UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE
DE RIBEIRÃO PRETO
DEPARTAMENTO DE ADMINISTRAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO DAS
ORGANIZAÇÕES

CRISTIANO PEREIRA ARRUDA

**Análise do Sistema de Avaliação das Escolas Municipais (SAEM)
aplicado na rede municipal de Educação Básica de Uberaba-MG:
um estudo de caso.**

Orientadora: Profa. Dra. Adriana Backx Noronha Viana

Ribeirão Preto
2011

Prof. Dr. João Grandino Rodas
Reitor da Universidade de São Paulo

Prof. Dr. Sigismundo Bialoskorski Neto
Diretor da Faculdade de Economia, Administração e Contabilidade de
Ribeirão Preto

Prof. Dr. Marcos Fava Neves
Chefe do Departamento de Administração

CRISTIANO PEREIRA ARRUDA

**Análise do Sistema de Avaliação das Escolas Municipais (SAEM)
aplicado na rede municipal de Educação Básica de Uberaba-MG:
um estudo de caso.**

Dissertação apresentada á Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo como requisito para obtenção do título de Mestre em Ciências obtido no Programa de Pós-Graduação em Administração de Organizações.

Orientadora: Profa. Dra. Adriana Backx Noronha Viana

VERSÃO CORRIGIDA. A ORIGINAL ENCONTRA-SE NO SERVIÇO DE
PÓS-GRADUAÇÃO DA FEARP-USP.

RIBEIRÃO PRETO
2011

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Arruda, Cristiano Pereira

Análise do Sistema de Avaliação das Escolas Municipais (SAEM) aplicado na rede municipal de Educação Básica de Uberaba-MG: um estudo de caso. Ribeirão Preto, 2011.

208 p. : il. ; 30 cm

Dissertação de Mestrado, apresentada à Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo.

Orientador: Viana, Adriana Backx Noronha .

1. Avaliação da educação. 2. SAEM. 3. Educação Básica. 4. Políticas públicas.

FOLHA DE APROVAÇÃO

Nome: ARRUDA, Cristiano Pereira.

Título: Análise do Sistema de Avaliação das Escolas Municipais (SAEM) aplicado na rede municipal de Educação Básica de Uberaba-MG: um estudo de caso.

Dissertação apresentada á Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo como requisito para obtenção do título de Mestre em Ciências obtido no Programa de Pós-Graduação em Administração de Organizações.

Aprovada em: ____ / ____ / ____

BANCA EXAMINADORA

Prof. Dr. _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Prof. Dr. _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Prof. Dr. _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Ao meu pai, Edison (in memoriam),
companheiro das madrugadas e das
estradas.

À minha esposa, Patrícia, sem a qual não
estaria escrevendo essas linhas, com
amor.

Aos meus filhos, Caio e Rafael, pacientes
com a ausência momentânea do pai.

AGRADECIMENTOS

A Deus, pelo dom da Vida e pelas linhas tortas que têm escrito uma boa história de vida.

A Universidade de São Paulo, em especial à Faculdade de Economia, Administração e Economia de Ribeirão Preto, pela acolhida e pela oportunidade de realização do curso de Mestrado e do desenvolvimento desse trabalho.

Aos professores do Programa de Pós-Graduação em Administração das Organizações, pela significativa contribuição com a minha formação intelectual e científica, além do exemplo profissional.

A minha orientadora, professora doutora Adriana Back Noronha Viana, por ter aceitado o projeto desta pesquisa e pela prática efetiva do que se entende por orientação.

A Prefeitura Municipal de Uberaba, em especial à equipe gestora da Secretaria Municipal de Educação, liderada pelo secretário professor José Wandir de Oliveira, pela permissão em utilizar os resultados da Rede Municipal de Ensino no SAEM neste trabalho e o pronto atendimento às solicitações enviadas.

Aos diretores, coordenadores, professores, alunos e pais de alunos das escolas da Rede Municipal de Ensino, pela atenção em participar das entrevistas realizadas para coleta de dados deste trabalho.

Ao Colégio Cenecista Dr. José Ferreira, pela oportunidade de trabalho, de formação intelectual e moral e de poder realizar o presente trabalho. Em especial, agradeço ao professor Danival Roberto Alves pelos ensinamentos diários e o exemplo de Vida dedicada à educação e ao próximo.

A minha família, pelo carinho e pela confiança, nunca permitindo que eu abrisse a porta do fracasso. A minha irmã Jussara, amiga e companheira desde os meus primeiros rabiscos. A minha prima Cristina, pela oferta de “pouso” durante minha estadia em Ribeirão Preto e pelo incansável incentivo para a realização deste trabalho. Em especial, agradeço à minha mãe, Maria Célia, por tudo o que ela já fez para que eu chegasse até aqui.

Obrigado!

Ai daqueles que pararem com sua capacidade de sonhar, de invejar sua coragem de anunciar e denunciar. Ai daqueles que, em lugar de visitar de vez em quando o amanhã pelo profundo engajamento com o hoje, com o aqui e o agora, se atrelarem a um passado de exploração e de rotina.

Paulo Freire

RESUMO

ARRUDA, Cristiano Pereira. **Análise do Sistema de Avaliação das Escolas Municipais (SAEM) aplicado na rede municipal de Educação Básica de Uberaba-MG**: um estudo de caso. 2011. 208 f. Dissertação(Mestrado) – Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto-SP, 2011.

Este trabalho teve como objeto de estudo o SAEM (Sistema de Avaliação das Escolas Municipais), desenvolvido pelo Colégio Cenecista Dr. José Ferreira – onde o mestrando trabalha como professor e coordenador de avaliações – e aplicado na rede municipal de ensino de Uberaba-MG, desde 2006, como uma ferramenta para mensuração do desempenho acadêmico dos alunos do 1º ao 9º ano do Ensino Fundamental no referido município. A análise explorou as quatro faces definidas para o SAEM: Teórica, Desenvolvimento, Aplicação e Expansão. Na face Teórica foi apresentada uma análise da adequação do SAEM aos padrões de qualidade da avaliação do *Joint Committee Standards for Educational Evaluation* (JCSEE) e uma análise comparativa entre as características da Prova Brasil/SAEB e do SAEM. A face Desenvolvimento discorreu sobre os passos para a elaboração de uma avaliação do SAEM, sendo que foi comprovada a validade de conteúdo da avaliação. Na face Aplicação foram descritas as etapas de aplicação de uma avaliação aos alunos da rede municipal de ensino de Uberaba-MG. Os relatórios produzidos a partir dos resultados de uma avaliação e os atuais usos desses resultados também são apresentados. Ainda dentro da face Aplicação, foi possível comprovar a adequação do SAEM aos conceitos de confiabilidade e de validade de construto e de critério de um sistema de avaliação. Finalmente, a face Expansão identificou possíveis pontos de melhoria na metodologia atual do SAEM e novos usos para os resultados de uma avaliação. Assim, foi possível atingir o objetivo principal deste trabalho, que consistiu na análise das características do SAEM, identificando suas vantagens e limitações como ferramenta de aferição do desempenho acadêmico de alunos de uma rede de ensino.

Palavras-chave: Avaliação da educação. SAEM. Educação Básica. Políticas públicas.

ABSTRACT

ARRUDA, Cristiano Pereira. **Analysis of SAEM – Evaluation System for Municipal Schools - Applied from Elementary to Middle Municipal Schools in Uberaba-MG**: a case study. 2011. 208 f. Dissertation - Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto-SP, 2011.

The objective of this study was to analyse SAEM - Evaluation System for Municipal Schools – developed by Colégio Cenecista Dr. José Ferreira - where the Masters student works as a coordinator of evaluation and a teacher – and applied to the municipal school chain in Uberaba-MG, since 2006, as a tool to measure the students' academic performance from elementary to middle school levels, in the mentioned city. The analysis comprised of the four aspects, defined for SAEM: Theory, Development, Application and Expansion. In the Theory aspect, an analysis of SAEM adequacy to the evaluation quality standards from the *Joint Committee Standards for Educational Evaluation* (JCSEE) and a contrasting analysis between the characteristics present in Prova Brasil/SAEB and SAEM. The Development aspect was about the steps to develop a SAEM evaluation, in which the premise of content validity for the evaluation was proved. In the Application aspect, the steps followed to apply an evaluation to students in the municipal school system in Uberaba-MG were described. Besides these data, reports produced as the result of an evaluation and the present uses of these results were submitted. Still in the Application aspect, it was possible to confirm that SAEM meets the standards for reliability and construct and criteria validity of an evaluation system. Finally, the Expansion aspect identified some possibilities for improvement in SAEM's present methodology and new usages for the evaluation results. It was then, possible to achieve the main objective of this study, which was to analyze SAEM's characteristics, identifying its advantages and limitations as a tool to measure the students' academic performance of a school system.

Keywords: Evaluation of education. SAEM. Basic Education. Public policies.

LISTA DE FIGURAS

Figura 01 - Histórico do desempenho do Brasil no PISA nas três áreas avaliadas	25
Figura 02 - Esquema representativo do trabalho	33
Figura 03 - Organização do Referencial Teórico	35
Figura 04 - Organização da Metodologia da pesquisa	77
Figura 05 - Gráfico de abstenção nas avaliações	79
Figura 06 - Visão global da pesquisa	84
Figura 07 - Organização dos Resultados	95
Figura 08 - Etapas de desenvolvimento de uma avaliação do SAEM	112
Figura 09 - Organograma (parte) da Secretaria Municipal de Educação de Uberaba	124
Figura 10 - Gráfico de dispersão – 4ª série do Ensino Fundamental	147
Figura 11 - Gráfico de dispersão – 8ª série do Ensino Fundamental	147

LISTA DE QUADROS

Quadro 01 - Publicações dos <i>Standards</i> do JCSEE	55
Quadro 02 - Os <i>Standards</i> do JCSEE por categoria	56
Quadro 03 - Objetivos do SAEB 2011	68
Quadro 04 - Disciplinas avaliadas em cada um dos ciclos do SAEB	70
Quadro 05 - Sistema Nacional de Avaliação da Educação Básica	72
Quadro 06 - Protocolo dos Roteiros de entrevistas	81
Quadro 07 - Resumo do delineamento da pesquisa	92
Quadro 08 - Dimensões de análise dos dados qualitativos	94
Quadro 09 - Análise da adequação do SAEM aos <i>Standards</i> da JCSEE.....	96
Quadro 10 - Distribuição dos resultados de adequação.....	101
Quadro 11 - Diferenças entre a Prova Brasil e o SAEB.....	102
Quadro 12 - Distribuição do número de questões por classe/série e disciplina	103
Quadro 13 - Comparação entre o SAEM e a Prova Brasil	104
Quadro 14 - Lista de temas do SAEM	114
Quadro 15 - Análise da validade de conteúdo do SAEM	121
Quadro 16 - Propostas de adequação do SAEM aos <i>Standards</i>	149
Quadro 17 - Comparação entre os resultados esperados e obtidos	162
Quadro 18 - Análise das informações coletados nas entrevistas.....	165

LISTA DE TABELAS

Tabela 01 - Resultados do cálculo do coeficiente de confiabilidade	135
Tabela 02 - Coeficientes de correlação entre os desempenhos por disciplina da Classe de 10 anos na 2ª avaliação de 2007.....	140
Tabela 03 - Coeficientes de correlação entre os desempenhos por disciplina da 8ª série na 2ª avaliação de 2007.....	140
Tabela 04 - Coeficientes de correlação entre os desempenhos por disciplina da Classe de 10 anos na 1ª avaliação de 2010.....	140
Tabela 05 - Coeficientes de correlação entre os desempenhos por disciplina da 8ª série na 1ª avaliação de 2010.....	141
Tabela 06 - Desempenho das escolas na Prova Brasil em 2007. 4ª série do Ensino Fundamental	143
Tabela 07 - Desempenho das escolas na Prova Brasil em 2007. 8ª série do Ensino Fundamental	144
Tabela 08 - Desempenho médio das escolas no SAEM em 2007. 4ª série do Ensino Fundamental	144
Tabela 09 - Desempenho médio das escolas no SAEM em 2007. 8ª série do Ensino Fundamental	145
Tabela 10 - Resultados do cálculo da correlação de Pearson para a 4ª série do Ensino Fundamental	146
Tabela 11 - Resultados do cálculo da correlação de Pearson para a 8ª série do Ensino Fundamental	146
Tabela 12 - Faixas de desempenho do SAEM 2011	154
Tabela 13 - Projeção de SAEM para 2011 – séries iniciais	158
Tabela 14 - Projeção de SAEM para 2011 – séries finais	158
Tabela 15 - Metas intermediárias do SAEM	158

SUMÁRIO

1.	INTRODUÇÃO	21
1.1.	PROBLEMA DE PESQUISA	31
1.2.	JUSTIFICATIVA	31
1.3.	OBJETIVO GERAL	32
1.3.1.	OBJETIVOS ESPECÍFICOS.....	33
1.4.	ORGANIZAÇÃO DO TRABALHO	33
2.	REFERENCIAL TEÓRICO	34
2.1.	O QUE É AVALIAR?	35
2.2.	QUALIDADE DA EDUCAÇÃO	40
2.3.	QUALIDADE DA AVALIAÇÃO	53
2.3.1.	OS PADRÕES DE QUALIDADE DA AVALIAÇÃO DO JCSEE	54
2.3.2.	VALIDADE E CONFIABILIDADE DE UMA AVALIAÇÃO	59
2.4.	O SISTEMA DE AVALIAÇÃO DA EDUCAÇÃO BÁSICA	64
2.4.1.	A PROVA BRASIL	66
2.4.2.	OBJETIVOS DO SAEB	66
2.4.3.	ABRANGÊNCIA: SÉRIES/ANOS, DISCIPLINAS E REDES	69
2.4.4.	MATRIZ DE REFERÊNCIA	70
2.4.5.	INSTRUMENTOS CONTEXTUAIS	71
2.4.6.	INSTRUMENTOS COGNITIVOS	73
2.4.7.	RESULTADOS DO SAEB	75
3.	METODOLOGIA.....	77
3.1.	CONTEXTOS METODOLÓGICOS DA PESQUISA	77
3.1.1.	POPULAÇÃO / AMOSTRA	79
3.1.2.	COLETA DE DADOS	80
3.1.3.	VARIÁVEIS	82
3.1.4.	PROPOSIÇÕES e HIPÓTESES	83
3.2.	DELINEAMENTO DA PESQUISA	84
3.2.1.	FACE TEÓRICA.....	86
3.2.2.	FACE DESENVOLVIMENTO	87
3.2.3.	FACE APLICAÇÃO.....	88
3.2.4.	FACE EXPANSÃO	91
3.3.	VISÃO GERAL E ANÁLISE DOS RESULTADOS	91
4.	RESULTADOS	95
4.1.	FACE TEÓRICA.....	95
4.2.	FACE DESENVOLVIMENTO	111
4.2.1.	DEFINIÇÕES INICIAIS.....	112

4.2.2.	CONTRATO	113
4.2.3.	ESCOLHA DO TEMA.....	114
4.2.4.	DEFINIÇÃO DOS CONTEÚDOS	115
4.2.5.	ELABORAÇÃO DOS DESCRITORES	116
4.2.6.	ELABORAÇÃO DAS QUESTÕES	117
4.2.7.	MONTAGEM DA AVALIAÇÃO	118
4.2.8.	ORGANIZAÇÃO E DISTRIBUIÇÃO DAS AVALIAÇÕES	119
4.2.9.	VALIDADE DE CONTEÚDO	120
4.3.	FACE APLICAÇÃO	122
4.3.1.	APLICAÇÃO DAS PROVAS	122
4.3.2.	CORREÇÃO E GERAÇÃO DOS RELATÓRIOS	127
4.3.3.	USO DOS RESULTADOS.....	129
4.3.4.	VALIDADE E CONFIABILIDADE	133
4.4.	FACE EXPANSÃO	148
4.4.1.	PROPOSTAS DE MUDANÇAS	149
4.4.2.	NOVOS USOS DOS RESULTADOS DO SAEM	156
4.5.	VISÃO GERAL DOS RESULTADOS	162
5.	CONCLUSÕES	168
6.	REFERÊNCIAS.....	173
	APÊNDICE A – AUTORIZAÇÕES DE USO DO NOME E DOS DADOS	182
	APÊNDICE B – ROTEIROS PARA AS ENTREVISTAS.....	184
	ANEXO A – RELATÓRIO DO SAEB	188
	ANEXO B – RELATÓRIOS DO SAEM	189
	ANEXO C – RELATÓRIOS REMODELADOS/NOVOS DO SAEM	197
	ANEXO D – EXEMPLO DE QUESTÕES DE UMA AVALIAÇÃO TEMÁTICA	203

1. INTRODUÇÃO

Dentro do universo da educação escolar, a avaliação tem um papel relevante, já que pode mensurar, a partir do desempenho dos alunos e de outros indicadores, a qualidade da educação ofertada pelos sistemas de ensino vigentes. Na medida em que aferem o desempenho acadêmico dos alunos, as ferramentas de avaliação assumem também o papel de indicar pontos críticos dentro da rede de ensino os quais devem ser mais bem trabalhados além de medir também a eficiência das políticas públicas intervencionistas adotadas. Segundo Fernandes (2009), as avaliações são um guia de orientação para as mudanças que devem ser implementadas nas redes de ensino, não somente no sentido de detectar falhas, mas também como uma referência do que se deve transmitir aos alunos e de como deve ser avaliada a produção do conhecimento. Perrenoud (1999) defende que a avaliação educacional deve atender a duas lógicas: a da seleção e a da aprendizagem, uma vez que, não necessariamente, essas lógicas são antagônicas.

O Governo Federal, através do Ministério da Educação, implantou, a partir do final década de 80, a “cultura das avaliações” (VIANNA, 1995) dos sistemas educacionais existentes no Brasil (FREITAS, 2007; BONAMINO, 2002). O que se tem visto é a criação de, cada vez mais, “siglas” e sistemas de avaliação em todos os níveis da educação, da Educação Básica ao Ensino Superior, estendendo-se, além do Governo Federal, para alguns estados e municípios. Pode-se considerar que o processo teve início com a criação do Sistema de Avaliação da Educação Básica (SAEB), em 1991, com o objetivo de avaliar, por amostragem, os alunos concluintes da 4ª e da 8ª séries (5º e 9º ano) do Ensino Fundamental e da 3ª série do Ensino Médio (PESTANA, 1998). O Índice de Desenvolvimento da Educação Básica (IDEB), instituído em 2007 (FERNANDES, 2007) o qual leva em conta o índice de aprovação e de evasão das escolas, além do desempenho de seus alunos na Prova Brasil/SAEB – é a mais recente ação do Governo Federal que contribui para a implantação da “cultura das avaliações”.

