepção da família do estudante.	Negativo	Quantitativa
II. Trajetórias educacionais e experiências escolares.	repi	Repetência: Estudante repetente pelo menos uma vez	Negativo	Dummy
	ini	Educação pré-escolar: Aluno frequentou regularmente alguma escola ou creche entre 3 e 6 anos	Positivo	Dummy
	falta	Frequência de falta: Com que frequência o estudante faltou à escola nos últimos seis meses? (0) Duas vezes no semestre ou menos; (1) Duas vezes por mês ou mais.	Negativo	Dummy
	bul_def	IBUL: Índice de violência, <i>bullying</i> ou maus-tratos entre colegas na escola.	Negativo	Quantitativa
III. Expectativas e envolvimento dos pais na vida escolar	expect1	Expectativa educacional dos pais: Nível de escolaridade mais alto que os pais acham que o aluno completará: (1) Tecnólogo ou licenciatura de curta duração ou superior. (0) Ensino Médio como máximo.	Positivo	Dummy
	usainfo	Para que você usa a informação que recebe da escola: (0) Apenas para conhecer o progresso do rendimento do estudante na escola; (1) Para apoiar, chamar a atenção ou parabenizar o estudante segundo as notas obtidas.	Positivo	Dummy
	isae_def	ISAE : Índice de Supervisão de Estudos e Atividades Educativas na Casa.	Positivo	Quantitativa
IV. Percepções e experiências em relação ao trabalho do professor.	idga_def	IDGA: Índice de Disciplina e Gestão da sala de Aula.	Positivo	Quantitativa
	ippe_def	IPPE: Índice de Práticas Pedagógicas de Ensino.	Positivo	Quantitativa

Quadro 6.2. Descrição dos indicadores, níveis e efeito esperado na proficiência (sem distinção de disciplina) no modelo empírico. Nível hierárquico ESCOLA.

Dimensão	Indicador	Descrição	Efeito esperado	Tipo de variável
V. Contexto social, escolar e da localização	tamloc3	Tamanho da cidade (segundo a população e divisão político-administrativa)	Positivo	Categórica
	nse_m	Média na turma/escola do Índice do Nível Socioeconômico e cultural do estudante da turma (a)	Positivo	Quantitativa
	maypa_m	Proporção de estudantes da turma que moram com os pais. (a)	Positivo	Quantitativa
	ivb_m	Média da turma para o Índice de Violência no Bairro. (a)	Negativo	Quantitativa
	isae_m	Média da turma para o Índice de Supervisão de Estudos e Atividades Educativas na Casa (a)	Positivo	Quantitativa
	infra	Índice de Infraestrutura do local escolar	Positivo	Quantitativa
	traba2_m	Proporção de estudantes da turma/escola que trabalham (inclui declaração de trabalho remunerado e não remunerado dentro e fora de casa). (a)	Negativo	Quantitativa
	lni_m	Proporção de estudantes da turma que frequentaram regularmente alguma escola ou creche entre 3 e 6 anos (a)	Positivo	Quantitativa
	expect1_m	Proporção de estudantes da turma que seus pais acham que completaram um estudo de Tecnólogo ou licenciatura de curta duração ou superior. (a)	Positivo	Quantitativa
	bul_def	Média da turma para o Índice de violência, bullying ou maus-tratos entre colegas na escola. (a)	Negativo	Quantitativa
VI. Dependência administrativa	tc	Tipo de escola pública: (1) Tempo Integral (0) Tempo parcial	Positivo	Dummy
	priva	Tipo de gestão ou administração da escola: (1) Particular, (0) pública	Positivo	Dummy
VII. Práticas e interação entre docentes e discentes	idga_m	Média da turma para o Índice de Disciplina e Gestão da sala de Aula (a)	Positivo	Quantitativa
	ippe_m	Média da turma para o Índice de Práticas Pedagógicas de Ensino (a)	Positivo	Quantitativa
VIII. Relações interpessoais entre professores e avaliação da gestão do Diretor	lid	Média da turma para o Índice de Liderança do Diretor	Positivo	Quantitativa
	cli	Média da turma para o Índice de Clima Escolar.	Positivo	Quantitativa
IX. Características sociodemográficas dos professores e Diretores	mtro_sex	Gênero do Professor: (1) Mulher; (0) Homem.	Positivo	Dummy
	mtro_exp	Anos de experiência no cargo de Professor	Positivo	Quantitativa
	mtro_antig_esc	Anos de trabalho do Professor nessa escola	Positivo	Quantitativa
	dir_sex	Gênero do Diretor: (1) Mulher; (0) Homem.	Positivo	Dummy
	dir_exp	Anos de experiência no cargo de Diretor	Positivo	Quantitativa
	dir_antig_esc	Anos de trabalho do Diretor nessa escola	Positivo	Quantitativa
	forma	Formação profissional do professor: Você participou de alguma das seguintes atividades de aperfeiçoamento profissional nos últimos 2 anos? (1) Mestrado o curso de especialização; (0) Não realizei nenhuma atividade de aperfeiçoamento profissional ou somente curso de aperfeiçoamento de curta duração.	Positivo	Dummy

Nota: (a) Indica variáveis agregadas para o nível da turma/escola.

É importante dizer que, no entanto, uma vez que apenas uma turma foi pesquisada por escola, foi considerado mais conveniente que as variáveis construídas ao nível de turma passassem a ser remetidas ao nível das escolas, em nível 2.

6.1. DECOMPOSIÇÃO DA VARIÂNCIA NOS NÍVEIS HIERÁRQUICOS: EFEITO ESCOLA

O primeiro passo da análise hierárquica por modelo linear consiste em estimar qual é a proporção da variância total da variável dependente de interesse que pode ser atribuída a cada um dos níveis na análise (RAUDENBUSH e BRYK, 2002).

O Modelo Nulo (M_0), que não tem variáveis explicativas, serve para avaliar a influência de cada nível, ou seja, quanto a escola (nível 2) influencia no desempenho do estudante e a influência do próprio estudante (nível 1) em sua escola.

Assim, a decomposição da variância total de aprendizagem permite-nos conhecer que proporção da variação corresponde a cada nível de análise. O percentual de variância explicada pela escola (CCI) se traduz na parcela de responsabilidade atribuída às características internas das escolas no desempenho dos seus estudantes. Ou, em outras palavras, qual é a magnitude da variação média entre escolas (o que a literatura chama de *between variance*) que se refere às propriedades contextuais e organizacionais da instituição. É, em suma, uma medida de heterogeneidade das escolas, por se tratar de uma medida de grupo (ANDRADE e SOARES, 2008).

De acordo com Fernández (2004), conceitualmente este primeiro resultado é de fundamental importância, pois permite empiricamente distinguir o impacto de vários níveis teóricos na determinação de um resultado individual e, assim, estabelecer um mecanismo de análise para a ligação micro-macro.

O modelo nulo se usa como base para a estimação da variância explicada versus variância não explicada, em comparação aos modelos condicionais estimados a posteriori.

Em primeiro lugar, na Tabela 6.1 confirma-se que as maiores diferenças de desempenho se devem às características individuais dos estudantes. O resultado sugere que a maior parte da variância total corresponde às diferenças entre os indivíduos (estudantes). A variância inter-escolar representa uma proporção importante, embora reduzida. A CCI correspondente à Leitura é de 26%, e muito similar em Matemática e Ciências Naturais.

Os resultados estão dentro do esperado, dado os antecedentes das pesquisas nacionais para o nível de educação básica (ANEP, 2009; 2015a; FERNÁNDEZ, 2009), e bem como em estudos internacionais como o PISA, onde o CCI ainda é maior (WILLMS, 2006; ANEP, 2009; 2013; 2014; FERNÁNDEZ, 2009; FERNÁNDEZ e BLANCO, 2008).

Tabela 6.1. Modelo incondicional (Modelo Nulo) e partição da variância.

	Leitura	Matemática	Ciências Naturais
Variância do nível 2 - Escola ($\sigma^2 v_{0j}$)	3163.6	3151.9	3058.9
Variância do nível 1 - Estudante ($\sigma^2 e$)	8936.4	8894.4	8800.4
Total da variância	12100.0	12046.3	11859.2
CCI: Coeficiente de Correlação Intra-classe	0.2615	0.2616	0.2579

Fonte: Elaborado pelo autor com base nos dados do Uruguai, no estudo TERCE (UNESCO 2013).

Pode-se afirmar que, a princípio, parte importante da variabilidade das proficiências dos estudantes, nestas áreas avaliadas, tem como principal causa as diferenças entre as escolas, o que a literatura chama de efeito escola, o seja, que a variabilidade pode ser explicada por características das escolas. Seu complemento, aproximadamente os 75% restantes da variação no desempenho, deve-se às diferenças entre estudantes dentro das escolas, que, em princípio, podem estar relacionados a fatores de natureza individual.

Porém, segundo Blanco (2008), há uma tendência a interpretar o CCI como um indicador da influência das escolas (nível 2) na aprendizagem dos estudantes. Isto constitui, por duas razões, uma atribuição equivocada. Primeiro, porque não considera que pelo menos parte dessas diferenças entre as médias das escolas se deve às características da seletividade e contextuais, sobre as quais as escolas não têm controle. Segundo, porque não considera que as escolas podem influenciar as

diferenças entre os estudantes, aumentando-as ou diminuindo-as de acordo com os processos escolares. Em palavras do autor,

"Interpretar o CCI como um indicador da influência de escolas peca tanto por excesso e por defeito: exagera a influência direta que as organizações escolares têm na aprendizagem, e não leva em conta que as desigualdades individuais não são "dadas", já que podem ser, pelo menos em parte, o produto dos processos escolares "(BLANCO, 2008: 67)

Como assinala Soares (2004), controlar estatisticamente o efeito das variáveis relacionadas à seletividade e composição escolar é absolutamente necessário, caso contrário, estaríamos atribuindo às escolas um efeito que não é delas, e sim do capital econômico e cultural das famílias dos estudantes.

Para avaliar a validade dessa hipótese, diferentes fatores do contexto socioeconômico e cultural da escola podem ser introduzidos no modelo. Nesse sentido, segundo a literatura, parte importante dessa variação pode ser explicada estatisticamente pela composição socioeconômica das escolas (FERNÁNDEZ e BLANCO, 2004; CERVINI, 2002; 2004, WILLMS, 2006). A composição socioeconômica e cultural da escola, resulta da média das características socioculturais dos estudantes. Ao "subtrair" esse efeito, pode-se estimar com mais precisão qual é a porcentagem da variação na aprendizagem que pode ser atribuída a fatores estritamente escolares.

De acordo com o trabalho clássico de Willms (2010), os fatores escolares associados à proficiência dos estudantes podem ser classificados em, pelo menos, dois tipos principais: os de composição escolar e os que se referem ao contexto. Este último refere-se ao ambiente onde a aprendizagem ocorre. Assim, pode-se dar um passo a mais na análise e incorporar outros fatores contextuais associados.

As seguintes variáveis de fatores contextuais foram introduzidas ao nível da escola: (i) a porcentagem de estudantes que trabalham; (ii) a porcentagem de famílias completas (presença de pai e mãe); (iii) a porcentagem que cursou a educação pré-escolar; (iv) a porcentagem de famílias que têm altas expectativas

sobre o futuro educacional de seus filhos; (v) a média do índice de violência no bairro.

Neste sentido, o Quadro 6.3 descreve as variáveis que são incorporadas ao modelo, seguindo a sequência das variáveis composicionais e contextuais.

Quadro 6.3. Modelos e variáveis composicionais e contextuais para o estudo da de decomposição da variância.

Modelo	Descrição do Modelo	Variáveis	Indicadores
M0	Modelo Nulo	Nenhuma / vazio	
M1	Composição socioeconômica da escola	Média na turma/escola do NSE (a)	nse_m
M2	Composição socioeconômica e fatores contextuais da escola	Média na turma/escola do NSE (a)	nse_m
		Proporção de estudantes na turma/escola que trabalham (a)	traba2_m
		Proporção de estudantes na turma/escola com integração familiar completa (a)	maypa_m
		Proporção de estudantes na turma/escola de estudantes com educação pré-escolar (a)	ini_m
		Proporção de estudantes na turma/escola com famílias com altas expectativas sobre o futuro educacional de seus filhos (a)	expect1_m
M3	Composição socioeconômica no nível escolar e estudante (variáveis socioeconômicas de nível 1 e nível 2) e fatores contextuais ao nível escolar	Média na turma/escola do IVB (a)	ivb_m
		Média na turma/escola do NSE (a)	nse_m
		NSE: Índice de Nível Socioeconômico e Cultural dos estudantes	nse_def
		Proporção de estudantes na turma/escola que trabalham (a)	traba2_m
		Proporção de estudantes na turma/escola com integração familiar completa (a)	maypa_m
		Proporção de estudantes na turma/escola de estudantes com educação pré-escolar (a)	ini_m
		Proporção de estudantes na turma/escola com famílias com altas expectativas sobre o futuro educacional de seus filhos (a)	expect1_m

Fonte: Elaborado pelo próprio autor. Nota: (a) Indica variáveis agregadas para o nível da turma/escola.

Conforme apresentado na Tabela 6.2, pode-se verificar no modelo M1, com inserção unicamente da variável de controle de composição socioeconômica escolar (nse_m), a CCI corrigida é visivelmente reduzida. Ou seja, apenas o efeito composicional do ambiente sociocultural é suficiente para "explicar" estatisticamente, dependendo da área, entre 35% e 59% da variância entre escolas. Ou visto de outra

forma, na média para as três áreas, 15% da variância do desempenho pode ser atribuída, especificamente, às escolas.

Um elemento interessante que emerge dessa análise é que, o valor do CCI parece afetar de maneira diferente conforme a área de avaliação. Por exemplo, na Leitura, o peso do M1 (Composição) é muito importante e o CCI baixa a 11%, enquanto nas Ciências Naturais a redução não é tão sensível (17%). Com apenas essa informação, e de forma exploratória, pode-se supor que o contexto socioeconômico e cultural da escola é mais relevante na proficiência em Leitura do que nas Ciências Naturais.

Como mostrado na Tabela 6.2, a introdução dessas variáveis contextuais (M2) reduz ainda mais a variabilidade atribuível a fatores intra-escolares. É claro, então, que nem todas as diferenças entre os resultados escolares são atribuíveis aos processos e práticas das próprias escolas. Fatores de composição e contextuais representam uma parte muito significativa da variação inicial no nível escolar, deixando relativamente pouco espaço para a ação da escola.

