

## QUESTIONÁRIOS CONTEXTUAIS DO SAEB: ANÁLISE QUALITATIVA BASEADA NOS RESULTADOS DA ANÁLISE FATORIAL

*Raíssa Rauter, Jacob Arie Laros y Girlene Ribeiro de Jesus*

### 1. INTRODUÇÃO

Implantado em 1990, o SAEB vem evoluindo tecnicamente de maneira significativa a cada ciclo. Em 2001 a construção dos questionários contextuais foi baseada em um quadro de referência, definido e explicitado em documento divulgado pelo Instituto Nacional de Estudos e Pesquisas Educacionais (INEP) (INEP, 2001). No mesmo documento, o INEP afirma que a divulgação do referencial teórico e quadro de referência têm o propósito de possibilitar o debate e viabilizar o processo de aprimoramento dos questionários contextuais.

O quadro de referência para os questionários contextuais indica os construtos que se pretende investigar e as questões elaboradas para medir cada um deles. Para cada nível do sistema de ensino (escola, turma e aluno), há um quadro de referência com as características a serem medidas e os itens do questionário relacionados a cada uma delas. O quadro relativo às características da escola, ou aos construtos relacionados a elas, indica que foram elaboradas 163 questões para medir 12 construtos teóricos abrangentes. Embora estejam esses construtos previamente definidos, somente uma análise empírica dos resultados das questões pode mostrar se os construtos teóricos estão sendo efetivamente medidos na prática.

A opção por estudar os construtos relacionados a características da escola surge com o intuito de contribuir com a busca dos fatores escolares que incidem no rendimento escolar dos alunos. Na bibliografia pesquisada, há poucos estudos que mencionam o método de agregação de variáveis nos fatores, seja ele a teoria ou algum método estatístico, como a análise fatorial. A tendência é a formação de fatores com base na teoria, reunindo questões com conteúdo semelhante. Muitos estudos feitos com dados do SAEB usam fatores obtidos dessa forma. Há registros de emprego de análise fatorial, especialmente para obtenção do índice de nível socioeconômico dos alunos (Albernaz *et al.*, 2002; Soares y Alves, no prelo; Barbosa y Fernandes, 2001).

Barbosa (2000) utilizou análise fatorial, coordenando um trabalho de redução da dimensão de todas as bases de dados do SAEB 1999. Nesse estudo, foi utilizada análise fatorial aplicada em blocos de variáveis homogêneas quanto ao conteúdo, nos questionários dos alunos, professores e diretores, separadamente. Primeiramente, a teoria orientou o agrupamento de variáveis relacionadas a uma mesma característica. Em seguida, cada grupo de variáveis foi submetido à análise fatorial, para a redução do número de variáveis visando a análise de cada questionário do SAEB 1999 individualmente.

Um dos objetivos desse estudo é reduzir o grande número de variáveis relacionadas à escola, que compõem os questionários contextuais do SAEB 2001 aplicados na 4ª série do ensino fundamental, por meio da análise fatorial. Ademais, o objetivo principal é identificar possíveis razões pelas quais algumas questões não se inserem na estrutura fatorial empírica. Essas razões podem estar relacionadas ao cumprimento ou não de critérios de elaboração de itens (Pasquali, 1999; Günther, 1999; Clark y

Watson, 1995). A identificação de problemas orienta a reformulação das questões para próximos ciclos do SAEB, o que, por sua vez, possibilita o aprimoramento da estrutura fatorial.

Para tanto, foi escolhida a amostra de escolas com alunos de 4ª série do Ensino Fundamental, que participaram do SAEB em 2001, fazendo provas de Matemática. A escolha de escolas com turmas de 4ª série do Ensino Fundamental foi motivada pelo fato de serem essas as escolas que refletem o ambiente onde os alunos brasileiros concluem o ciclo mais básico e fundamental de ensino. Os primeiros anos de escolarização têm sido apontados por muitos estudos como decisivos para o progresso e desempenho escolar e profissional subseqüentes (Rosales, 2000; Teddlie y Reynolds, 2000).

## 2. MÉTODO

### 2.1. Amostra

Foram considerados os dados referentes a 4.065 escolas nas quais 57.258 alunos de 4ª série do Ensino Fundamental fizeram testes de Matemática. Esses números se referem à amostra efetiva, representando os alunos que preencheram pelo menos um item da prova de Matemática e as escolas a eles relacionadas.

A fim de verificar a replicabilidade dos resultados da análise, a amostra total de escolas foi dividida em duas sub-amostras independentes, de tamanhos e perfis muito semelhantes: sub-amostra de derivação e sub-amostra de verificação. A análise fatorial realizada na sub-amostra de derivação foi repetida na sub-amostra de verificação e a composição dos fatores obtidos foi comparada.

Como a amostra total é estratificada, a simples divisão aleatória dos elementos poderia gerar um viés nas características de cada sub-amostra. Por isso, através de uma programação em SAS (versão 8.02, de 2001), os totais de escolas dentro de cada um dos 147 estratos foram aleatoriamente divididos em dois grupos. A sub-amostra de derivação foi composta de 2.063 escolas e a sub-amostra de verificação foi de 2.002 escolas.