A criação de indicadores como o IDEB é uma saída para que se possa acompanhar a evolução da qualidade da educação, mesmo sabendo que a mesma

não deve ser mensurada a partir de alguns poucos parâmetros, como a avaliação de desempenho de alunos. Alguns dos principais documentos que discutem a questão da qualidade da educação: o documento da Unesco “Educação de Qualidade para Todos” (EDUCAÇÃO, 2007); o documento final da CONAE 2010 (Conferência Nacional de Educação) (BRASIL, 2010); e o documento do INEP “Indicadores da qualidade na educação” (INDICADORES, 2007) ressaltam a importância da avaliação como uma das dimensões da qualidade da educação.

O ainda pequeno período de pesquisas e desenvolvimento das avaliações de larga escala no Brasil – a partir dos anos 90 – traz vários desafios. Ainda no final da década de 80, Vianna (1989) já previa alguns desses desafios: a falta de pesquisadores com formação específica na área; a falta de uma teoria perfeitamente estruturada representativa do consenso dos educadores e a deficiência na interpretação e comunicação dos resultados.

Apesar de o fenômeno ser recente nacionalmente, pode-se considerar que, no período de 1930 a 1988, já havia iniciativas de medidas avaliativas para aferição da qualidade educacional, haja vista a própria criação do INEP (Instituto Nacional de Estudos Pedagógicos¹) no final da década de 1930 (FREITAS, 2007; INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA, 2009). Ainda segundo a autora, no período do Estado Novo, compreendido entre 1937 e 1945, houve um investimento na “quantificação” da educação para fins de seu controle.

Na década de 40 e no início da década de 50 do século passado, o processo foi dominado por pesquisas de natureza psicopedagógica. À época, Lourenço filho e Hildebrand² (1945 apud FREITAS, 2007, p. 10) defenderam a idéia de mudança da avaliação “grosseira” para a avaliação como “medida”, além de também defenderem o uso dos chamados “testes pedagógicos”, que, segundo eles, permitiriam avaliar o trabalho do aluno, do professor e das escolas.

¹ A partir de 1972, o INEP é passou a ser chamado de Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira

² LOURENÇO FILHO, Manuel B.; HILDEBRANDO, Armando. São necessários os exames escolares?. **Revista Brasileira de Estudos Pedagógicos**, v. 4, n. 10, p. 51-54, abr. 1945.

A década de 50 se destaca pela entrada de Anísio Teixeira no INEP e pelas suas idéias de reforma da educação, de criação do CBPE (Centro Brasileiro de Pesquisas Educacionais) e de centros regionais (FREITAS, 2007). Os anos 60 tiveram, no Brasil, um direcionamento para a produção científica de caráter sociológico e também para os estudos ligados aos processos seletivos de acesso ao Ensino Superior (VIANNA, 1995, GATTI, 2002).

A concomitância entre a evolução da avaliação educacional internacionalmente e o momento político/histórico brasileiro da década de 80 – reconquista da democracia, nova Constituição etc – culminou na criação do Sistema de Avaliação da Educação Básica (SAEB) (BONAMINO, 2002). A década de 80 é considerada como o período do surgimento das avaliações em larga escala no país. Nesse período, além das iniciativas do Governo Federal como o SAEB, o Exame Nacional do Ensino Médio (ENEM) e o Exame Nacional de Cursos, vários governos estaduais implantaram sistemas de avaliação em suas redes de ensino (VIANNA, 1995; COELHO, 2008).

Internacionalmente, o fenômeno tem início no século XVIII, com os exames de seleção de funcionários públicos na Prússia, atual Alemanha (FERNANDES, 2009). Entretanto, o *boom* das avaliações em larga escala ocorreu na década de 60, nos Estados Unidos, conforme cita Vianna (1995, p. 7), devido ao “trauma provocado pela constatação da deficiência tecnológica associada à carência educacional, no mundo ocidental, especialmente nos Estados Unidos, com o lançamento do Sputnik, no dia 04 de outubro de 1957”. A partir de então, nomes como Michael Scriven, Robert Stake, Daniel Stufflebeam, David Hamilton e Malcom Parlett passam a dar os fundamentos da nova avaliação educacional, afastando-a do modelo positivista vigente à época (VIANNA, 1995; SOUZA, 2005).

Em paralelo aos avanços locais no final da década de 80, conforme já foi citado, o país inicia também a sua participação em projetos internacionais de avaliação educacional. Vianna (1995) e Gatti (2002) relatam que a primeira iniciativa se deu com crianças de 13 anos das cidades de Fortaleza e de São Paulo, em Matemática e em Ciências. O projeto foi conduzido pela *Internacional Assessment of Educational Progress* (IAEP) e contou com a participação de 20 países. Broke (2006) acrescenta que esse projeto foi o segundo promovido pela IAEP e que incluía

crianças de 9 e de 13 anos. Bonamino (2002) destaca também a participação, a partir de 1998, no TIMSS (*Third International Mathematics and Science Study*) promovido pela *Internacional Association for the Evaluation of Educational Achievement* (IEA) e a avaliação da UNESCO, intitulada “*Oficina Regional para La Educación em América Latina u el Caribe*” (OREALC).

A partir do ano 2000, o principal exame internacional do qual o país tem participado é o Programa Internacional de Avaliação de Alunos (PISA), da Organização para a Cooperação e Desenvolvimento Econômico (OCDE). Segundo Fernandes (2009), as avaliações são aplicadas a cada três anos e afere os conhecimentos dos alunos com 15 anos de idade em Leitura, Matemática e Ciências, sendo que a cada avaliação é dada ênfase a uma das três áreas. Em 2000 foi a Leitura, em 2003, a Matemática e, em 2006, a área privilegiada com dois terços do total de questões, foi a de Ciências. Já em 2009 (OECD, 2009), a avaliação voltou a enfatizar a Leitura. Oliveira e Araújo (2005), ao discutirem a qualidade da educação brasileira, apresentam o desempenho dos alunos brasileiros na avaliação do PISA. Em 2000, o País ficou em último lugar dentre os 32 países participantes, com 56% dos alunos com baixo desempenho em leitura. Completam os autores (p. 15):

Assim, 56% dos estudantes brasileiros que participaram do PISA-2000 demonstraram que, ao final da etapa elementar de escolarização, aquela cujo objetivo fundamental é oferecer as bases para a vida em sociedade e para o prosseguimento dos estudos, conseguem, no máximo, compreender textos simples, não tendo se beneficiado suficientemente das oportunidades educacionais e não adquirindo os conhecimentos e habilidades esperadas para essa etapa de escolarização.

A edição do relatório completo dos resultados do ano 2000 (OECD, 2001) apresenta o Brasil também em último lugar em Matemática e em Ciências, dentre os mesmos 32 países participantes.

Na avaliação de 2003 do PISA, o país ficou na 37ª posição em relação às habilidades com a leitura e na 40ª posição nas provas de Ciências e de Matemática, dentre 41 países participantes (OCDE, 2004; BROKE, 2006). Já na avaliação de 2006, o país atingiu a 52ª posição em Ciências e a 54ª posição em Matemática, dentre 57 países, além da 48ª posição em Leitura, dentre 56 países participantes (INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA, 2008; OCDE, 2007).

Nos resultados de 2009, quando o foco da avaliação voltou a ser a leitura, o Brasil teve uma melhora de pontuação. Em leitura, obteve a 53ª posição dentre 65 países participantes, posição também obtida em Ciências, enquanto, em Matemática, o País não passou da 57ª posição (OCDE, 2010). Vale ressaltar que a amostra selecionada para a avaliação considera alunos de 15 anos que estejam cursando pelo menos o 8º ano do Ensino Fundamental de todas as regiões do País, matriculados tanto na rede pública ou privada. A Figura 1 mostra o desempenho do país em todas as edições do PISA, nas três áreas avaliadas. Analisando os dados, pode-se verificar que, desde 2000, houve crescimento do desempenho nas três áreas avaliadas, com destaque para a Matemática. Entretanto, o país ainda se encontra abaixo da média dos demais países participantes.

Figura 1: Histórico do desempenho do Brasil no PISA nas três áreas avaliadas
Fonte dos dados: OCDE, 2001; OCDE, 2004; OCDE, 2007; OCDE, 2010

As avaliações internacionais como o PISA também têm suas desvantagens. Ferreiro (2005) inicia suas críticas questionando a legitimidade dos órgãos responsáveis, Banco Mundial e OCDE, não criados para esse fim, em promover tais exames. A grande repercussão dos resultados desses exames traz consequências ruins, como a redução do currículo e o foco no desenvolvimento de habilidades e não mais nos conteúdos, segundo a autora. Fernandes (2009), ao abordar a participação dos alunos portugueses nos exames internacionais, afirma que, apesar

dos quase 30 anos de participação nesses exames, desconhecem-se consequências significativas. O autor relaciona oito problemas relativos ao modelo de avaliação adotado nos exames internacionais: 1) diversidade de currículos e contextos entre os países; 2) diferenças estruturais e organizacionais, dificultando a equivalência de amostras; 3) dificuldade em garantir a mesma interpretação com a tradução para várias línguas; 4) resultado apresentado somente pela nota final; 5) não garantia da mesma relevância curricular dada pelo país; 6) problemas de inconsistência ao comparar os resultados com outros exames, 7) divergências nos procedimentos de aplicação nos diferentes países e 8) ausência de estudos longitudinais.

A evolução das pesquisas na área da avaliação educacional apresenta um panorama no qual se nota o seu desmembramento em distintos setores, conforme o objeto em estudo. A avaliação da aprendizagem procura avaliar o processo de ensino e aprendizagem dos alunos; a avaliação de currículos e programas e as avaliações em larga escala buscam avaliar as redes de ensino. As avaliações que levam em conta fatores como a estrutura física das escolas, as características sociais dos alunos e a qualificação dos professores são chamadas de avaliações sistêmicas.

Como a prática das avaliações de redes está cada vez mais difundida, torna-se imprescindível a criação de ferramentas que permitam aos gestores escolares proceder a análise dos resultados obtidos por sua escola, a fim de que identifiquem corretamente os fatores que os levaram àquele resultado (VIANNA, 2005; SOARES, 2008).

A constituição brasileira de 1988 (BRASIL, 1988), no artigo 30, transferiu para os municípios a responsabilidade de oferta de Educação Infantil e do Ensino Fundamental “com cooperação técnica e financeira da União e do Estado” (BRASIL, 1988). Já a Lei de Diretrizes e Bases (LDB) de 1996, em seu artigo 9º, determina a obrigatoriedade da avaliação dos diferentes níveis do sistema educacional brasileiro pela União (BRASIL, 1996). Entretanto, Freitas et al. (2009, p. 48) defendem que as avaliações sistêmicas seriam mais eficazes se fossem planejadas e conduzidas no nível dos Municípios. “É preciso que o dado seja ‘reconhecido’ como ‘pertencendo’ à escola” completam os autores.

A construção de um sistema de avaliação possível de ser aplicado com facilidade em qualquer escola ou rede educacional é um grande passo rumo à independência dessas instituições em relação aos processos avaliativos adotados e à massificação das avaliações sistêmicas, o que justifica fortemente o presente estudo.

Em consonância com esses mecanismos de avaliação de desempenho, o Colégio Cenecista Dr. José Ferreira (CCDJF) – colégio onde o mestrando trabalha como professor e coordenador de avaliações – desenvolveu o Sistema de Avaliação das Escolas Municipais (SAEM), que tem por objetivo, a partir de avaliações periódicas em todos os alunos de todas as classes/séries do Ensino Fundamental da rede pública municipal, aferir o desempenho acadêmico dos alunos da Rede Municipal de Ensino. Desde 2006, ininterruptamente, o SAEM tem sido aplicado à rede municipal de Uberaba, Minas Gerais. Desde então, esse sistema de avaliação tem possibilitado a adoção de estratégias coletivas ou individuais, visando a atingir as metas estabelecidas no Contrato de Gestão que cada escola firmou com o governo do município. A cada avaliação, cerca de 17 mil alunos das classes de 6 anos (1º ano) à 8ª série (9º ano) do Ensino Fundamental são avaliados nas disciplinas de Língua Portuguesa, Matemática, Ciências, História e Geografia.

Tanto o Colégio Cenecista Dr. José Ferreira, através de seu diretor, professor Danival Roberto Alves, quanto a Secretaria Municipal de Educação e Cultura de Uberaba, através de seu secretário municipal, professor José Vandir de Oliveira, autorizaram o uso do nome do SAEM e dos resultados das avaliações já aplicada, no desenvolvimento do presente trabalho de pesquisa, conforme pode ser visto no apêndice A.

O acompanhamento da implantação do SAEM, em 2006, e da sua evolução até os dias de hoje pelo mestrando possibilitou um relato histórico das dificuldades encontradas ao longo do tempo. Tão logo a parceria entre a Secretaria e o Colégio foi confirmada, em março de 2006, houve uma reunião entre os gestores das escolas (diretores, vice-diretores e coordenadores) com a equipe do Colégio José Ferreira. A reunião ocorreu no auditório da Escola Municipal Frei Eugênio. Ao apresentar a proposta do SAEM, muitos gestores se posicionaram contra ela.

De modo geral, as resistências estavam concentradas: 1) no temor de que os resultados das escolas se tornariam públicos; 2) no boato de que os resultados definiriam os índices de reajustes salariais dos professores e dos gestores; e 3) na falta de cultura dos alunos, principalmente os mais novos, para o preenchimento do cartão de respostas (usado a partir das classes de 8 anos). Chegaram, inclusive, a falar até que as crianças não sabiam usar caneta.

O professor Danival, diretor do Colégio, procurou conter os ânimos e explicou que o propósito principal do SAEM era o de conhecer as deficiências de conteúdos de cada um dos alunos e, conseqüentemente, das turmas e das escolas. A partir dessas informações, seriam traçadas estratégias de recuperação desses alunos e de orientação dos professores e dos gestores.

Outro problema encontrado foi a não coincidência dos conteúdos trabalhados pelas disciplinas nas escolas e mesmo em turmas diferentes de uma mesma escola. Chegou-se ao extremo de se encontrar uma escola em que o professor de uma turma trabalhava História do Brasil, e o de outra turma, História Geral. Ao final dessa reunião, ficou definido que, para aquele primeiro ano do SAEM, seriam cobrados os conteúdos de maior incidência nos programas das disciplinas que estavam sendo trabalhados nas escolas e que, para o próximo ano, seriam usadas as diretrizes curriculares do município.

A primeira reunião para apresentação dos resultados demonstra bem a visão que a maioria tinha, à época, sobre o uso dos resultados. No mesmo auditório das reuniões anteriores, além dos representantes da Secretaria de Educação e dos gestores das escolas, havia jornalistas dos principais jornais e rádios da cidade, além de duas emissoras de televisão capturando imagens do evento. Conforme já havia sido falado, o foco do SAEM não seria o ranqueamento das escolas. Foram apresentados, portanto, os resultados gerais da rede para cada classe/série avaliada, o que frustrou muito os presentes, ávidos para saber, principalmente, qual era a pior escola do município. Para os representantes da Secretaria foi elaborado um relatório com a relação das escolas e seus desempenhos por classe/série, mas propositalmente ordenado em ordem alfabética e não ordem decrescente de desempenho. Essa postura do Colégio em não dar foco ao ranqueamento foi decisiva

para que os diretores das escolas passassem a confiar na proposta do SAEM, principalmente quanto ao caráter não punitivo da mesma.

Após esses cinco anos de avaliação, a reunião para apresentação dos resultados do SAEM passou a ser uma reunião praticamente técnica, na qual se discutem os resultados obtidos, influências positivas e negativas que tenham ocorrido, além de se acolherem as propostas para a próxima avaliação. O trabalho principal está nas escolas, onde os resultados são analisados e as estratégias de crescimento são traçadas.

Por vários SAEMs, o Colégio enviou os resultados gerais das escolas, em um relatório exclusivo, para a Secretaria Municipal de Educação. Entretanto, ao chegar para a apresentação dos resultados de mais um SAEM, o mestrando se deparou com pessoas da equipe da Secretaria distribuindo cópias do relatório, enviado antecipadamente, já em ordem decrescente de desempenho médio das escolas. Ao questionar a divulgação, informaram que a prática de tirar cópias do relatório para todos já era feita há algum tempo e que os diretores aprovavam a iniciativa.

O quadro atual é bastante interessante, uma vez que as escolas conhecem seus resultados em comparação com as outras e os mesmos têm servido como motivação para o seu crescimento. O relatório geral de desempenho da rede é atualmente enviado a todas as escolas. A postura atual da imprensa é a de divulgar os resultados, quando informados via Secretaria de Educação, sem destaques sensacionalistas. A impressão que se tem é a de que todos os envolvidos – alunos, professores, diretores e Secretaria – tomaram consciência das vantagens do SAEM, uma vez que os resultados são usados para a melhoria da educação municipal, eliminando a visão punitiva da avaliação externa.

Em seu desenvolvimento, o SAEM é elaborado a partir das diretrizes municipais de ensino, documento que orienta as habilidades que devem ser desenvolvidas nos conteúdos ministrados em sala. Inicialmente, cada equipe de professores responsáveis pela elaboração da avaliação em cada classe/série elabora uma matriz de descritores. Essa Matriz de descritores é enviada às escolas da rede municipal com antecedência, para que os professores possam preparar melhor seus alunos. Nessa matriz, cada descritor determina a forma com que serão abordados os conteúdos propostos. A avaliação possui 20 questões discursivas para

as classes de 6 e 7 anos. A partir da classe de 8 anos, os alunos respondem a questões objetivas, com 4 alternativas cada, sendo 20 questões para a classe de 8 anos e 30 questões para todas as demais classes/séries.

O modelo de avaliação proposto pelo SAEM é de uma Avaliação Temática e Interdisciplinar, ou seja, a partir de um tema escolhido, como, por exemplo, “Esporte: unindo povos, disseminando culturas”. As questões são elaboradas com uma abordagem contextualizada que liga o conteúdo de cada questão, de todas as disciplinas, ao tema proposto. A cada nova avaliação, é escolhido um novo tema e questões inéditas são elaboradas pelas equipes de professores do Colégio Cenecista Dr. José Ferreira – há uma equipe para cada série – para que, de acordo com esse tema, sejam avaliados os conteúdos trabalhados durante o semestre letivo, o que provoca a não organização sequencial das questões por disciplinas. Assim, a partir de um gênero textual³ em comum, referente ao tema da avaliação, algumas questões de disciplinas diferentes são apresentadas para resolução. A avaliação contém vários gêneros textuais aos quais as questões se relacionam.