Tabela 6.2. Decomposição da variância, controladas pela composição socioeconômica e cultural da escola e fatores contextuais.

	M0: Nulo	M1: Composição socioeconômica da escola		M2: Composição socioeconômica e Fatores Contextuais da escola		M3: Composição socioeconômica da escola e estudantes (N1 e N2) e Fatores Contextuais da escola	
	Variância	Variância	Redução	Variância	Redução	Variância	Redução
Leitura							
Nível 2 - Escola ($\sigma^2 v_{0j}$)	3163.6	1076.3	66.0%	796.0	74.8%	813.9	74.3%
Nível 1 - Estudante ($\sigma^2 e$)	8936.4	8939.5	0.0%	8941.9	-0.1%	8389.1	6.1%
Total	12100.0	10015.8	17.2%	9737.9	19.5%	9203.0	23.9%
CCI (p)	0.2615	0.1075	58.9%	0.0817	68.7%	0.0884	66.2%
Matemática							
Nível 2 - Escola ($\sigma^2 v_{0j}$)	3151.9	1391.2	55.9%	965.3	69.4%	960.1	69.5%
Nível 1 - Estudante ($\sigma^2 e$)	8894.4	8895.7	0.0%	8917.6	-0.3%	8420.3	5.3%
Total	12046.3	10286.9	14.6%	9883.0	18.0%	9380.4	22.1%
CCI (p)	0.2616	0.1352	48.3%	0.0977	62.7%	0.1024	60.9%
Ciências Naturais							
Nível 2 - Escola ($\sigma^2 v_{0j}$)	3058.9	1787.7	41.6%	1256.2	58.9%	1254.5	59.0%
Nível 1 - Estudante ($\sigma^2 e$)	8800.4	8801.9	0.0%	8805.3	-0.1%	8183.3	7.0%
Total	11859.2	10589.6	10.7%	10061.5	15.2%	9437.8	20.4%
CCI (p)	0.2579	0.1688	34.6%	0.1249	51.6%	0.1329	48.5%

Fonte: Elaborado pelo autor em base de dados do Uruguai no estudo TERCE (UNESCO 2013).

Em suma, surgem duas considerações relevantes. Em primeiro lugar, confirma a hipótese de que as diferenças de desempenho entre escolas se devem, principalmente, às características composicionais dos estudantes. Em segundo lugar, as evidências mostram que entre 8% e 13% das diferenças de desempenho entre as escolas podem ser consideradas estatisticamente independentes da seletividade, composição social dos estudantes e fatores contextuais.

Nesse sentido, os resultados do CCI corrigido mostram que parte da variação dos desempenhos deve-se a um conjunto de fatores não observados e que são propriamente escolares, ou seja, da organização, a gestão e as práticas educativas que operam dentro das escolas.

6.2. AJUSTE DE MODELOS LINEARES HIERÁRQUICOS COM VARIÁVEIS EXPLICATIVAS NOS NÍVEIS INDIVIDUAIS E ESCOLARES

Na seção anterior, diferentes modelos foram introduzidos para mostrar o efeito dos fatores de seletividade, composição e contextuais do nível escolar associados ao desempenho de aprendizagem. Agora, o objetivo desta seção é apresentar os efeitos dos fatores de nível individual, ou seja, aqueles que são característicos dos alunos e suas famílias e dos fatores escolares.

6.2.1. Etapas do ajuste do modelo hierárquico linear

As restrições impostas pelo tamanho das amostras atingem o modo pelo qual as variáveis são inseridas para o ajuste dos modelos: inserir todas ao mesmo tempo causaria um processo iterativo muito longo ou um bloqueio devido à multicolinearidade. Conseqüentemente, é necessário pensar em uma estratégia "fase a fase" em que cada estágio foi constituído por um bloco de variáveis. A incorporação de variáveis por níveis (aluno e escola) foi seguida para esse fim.

Os ajustes dos modelos são feitos em etapas sucessivas, composto de todas as variáveis identificadas como teórica e empiricamente relevantes, conforme descrito no capítulo 3.

Para respeitar a sequência de análise realizada até o momento, em primeiro lugar, são considerados os modelos M0 até M3, ou seja, o modelo nulo (M0) e o modelo que incorpora as variáveis de controle dos efeitos de seletividade, composição e contexto escolar (M3).

O quarto modelo (M4) será ajustado com todas as variáveis do nível individual que a literatura considera relevante. Isso permite que as estimativas controladas sejam mais robustas.

Como critério para a conservação das variáveis no modelo, foi seguido os critérios indicados por Fernández, que afirma que:

"Uma questão importante é determinar os níveis máximos de erro de tipo I que serão aceitos (...) Em termos substantivos o tipo I refere-se

ao erro cometido a aceitar uma teoria explicativa do aprendizado, quando na verdade seus efeitos são zero. As consequências de testes de hipóteses são diametralmente opostas pelo estado da arte no problema analisado. É razoável que um investigador não esteja disposto a descartar uma teoria que está começando a ser explorada em um campo onde há (relativamente) pouco de acumulação. Tendera-se a ser mais "liberal" em seu nível de significância. Pelo contrário, se é para testar uma teoria madura em um campo com um monte de acumulação teórica, é razoável a adotar um alto nível de exigência e um pequeno erro tipo I." (FERNÁNDEZ, 2004: 451)

Uma vez que o ajuste M4 foi feito, um grande número de termos não significantes foi encontrado. Como consequência, adotou-se a estratégia de remover primeiro os termos não significantes com $p > 0,10$ e reestimar o modelo (M5). Um nível máximo de 10% foi adotado devido ao contexto de análise exploratória desta fase.

Em qualquer um dos modelos sucessivos aplicados, além do critério empírico, existem variáveis que, por sua relevância teórica, são mantidas em todos os modelos, embora não atendam aos critérios estatísticos mencionados acima.

Para manter a lógica da análise, elas serão realizadas de forma independente por disciplinas.

6.2.2. Ajuste de modelos em Leitura

A Tabela 6.3 apresenta os resultados dos modelos hierárquicos ajustados para descrever a variação da proficiência em Leitura. São apresentados os Modelo 4 a 6, com as variáveis de controle (nível socioeconômico individual e agrupado por turma/escola) que foram significantes nos modelos do estudo da decomposição da variância (M1 a M3). Assim, M4 incorpora as variáveis relevantes do nível 1 (estudante) e M5 do nível 2 (escola), somente no intercepto (β_{0j}). O Modelo 6 apresenta as variáveis definitivas no modelo final.

No Modelo 4, depois da inserção das 13 variáveis do nível 1, numeradas de [8] até [20] na Tabela 6.3, têm-se resultados que merecem destaque. Para o comentário de cada variável considera-se que todas as demais variáveis permaneçam fixas.

Em termos gerais, 6 das 13 variáveis incorporadas no Modelo 4 apresentaram associação estatisticamente significativa ao 90% de confiança ($p < 0,10$ ou seja um valor absoluto de $T > 1.645$) com o aprendizado em Leitura.

Das variáveis consideradas no modelo de controle (M3), [1] e [2] ainda são estatisticamente significantes (ao 90% de confiança), uma vez que as variáveis do nível 1 são incorporadas.

As variáveis associadas ao desempenho escolar dos alunos que merecem destaque são: Repetência [12], Expectativas educacional dos pais [16] e trabalhar [10]. As variáveis [12] e [10] tem um efeito negativo sobre o desempenho dos alunos em Leitura, mesmo depois da inserção das variáveis nível socioeconômico individual [1] e da escola [2]. Verifica-se que nos estudantes com distorção idade/série, o desempenho médio diminuiu em 60 pontos (mais de meio desvio padrão na escala de proficiência TERCE).

As linhas da tabela com as variáveis [8] a [11] contêm as estimativas feitas para as variáveis relacionadas à Dimensão I: Características sociodemográficas, econômicas e familiares. O ajuste do Modelo 4 mostra que a variável [9] que indica, que o estudante mora com ambos pais, não está estatisticamente associada ao aprendizado na área de Leitura (valor $T < 1.645$).

A variável gênero feminino [8], mesmo quando não apresenta uma forte associação, permanece no modelo. Veremos mais adiante que, com uma melhor especificação, a variável é significativa.

Em relação à Dimensão II, a frequente ausência às aulas [14] não parece ser um fator relevante em termos de aprendizagem. O mesmo acontece com o Índice de bullying entre colegas na escola, na linha [15].

Com relação à terceira dimensão, que se refere às expectativas educacionais dos pais e seu nível de envolvimento, observa-se que os desempenhos de alunos cujos pais declaram ter altas expectativas quanto à educação futura de seus filhos, são estatisticamente superiores aos daqueles que não tem essa expectativa. Pode-se considerar que neste indicador, está contido a projeção educacional de seus filhos, mas também dá conta da importância que a educação apresenta aos pais.

Algo similar acontece com a variável [17], onde se observa que em relação às informações que os pais recebem da escola, aqueles que apoiam a seus filhos parece ser um fator positivamente associado e muito relevante.

Finalmente, os dois índices [19] e [20] criados para dar conta da Dimensão IV sobre percepções e experiências em relação ao trabalho do professor, não parecem estar associados à aprendizagem, quando são controlados pelas outras variáveis.

Como preparação do modelo seguinte se faz o ajuste final para as variáveis significantes do nível 1 ($p < 0.10$). Assim, estas 7 variáveis (6 estatisticamente significantes mais o gênero que se considera teoricamente relevante), sem considerar as ajustadas como controle [1] e [2], são aquelas que serão mantidas para o ajuste do Modelo 5 que incorpora as variáveis do nível 2.

Em seguida, no Modelo 5 (M5), além das variáveis e indicadores do modelo anterior (M4), foram consideradas as variáveis explicativas associadas ao nível 2 (escola). As dimensões e variáveis incluídas nesse modelo foram as descritas no Quadro 6.2. No modelo M5 foram inseridas as variáveis relativas ao nível 2 (escola) somente no intercepto, por simplicidade.

No M5, depois da inserção das 17 variáveis do nível 2, [21] até [37], é possível ressaltar que um número significativo de variáveis da Dimensão V, do contexto social, escola e localização, foram incluídas como variáveis de controle e não são significantes, exceto, conforme indicado pela teoria, a Média na turma/escola do Índice de Nível Socioeconômico e Cultural do estudante.

Em termos gerais, somente dois das dezessete variáveis incorporadas no M5 apresentou associação estatisticamente significativa ($p < 0,10$ ou seja um valor absoluto de $T > 1.645$) com o aprendizado em Leitura [24] e [37].

No modelo final (M6) foi testado o efeito randômico dos coeficientes de inclinação das variáveis do nível do aluno que já haviam entrado no modelo. Em outras palavras, procurou-se verificar se os efeitos dessas variáveis se comportam diferentemente entre as escolas. Verificou-se que as variáveis inseridas no M6 não apresentaram coeficiente de regressão com efeito randômico, ou seja, seus efeitos sobre o desempenho dos estudantes não se comportam diferentemente entre as escolas.

O último passo foi o de verificar efeitos de interação entre variáveis do mesmo nível e entre os níveis (efeito cross-level). Uma interação entre níveis na modelagem multinível refere-se ao efeito conjunto de uma variável de nível micro com uma variável de nível macro. Como assinala Hox (2002), os efeitos da interação são de difícil compreensão e os pesquisadores devem tratar deles com cuidado. Dessa forma, os resultados aqui encontrados devem ser confrontados com resultados de outras pesquisas, a fim de que sejam corroborados ou refutados. Os resultados dessa análise indicam que não foram encontrados efeitos de interação entre variáveis do mesmo nível e entre os níveis.

Tabela 6.3. Ajuste dos modelos hierárquicos em Leitura. Estimativas para os coeficientes, erros-padrão e estatística T.

Variáveis Explicativas	Modelo 4 (M4)			Modelo 5 (M5)			Modelo 6 (M6)		
	<i>Depois da inserção das variáveis do nível 1</i>			<i>Depois da inserção das variáveis do nível 2</i>			<i>Variáveis no modelo final</i>		
EFEITO FIXO	Coef.	E.P.	Est. T	Coef.	E.P.	Est. T	Coef.	E.P.	Est. T
Intercepto	704.47	23.39		720.63	27.27		703.69	12.27	
<i>Variáveis de controle</i>									
[1] Média na turma/escola do NSE ^(a)	40.84	11.05	3.70	41.07	9.72	4.22	34.58	7.98	4.33
[2] NSE	15.52	5.83	2.66	16.33	5.70	2.86	15.98	5.55	2.88
[3] Proporção de estudantes na turma/escola que trabalham ^(a)	-1.98	17.47	-0.11						
[4] Proporção de estudantes na turma/escola com integração familiar completa ^(a)	4.67	22.52	0.21						
[5] Proporção de estudantes na turma/escola de estudantes com educação pré-escolar ^(a)	31.76	27.73	1.15						
[6] Proporção de estudantes na turma/escola com famílias com altas expectativas sobre o futuro educacional de seus filhos ^(a)	-26.75	26.62	-1.00						
[7] Média na turma/escola do IVB ^(a)	-7.80	9.33	-0.84						
<i>Variáveis do nível 1 - Estudante</i>									
Dimensão I: Características sociodemográficas, econômicas e condições familiares									
[8] Gênero feminino	7.91	6.13	1.29	10.90	6.27	1.74	11.13	6.37	1.75
[9] Integração familiar (Mora com os pais)	-6.10	6.97	-0.88						
[10] Trabalha	-21.96	6.99	-3.14	-21.14	6.65	-3.18	-22.98	6.99	-3.29
[11] IVB - Índice de Violência no Bairro na percepção da família do estudante	7.94	4.40	1.81	5.98	3.81	1.57			
Dimensão II: Trajetórias educacionais e experiências escolares									
[12] Repetência	-59.50	10.48	-5.68	-57.92	9.80	-5.91	-60.23	10.16	-5.93
[13] Educação pré-escolar	13.70	6.90	1.98	13.58	7.12	1.91	15.22	6.52	2.33
[14] Falta com frequência à escola	9.58	6.52	1.47						
[15] IBUL - Índice de <i>bullying</i> entre colegas na escola	7.62	4.86	1.57						
Dimensão III: Expectativas e envolvimento dos pais na vida escolar									
[16] Expectativas educacional dos pais	25.99	8.53	3.05	21.46	8.75	2.45	20.48	8.98	2.28
[17] Uso familiar da informação escolar	17.18	8.57	2.00	17.30	8.27	2.09	18.36	8.39	2.19
[18] ISAE - Índice de Supervisão de Estudos e Atividades Educativas na Casa	-2.99	4.63	-0.65						
Dimensão IV: Percepções e experiências em relação ao trabalho do professor									
[19] IDGA - Índice de Disciplina e Gestão da Aula	6.12	5.06	1.21						
[20] IPPE - Índice de Práticas Pedagógicas de Ensino.	1.32	3.48	0.38						
<i>Variáveis do nível 2 - Escola</i>									
Dimensão V: Contexto social, escolar e da localização									
[21] Tamanho da cidade (vs. Grande Montevidéu):									
Capitais departamentais				-10.92	11.11	-0.98			
Cidades não capitais > 5.000 hab.				-7.80	11.79	-0.66			
Cidades não capitais < 5.000 hab., vilas e rurais				2.92	13.05	0.22			
[22] ISAE - Índice de Supervisão de Estudos e Atividades Educativas na Casa (a)				-3.07	11.53	-0.27			
[23] INFRA - Índice de Infraestrutura do local escolar				-1.98	5.32	-0.37			
[24] IBUL - Índice de <i>bullying</i> entre colegas na escola ^(a)				21.10	12.64	1.67			