### 2.2. Os instrumentos do SAEB 2001

Nos questionários aplicados havia 163 questões para medir os construtos relacionados à escola, distribuídas conforme mostra a Tabela 1.

TABELA 1. QUESTÕES DO SAEB 2001 DESTINADAS A MEDIR CONSTRUTOS RELACIONADOS À ESCOLA

Questionário	Número de questões	Questões
Escolas	36	E1.1 a E8
Diretores	72	D01 a D72
Turmas	5	T3, T4, T5, T7 e T8
Professores	47	P49 a P95
Alunos	3	A34 a A36 do questionário 2
Total	163	

### 2.3. Análise dos dados

Quanto melhor a qualidade das variáveis que entram na análise fatorial, melhor a qualidade dos resultados obtidos. O processo de construção das questões deve ser muito rigoroso para que as análises resultantes dos dados sejam pertinentes. Como aqui se tratam de dados secundários, foi feita

uma extensa análise exploratória das questões, para assegurar, na medida do possível, que somente questões de alta qualidade fossem inseridas na análise.

Para seleção das questões, entre as 163 totais, cinco critérios foram adotados: (a) escala de medida; (b) distribuição das respostas nas alternativas; (c) comportamento na análise fatorial prévia; (d) correlação ítem-total corrigida dentro do fator; e (e) similaridade entre o conteúdo da questão e o domínio teórico do construto.

A qualidade dos dados para análise fatorial se manifesta também em função da existência de correlações no mínimo moderadas ( $r$  de aproximadamente 0,20) entre as variáveis. Porém, considerando-se o número de variáveis, tratam-se de 11 mil correlações bivariadas a serem inspecionadas. Quando o número de questões é muito alto, alguns autores recomendam o uso da análise fatorial na seleção prévia dos itens. Questões com propriedades psicométricas fracas são naturalmente eliminadas (Laros, no prelo; Floyd y Widaman, 1995).

Um processo iterativo de análises fatoriais foi feito para seleção das questões. O processo parou quando não havia mais variáveis que se encaixassem em algum dos critérios de eliminação estabelecidos, que foram: (a) carga menor que 0,33, em valores absolutos, em todos os fatores (explicando menos de 10% da variância); (b) carga de pelo menos 0,33, em valores absolutos, em dois ou mais fatores, com diferença entre elas menor que 0,10; e (c) ser a única variável com carga de pelo menos 0,33, em valores absolutos, em um fator.

Em todas as análises desse processo iterativo, foram usados os mesmos métodos de análise da fatorabilidade da matriz de correlações, definição do número de fatores, extração e rotação dos fatores. Para maiores detalhes sobre os métodos de análise fatorial empregados, consultar Johnson y Wichern (1998), Kim y Mueller (1987a e 1987b), Tabachnick y Fidell (1996) ou Floyd y Widaman (1995).

Como primeiro passo, o método de análise de componentes principais foi usado para análise da fatorabilidade da matriz  $R$  de correlações e para definição do número de fatores a serem extraídos. A matriz  $R$  foi analisada em termos do percentual de correlações maiores ou iguais a 0,30, em números absolutos, e do teste de adequação da amostra de Kaiser-Meyer-Olkin (KMO).

Para variáveis categóricas, a matriz de correlações não-paramétricas *Kendall's tau b* seria mais adequada, mas o SPSS não retornou essa matriz devido ao grande número de variáveis. Foi gerada uma matriz de correlações *Kendall's tau b* e também de *Pearson* entre uma parte das variáveis. A correlação média entre os dois tipos de correlação foi de 0,985, variando de 0,970 a 0,998. Diante disso, decidiu-se utilizar a matriz de correlações *Pearson*, gerada automaticamente no comando para análise fatorial ou componentes principais.

A definição do número de fatores a extrair é considerada a mais importante etapa de decisão em uma análise fatorial. Entre os vários métodos, a análise visual do gráfico *scree* (número de fatores por valores dos *eigenvalues*) é considerada um método bastante eficiente (Zwick y Velicer, 1985; Zoski & Jurs, 1996; Laros, no prelo), com a restrição, porém, do risco de subjetividade na avaliação visual por diferentes pesquisadores. Além disso, quando o número de variáveis é muito grande, a clareza na identificação visual do ponto que define o número de fatores no gráfico fica bastante comprometida.

Como alternativa para o método de análise visual, Zoski y Jurs (1996) propuseram o Método *Scree* do Erro Padrão (*Standard Error Scree Method*), que foi adotado neste estudo.

Definido o número de fatores, a extração foi feita usando-se a fatorização dos eixos principais (*Principal Axes Factoring* – PAF). Para rotação dos fatores e, portanto, para simplificação da solução fatorial, adotou-se o método *Promax*.