Após a leitura e a correção das avaliações, são gerados relatórios contendo informações direcionadas a seus respectivos públicos-alvo: os gestores (secretaria e direção escolar), os professores e os pais. Há a preocupação de que os dados relativos a uma turma, série ou escola fiquem restritos aos responsáveis por esses grupos. Desse modo, os dados relativos à toda a rede municipal de ensino são exclusivos dos gestores da secretaria de educação.

A aplicação do SAEM, desde 2006, gerou um volume de dados ricos de informações pouco exploradas até o presente momento. Devido à sua importância como política pública da Rede Municipal de Ensino e como ferramenta de apoio ao trabalho do professor e da gestão escolar, o estudo aprofundado das características do SAEM é de extrema importância para se conhecer, em detalhes, o que o modelo de avaliação é capaz de produzir e, principalmente, para se vislumbrar caminhos que o sistema de avaliação deve direcionar.

³ Gêneros Textuais é todo e qualquer tipo de texto de qualquer natureza, literários ou não, tais como: bilhete, anúncio, poema, charges, letras de música etc. Para saber mais sobre o assunto, consultar MARCUSCHI (2003).

1.1. PROBLEMA DE PESQUISA

A questão que se pretende estudar neste trabalho é: **Quais são as limitações e as vantagens do SAEM como ferramenta de avaliação do desempenho acadêmico dos alunos de uma rede municipal de educação básica?**

1.2. JUSTIFICATIVA

A implantação de sistemas de avaliação que buscam mensurar o nível de qualidade da educação tem grande importância para o sistema educacional como um todo, na medida em que esses sistemas são geradores de indicadores nas quais são baseadas as decisões para a implementação de estratégias de melhora no nível de ensino (PESTANA, 1998). Segundo Freitas (2008) e Dourado, Oliveira e Santos (2007), a discussão sobre o tema é ampla e não está esgotada, haja vista que o tema “qualidade da educação” entrou em pauta somente nos anos 90, não por acaso, no mesmo período do surgimento dos Sistemas de avaliação no Brasil.

O SAEM, pela sua facilidade de produção, permite a aplicação do seu modelo em diferentes redes de ensino, com rapidez de apuração dos resultados. O sistema de avaliação faz uso dos desempenhos percentuais médios dos alunos, tanto na classificação geral quanto por disciplina, dispensando qualquer ajuste estatístico mais elaborado dos resultados. Como argumenta Freitas et al. (2009), um sistema de avaliação a nível municipal é o que mais gera o envolvimento dos principais agentes do processo de mudança pretendido – diretores, professores, pais e alunos, o que justifica ainda mais a sua importância, não no sentido de substituir os exames a nível estadual e federal, como a Prova Brasil/SAEB, mas sim de contribuir para a efetiva melhora da formação dos alunos.

A adoção de um modelo de avaliação temática se justifica pelas características de contextualização e interdisciplinaridade presentes nesse modelo

de avaliação, tornando-a mais atrativa para os alunos. A escolha dos temas leva em conta a atualidade e significância dos mesmos para a realidade e o crescimento cultural e científico que se deseja propiciar aos alunos. Os temas são divulgados com bastante antecedência – no início do ano letivo, com a avaliação ocorrendo no final do semestre e do ano letivo – possibilitando que os professores o desenvolvam com os alunos, trabalhando principalmente a leitura de textos de diversos gêneros.

Assim, ao escolher como tema a Amazônia, os professores buscam mostrar a seus alunos todas as relações entre o conteúdo de sua disciplina e o tema, trazendo para a sala de aula textos, reportagens e outros materiais para aprofundar o conhecimento.

De acordo com informações colhidas por professores e alunos durante a realização das entrevistas para esse trabalho, o modelo temático, com a necessidade de leitura de textos para a resolução das questões e o trabalho dos professores com textos sobre o tema, tem contribuído para o aumento do hábito da leitura entre os alunos.

Em relação às motivações pessoais, o presente trabalho se justifica pelo envolvimento direto do mestrando com o tema Avaliação Educacional já há alguns anos, ocupando a função de coordenador de avaliações do Colégio Cenecista Dr. José Ferreira e trabalhando diretamente na implantação do SAEM e na sua coordenação desde 2006.

1.3. OBJETIVO GERAL

Como objetivo geral, o presente trabalho visa a analisar as características do SAEM, ao identificar suas vantagens e limitações como ferramenta de aferição do desempenho acadêmico de alunos de uma rede de ensino e ao ressaltar pontos de melhoria que possam ser introduzidos ao processo.

1.3.1. OBJETIVOS ESPECÍFICOS

- a) Apresentar o modelo de avaliação proposto pelo SAEM, verificando sua adequação aos padrões de qualidade conforme preconiza a literatura e comparando-o com a Prova Brasil/SAEB.
- b) Descrever os passos de desenvolvimento e de aplicação de uma avaliação temática do SAEM, desde a escolha do tema à entrega dos resultados.
- c) Verificar se o SAEM atende aos quesitos de validade e confiabilidade de um sistema de avaliação à luz da teoria avaliativa.
- d) Identificar pontos de melhoria e oportunidades de novos usos dos resultados, ainda não explorados.

1.4. ORGANIZAÇÃO DO TRABALHO

Como forma de facilitar a compreensão, a Figura 2 apresenta a forma como o presente trabalho está organizado:

Figura 2. Esquema representativo do trabalho

2. REFERENCIAL TEÓRICO

O presente referencial teórico abordou as bases conceituais relativas à avaliação educacional e à avaliação em larga escala de redes de ensino, a fim de auxiliar a compreensão do objeto de estudo, o que é o SAEM.

Inicialmente, a pergunta “o que é avaliar?” foi respondida com uma definição mais genérica: Avaliação como emissão de um juízo de valor, compreendendo a medição de variáveis latentes (conhecimento) e aprofundando a discussão da coexistência da avaliação como medida e como uma prática social, dentro do campo da avaliação educacional.

A seguir, partiu-se para uma discussão do que se entende por qualidade da educação, tendo como base a visão da academia por meio de alguns autores que trabalham com o tema e uma pesquisa documental em artigos de lei e projetos do Ministério da Educação que abordam a qualidade da educação.

Após a compreensão do que se entende por qualidade da educação, tornou-se necessário explorar o que se tem de referencial teórico sobre a qualidade da avaliação. Com base em um dos principais autores que discorrem sobre o tema, Heraldo Marelim Vianna (VIANNA, 2000), chegou-se aos *Standards* sobre qualidade da avaliação definidos pelo *Joint Committee on Standards for Educational Evaluation (JCSEE)*.

Dentro do tema de desenvolvimento e de aplicações de sistemas de avaliação de ensino, o próximo passo foi conhecer as bases teóricas que alicerçam o principal sistema de avaliação em vigência, que é o Sistema de Avaliação da Educação básica (SAEB), de responsabilidade do Governo Federal. Uma das avaliações que compõem o SAEB, a Prova Brasil, avalia de forma censitária os alunos da rede pública na Educação Básica. Como o SAEM envolve os alunos da rede pública municipal de ensino de Uberaba, o modelo da Prova Brasil foi usado como referência para os estudos sobre o SAEM.

O referencial teórico está organizado conforme a Figura 3, a seguir:

Figura 3 - Organização do Referencial Teórico

2.1.O QUE É AVALIAR?

Segundo o dicionário Michaelis (2007), as acepções da palavra avaliar são:

a.va.li.ar (*a*¹+*valia*+*a*²) *vtd* e *vti* **1** Calcular ou determinar o valor, o preço ou o merecimento de. *vtd* **2** Reconhecer a grandeza, a intensidade, a força de: *Avaliar a dor, a mágoa.* *vtd* **3** Apreciar: *Avaliar a força, os costumes.* *vtd* **4** Computar, orçar: *Avaliar a riqueza de um povo. Avaliam a herança em dois milhões de dólares.*

Nota-se que são usados os verbos calcular, determinar e computar, todos relacionados ao ato primordial de uma avaliação que é o de atribuir um valor, de medir. A palavra avaliar vem do latim “a + valere”, que significa atribuir um valor, um juízo de valor ou mérito ao objeto – ou objetos – em estudo (LUCKESI, 1995). Barlow (2006, p. 12) define: “avaliar é emitir um julgamento preciso ou não sobre uma realidade quantificável ou não depois de ter efetuado ou não uma medição”. Com essa definição, o autor ressalta que avaliar também possui um caráter de estimativa. Assim, de modo quantitativo, avaliar é, por exemplo, medir a altura dos alunos de uma classe com o uso de um instrumento de medição, uma fita métrica, atribuindo um valor para cada medição. De modo qualitativo, avaliar é escolher as

maças mais “vistas” dentre todas as que estão na bancada do supermercado. Freud⁴ (1975 apud LIMA, 2002, p. 88) considera avaliar como o processo contínuo de realizar escolhas do sujeito em interação com o meio. Ou seja, desde o momento em que se está acordado, estão sendo feitas avaliações, julgamentos e medições sobre tudo o que ocorre ao nosso redor.

A partir da definição do verbo avaliar chega-se ao substantivo avaliação, definida pelo dicionário Michaelis (2007) como o “ato de avaliar”. As duas palavras, avaliar e avaliação, estão diretamente relacionadas à ação. Barlow (2006, p. 14) considera a avaliação como “o eco da ação”. Por esse motivo, nota-se que nessas definições generalistas sobre avaliar e avaliação surge uma pergunta que auxilia o entendimento das várias definições que esses termos assumem, conforme a situação em que são usadas, como: Avaliar para quê? Quais as funções de uma avaliação?

Ainda de modo generalista, a principal função de uma avaliação é fornecer informações para a tomada de decisões. A professora mede a altura de seus alunos para formar uma fila, ou seja, ela precisa saber a altura de cada um para decidir a posição deles numa fila em ordem crescente de altura, por exemplo. A pessoa avalia as maçãs para decidir quais ela vai comprar.

O ato de avaliar ou a avaliação, ao emitir um juízo de valor ou mesmo uma medida referente a um objeto, permite que diferentes objetos possam ser comparados, estabelecendo, assim, mais uma das funções – generalistas – de uma avaliação: a comparação. Entretanto, deve-se atentar para o fato de que a comparação só é possível quando a avaliação é feita a partir da mesma unidade de medida. As ciências da natureza possuem alguns exemplos claros sobre comparações a partir de unidades de medidas diferentes. Um carro a 90 Km/h tem maior velocidade que outro a 25 m/s? Uma pessoa com temperatura de 104 graus Fahrenheit está com febre?⁵ Nas ciências humanas o problema se agrava, uma vez que as unidades de medida podem não ser tão diretas. Qual a unidade de medida para a felicidade, a inteligência e o amor?

⁴ FREUD, S. O Ego e o Id. Rio de Janeiro: Imago, 1975

⁵ Uma velocidade de 90Km/h é igual a velocidade de 25 m/s. A temperatura de 104º Fahrenheit corresponde a 40º Celsius.

Focando um pouco mais a pesquisa para o tema avaliação educacional, deve-se destacar que o uso do termo é recente. Segundo Barlow (2006), somente na década de 70, na França, a palavra avaliação passou a ser usada no ambiente educacional pelos responsáveis da formação contínua de profissionais, em substituição a outras palavras mais ligadas à educação básica. No Brasil, nota-se que, na década de 60, quando os processos de avaliação definitivamente surgiram no País, a palavra mais usada para denominar esses processos era “teste”. Um exemplo disso é o título do primeiro livro de Heraldo Marelím Vianna “Testes em educação”, cuja primeira edição data de 1972. Outra prova é a que a definição do termo “testes educacionais” é de 1968, registrada por Novaes e Martins⁶ (1968, apud THESAURUS BRASILEIRO DA EDUCAÇÃO) como “testes destinados a avaliar a aprendizagem do examinando em relação a determinadas matérias de ensino”. Nota-se que o uso do termo “testes educacionais” tinha, inicialmente, o propósito de separá-los dos testes psicológicos, dos quais são originários.

Há também certa confusão em relação ao uso das palavras prova e avaliação. Para tanto, Luckesi (1995, p.169) define:

Provas/Exames têm por finalidade, no caso da aprendizagem escolar, verificar o nível de desempenho do educando em determinado conteúdo, entendendo por conteúdo o conjunto de informações, habilidades motoras e mentais, convicções, criatividade, e classificá-lo em termos de aprovação/reprovação [...]

Já avaliação é definida por Luckesi (1995, p. 172) como: “um ato amoroso, no sentido de que avaliação, por si, é um ato acolhedor, integrativo, inclusivo”. De um modo mais simples, pode-se definir que a prova é apenas um dos instrumentos que podem ser usados em uma avaliação.

Voltando à discussão sobre avaliação e medida, agora no âmbito da avaliação educacional, nota-se mais facilmente um distanciamento entre as duas definições. Hadji (2001), apesar de entender que a ideia de que a avaliação é uma medida dos desempenhos dos alunos está enraizada na mente de professores e alunos, defende veemente que a avaliação não é uma medida. O autor argumenta que, como o instrumento da avaliação é o professor (ou uma banca de professores), a grande variabilidade das notas dadas por diferentes grupos de professores para o

⁶ NOVAES, M. H.; MARTINS, O. Glossário de termos referente aos testes e medidas psicológicas. Rio de Janeiro: FGV, 1968.

mesmo conjunto de avaliações, comprovadas por diferentes pesquisas, mostram que os resultados não são confiáveis. Desse modo, as avaliações não podem ser consideradas como uma medida, segundo o autor.

Numa outra perspectiva, a relação entre medida e avaliação é definida como complementar por Popham⁷ (1983 apud DEPRESBITERIS, 1998), uma vez que o autor define que o processo avaliativo inclui a medida, mas nela não se esgota. Worthen⁸ (1982 apud DEPRESBITERIS, 1998) traz um bom exemplo da diferença entre medida e avaliação:

[...] a competição de saltos da qual participam vários esportistas. A medida responde à pergunta: "Qual a altura que cada esportista conseguiu saltar?" - é o simples ato de determinar a altura máxima do salto de cada indivíduo. Já a avaliação responde às perguntas: "Dado um critério para a altura do salto, que rapazes conseguiram alcançar esse critério?" e "O programa adotado por determinado instrutor foi satisfatório?".

Desse modo, o autor entra numa definição mais específica de avaliação educacional, que é a da avaliação da aprendizagem. Vianna (1989) expõe que os estudos específicos sobre avaliação da aprendizagem têm início, internacionalmente, na década de 40 e cita um dos principais pesquisadores do tema, Ralph W. Tyler. A Tyler⁹ (1942 apud VIANNA, 1989, p. 20) é creditada a definição de avaliação como "um processo de comparação com os dados de desempenho e os objetivos instrucionais preestabelecidos". Vianna (1989) relata ainda que o campo da avaliação educacional se desenvolveu a partir da década de 60, com os trabalhos, principalmente, de Lee J. Cronbach, Michael Scriven e Robert E. Stake. A avaliação do processo de ensino e aprendizagem – também chamada de avaliação interna – é, sem sombra de dúvidas, o tema mais pesquisado dentro da avaliação. Entretanto, o tema não faz parte do objeto de estudo desse trabalho, que se dedica às avaliações externas em larga escala.

Numa perspectiva mais ampla dentro da avaliação educacional, além da avaliação da aprendizagem, está a avaliação de programas e de currículos, que, segundo Depresbiteris (1989, p. 5), "visa ao programa como um todo, expandindo-se

⁷ POPHAM, W.J. Avaliação educacional. Porto Alegre, Globo, 1983.

⁸ WORTHEN, B.R. Visão geral do mosaico formado pela avaliação e controles educacionais. In: GOLDBERG, M.A.; SOUZA, C.P., orgs. Avaliação de programas educacionais: vicissitude, controvérsias e desafios. São Paulo, EPU, 1982. p.4-9.

⁹ TYLER, R. W. *General statement on evaluation. Journal of Educational Research*. v.35. 1942.

aos propósitos educacionais, aos currículos e incluindo relações com outros sistemas: social, econômico e político”. Pophan¹⁰ (1983 apud DEPRESBITERIS, 1989) considera que o início da avaliação de programas está relacionado à derrota do governo americano na guerra pela conquista do espaço, com o lançamento do primeiro satélite artificial – o Sputnik – pelos russos. O sistema educacional americano foi severamente questionado e, a partir de então, verificou-se um grande investimento em sua melhoria e, por conseguinte, em sistemas de avaliação de currículos e programas.

Um dos modelos de avaliação de programas mais citados é o da avaliação responsiva, proposto por Robert E. Stake (DEPRESBITERIS, 1989; VIANNA, 2000). Vianna (2000) apresenta um resumo dos procedimentos para realização de uma avaliação responsiva, segundo Stake¹¹ (1993 apud VIANNA, 2000, p. 39)

1) discutir com os responsáveis pelo programa; 2) identificar o escopo do programa; 3) analisar as atividades do programa; 4) identificar propósitos e preocupações; 5) conceituar questões e problemas; 6) identificar necessidades de dados; 7) selecionar observadores, juízes e até mesmo instrumentos, se necessários; 8) observar antecedentes, interações e resultados, como no modelo *counthance*; 9) preparar descrições, estudos de casos; 10) validar – confirmar ou “desconfirmar”; 11) elaborar informações para vários públicos; e 12) apresentar relatório formal, se for o caso.

Depresbiteris (1989, p. 16) aborda com mais clareza o teor do item 8, acima.

Os antecedentes são todas as condições anteriores ao programa e que se relacionam com os resultados, tais como: interesses, recursos e experiências anteriores. As transações são todas as atividades realizadas para o alcance dos objetivos, incluindo relação e interações entre aluno e professor. Os resultados são os produtos da aprendizagem e que podem ser identificados como habilidades, realizações, atitudes, expectativas cumpridas, transferência de aprendizagem.

Finalmente, Vianna (2000) e Depresbiteris (1989), mesmo em épocas distantes, destacaram que uma avaliação é responsiva se sua orientação está mais voltada para as atividades do programa e menos para os seus objetivos.