Variáveis Explicativas	Modelo 4 (M4)	Modelo 5 (M5)	Modelo 6 (M6)
	<i>Depois da inserção das variáveis do nível 1</i>	<i>Depois da inserção das variáveis do nível 2</i>	<i>Variáveis do modelo final</i>
EFEITO FIXO	Coef. E.P. Est. T	Coef. E.P. Est. T	Coef. E.P. Est. T
Dimensão VI: Dependência administrativa			
[25] Escola Particular		8.34 12.76 0.65	
[26] Escola de Tempo Integral		4.63 8.72 0.53	
Dimensão VII. Práticas e interação entre docentes e discentes			
[27] IDGA - Índice de Disciplina e Gestão da Aula ^(a)		6.07 6.68 0.91	
[28] IPPE - Índice de Práticas Pedagógicas de Ensino ^(a)		6.70 10.48 0.64	
Dimensão VIII. Relações interpessoais entre professores e avaliação da gestão do Diretor			
[29] LID - Índice de Liderança do Diretor		-3.46 3.47 -1.00	
[30] CLI - Índice de Clima Escolar		-0.37 3.92 -0.10	
Dimensão IX. Características sociodemográficas dos professores e Diretores			
[31] Gênero do Professor		4.33 10.07 0.43	
[32] Anos de experiência como Professor		-0.54 0.34 -1.62	
[33] Anos de trabalho do Professor nessa escola		0.38 0.55 0.70	
[34] Gênero do Diretor		-8.31 9.40 -0.88	
[35] Anos de experiência como Diretor		0.21 0.30 0.68	
[36] Anos de trabalho do Diretor nessa escola		-0.51 0.52 -0.99	
[37] Formação profissional do professor:			
Curso de aperfeiçoamento		2.09 7.63 0.27	0.49 8.16 0.06
Mestrado e curso de especialização		56.77 30.84 1.84	96.22 48.17 2.00

NOTA: ^(a) Agregação de respostas dos estudantes da turma/escola.

Fonte: Elaborado pelo autor em base a dados da 6ª série do ensino fundamental do Uruguai no estudo TERCE (UNESCO 2013). Estudantes (n=2310). Escolas (n=167)

6.2.3. Ajuste de modelos em Matemática

Para ajustar os modelos em Matemática, o procedimento é semelhante ao anterior. Os Modelos 4 a 6 são apresentados com as variáveis de controle, utilizadas para a análise do ponto 6.1.

O M4 incorpora as variáveis de nível 1 (estudante), M5 adiciona as de nível 2 (escola) e, finalmente, M6 é o modelo definitivo com as variáveis significantes dos níveis 1 e 2. Como critério estatístico geral, são conservadas nos modelos as variáveis que apresentaram associação estatisticamente significativa ao 90% de confiança ($p < 0,10$ ou seja um valor absoluto de $T > 1.645$). Como procedimento final, possíveis efeitos aleatórios e de interação são analisados.

No M4, as 13 variáveis que se referem ao nível 1 são modeladas. Os resultados desse modelo (M4) podem ser observados na Tabela 6.4.

Além das variáveis de controle, [1] e [2], que são estatisticamente significantes, observa-se que o gênero feminino [8] tem efeito significativo (ao 90% de confiança) e negativo. Isso se alinha com as teorias que indicam o efeito regressivo do gênero em relação ao conhecimento matemático.

Em segundo lugar, tal como o modelo de Leitura, repetência [12] e participação no ensino pré-escolar [13], têm efeitos significantes e opostos.

Em terceiro lugar, observa-se que os desempenhos dos alunos cujos pais têm altas expectativas quanto ao futuro educacional [16] apresentam, em média, melhores desempenhos em Matemática.

Mantendo as variáveis significantes do M4, no M5 as variáveis do nível 2 são incorporadas. As 17 variáveis [21] a [37] do nível 2 foram modeladas, observando que sem considerar a variável de controle [1], nenhuma variável incorporada no M5 apresentam associação estatisticamente significativa ao 90% de confiança (valor absoluto de $T > 1.645$).

Mais dois procedimentos foram aplicados ao modelo final (M6). Em primeiro lugar, verificou-se que as variáveis inseridas no M6 não apresentaram coeficiente de regressão com efeito randômico, ou seja, não comportam-se diferentemente entre as escolas.

Os resultados da modelagem indicam, finalmente, que não foram encontrados efeitos de interação entre variáveis do mesmo nível e entre os níveis.

Tabela 6.4. Ajuste dos modelos hierárquicos em Matemática. Estimativas para os coeficientes, erros-padrão e estatístico T.

Variáveis Explicativas	Modelo 4 (M4)			Modelo 5 (M5)			Modelo 6 (M6)		
	<i>Depois da inserção das variáveis do nível 1</i>			<i>Depois da inserção das variáveis do nível 2</i>			<i>Variáveis no modelo final</i>		
EFEITO FIXO	Coef.	E.P.	Est. T	Coef.	E.P.	Est. T	Coef.	E.P.	Est. T
Intercepto	736.94	15.54		772.44	34.21		747.93	10.84	
<i>Variáveis de controle</i>									
[1] Média na turma/escola do NSE (a)	44.52	13.44	3.31	31.07	10.78	2.88	36.43	9.26	3.93
[2] NSE	13.50	4.15	3.25	17.57	4.42	3.97	16.53	3.97	4.16
[3] Proporção de estudantes na turma/escola que trabalham (a)	-0.88	15.43	-0.15						
[4] Proporção de estudantes na turma/escola com integração familiar completa (a)	6.81	21.25	0.25						
[5] Proporção de estudantes na turma/escola de estudantes com educação pré-escolar (a)	22.76	17.30	1.21						
[6] Proporção de estudantes na turma/escola com famílias com altas expectativas sobre o futuro educacional de seus filhos (a)	-16.87	19.26	-0.96						
[7] Média na turma/escola do IVB (a)	-6.98	9.33	-0.77						
<i>Variáveis do nível 1- Estudante</i>									
Dimensão I: Características sociodemográficas, econômicas e condições familiares									
[8] Gênero feminino	-12.52	6.16	-2.03	-12.97	5.83	-2.23	-12.23	5.86	-2.09
[9] Integração familiar (Mora com os pais)	1.84	16.87	0.11						
[10] Trabalha	-12.38	9.21	-1.34	-5.68	4.74	-1.20	-15.85	8.33	-1.90
[11] IVB - Índice de Violência no Bairro na percepção da família do estudante	5.09	5.97	0.85						
Dimensão II: Trajetórias educacionais e experiências escolares									
[12] Repetência	-62.71	9.77	-6.42	-58.25	8.83	-6.60	-59.37	9.28	-6.40
[13] Educação pré-escolar	20.04	7.22	2.78	20.96	6.71	3.12	20.82	6.81	3.06
[14] Falta com frequência à escola	-7.85	7.61	-1.03						
[15] IBUL - Índice de <i>bullying</i> entre colegas na escola	5.57	5.29	1.05						
Dimensão III: Expectativas e envolvimento dos pais na vida escolar									
[16] Expectativas educacional dos pais	28.13	8.46	3.33	25.68	7.78	3.30	25.23	7.82	3.23
[17] Uso familiar da informação escolar	15.03	10.86	1.38						
[18] ISAE - Índice de Supervisão de Estudos e Atividades Educativas na Casa	-4.75	4.90	-0.97						
Dimensão IV: Percepções e experiências em relação ao trabalho do professor									
[19] IDGA - Índice de Disciplina e Gestão da Aula	3.63	5.55	0.65						
[20] IPPE - Índice de Práticas Pedagógicas de Ensino.	7.17	4.69	1.53						
<i>Variáveis do nível 2 - Escola</i>									
Dimensão V: Contexto social, escolar e da localização									
[21] Tamanho da cidade (vs. Grande Montevideú):									
Capitais departamentais				-4.32	9.59	-0.45			
Cidades não capitais > 5.000 hab.				-5.87	10.02	-0.59			
Cidades não capitais < 5.000 hab., vilas e rurais				4.17	10.71	0.39			
[22] ISAE - Índice de Supervisão de Estudos (a)				10.23	15.17	0.67			

Variáveis Explicativas	Modelo 4 (M4)			Modelo 5 (M5)			Modelo 6 (M6)		
	<i>Depois da inserção das variáveis do nível 1</i>			<i>Depois da inserção das variáveis do nível 2</i>			<i>Variáveis do modelo final</i>		
EFEITO FIXO	Coef.	E.P.	Est. T	Coef.	E.P.	Est. T	Coef.	E.P.	Est. T
[23] INFRA - Índice de Infraestrutura do local escolar				3.46	5.73	0.60			
[24] IBUL - Índice de <i>bullying</i> entre colegas na escola (a)				2.39	12.95	0.18			
Dimensão VI: Dependência administrativa									
[25] Escola Particular				-2.34	11.71	-0.20			
[26] Escola de Tempo Integral				-3.70	8.18	-0.45			
Dimensão VII. Práticas e interação entre docentes e discentes									
[27] IDGA - Índice de Disciplina e Gestão da Aula (a)				9.13	9.83	0.93			
[28] IPPE - Índice de Práticas Pedagógicas de Ensino (a)				6.44	7.03	0.92			
Dimensão VIII. Relações interpessoais entre professores e avaliação da gestão do Diretor									
[29] LID - Índice de Liderança do Diretor				-0.71	3.28	-0.22			
[30] CLI - Índice de Clima Escolar				-1.77	4.51	-0.39			
Dimensão IX. Características sociodemográficas dos professores e Diretores									
[31] Gênero do Professor				-1.15	9.04	-0.13			
[32] Anos de experiência como Professor				-0.14	0.41	-0.33			
[33] Anos de trabalho do Professor nessa escola				0.47	0.58	0.80			
[34] Gênero do Diretor				-8.15	12.96	-0.63			
[35] Anos de experiência como Diretor				0.27	0.31	0.88			
[36] Anos de trabalho do Diretor nessa escola				-1.00	0.69	-1.45			
[37] Formação profissional do professor:									
Curso de aperfeiçoamento				4.30	7.19	0.60			
Mestrado e curso de especialização				38.19	32.01	1.19			

(a) Agregação de respostas dos estudantes da turma/escola.

Fonte: Elaborado pelo autor em base a dados da 6ª série do ensino fundamental do Uruguai no estudo TERCE (UNESCO 2013). Estudantes (n=2322). Escolas (n=167)

6.2.4. Ajuste de modelos em Ciências Naturais

Para ajustar os modelos em Ciências Naturais, o procedimento é o mesmo das outras disciplinas. Os modelos são apresentados com as variáveis de controle (M3), utilizadas para a análise da secção 6.1.

Em M4 incorpora-se as variáveis de nível 1 (estudante), em M5 as de nível 2 (escola) e, finalmente, M6 é o modelo definitivo com as variáveis significantes dos níveis 1 e 2. No final, possíveis efeitos aleatórios e de interação são analisados. Os resultados da modelagem são apresentados na Tabela 6.5.

O M4 mostra o ajuste do modelo para as variáveis do nível 1. Elas foram incorporadas, como nas disciplinas anteriores: 13 variáveis que indicam 4 dimensões associadas ao aprendizado.

Em primeiro lugar, observa-se que o gênero feminino [8], como na Matemática, apresenta um efeito regressivo na aprendizagem. Ainda controlada pelas demais variáveis ajustadas, as meninas tem, em média, um desempenho 15 pontos inferior à média dos homens.

Estas informações reforçam as evidências internacionais que mostram que há desigualdades no desempenho acadêmico entre meninos e meninas, que seriam influenciadas pelo tipo de relações de gênero predominantes que permeiam os processos de socialização e educação.

A análise da variável [10] mostra que os estudantes que trabalham (remunerado ou não) possuem 19 pontos a menos que seus pares que não trabalham, mesmo considerando o NSE.

Em relação às variáveis da trajetória educacional (Dimensão II) a repetição escolar [12] é o fator associado ao desempenho mais significativo, após o controle do NSE do aluno.

Nesta mesma dimensão, observa-se que as crianças que frequentaram o ensino pré-escolar [13] obtiveram, em Ciências Naturais e em todas as áreas avaliadas, melhores resultados de aprendizagem nos testes aplicados.

Em relação às expectativas e envolvimento dos pais na vida escolar, dimensão III, observa-se que as variáveis [16] e [17] apresentam efeito direto e significativo (ao 90% de confiança) sobre a proficiência em Ciências Naturais.

É importante destacar que, nas Ciências Naturais, é a única área em que o IDGA - Índice Disciplinar e Gestão da Aula [19] apresenta um efeito positivo e significativo na aprendizagem.

No M5 as variáveis do nível 2 são incorporadas e alguns comportamentos relevantes são observados:

i) A variável do tamanho da localidade onde a escola está localizada [21], torna-se significativa para as escolas das capitais departamentais. Este modelo é o primeiro a mostrar a relevância desse fator. Portanto, neste ponto, deve-se ter cautela sobre sua interpretação, mas os dados sugerem que, quando as escolas estão localizadas em capitais departamentais, exceto Montevideú, o desempenho médio de seus alunos tende a ser menor que a média nacional.

ii) Já observamos que, na área de Leitura, a formação de professores foi um fator relevante, sendo seu efeito particularmente importante entre aqueles com mestrado ou cursos superiores de especialização. Neste caso, são os cursos de aperfeiçoamento [37] que parecem mostrar um comportamento estatisticamente significativo (ao 90% de confiança). Como no ponto i) é importante manter a precaução sobre a interpretação. Assim, no ajuste do modelo definitivo (M6) o fator deixa de ser significativo.

