

3. Metodologia

3.1 A construção dos questionários

Os questionários aplicados pelo Geres incluem medidas sobre as características promotoras de eficácia escolar. Essas características foram definidas a partir da literatura sobre o tema. No âmbito dos estudos internacionais alguns pesquisadores elaboraram sínteses dos fatores escolares promotores de eficácia escolar. Willms (1992 apud BROOKE; BONAMINO 2011), por exemplo, propôs um modelo de geração de dados para responder a quatro perguntas sobre eficácia e equidade: 1) até que ponto os resultados variam de uma escola para outra; 2) até que ponto os resultados variam entre alunos de classes sociais diferentes; 3) quais são as práticas e as políticas escolares que contribuem para melhorar os resultados escolares; 4) quais são as práticas e as políticas escolares que contribuem para diminuir as desigualdades nos resultados entre alunos de diferentes classes sociais. Esse modelo proposto por Willms baseia-se no fato de que os resultados escolares estão relacionados tanto à trajetória escolar prévia e às influências da família, quanto às experiências escolares dos alunos que são influenciados pelas práticas, políticas, estruturas organizacionais, normas da sala de aula, da escola e da rede ou distrito no qual a escola está inserida. Assim, Willms elaborou alguns construtos relevantes para a avaliação de um sistema escolar. Esses construtos estão diretamente relacionados aos principais atores do processo ensino aprendido: escola, alunos e atitudes dos professores.

Outro estudo que também contribuiu para a elaboração dos questionários Geres foi o trabalho de síntese de Lee, Brynk e Smith (1993 apud BROOKE; BONAMINO, 2011) que propôs um modelo conceitual heurístico de organização da escola secundária e da influência dessa organização sobre os alunos, sobre os professores, sobre as relações entre eles e sobre o desempenho dos alunos. Estes estudiosos dividiram os fatores escolares em dois grandes grupos – fatores externos e fatores internos. Os fatores externos dizem respeito às características dos alunos e das famílias dos alunos (raça, nível socioeconômico, interações sociais, vínculos escola-família, envolvimento dos pais no aprendizado, ação da comunidade etc.). Os fatores internos dizem respeito à organização interna da

escola (tipo de gestão, admissão e demissão de professores, admissão e expulsão de estudantes, autoridade do corpo docente, administração dos recursos etc.), à organização formal do trabalho (função dos departamentos, função dos professores, currículo, atividades acadêmicas dos estudantes) e à organização social da escola (relacionamento docente, influência dos pares, relacionamento professores-alunos etc.).

Um terceiro estudo elaborado por Sammons, Hillman e Mortimore (1995 apud BROOKE; BONAMINO, 2011), baseado em dados de diversos países, faz uma síntese das principais características que melhor descrevem escolas bem-sucedidas, ou seja, escolas eficazes. O estudo aponta que há pelo menos onze características (com subitens) que agregam valor aos seus alunos. São elas: 1) Liderança profissional; 2) Visão e metas compartilhadas; 3) Um ambiente de aprendizado; 4) Concentração no ensino e na aprendizagem; 5) Ensino com propósitos definidos; 6) Altas expectativas; 7) Reforço positivo; 8) Monitoramento do progresso; 9) Direitos e responsabilidades dos alunos; 10) Relacionamento família-escola; 11) Organização orientada para a aprendizagem.

Essas onze características tiveram uma grande repercussão e apareceram como referência em diversos trabalhos produzidos posteriormente.

No âmbito das pesquisas nacionais, ainda são raros os estudos que apresentam uma síntese dos principais fatores presentes em escolas eficazes. No entanto, algumas pesquisas já caminharam nesta direção. O estudo realizado por Guiomar Namó de Mello (1994 apud BROOKE; BONAMINO, 2011) faz uma síntese das principais características das escolas eficazes. Esta síntese inclui alguns fatores citados nos estudos internacionais. De acordo com Mello, os principais fatores de escolas eficazes são: 1) Planejamento curricular da escola; 2) Estrutura e organização curricular; 3) Domínio do conteúdo pelo professor; 4) Estratégias de organização da sala de aula; 5) Métodos de ensino.

Na década de 1990 o Ministério da Educação e do Desporto – MEC elaborou um documento denominado *Referenciais para Formação de Professores*, no qual discute o papel da escola e do professor e apresenta itens que sintetizam as condições necessárias para assegurar uma educação de qualidade: 1) Apoio administrativo ao projeto educativo; 2) Existência de um projeto educativo explícito e compartilhado pelos diferentes segmentos da escola; 3) Formas ágeis e flexíveis de organização institucional; 4) Planejamento coletivo do trabalho; 5)

Adequação do espaço físico e das instalações; 6) Qualidade dos recursos didático-pedagógicos; 7) Existência de biblioteca e acervo de materiais diversificados; 8) Tempo adequado de permanência dos alunos na escola; 9) Qualidade de formação inicial dos professores; 10) Desenvolvimento profissional contínuo; 11) Proporção apropriada alunos-professor; 12) Condições apropriadas de trabalho e salário; 13) Quadro estável de profissionais.

Analisando os fatores elencados pelo MEC pode-se perceber que estes itens estão diretamente relacionados com a realidade educacional brasileira, sendo que muitos destes itens não se aplicam a outras realidades. Um exemplo disso é o item oito: tempo adequado de permanência dos alunos na escola. Nenhum estudo sobre eficácia escolar aponta este item como fator de promoção de eficácia escolar, pois o tempo que os alunos ficam na escola, nos demais países, é um tempo adequado. No entanto, no Brasil, este item é considerado um fator importante para a melhoria da qualidade do ensino, pois os alunos brasileiros permanecem um período de tempo curto na escola, sendo, portanto, necessário aumentar a carga horária escolar.

Os trabalhos nacionais começaram por incorporar em seus estudos a relação existente entre os fatores das escolas eficazes e o desempenho dos alunos no Saeb. Estes trabalhos visam identificar quais são os fatores que contribuem para que o desempenho dos alunos melhore.