Quando aplicadas aos sistemas de ensino, as avaliações em larga escala são denominadas de avaliações sistêmicas (VIANNA, 2003). O autor justifica a importância de tais avaliações a fim de: 1) elevar o nível de ensino ofertado; 2)

¹⁰ POPHAM, W.J. Avaliação educacional. Porto Alegre, Globo, 1983.

¹¹ STAKE, R. E. *Program Evaluation, Particularly Responsive Evaluation. (Paper) Conference on New Trends in Evaluation*, Göteborg, Suécia, 1973. In: MADDAUS, G. F. et al. *Evaluation Models*. Boston, Kluwer-Nijhoff Publishing, 1993.

ajustar os processos de ensino e aprendizagem; 3) contribuir com a formação de cidadãos capazes de enfrentar a complexidade da sociedade tecnológica atual; e 4) proporcionar aos gestores o *feed-back* necessário para apoiar a tomada de decisões. Em complemento, Fernandes e Gremaud (2009) apresentam que as avaliações de sistemas têm incorporado os objetivos do chamado *accountability*, termo da língua inglesa que, em tradução livre, significa “prestação de contas”. Em relação à avaliação educacional, *accountability* representa o papel da avaliação sistêmica de “prestar contas” à sociedade dos serviços educacionais ofertados pelos sistemas de ensino.

Fernandes (2009) usa a expressão Avaliação Externa para identificar os exames e os estudos internacionais promovidos pelos Ministérios da Educação de diversos países, tais como Estados Unidos, Inglaterra, Holanda, França e Portugal. Dentre as funções de uma avaliação externa estão: 1) certificação, que pode conferir algum certificado de conclusão de curso ao aluno; 2) seleção, determinando o acesso dos alunos ao ensino superior, por exemplo; 3) controle, para que os governos atestem que os conteúdos previstos têm sido ensinados; 4) monitoração, que faz a função do *accountability* e gera rankings de alunos e escolas; e 5) motivação, na medida em que os exames têm efeito sobre o progresso escolar do aluno.

Assim, avaliar, no sentido que foi usado neste trabalho, é desenvolver um sistema de avaliação, com o uso de provas temáticas e a partir de um conteúdo predeterminado, visando a aferir o desempenho dos alunos de uma rede de ensino e, de posse dos resultados, informar à sociedade a qualidade do ensino que é ofertado por aquela rede. Além disso, objetiva-se fornecer informações para a definição de políticas públicas que resultem na melhoria da aprendizagem dos alunos.

2.2. QUALIDADE DA EDUCAÇÃO

O termo “qualidade da educação” tem sido usado frequentemente em referência a toda e qualquer dimensão da educação. Um gestor público, ao

implantar uma nova política educacional, pretende melhorar a qualidade da educação. Um professor, ao ministrar boas aulas a seus alunos, busca melhorar a qualidade da educação. Um diretor escolar, ao investir na compra de equipamentos, o faz para melhorar a qualidade da educação. Entretanto, a definição de qualidade da educação envolve nada menos do que todos os processos educativos. Segundo Dourado, Oliveira e Santos (2007, p. 9), a qualidade da educação é um fenômeno complexo e abrangente, “não podendo ser apreendido apenas por um reconhecimento da variedade e das quantidades mínimas de *insumos* considerados indispensáveis ao desenvolvimento do processo de ensino aprendizagem”.

Desse modo, é necessário que se determinem as dimensões e os fatores, intraescolares e extraescolares, que influenciam no processo educativo. Além disso, é importante que também se determinem quais dessas dimensões e fatores são possíveis de ser mensuradas, a fim de que se tornem indicadores representativos da qualidade educacional.

Antes disso, entretanto, é importante analisar as condições necessárias para que a avaliação contribua para a qualidade da educação. Murillo (2011, p. 95) questiona se essa onda de avaliar “qualquer coisa que se mexe” tem realmente trazido bons resultados para a educação. O autor pondera que, para a avaliação atender à qualidade da educação, ela deve considerar cinco postulados: 1) A educação está conformada por uma série de componentes inter-relacionados e dependentes entre si formando um todo global e inseparável; 2) A educação é tarefa compartilhada por famílias, docentes, comunidade escolar no seu conjunto e sociedade; 3) A educação é uma tarefa para se desenvolver ao longo da vida toda; 4) A educação, para que seja de qualidade, deve buscar o desenvolvimento de todos e cada um dos estudantes; e 5) A finalidade da educação é o desenvolvimento integral das pessoas.

Os dois primeiros postulados já comprometem fortemente os atuais sistemas de avaliação existentes. Se a educação possui vários componentes e é tarefa de todos, uma avaliação centrada no desempenho dos alunos não é capaz de indicar a qualidade da educação. Murillo (2011) propõe que as avaliações passem a considerar esses outros componentes e agentes.

Ao abordar o terceiro postulado, o autor defende a ideia de que os sistemas de avaliação não devem se concentrar somente no Ensino Fundamental e Médio, centrando atenção na primeira infância e, também, na formação permanente de adultos.

O quarto postulado, ao tratar da educação de todos e de cada um, defende que a avaliação deve contribuir para a inclusão, levando em conta a diversidade cultural, socioeconômica, étnica, de gênero, etc., reconhecendo as necessidades de cada grupo.

Desenvolvimento integral dos estudantes é o quinto postulado analisado por Murillo (2011). Tão importante quanto o cognitivo, o desenvolvimento socioafetivo e ético é fundamental para a qualidade da educação e, por seu aspecto de inter-relação, facilita a aprendizagem. O autor reconhece o poder da avaliação em pressionar os sistemas, já que o que é avaliado passa a ter maior relevância. Assim, a inclusão da avaliação das outras habilidades, não cognitivas, contribuiria fortemente para a melhora da qualidade da educação.

Dourado e Oliveira (2009) ressaltam que qualidade é um conceito histórico, ou seja, que varia no tempo e no espaço conforme as demandas e as exigências sociais de um determinado período. Assim, o que se entende hoje por qualidade da educação, não é o que se entendia em outros períodos históricos. Na década de 90, conforme determinação constitucional (BRASIL, 1988), o Governo Federal priorizou a questão da universalização do Ensino Fundamental. Posteriormente, após o acesso ter atingido cerca de 97% de cobertura, o foco passou a ser a permanência e a progressão dos alunos no sistema. Atualmente, com entrada das avaliações externas, a qualidade passa a se vincular a várias outras dimensões, principalmente à aprendizagem dos alunos (OLIVEIRA, 2007, 2011).

Dourado, Oliveira e Santos (2007) apresentam que a visão da UNESCO e da OCDE¹² sobre qualidade da educação, no início da década de 2000, tinha como paradigma a relação entre insumos (materiais e humanos), processos (o que ocorre na sala de aula) e resultados, defendendo que a qualidade pode ser definida pelo desempenho dos alunos.

¹² Organização para a Cooperação e Desenvolvimento Econômico, responsável pela Avaliação Internacional PISA.

Freitas (2008) mostra a mudança da visão da UNESCO em relação ao tema “qualidade da educação”. Segundo a autora, de uma visão de educação integrada e abrangente do Relatório Delors¹³, de 1995, a UNESCO passa para uma visão de qualidade como direito humano para todas as atividades educacionais, no documento “Educação de qualidade para todos”¹⁴ (EDUCAÇÃO, 2008).

O documento reforça a questão da polissemia do termo “qualidade da educação” (p. 29):

Trata-se de um conceito com grande diversidade de significados, com frequência não coincidentes entre os diferentes atores, porque implica um juízo de valor concernente ao tipo de educação que se queira para formar um ideal de pessoa e de sociedade. As qualidades que se exigem do ensino estão condicionadas por fatores ideológicos e políticos, pelos sentidos que se atribuem à educação num momento dado e em uma sociedade concreta, pelas diferentes concepções sobre o desenvolvimento humano e a aprendizagem, ou pelos valores predominantes em uma determinada cultura. Esses fatores são dinâmicos e mutantes, razão por que a definição de uma educação de qualidade também varia em diferentes períodos, de uma sociedade para outra e de alguns grupos ou indivíduos para outros.

Além disso, o documento dá uma grande contribuição ao recomendar onze propostas de políticas públicas para cumprir com os objetivos de educação de qualidade para todos; 1) Assegurar o papel do Estado como fiador e regulador do direito universal a uma educação de qualidade; 2) Garantir o direito das pessoas de aprender durante toda a vida; 3) Avançar a partir de enfoques homogêneos e padronizados rumo a políticas educacionais integrais que consideram a diversidade com coesão social; 4) Enfatizar as políticas destinadas a garantir a inclusão; 5) Distribuição equitativa de recursos para uma escola pública gratuita e de qualidade para todos; 6) Melhorar o equilíbrio na distribuição de tarefas e responsabilidades entre os diferentes níveis de gestão, fortalecendo o protagonismo dos atores locais; 7) Propor políticas integrais para o fortalecimento da profissão de docente; 8) Desenhar e desenvolver currículos relevantes e pertinentes para todo o alunado; 9) Implantar políticas que tenham como centro a transformação das escolas para que sejam mais inclusivas e consigam melhores aprendizagens; 10) Propor políticas articuladas de avaliação educativa orientadas para a melhoria progressiva da

¹³ Relatório para a UNESCO da Comissão Internacional sobre Educação para o século XXI. Publicado no Brasil com o título “Educação: um tesouro a descobrir” (Editora Cortez, 1999).

¹⁴ Documento elaborado a partir das deliberações da Segunda Reunião Intergovernamental de Ministros de Educação do Projeto Regional de Educação para a América Latina e o Caribe (PRELAC) ocorrida em março de 2007, em Buenos Aires (EDUCAÇÃO, 2008).

qualidade da educação e do funcionamento dos sistemas educacionais 11) Construir sistemas integrados de informação educativa que alimente os processos de tomada de decisão e de debate público sobre a educação.

A evolução histórica do que se entende por qualidade da educação pode ser também analisada pelos documentos oficiais disponíveis. A partir da Constituição de 1988 (BRASIL, 1988), algumas leis buscam garantir o direito a uma educação de qualidade. A Carta Magna, em seu artigo 206º, que trata dos princípios do ensino, tem, no inciso VII, a “garantia do padrão de qualidade”. Já o artigo 211º, que trata da organização dos sistemas de ensino em regime de colaboração entre União, Estados, Distrito Federal e Municípios, em seu primeiro parágrafo, define que a União tem função de garantir um padrão mínimo de qualidade do ensino. Entretanto, nota-se que não há definição do que se entende por qualidade da educação.

A Lei de Diretrizes e Bases da Educação Nacional (BRASIL, 1996), reeditada em 1996, no artigo 4º, inciso IX, define como dever do Estado a garantia de “padrões mínimos de qualidade de ensino definido como a variedade e quantidade mínimas, por aluno, de insumos indispensáveis ao desenvolvimento do processo de ensino e aprendizagem”. Desse modo, ao contrário do que ocorre na Constituição, percebe-se a existência de uma definição de qualidade educacional vinculada aos insumos e atrelada ao processo de ensino e aprendizagem.

A Constituição Federal ainda aborda, em seu artigo 214º, a questão da qualidade da educação ao instituir a criação de um Plano Nacional de Educação, com objetivo de estabelecer metas para garantir a manutenção e desenvolvimento do ensino por meio de ações integradas dos poderes públicos que conduzam a melhora da qualidade de ensino (BRASIL, 1988). O Plano Nacional de Educação vigente, Lei nº 10.172 de 9 de janeiro de 2001, tem em seus objetivos e prioridades a questão da melhoria da qualidade do ensino (BRASIL, 2001). Mais especificamente, o Plano, ao tratar da questão da formação de professores da educação básica, vincula claramente a qualidade da educação à valorização do magistério. Além disso, o Plano avança em relação à universalização da Educação Infantil e do Ensino Médio, considerada obrigação do Estado somente a partir de Emenda Constitucional nº 59, de 2009.

O Plano Nacional de Educação tem vigência de 10 anos. Assim, conforme estabelece a Constituição Federal, o Ministério da Educação organizou Conferências Municipais, Estaduais e a Conferência Nacional de Educação (CONAE), ocorrida em abril de 2010, para discutir as metas que iriam compor o novo Plano Nacional de Educação¹⁵, com vigência prevista de 2011 a 2020. O documento final da CONAE 2010 (BRASIL, 2010) aborda a questão da qualidade da educação em dois dos seis temas existentes. No primeiro tema, denominado Papel do Estado na Garantia do Direito à Educação de Qualidade: Organização e Regulação da Educação Nacional, o documento analisa a questão da qualidade sob o ponto de vista da garantia do acesso, da importância da consolidação de um Sistema Nacional de Educação e da efetivação do Regime de Colaboração entre os sistemas de ensino federais, estaduais, do Distrito Federal e dos municípios. Além disso, o documento (BRASIL, 2010, p. 30) ainda destaca que:

Ao eleger a **qualidade** como parâmetro de suas diretrizes, metas, estratégias e ações e conferindo a ela uma dimensão social e histórico-política e, portanto, inclusiva, a constituição do referido Sistema Nacional de Educação significará investimento na educação e envolverá questões como: financiamento; inclusão social; reconhecimento e valorização à diversidade; gestão democrática e formação e valorização dos profissionais da educação, dentre outros.

O segundo eixo do documento final da CONAE é denominado Qualidade da educação, gestão democrática e avaliação. Nessa parte, o documento aborda profundamente a questão da qualidade da educação. Primeiro, resgata-se o que se entende por educação como sendo “elemento partícipe das relações sociais, contribuindo, contraditoriamente, para a transformação e a manutenção dessas relações” (BRASIL, 2010, p. 46). De acordo com o texto, o estabelecimento de parâmetros de qualidade nacionais é dificultado pela estrutura descentralizada de oferta de ensino por todos os entes federados. Mesmo assim, o documento avança ao definir as diretrizes e bases para a efetivação de políticas de Estado voltadas à qualidade da educação, relacionados a seguir: a) articulação das dimensões intra e extraescolares; b) consideração da dimensão socioeconômica e cultural; c) discussão sobre o custo aluno-qualidade; d) promoção e atualização histórico-

¹⁵ O novo Plano Nacional de Educação está em tramitação no Congresso Nacional e tem sua votação prevista para o segundo semestre de 2011. A minuta do projeto de lei pode ser obtida pelo endereço: <http://www.todospelaeducacao.org.br/arquivos/biblioteca/pne_15_12_2010.doc>

cultural visando à formação ética e solidária; e) discussão das práticas escolares, dos processos formativos, do planejamento pedagógico, dos processos de participação e da dinâmica das avaliações; f) relações de alunos por turma, alunos por professores e alunos por funcionário técnico-administrativo; g) garantia de financiamento público da educação; h) atenção para a estrutura física da escola; i) organização sindical e estudantil livre; j) identificação dos professores com a escola, vinculada aos processos de formação inicial e continuada; e k) engajamento de alunos e professores no processo político-pedagógico e de ensino-aprendizagem.

O documento final da CONAE (BRASIL, 2010) também define o que devem ser considerados como dimensões, fatores e condições para a educação de qualidade. A dimensão extraescolar é analisada em termos do espaço social e das obrigações do Estado. Espaço social refere-se à dimensão socioeconômica e cultural, cujo acúmulo tem influências positivas no processo ensino-aprendizagem. Refere-se também, principalmente, à consideração da identidade individual e social dos alunos, com vistas ao seu desenvolvimento integral e à sua permanência no espaço escolar. A outra dimensão extraescolar são as obrigações do Estado, que envolve desde as questões de garantia da oferta de ensino com igualdade de acesso e permanência até a implementação de programas de fornecimento de livros didáticos, alimentação e saúde dos alunos, transporte escolar, recursos tecnológicos e segurança.

As dimensões intraescolares são analisadas a partir da divisão em seus quatro principais elementos. Assim, tem-se, primeiramente, um plano do sistema, que compreende a oferta de um ambiente educacional apropriado à realização de atividades de ensino, que envolve desde os equipamentos da sala de aula, passando pelos laboratórios, brinquedoteca e salas multifuncionais, até a biblioteca, com livros em quantidade suficiente e acesso ao acervo *on line*.

O segundo plano dentro da dimensão intraescolar definido pelo documento final da CONAE (BRASIL, 2010) é o plano da instituição. Nesse plano, estão as ações de gestão e de organização do trabalho educativo, que trata, dentre outras coisas: do planejamento, do monitoramento e da avaliação dos programas e projetos; da disponibilidade de docentes para as atividades curriculares; da gestão democrática que estimule a participação de toda a comunidade escolar e da

construção e condução do projeto político-pedagógico, com explicitação de seus fins sociais e pedagógicos.

O professor é o elemento principal do terceiro plano da dimensão intraescolar da qualidade da educação. Seu trabalho deve estar focado, dentre outras coisas, para a qualificação adequada ao exercício profissional e formação continuada para progressão na carreira; a efetiva dedicação a somente uma instituição educativa; e a existência de carga horária para as atividades de planejamento, reuniões pedagógicas, atendimento a pais e alunos.

O último plano da dimensão intraescolar é o do aluno, que leva em conta as questões de acesso, permanência e desempenho. Nesse sentido: devem-se focar as questões relativas à existência de condições de permanência e desenvolvimento educacional dos alunos independentemente de suas condições socioeconômicas, etnicoraciais, de gênero ou culturais, bem como de necessidades especiais; à visão de qualidade que os alunos e seus pais têm da instituição educativa e à valorização dos processos avaliativos centrados na melhoria das condições de aprendizagem.

Em abril de 2007, o Governo Federal lançou o Plano de Desenvolvimento da Educação (PDE) (BRASIL, 2007c), com objetivo de concentrar ações para enfrentar o desafio de melhorar a qualidade da educação (SAVIANI, 2007). Segundo o autor, o PDE reuniu antigos e novos programas de governo, inclusive fora do Ministério da Educação, como é o caso do programa Luz para Todos, tentando dar-lhes um sentido de inter-relação, que, na visão do autor, é inexistente.

Além do documento do PDE (BRASIL, 2007c), que reúne os trinta programas governamentais voltados para a qualidade da educação, o Governo Federal editou um decreto, nº 6094/2007, que implementou o Plano de Metas Compromisso Todos Pela Educação (BRASIL, 2007c). Conforme expõe o texto do PDE, o Plano de Metas contém 28 diretrizes que devem ser seguidas pelo Governo Federal, em regime de colaboração com os governos municipais, estaduais e Distrito Federal, que, voluntariamente, assinarem a adesão ao Plano de Metas. Segundo o PDE, essas diretrizes surgiram de dois estudos em escolas e redes de ensino cujos alunos apresentaram desempenho acima do previsto na Prova Brasil de 2005.