Tabela 6.5. Ajuste dos modelos hierárquicos em Ciências Naturais. Estimativas para os coeficientes, erros-padrão e estatístico T.

Variáveis Explicativas	Modelo 4 (M4)			Modelo 5 (M5)			Modelo 6 (M6)		
	<i>Depois da inserção das variáveis do nível 1</i>			<i>Depois da inserção das variáveis do nível 2</i>			<i>Variáveis no modelo final</i>		
EFEITO FIXO	Coef.	E.P.	Est. T	Coef.	E.P.	Est. T	Coef.	E.P.	Est. T
Intercepto	707.73	11.68		723.87	33.39		723.02	9.47	
<i>Variáveis de controle</i>									
[1] Média na turma/escola do NSE ^(a)	32.41	6.37	5.08	38.02	10.01	3.80	30.85	6.28	4.91
[2] NSE	20.59	4.89	4.21	20.58	5.17	3.98	21.06	4.94	4.26
[3] Proporção de estudantes na turma/escola que trabalham ^(a)	-26.71	17.42	-1.53						
[4] Proporção de estudantes na turma/escola com integração familiar completa ^(a)	-4.42	21.73	-0.20						
[5] Proporção de estudantes na turma/escola de estudantes com educação pré-escolar ^(a)	12.57	9.95	0.26						
[6] Proporção de estudantes na turma/escola com famílias com altas expectativas sobre o futuro educacional de seus filhos ^(a)	32.10	29.26	1.10						
[7] Média na turma/escola do IVB ^(a)	-6.11	7.48	-0.82						
<i>Variáveis do nível 1- Estudante</i>									
Dimensão I: Características sociodemográficas, econômicas e condições familiares									
[8] Gênero feminino	-14.54	6.89	-2.11	-14.80	7.05	-2.10	-14.86	6.82	-2.18
[9] Integração familiar (Mora com os pais)	-2.25	5.77	-0.39						
[10] Trabalha	-18.83	7.58	-2.48	-21.14	7.98	-2.65	-20.43	7.75	-2.64
[11] IVB - Índice de Violência no Bairro na percepção da família do estudante	3.98	4.40	0.90						
Dimensão II: Trajetórias educacionais e experiências escolares									
[12] Repetência	-60.58	8.72	-6.95	-56.47	8.11	-6.97	-58.09	7.83	-7.42
[13] Educação pré-escolar	18.04	8.85	2.04	18.80	8.13	2.31	18.11	7.81	2.32
[14] Falta com frequência à escola	-0.94	9.27	-0.10						
[15] IBUL - Índice de <i>bullying</i> entre colegas na escola	3.86	4.92	0.78						
Dimensão III: Expectativas e envolvimento dos pais na vida escolar									
[16] Expectativas educacional dos pais	28.59	6.85	4.17	27.98	7.12	3.93	28.60	6.69	4.27
[17] Uso familiar da informação escolar	14.27	8.66	1.65	1.91	7.22	0.16			
[18] ISAE - Índice de Supervisão de Estudos e Atividades Educativas na Casa	-5.63	4.42	-1.27						
Dimensão IV: Percepções e experiências em relação ao trabalho do professor									
[19] IDGA - Índice de Disciplina e Gestão da Aula	11.78	4.92	2.39	7.49	5.87	1.28	12.89	4.49	2.87
[20] IPPE - Índice de Práticas Pedagógicas de Ensino.	4.26	4.17	1.02						
<i>Variáveis do nível 2 - Escola</i>									
Dimensão V: Contexto social, escolar e da localização									
[21] Tamanho da cidade (vs. Grande Montevideú):									
Capitais departamentais				-21.41	9.89	-2.16	-14.44	7.99	-1.81
Cidades não capitais > 5.000 hab.				-1.51	10.57	-0.14			
Cidades não capitais < 5.000 hab., vilas e rurais				-9.74	11.32	-0.86			
[22] ISAE - Índice de Supervisão de Estudos ^(a)				-11.16	14.57	-0.77			
[23] INFRA - Índice de Infraestrutura do local escolar				-9.92	6.28	-1.58			

Variáveis Explicativas	Modelo 4 (M4)			Modelo 5 (M5)			Modelo 6 (M6)		
	<i>Depois da inserção das variáveis do nível 1</i>			<i>Depois da inserção das variáveis do nível 2</i>			<i>Variáveis do modelo final</i>		
EFEITO FIXO	Coef.	E.P.	Est. T	Coef.	E.P.	Est. T	Coef.	E.P.	Est. T
[24] IBUL - Índice de <i>bullying</i> entre colegas na escola ^(a)				7.90	12.25	0.65			
Dimensão VI: Dependência administrativa									
[25] Escola Particular				14.03	12.16	1.15			
[26] Escola de Tempo Integral				9.33	8.25	1.13			
Dimensão VII. Práticas e interação entre docentes e discentes									
[27] IDGA - Índice de Disciplina e Gestão da Aula ^(a)				4.62	10.35	0.45			
[28] IPPE - Índice de Práticas Pedagógicas de Ensino ^(a)				9.13	7.61	1.20			
Dimensão VIII. Relações interpessoais entre professores e avaliação da gestão do Diretor									
[29] LID - Índice de Liderança do Diretor				-1.17	3.87	-0.30			
[30] CLI - Índice de Clima Escolar				3.90	4.80	0.81			
Dimensão IX. Características sociodemográficas dos professores e Diretores									
[31] Gênero do Professor				-0.85	11.19	-0.08			
[32] Anos de experiência como Professor				-0.56	0.39	-1.45			
[33] Anos de trabalho do Professor nessa escola				0.60	0.63	0.96			
[34] Gênero do Diretor				-4.57	12.47	-0.37			
[35] Anos de experiência como Diretor				0.38	0.37	1.01			
[36] Anos de trabalho do Diretor nessa escola				-0.59	0.65	-0.90			
[37] Formação profissional do professor:									
Curso de aperfeiçoamento				15.94	7.27	2.19	4.25	5.77	0.31
Mestrado e curso de especialização				28.16	41.21	0.68			

^(a) Agregação de respostas dos estudantes da turma/escola.

Fonte: Elaborado pelo autor em base a dados da 6ª série do ensino fundamental do Uruguai no estudo TERCE (UNESCO 2013). Estudantes (n=2324). Escolas (n=167)

6.3. DISCUSSÃO DOS RESULTADOS DOS MODELOS FINAIS

O Quadro 6.4 foi preparado como síntese para mostrar, de forma panorâmica, os principais fatores associados aos desempenhos de aprendizagem nas três áreas avaliadas. O objetivo é resumir a intensidade e o significado da associação entre os fatores associados à aprendizagem nas diferentes áreas avaliadas.

Para cada área, são considerados os ajustes para os modelos completos (M6), com todos os controles do nível 1 e 2. Em cada quadrante são apresentados os sinais dos coeficientes e seu nível de significância ($p < 0.10$; $p < 0.05$; $p < 0.01$).

Em termos gerais, deve-se ressaltar que, independentemente da área avaliada, há um conjunto de fatores que estão associados, com nuances de intensidade, aos desempenhos dos alunos. Assim, percebe-se que as variáveis de

influência socioeconômica, consideradas como controle, apresentam forte associação aos desempenhos de qualquer das áreas avaliadas.

Quadro 6.4. Ajuste dos Modelos Finais (M6) para as diferentes áreas avaliadas. Resumo conceitual da intensidade e sentido da associação.

	Leitura	Matemática	Ciências Naturais
<i>Variáveis de controle</i>			
[1] Média na turma/escola do NSE (a)	+++	+++	+++
[2] NSE	+++	+++	+++
<i>Variáveis do nível 1- Estudante</i>			
<i>Dimensão I: Características sociodemográficas, econômicas e condições familiares</i>			
[8] Gênero feminino	+	--	--
[10] Trabalha	---	-	---
<i>Dimensão II: Trajetórias educacionais e experiências escolares</i>			
[12] Repetência	---	---	---
[13] Educação pré-escolar	++	+++	++
<i>Dimensão III: Expectativas e envolvimento dos pais na vida escolar</i>			
[16] Expectativas educacional dos pais	++	+++	+++
[17] Uso familiar da informação escolar	++		
<i>Dimensão IV: Percepções e experiências em relação ao trabalho do professor</i>			
[18] IDGA - Índice de Disciplina e Gestão da Aula			+++
<i>Variáveis do nível 2 - Escola</i>			
<i>Dimensão V: Contexto social, escolar e da localização</i>			
[21] Tamanho da cidade			
Capitais departamentais			-
[24] IBUL - Índice de bullying entre colegas na escola (a)	+		
<i>Dimensão IX. Características sociodemográficas dos professores e Diretores</i>			
[36] Anos de trabalho do Diretor nessa escola		-	
[37] A formação profissional do professor			
Mestrado e curso de especialização	++		
Estudantes:	2310	2322	2324
Escolas:	167	167	167

Fonte: Elaborado pelo autor com base em dados da 6ª série do ensino fundamental do Uruguai, no estudo TERCE (UNESCO 2013).

Nota sobre o nível de significância (“+” / “-” p<0.10; “++” / “--” p<0.05; “+++” / “---” p<0.01).

Em primeiro lugar, encontra-se um forte efeito da composição sociocultural da escola, capturado a partir do valor médio na escola do índice de NSE de estudantes que participaram dos testes.

O "efeito ambiental" tem uma magnitude consideravelmente maior do que o status socioeconômico individual (NSE) da família do estudante, um resultado que tem sido repetidamente documentado em estudos anteriores para diferentes países e para diferentes níveis de ensino. Além disso, o efeito composicional é independente e adicional para o NSE de cada aluno.

Embora a importância da composição seja quase indiscutivelmente reconhecida, os mecanismos subjacentes que explicam os impactos da composição sociocultural ou socioeconômica da escola no desempenho individual são menos claros. Uma interpretação está ligada à incidência de interação entre colegas, ou seja, os efeitos dos pares (peer effects) na trajetória escolar pessoal. Como indica De Melo (2011) o efeito dos pares poderia operar de diferentes maneiras (por exemplo, no capital social ou nas áreas de desenvolvimento, etc). Em qualquer caso, convém chamar a atenção para o fato de que a composição escolar, ou seja, as características médias de estudantes que compartilham a escola, são para a escola um "dado" ou atributo fixo, ou seja, não são parte do âmbito de ação da escola e da política educacional. No Uruguai, particularmente no setor público, eles refletem os mecanismos institucionais para designar alunos às escolas, baseados essencialmente em um critério territorial. Nesse sentido, a segmentação escolar é um produto, em grande parte, da segmentação social e residencial.

Em segundo lugar, os resultados mostram o efeito de fatores, como o gênero, que operam de maneiras diferentes, dependendo da área avaliada. Os dados sugerem que as meninas têm melhor desempenho em Leitura e pior em Matemática e Ciências Naturais. Estas informações reforçam as evidências internacionais que mostram que há desigualdades no desempenho acadêmico entre meninos e meninas, que seriam influenciadas pelo tipo de relações de gênero predominantes que permeiam os processos de socialização e educação.

Em terceiro lugar, a condição de executar tarefas que, a partir de sua própria percepção, são consideradas de trabalho, seja dentro ou fora de casa e com

pagamento ou não, tem impacto negativo no desempenho de aprendizagem nas três áreas avaliadas.

Em relação aos aspectos relacionados à trajetória escolar, a repetência escolar e a escolarização no nível pré-escolar nas idades de 4 a 6 anos são dois aspectos relevantes.

Os dados sugerem que, de acordo com a literatura, a repetição é um fator mais poderoso nos resultados da aprendizagem e semelhantes nas três áreas avaliadas. Em outras palavras: o fato de reprovar pelo menos uma série escolar tem efeitos negativos e significantes sobre os desempenhos em todas as áreas avaliadas. Esse efeito negativo associado a eventos de repetição merece considerações de vários tipos. A primeira é que além de ser um efeito estatisticamente significativo, é também de considerável magnitude. De acordo com o ajuste dos modelos completos (M6), a reprovação de pelo menos um ano escolar resulta, em média, em desempenho de, aproximadamente, 60 pontos a menos na escala TERCE (0,6 desvio padrão), mesmo após controlar o resto das variáveis individuais ou escolares.

Uma segunda leitura é que os repetidores constituem, por assim dizer, um grupo auto-selecionado de "maus alunos" (e é por isso que eles teriam repetido de série), então não deveria ser surpreendente que eles obtivessem desempenhos médios menores. Embora este raciocínio seja, em princípio, correto, faltaria mais evidências sobre como funciona um outro conjunto de características extra-acadêmicas, não observadas diretamente, como hábitos de trabalho, esforço, comportamento em sala de aula, o que poderia estar explicando tanto a repetição quanto o desempenho nas provas.

A terceira observação refere-se às implicações da repetição como instrumento de "política educacional" diante da situação dos estudantes que não alcançam o progresso esperado. O forte impacto da repetição no desempenho, mesmo depois de controlar uma ampla gama de fatores, requer pelo menos reconsiderar a eficácia desse instrumento no objetivo final de alcançar uma aprendizagem de qualidade para todos os alunos. A hipótese que com mais tempo

para o ensino o aluno conseguirá igualar o desempenho aos demais alunos, não é verificada em nenhum estudo empírico consultado e nem neste.

Evidentemente, a repetição pode ser interpretada mais como um sintoma de alguma dificuldade de aprendizagem do que como uma experiência educacional na trajetória do aluno. A evidência indica que a mera repetição é geralmente um mecanismo ineficaz para melhorar a aprendizagem dos alunos atrasados, ou seja, insuficiente para obter uma mudança positiva na proficiência e, portanto, é necessário explorar novos mecanismos de apoio e acompanhamento para esse grupo de alunos.

O outro aspecto educacional fortemente associado às proficiências é a escolarização em níveis pré-escolares. As crianças que frequentaram o ensino pré-escolar entre as idades de 4 e 6 anos obtiveram, de forma sistemática em todas as áreas avaliadas, melhores resultados.