Foi com base nestes estudos que o Geres elaborou os questionários contextuais. No âmbito das pesquisas em avaliação educacional em larga escala, a explicação dos resultados cognitivos dos alunos nos testes padronizados se dá em função de indicadores contextuais, ou seja, das informações presentes nos questionários aplicados pelas pesquisas. A pesquisa Geres elaborou seis questionários no total, para: alunos, pais dos alunos, professor, diretor, escola e turma.

O pesquisador deve dispor de boas medidas sobre as características da escola e do professor, não apenas para viabilizar o controle estatístico dos resultados, mas ainda para gerar conhecimento sobre aquelas características que maximizam o aprendizado médio dos alunos e, ao mesmo tempo, minimizam as influências de origem sociocultural. É necessário, também, contar com boas medidas sobre as características econômicas, culturais e sobre os hábitos de estudo dos alunos, características profissionais de seus professores e diretores e dados sobre a escola em que estudam. Esses dados são obtidos por meio da administração de questionários contextuais. (BROOKE; BONAMINO 2011, p.45)

Os questionários dos pais e dos alunos buscaram identificar o nível educacional e a profissão dos pais e/ou responsáveis dos alunos. Além disso, esses questionários apresentam itens sobre objetos de consumo existentes nas casas dos alunos. Essas informações são necessárias para a confecção da variável nível socioeconômico das famílias. Estes questionários foram utilizados para a construção do nível socioeconômico dos alunos e da escola. O questionário da escola, preenchido por um aplicador (membro da pesquisa Geres), contém itens sobre a infraestrutura da escola. O questionário do diretor possui informações sobre sua experiência e formação profissional, características do financiamento da escola (em caso de escolas particulares), processo de nomeação do diretor (escolas públicas), relacionamento com a comunidade e critérios para a formação de turmas. O questionário do professor é composto por itens que tratam das expectativas do professor sobre o desempenho de seus alunos, estilo e liderança do diretor, colaboração entre pares, práticas em sala de aula, formação profissional, práticas culturais. O questionário da turma traz itens sobre o conforto e a existência de determinados recursos pedagógicos na sala de aula.

O primeiro passo dado na construção de escalas de medição estatística é a definição da variável latente ou conceito teórico a ser medido. De um modo geral não existe um consenso sobre o que é exatamente uma variável latente. Alguns pesquisadores a definem como uma variável que não pode ser observada diretamente e, por isso, necessita de indicadores diretamente observáveis que determinam o seu valor.

Os conceitos ou as variáveis que mais interessam aos pesquisadores sociais nem sempre podem ser observados diretamente. Muitas vezes o interesse está focalizado nos que oferecem “riqueza de significados”, tais como nível socioeconômico, preconceitos, sentimentos, satisfações etc. Os conceitos evocam diferentes imagens e noções para diferentes pesquisadores, por isso são classificados como latentes. O pesquisador precisa especificar as manifestações evocadas por eles e, ao fazer isso, estará medindo aquilo que ele convencionou ser seu conceito, ou variável latente (BROOKE; BONAMINO 2011, p.45).

Assim, o primeiro passo para a construção dos questionários foi definir a variável latente ou o conceito teórico a ser medido. Em seguida, as manifestações observáveis desta variável latente foram listadas. Essas manifestações foram, então, formuladas como itens de questionário. A combinação desses itens deve

formar uma escala, mas para saber se isso acontece com esses itens é preciso fazer uso de ferramentas estatísticas adequadas.

As escalas são mecanismos de redução de dados, de forma que as várias respostas de um único respondente podem ser reduzidas a um único valor. De acordo com Cazelli (2005), um dos aspectos relevantes para se construir uma escala é a busca pela unidimensionalidade, isto é, vários itens representam, grosso modo, apenas uma dimensão. Por exemplo, no questionário da escola há um bloco de perguntas com cinco itens. A partir desses cinco itens foi feita uma escala denominada *Existência de espaços didático-pedagógicos*. Ou seja, esses cinco itens representam uma única dimensão que trata da existência desses espaços nas escolas. Assim, ao invés de se analisar item por item, foram construídas escalas de forma que os itens que representam uma dimensão foram agrupados. Para a construção das escalas, utilizou-se o *software Mokken Scale Analysis for Polytomous Items – MSP, version 5.0 for Windows*.

Para que uma escala seja formada é preciso que os itens tenham correlações entre si. Essa correlação é verificada pelo Coeficiente de Loevinger (H) fornecida pelo MSP. Itens que tem o H inferior ao valor de referencia determinado pelo pesquisador (normalmente 0.3) são automaticamente rejeitados da escala. Os itens que possuem o H com valor igual ou menor a 0.3 possuem uma baixa correlação com os demais itens e por isso não contribuem para a escala.

O programa MSP também oferece uma medida denominada *Rho de Mokken* que afere a fidedignidade da escala. Escalas que possuem *Rho de Mokken* com valores abaixo de 0.6 são consideradas escalas de baixa fidedignidade; se esse valor estiver entre 0.6 e 0.7, média fidedignidade; e valores acima de 0.7, alta fidedignidade.

Embora a pesquisa já tivesse elaborado uma série de escalas, optou-se por fazer novas escalas utilizando apenas as variáveis de interesse sobre a infraestrutura e a gestão escolar. No total foram construídas sete escalas e três itens analisados individualmente.