Saviani (2007), ao estudar a elaboração do PDE, cita a influência do Movimento Todos pela Educação¹⁶ na elaboração do PDE e do decreto nº 6094/2007, haja vista a proposital inclusão do nome do movimento no nome do Plano de Metas.

Após a discussão das várias dimensões envolvidas na definição da qualidade da educação, torna-se necessário lançar olhares para a questão de como se afere a qualidade da educação, ou seja, para a questão da construção e o uso dos indicadores de qualidade. De modo geral, indicadores são sinais que revelam aspectos de determinada realidade (INDICADORES, 2007). A temperatura corporal é um indicador de febre. A inflação é um indicador da Economia. Esses indicadores são estabelecidos a partir da definição de quais dimensões, dentre todas as mensuráveis, serão levadas em conta em seu cálculo.

A primeira discussão que surge é relativa à capacidade de um indicador, que não inclui todas as dimensões da qualidade da educação, em mensurar, com apresentação de resultados confiáveis, essa qualidade. Fernandes e Gremaud (2009) mostram que não há unanimidade em relação a essa questão. Os autores relatam que existem estudos que mostram que há aumento no desempenho dos alunos, simplesmente a partir da adoção de indicadores. Entretanto, os autores alertam que esse aumento no valor dos indicadores pode não estar associado à melhora da qualidade da educação, uma vez que as escolas podem adotar práticas de treinamento de alunos para o teste – denominado *gaming* – ou mesmo a escolha dos melhores alunos. Além disso, ao definir indicadores relacionados ao desempenho acadêmico de alunos, há o risco de reducionismo do currículo, que tende a ficar limitado ao currículo cobrado pelo exame.

Oliveira (2011) reforça os problemas citados, relacionados à avaliação, quando a mesma se mantém restrita à avaliação de resultados, e defende a avaliação centrada na tríade insumos – resultados – processos, de modo semelhante à posição da Unesco apresentada por Dourado, Oliveira e Santos (2007). Entretanto, Oliveira (2011) considera que um indicador de resultados como é

¹⁶ O Movimento Todos pela Educação é uma organização não governamental criada em 2006 com objetivo de contribuir para o acesso à educação e para o sucesso escolar. É financiada exclusivamente pela iniciativa privada.

o caso do IDEB não pode ser considerado um indicador de qualidade. Completa o autor (p. 127):

Entendo que considerá-lo um indicador de qualidade, como tem sido apresentado, além de um equívoco, é um risco. Aqui reside um problema importante, pois se ele se generaliza como indicador de qualidade, o risco de modelar políticas buscando apenas elevá-lo cria problemas conceituais e práticos complicados.

Ao discutir a garantia ao direito à educação com padrões mínimos de qualidade, presente nos documentos legais, Oliveira (2011) defende a necessidade de indicadores de insumos e vai além, indicando que o não cumprimento de alguns deles, como o custo aluno qualidade, por parte dos gestores públicos, gere consequências jurídicas para os mesmos.

Em relação aos indicadores de processos, Oliveira (2011) faz uma observação importante em relação a quais processos pedagógicos devem ser escolhidos. A priori, processos que influenciam no desempenho cognitivo dos alunos não devem ser levados em conta, uma vez que a influência dos mesmos já estará registrada pelos indicadores de resultados. Na medida em que a concepção da qualidade da educação não se restringe à transmissão de conhecimento, o autor reforça a importância de se considerarem outros objetivos como a educação para a diversidade, a ética e os valores e considera o desenvolvimento desse indicador como “questão de fronteira da pesquisa em política educacional no Brasil hoje” (p. 132).

Conforme o decreto nº 6094/2007 (BRASIL, 2007b), o Índice de Desenvolvimento da Educação Básica (IDEB) é o indicador oficial para a verificação do cumprimento das metas do Compromisso Todos pela Educação. Fernandes e Gremaud (2009) defendem a criação do IDEB como um indicador capaz de aferir a qualidade da educação e também introduzir os objetivos de *accountability* no sistema educacional brasileiro, (p. 217) “cujo cerne é considerar não apenas os alunos, mas escolas, professores, diretores e gestores como responsáveis pelo desempenho dos estudantes” explicam os autores.

Os programas de *accountability* possuem três características comuns: 1) ênfase no desempenho acadêmico, podendo incluir outras dimensões como frequência e evasão; 2) Co-responsabilidade de professores, diretores e gestores; e 3) Presença de sistema de incentivos e metas.

Ao discutirem os prós e contras da introdução dos objetivos de *accountability*, os autores destacam que um dos riscos existentes é o agravamento da repetência e da exclusão de alunos de baixo rendimento. Desse modo, os autores mostram que o IDEB, ao levar em conta também a taxa média aprovação, busca minimizar esse risco.

Em termos de abrangência, o IDEB, por utilizar os resultados de proficiência dos alunos na Prova Brasil, está limitado às redes públicas de Ensino Fundamental. Assim, não existe o indicador para as escolas particulares nem mesmo para as escolas rurais. Para o Ensino Médio, o indicador que tem sido usado é o ENEM (Exame Nacional do Ensino Médio). Para introduzir os objetivos de *accountability* no ENEM, a partir de 2005, os resultados passaram a ser divulgados por Estado, sistema de ensino (municipal, estadual ou federal) e, principalmente, por escola (FERNANDES e GREMAUD, 2009). Entretanto, o ENEM tem diferenças importantes em relação ao IDEB. Em termos de abrangência, o ENEM cobre tanto a rede pública quanto a rede privada. Em relação às dimensões, o índice do exame leva em conta somente o desempenho dos alunos, favorecendo os riscos de evasão e de exclusão já citados. Por último, o ENEM não é nem censitário nem amostral, o que aumenta consideravelmente os riscos de *gaming* pela participação dos melhores alunos¹⁷. O uso desses dois indicadores – IDEB e ENEM – fez com que o Brasil, a partir de 2005, passasse a ter um sistema nacional de *accountability* da Educação Básica (FERNANDES e GREMAUD, 2009).

Mesmo com o foco no uso de indicadores de resultados para a mensuração da qualidade da educação, o Governo Federal, através do INEP (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira), lançou, em 2004, com uma nova edição – ampliada – em 2007, o documento Indicadores da Qualidade na Educação (INDICADORES, 2007), em parceria com a ONG Ação Educativa, a Unicef (Fundo das Nações Unidas para as Crianças) e o PNUD (Programa das Nações Unidas para o Desenvolvimento).

¹⁷ A partir de 2009, o MEC lançou o NOVO ENEM que, dentre outras mudanças, passou a ser usado como único processo de acesso ao Ensino Superior de várias Universidades Federais. Essa ação tende a diminuir os problemas de participação somente dos melhores alunos.

O documento (INDICADORES, 2007), voltado para o Ensino Fundamental, estabelece sete dimensões fundamentais para que uma escola possa se basear ao analisar a qualidade da educação praticada. Não há avaliação de desempenho de alunos. O seu objetivo é orientar os gestores escolares a realizarem um diagnóstico da escola e obterem os indicadores de cada dimensão proposta para que, a partir do conhecimento da sua realidade, proponham mudanças com vistas à melhoria da qualidade. O documento é, na verdade, um guia tutorial que orienta passo a passo como proceder para avaliar a qualidade da escola, envolvendo toda a comunidade escolar.

A primeira dimensão aborda a questão do Ambiente Educativo, que envolve, dentre outros aspectos, as questões sociais de respeito, amizade, solidariedade e o combate à discriminação. O tema da segunda dimensão é Prática Pedagógica e Avaliação, envolvendo tanto as questões de planejamento das aulas e de conhecimento e atendimento às dificuldades quanto às questões relativas aos processos de avaliação dos alunos, defendendo claramente a proposta formativa de avaliação.

A terceira dimensão considerada dá ênfase ao Ensino e Aprendizagem da Leitura e da Escrita. Assim, são avaliadas as práticas da escola em relação à alfabetização inicial, à orientação dada aos pais para o incentivo à leitura em casa, acesso à biblioteca e, também, como o Projeto Político Pedagógico aborda essas questões.

Gestão Democrática é o tema da quarta dimensão abordada pelo documento Indicadores da Qualidade na Educação (INDICADORES, 2007), envolvendo, como não poderia deixar de ser, as ações dos gestores em termos de compartilhamento de informações e decisões e a transparência no uso dos recursos financeiros.

A quinta dimensão avalia a Formação e Condições de Trabalho dos Profissionais da Escola, a partir do questionamento sobre o funcionamento da formação continuada, a quantidade de profissionais e a assiduidade dos mesmos. Já a sexta dimensão aborda o Ambiente Físico Escolar. Nessa dimensão, os indicadores foram divididos em três níveis: 1) Suficiência, que verifica a existência do material; 2) Qualidade, para analisar se o material é bem cuidado; e 3) Bom aproveitamento, que analisa se o material é valorizado e usado eficientemente.

A última dimensão trata da questão do Acesso e Permanência dos Alunos na Escola. Desse modo, são avaliados como a escola trabalha com a questão das faltas dos alunos e das estratégias de recuperação dos alunos com deficiência de aprendizagem. Finalizando, o documento apresenta uma sugestão metodológica, bem prática, para trazer de volta alunos que abandonaram a escola.

De todos os aspectos discutidos como determinantes para se obter uma educação de qualidade, é interessante observar que, ao tratar da importância da gestão, o foco principal se dá na questão da democratização da gestão, da participação da comunidade (professores, alunos e famílias) nas decisões que os afetam. Entretanto, conforme destaca Loschpe (2011), a qualificação dos gestores, especialmente a do diretor da escola, tem uma clara influência na qualidade da educação ofertada pela escola. A condução de todos os processos que impactam na qualidade da educação, da manutenção dos insumos materiais e da formação continuada dos insumos humanos à construção de um ambiente educacional que propicie a aquisição do saber depende de um gestor com habilidades administrativas, principalmente relativas à gestão de equipes e com visão educacional que valorize os fatores que impactam na qualidade da educação.

Mesmo assim, ao ler as 20 metas definidas para o novo Plano Nacional de Educação (PNE) (MEC, 2010), verifica-se que não há nenhuma meta que foque a qualificação dos gestores. A meta 15, única a tratar da gestão, busca garantir que a eleição do diretor seja realizada por critérios de mérito e com a participação da comunidade.

Finalmente, a pesquisa por documentos¹⁸ e a posição da academia através da apresentação de alguns autores¹⁹ que discutem sobre a qualidade da educação mostram que, apesar de a definição sobre o tema ainda estar em construção e ser mutável ao longo do tempo, existe uma série de fatores que podem ser associados à

¹⁸ Em ordem de citação no texto: Constituição Federal (BRASIL, 1988), Educação de qualidade para todos (EDUCAÇÃO, 2008), Lei de Diretrizes e Bases da Educação Nacional (BRASIL, 1996), Plano Nacional de Educação (BRASIL, 2001), Documento final da CONAE 2010 (BRASIL, 2010), Plano de Desenvolvimento da Educação (PDE) (BRASIL, 2007c), Plano de Metas Compromisso Todos Pela Educação (BRASIL, 2007b) e Indicadores da Qualidade na Educação (INDICADORES, 2007).

¹⁹ Em ordem de citação no texto: (DOURADO; OLIVEIRA; SANTOS, 2007), (MURILLO, 2011), (DOURADO; OLIVEIRA, 2009), (OLIVEIRA, 2007, 2011), FREITAS (2008), (SAVIANI, 2007) e (FERNANDES E GREMAUD, 2009).

melhora da qualidade e, na medida em que são passíveis de ser mensurados, devem compor a construção de indicadores de qualidade, em complemento aos resultados de desempenho acadêmico dos alunos.

Embora se deseje a entrada de novos componentes na mensuração da qualidade da educação, o principal indicador usado é o rendimento escolar, obtido a partir de avaliações externas de larga escala. Desse modo, a discussão sobre qualidade deve avançar da qualidade da educação para a qualidade da avaliação.

2.3. QUALIDADE DA AVALIAÇÃO

A grande evolução que as avaliações passaram desde o início de sua disseminação, no final da década de 80, teve como um dos seus grandes benefícios a preocupação com a sua qualidade. Ao adotar indicadores de qualidade da educação obtidos por avaliações e seus resultados como retrato da realidade educacional (OLIVEIRA, 2011), há uma grande responsabilidade na realização de uma avaliação de qualidade.

O pesquisador Heraldo Marelin Vianna discute a questão da qualidade da avaliação em vários de seus artigos (VIANNA, 1998, 2000, 2001, 2005); algumas vezes de forma explícita, pontuando aspectos decisivos para o sucesso de uma avaliação; outras vezes de forma implícita, quando usa toda a sua experiência na área para emitir seus conselhos. Em *Avaliação Educacional: vivências e reflexões* (VIANNA, 1998), o autor procura apresentar sua visão acerca dos problemas considerados e vivenciados por ele no período de 1962 a 1995.

No livro *Avaliação Educacional: teoria, planejamento e modelos*, Vianna (2000) faz algumas considerações teóricas sobre a avaliação, apresenta os modelos de avaliação de Tyler, Cronbach, Scriven, Stufflebeam e Stake e, no último capítulo, faz uma discussão imprescindível para a qualidade da avaliação, que é a questão da formação dos avaliadores.

Ao abordar a avaliação em larga escala, Vianna (2001) tece algumas considerações diretas sobre a questão da qualidade da avaliação. Devido às

proporções de uma avaliação em larga escala, o autor alerta para a necessidade da existência de estrutura básica de funcionamento e, principalmente, a presença de um *controller* que interaja com todas as equipes, “de forma a garantir uma produção de qualidade em todos os estágios do seu desenvolvimento” (p. 94).

Em 2005, o autor volta a tratar da qualidade da avaliação no artigo *Avaliação de Programas: duas questões* (VIANNA, 2005). A primeira questão se refere à natureza democrática do processo da avaliação, onde os interesses de todas as partes envolvidas devem ser negociados e definidos para que a avaliação possa atingir os objetivos propostos. A segunda questão se refere aos elementos resultantes de uma avaliação de programas. O autor destaca que a divulgação dos resultados precisa considerar todas as partes envolvidas, selecionando informações relevantes a cada uma, inclusive por meio de distintos relatórios, caso seja necessário.

Para explorar melhor as questões que envolvem a qualidade da avaliação foi feito um aprofundamento teórico no que são definidos como padrões de qualidade de uma avaliação, pelo *Joint Committee Standards for Educational Evaluation* (JCSEE). Posteriormente, foram abordadas as questões relativas à validade e a confiabilidade das avaliações.

2.3.1. OS PADRÕES DE QUALIDADE DA AVALIAÇÃO DO JCSEE

Retornando ao livro *Avaliação Educacional: teoria, planejamento e modelos*, ao apresentar o modelo de avaliação de Daniel L. Stufflebeam, Vianna (2000) apresenta o que são chamados Padrões de qualidade da Avaliação, definidos pelo *Joint Committee on Standards for Educational Evaluation* (JCSEE), comitê formado por renomados pesquisadores e instituições norte-americanas ligadas à avaliação

educacional e que, à época do lançamento da primeira edição dos *Standards*²⁰, era dirigido por Stufflebeam.

Vianna (2000) e Rodrigues (2009) destacam o esforço coletivo para que, contando com a participação de um grande grupo de pesquisadores, todos chegassem a um consenso do que seriam alguns padrões de qualidade de uma avaliação. Um reflexo das dificuldades encontradas foi a duração do projeto, de 1975 a 1981, quando da publicação de sua primeira edição.

Rodrigues (2009) informa que, de 1981 até o ano de 2009, houve a publicação de novas edições dos *Standards*, mais específicas quanto ao seu público. O *website* do Comitê (JOINT, 2011) apresenta a relação dos *Standards* das três distintas edições, sendo que a publicação dos *Standards* de Programas teve uma nova edição, a terceira, publicada em 2011. O Quadro 1 apresenta as publicações dos *Standards* disponíveis e seus anos de publicação.

Título	Descrição	1ª edição	Reedições
<i>Standards for evaluations of educational programs, projects and materials</i> ²¹	<i>Standards</i> voltados para a Avaliação de Programas	1981	1994 (2ª ed.) 2011 (3ª ed.)
<i>The Personnel Evaluation Standards</i>	<i>Standards</i> de avaliação do pessoal da educação	1988	2009 (2ª ed.)
<i>The Student Evaluation Standards</i>	<i>Standards</i> para as avaliações de alunos	2003	

Quadro 1. Publicações dos *Standards* do JCSEE
Fonte dos dados: VIANNA, 2000; RODRIGUES, 2009; JOINT, 2011

Rodrigues (2009) mostra que, para as três edições distintas, os *Standards* são agrupados em quatro categorias, que revelam os principais atributos de uma avaliação de qualidade: 1) Adequação Ética; 2) Utilidade; 3) Exequibilidade (viabilidade); e 4) Rigor/Exatidão (precisão).

A terceira edição do *Standards* de Programas, disponível no *website* do JCSEE (JOINT, 2011) apresenta uma nova categoria – a quinta – denominada *Evaluation Accountability Standards* que, na verdade, reúne três *Standards* que

²⁰ JOINT COMMITTEE ON STANDARDS FOR EDUCATIONAL EVALUATION. *Standards for evaluations of educational programs, projects and materials*. New York: McGraw-Hill, 1981.

²¹ A partir da edição de 1994, o título passou a ser *The Program Evaluation Standards*.

anteriormente pertenciam ao grupo Rigor/Exatidão e que estão relacionados às questões de documentação e de meta-avaliação. A partir do quadro apresentado por Rodrigues (2009, p. 117) e da nova organização dos *Standards* de Programas (JOINT, 2011), pode-se montar o Quadro 2, com a relação de *Standards* agrupados em suas categorias.