Em relação às características institucionais das escolas, os resultados indicam, em primeiro lugar, que, controlando as variáveis de nível 1 e nível 2, os desempenhos dos alunos das escolas particulares não diferem significativamente das escolas da rede pública. Este aspecto, em termos gerais para o caso uruguaio, tem sido documentado desde que existem avaliações padronizadas de aprendizado. A diferença na aprendizagem observada nas escolas particulares deve-se, em maior medida, a um efeito de seletividade socioeconômica dos estudantes do que a questões específicas de gestão escolar.

Além disso, não há efeitos significantes para a categoria Tempo Integral (Tiempo Completo, em espanhol). Esse achado é consistente com os antecedentes nacionais, onde em avaliações em larga escala não existem diferenças significativas entre as escolas em Tempo Integral e os demais tipos de escolas públicas (ANEP, 2010; 2016).¹²

Os dados sugerem que pais com altas expectativas sobre a educação de seus filhos é um fator altamente relevante, independentemente da área avaliada. Foram encontradas evidências sólidas sobre a importância de gerar altas

¹² Um estudo recente sobre o impacto das escolas Tempo Integral (TC) foi capaz de captar, em primeiro lugar, que, no caso da matemática, há um efeito favorável e significativo, ou seja, os estudantes das TC avançaram mais entre o 3º e o 6º do que seus pares de escolas que não pertencem a essa modalidade. Em contraste, na leitura não foram achados impactos específicos associados à modalidade de TC (ANEP, 2016).

expectativas em relação à aprendizagem. A literatura especializada concorda que a capacidade de incutir altas expectativas para todos os alunos é um aspecto fundamental em qualquer processo de aprendizagem.

Por outro lado, há um conjunto de fatores que parecem ter um impacto diferencial nas áreas avaliadas. Nesse sentido, a formação profissional do professor parece ser um fator relevante, uma vez que mostra associação com o desempenho em Leitura e Matemática, mas não com Ciências Naturais. É verdade que a população de professores com estudos pós-terciários é muito escassa, mas além do fato específico sobre a sua escolaridade também pode ser considerado como um indicador relevante sobre a atitude de melhoria e relevância da educação que possui a formação para esses docentes. Nesse sentido, deve ser considerada que a profissionalização do corpo docente é um dos elementos fundamentais da teoria dos "suportes essenciais para o aprimoramento escolar", elaborada por Anthony Bryk (BRYK, 2008; BRYK et al., 2001; BRYK e RAUDENBUSH, 1988).

Esse apoio inclui a capacidade das escolas de recrutar e reter seus professores, oferecer oportunidades de crescimento profissional na escola e desenvolver trabalho de maneira colaborativa. Embora seja verdade que no Uruguai há pouca variação entre as escolas em qualquer um desses aspectos, devido ao escopo limitado de atuação dos diretores na contratação de pessoal docente e os altos níveis de rotação de professores entre as escolas, os dados mostram uma pista onde continuar aprofundando os efeitos desses fatores sobre a aprendizagem.

Outro conjunto de fatores parece afetar algumas disciplinas específicas. Os resultados de alunos, cujos pais parecem utilizar as informações recebidas das escolas para estimular a seus filhos, parecem apresentar melhor desempenho na área de Leitura. Esta tendência não foi observada nas outras áreas avaliadas. Algo semelhante acontece com o IDGA (Índice de Disciplina e Gestão da Aula), onde se observou que é um fator significativo e positivo para os estudantes de Ciências Naturais, e não para as outras áreas.

6.4. TESTE EMPÍRICO SOBRE O USO DE VALORES PLAUSÍVEIS EM MODELOS LINEARES HIERÁRQUICOS.

Nesta seção final, nos concentraremos em mostrar empiricamente as razões que tornam necessário usar o conjunto de valores plausíveis (VP) nos modelos MLH.

O objetivo é demonstrar o papel dos VP em avaliações em larga escala quando modelos multiníveis são executados. Para isso, seguiremos os procedimentos de Laukaityte e Wiberg (2017), onde diferentes estratégias são utilizadas na estimação de médias e variâncias. Vamos tentar mostrar que não considerar o uso do conjunto completo de todos os VP pode levar a estimativas incorretas.

6.4.1. Componentes do erro padrão das estimativas de proficiência nas avaliações em larga escala

Foi discutido no capítulo 4 que as avaliações em larga escala apresentam o desafio de administrar um grande número de itens a um pequeno grupo de alunos em um tempo relativamente curto.

Portanto, é importante não calcular os erros padrão usando a fórmula usual da variância, uma vez que os erros padrão das estimativas das avaliações em larga escala têm dois componentes. O primeiro, é o erro amostral resultante da amostragem de estudantes de uma população em uma amostra estratificada em múltiplos estágios. O segundo, é o erro de medição (especificamente, o erro de imputação) que resulta dos itens efetivamente aplicados a cada estudante na avaliação do universo de itens possíveis.

Para corrigir o primeiro componente, há uma variedade de abordagens, as mais comuns são os procedimentos de estimativa de variação baseados na replicação, como a replicação repetida de *Jackknife* (Jackknife Repeated Replication ou JRR) e a *Replicação de Repetição Balanceada* (Balanced Repeated Replication ou BRR). Todos esses métodos funcionam estimando subamostras do conjunto de dados completos, calculando as estatísticas de interesse para cada subamostra e, em seguida, usando a variância entre subamostras para estimar a variação da amostragem (CHMIELEWSKI e DHUEY, 2017).

O segundo componente do erro é o erro de medição na estimativa de proficiência. A fim de avaliar a proficiência dos alunos de forma confiável, um grande número de itens é necessário. Este é um pré-requisito geral para minimizar o erro de medição. O erro de medição reflete a imprecisão da medida. Em geral, quanto mais extensa for uma prova, em número de itens, mais precisa e confiável é sua medição. As avaliações em larga escala não se destinam a fornecer estimativas precisas individuais porque os erros de medição são relativamente grandes.

Para cada estudante é aplicado apenas uma fração de todos os ítemes (tanto a múltipla escolha quanto a resposta aberta). As diferentes versões de cadernos de prova se vinculam por um bloco comum de itens. É por estas razões que a medição da proficiência individual é feita com erros de medição (von DAVIER et al., 2009).

No entanto, essa abordagem gera alguma incerteza, pois cada aluno responde apenas às perguntas do caderno que ele recebe e não ao grupo inteiro de itens. As respostas dos alunos ao conjunto completo de itens podem ser estimadas, mas como essas estimativas são probabilísticas, há alguma incerteza quanto à validade dos escores finais. Essa incerteza, a variância de medição (também chamada de variância de imputação), é um segundo componente de erro dos resultados das avaliações em larga escala (FOY et al., 2008).

Para refletir a incerteza dessa medida, vários scores ou imputações, chamadas valores plausíveis, são geradas para cada aluno. De acordo com Laukaityte e Wiberg (2017), valores plausíveis têm sido usados com sucesso para melhorar a inferência sobre variáveis latentes em avaliações em larga escala.

Pesquisadores que usam os dados das avaliações em larga escala devem usar a variabilidade entre os valores plausíveis fornecidos para estimar a variância da imputação. Segundo Chmielewski (2017) na literatura mostra diversos exemplos de trabalho que reportados usando todos os valores plausíveis para estimar os erros padrões (AMMERMUELLER, 2013; AMMERMUELLER e PISCHKE, 2009; CHMIELEWSKI, 2014; CHMIELEWSKI et al., 2013; CHUDGAR e LUSCHEI, 2009; GUSTAFSSON, 2016; ROSEN e GUSTAFSSON, 2016; D. RUTKOWSKI et al., 2012; SCHMIDT et al., 2015).

Um pequeno número de trabalhos relatou não usar todos os valores plausíveis, o que produz estimativas viesadas. Vários artigos relatam que usam apenas um dos valores plausíveis, o que produz estimativas imparciais, mas subestima os erros padrão e, portanto, é considerado apropriado apenas para análises exploratórias. Também um pequeno número de artigos parece ter a ideia errônea de que valores plausíveis só devem ser usados quando a proficiência é a variável dependente; no entanto, os valores plausíveis também devem ser usados quando a proficiência é uma variável independente (Chmielewski, 2017).

6.4.2. Comparação de estimativas via modelos lineares hierárquicos com valores plausíveis

O objetivo deste exercício comparativo é mostrar as diferenças nas estimativas quando é realizado com valores plausíveis (VP). A ideia é mostrar a importância de utilizar adequadamente o set de VP em Modelos Lineares Hierárquicos (MLH). Para isso, os resultados dos modelos são comparados de quatro formas: i) considerando cada um dos 5 VP individualmente, ii) usando a média das estimativas de cada VP, iii) usando o set completo dos 5 VP disponíveis e iv) usando o valor único de proficiência para cada estudante pelo método Warm's Mean Weighted Likelihood Estimates (WLE).

Os dados utilizados são os próprios do TERCE, onde 5 valores plausíveis são gerados para cada área avaliada. Também na base de dados do TERCE foram geradas as variáveis de proficiência, para cada uma das três disciplinas avaliadas, pela estimação WLE.

Por simplicidade, apenas duas variáveis de controle foram selecionadas para ajustar os MLH. Dessa forma, e por ser as variáveis de controle mais importantes, foram selecionadas as variáveis de Nível Socioeconômico (NSE), tanto individualmente quanto no nível da turma/escola, ou seja, via a média agregada das turmas/aulas avaliadas em cada escola (NSEa). É importante mencionar que o interesse é focar-se em comparar como muda o ajuste dos modelos utilizando diferentes métodos de estimação e formas de uso de VP.

A Tabela 6.6 mostra os resultados obtidos. Em primeiro lugar, os coeficientes e erros padrões, são diferentes para cada um dos VPs considerados

isoladamente. Ou seja, quando somente um dos VP são usados, os parâmetros estimados, coeficientes e os erros padrão dos MLH, são menores ou maiores do que aqueles obtidos através do conjunto completo dos 5 VP, dependendo do VP que é observado.

No uso das médias dos VP, os resultados obtidos para os coeficientes são idênticos às estimações dos modelos usando os conjuntos de 5 VP. Nesse sentido, para as três áreas avaliadas, observa-se que os coeficientes estimados nos modelos de MLH, são iguais para o cálculo baseado nas médias dos VP e o conjunto de 5 VP.

Por outro lado, usando as médias dos VP os erros-padrão são subestimados. Para as três áreas avaliadas, os erros padrões das médias dos VP são menores, em comparação com aqueles obtidos através do conjunto completo dos 5 VP.

No final, a comparação com a estimativa da proficiência por WLE mostra que, para as três áreas avaliadas, as estimativas dos erros padrão estão sobrestimadas do que usando o conjunto completo de 5 VP. Além disso, as variâncias intra e entre escolares também são sobrestimadas com WLE do que usando os 5 VP, em particular, a variância entre as escolas (que se refere ao nível 2).

Tabela 6.6. Estimativas dos modelos MLH (coeficientes, erros padrões e variâncias) usando diferentes estratégias dos valores plausíveis para as três áreas avaliadas

	5 VPs	1° VP	2° VP	3° VP	4° VP	5° VP	Média dos 5 VPs	WLE
Leitura								
Intercepto	720.78 (3.179)	721.55 (2.897)	719.88 (2.930)	720.78 (2.941)	719.76 (2.934)	721.94 (3.258)	720.78 (2.992)	720.00 (3.515)
Efeitos Fixos								
NSE	35.77 (3.708)	34.92 (3.100)	35.15 (2.828)	39.46 (3.003)	34.28 (2.743)	35.03 (2.901)	35.77 (2.915)	35.01 (4.284)
NSEa	37.34 (5.648)	37.49 (5.294)	39.20 (5.174)	34.55 (5.358)	36.98 (4.983)	38.48 (5.665)	37.34 (5.295)	31.27 (6.780)
Variância								
Nível 2 - Escola ($\sigma^2 v_{0j}$)	8077.10	8117.85	7942.05	8186.02	8002.73	8139.40	8077.61	8456.11
Nível 1 - Estudante ($\sigma^2 e$)	436.56	381.14	432.31	413.47	436.36	533.40	439.34	1149.83
Matemática								
Intercepto	755.66 (3.427)	755.48 (3.182)	754.12 (3.317)	756.54 (3.273)	756.07 (3.308)	756.10 (3.258)	755.66 (3.268)	752.39 (3.716)
Efeitos Fixos								
NSE	30.95 (3.306)	31.48 (3.103)	31.22 (3.199)	30.84 (3.202)	31.62 (3.099)	29.61 (3.317)	30.95 (3.184)	32.18 (4.898)
NSEa	37.28 (5.792)	37.92 (5.227)	40.69 (5.104)	36.12 (5.397)	36.38 (5.287)	35.31 (5.494)	37.28 (5.302)	29.22 (6.615)
Variância								
Nível 2 - Escola ($\sigma^2 v_{0j}$)	8044.13	8097.48	7921.37	8176.96	8054.98	7972.39	8044.64	8492.48
Nível 1 - Estudante ($\sigma^2 e$)	663.01	613.91	693.85	705.25	668.40	638.04	663.89	1449.18
Ciências Naturais								
Intercepto	723.96 (3.969)	724.66 (3.768)	724.71 (3.968)	723.70 (3.749)	724.66 (3.582)	722.06 (3.755)	723.96 (3.764)	728.10 (4.364)
Efeitos Fixos								
NSE	37.34 (3.556)	38.41 (3.026)	37.12 (3.387)	35.33 (3.460)	38.57 (3.257)	37.27 (3.145)	37.34 (3.255)	37.19 (5.765)
NSEa	28.39 (6.607)	27.53 (5.701)	27.86 (6.354)	31.63 (6.294)	25.38 (6.013)	29.55 (6.069)	28.39 (6.086)	12.98 (8.124)
Variância								
Nível 2 - Escola ($\sigma^2 v_{0j}$)	8007.74	7701.72	8173.16	8328.08	7696.74	8160.59	8012.06	8261.79
Nível 1 - Estudante ($\sigma^2 e$)	975.84	977.62	1100.43	978.12	869.77	966.88	978.56	1873.32

Fonte: Elaborado pelo autor em base a dados da 6ª série do ensino fundamental do Uruguai no estudo TERCE (UNESCO 2013).

Assim como é importante considerar a estrutura hierárquica dos dados quando as estimativas são feitas usando modelos MLH, é muito importante fazer uso adequado de todos os VP que informam a proficiência nas áreas avaliadas conhecimento.