O passo seguinte foi a análise da consistência interna de cada fator, a partir de dois critérios: o coeficiente alfa ( $\lambda^2$  de Guttman) e as correlações item-total. Foi considerada pobre a questão que apresentasse correlação item-total menor que 0,30. Quanto ao valor mínimo aceitável para o alfa, usou-se um critério de 0,60.

Finalmente, é importante que se faça distinção entre questões que são protótipos do construto medido pelo fator e questões que estão altamente correlacionados com o construto que o fator está medindo (Smith y McCarthy, 1995). Algumas questões não se mostraram interpretáveis no contexto do fator, apesar de apresentarem correlação significativa com as demais questões do grupo. Para garantir a validade do fator, decidiu-se por eliminá-las.

Com exceção do método de definição do número de fatores a serem extraídos, todos os procedimentos de análise fatorial para encontrar os fatores de segunda ordem foram idênticos aos adotados na análise fatorial que gerou os fatores de primeira ordem. Para definir o número de fatores de segunda ordem a serem extraídos, usou-se o método de análise paralela (Horn, 1965).

A replicabilidade da estrutura fatorial obtida na sub-amostra de derivação foi checada na sub-amostra de verificação.

### 3. RESULTADOS

Das 163 questões construídas para medir construtos relacionados à escola, apenas 83 compõem a estrutura fatorial resultante da análise. A fase de seleção das questões que entram na estrutura fatorial é crucial para se ter um produto de qualidade no final da análise. Itens problemáticos, com propriedades psicométricas fracas, podem prejudicar muito os resultados.

Como foi descrito no Método, foram aplicados cinco critérios na seleção das questões para a estrutura fatorial. As seguintes questões não foram incluídas na análise: (1) questões nominais, sem possibilidade de recodificação das alternativas; (2) questões com baixa variabilidade de respostas (mais de 90% delas concentradas em uma alternativa); (3) questões com comportamento inadequado na análise fatorial inicial; (4) questões com correlação item-total corrigida menor que 0,30 dentro do fator; e (5) questões com falta de similaridade entre o seu conteúdo e o domínio teórico do construto medido pelo fator. A Tabela 2 resume o número de variáveis que não foram incluídas, a partir da aplicação de cada um desses critérios.

TABELA 2. NÚMERO DE QUESTÕES NÃO-INERIDAS NA ESTRUTURA FATORIAL POR CRITÉRIO ADOTADO

Critério de não-inclusão de questões	Número de questões excluídas
1. Variáveis nominais	8
2. Baixa variabilidade de respostas	6
3. Comportamento inadequado na análise fatorial inicial	60
4. Baixa correlação item-total no fator	4
5. Falta de similaridade com conteúdo do construto	2
Total	80

### 3.1. Possíveis razões de não-inclusão de questões na estrutura fatorial

Depois da exclusão das questões pela aplicação dos critérios mencionados, foi verificada uma provável razão pela qual essas questões apresentaram características psicométricas insatisfatórias. A análise do texto, das alternativas e da estrutura das questões sugere motivos pelos quais não houve desempenho adequado no processo de seleção das variáveis. Em alguns casos, não há motivo aparente, mas, em muitos deles, pode-se apontar problemas de construção dos itens que, se forem reelaborados, podem vir a medir os atributos com mais precisão e aprimorar a validade da estrutura fatorial.

Os procedimentos teóricos para a construção de instrumentos de mensuração envolvem a criteriosa construção de itens e suas alternativas, bem como das instruções para preenchimento. Foram usados, como referência para avaliação, os critérios de construção de itens definidos por Pasquali (1999) e as recomendações encontradas em Clark y Watson (1995) y Günther (1999).

A Tabela 3 indica seis critérios teóricos definidos para análise das questões, bem como a incidência de falha em cada um deles, por questionário. Não há referência ao questionário de alunos porque as três questões que medem construtos da escola por esse questionário se mantiveram na estrutura fatorial. Observe-se que não se pôde observar problemas de construção em todas as questões. Algumas delas, apesar de não terem se inserido na estrutura fatorial, não apresentaram problema aparente.

A maior incidência foi de questões com falta de simplicidade na construção. Segundo esse critério, a questão deve ser formulada da forma mais objetiva possível. Frases longas e detalhamentos desnecessários cansam a leitura e, principalmente, tornam o questionário muito extenso. A conseqüente falta de disposição para responder o instrumento pode comprometer a validade das respostas, especialmente nos últimos itens do questionário. Um exemplo é a questão E4.6: “*As salas onde são guardados os equipamentos mais caros (computadores, projetores, televisão, vídeo etc.) têm dispositivos para serem trancadas (cadeados, grades, travas, trancas etc)?*” Tirando-se as observações entre parênteses, percebe-se que a questão permanece com o mesmo sentido e se tornaria bem mais leve: “*As salas com os equipamentos mais caros têm dispositivos para serem trancadas?*”

TABELA 3. CRITÉRIOS DE ANÁLISE DAS QUESTÕES NÃO INCLuíDAS NA ESTRUTURA FATORIAL FINAL E INCIDÊNCIA DE CASOS COM PROBLEMA, POR QUESTIONÁRIO

Critério	Número de questões com problema, por questionário				
	Escola	Diretor	Professor	Turma	Total
Número de questões excluídas, com problema*	6	28	13	0	47
Simplicidade	3	18	10	0	31
Precisão	1	16	8	0	25
Relevância	0	5	0	0	5
Clareza	2	2	0	0	4
Variedade	0	0	2	0	2
Modalidade	0	0	1	0	1
Número de questões excluídas, sem problema aparente	4	21	4	5	34
Total de questões excluídas	10	49	16	5	80

\* Há questões com mais de um tipo de problema de formulação, tendo sido contabilizadas duas vezes na Tabela. Por isso, o total de questões com problema pode ser inferior à soma do número de questões em cada categoria.