<i>Standards</i> de Avaliação de Estudantes	<i>Standards</i> de Avaliação do Pessoal	<i>Standards</i> de Avaliação de Programas
<i>Standards</i> de Adequação Ética		
Serviço aos estudantes	Orientação de Serviço	Orientação responsiva e inclusiva
Políticas e Procedimentos Adequados	Políticas e Procedimentos Adequados	Acordo Formal
Acesso à informação avaliativa	Acesso à informação avaliativa	Transparência na divulgação
Tratamento dos estudantes	Interação com os avaliados	
Direito dos estudantes	Viabilidade legal	Direitos humanos e respeito
Avaliação equilibrada	Avaliação abrangente	Avaliação clara e justa
Conflito de interesses	Conflito de Interesses	Conflito de Interesses
		Responsabilidade fiscal
<i>Standards</i> de Utilidade		
Orientação construtiva	Orientação construtiva	Propósitos negociados
Usos e utilizadores definidos	Usos definidos	Atenção aos <i>Stakeholders</i>
Alcance da informação		Relevância da informação
Competência do avaliador	Competência do avaliador	Credibilidade do avaliador
Valores explícitos	Crítérios explícitos	Valores explícitos
Relatórios eficazes	Relatórios funcionais	Relatórios apropriados e adequados
		Processos e produtos significativos
<i>Follow-up</i> (acompanhamento)	<i>Follow-up</i> e desenvolvimento profissional	Preocupação com as influências e consequências
<i>Standards</i> de Viabilidade		
		Gerenciamento de projeto
Orientação prática	Procedimentos práticos	Procedimentos práticos
Viabilidade política	Viabilidade política	Viabilidade contextual
Apoios à avaliação	Viabilidade fiscal	Uso de recursos
<i>Standards</i> de Rigor/Exatidão		
Orientação para a validade	Pareceres válidos	Informação válida
Expectativas definidas para os estudantes	Expectativas definidas	

Standards de Avaliação de Estudantes	Standards de Avaliação do Pessoal	Standards de Avaliação de Programas
Análise do contexto	Análise do contexto	
Procedimentos Documentados	Propósitos e procedimentos documentados	Descrição de programas e contextos documentados
Informação defensável	Informação defensável	
Informação confiável	Informação confiável	Informação confiável
Identificação e Gestão de tendências/deturpações	Identificação e gestão de tendências/deturpações	Comunicação e relatórios
Processamento da Informação e Controle da qualidade	Controle Sistemático dos dados	Projeto e análise sólidos
Análise da informação	Análise da informação	Gerenciamento da informação
Conclusões justificadas	Conclusões justificadas	Conclusões e decisões justificadas
		Racionalidade da avaliação explícita
Meta-avaliação	Meta-avaliação	
Standards de Responsabilização		
		Documentação da avaliação
		Meta-avaliação interna
		Meta-avaliação externa

Quadro 2. Os *Standards* do JCSEE por categoria
Fonte: Adpatado de Rodrigues (2009)

Nota: conforme o que foi proposto por Rodrigues (2009), não foi seguida a sequência numérica dos *Standards* para que pudesse ser apresentada uma relação, mesma linha, entre *Standards* iguais ou próximos nas distintas publicações.

As categorias representam os principais atributos de uma avaliação (VIANNA, 2000), mas isso não significa que não haja uma relação entre os *Standards* de categorias diferentes (RODRIGUES, 2009).

Na categoria Adequação Ética, estão reunidos os *Standards* que tratam das questões legais, éticas e do bem estar de todos os envolvidos, buscando garantir as necessidades educativas dos estudantes e da sociedade (RODRIGUES, 2009; VIANNA, 2000).

A categoria Utilidade agrupa os *Standards* relativos ao uso dos resultados pelos seus utilizadores (*stakeholders*), na medida em que a avaliação gera respostas aos questionamentos solicitados e define critérios até para o acompanhamento (follow-up) das ações propostas (RODRIGUES, 2009; VIANNA, 2000).

Os *Standards* de Viabilidade (exequibilidade), chamados de praticabilidade por Vianna (2000), reúnem os cuidados que se devem ter para que a avaliação seja planejada e disponha de recursos que garantam a realização do que foi proposto no seu projeto, o que inclui que a avaliação seja eficaz e eficiente na utilização desses recursos (RODRIGUES, 2009; VIANNA, 2000).

A precisão é o aglutinador dos *Standards* da categoria de Rigor/Exatidão. Aqui são abordadas as questões mais técnicas, a fim de garantir que a avaliação produza resultados válidos, confiáveis e, assim, representativos do objeto em estudo (RODRIGUES, 2009; VIANNA, 2000).

Como já foi dito, os padrões de qualidade foram criados por um Comitê formado por instituições norte-americanas. Entretanto, esses padrões têm sido adotados ou adaptados por pesquisadores em vários países (RODRIGUES, 2009).

No Brasil, Dannemann, Firme e Letichevsky (2005) relatam sobre a criação de uma associação nacional voltada para a avaliação, denominada Rede Nacional de Avaliação²². As autoras defendem que o bom conhecimento e aplicação dos *Standards* e suas categorias são o mínimo que se deve ter para a evolução da área, sendo papel da instituição, a Rede Nacional de Avaliação, a divulgação dos mesmos em todos os níveis de público de interesse, principalmente as instituições de ensino.

Assim, pode-se concluir que as autoras Dannemann, Firme e Letichevsky (2005) defendem o uso pleno dos *Standards* norte-americanos sem qualquer tipo de adaptação à realidade brasileira.

Em um artigo anterior, Firme e Letichevsky (2002) descrevem a realização de uma meta-avaliação numa instituição universitária. Segundo as autoras, (p.289) “a meta-avaliação verificou até que ponto a autoavaliação em questão correspondeu às quatro categorias de *Standards* ou padrões de programas avaliativos”. Desse modo, confirma-se a opção em aplicar os *Standards* norte-americanos para a verificação da qualidade de uma avaliação.

²² As pesquisas eletrônicas sobre o termo “Rede Nacional de Avaliação” retornaram praticamente somente o artigo referido (DANNEMANN, FIRME e LETICHEVSKY, 2005). Como o endereço eletrônico citado em algumas referências, www.avaliabrasil.org.br, está fora do ar, tem-se a impressão de que a associação não está ativa.

Em um artigo mais recente, Ribeiro e Gusmão (2010) fazem um estudo sobre o uso dos Indicadores da Qualidade na Educação (INDICADORES, 2007). Para tanto, um dos critérios usados pelos autores foram as categorias Rigor/Precisão, Utilidade e Factabilidade/Viabilidade, seguindo a descrição apresentada por Vianna (2000). Na análise apresentada, Ribeiro e Gusmão (2010) não justificam por que eles não usaram a categoria Propriedade/Adequação Ética. Os autores usaram ainda alguns outros critérios, como a Simplicidade, apoiados por Fernandes²³ (2007 apud RIBEIRO; GUSMÃO, 2010, p. 829).

Conforme exposto, não foi encontrada uma padronização nacional para a aferição da qualidade da avaliação. Os autores brasileiros pesquisados fazem referencia aos *Standards* definidos pela *Joint Committee Standards for Educational Evaluation* em seus trabalhos. Rodrigues (2009) afirma que esses padrões têm sido usados por pesquisadores de vários países.

Desse modo, a qualidade da avaliação do SAEM foi analisada segundo os padrões definidos para a Avaliação de Programas, seguindo os critérios de Adequação Ética, Utilidade, Viabilidade, Rigor/Precisão e Responsabilização.

Dentro dos padrões de qualidade da categoria Rigor/Precisão, há a presença dos requisitos sobre a validade e a confiabilidade da avaliação. Assim, faz-se necessário apresentar um aprofundamento teórico a respeito desses dois conceitos.

2.3.2. VALIDADE E CONFIABILIDADE DE UMA AVALIAÇÃO

Uma das características básicas que se espera de um sistema de avaliação é que o mesmo seja válido e fidedigno, ou seja, confiável (VIANNA, 1978; PASQUALI, 2003; MARTINS, 2006).

Em termos de validade, um teste – nome de uma avaliação na Psicometria – é válido quando este de fato mede o que se deve medir (PASQUALI, 2003; MARTINS, 2006). Martins (2006) ainda destaca que a validade de um instrumento

²³ FERNANDES, D. *Avaliação de programas e projetos*. Lisboa, [s/d]. Mimeo. [Texto de aula oferecida na condição de convidado, Feusp, maio 2009]

de medida nunca é absoluta, uma vez que ela dependerá do uso que faz do instrumento. O autor usa como exemplo a medição do tempo. Pode-se usar a posição do sol como instrumento de medida para se determinar a hora do almoço. Entretanto, esse mesmo instrumento não retornará medidas válidas para a medição de uma partida de futebol, argumenta Martins (2006).

Segundo Pasquali (2003), a medida da validade de um teste é obtida por três processos distintos: a) validade de construto, b) validade de critério e c) validade de conteúdo. Sampieri, collado e Lucio²⁴ (1996 apud MARTINS, 2006) definem como validade total a soma das validades de construto, de critério e conteúdo.

Segundo Pasquali (2003, p. 164),

A validade de construto ou de conceito é considerada a forma mais fundamental de validade [...] dado que ela constitui a maneira direta de verificar a hipótese da legitimidade da representação comportamental do traço latente [...]

O autor relaciona alguns processos que podem ser aplicados para se obter a validade de construto: a) erro de estimação, b) análise da representação, c) análise por hipótese e d) a curva de informação da TRI. A técnica de validade de construto pelo cálculo do erro de estimação usa o escore de um teste paralelo para, a partir do cálculo do erro de estimação entre os escores, validar o próprio teste.

A validade de construto por análise da representação pode ser obtida a partir de duas técnicas: a análise de consistência interna do teste e a análise fatorial. A análise de consistência interna implica no cálculo das correlações de cada questão individualmente com o restante do teste, não constituindo prova definitiva da validade de construto, conforme orienta Pasquali (2003). Já a análise fatorial tem como lógica verificar quantos construtos comuns são necessários para explicar as covariâncias dos itens, sendo, assim, mais indicada que a análise de consistência interna.

A análise por hipótese pode ser obtida a partir de quatro técnicas normalmente utilizadas, segundo Pasquali (2003): a) validação convergente-discriminante, b) idade, quando o traço latente busca discriminar grupos de idades diferentes, c) outros testes do mesmo construto e d) a experimentação. Desses o

²⁴ SAMPIERI, R. H.; COLLADO, C. F.; LUCIO, P. B. **Metodologia de La Investigacion**. México: MCGRAW HILL, 1996.

autor destaca o uso de outros testes a partir da determinação da correlação entre eles, e da experimentação, que consiste em verificar se o teste discrimina grupos-critério definidos em termos do traço objeto da medida.

Finalmente, em relação à validade de construto, Pasquali (2003) defende que seja usada mais de uma das técnicas citadas para demonstrar a validade do teste “dado que a convergência de resultados das várias técnicas constitui garantia para a validade do instrumento” (p. 185).

A validade de critério é definida por Kaplan²⁵ (1975 apud MARTINS, 2006), como a comparação do instrumento de medida com algum critério externo. Para Pasquali (2003), a validade de critério é a medida do grau de eficácia do teste em prever o desempenho de um sujeito, no caso, um aluno. Um dos critérios que pode ser usado é a existência de um teste comprovadamente validado, contra o qual se pode validar o novo teste. O autor define que, caso seja usado como critério de validade outro teste, a correlação entre eles deve ser, no mínimo, 0,75.

Há dois tipos de validade de critério: a) validade preditiva e b) validade concorrente, que se diferenciam em relação ao tempo entre o teste e a definição do critério. Martins (2006) exemplifica a validade preditiva com o caso de uma pesquisa eleitoral realizada durante a campanha, pode ser validada pelos posteriores resultados da eleição (critério). A validade concorrente, como o nome já diz, ocorre quando a coleta dos dados do teste e do critério são muito próximas (de tempo) (PASQUALI, 2003).

Finalmente, ainda segundo Pasquali (2003), a validade de conteúdo é garantida pela técnica de construção do teste, de forma que estejam definidos os conteúdos, os objetivos e as proporções de representação dos conteúdos no teste. Martins (2006) cita como exemplo a construção de uma prova de operações aritméticas. Ela não terá validade de conteúdo, caso ela possua somente questões de adição e não aborde questões de subtração, multiplicação e divisão. Pasquali (2003) descreve os passos para a validação de conteúdo como:

- 1) Definição do domínio cognitivo: definir os objetivos que se quer avaliar. Em termos educacionais, significa a construção dos descritores da avaliação.

²⁵ KAPLAN, A. **A conduta na pesquisa**. São Paulo: EDUSP, 1975.

- 2) Definição do universo do conteúdo: selecionar os conteúdos que serão cobrados na avaliação.
- 3) Definição da representatividade de conteúdo: definir a proporção de cada conteúdo cobrado na avaliação.
- 4) Tabela de especificação: tabela na qual serão relacionados os conteúdos com os processos cognitivos definidos
- 5) Construção do teste: elaboração das questões propriamente dita.
- 6) Análise teórica dos itens: verificação da compreensão das tarefas propostas nas questões (análise semântica) e pertinência da questão ao conteúdo selecionado e ao processo cognitivo proposto.
- 7) Análise empírica dos itens: determinação dos índices de dificuldade e de discriminação das questões para uso futuro.

Além das validades de construto, critério e conteúdo, Martins (2006) acrescenta ainda a validade aparente com uma das formas de validade de um instrumento de medida. Considerada a mais simples das validades, a validade aparente avalia se um instrumento, aparentemente, mede a variável em estudo. Assim, essa validade é um processo subjetivo, realizado por um juiz ou um grupo de juízes, afirma o autor.

Vianna (2003), ao escrever sobre o Sistema Nacional de Avaliação da Educação Básica, declara que uma das características dos instrumentos de medida é a validade consequential, que “se refere ao impacto da avaliação sobre o sistema, determinando mudanças de pensamento, gerando novos comportamentos, formando novas atitudes e promovendo novas ações” (VIANNA, 2003, p. 54). O autor complementa que a validade consequential mede o diferencial da avaliação para a comunidade.

Além da validade e de suas sub-divisões, uma avaliação de qualidade tem também como característica a confiabilidade. Entende-se como confiabilidade, também chamada de fidedignidade ou precisão, a capacidade do teste ser o mais preciso possível do resultado real que se deseja medir (PASQUALI, 2003). Martins (2006) define confiabilidade como a capacidade de, ao se aplicar o mesmo teste para os mesmos sujeitos em ocasiões diferentes, obter resultados iguais, o que Pasquali (2003) define como medir sem erros.

A análise do grau de confiabilidade de um instrumento é determinada pelo **coeficiente de fidedignidade** do teste, que é a razão entre variância verdadeira do teste e a variância total do teste (PASQUALI, 2003).

Há a necessidade de duas coletas de dados para que se calcule o coeficiente de fidedignidade de um teste. Para tanto, Pasquali (2003) define três delineamentos possíveis para as duas coletas:

- 1) Uma amostra de sujeitos, um mesmo teste e uma única ocasião: a questão das duas coletas é superada dividindo o teste em duas partes, chamado de duas metades, cada uma com a metade dos itens do teste todo.
- 2) Uma amostra de sujeitos, dois testes e uma única ocasião: aplicam-se dois testes paralelos (ambos com validade para o mesmo construto) e analisa-se a correlação entre as distribuições dos dois testes
- 3) Uma amostra de sujeitos, um mesmo teste e duas ocasiões: chamada de teste-reteste, o procedimento consiste em aplicar o mesmo teste aos mesmos alunos, com certo tempo entre as aplicações.

Além dos delineamentos, Pasquali (2003) argumenta que há basicamente duas técnicas estatísticas para o cálculo do coeficiente de fidedignidade: a correlação simples e a(s) técnica(s) alfa.

Coletados os dados a partir de um dos três procedimentos anteriores, o coeficiente de fidedignidade será igual à correlação entre os dois testes. Quanto mais próximo de 1 for o coeficiente, mais confiáveis são os testes (Martins, 2006). No caso da técnica das duas metades, Pasquali (2003) alerta para a necessidade de se aplicar a correção de Spearman-Brown, que leva em conta a divisão do número de itens à metade. Pasquali (2003) argumenta que coeficientes de fidedignidade aceitáveis devem ter valores superiores a 0,70, preferencialmente superiores a 0,8, e normalmente esperados maiores que 0,9. Vianna (1978) apresenta o cálculo do coeficiente de fidedignidade esperado, que depende no número de itens que compõe o teste e o número de alternativas dos itens.

Outra técnica para o cálculo do coeficiente de fidedignidade é através da análise da consistência interna do teste a partir da consistência interna dos seus

itens. O cálculo mais comum desse tipo de análise é pelo coeficiente alfa de Cronbach (PASQUALI, 2003; MARTINS, 2006).

Após discorrer sobre os fatores que influenciam na qualidade de um sistema de avaliação, é importante que se faça um estudo do modelo teórico adotado pela Prova Brasil/SAEB, uma vez que esse modelo de avaliação será usado como referência de comparação com o SAEM

2.4. O SISTEMA DE AVALIAÇÃO DA EDUCAÇÃO BÁSICA

A Prova Brasil e o SAEB são dois exames complementares que compõem o Sistema de Avaliação da Educação Básica desde 2005 (INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA, 2010). A Portaria Ministerial nº 931, de 21 de março de 2005 (BRASIL, 2005), instituiu o Sistema de Avaliação da Educação Básica, SAEB, com dois exames. O modelo de avaliação já existente passou a se chamar ANEB (Avaliação Nacional da Educação Básica). O novo exame foi denominado ANRESC (Avaliação Nacional do Rendimento Escolar), mais conhecido por Prova Brasil (INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA, 2010).

Entretanto, as discussões iniciais para a implantação de um Sistema de avaliação em larga escala no Brasil remontam à década de 80 (BRASIL, 2008). Bonamino e Franco (1999) relatam que esse período caracterizou-se por diversas reformas educacionais com mudanças profundas relativas ao financiamento, ao currículo e também à avaliação educacional tanto no Brasil como em vários países da América Latina, principalmente devido aos processos de redemocratização pelos quais esses países passaram.

Bonamino (2002) mostra também que o período da década de 80 foi marcado pelo crescimento vegetativo da população e pela aceleração dos processos de urbanização. Esses fatores contribuíram para o aumento da demanda por serviços públicos básicos como a saúde e a educação, o que pressionou a municipalização desses serviços.

A década de 80 acompanha também a entrada das agências de financiamento, principalmente o Banco Mundial, como atores dos projetos de avaliação educacional (BRASIL, 2008; BONAMINO e FRANCO, 1999; BONAMINO, 2002). No Brasil, a primeira experiência de avaliação em larga escala, o EDURURAL²⁶, teve o apoio financeiro do Banco Mundial. A partir dessa experiência, o Ministério da Educação institui o SAEP (Sistema de Avaliação da Educação Primária), que teve uma aplicação piloto, em 1988, em alunos das séries iniciais dos estados do Paraná e do Rio Grande do Norte (BRASIL, 2008; BONAMINO, 2002).