3.2 Os aspectos da Infraestrutura

O Questionário da Escola foi preenchido por um aplicador da equipe Geres e nele encontram-se questões sobre a infraestrutura física das escolas. Este questionário foi dividido basicamente em quatro blocos: o primeiro aborda o estado de conservação dos diferentes itens que fazem parte do prédio escolar – telhados, paredes, piso, portas e janelas, banheiros, cozinha, refeitório, cantina, área de recreação/pátio, auditório, instalações hidráulicas e instalações elétricas. O segundo bloco traz perguntas sobre a limpeza dos espaços escolares – entrada do prédio, paredes externas, portas e janelas, área de recreação/pátio, corredores, salas de aula, banheiros, cozinha, cantina e refeitório. Para estes dois conjuntos de variáveis não se fez escalas, pois a literatura específica sobre o tema considera que a conservação e a limpeza do prédio escolar não é um fator de eficácia escolar. No Brasil, muitas instituições escolares encontram-se depredadas, sujas, pichadas, com péssimas condições de conservação de suas dependências. Apesar desta realidade, optou-se por não analisar este construto, pois as pesquisas apontam que sua influência no desempenho cognitivo dos alunos é pouco relevante até mesmo em nosso contexto. Assim, este trabalho privilegiou aspectos da infraestrutura considerados importantes para a qualidade da educação e do aprendizado dos alunos.

O terceiro bloco aborda a existência de determinados espaços didático-pedagógicos – horta, laboratório de ciências, laboratório de informática, quadra de esportes e sala para atividade de artes. Com este conjunto de variáveis formou-se uma escala. Na primeira escala formaram todos os itens. Ou seja, todos os itens apresentaram o Coeficiente de Loevinger (H) acima de 0.3 e o *Rho de Mokken* de 0.7. O pior item desta escala foi o item horta, que apresentou o H de 0.36. Sabe-se que a maioria das escolas brasileiras situa-se em áreas urbanas e não possui em suas dependências um espaço específico para a horta. Apenas algumas poucas escolas rurais possuem este espaço. De acordo com os dados do Geres, das escolas que participaram da pesquisa, 87,3% responderam que não possuem horta e 4,7% têm, mas não a utilizam. Ou seja, no total, 92% das escolas Geres não trabalham com hortas. A horta é um espaço didático-pedagógico pouco presente nas escolas e por isso optou-se por não utilizar esse item. A escala *Existência de*

espaços didático-pedagógicos (Tabela 1) foi formada, portanto, com apenas quatro itens.

Tabela 1: Média e Coeficiente de Loevinger para os itens relacionados à existência de espaços didático-pedagógicos

<i>Itens considerados</i>	Média	Coeficiente de Loevinger ⁵
Laboratório de ciências	1.46	0.58
Sala de artes	1.51	0.53
Laboratório de informática	2.14	0.47
Quadra de esportes	2.36	0.55

Fidedignidade: Rho de Mokken⁶ = 0.70. Escala de alta fidedignidade

O quarto bloco traz perguntas sobre a existência de determinados equipamentos/recursos didático-pedagógicos – material concreto de Matemática, mapas geográficos/ globos, terrário/aquário, diagramas (representações) do corpo humano/modelos anatômicos, material para atividades lúdicas, fitas de vídeo/DVD (lazer), fitas de vídeo/DVD (educativas), livros didáticos do professor, livros de consulta para os professores, computador para uso dos professores, acesso à internet para uso dos professores, computador para uso administrativo, aparelho de som, retroprojektor, televisão, vídeo cassete/DVD, canhão multimídia, mimeógrafo, máquina fotocopadora, impressora e scanner.

Essas variáveis foram reagrupadas formando dois subgrupos: um com os itens que dizem respeito aos equipamentos da escola e outro com os itens que dizem respeito aos equipamentos do professor, sendo que este último conjunto não foi analisado. Assim, fez-se uma única escala denominada *Equipamentos da escola* (Tabela 2), na qual não constam os itens relacionados aos equipamentos do

⁵ O Coeficiente *de Loevinger* é uma medida que expressa o quanto a escala se afasta da escala perfeita de Guttman, tendo como referência o modelo de Mokken. O “H” de Loevinger mede o grau de associação entre um par de itens e, entre o item e os outros que formam a escala. Cada item possui seu coeficiente e também a escala como um todo.

⁶ O Coeficiente de confiabilidade Rho de Mokken está relacionado ao número de itens na escala e é associado à confiabilidade da escala em relação ao traço latente. Ou seja, se uma escala possui coeficiente igual a 0,67, isto significa que do total da variância encontrada nas correlações entre os pares de itens, 67% está captando o traço latente e o restante é considerado ruído na mensuração. Rho de Mokken caracteriza as escalas em baixa, média e alta fidedignidade, de acordo com a segurança com que se podem ordenar os itens em relação ao traço latente que ele se propõe medir.

professor (livros didáticos do professor, livros de consulta para os professores, computador para uso dos professores, acesso à internet para uso dos professores).

Na escala formada, dois itens foram excluídos: o mimeógrafo que apresentou correlação negativa ($H = -0.21$) e o item material para atividade lúdica que apresentou fraca correlação com os demais itens ($H = 0.29$).

Tabela 2: Média e Coeficiente de Loevinger (H) para os itens relacionados aos equipamentos da escola

Itens considerados	Média	Coeficiente de Loevinger
Canhão multimedia	1.31	0.54
Terrário/Aquário	1.35	0.50
Scanner	2.18	0.43
Diagramas (representações) do corpo humano	2.36	0.44
Retroprojeter	2.37	0.35
Maquina fotocopadora	2.50	0.38
Fitas de vídeo/DVD (lazer)	2.53	0.38
Material concreto	2.68	0.37
Fitas de vídeo/DVD (educativas)	2.71	0.50
Vídeo cassette	2.85	0.54
Impressora	2.85	0.50
Mapas geográficos/globos	2.87	0.45
Computador para uso administrativo	2.89	0.44
Televisão	2.91	0.46
Aparelho de som	2.92	0.48
Itens rejeitados $H < 0.3$		
Mimeógrafo	2.50	-0.21
Material para atividade lúdica	2.85	0.29

Fidedignidade: Rho de Mokken = 0.83. Escala de alta fidedignidade

Além dessas duas escalas, foram analisados os itens: existência e utilização da biblioteca, existência e utilização da sala de leitura e existência e utilização do laboratório de informática com computadores ligados à internet.