Pelo critério da precisão, a questão deve medir apenas um atributo, de modo distinto de qualquer outro item do instrumento. Falhas nesse critério também foram bastante frequentes. Há casos com mais de uma pergunta no texto da questão ou mais de uma resposta em cada alternativa. Questões desse tipo são muito complexas, dificultam a análise das respostas e o respondente pode ter dificuldade de se posicionar diante delas.

As questões P86 a P95 são, por exemplo, pouco precisas. Essas questões formam um bloco sobre o uso de recursos pedagógicos pelos professores e cujas alternativas de respostas são: “(A) Sim, uso; (B) Não utilizo, porque a escola não tem, mas acho necessário; (C) Não utilizo, porque não acho necessário”. Como se vê, a questão mede três atributos: se os recursos pedagógicos são utilizados, se a escola possui esses recursos e se o professor acha necessária a sua utilização. Além da falta de precisão de medida, as alternativas não seguem regra básica na construção de instrumentos de pesquisa por não cobrirem todas as opções possíveis de respostas. Caso o professor ache necessário o uso e a escola tenha o recurso, mas, por algum motivo, o professor não use, não há opção de resposta para preenchimento.

Outro exemplo de questão pouco precisa é a E2.3: “*O volume dos ruídos externos prejudica a concentração dos alunos? É preciso falar em voz muito alta para se fazer ouvir?*” O respondente está exposto a duas perguntas diferentes em uma mesma questão. Caso a resposta a uma delas seja diferente da resposta à outra, pessoas com a mesma posição podem marcar alternativas diferentes. Esse item foi eliminado na primeira análise fatorial efetuada para seleção de variáveis, por ter apresentado correlação muito fraca com todos os demais. Isso significa que pessoas que respondem de modo semelhante a outros itens têm comportamentos variados diante desse, confirmando empiricamente a análise teórica dessa questão.

Pelo critério da relevância, o conteúdo da questão deve estar relacionado com um atributo do construto a ser medido. Questões pouco relevantes trazem pouca informação e devem ser evitadas. Dessa forma, o questionário fica mais sucinto, aprimorando a validade da medida de outros atributos mais centrais para o construto.

Há questões em que não se identificou relação clara com fatores escolares e foram consideradas pouco relevantes. Por exemplo, a questão D67: “*No início deste ano letivo, houve uma procura maior que a oferta de vagas nesta escola?*” Considerando os objetivos a que o SAEB se propõe<sup>1</sup>, não se pode identificar a relevância dessa pergunta no questionário. Entre as 72 questões totais para o diretor, algumas perguntas a menos fariam uma diferença positiva na validade das medidas.

Em quatro situações, pode-se verificar problemas de clareza nos termos empregados. As regras de construção de instrumentos incluem o uso de termos simples, compreensíveis por todos os estratos da população-alvo e de forma igual por todos, sem viés para grupos específicos. Questões formuladas em forma de negação devem ser evitadas.

Entre as questões com pouca clareza, está E1.8, em que o respondente deve manifestar sua avaliação sobre as “*instalações hidráulicas*” da escola. Em casos como esse, pode haver problemas de compreensão do termo por algumas pessoas e, nesse caso específico, de como acessar as instalações hidráulicas da escola para que sejam avaliadas.

---

<sup>1</sup> Ver INEP, 2001, p. 45.

O critério da variedade é aplicável a conjuntos de questões e não a questões individuais. Deve-se variar a linguagem. Estruturas de frases e termos repetidos em um bloco com várias questões são desmotivantes para o respondente. As questões P59 a P61 são um exemplo desse caso, em que a expressão “*O (A) diretor(a) dá atenção especial a aspectos relacionados com*” é repetida três vezes, seqüencialmente.

Como mencionado anteriormente, algumas questões não apresentaram variabilidade suficiente de respostas para entrarem na análise fatorial. A concentração de mais de 90% das respostas em apenas uma alternativa pode representar uma informação verdadeira ou, ao contrário, os respondentes podem ter tido a tendência a optar pela resposta socialmente mais desejável. A possibilidade de ocorrência dessa última situação é uma limitação dos *surveys* (ANEP, 1999). Pesquisas feitas por observação podem ser mais fidedignas para avaliar a qualidade do ensino, mas, no caso do SAEB, essa técnica seria pouco viável por razões de cobertura, custo e tempo.