Com a Constituição de 1988 (BRASIL, 1988), o SAEP passa a se chamar Sistema Nacional de Avaliação da Escola Básica²⁷, o SAEB. Bonamino e Franco (1999) relatam que problemas de ordem financeira fizeram com que o primeiro ciclo do SAEB ocorresse somente em 1990.

Pestana (1998) relata que o primeiro ciclo contou com a adesão de 23 estados. A avaliação, a mesma para todo o país em cada série, foi aplicada a uma amostra de alunos da 1ª, 3ª, 5ª e 7ª séries do Ensino Fundamental. As disciplinas avaliadas foram Língua Portuguesa, Matemática e Ciências (BONAMINO; FRANCO, 1999). A definição dos conteúdos mínimos comuns foi feita a partir de consulta aos currículos estaduais e posterior discussão entre as respectivas equipes estaduais para a escolha final (PESTANA, 1998).

O segundo ciclo, realizado em 1993, manteve as mesmas características da coleta de 1990 (PESTANA, 1998; BONAMINO; FRANCO, 1999) e, desde então, vem ocorrendo, ininterruptamente, a cada dois anos (BRASIL, 2008). A partir de 1995, várias alterações foram implementadas no modelo de avaliação do SAEB. Assim, é importante que se apresente uma síntese das principais alterações ocorridas desde então. Para 2011, está prevista a aplicação do décimo primeiro ciclo do SAEB (BRASIL, 2011).

²⁶ EDURURAL – Programa de Expansão e Melhoria do Ensino no Meio Rural do Nordeste Brasileiro. A pesquisa envolveu alunos da 2ª e da 4ª séries do Ensino Fundamental, em 60 escolas rurais dos estados do Ceará, Piauí e Pernambuco, durante os anos de 1981, 1983 e 1985 (BONAMINO, 2002).

²⁷ Não foi encontrada dentre os autores pesquisados, a partir de quando o SAEB deixa seu nome original – Sistema Nacional de Avaliação da Escola Básica – para o nome atual – Sistema Nacional de Avaliação da Educação Básica.

2.4.1. A PROVA BRASIL

A principal alteração promovida no SAEB foi a inclusão de uma nova avaliação, a ANRESC, que passou a ser conhecida por Prova Brasil, em 2005. Fernandes e Gremaud (2009) afirmam que, devido à descentralização e à municipalização dos sistemas de ensino no período pós-redemocratização, havia a necessidade de “aumentar o conteúdo informacional da avaliação diagnóstica e permitir que todos os sistemas educacionais públicos se ‘enxergassem’ na avaliação” (p. 226). Os autores destacam ainda a possibilidade de agregar as noções de *accountability* à avaliação.

Assim, a principal diferença entre as duas avaliações – Prova Brasil e SAEB – está na abrangência das mesmas. Enquanto a avaliação do SAEB contempla amostras de alunos de toda a rede de ensino: pública e privada, urbana e rural, a Prova Brasil é censitária, mas, inicialmente, limitada às redes públicas urbanas (BRASIL, 2008). Nas seções seguintes as demais diferenças serão apresentadas.

2.4.2. OBJETIVOS DO SAEB

Bonamino e Franco (1999) apresentam uma análise da evolução dos objetivos do SAEB nos seus quatro primeiros ciclos (1990, 1993, 1995 e 1997), a partir dos documentos oficiais disponíveis. Os autores analisam que para os primeiros ciclos havia uma proposta descentralizada e participativa que foi sendo substituída por uma proposta mais centralizadora no âmbito do Ministério da Educação. Complementam os autores (BONAMINO; FRANCO, 1999, p. 112):

Em contraste com os objetivos dos dois primeiros ciclos e suas ênfases em aspectos processuais, os objetivos dos demais ciclos caracterizam-se pela ênfase em produção de resultados que possam contribuir para a monitoração da situação educacional brasileira e para subsidiar os formuladores de políticas públicas.

Em 2005, a portaria nº931 (BRASIL, 2005, p.1), a qual criou a Prova Brasil, define que os objetivos do SAEB seriam mantidos na ANEB, que o substituiu.

§ 1º A ANEB manterá os objetivos, características e procedimentos da avaliação da educação básica efetuada até agora pelo SAEB realizado por meio de amostras da população, quais sejam:

a) a ANEB tem como objetivo principal avaliar a qualidade, equidade e a eficiência da educação brasileira; [...]

Como se pôde notar, foi mantido o objetivo principal da avaliação de 1997, como exposto por Bonamino e Franco (1999).

A portaria nº 47, de maio de 2007, traz os objetivos das duas avaliações – ANEB e Prova Brasil separadamente, mantendo somente para a ANEB a responsabilidade de produzir informações sobre o desempenho dos alunos e de avaliar a qualidade, a equidade e a eficiência da educação brasileira (BRASIL, 2007a).

A partir de 2009, os objetivos passam a ser apresentados de forma comum para as duas avaliações, sendo ampliados para a edição de 2011, previstas para serem aplicadas no período de 07 a 18 de novembro de 2011 (BRASIL, 2009, 2011).

O Quadro 3 apresenta os objetivos gerais e específicos do Sistema de Avaliação da Educação Básica das duas avaliações para 2011.

De um modo geral, os objetivos estão cada vez mais diversificados, cobrindo desde o que pode ser considerado como o objetivo principal, que é o de “Oferecer subsídios à formulação [...] de políticas públicas [...]”, passando por objetivos de *accountability*, ao “Proporcionar [...] à sociedade, uma visão dos resultados [...] e das condições em que são desenvolvidos.”, até objetivos secundários – visto que são consequências independentes do processo – como o de “Desenvolver competência técnica e científica na área de avaliação educacional, ativando o intercâmbio entre instituições educacionais de ensino e pesquisa”.

Vale ressaltar também que o objetivo de “Fornecer dados para o cálculo do IDEB” foi inserido somente na edição de 2009, mesmo com o IDEB sendo lançado em 2007 (FERNANDES, 2007).

Objetivos	Descrição
Gerais	<p>I. Oferecer subsídios à formulação, reformulação e monitoramento de políticas públicas e programas de intervenção ajustados às necessidades diagnosticadas;</p> <p>II. Identificar problemas e diferenças regionais na Educação Básica;</p> <p>III. Produzir informações sobre os fatores do contexto socioeconômico, cultural e escolar que influenciam o desempenho dos estudantes;</p> <p>IV. Proporcionar aos agentes educacionais e à sociedade uma visão dos resultados dos processos de ensino e aprendizagem e das condições em que são desenvolvidos;</p> <p>V. Desenvolver competência técnica e científica na área de avaliação educacional, ativando o intercâmbio entre instituições educacionais de ensino e pesquisa;</p> <p>VI. Aplicar testes de Matemática e de Língua Portuguesa, com foco em resolução de problemas e em leitura, respectivamente, definidos nas Matrizes de Referência do Sistema de Avaliação da Educação Básica;</p> <p>VII. Produzir informações sobre o desempenho dos estudantes, assim como sobre as condições intra e extraescolares que incidem sobre o processo de ensino e aprendizagem, no âmbito das redes de ensino e unidades escolares;</p> <p>VIII. Fornecer dados para cálculo do IDEB;</p> <p>IX. Avaliar a qualidade, a equidade e a eficiência dos sistemas e redes de ensino brasileiras;</p> <p>X. Manter a construção de séries históricas, permitindo comparabilidade entre anos e entre séries escolares.</p>
Prova Brasil	<p>I. Aplicar instrumentos (provas e questionários) nas escolas da rede pública de ensino das zonas urbana e rural que possuam pelo menos 20 estudantes matriculados nos 5º e 9º anos do ensino fundamental regular que estejam organizadas no regime de 09 anos, e nas 4ª e 8ª séries do ensino fundamental regular de 08 anos;</p> <p>II. Fornecer informações sobre as unidades escolares, que sejam úteis aos gestores da rede a qual pertençam às escolas avaliadas.</p>
ANEB	<p>I. Aplicar instrumentos (provas e questionários) em uma amostra representativa de estudantes de 5º e 9º anos (4ª e 8ª) séries do ensino fundamental regular e de 3ª série do ensino médio regular das escolas das redes pública e privada, localizadas na zona urbana, distribuídas nas 27 unidades da federação. No ensino fundamental da rede pública, esta amostra complementará as informações da ANRESC 2011.</p>

Quadro 3. Objetivos do SAEB 2011

Fonte dos dados: BRASIL, 2011

2.4.3. ABRANGÊNCIA: SÉRIES/ANOS, DISCIPLINAS E REDES

Em relação às séries/anos avaliados, as duas primeiras edições do SAEB envolveram alunos da 1ª, 3ª, 5ª e 7ª séries do Ensino Fundamental (BONAMINO; FRANCO, 1999). A partir de 1995, passou-se a avaliar somente os alunos das séries “conclusivas” cada ciclo: 4ª e 8ª série (5º e 9º ano) do Ensino Fundamental e a 3ª série do Ensino Médio.

De 1995 até a avaliação prevista para 2011, as séries/anos envolvidos não têm se mantido constantes. A Prova Brasil envolve as mesmas séries do Ensino Fundamental, mas não avalia os alunos do Ensino Médio (BRASIL, 2008; FERNANDES; GREMAUD, 2009).

A rede privada de educação básica passa a ser avaliada pelo SAEB também a partir de 1995 (BONAMINO; FRANCO, 1999). Entretanto, a Prova Brasil avalia somente as escolas da rede pública de ensino (BRASIL, 2008).

Em relação ao zoneamento, o SAEB sempre avaliou alunos de escolas da zona urbana e rural (BRASIL 2008, 2009, 2011). A Prova Brasil, nas suas duas primeiras edições, avaliou somente alunos de escolas da zona urbana. Desde 2009, a avaliação abrange escolas da zona urbana e rural²⁸ (BRASIL 2007a, 2009, 2011).

O foco da avaliação cognitiva do SAEB é o domínio da leitura em Língua Portuguesa e o da resolução de problemas em Matemática (BRASIL, 2007a, 2009, 2011; INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA, 2010). Todavia, Bonamino e Franco (1999) relatam que as primeiras edições abrangiam também a disciplina de Ciências. Houve avaliação dos conteúdos de História e Geografia em 1999 (INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA, 2010). O Quadro 4 apresenta as disciplinas envolvidas em cada um dos ciclos do SAEB.

²⁸ Em todos os casos, são consideradas somente escolas com mais de 20 alunos matriculados em cada uma das séries/anos avaliados. Em 2005, esse número era de 30 alunos (FERNANDES; GREMAUD, 2009)

Ano	Série/ano	Disciplinas
1990	1ª e 3ª série do EF	Língua Portuguesa, Matemática e Ciências
	5ª e 7ª série do EF	Língua Portuguesa, Matemática, Ciências e Redação
1993	1ª e 3ª série do EF	Língua Portuguesa, Matemática e Ciências
	5ª e 7ª série do EF	Língua Portuguesa, Matemática, Ciências e Redação
1995	4ª e 8ª série do EF	Língua Portuguesa e Matemática
	3ª série do EM	Língua Portuguesa e Matemática
1997	4ª e 8ª série do EF	Língua Portuguesa, Matemática e Ciências
	3ª série do EM	Língua Portuguesa, Matemática e Ciências (Física, Química e Biologia)
1999	4ª e 8ª série do EF	Língua Portuguesa, Matemática e Ciências
	3ª série do EM	Língua Portuguesa, Matemática, Ciências (Física, Química e Biologia), História e Geografia
2001 a 2011	4ª e 8ª série (5º e 9º ano) do EF e 3ª série do EM	Língua Portuguesa e Matemática

Quadro 4. Disciplinas avaliadas em cada um dos ciclos do SAEB

Fonte dos dados: INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA, 2010; BONAMINO; FRANCO, 1999; BRASIL, 2008, 2011

Uma análise nas disciplinas mostra que a opção de envolver cada vez mais disciplinas foi abandonada a partir de 2001, permanecendo inalterada desde então. Bonamino e Franco (1999) já criticavam que a escolha de somente Língua Portuguesa, Matemática e Ciências indicava “a adoção de uma visão estreita da educação” (p. 112), incorrendo no problema do reducionismo curricular.

2.4.4. MATRIZ DE REFERÊNCIA

O ciclo do SAEB de 1997 foi marcado pela introdução das Matrizes Curriculares de Referência na avaliação. O INEP contratou consultores e especialistas educacionais para a elaboração das Matrizes (BONAMINO, 2002), que são a descrição das competências e habilidades as quais os alunos de cada série

deveriam dominar (BRASIL, 2008). Bonamino (2002, p. 153) relata a dificuldade e a ousadia dos gestores do INEP à época para modificar “as orientações seguidas na formulação dos instrumentos de avaliação, visando avaliar níveis cognitivos dos alunos”.

Para Pestana (1998, p. 70), as Matrizes “devem cumprir dois objetivos: dar transparência à avaliação e diminuir um dos maiores riscos desta, que é o da redução daquilo que é ensinado ao que é cobrado”.

Em relação à formação da equipe que trabalhou na elaboração das Matrizes, enquanto Pestana (1998) comenta sobre o caráter democrático do processo, sendo que as propostas foram enviadas às secretarias estaduais de educação para análise e envio de contribuições, Bonamino (2002) destaca que a equipe contratada tinha relação com a equipe responsável pela elaboração dos Parâmetros Curriculares Nacionais (PCN), lançados também em 1997, o que trouxe um caráter construtivista às Matrizes.

2.4.5. INSTRUMENTOS CONTEXTUAIS

Apesar da mudança de foco durante todos esses anos, o SAEB, desde o seu primeiro ciclo, em 1990, coletou informações contextuais através da aplicação de instrumentos (questionários) diversificados conforme o respondente. A existência desses instrumentos está relacionada a alguns dos objetivos do SAEB: “Produzir informações sobre os fatores do contexto socioeconômico, cultural e escolar que influenciam o desempenho dos estudantes” e “Produzir informações sobre [...] as condições intra e extraescolares que incidem sobre o processo de ensino e aprendizagem” (BRASIL, 2011).

Na primeira avaliação, informam Bonamino e Franco (1999), buscou-se coletar informações sobre três eixos: 1) a universalização do ensino; 2) valorização do magistério; e 3) democratização da gestão. Foram criados questionários para alunos, professores e diretores escolares.

Somente no terceiro ciclo do SAEB é que foram introduzidas questões relativas às condições socioeconômicas, culturais e sobre hábitos de estudos dos

alunos, voltando os olhos para os fatores extraescolares (BONAMINO; FRANCO, 1999).

Sobre a importância dos instrumentos contextuais na avaliação do SAEB, Pestana (1998, p. 68) afirma:

No SAEB, parte-se do princípio de que o desempenho do aluno sofre múltiplos condicionamentos, os quais devem ser considerados pela avaliação. Para tanto, deve-se somar ao desempenho do aluno, pelo menos, mais três dimensões, quais sejam: a determinação de contextos em que ocorrem o ensino e a aprendizagem; a identificação de processos de ensino e aprendizagem e, por fim, o dimensionamento dos insumos utilizados.

A autora apresenta essas dimensões, associadas a seus indicadores e instrumentos, através do Quadro 5 (PESTANA, 1998, p. 68).

ENFOQUE	DIMENSÃO	INDICADOR/VARIÁVEL	INSTRUMENTO
Acesso ao Ensino Básico	Atendimento	<ul style="list-style-type: none"> • Taxas de Acesso • Taxas de Escolarização 	Questionário do CENSO ESCOLAR ²⁹
	Eficiência	<ul style="list-style-type: none"> • Taxas de produtividade • Taxas de transição • Taxas de eficiência interna 	
Qualidade, Eficiência e Equidade do Sistema de Ensino Básico	Produto	<ul style="list-style-type: none"> • Desempenho do aluno, em termos de: <ul style="list-style-type: none"> • Aprendizagem de conteúdos • Desenvolvimento de habilidades e competências 	Questões das provas
	Contexto	<ul style="list-style-type: none"> • Nível socioeconômico dos alunos • Hábitos de estudo dos alunos • Perfil e condições de trabalho dos docentes e dos diretores • Tipo de escola • Grau de autonomia da escola • Matriz organizacional da escola 	Questionários para: <ul style="list-style-type: none"> • alunos • professores • diretores
	Processo	<ul style="list-style-type: none"> • Planejamento do ensino e da escola • Projeto Político Pedagógico • Relação entre conteúdos propostos/ensinados e conteúdos ensinados/aprendidos • Utilização do tempo pedagógico • Estratégias e técnicas de ensino utilizadas 	
	Insumos	<ul style="list-style-type: none"> • Infraestrutura (adequação, manutenção, conservação) • Espaço físico e instalações • Equipamentos • Recursos e materiais didáticos 	Questionário sobre condições da escola

Quadro 5. Sistema Nacional de Avaliação da Educação Básica

Fonte: PESTANA, 1998

²⁹ Atualmente a coleta dessas informações é feita pelo programa EDUCACENSO, para o qual as escolas enviam as informações sobre seus alunos via Internet. Mais informações consultar: <http://www2.inep.gov.br/web/educacenso>

A visão proporcionada pela leitura do quadro das dimensões do SAEB é de extrema importância para entender que o mesmo não se restringe a uma avaliação de desempenho cognitivo dos alunos.

2.4.6. INSTRUMENTOS COGNITIVOS

A base de todo sistema de avaliação que busca aferir desempenho cognitivo é a prova – também chamada de teste – cujas questões – chamadas de itens e que formam o teste – são respondidas pelos alunos. Assim, a elaboração dos itens tem uma grande influência na qualidade da avaliação.

Na primeira edição do SAEB, em 1990, procurou-se estabelecer, inicialmente, um conteúdo comum dentre os estados participantes para, posteriormente, elaborar as questões (BONAMINO, 2002). Segundo a autora, o mesmo grupo de professores que trabalhou na definição dos conteúdos, também foi responsável pela elaboração das questões.

Para o segundo ciclo, foram analisadas as observações e as críticas recebidas na primeira edição. Bonamino e Franco (1999) apontam que as solicitações se concentravam em: a) medição de outras habilidades, além da memorização; b) vocabulário adequado às diversas regiões; c) enunciados mais claros; e d) cuidado com a qualidade das ilustrações. Os conteúdos também foram revistos, a partir da participação das secretarias de educação. Bonamino e Franco (1999) ainda destacam que os itens passaram por pré-testagem nas escolas de aplicação da Universidade Federal de Minas Gerais (UFMG) e na Universidade Estadual de Londrina.