O Questionário da Turma também apresenta questões relacionadas à infraestrutura da escola, mais especificamente, dos recursos e equipamentos da sala de aula. A partir dos itens deste questionário construiu-se uma escala denominada *Conforto da sala de aula*, englobando os itens *Isolamento do barulho externo*, *Ventilação*, *Iluminação* e *Espaço para circulação entre as mesas e*

cadeiras. Apesar de os resultados apontarem que os itens são bem correlacionados, a escala apresentou um *Rho de Mokken* insatisfatório, sendo esta, uma escala de baixa fidedignidade. Por isso esta escala não foi utilizada na análise.

Os itens existência de cantinho de leitura, quadro de giz e/ou quadro branco; tipo de mobiliário do aluno e tipo de mobiliário do professor também foram analisados, mas não serão discutidos neste trabalho pois não mostraram variações: 99,7% das escolas possuem quadro de giz e/ou quadro branco; 98% das escolas afirmaram que oferecem aos alunos mesas e cadeiras móveis e 91,4% das escolas disseram que oferecem aos professores mesas e cadeiras móveis. Esses percentuais permanecem semelhantes quando analisados por rede de ensino, por município e por nível socioeconômico, não variando e não discriminando os diferentes estratos da população estudada.

O item cantinho de leitura não foi analisado, pois 81,3% foram respostas inválidas. Além disso, das respostas válidas (18,7%), todas afirmaram possuir cantinho de leitura, sendo que apenas 6,4% dos respondentes disseram utilizar com frequência este espaço. Como este item apresentou um alto percentual de respostas inválidas também não foi analisado.

O construto infraestrutura foi composto, portanto, por cinco variáveis: 1) existência de espaços didático-pedagógicos (existência); 2) equipamentos da escola (*equipa_escola*); 3) existência e utilização da biblioteca (*biblioteca*); 4) existência e utilização da sala de leitura (*sala de leitura*) e 5) existência e utilização do laboratório de informática com computadores ligados à internet (*computador*).

3.3 Os aspectos da Gestão Pedagógica

O construto gestão pedagógica foi formado por itens que se encontram nos questionários do professor e do diretor. O primeiro foi aplicado aos professores responsáveis por turmas regulares e teve sua configuração alterada entre os anos da pesquisa. De um ano para o outro o questionário sofreu alterações, que fizeram com que algumas questões fossem excluídas e outras acrescentadas. No entanto, a

maioria dos itens permaneceu nas quatro versões (onda 2, onda 3, onda 4 e onda 5). Todos os itens analisados e que compõem as escalas estão presentes nos quatro instrumentos.

A primeira escala formada chama-se *Colaboração entre os professores* (Tabela 3). Ela foi formada por sete itens todos com boa correlação. Ou seja, nenhum item foi rejeitado da escala. Além disso, a escala apresenta o valor de *Rho de Mokken* igual a 0.86, indicando que esta é uma escala com alta fidedignidade.

Tabela 3: Média e Coeficiente de Loevinger para os itens relacionados à colaboração entre os professores

Itens considerados	Coeficiente de	
	Média	Loevinger
O projeto educacional desta é consequência da troca de ideias entre os professores	3.96	0.46
O ensino que a escola oferece aos alunos é muito influenciado pela troca de ideias entre os professores	4.11	0.56
Os professores desta escola se esforçam para coordenar o conteúdo das matérias entre as diferentes séries	4.23	0.54
A equipe de professores leva em consideração as minhas ideias	4.29	0.53
Os diretores, professores e os demais membros da equipe da escola colaboram para fazer esta escola funcionar bem	4.40	0.54
Eu levo em consideração as ideias de outros colegas	4.47	0.51
Participo das decisões relacionadas com o meu trabalho	4.48	0.46

Fidedignidade: *Rho de Mokken* = 0.86. Escala de alta fidedignidade

A segunda escala formada chama-se *Interrupção das aulas* (Tabela 4) e o item anúncios ou comunicações da direção, coordenação e/ou secretaria foi rejeitado por apresentar baixa correlação com os demais itens – o Coeficiente de Loevinger teve o valor de 0.29. Além disso, o *Rho de Mokken* apresentou o valor de 0.67, ou seja, esta é uma escala de média fidedignidade. Estes itens estão no questionário do professor.

Tabela 4: Média e Coeficiente de Loevinger para os itens relacionados à interrupção das aulas

	Média	Coeficiente de Loevinger
<i>Itens considerados</i>		
Estudantes atrasados (na entrada e na volta do recreio)	2.42	0.47
Barulho no corredor	2.57	0.41
Bagunça dos alunos	3.52	0.42
<i>Itens rejeitados</i>		
Anúncios ou comunicações da direção, coordenação e/ou secretaria	1.98	0.29

Fidedignidade: Rho de Mokken = 0.67. Escala de média fidedignidade

A terceira escala elaborada trata da *Liderança do diretor* (Tabela 5). Os itens que compõem esta escala estão no questionário do professor. Ou seja, a liderança do diretor, neste trabalho, é analisada do ponto de vista dos professores. Todos os itens apresentaram alta correlação e por isso nenhum item foi excluído da escala. Vale observar que o *Rho de Mokken* apresentou o valor de 0.95. Isso significa que esta é uma excelente escala, pois além dos itens apresentarem alta correlação, é altamente fidedigna.