Como Laukaityte e Wiberg (2017) argumentam, talvez por causa da complexidade da modelagem ou algumas limitações dos softwares, alguns pesquisadores não fazem uso adequado dos VP, que pode levar a resultados viesados ou conclusões inválidas.

As análises realizadas mostram que são obtidos resultados viesados se os VP são utilizados de forma inadequada. Quando apenas um dos VP disponíveis é usado, os parâmetros variam e o resultado final depende de qual VP é escolhido.

Como os trabalhos do Laukaityte e Wiberg (2017) e Carstens e Hastedt (2010), os resultados com dados do TERCE mostram que as estimativas, para as três áreas avaliadas, dos coeficientes MLH utilizados pela média dos VP considerados individualmente e usando os 5 VPs são idênticas, mas os erros padrão estão subestimados. Por outro lado, na estimativa da proficiência por meio do WLE, são obtidos erros padrão e variâncias superestimadas.

7. CONSIDERAÇÕES FINAIS

O estudo da variação dos resultados escolares tenta responder a uma questão simples, mas muito relevante: por que as escolas apresentam resultados diferentes umas das outras? Este trabalho procurou mostrar o efeito de um conjunto de características que influenciam o aprendizado. Com esse objetivo, a estratégia de pesquisa consistiu em identificar os fatores associados ao desempenho (proficiência) em Leitura, Matemática e Ciências Naturais dos estudantes do 6º ano das escolas do Uruguai, no ano de 2013.

O estudo de fatores associados à aprendizagem é assunto bastante explorado por vários autores, contudo, nem sempre as conclusões apontam em uma mesma direção. Partiu-se, então, com a ideia de poder comprovar ou refutar os principais fatores que, de acordo com a literatura nacional e internacional, estão associados à aprendizagem, a fim de mostrar as particularidades de nossa população de estudo, e trazer novas considerações para a discussão dos fatores que afetam a aprendizagem, tentando esclarecer eventuais pontos controversos e avaliando a significância de cada fator na aprendizagem.

Foi utilizada a modelagem multinível de dois níveis, nível 1 (estudante) e nível 2 (escola). Os modelos lineares hierárquicos são especialmente adequados para isso, pois dividem a variância do desempenho dos alunos em dois componentes, associados aos estudantes e às escolas.

Em primeiro lugar, a análise mostrou que a estrutura dos fatores associados analisados permite verificar a importância de um conjunto de variáveis relacionadas ao desempenho dos alunos que podem ser identificadas em outros estudos nacionais, bem como em estudos realizados na América Latina e internacionalmente. Por outro lado, na análise existem fatores associados ao desempenho escolar que respondem às características socioculturais e econômicas do funcionamento do sistema educacional uruguaio que merecem destaque.

Um primeiro eixo de análise consistiu em estudar a variabilidade entre escolas explicadas pelos diferentes modelos, ou seja, a redução da variância não explicada na sequência de modelos especificados. Os modelos nulos ajustados para

cada área formam a "linha de base", na qual se comparam ajuste de os modelos subsequentes. Esses modelos permitem avaliar a variabilidade do desempenho entre os dois níveis de análise (alunos e escolas) e estimar o quanto da variabilidade entre as escolas pode ser explicada pelas variáveis escolares.

De acordo com os resultados, aproximadamente uma quarta parte (26%) da variação dos escores nas provas correspondem às diferenças entre as escolas, determinando o possível "teto máximo" do "efeito escolar" na aprendizagem, já que o restante da variação no desempenho corresponde às diferenças entre os alunos das mesmas escolas. Mas essa variação cai mais da metade quando os fatores contextuais e composicionais são incorporados.

A magnitude dos "efeitos escola" depende, em teoria, de dois tipos de aspectos diferentes. De um lado, daqueles fatores que, ao longo de nossa análise, chamamos de "composição" e que não dizem respeito diretamente à ação da escola, como seu ambiente sociocultural, o formato institucional ou sua localização geográfica. No Uruguai, as escolas não têm influência direta nem controlam nenhuma dessas dimensões. Algumas delas, como o formato institucional, podem e são, de fato, objeto de política educacional, mas dependem de definições centrais e não das próprias escolas. Como foi discutido acima, a margem de ação escolar sobre esses fatores é geralmente muito pequena. A pouca variação entre as escolas não deveria nos surpreender, pois indica a centralidade do sistema educacional, caracterizado por ter um currículo único, onde nem o tempo de instrução nem o treinamento básico dos professores, nem a forma de contratação variam substancialmente entre as escolas.

Assim, a primeira conclusão que emerge das análises anteriores é que as escolas uruguaias apresentam margens relativamente baixas para influenciar a aprendizagem e a melhoria escolar, além das características do ambiente familiar das crianças e os efeitos da composição escolar. Em termos gerais, a capacidade das escolas de influenciar a aprendizagem é consideravelmente limitada em comparação com o peso das condições socioculturais dos alunos e do ambiente da própria escola. No entanto, essa capacidade, apesar de ser restrita, deve ser considerada um recurso para melhorar a qualidade e a equidade da educação.

Em relação aos fatores associados à aprendizagem, a análise possibilitou observar a importância do papel do nível socioeconômico familiar, que está diretamente relacionado ao desempenho acadêmico. Este é um indicador claro da distribuição desigual da educação observada nos países latino-americanos.

Os problemas de qualidade nos sistemas educacionais tendem a ser problemas de equidade. De fato, além dos problemas de qualidade, a educação uruguaia apresenta sérios problemas em termos de equidade. Um sistema educacional não é equitativo quando os resultados educacionais dos alunos mais pobres são, em geral, piores do que os dos estudantes com maiores recursos (REIMERS, 2000).

A distribuição da desigualdade não ocorre apenas entre indivíduos, mas está concentrada em alguns grupos específicos de alunos em determinados contextos ou em escolas específicas. Em outras palavras, há menos variabilidade entre o status socioeconômico dos estudantes que frequentam a mesma escola do que entre os estudantes que frequentam escolas diferentes. Assim, não é possível entender o que se acontece em Uruguai (e em qualquer sistema educacional nesta região), sem a consideração explícita do nível socioeconômico de seus alunos. Este não é um achado em si, mas reduzir seu forte condicionamento também é, ao mesmo tempo, a principal preocupação de qualquer sistema educacional que busque melhorar a qualidade e equidade do aprendizado de seus alunos.

É importante destacar que, no nível individual, o efeito do nível socioeconômico ocorreria diretamente, mas também através da intermediação de outras variáveis. Os principais fatores intermediários seriam a condição de emprego do aluno, as aspirações educacionais das famílias e sua própria trajetória escolar.

A análise confirma alguns efeitos associados à trajetória escolar dos alunos e suas características sócio-familiares, como sexo, origem familiar e o impacto da distorção idade/série.

Esses achados mostram que a relação entre posição social e desempenho acadêmico é direta, mas também se configura através de uma série de experiências, representações e práticas individuais e familiares que, eventualmente, podem ser influenciadas pela dinâmica escolar.

Substancialmente, após o capital econômico e cultural disponível para um aluno através de sua casa, é importante considerar que um dos componentes mais

importantes da estrutura da desigualdade de aprendizagem é o fato de que o aluno trabalha enquanto completa seus estudos no nível primário. Um padrão consistente de efeitos negativos foi observado nas três disciplinas avaliadas. É razoável pensar que para se ter uma medida mais refinada e precisa do trabalho infantil, capturando as especificidades de gênero ou diferenças rurais/urbanas, a magnitude estimada dos efeitos seria ainda maior.

Em relação aos fatores de nível escolar, o resultado mais notável é que, em comparação com as características do ambiente e contextuais, os fatores organizacionais representaram uma parte menor da variância total explicada. Da mesma forma que acontece no nível individual, uma parte muito importante da variação dos resultados está associada à incidência do ambiente socioeconômico e cultural da escola.

Outro achado digno de nota é que o setor privado, as escolas particulares, não apresentam efeitos significantes na aprendizagem, uma vez que seja estatisticamente controlado por outras variáveis. Tampouco os efeitos das escolas de Tempo Integral são observados em nenhuma dessas áreas, o que indica que a maior extensão do tempo pedagógico não deve ser entendida linearmente como melhorias diretas nessas disciplinas, embora isso não signifique que esses modelos de escola não apresentem benefícios de outro tipo.

Com relação aos fatores organizacionais, a maioria daqueles que foram significantes nos modelos pertencem, genericamente, às características dos professores e não das escolas. Não foram encontradas relações positivas entre a aprendizagem e a infraestrutura escolar, nem com a experiência e estabilidade dos professores. No entanto, em Leitura e Matemática, relações positivas foram encontradas com o fato de que eles tiveram mestrado ou cursos de especialização de média duração.

Outro dado digno de nota é que nenhuma das variáveis de gestão ou práticas de ensino é significativa. Nesse sentido, as relações, representações e práticas dos professores quase não mostraram efeitos diretos.

O clima escolar, que segundo a teoria é um fator amplamente associado à aprendizagem, não apresentou efeito significativo na aprendizagem, uma vez ajustadas as variáveis de controle.

Na seção final da análise, os métodos de estimação de variância foram considerados teórica e empiricamente para obter erros padrão corretos. Foi demonstrado que na estimação de modelos multiníveis (MLH) nas avaliações em larga escala, deve ser utilizada a variabilidade entre os valores plausíveis fornecidos para estimar a variância da imputação. No entanto, apesar da importância das técnicas de estimativa de variância para corrigir erros padrão, em um grande número de artigos e trabalhos revisados, eles não usam o método de estimativa de variância corretamente ou não fornecem detalhes sobre seus procedimentos.

Seguindo os procedimentos de Laukaityte e Wiberg (2017), foram encontrados resultados viesados se os valores plausíveis são utilizados de forma inadequada. Para as três áreas avaliadas, as estimativas dos MLH que não usam o conjunto dos valores plausíveis são iguais para os coeficientes dos modelos, mas os erros padrão estão subestimados. Por outro lado, na estimativa da proficiência via o método do WLE, são obtidos erros padrão e variâncias within e between escolas superestimadas.

Para terminar, é importante mencionar um conjunto de reflexões baseadas nos resultados encontrados. O primeiro, e possivelmente o fundamental, é que as recomendações de políticas públicas devem ser apoiadas por evidências.

Nos últimos 25 anos, a maioria dos países latino-americanos avançou na discussão sobre a avaliação da qualidade educacional como um eixo fundamental do funcionamento dos sistemas educacionais. Assim, uma vez que os sistemas nacionais de avaliação foram instalados, vários países se vincularam a programas de avaliação nos níveis regional e internacional. Estudos em larga escala como o TERCE, estão dentro dos esforços dos países para construir medidas válidas e confiáveis sobre o desempenho educacional na área de aprendizagem. Nesse sentido, os esforços dos países e das unidades nacionais de avaliação parecem estar caminhando na direção certa.

No entanto, mesmo quando os dados do TERCE estão alinhados com os antecedentes encontrados em estudos nacionais e internacionais, as evidências apresentadas nesta dissertação não podem ser consideradas definitivas, seja por

considerar apenas uma série escolar (6° ano) ou por se focar no estudo de algumas habilidades específicas. Certamente, o uso de dados longitudinais permitiria uma análise mais robusta da aprendizagem no sistema educacional que permita avaliar as desigualdades de acesso aos níveis educacionais superiores e no mercado de trabalho.

Para trabalhos futuros será necessário aprofundar o estudo dos possíveis fatores institucionais que podem influenciar o desempenho dos estudantes. Ainda que a pesquisa por fatores associados à aprendizagem realizada nesta dissertação tenha sido bastante ampla e se tenha procurado explorar ao máximo a base de dados disponível, não se pode afirmar que os fatores aqui apontados sejam, de fato, os únicos que afetam a aprendizagem. Nesse sentido, não se observa na literatura nacional ampla pesquisa qualitativa no sentido de identificar os potenciais fatores associados à aprendizagem no contexto nacional. Propõe-se como continuidade e melhoria do presente estudo, a realização de pesquisa qualitativa ampla que atinja alunos, pais, professores e diretores de escolas e inspetores dentre outros profissionais de educação, para aprofundar em fatores associados significantes que ainda não foram desenvolvidos em questionários contextuais.

REFERÊNCIAS

ALBERNAZ, A., FERREIRA, F.H.G. e FRANCO, C. (2002). Qualidade e equidade na educação fundamental brasileira. Rio de Janeiro: PUC-RIO, (Texto para discussão, n. 455).

ALVES, M. T. G.; SOARES, J. F.(2007). As pesquisas sobre o efeito das escolas: contribuições metodológicas para a sociologia da educação. Sociedade e Estado, Brasília, v. 22, n. 2, p. 435-473.

AMMERMUELLER, A. (2013). Institutional Features of Schooling Systems and Educational Inequality: Cross-Country Evidence From PIRLS and PISA. German Economic Review, 14(2), 190-213.

AMMERMUELLER, A., e PISCHKE, J.-S. (2009). Peer Effects in European Primary Schools: Evidence from the Progress in International Reading Literacy Study. Journal of Labor Economics, 27(3), 315-348.

ANDERS. I., ROßBACH.H.G e KUGER.S. (2016) Early Childhood Learning Experiences Em: S. Kuger et al. (eds.), Assessing Contexts of Learning, Methodology of Educational Measurement and Assessment, DOI 10.1007/978-3-319-45357-6_7

ANDRADE, D. F. et al. (2000). Teoria da resposta ao item: conceitos e aplicações. São Paulo: ABE Associação Brasileira de Estatística.

ANDRADE, R. J.; SOARES, J. F. (2008). O efeito da escola básica brasileira. Estudos em avaliação educacional, p. 379-406.

ANEP (1997). Evaluación nacional de aprendizajes en lengua materna y matemática. 6° año de enseñanza primaria-1996. Primer informe de difusión pública de resultados, ANEP-MECAEP, Montevideo.

ANEP (1999). Estudio de los factores institucionales y pedagógicos que inciden en los aprendizajes en escuelas primarias de contextos sociales desfavorecidos en el Uruguay, ANEP, MECAEP-UMRE, Montevideo.

ANEP (2002). Evaluación nacional de aprendizajes en ciencia y matemática. 6° año de enseñanza primaria/2002. Primer informe de devolución de resultados de la muestra nacional, ANEP- MECAEP, Montevideo.