Ao se elaborar o instrumento para o *survey*, o critério da modalidade deve ser seguido para se evitar, na medida do possível, a indução da resposta na direção da alternativa mais desejável. Na questão P51, a concentração das respostas pode estar associada à falha na aplicação desse critério, quando pergunta se o professor “*muda o planejamento das aulas, de modo a viabilizar a aprendizagem de conceitos relevantes que não foram apreendidos pelos alunos*” (grifo meu).

Com relação às demais questões com baixa variabilidade de respostas, não parece haver indução de resposta. A não-inclusão dessas questões na análise fatorial não significa necessariamente que não devam ser incluídas no instrumento, desde que haja adequação delas ao conteúdo do construto medido (Pasquali, 1999). Naturalmente que se o objetivo for ter variabilidade de resposta para comparação entre diferentes grupos, a questão não está sendo funcional. Mas se o propósito for simplesmente a caracterização da população, a concentração das respostas em uma alternativa não deixa de ser um dado.

No caso do SAEB, porém, os questionários são bastante longos, o que, como já foi dito, pode comprometer a validade das medidas. Por isso, aconselha-se que não sejam mantidas questões com baixa variabilidade de respostas, de modo que se possam medir fatores mais relevantes e com maior qualidade.

Com relação à tipicidade e à credibilidade, todas as questões parecem ter satisfeito essas regras. O critério da tipicidade refere-se ao uso de termos que façam parte do universo do atributo, que, no caso, é o meio escolar. O critério da credibilidade, por sua vez, prega que não sejam usadas expressões infantis ou excessivamente fáceis. Esse último é o extremo do critério da clareza, podendo irritar ou ofender o respondente.

Analisando-se os questionários como um todo, pode-se apontar também algumas falhas relativas às regras básicas de elaboração de instrumentos de pesquisa. Há, por exemplo, problemas gerais na simplicidade das orientações. Repetidas indicações de flexibilização do gênero das palavras (“*Sinto-me respeitado(a) pelo(a) diretor(a)*”) são dispensáveis e poluem o questionário. Basta uma simples menção prévia de que o gênero masculino ou o feminino será adotado em todo o instrumento, solicitando-se que o respondente flexibilize o gênero, caso não se aplique o adotado no instrumento.

Em todos os questionários, há trechos em que vários itens contêm a observação “Marque apenas UMA alternativa”. A repetição alonga o questionário, polui sua aparência, desmotiva e, eventualmente, pode até ofender o respondente. Há blocos de itens do formato SIM/NÃO com a

observação “*Marque SIM ou NÃO em cada linha*”. A indicação é desnecessária e também pode ser ofensiva, especialmente quando o respondente é o diretor ou o professor.

A orientação prévia ao bloco das questões P55 a P69, além de dispensável, não está correta. As questões têm quatro alternativas e a orientação é que seja preenchida “*uma das duas alternativas*” (grifo meu). Erros como esse devem ser evitados ao máximo, porque podem comprometer a credibilidade da pesquisa e a validade das respostas.

A Tabela 4 resume o número de questões excluídas e, entre elas, as questões onde foram identificados problemas de formulação. Os percentuais indicam que o questionário dos diretores é o que apresenta maiores problemas: 68% das questões foram excluídas e puderam ser identificados problemas em 39% delas.

As três questões do questionário de turmas não apresentaram correlação com as demais questões que medem fatores da escola, de modo que nenhuma delas se manteve na estrutura fatorial. Mesmo assim, não se observou nenhum problema de formulação nessas questões.

TABELA 4. QUESTÕES EXCLUÍDAS: TOTAL E QUESTÕES COM PROBLEMAS DE FORMULAÇÃO, POR QUESTIONÁRIO

Questionário	Total de questões	Questões excluídas			
		Total	Total % sobre o total de questões	Total	Questões excluídas, com problema % de questões com problema sobre o total de excluídas
Escola	36	10	28%	6	60%
Diretor	72	49	68%	28	57%
Professor	47	16	34%	13	81%
Turma	5	5	100%	0	0%
Aluno	3	0	0%	0	-
Total	163	80	49%	47	59%

### 3.2. Estrutura fatorial

Foram identificados 19 fatores de primeira ordem subjacentes às 83 questões incluídas na análise fatorial. Esses 19 fatores respondem por 63% da variância nas variáveis observadas. O nome de cada fator é uma referência curta ao seu conteúdo. Os fatores estão listados a seguir.