Tabela 5: Média e Coeficiente de Loevinger para os itens relacionados à liderança do diretor

Itens considerados	Coeficiente de	
	Média	Loevinger
O (a) diretor (a) me anima e motiva para o trabalho	4.05	0.77
O (a) diretor (a) consegue que os professores se comprometam com a escola	4.11	0.74
O (a) diretor (a) dá atenção especial a aspectos relacionados com a aprendizagem dos alunos	4.16	0.75
O (a) diretor (a) estimula as atividades inovadoras	4.21	0.76
Tenho plena confiança profissional no (a) diretor (a)	4.26	0.79
O (a) diretor (a) dá atenção especial aos aspectos relacionados com a manutenção da escola	4.38	0.66
O (a) diretor (a) dá atenção especial aos aspectos relacionados com as normas administrativas	4.45	0.65
Sinto-me respeitado (a) pelo (a) diretor (a)	4.51	0.73
Respeito o (a) diretor (a)	4.77	0.66

Fidedignidade: Rho de Mokken = 0.95. Escala de alta fidedignidade

A escala *Falta dos alunos* (Tabela 6) foi formada com os itens presentes no questionário do diretor e teve apenas um item excluído por apresentar baixa correlação com os demais. É importante observar que esta escala tem o valor do *Rho de Mokken* de 0,56, sendo, portanto, inferior a 0.6. Ou seja, é uma escala pouco fidedigna. Embora tenha apresentado este índice, este item é importante para analisar a gestão pedagógica da escola. A falta dos alunos às aulas é um aspecto extremamente relevante que interfere no desempenho cognitivo dos estudantes. Assim, investigar que tipo de providências os diretores tomam com relação à falta dos alunos às aulas é pertinente na medida em que irá informar se a gestão está focada no aluno.

Tabela 6: Média e Coeficiente de Loevinger para os itens relacionados às providências que os professores tomam para diminuir as falta dos alunos

	Média	Coeficiente de Loevinger
<i>Itens considerados</i>		
Para diminuir as faltas dos alunos, os pais foram chamados à escola?	1.02	0.47
Para diminuir as faltas dos alunos, os professores foram orientados a falar com eles?	1.04	0.47
Para diminuir as faltas dos alunos, os pais foram avisados por escrito?	1.11	0.42
Para diminuir as faltas dos alunos, o assunto foi tratado com os pais nas reuniões escolares?	1.17	0.34
<i>Itens rejeitados</i>		
Para diminuir as faltas dos alunos, os pais foram avisados por telefone ou por um mensageiro da escola?	1.11	0.22

Fidedignidade: Rho de Mokken = 0.56. Escala de baixa fidedignidade

Além dessas escalas, o item frequência com que o professor passa dever de casa, presente no questionário do professor, também foi analisado.

No total, o construto gestão pedagógica foi formado por: 1) frequência com que o professor passa dever de casa (dever de casa), colaboração docente (colaboração), interrupção das aulas (interrupção), liderança do diretor (liderança) e falta dos alunos (falta_alunos).

3.4 A construção do nível socioeconômico

O indicador nível socioeconômico é uma variável latente, ou seja, é um construto teórico que não pode ser observado diretamente. Por isso, os questionários do aluno e dos pais possuíam itens que informam sobre a escolaridade dos pais, a ocupação dos pais e a renda da família. Com base nestes três aspectos criou-se o nível socioeconômico do aluno.

O primeiro passo para a construção do nível socioeconômico dos alunos foi a construção de um índice para os pais. Na segunda onda (final de 2005) foi enviado aos pais dos alunos um questionário com itens sobre os três aspectos necessários para a construção do NSE. O construto escolaridade foi mensurado considerando-se a escolaridade da mãe do aluno; para o construto ocupação dos

pais considerou-se apenas a medida com maior status socioeconômico. Ou seja, o questionário possui duas questões: uma sobre a ocupação da mãe e a outra sobre a ocupação do pai, mas para a construção do NSE utilizou-se apenas uma medida da mãe ou do pai do aluno que indicava maior status socioeconômico. A renda foi medida indiretamente através das informações sobre os itens de conforto (televisão a cores, DVD, TV por assinatura, máquina de lavar etc.) na casa dos alunos.

Estes três indicadores (renda, ocupação dos pais e escolaridade da mãe) foram agrupados segundo a mesma metodologia, obtendo-se, assim, um indicador de nível socioeconômico dos pais.

A segunda etapa para a construção do nível socioeconômico dos alunos foi a construção de um índice para os alunos. Para isso, replicou-se a metodologia utilizada para o cálculo do NSE dos pais, mas utilizando apenas dois indicadores: renda e escolaridade dos pais. Os itens sobre renda e escolaridade dos pais presentes no questionário dos alunos são equivalentes aos itens presentes no questionário dos pais.

Para a construção do nível socioeconômico dos alunos do Projeto Geres utilizaram-se todas as informações obtidas, tanto dos questionários dos alunos quanto dos pais. Para a construção do indicador final utilizaram-se quatro indicadores: ocupação dos pais (apenas questionário dos pais), renda (questionário dos alunos e dos pais) e dois indicadores de escolaridade dos pais (questionário dos alunos e dos pais). Assim, construiu-se uma base com quatro variáveis: as duas informações sobre escolaridade da mãe, índice estimado de renda e índice que representa a ocupação de maior status socioeconômico entre os pais. Estas variáveis foram levadas para o *software* Multilog que estimou o NSE final dos alunos Geres.

Com base neste NSE dos alunos, fez-se a média do NSE das escolas. Para analisar a distribuição das características escolares utilizamos o construto NSE médio das escolas. Na análise realizada para identificar quais são as características que possuem maior impacto na proficiência dos alunos utilizaremos o NSE dos alunos como controle.

3.5 Regressão Linear

Além de verificar a distribuição das características escolares por rede de ensino, município e nível socioeconômico médio da escola, este estudo também analisou quais são os principais fatores que impactam na proficiência dos alunos. Ou seja, pretende-se verificar quais são as características escolares que afetam o desempenho escolar dos alunos. Para isso foi criado um modelo de regressão linear. A regressão pode ser definida como uma metodologia estatística que utiliza a relação entre duas ou mais variáveis quantitativas ou qualitativas de tal forma que uma variável pode ser predita a partir da outra (ou outras). Ou seja, a variável dependente do modelo, pode ser explicada por uma ou mais variáveis explicativas. Em nosso estudo, a variável dependente é a proficiência dos alunos e os fatores que impactam ou não a proficiência são as variáveis independentes.