ANEP (2005). Panorama de la educación en el Uruguay. Una década de transformaciones (1992- 2004). ANEP-CODICEN, Montevideo.

ANEP (2007). Evaluación nacional de aprendizajes en lenguaje y matemática. 6° año de enseñanza primaria/2005, ANEP-CODICEN, Montevideo.

ANEP (2009). Uruguay en el segundo estudio regional comparativo y explicativo (SERCE). Informe nacional, ANEP-CODICEN, Montevideo.

ANEP (2010). Evaluación de aprendizajes en lengua, matemática y ciencias. 6° año de enseñanza primaria. 2009, ANEP-CODICEN, Montevideo.

ANEP (2013). Primer Informe Uruguay en PISA 2012, ANEP-PISA, Montevideo.

ANEP (2014). Uruguay en PISA 2012. Primer Informe, ANEP, Montevideo.

ANEP (2015a). Evaluación Nacional de 6° año. En Matemática, Ciencias y Lengua. 2013. Primer informe, ANEP, Montevideo.

ANEP (2015b). Proyecto de Presupuesto Período 2015-2019. Tomo I, Exposición de Motivos, ANEP- CODICEN, Montevideo.

ANEP (2016). Evaluación de impacto de las escuelas de Tiempo Completo en Uruguay 2013-2016, ANEP- CODICEN, Montevideo.

ARAUJO, M. C., CARNEIRO, P., CRUZ-AGUAYO, e SCHADY, N. (2014). A Helping Hand? Teacher Quality and Learning Outcomes in Kindergarten. Washington, DC: Inter-American Development Bank.

AUCEJO, E. M., e ROMANO, T. F. (2014). Assessing the effect of school days and absences on test score performance. London: Centre for Economic Performance. <http://cep.lse.ac.uk/pubs/download/dp1302.pdf> . Accessed 25 May 2018.

BACCHETTO, J. G. (2016). O Pisa e o custo da repetência no Fundeb. Ensaio: Avaliação e Políticas Públicas em Educação, 24(91), 424–444. <https://doi.org/10.1590/S0104-40362016000200008>

BARBETTA, P.; TREVISAN, L.; TAVARES, H. e AZEVEDO, T. (2014). Aplicação da Teoria da Resposta ao Item Uni e Multidimensional. Revista Estudos em Avaliação Educacional, São Paulo: FCC, v.25, n. 57, p. 280-302, jan./abr. 2014.

BARBETTA, P.A.; ANDRADE, D. F.; TAVARES, H. R. (2018). Estudo de fatores associados através de regressão quantílica hierárquica. Estudos em Avaliação Educacional, São Paulo, v. 29, n. 71, p. 320-349, maio/ago. 2018. Disponível em: <http://publicacoes.fcc.org.br/ojs/index.php/eae/article/view/4973/3634>

BARNETT, W. S., e MASSE, L. N. (2007). Comparative benefit-cost analysis of the Abecedarian pro- gram and its policy implications. Economics of Education Review, 26, 113–125.

BARBOSA, M.E.F, FERNANDES, C. (2001). A escola brasileira faz diferença? Uma investigação dos efeitos da escola na proficiência em matemática dos estudantes da 4ª série. In: FRANCO, C. (Org.) Avaliação, ciclos e promoção na educação. Porto Alegre: ARTMED, p.155-172.

BERLINER, D. (1990). What's all the fuss about instructional time? In M. Ben-Peretz & R. Bromme (Eds.), *The nature of time in schools* (pp. 3–35). New York: Teachers College Press.

BLANCO, E. (2008). Factores escolares asociados a los aprendizajes en la educación primaria mexicana: un análisis multinivel. *REICE - Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 6(1), 1–8.

BLANCO, E. (2011). *Los límites de la escuela: educación, desigualdad y aprendizajes en México*. El Colegio de México, Centro de Estudios Sociológicos, 2011. México, D.F.

BLOSSFELD, H.-P., SCHNEIDER, T., e DOLL, J. (2009). Methodological Advantages Of Panel Studies: Designing The New National Educational Panel Study (Neps) In Germany. *Journal For Educational Research Online*, 1(1), 10–32.

BOADO, M. e FERNÁNDEZ, T. (2010). *Trayectorias Académicas Y Laborales De Los Jóvenes En Uruguay*. El Panel PISA 2003-2007, Udelar, FCS, Montevideo.

BORGES JWP, MOREIRA TMM, SCHMITT J, ANDRADE DF, BARBETTA PA, SOUZA ACC, et al. (2017). Medição da qualidade de vida em hipertensão arterial segundo a Teoria da Resposta ao Item. *Rev Saude Publica*;51:45.

BOWEN, G. L., e CHAPMAN, M. V. (1996). Poverty, Neighborhood Danger, Social Support, and the Individual Adaptation among at-risk Youth in Urban Areas. *Journal of Family Issues*, 17(5), 641-666.

BOWEN, N. K, e BOWEN, G. L. (1999). Effects of Crime and Violence in Neighborhoods and Schools on the School Behavior and Performance of Adolescents. *Journal of Adolescent Research*, 14(3), 319-342. doi: 10.1177/0743558499143003

BRAY, M., MAZAWI, A. E., e SULTANA, R. G. (2013). *Private tutoring across the Mediterranean—power dynamics and implications for learning and equity*. Rotterdam: Springer. doi: 10.1007/978-94-6209-237-2

BROOKOVER, W.B. et al (1979). *School Systems and Student Achievement: Schools Make a Difference*. New York: Praeger.

BRUNS, B. e LUQUE, J. (2014). *Great Teachers: How to Raise Student Learning in Latin America and the Caribbean*. World bank group.

BRYK, A. (2008). “Organizando escuelas para la mejora: marcos teóricos, resultados empíricos y métodos de investigación”. Stanford University . UCA- Buenos Aires.

BURDICK-WILL, J. (2016). Neighborhood Violent Crime and Academic Growth in Chicago: Lasting Effects of Early Exposure. *Social Forces*, 95(1), 133-157. doi:10.1093/sf/sow041

CARDOZO, S. (2010). El Comienzo Del Fin: Las Decisiones De Abandono Durante La Educación Media Y Su Incidencia En Las Trayectorias, En Tabaré Fernández (Coord.), *La Desafiliación En La Educación Media Y Superior De Uruguay: Conceptos, Estudios Y Políticas*, Udelar, Montevideo.

CARDOZO, S. (2016). Trayectorias educativas en la educación media PISA-L 2009-2014, INEED - Grupo de estudios sobre Transiciones Educación-Trabajo (TET), Montevideo.

CARDOZO, S. e FERNÁNDEZ, T. (2013). “Los Proyectos PISA Longitudinales. Un Panorama Sobre Australia, Canadá, Dinamarca, República Checa, Suiza y Uruguay”, Material Elaborado Para El Curso De Educación Permanente “Estudios Longitudinales En Educación Y Los Proyectos PISA-L”, Udelar- CEPE.

CARRASCO OGAZ, D. A. (2016) Multivariate approaches to school climate factors and school outcomes, PhD Thesis, University Of Sussex.

CARSTENS, R., e HASTEDT, D. (2010). The effect of not using plausible values when they should be: An illustration using TIMSS 2007 Grade 8 mathematics data. In *Proceedings of the 4th IEA International Research Conference*.

CASTRO, R. F. (2009) Factores asociados ao desempenho escolar na 4ª série do ensino fundamental. Ed. UFBA,. Disponível em: <<http://books.scielo.org/id/wd/pdf/lordelo-9788523209315-13.pdf>> Acesso em: 02 junho 2018.

CEPAL (1992). *¿Aprenden los estudiantes? El Ciclo Básico de Educación Media*. Montevideo: CEPAL.

CEPAL (2012), *Panorama Social de América Latina, 2012 (LC/G.2481-P)*, Santiago de Chile. Publicación de las Naciones Unidas, N° de venta: S.11.II.G.6.

CERVINI, R. (1999). *Calidad Y Equidad En La Educación Básica De Argentina. Serie Factores Asociados Al Logro Escolar*. Buenos Aires.

CERVINI, R. (2002). Desigualdades En El Logro Académico Y Reproducción Cultural En Argentina. Un Modelo De Tres Niveles. *Revista Mexicana De Investigación Educativa*, 7(16), 445–500. Recuperado De <Http://Www.Redalyc.Org/Html/140/14001604/>

CHARLIN, V., TORRES, A., e CAYUMÁN, C. (2016). Expectativas de género y logro de los estudiantes en TERCE. Midevidencias, 9, 1-8. <http://www.mideuc.cl/wp-content/uploads/2016/MidEvidencias-N9.pdf>

CHMIELEWSKI, A., e DHUEY, E. (2017). The Analysis of International Large-Scale Assessments to Address Causal Questions in Education Policy. University of Toronto.

CHUBB, J. e MOE, T. M., (1990). "Politics, Markets and America's Schools". The Brookings Institution, Washington, D.C.

CLAVIJO, J., FRANCIA, T. e RETAMOSO, A. (2005). "Escuelas de Tiempo Completo: 1997 a 2005". ANEP-MECAEP, Montevideo.

COLEMAN, J., CAMPBELL, E., HOBSON, C., MCPARTLAND, J., MOOD, A., WEINFELD, F., E YORK, R. (1966). Equality Of Educational Opportunity. Washington, DC: National Center For Educational Statistics.

DANIELSON, C. (2011). Enhancing professional practice: A framework for teaching. ASCD. Danielson, C. (2013). The Framework for Teaching. Evaluation Instrument. 2013 Edition.

DOWNER, J. T., STUHLMAN, M., SCHWEIG, J., MARTÍNEZ, J. F., e RUZEK, E. (2015). Measuring effective teacher-student interactions from a student perspective: A multi-level analysis. The Journal of Early Adolescence, 35(5-6), 722-758.

DURYEA, S., e ROBLES M. (2016). Pulso Social de América Latina y el Caribe: Realidades y Perspectivas. Banco Interamericano de Desarrollo.

De ARMAS, G. e RETAMOSO, A. (2010). La Universalización De La Educación Media En Uruguay: Tendencias, Asignaturas Pendientes Y Retos A Futuro, UNICEF, Montevideo.

De LEEUW, J., e KREFT, I. (1986). Random Coefficient Models For Multilevel Analysis. Journal Of Educational Statistics, 11(1), 57-85. <https://doi.org/10.3102/10769986011001057>

DE MELO, G. (2011). Peer effects identified through social networks. Evidence from Uruguayan schools, Documento de Trabajo, Decon-UdelaR, Montevideo.

EDMONDS, R. (1982). Effective Schools for the Urban Poor. Educational Leadership, 37, pp. 15-27.

FAUTH, B., DECRISTAN, J., RIESER, S., KLIEME, E., e BÜTTNER, G. (2014). Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. Learning and Instruction, 29, 1-9.

FERNÁNDEZ, T. (1999) Efectividad en Educación. Revista de Ciencias Sociales, 16. Departamento de Sociología. Universidad de la República. Montevideo.

FERNÁNDEZ, T. (2004). De las Escuelas Eficaces a las reformas educativas de segunda generación. *Estudios Sociológicos*, 22(65), 377-408.

FERNÁNDEZ, T. (2004). Distribución del conocimiento escolar: clases sociales, escuelas y sistema educativo en América Latina. *El Colegio de México*. México, D.F.

FERNÁNDEZ, T. (2009). La desafiliación en la educación media de Uruguay. Una aproximación con base en el panel de estudiantes evaluados por PISA 2003. *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 7(4), 165-179.

FERNÁNDEZ, T., ALONSO, C., BOADO, M., CARDOZO, S. e MENESE, P. (2013). Reporte Técnico PISA-L (2003-2012). Metodología de la Segunda Encuesta de seguimiento a los jóvenes evaluados por PISA en 2003, Departamento de Sociología de la Facultad de Ciencias Sociales de la Universidad de la República, Montevideo.

FERNÁNDEZ, T., BOADO, M. e BONAPELCH, S. (2008). “Reporte Técnico del Estudio Longitudinal de los estudiantes evaluados por PISA 2003 en Uruguay”, Informe de Investigación n° 40, DS-Udelar, Montevideo.

FORQUIN, J. C. (1995). *Sociologia Da Educação: Dez Anos De Pesquisa*. Petrópolis: Vozes.

FOY, P., GALIA, J., e LI, I. (2008). Scaling the data from the TIMSS 2007 mathematics and science assessments. In J. F. Olson, M. O. Martin, & I. V. Mullis (Eds.), *TIMSS 2007 technical report* (pp. 225–280). Chestnut Hill, MA: Boston College.

HANUSHEK, E. A. (2011). The Economic Value Of Higher Teacher Quality. *Economics Of Education Review*, 30(3), 466–479.
<https://doi.org/10.1016/j.econedurev.2010.12.006>

HANUSHEK, E. A. (2013). Economic Growth In Developing Countries: The Role Of Human Capital. *Economics Of Education Review*, 37, 204–212.
<https://doi.org/10.1016/j.econedurev.2013.04.005>

HARDING, D. (2009). Violence, Older Peers, and the Socialization of Adolescent Boys in Disadvantaged Neighborhoods. *American Sociological Review*, 74(3), 445-464. doi: 10.1177/000312240907400306

HATTIE, J. A. (1999). Influences on student learning (Inaugural professorial address, University of Auckland, New Zealand).
<http://www.arts.auckland.ac.nz/staff/index.cfm?P=8650>

HERTEL, S. e JUDE, N. (2016). Parental Support and Involvement in School En: S. Kuger et al. (eds.), *Assessing Contexts of Learning, Methodology of Educational Measurement and Assessment*, DOI 10.1007/978-3-319-45357-6_7

HILL, N. E. (1997). Does parenting differ based on social class?: African American women's perceived socialization for achievement. *American Journal of Community Psychology*, 25(5), 675-697.

HILL, N. E., CASTELLINO, D. R., LANSFORD, J. E., NOWLIN, P., DODGE, K. A., BATES, J. E., e PETTIT, G. S. (2004). Parental academic involvement as related to school behavior, achievement, and aspirations: Demographic variations across adolescence. *Child Development*, 75, 1491–1509.

HO, A. D., e KANE, T. J. (2012). *The reliability of classroom observations by school personnel*. Seattle, WA: Bill & Melinda Gates Foundation.

HOX, J.J. (2010). *Multilevel Analysis: Techniques and Applications*. 2^a ed. New York, Routledge.