TABELA 5. FATORES DE PRIMEIRA ORDEM

Fatores	Nº de itens	$\lambda_2$
F1 - Instalações físicas	15	0,92
F2 - Liderança do diretor	8	0,91
F3 - Recursos materiais e didáticos	10	0,88
F4 - Comprometimento dos professores (P)	5	0,91
F5 - Comprometimento dos professores (D)	5	0,89
F6 - Atitude positiva do alunado	3	0,87
F7 - Clima disciplinar	5	0,63
F8 - Formação continuada do diretor	3	0,79
F9 - Experiência do diretor	2	0,89
F10 - TV e vídeo	2	0,90
F11 - Segurança	3	0,71
F12 - Vídeos didáticos	3	0,67
F13 - Reforço escolar	2	0,77
F14 - Recursos financeiros e pedagógicos	4	0,64
F15 - Colaboração entre professores	3	0,78
F16 - Equipe pedagógica	3	0,64
F17 - Políticas contra repetência e abandono	2	0,69
F18 - Estabilidade e assiduidade dos professores	3	0,60
F19 - Não ocorrência de roubos e depredações	2	0,70

A matriz de correlação entre os 19 fatores contém cerca de 16% das correlações maiores que 0,25, sendo 10% maiores que 0,30, em valores absolutos. Isso sugeriu a existência de fatores mais gerais, de ordem superior, interpretados como causa da covariância entre esses 19 fatores, chamados de fatores de segunda ordem.

Como os fatores de primeira ordem, além de numerosos, são, em sua maioria, muito específicos, seria muito interessante que se pudesse identificar fatores mais abrangentes, subjacentes a eles. Com isso, a informação poderia, inclusive, ser ainda mais condensada.

Escores fatoriais dos 19 fatores de primeira ordem foram formados pela média dos valores padronizados das questões que compunham cada fator. Os escores padronizados foram as variáveis submetidas à nova análise fatorial.

Na matriz de correlação entre os escores fatoriais, observou-se 11% de correlações maiores que 0,25. O KMO foi de 0,75. A análise paralela apontou a existência de cinco fatores de segunda ordem. Nessa primeira extração, F8, F13 e F18 foram eliminados pelos mesmos critérios adotados na seleção das questões para fatores de primeira ordem.

Excluindo-se esses três escores, o índice KMO permaneceu em 0,75. Cinco fatores de segunda ordem foram novamente extraídos. Ainda três escores (F1, F11 e F17) foram eliminados. Uma terceira e última extração de cinco fatores de segunda ordem foi realizada.

Os cinco fatores, que explicaram 61% da variância dos dados, são os seguintes:

TABELA 6. FATORES DE SEGUNDA ORDEM

Fatores	$\lambda_2$	Nº de fatores	Nº de itens
F1 - Confiança, cooperação e comprometimento do corpo docente	0,91	3	16
F2 - Recursos materiais e didáticos	0,88	3	15
F3 - Pessoal e recursos pedagógicos	0,69	2	7
F4 - Clima disciplinar	0,68	2	7
F5 - Clima acadêmico	0,83	3	10

### 3.3. Validação cruzada: verificando a replicabilidade da estrutura fatorial

A replicabilidade dos fatores de primeira ordem foi verificada pela extração de 19 fatores das 83 questões incluídas na estrutura fatorial, em uma sub-amostra independente (sub-amostra de verificação). Os mesmos métodos adotados na sub-amostra de derivação foram utilizados na extração dos fatores na sub-amostra de verificação.

Observou-se uma replicabilidade muito boa do primeiro nível da estrutura fatorial. Dos 19 fatores identificados na sub-amostra de derivação, 12 apresentaram exatamente a mesma composição de questões, 6 tiveram uma estrutura muito semelhante e apenas um, o fator F17, não se mostrou replicável. Os fatores com composição idêntica foram: F1, F2, F4, F5, F6, F8, F9, F10, F12, F13, F15 e F18.

Com relação aos fatores com estrutura semelhante, observou-se uma pequena variação na distribuição de questões entre fatores de primeira ordem que, na sub-amostra de derivação, agregaram-se em fatores de segunda ordem. Isso significa que a redistribuição das variáveis, na verificação, está associada a uma correlação entre elas que, na derivação, se manifestou no segundo nível da estrutura fatorial.

## 4. DISCUSSÃO

A estrutura fatorial resultante das análises pretende ser um passo a mais para o aprimoramento das medidas de fatores escolares pelos questionários contextuais do SAEB. Esses questionários são instrumentos jovens, que ainda precisam ser aperfeiçoados. Esse fato se reflete em algumas limitações encontradas nos dois níveis da estrutura fatorial.

Primeiramente, a estrutura fatorial capta a informação de pouco mais da metade das questões elaboradas para medir os construtos relacionados à escola no SAEB 2001. Algumas delas são nominais, o que já é uma razão natural de exclusão da análise fatorial. Ainda assim, pode-se observar problemas de formulação em quase 60% das questões excluídas. Essas questões têm chance de serem assimiladas pela estrutura fatorial, caso sejam reformuladas.

Outra limitação refere-se à abrangência do conteúdo. Os fatores mais abrangentes e com melhores propriedades psicométricas, em geral, medem construtos relacionados a aspectos físicos e materiais da escola. Uma boa medida desses construtos é um resultado positivo, já que esse fator é enfatizado nos estudos brasileiros e latino-americanos e é importante que seu impacto seja verificado e acompanhado. Por outro lado, é altamente desejável que se possa medir também com qualidade fatores

relacionados à gestão escolar e outros aspectos não restritos aos recursos materiais. O efeito desses fatores poderia, assim, ser verificado com maior precisão, abrindo a possibilidade de se perceber um espaço de atuação da escola na qualidade do ensino para além dos recursos materiais de que dispõe.