Em uma regressão, o primeiro coeficiente a ser analisado é o *R square* (R^2). Sua função é medir a qualidade do ajuste do modelo de regressão e explicar a proporção da variação de Y (variável dependente) em relação à média das variáveis que são incluídas no modelo de regressão. Quanto maior for o R^2 , melhor será a qualidade do ajuste do modelo. O valor de R^2 oscila entre 0 e 1. Se o R^2 for igual a 0 isto significa que as variáveis independentes não explicam a variável dependente. Por outro lado, se o R^2 for igual 1, isso indica que o modelo tem um ajuste perfeito, ou seja, as variáveis independentes explicam em 100% a variável dependente. Para as ciências exatas, o R^2 maior que 0.9 é aceito como indicador de bom ajustamento do modelo. Para as ciências sociais, esse valor é bem mais flexível, sendo aceito um valor de R^2 igual ou superior a 0.5. Para tornar a explicação mais clara, elaborou-se um modelo de regressão que ilustra a função do R^2 .

$$Profic_Aluno2_i = \beta_0 + \beta_1 NSE_Aluno + \beta_2 Profic_1 + e_i$$

Onde:

Profic_Aluno2 é a variável dependente, ou seja, a proficiência dos alunos em Português (Leitura), onda 2;

β_0 , β_1 , β_2 ... β_k são os coeficientes da regressão;

NSE_Aluno é a variável independente de controle – nível socioeconômico dos alunos;

Profic_1 é a variável independente;

e_i é o termo de erro.

A Tabela 7 informa que o R^2 tem o valor de 0.626. Isto significa que as variáveis incluídas no modelo (proficiência em Português na onda 1 e NSE dos alunos) explicam 62,6% da proficiência dos alunos em português na onda 2. No entanto, para fins de análise, utiliza-se sempre o R^2 ajustado (*Adjusted R Square*). No exemplo da Tabela 7, o R^2 ajustado tem o mesmo valor que o R^2 .

Os dados do Geres foram colhidos a partir de uma amostra. Para que seja possível generalizar os resultados para a população, o R^2 ajustado ajusta o número de variáveis do modelo de regressão para fornecer uma estimativa mais realista. Sempre que há a inclusão de uma nova variável independente no modelo, o R^2 ajustado ajusta o efeito desta variável sobre a variável dependente. Por exemplo, quando incluímos uma variável que tem efeito zero sobre a variável dependente, ou seja, uma variável que não explica a variável dependente, o R^2 irá aumentar automaticamente, enquanto que o R^2 ajustado irá corrigir este aumento, indicando que a variável incluída não ajuda a explicar a variável dependente. Assim, a utilização do R^2 ajustado é mais adequada, pois fornece uma estimativa mais realista e permite a generalização dos dados. Este trabalho irá sempre considerar o valor obtido no R^2 ajustado por ser considerado o melhor indicador da qualidade do ajustamento do modelo.

Tabela 7: Exemplo de Sumário de um Modelo de Regressão

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.791(a)	.626	.626	14.01530

a Predictors: (Constant), Zscore: Nível socioeconômico, Proficiência em português na onda 1?

A Tabela 8 traz os coeficientes estimados do modelo de regressão. Para a análise da significância estatística dos coeficientes estimados foi utilizado o *teste*

t. A estatística *t* depende do tamanho da amostra e é dada pela divisão entre o coeficiente estimado (*beta*) e o erro amostral (*St. Error*). Através do teste *t* é possível identificar se o parâmetro estimado é estatisticamente diferente de zero. Definindo um intervalo de confiança de 95%, teremos um nível de significância de 5%. O teste *t* consiste na hipótese nula de que o coeficiente da variável é zero. Portanto, se o *p-valor* (Sig.) é menor que 5%, rejeitamos a hipótese nula, ou seja, a variável é significativa para explicar o modelo.

De acordo com o exemplo, as duas variáveis apresentam significância estatística, pois todos os valores do *p-valor* são inferiores a 0.05. Assim, este modelo de regressão inclui variáveis significativas para explicar a proficiência em português na onda 2.

Tabela 8: Exemplo de Coeficientes de um Modelo de Regressão
Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta	B	Std. Error
1	(Constant)	60.835	.691		88.054	.000
	Proficiência em português na onda 1?	.611	.006	.700	98.022	.000
	Zscore: Nível socioeconômico	3.633	.162	.160	22.427	.000

a Dependent Variable: Proficiência em português na onda 2?

Os dados da Tabela 8 são interpretados da seguinte maneira: o aumento de 1 unidade na proficiência em português na onda 1 está associado ao aumento de 0,61 pontos na proficiência em português na onda 2. Com relação à variável NSE dos alunos, o aumento de uma unidade na escala do NSE, associa-se ao aumento de 3.63 pontos na proficiência em português na onda 2. Este resultado indica que tanto o conhecimento prévio quanto o nível socioeconômico dos alunos influenciam na proficiência dos estudantes, embora a magnitude dessa influência seja diferente.

Neste trabalho, os modelos de regressão vão tentar explicar quais são os fatores que influenciam a proficiência em todas as ondas (1, 2, 3, 4 e 5), em Matemática e Português (Leitura). Todos os modelos incluem a variável NSE dos alunos. A literatura sobre eficácia escolar aponta que o principal fator que explica a proficiência é a condição social dos alunos. Por isso, qualquer estudo que busca averiguar os fatores estritamente escolares que influenciam a proficiência deve inserir a variável NSE como controle. Os modelos foram rodados considerando apenas duas bases de dados.