HUEBENER, M., KUGER, S., e MARCUS, J. (2016). Increased instruction hours and the widening gap in student performance . Berlin: DIW Berlin German Institute for Economic Research. <http://hdl.handle.net/10419/130217> . Accessed 25 May 2016.

INEED (2014). *Informe Sobre El Estado De La Educación En Uruguay 2014*, Ineed, Montevideo.

INEED (2015). *Uruguay En El TERCE: Resultados Y Prospecciones*, Ineed, Montevideo.

INEED (2017). *Informe Sobre El Estado De La Educación En Uruguay 2015-2016*, Ineed, Montevideo.

KANE, T. J., MCCAFFREY, D. F., e STAIGER, D. O. (2010). *Learning about teaching: Initial findings from the Measures of Effective Teaching Project*. Seattle, WA: Bill & Melinda Gates Foundation.

KLEIN, R. e FONTANIVE, N. (1995). *Avaliação Em Larga Escala: Uma Proposta Inovadora*. Em *Aberto*, Brasília, V. 66, P. 29-35.

KUGER, S. (2016). *Curriculum and Learning Time in International School Achievement Studies* En: S. Kuger et al. (eds.), *Assessing Contexts of Learning, Methodology of Educational Measurement and Assessment*, DOI 10.1007/978-3-319-45357-6_7

LAUKAITYTĖ, I. (2016). *Statistical modeling in international large-scale assessments*. 2016. 28p. Doctoral Thesis – Umea School of Business and Economics, Umea, 2016.

LAUKAITYTĖ, I., e WIBERG, M. (2017). Using plausible values in secondary analysis in large-scale assessments. *Communications in Statistics - Theory and Methods*, 46(22), 11341–11357. <https://doi.org/10.1080/03610926.2016.1267764>

LLAMBÍ, M. C. (2013). El efecto causal de la política de tiempo completo sobre los resultados educativos en la enseñanza media: análisis mediante la aplicación de cuatro métodos no experimentales e identificación de posibles sesgos. Universidad de la República.

LITTELL, R. C., MILLIKEN, G., STROUP, W. W., WOLFINGER, R. D., e SCHABENBERGER, O. (1996). *Sas For Mixed Models*. (Nc, Org.). Sas Institute Inc.

LONGFORD, N. T. (1993). *Regression Analysis Of Multilevel Data With Measurement Error*. *British Journal Of Mathematical And Statistical Psychology*. United Kingdom: British Psychological Society. <https://doi.org/10.1111/J.2044-8317.1993.Tb01018.X>

MACEDO ALVES, G. (2004). *Fatores Associados Ao Rendimento Escolar De Alunos Da 5a Série (2000) - Uma Abordagem Do Valor Adicionado*. Xiv Encontro Nacional De Estudos Populacionais, Abep.

MACHADO, D. C. O. *Análise de Fatores Associados ao Desempenho Escolar de Alunos do Quinto Ano do Ensino Fundamental com Base na Construção de Indicadores*. 2014. Dissertação (Métodos e Gestão em Avaliação) - Universidade Federal de Santa Catarina.

MANCIBO, M.E. e TORELLO, M. (2000). *El sistema educativo uruguayo: Estudio de diagnóstico y propuesta de políticas públicas para el sector*. IADB. Washington D.C.

MARCOTTE, D. E. (2007). *Schooling and test scores: A mother-natural experiment*. *Economics of Education Review*, 26 (5), 629–640. doi: 10.1016/j.econedurev.2006.08.001 .

MARTIN, M. O., MULLIS, I. V. S., FOY, P., e HOOPER, M. (2016). *TIMSS 2015 International Results in Science*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <http://timssandpirls.bc.edu/timss2015/international-results/>

MENESE, P. (2012). *Trayectorias educativas y los factores asociados a la acreditación de la Educación Media Superior en Montevideo y Maldonado*, FCS, Montevideo.

MISLEVY, R. J. (1991). *Randomization-based inference about latent variables from complex samples*. *Psychometrika*, 56(2), 177–196.

MISLEVY, R. J., BEATON, A. E., KAPLAN, B., e SHEEHAN, K. M. (1992). *Estimating population characteristics from sparse matrix samples of item responses*. *Journal of Educational Measurement*, 29(2), 133–161

MISLEVY, R. J., JOHNSON, E. G., e MURAKI, E. (1992). Chapter 3: Scaling procedures in NAEP. *Journal of Educational and Behavioral Statistics*, 17(2), 131–154. <http://doi.org/10.3102/10769986017002131>

MORO, A. (2018). *A Construção E As Evidências De Validade De Instrumentos De Medida Para Avaliar O Clima Escolar*. Universidade De Campinas.

MULLIS, I. V. S., MARTIN, M. O., FOY, P., e DRUCKER, K. T. (2012). *PIRLS 2011 International Results in Reading*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <https://timssandpirls.bc.edu/pirls2011/international-results-pirls.html>

MULLIS, I. V. S., MARTIN, M. O., FOY, P., e HOOPER, M. (2016). *TIMSS 2015 International Results in Mathematics*. Recuperado de Boston College, TIMSS & PIRLS International Study Center website: <http://timssandpirls.bc.edu/timss2015/international-results/>

MORDUCHOWICZ, A. (2004): *Discusiones de economía de la educación*, Losada, Buenos Aires. OECD.

NASCIMENTO, P. A. M. M. (2012). Fatores Associados Ao Desempenho Escolar: Resultados De Um Modelo Multinível De Valor Adicionado. *Estudos Em Avaliação Educacional*, 23(51), 16–39. Recuperado De <Http://Publicacoes.Fcc.Org.Br/Ojs/Index.Php/Eae/Article/View/1943>

OECD. (2013). *PISA 2012 Results: What Makes Schools Successful? Resources, Policies and Practices (Volume IV)*. PISA, OECD Publishing. <http://dx.doi.org/10.1787/9789264201156-en>

OECD (2016). *PISA 2015 Results (Volume I): Excellence and Equity in Education*. PISA, OECD Publishing, Paris. <http://dx.doi.org/10.1787/9789264266490-en>

OECD. (2017). *Starting Strong 2017: Key OECD Indicators on Early Childhood Education and Care*. OECD Publishing, Paris. <http://dx.doi.org/10.1787/9789264276116-en>

ÖZDEMİR, C. (2016). Equity In The Turkish Education System: A Multilevel Analysis Of Social Background Influences On The Mathematics Performance Of 15-Year-Old Students. *European Educational Research Journal*, 15(2), 193–217. <https://doi.org/10.1177/1474904115627159>

POST, D. (2001). Region, Poverty, Sibship, and Gender Inequality in Mexican Education: Will Targeted Welfare Policy Make a Difference for Girls?. *Gender & Society*, 15(3), 468–489.

POST, D. e PONG, S. (2000). “Employment during middle school: the effects on academic achievement in the us and abroad”, en *Educational Evaluation and Policy Analysis*, vol. 22, núm. 3.

RAVELA, P. (2006). Para comprender las evaluaciones educativas. Fichas didácticas. Preal Documentos. Washington. Recuperado de <http://dide.minedu.gob.pe/handle/123456789/4999>

RAUDENBUSH, S. W. (1993). Hierarchical Linear Models And Experimental Design. *Applied Analysis Of Variance In Behavioral Science*, 459–496.

RAUDENBUSH, S., e BRYK, A. S. (1986). A Hierarchical Model For Studying School Effects Source. *American Sociological Association*, 59(1), 1–17.

RAUDENBUSH, S. y BRYK, A. (2002). *Hierarchical Linear Models . Second Edition*. Ed. Sage. Thousand Oaks. CA.

RAUDENBUSH, S. W. e WILLMS, D. (1995). The Estimation of School Effects, *Journal of Educational and Behavioral Statistics* Winter 1995, Vol. 20, No. 4, pp. 307-335.

REIMERS, F. (2000) (ed.). *Unequal Schools, Unequal Chances. The Challenges to Equal Opportunities in the Americas*. Cambridge: The David Rockefeller Center Series on Latin American Studies, Universidad de Harvard.

RÍOS, Á. (2014). “Perfiles de riesgo educativo y trayectorias de los jóvenes durante la educación media superior”, en *Páginas de educación*, vol. 6, n° 2, Montevideo.

RITZEN, J. (2013). International Large-Scale Assessments As Change Agents. In Y. K. Von Davier M., Gonzalez E., Kirsch I. (Org.), *The Role Of International Large-Scale Assessments: Perspectives From Technology, Economy, And Educational Research*. Dordrecht: Springer. <https://doi.org/10.1007/978-94-007-4629-9>

RIVAS, A. (2015). *América Latina Después De PISA. Lecciones Aprendidas De La Educación En Siete Países (2000-2015)*, CIPPEC, Buenos Aires.

ROBAINA, S. (2012). ¿Qué factores determinan la transición a la Educación Terciaria? Una mirada a las trayectorias académicas de los jóvenes en Maldonado a partir de la instalación del CURE, FCS, Montevideo.

ROCKOFF, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *The American Economic Review*, 94(2), 247-252. Doi: 10.1257/0002828041302244

RUBIN, D. B. (1987). *Multiple imputation for nonresponse in surveys*. Hoboken, NJ: Wiley.

RUTKOWSKI, L., GONZALEZ, E., JONCAS, M., & von DAVIER, M. (2010). International large-scale assessment data: Issues in secondary analysis and reporting. *Educational Researcher*, 39(2), 142–151.

SAMEJIMA, F. A. (1996). Estimation of latent ability using a response pattern of graded scores. *Psychometric Monograph*, 17.

SHONKOFF, J. P., e PHILLIPS, D. A. (2000). *From Neurons to Neighborhoods: The Science of Early Childhood Development*. Washington, D.C.: National Academy Press.

SOARES, J.F, CÉSAR, C.C., MAMBRINI, J. (1997). Determinantes Do Desempenho Dos Estudantes Do Ensino Básico Brasileiro: Evidências Do SAEB De 1997. In: FRANCO, C. (Org.) *Avaliação, Ciclos E Promoção Na Educação*. Porto Alegre: Artmed, 2001. P.121-154

SOARES, J.F. (2003). Quality And Equity In Brazilian Basic Education: Facts And Possibilities. In: SEMINÁRIO IASI DE ESTATÍSTICA APLICADA: ESTATÍSTICA NA EDUCAÇÃO E EDUCAÇÃO EM ESTATÍSTICA, 9, 2003, Rio De Janeiro. Conferência Convidada: Quality And Equity In Brazilian Basic Education: The Hlm Answers. Rio De Janeiro: IASI.

SOARES, J.F. (2004). O Efeito Da Escola No Desempenho Cognitivo De Seus Estudantes. *REICE: Revista Electrónica Iberoamericana Sobre Calidad, Eficacia Y Cambio En Educación*, Vol. 2, Nº. 2.

TAYLOR, L. C., CLAYTON, J. D., e ROWLEY, S. J. (2004). Academic socialization: Understanding parental influences on children's school-related development in the early years. *Review of general psychology*, 8(3), 163.

TRAUTWEIN, U., KÖLLER, O., SCHMITZ, B., e BAUMERT, J. (2002). Do homework assignments enhance achievement? A multilevel analysis in 7th-grade mathematics. *Contemporary Educational Psychology*, 27 (1), 26–50. doi: 10.1006/ceps.2001.1084.

UNESCO (2007). *Situación Educativa De América Latina Y El Caribe: Garantizando La Educación De Calidad Para Todos*. Santiago: OREALC/UNESCO Santiago.

UNESCO (2008). *Una mirada al interior de las escuelas primarias. Estudio comparativo realizado en el marco del proyecto de Indicadores Mundiales de Educación. Resumen Ejecutivo*. Ed UNESCO. Montreal, Québec.

UNESCO (2010). *Factores asociados al logro cognitivo de los estudiantes de América Latina y el Caribe*, UNESCO, Santiago de Chile.

UNESCO. (2016a). *Reporte Técnico. Tercer Estudio Regional Comparativo Y Explicativo, Terce*. Santiago.

UNESCO (2016b). *Informe De Resultados Del Tercer Estudio Regional Comparativo Y Explicativo. Factores Asociados*. Santiago De Chile: UNESCO.

UNESCO (2016c). *Informe De Resultados Del Tercer Estudio Regional Comparativo Y Explicativo. Logros De Aprendizaje*. Santiago De Chile: UNESCO.

VON DAVIER, M., GONZALEZ, E. J., e MISLEVY, R. J. (2009). What Are Plausible Values And Why Are They Useful? Ieri Monograph Series: Issues And Methodologies In Large-Scale Assessments, 2, 9–36. Recuperado De [Http://Www.IerInstitute.Org/Fileadmin/Documents/Ieri_Monograph/Ieri_Monograph_Volume_02_Chapter_Introduction.Pdf](http://www.ierinstitute.org/fileadmin/Documents/Ieri_Monograph/Ieri_Monograph_Volume_02_Chapter_Introduction.Pdf)

VON DAVIER, M., GONZALEZ, E., KIRSCH, I., e YAMAMOTO, K. (Orgs.). (2013). The Role Of International Large-Scale Assessments: Perspectives From Technology, Economy, And Educational Research. Dordrecht: Springer Netherlands. [Https://Doi.Org/10.1007/978-94-007-4629-9](https://doi.org/10.1007/978-94-007-4629-9)

WILD, E., e LORENZ, F. (2010). Elternhaus und Schule [Parents and school]. Paderborn: Schöningh/ UTB.

WILLMS, D. (2006), Learning Divides: Ten Policy questions about the performance and equity of schools and schooling systems, UNESCO, Institute for Statistics, Montreal.

WILLMS, D. (2010), “School Composition and Contextual Effects on Student Outcomes”, Teachers College Record, No. 112, Vol. 4.

WU, M. (2005). The role of plausible values in large-scale surveys. *Studies in Educational Evaluation*, 31, 114–128.

WU, M. (2010). Measurement, sampling, and equating errors in large-scale assessments. *Educational Measurement: Issues and Practice*, 29(4), 15–27. <http://doi.org/10.1111/j.1745-3992.2010.00190.x>

YANG HANSEN, K., e STRIETHOLT, R. (2018). Does Schooling Actually Perpetuate Educational Inequality In Mathematics Performance? A Validity Question On The Measures Of Opportunity To Learn In PISA. *ZDM - Mathematics Education*, 50(4), 643–658. [Https://Doi.Org/10.1007/S11858-018-0935-3](https://doi.org/10.1007/S11858-018-0935-3)