O que se percebe, porém, é que a maior parte dos 19 fatores de primeira ordem tem conteúdo muito específico, sendo formados por poucas questões, com perguntas muito semelhantes em um mesmo questionário, ou iguais em questionários diferentes. Alguns desses fatores apresentam um alfa alto, de 0,80 ou mais, indicando alta consistência interna. Mas esse valor reflete a correlação entre as questões que compõem o fator e, muitos deles têm questões com conteúdo muito similar. O que se deseja não é somente uma boa consistência interna, mas também uma boa validade da medida, com um número suficiente de questões claramente distintas para que sejam abrangentes no conteúdo do construto.

Não obstante as limitações expostas, pode-se observar uma boa replicabilidade dos fatores de primeira ordem, manifestada na sub-amostra de verificação, especialmente no que diz respeito aos fatores de primeira ordem. A replicação dos fatores de primeira ordem só não foi perfeita, possivelmente, devido à ocorrência de baixas correlações entre as questões.

Quanto ao segundo nível da estrutura fatorial, a replicabilidade não foi tão boa. Como os fatores de segunda ordem são resultado da análise fatorial dos escores fatoriais de primeira ordem, a falta de estabilidade pode estar relacionada com o chamado grau de indeterminação dos escores fatoriais, para o qual Grice (2001) chama atenção.

Outra sugestão diz respeito à técnica de identificação de variáveis latentes subjacentes às questões. A análise fatorial é vastamente utilizada no campo da Psicologia e gera bons resultados; embora os dados do presente estudo violem o pressuposto da normalidade requerido por essa análise, Tabachnick e Fidell (1996) referem que ela é robusta quanto à fuga desse pressuposto. Neste estudo, a validação cruzada mostrou que os fatores observados apresentam uma boa replicabilidade. Há, porém, outras abordagens para a identificação de variáveis latentes que poderiam ser aplicadas ao mesmo conjunto de dados para a comparação dos resultados com os obtidos nessa análise. Alguns estudos sugerem, por exemplo, que, para itens de escala Likert, a análise fatorial seja baseada em correlações específicas para esse tipo de escala, como as correlações policóricas (Flora, Finkel y Foshee, 2003). Ademais, vale ressaltar que estudos futuros com os questionários contextuais do SAEB também podem verificar a validade dos seus construtos por meio de análise dos subconjuntos de itens agrupados em função de hipóteses teóricas.

A estrutura fatorial é uma ferramenta dinâmica de estudo. Deve-se acompanhar o comportamento dessa estrutura em cada ciclo do SAEB, porque a qualidade psicométrica das questões pode ser verificada, os problemas identificados e as questões reformuladas, acrescentadas ou excluídas. A análise da formulação das questões mostra que o problema mais freqüente são textos longos e explicações desnecessárias, que tornam o questionário muito cansativo. Questões com concentração de respostas em uma só alternativa também são freqüentes. Por contribuírem com pouca informação, sua manutenção nos questionários deve ser reconsiderada. Quanto menor o número de questões, menos cansativo é o questionário e a validade da medida de construtos importantes tende a ser melhor.

Quase todos os construtos planejados no quadro de referência do SAEB, ou suas especificações, tiveram fatores equivalentes na estrutura fatorial, ainda que com um número menor de questões ou com uma distribuição um pouco diferente. Apenas um dos construtos teóricos (“Atividades

extracurriculares”) não foi representado na estrutura fatorial. Entre as especificações de “Condições de trabalho do diretor e da equipe”, somente “Carências da escola” foi medida, mas as questões que o formam estiveram distribuídas em vários fatores. A especificação “Direção: apoio e interferências” teve todas as questões excluídas.

O quadro de referência preparado para elaboração dos questionários contextuais do SAEB 2001 é uma evolução muito positiva no aparato técnico do sistema de avaliação (Instituto Nacional de Estudos e Pesquisas Educacionais, 2001). O levantamento bibliográfico e o uso intensivo do quadro de referência neste estudo permitiram a elaboração de algumas sugestões de aprimoramento dessa estrutura.

Primeiramente, o construto “Trabalho colaborativo” tem três especificações definidas, incluindo “Instâncias de gestão da escola”. As questões indicadas para essa especificação medem um construto muito importante, relacionado com a participação da equipe escolar e dos pais na gestão da escola. Na literatura, essa é uma característica de escolas onde a liderança é compartilhada, com abordagem participativa. Essa especificação poderia estar inserida no construto “Liderança do diretor”, com o nome de “Abordagem participativa da liderança”, por exemplo.

Infelizmente, manteve-se na estrutura fatorial apenas uma das onze questões elaboradas para medir o construto “Instâncias de gestão da escola”. Há possíveis explicações para a exclusão de algumas dessas questões, como textos longos e questões condicionadas a respostas anteriores, tornando o bloco mais complexo e cansativo. Dada a relevância desse construto, sugere-se que a estrutura das questões seja repensada para que se possa obter uma medida válida desse fator.