Inicialmente as variáveis estavam em bases diferentes: as variáveis de infraestrutura estavam na base da escola, as variáveis de gestão pedagógica estavam na base do professor e na base do diretor e as variáveis proficiência e nível socioeconômico estavam na base do aluno. Assim, haviam quatro bases. Através de um comando do SPSS denominado *merge*, as bases foram juntadas, de forma que todas as variáveis foram migradas para a base dos alunos, formando uma única base com todas as informações necessárias para a análise dos dados. Essa base foi então separada em duas: uma contendo as variáveis a serem analisadas e as proficiências em Matemática nas cinco ondas, e a outra, contendo as mesmas variáveis e as proficiências em Português nas cinco ondas. Ou seja, fez-se uma base com as proficiências em Matemática e outra com as proficiências em Português, ambas contendo as variáveis de interesse.

Todas as variáveis que são contínuas (existência de espaços didático-pedagógicos, equipamentos da escola, colaboração docente, interrupção das aulas, liderança do diretor e falta dos alunos) foram padronizadas, já que as variáveis possuíam escalas e dispersões muito diferentes entre si. Foi preciso colocar as variáveis em uma mesma escala para poder fazer a comparação do impacto de cada uma sobre a proficiência. As demais variáveis foram dicotomizadas. Para isso, elas foram transformadas em *dummies*. Por exemplo, o questionário do professor pergunta se o professor passa dever de casa. Cinco são as possibilidades de resposta: “diariamente”, “algumas vezes por semana”, “semanalmente”, “raramente” e “não passa dever de casa”. Para poder incluir esta variável no modelo de regressão, optou-se por transformar estas cinco respostas em apenas duas. Ou seja, as cinco respostas foram transformadas em *dummies*. As respostas “diariamente”, “algumas vezes por semana” e “semanalmente” foram agrupadas e recodificadas para “sim”. O mesmo procedimento aconteceu para as respostas

“raramente” e “não passa dever de casa”, que foram recodificadas para “não”. Assim, a pergunta do questionário “o professor passa dever de casa” passou a ter apenas duas respostas: sim ou não. As respostas inválidas permaneceram como inválidas. Na base de dados, “sim” recebeu o valor de 1 e “não” recebeu o valor de 0. Este procedimento também foi realizado para as variáveis biblioteca, sala de leitura, computador, rede de ensino e município. O Quadro 1 descreve as variáveis utilizadas nos modelos de regressão.

Quadro 1: Descrição das variáveis utilizadas nos modelos

Sigla	Variável	Tipo	Descrição
Variáveis dependentes			
Prof_mat.	Proficiência na onda 1	Contínua	Proficiência de Matemática na onda 1
Prof_mat.	Proficiência na onda 2	Contínua	Proficiência de Matemática na onda 2
Prof_mat.	Proficiência na onda 3	Contínua	Proficiência de Matemática na onda 3
Prof_mat.	Proficiência na onda 4	Contínua	Proficiência de Matemática na onda 4
Prof_mat.	Proficiência na onda 5	Contínua	Proficiência de Matemática na onda 5
Prof_port.	Proficiência na onda 1	Contínua	Proficiência de português na onda 1
Prof_port.	Proficiência na onda 2	Contínua	Proficiência de português na onda 2
Prof_port.	Proficiência na onda 3	Contínua	Proficiência de português na onda 3
Prof_port.	Proficiência na onda 4	Contínua	Proficiência de português na onda 4
Prof_port.	Proficiência na onda 5	Contínua	Proficiência de português na onda 5
Variáveis independentes ou explicativas			
NSE	Nível socioeconômico dos alunos	Contínua	Nível socioeconômico dos alunos
Rede de Ensino	Rede de Ensino em que o aluno está matriculado	Dicotômica (1= Rede privada e especial)	Rede de Ensino em que o aluno está matriculado
Município	Município em que a escola está localizada	Dicotômica (1=Rio de Janeiro)	Município em que a escola está localizada
Biblioteca	Existência e utilização da biblioteca	Dicotômica (1=Sim)	Existência e utilização da biblioteca da escola. Questionário da escola.

Sala_Leitura	Existência e utilização da Sala de Leitura	Dicotômica (1=Sim)	Existência e utilização da sala de leitura escola. Questionário da escola.
Existência	Existência de espaços didático-pedagógicos	Contínua	Escala com quatro itens. Obtida por TRI não paramétrica a partir dos itens do questionário da escola.
Equipa_Escola	Existência de equipamentos na escola	Contínua	Escala com quinze itens. Obtida por TRI não paramétrica a partir dos itens do questionário da escola.
Computador	Existência e utilização de laboratório de informática	Dicotômica (1=Sim)	Existência e utilização do laboratório de informática com computadores ligados à internet. Questionário da escola.
Dever_Casa	Frequência com que o professor passa dever de casa	Dicotômica (1= Sim)	Frequência com que o professor passa dever de casa para os alunos. Questionário do professor.
Colaboração	Colaboração entre os professores	Contínua	Escala com sete itens. Obtida por TRI não paramétrica a partir dos itens do questionário do professor.
Interrupção	Interrupção das aulas	Contínua	Escala com três itens. Obtida por TRI não paramétrica a partir dos itens do questionário do professor.
Liderança	Liderança da escola	Contínua	Escala com oito itens. Obtida por TRI não paramétrica a partir dos itens do questionário do professor.
Falta_alunos	Falta dos alunos	Contínua	Escala com quatro itens. Obtida por TRI não paramétrica a partir dos itens do questionário do diretor.

Como dito anteriormente, este trabalho irá analisar as proficiências obtidas pelos alunos Geres em Português (Leitura) e Matemática. De acordo com a literatura, a Matemática é uma disciplina mais escolar, pois é a escola o principal suporte para a apreensão das especificidades dessa disciplina. Os conhecimentos mais específicos da Matemática dificilmente são adquiridos em um ambiente que não seja o escolar. A Língua Portuguesa, por sua vez, embora seja muito influenciada pelo contexto socioeconômico dos alunos e pelo capital cultural das famílias e, por isso, considerada uma disciplina menos escolar, é a nossa língua

materna, ou seja, é através da Língua Portuguesa que aprendemos. Para se aprender Matemática, Ciências, Geografia e todos os demais tipos de conhecimentos, é preciso saber ler e escrever. É a língua que modela grande parte do nosso o aprendizado.

Portanto, as variáveis dependentes dos modelos de regressão serão as proficiências dos alunos em Matemática e em Português (Leitura) e as variáveis independentes ou explicativas serão àquelas relacionadas com a infraestrutura e a gestão pedagógica das escolas, já descritas acima.

As variáveis relativas à infraestrutura compõem o questionário da escola. Este questionário foi respondido uma única vez no ano de 2005. Assim, todas as regressões foram rodadas considerando-se as variáveis desse ano. Por exemplo, analisa-se a proficiência obtida no teste da onda 5, que foi realizado no ano de 2008, mas as variáveis de infraestrutura referem-se ao ano de 2005. Adotou-se este procedimento em razão de as características da infraestrutura serem aspectos estáveis, que não mudam ou mudam pouco ao longo do tempo. Por isso, analisar a proficiência dos alunos no ano de 2008 com variáveis (de infraestrutura) relativas ao ano de 2005 é plausível.

No entanto, o mesmo não acontece com as variáveis de gestão pedagógica (dever de casa, colaboração, interrupção, liderança e falta dos alunos). A cada ano os professores mudam. Ou seja, na 1ª série/2º ano, os alunos possuem um professor. Na 2ª série/3º ano, é outro professor e assim por diante. Pode acontecer de os professores acompanharem os alunos ao longo de alguns anos, mas isso é raro. Via de regra, os professores mudam ano a ano. As variáveis relativas à gestão pedagógica estão presentes nos questionários do professor e do diretor. As variáveis dever de casa, colaboração, interrupção e liderança compõem o questionário do professor. Este questionário foi respondido nas ondas 2, 3, 4 e 5. Assim, o modelo de regressão rodado para a onda 1, utilizou as variáveis da onda 2, que corresponde ao mesmo ano letivo (2005). Os demais modelos utilizaram as informações obtidas nos questionários das respectivas ondas.

A variável falta dos alunos compõe o questionário do diretor. Este questionário foi respondido no ano de 2005. No entanto, quando um novo diretor assumia o cargo, este questionário era respondido novamente pelo novo diretor. Como a mudança da direção é algo que acontece de tempos em tempos, ou seja,

não é uma mudança frequente, considerou-se apenas a variável do questionário aplicado no ano de 2005. O Quadro 2 ilustra a explicação acima.

Quadro 2: Variáveis utilizadas nos modelos de regressão

Variável	Questionário	Ano em que o questionário foi respondido
NSE	Questionário do aluno	2008
Existência	Questionário da escola	2005
Equipa_Escola	Questionário da escola	2005
Biblioteca	Questionário da escola	2005
Sala_Leitura	Questionário da escola	2005
Computador	Questionário da escola	2005
Dever_Casa_1	Questionário do professor – onda 2	2005
Dever_Casa_2	Questionário do professor – onda 2	2005
Dever_Casa_3	Questionário do professor – onda 3	2006
Dever_Casa_4	Questionário do professor – onda 4	2007
Dever_Casa_5	Questionário do professor – onda 5	2008
Colaboração_1	Questionário do professor – onda 2	2005
Colaboração_2	Questionário do professor – onda 2	2005
Colaboração_3	Questionário do professor – onda 3	2006
Colaboração_4	Questionário do professor – onda 4	2007
Colaboração_5	Questionário do professor – onda 5	2008
Interrupção_1	Questionário do professor – onda 2	2005
Interrupção_2	Questionário do professor – onda 2	2005
Interrupção_3	Questionário do professor – onda 3	2006
Interrupção_4	Questionário do professor – onda 4	2007
Interrupção_5	Questionário do professor – onda 5	2008
Liderança_1	Questionário do professor – onda 2	2005
Liderança_2	Questionário do professor – onda 2	2005
Liderança_3	Questionário do professor – onda 3	2006
Liderança_4	Questionário do professor – onda 4	2007
Liderança_5	Questionário do professor – onda 5	2008
Falta_alunos	Questionário do diretor – onda 2	2005

Estas variáveis foram agrupadas dando origem a uma nova base de dados, utilizada para gerar as informações e análises deste estudo.

O Geres é uma pesquisa que tem um desenho longitudinal, ou seja, permite acompanhar a evolução do aprendizado dos alunos ao longo do tempo. Por ter esta característica, os dados da pesquisa permitem analisar quais são os aspectos que mais impactam a proficiência nos anos iniciais do ensino fundamental e verificar quais são as características mais importantes em uma determinada série/ano escolar. O estudo irá mostrar, mais adiante, que certos

fatores escolares exercem uma influência maior na proficiência dependendo da série/ano escolar frequentado pelo aluno. Assim, foram rodadas, no total, oito regressões, uma para cada onda em Matemática e uma para cada onda em Português, a partir da onda 2.

As análises referem-se apenas aos alunos da Geração Geres, ou seja, aos alunos que possuem proficiência em todas as ondas da pesquisa. Em Matemática, 8796 alunos possuem proficiência em todas as ondas. Em português, este valor é um pouco menor, sendo 8737 o total de alunos que possuem proficiência nas cinco ondas. Optou-se por analisar apenas os alunos que possuem proficiência em todas as ondas, pois estes alunos integraram o painel da Geração Escolar 2005 e estiveram, ao longo das cinco ondas, submetidos às mesmas condições, tanto de infraestrutura quanto de gestão pedagógica. São alunos que estiveram sob a influência dos mesmos fatores escolares ao longo dos quatro primeiros anos do Ensino Fundamental. Na análise não há alunos que entraram no meio da pesquisa, ou seja, que tivessem trazido consigo influências de outras escolas.

O próximo capítulo está dividido em duas seções. A primeira, de natureza descritiva, apresenta a distribuição das características escolares por rede de ensino, por município e por nível socioeconômico. A segunda seção analisa os resultados dos dois modelos de regressão.