Outra especificação do construto “Trabalho colaborativo” é a “Estabilidade ou rotatividade do corpo docente”, cujo conteúdo não é representativo do construto, apesar de se poder considerar uma relação teórica entre eles. Na literatura pesquisada, o vínculo contratual dos professores é mencionado, mas não está claramente inserido em nenhum fator mais abrangente. Sugere-se que, no quadro de referência, essa especificação constitua um construto independente ou, alternativamente, seja inserido em “Condições de trabalho do diretor e da equipe”.

Há ainda uma duplicidade com relação ao construto teórico sobre “recursos pedagógicos”. Esse construto é colocado tanto no nível mais abrangente do quadro de referência, como em uma especificação dentro de “Situação das instalações e equipamentos”. Sugere-se que essa dupla referência seja eliminada.

Por fim, é relevante ressaltar a importância de se ter boas medidas dos fatores que estão sendo estudados. Ainda que os métodos de análise dos dados sejam sofisticados, se as informações são coletadas sem precisão e qualidade e as medidas têm baixa validade, os resultados certamente serão distorcidos. Espera-se que esse trabalho tenha contribuído para o aprimoramento das medidas dos fatores escolares captados pelo SAEB, proporcionando ferramentas para estudos que nos aproximem mais da garantia ao direito a uma educação com qualidade e equidade no Brasil.

## REFERENCIAS BIBLIOGRÁFICAS

Administración Nacional de Educación Pública-ANEP. (1999). *Estudio de los factores institucionales y pedagógicos que inciden en los aprendizajes en escuelas primarias de contextos sociales desfavorecidos en el Uruguay*. Montevideo: ANEP/UMRE, 1999.

- Instituto Nacional de Estudos e Pesquisas Educacionais – INEP (2001). *SAEB 2001 - Novas perspectivas*. Brasília: INEP/MEC.
- Albernaz, A., Ferreira, F.H.G. y Franco, C. (2002). **Qualidade e equidade na educação fundamental brasileira**. Texto para discussão nº 45. Rio de Janeiro: Departamento de Economia - PUC – RIO.
- Barbosa, M.E.F. (2000). *Redução da dimensão da base de dados do SAEB-99*. Relatório técnico – INEP.
- Clark, L.A. y Watson, D. (1995). Constructing validity: basic issues in objective scale development. *Psychological Assessment*, 7(3), pp.309-319.
- Flora, D.B., Finkel, E.J. y Foshee, V.A. (2003). Higher order factor structure of a self-control test: evidence from confirmatory factor analysis with polychoric correlations. *Educational and Psychological Measurement*, 63, pp.112-127.
- Floyd, F.J. y Widaman, K.F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, 7(3), pp.286-299.
- Grice, J. (2001). Computing and evaluation factor scores. *Psychological Methods*, 6, pp.430-450.
- Gunther, H. (1999). Como elaborar um questionário. En: Pasquali, L. (Ed.). *Instrumentos Psicológicos: manual prático de elaboração*. Brasília: LabPAM/IBAPP.
- Horn, J.L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30, pp.179-185.
- Johnson, R.E. y Wichern, D. (1998). *Applied multivariate statistical analysis*. 4ªed. New Jersey: Prentice Hall.
- Kim, J. y Mueller, C.W. (1987). *Factor analysis: statistical methods and practical issues*. London: Sage.
- Kim, J. y Mueller, C.W. (1987). *Introduction to factor analysis: what it is and how to do it*. London: Sage.
- Laros, J.A. (no prelo). O uso da análise fatorial: algumas diretrizes para pesquisadores. In: Pasquali, L. (Ed.). *Análise fatorial para pesquisadores*. Rio de Janeiro: Vozes.
- Pasquali, L. (1999). Testes referentes a construto: teoria e modelo de construção. In: Pasquali, L. (Ed.). *Instrumentos Psicológicos: Manual prático de elaboração*. Brasília: LabPAM/IBAPP.
- Rosales, C.A.B. (2000). *Escolaridade, qualidade e equidade: convivência frustrada*. Anais do Encontro "Calidad y equidad de la educación" de diretores e altos funcionários dos Ministérios de Educação dos países ibero-americanos. Guatemala: Organización dos Estados Americanos – OEI.
- Smith, G.T. y McCarthy, D.M. (1995). Methodological considerations in the refinement of clinical assessment instruments. *Psychological Assessment*, 7, pp.300-308.
- Tabachnick, B.G. y Fidell, L.S. (1966). *Using multivariate statistics*. 3ªed. New York: Harper Collins.
- Teddlie, C. y Reynolds, D. (2000). *The international handbook os school effectiveness research*. London: Falmer Press.
- Zoski, K.W. y Jurs, S. (1996). An objective counterpart to the visual scree test for factor analysis: the standard error scree. *Educational and Psychological Measurement*, 56, pp.443-451.
- Zwick, W.R. y Velicer, W.F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, 99, pp.432-442.