A partir de 1995, o SAEB passa por várias mudanças, como já foi citado anteriormente, motivado, principalmente, pela troca de governo após as eleições de 1994 (BONAMINO, 2002). Dentre essas mudanças, talvez a principal tenha sido a adoção da Teoria da Reposta ao Item (TRI) para permitir a comparabilidade no tempo, entre avaliações de anos diferentes, e entre as séries (PESTANA, 1998).

Bonamino e Franco (1999) também destacam uma mudança na forma de execução do SAEB, com a terceirização de serviços para instituições privadas. Com relação à elaboração das questões, foi usado um banco de itens da Fundação Cesgranrio³⁰, já calibrado para a TRI, uma vez que a mesma já vinha trabalhando com a metodologia desde 1994 (BONAMINO, 2002).

Além da elaboração, a Fundação Cesgranrio e a Fundação Carlos Chagas³¹ ficaram responsáveis pela seleção da amostra para a pré-testagem, confecção dos cadernos de teste, padronização e elaboração de manuais e vídeo de treinamento do procedimento de aplicação dos instrumentos, distribuição dos materiais às sedes das secretarias estaduais de educação e o processamento e análise dos resultados (BONAMINO; FRANCO, 1999).

Em 1997, quarta edição do SAEB, houve a terceirização das etapas de elaboração, pré-testagem, aplicação das provas e análise dos resultados. Bonamino (1999, p. 152) faz a seguinte análise:

[...] a utilização da TRI, na medida em que tornou o SAEB mais dependente de agências externas que dominam essa técnica, expressa a consolidação da tendência, emergente desde o ciclo anterior, à terceirização de definições importantes a respeito dos sistema de avaliação da educação básica.

Conforme alertado pela autora, o site da Fundação Cesgranrio (FUNDAÇÃO CESGRANRIO, 2010) confirma que a instituição foi a vencedora da licitação para o fornecimento de itens para o Banco Nacional de Itens, mantido pelo INEP, nos anos de 1999, 2001 e 2003, além de participar da execução, sozinha ou na forma de consórcio, desses ciclos.

A criação da Prova Brasil, em 2005, promoveu a elaboração de instrumentos cognitivos diferentes, uma vez que os cadernos de testes do SAEB continham apenas uma das disciplinas (FERNANDES; GREMAUD, 2009). A partir do ciclo de 2007, alterou-se a organização dos cadernos do SAEB para que todos os alunos respondessem a itens de Língua Portuguesa e Matemática, possibilitando a “fusão” dos instrumentos. “Assim, voltamos a ter uma única avaliação (o SAEB), e o que

³⁰ A Fundação Cesgranrio é uma instituição privada que, desde 1973, presta serviços de avaliação, iniciando-se com os Vestibulares de acesso ao Ensino Superior, passando pelos Concursos Públicos e pelas avaliações da Educação Básica. Ver mais informações em www.cesgranrio.org.br.

³¹ Fundada em 1964, a Fundação Carlos Chagas é uma instituição privada sem fins lucrativos, reconhecida como de utilidade pública e dedicada à avaliação e à pesquisa na área de educação. Ver mais em www.fcc.org.br.

entendemos por Prova Brasil constitui-se, na verdade, em estratos censitários do SAEB” concluem Fernandes e Gremaud (2009, p. 226).

Além das questões relativas à adoção da TRI, Pestana (1998) cita também a vantagem da adoção dos Blocos Intercalados Balanceados (BIB), que permitiram a superação da dificuldade da pequena amplitude de conhecimentos avaliados em um teste de 30 questões. Segundo Conde e Laros (2007), antes de 1999, os blocos continham de 10 a 13 itens. A partir de 1999, foram formulados 13 blocos com 13 itens cada. A combinação desses blocos possibilita a montagem de 26 cadernos de testes diferentes. Assim, enquanto cada aluno respondia a somente 39 questões de uma disciplina, a avaliação toda conta com 169 itens diferentes (CONDE; LAROS, 2007).

2.4.7. RESULTADOS DO SAEB

A última característica do SAEB que passou por grandes alterações é a questão dos resultados. Pestana (1998) relata que, também em 1995 e em consequência da adoção da TRI, se iniciou o processo de elaboração e de interpretação de escalas de proficiência. A divulgação dos resultados em escalas de proficiência é que permite a comparação entre diferentes ciclos e séries/anos do SAEB.

Para o caso do SAEB, a escala adotada varia de 0 a 500 pontos, para cada uma das disciplinas (BRASIL, 2008). Assim, espera-se que um aluno da 4ª série (5º ano) do Ensino Fundamental tenha um desempenho mais baixo de que um aluno da 8ª série (9º ano) do Ensino Fundamental, e este, por sua vez, tenha um desempenho mais baixo do que um aluno da 3ª série do Ensino Médio. A Prova Brasil usa a mesma escala de proficiência que o SAEB (FERNANDES; GREMAUD, 2009).

Ainda em relação aos resultados, devido à questão do uso de estratos amostrais, o SAEB só apresenta resultados por unidade federativa, região e Brasil (BRASIL, 2008).

É para atender à demanda dos gestores públicos, dos educadores, dos pesquisadores e da sociedade em geral por informações sobre o ensino nas redes municipais e nas escolas que foi criada a Prova Brasil. (INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA, 2010).

Finalmente, essa breve apresentação das características do Sistema de Avaliação da Educação Básica (SAEB) e sua considerável evolução durante os já 20 anos de existência justificam a escolha desse sistema – mais especificamente a avaliação da Prova Brasil – como parâmetro de comparação com o SAEM, que é o objeto de estudo desse trabalho.

Assim, após o resgate dos referenciais teóricos necessários ao desenvolvimento da pesquisa, passa-se ao detalhamento da metodologia adotada.

3. METODOLOGIA

O capítulo sobre metodologia é dividido em três partes. A primeira parte descreve os aspectos metodológicos abordados nesta pesquisa, incluindo a definição de população e amostra, forma de coleta de dados, as variáveis a serem estudadas e as proposições e hipóteses que serão testadas. A segunda parte apresenta as etapas de desenvolvimento da pesquisa. Finalmente, a terceira parte apresenta uma visão geral das etapas e como será feita a análise dos resultados.

A Figura 4 apresenta a organização desse capítulo.

Figura 4. Organização da Metodologia da pesquisa

3.1. CONTEXTOS METODOLÓGICOS DA PESQUISA

As pesquisas científicas recebem algumas classificações diferentes, dependendo do critério adotado. Com relação aos objetivos da pesquisa, elas são classificadas como: exploratórias, descritivas ou explicativas (GIL, 2008). As pesquisas descritivas objetivam a descrição das características de uma determinada população e se caracterizam pela utilização de técnicas padronizadas de coleta de

dados. Desse modo, em relação aos objetivos da pesquisa, entende-se que o presente projeto foi desenvolvido a partir do uso de pesquisa descritiva.

Além da classificação quanto ao objetivo da pesquisa, é importante a sua classificação em relação aos procedimentos técnicos que serão utilizados, ou seja, o delineamento da pesquisa, uma vez que ela define o modelo conceitual e operativo que será usado (GIL, 2008). Dentre as várias classificações possíveis, o presente trabalho fez uso do método de Estudo de Caso, uma vez que se objetiva estudar a particularidade e complexidade de um simples caso, o SAEM. Segundo Yin (2010, p. 24), “o método de Estudo de Caso permite que os investigadores retenham as características holísticas e significativas dos eventos da vida real – como os ciclos individuais da vida, [...] o desempenho escolar, [...]”, o que justifica a escolha do referido método. Gil (2008, p. 54) define o estudo de caso como “estudo profundo e exaustivo de um ou poucos objetos, de maneira que permita seu amplo e detalhado conhecimento, tarefa praticamente impossível mediante outros delineamentos já considerados”.

Em relação à natureza do método, as pesquisas são qualificadas em qualitativas ou quantitativas. As pesquisas qualitativas têm por característica o trabalho com dados qualitativos, ou seja, que não permitem o uso de instrumental estatístico (RICHARDSON, 1989). Por sua vez, as pesquisas quantitativas extraem dados quantitativos, a partir do uso de procedimentos sistematizados, tentando explicar a ocorrência de algum fenômeno. Assim, quanto à natureza do método, a presente pesquisa teve, em uma primeira etapa, natureza qualitativa, visto que foi feita uma análise documental para apresentar o modelo de avaliação adotado pelo SAEM. Na etapa posterior, a pesquisa passou a ter uma natureza quantitativa, uma vez que foram trabalhados os resultados de algumas edições do SAEM, a fim de se extrair algumas características da população estudada.

3.1.1. POPULAÇÃO / AMOSTRA

O objeto de estudo deste trabalho é o SAEM, que é aplicado na rede municipal de Uberaba desde o ano de 2006. Até o ano de 2010, foram aplicadas duas avaliações, a primeira ao final do 1º semestre e a segunda ao final do 2º semestre letivo. A avaliação é aplicada a todos os alunos da rede, ou seja, envolve alunos do 1º ao 9º ano do Ensino Fundamental. A rede municipal de Uberaba, no Ensino Fundamental, é composta por 27 escolas com cerca de 17 mil alunos. Historicamente, como pode ser visto na Figura 5, os índices de abstenção têm se mantido próximos a 10%. A literatura consultada (ANDERSON, 2007; STEVENSON, 2001) apresenta expressões para cálculos do número mínimo de elementos de uma amostra, sendo que o valor base para uma população finita é de uma amostra de 5% da população. Assim, como serão usados dados de cerca de 90% da população, pode-se concluir que, quantitativamente, a amostra de dados do SAEM que usada neste trabalho é representativa da população da rede municipal de ensino de Uberaba e, desse modo, suficiente para o desenvolvimento do estudo de caso proposto.

Figura 5. Gráfico de abstenção nas avaliações

Fonte dos dados: Relatórios de resultados do SAEM gerados pelo Colégio Cenecista Dr. José Ferreira

3.1.2. COLETA DE DADOS

A coleta de dados de um estudo de caso deve abranger mais de uma técnica básica para a obtenção de dados. Gil (2008, p. 140) ressalta que “Isso [mais de uma técnica] constitui um princípio básico que não pode ser descartado”. Yin (2010, p. 124) define que a evidência do Estudo de Caso pode vir de seis fontes distintas: documentos, registros em arquivos, entrevistas, observação direta, observação participante e artefatos físicos.

Para proceder aos testes de validade de critério do SAEM foram usados dados secundários, referentes aos dados de desempenho das escolas no SAEM de novembro de 2007 e os de desempenho das referidas escolas na Prova Brasil em 2007, a partir do site do INEP (www.inep.gov.br). A Prova Brasil avalia os alunos da 4ª série (5º ano) e 8ª série (9º ano) do Ensino Fundamental. Além disso, os alunos das escolas rurais não eram submetidos à Prova Brasil. Já a rede municipal possui algumas escolas que só trabalham com o 1º ano do Ensino Fundamental e algumas outras que não possuem a 8ª série (9º ano) do Ensino Fundamental. Assim, foram usados os dados do SAEM somente das séries e escolas avaliadas pelo SAEM e pela Prova Brasil, totalizando 23 escolas da 4ª série e 21 escolas da 8ª série do Ensino Fundamental.

Para os testes de validade dos itens, de validade de construto e de confiabilidade foram usados também os dados da 1ª avaliação do SAEM 2010, uma vez que eram os resultados mais recentes à época da verificação desses itens.

Além do uso dos resultados do SAEM, utilizou-se de entrevistas com os representantes das partes responsáveis pela produção do SAEM – Colégio Cenecista Dr. José Ferreira e Secretaria de Educação e Cultura da Prefeitura Municipal de Uberaba -, além dos usuários do sistema de avaliação – diretores, coordenadores, pais e alunos. Dessas entrevistas, obtiveram-se informações suficientes para a montagem da descrição dos passos de produção de uma avaliação do SAEM e também para a identificação de novos usos e melhorias que o sistema possa vir a receber, dois dos objetivos específicos deste trabalho. As entrevistas foram realizadas seguindo o modelo da entrevista focada (YIN, 2010), que se caracteriza por ser rápida, seguindo um determinado conjunto de questões, conforme está apresentado no apêndice B deste documento, mas podendo se tornar

abertas a novos assuntos. Boni e Quaresma (2005) apresentam as formas mais utilizadas de entrevistas: a) entrevista estruturada; b) semiestruturada; c) aberta, d) grupos focais; e) história de vida e f) entrevista projetiva. Dentre as formas citadas, a que se identifica com a entrevista focada definida por Yin (2010) é a entrevista semiestruturada, como definem Boni e Quaresma (2005, p.75):

As entrevistas semiestruturadas combinam perguntas abertas e fechadas, onde o informante tem a possibilidade de discorrer sobre o tema proposto. O pesquisador deve seguir um conjunto de questões previamente definidas, mas ele o faz em um contexto muito semelhante ao de uma conversa informal.

O Quadro 6 apresenta um resumo com o protocolo dos roteiros de entrevistas que foram realizadas.

Roteiro	Foco	Objetivo	Análise	Resultados Esperados
B1	<ul style="list-style-type: none"> • Funcionários do Colégio Dr. José Ferreira. 	<ul style="list-style-type: none"> • Conhecer os passos da produção de uma avaliação do SAEM. 	<ul style="list-style-type: none"> • Levantamento inicial dos passos da produção de uma avaliação e cruzamento das informações contidas nas entrevistas. 	<ul style="list-style-type: none"> • Levantamento de dados para documentação das etapas do SAEM. • Indicação de problemas e possíveis melhorias no processo
B2	<ul style="list-style-type: none"> • Funcionários da Secretaria Municipal de Educação de Uberaba. 	<ul style="list-style-type: none"> • Saber mais sobre a participação da secretaria na produção do SAEM; • Conhecer os procedimentos de aplicação da avaliação; • Identificar atuais usos dos resultados; • Coletar sugestões de melhorias e novos usos do SAEM. 	<ul style="list-style-type: none"> • Cruzamento das informações dos entrevistados. 	<ul style="list-style-type: none"> • Continuação do levantamento de dados sobre a produção do SAEM; • Indicações sobre como os resultados do SAEM são usados; • Recebimento de sugestões de melhorias e novos usos do SAEM.
B3	<ul style="list-style-type: none"> • Diretores das escolas da rede municipal; • Coordenadores pedagógicos; • Professores. 	<ul style="list-style-type: none"> • Conhecer como a escola recebe e usa os resultados do SAEM; • Conhecer sobre como é aplicada a avaliação; • Coletar sugestões de melhorias e novos 	<ul style="list-style-type: none"> • Cruzamento das informações dos entrevistados 	<ul style="list-style-type: none"> • Continuação das indicações sobre como os resultados do SAEM são usados; • Recebimento de sugestões de melhorias e novos usos do SAEM.

Roteiro	Foco	Objetivo	Análise	Resultados Esperados
		usos do SAEM.		
B4	<ul style="list-style-type: none"> • Pais de alunos; • Alunos das escolas da rede municipal. 	<ul style="list-style-type: none"> • Conhecer como a família recebe e usa os resultados do SAEM; • Conhecer participação dos pais na aplicação da avaliação; • Coletar sugestões de melhorias e novos usos do SAEM. 	<ul style="list-style-type: none"> • Cruzamento das informações dos entrevistados. 	<ul style="list-style-type: none"> • Conhecimento de como os resultados do SAEM são usados; • Consolidação das sugestões de melhorias e novos usos do SAEM.

Quadro 6. Protocolo dos Roteiros de entrevistas

Finalmente, foi usada como fonte de coleta de dados a observação direta. Segundo Yin (2010, p. 136),

A evidência observacional é frequentemente útil para proporcionar informação adicional sobre o tópico sendo estudado. Se o estudo de caso for sobre uma nova tecnologia ou um currículo escolar, por exemplo, as observações da tecnologia ou do currículo em funcionamento são auxiliares valiosos para o entendimento dos seus verdadeiros usos ou de qualquer problema potencial encontrado.

Para atingir os objetivos da observação direta, o pesquisador acompanhou a aplicação da 2ª avaliação do SAEM, em 2010, no dia 24 de novembro.

3.1.3. VARIÁVEIS

Para a comparação dos desempenhos das escolas nos dois exames realizados, SAEM e Prova Brasil, as variáveis podem ser definidas como:

- Nome da escola: variável nominal;
- Série: variável nominal;
- Desempenho Prova Brasil: variável racional que representa o desempenho da escola nessa avaliação, numa escala que vai de 0 a 10. Vale lembrar que a Prova Brasil divulga os resultados também em escala de proficiência

- Desempenho SAEM: variável racional do desempenho da escola nessa avaliação, numa escala que vai de 0 a 100.

Para a realização da validade dos itens e de construto, foram considerados como sujeitos os desempenhos das turmas e não os desempenhos dos alunos. Assim, as variáveis usadas foram:

- Nome da escola: variável nominal;
- Série: variável nominal;
- Turma: variável nominal;
- Desempenhos por questão: variáveis racionais que representam o percentual de acerto dos alunos de uma turma em cada uma das 3 questões propostas;

3.1.4. PROPOSIÇÕES e HIPÓTESES

No presente estudo, desejou-se comprovar a possibilidade de o SAEM ser uma avaliação capaz de avaliar o desempenho acadêmico de uma rede de educação básica. Segundo GIL (2002), uma proposição é uma expressão verbal suscetível de ser declarada verdadeira ou falsa. Quando se relacionam a testes empíricos, as proposições são chamadas de hipóteses. Para tanto, foram formuladas as seguintes proposição e hipótese:

- **Proposição P1**: O SAEM é um sistema de avaliação que atende os padrões de qualidade quanto à adequação ética, utilidade, viabilidade e rigor, definidos pela JCSEE para as avaliações de estudantes.
- **Hipótese H1**: Não há diferença significativa entre os desempenhos das escolas da Rede Pública Municipal de Uberaba na Prova Brasil e no SAEM.

Desse modo, tanto a proposição P1 quanto a Hipótese H1 estão vinculadas ao objetivo específico **(a)** deste trabalho, que caracteriza o SAEM conforme preconiza a literatura sobre sistemas de avaliação e também ao objetivo geral desse trabalho, que era o de analisar as vantagens e limitações do SAEM.

3.2. DELINEAMENTO DA PESQUISA

Nesta parte, são descritos os procedimentos que foram seguidos na realização deste trabalho, iniciando pela apresentação da Figura 6, que contém a visão global da pesquisa, cujo objeto de estudo é o SAEM.

Figura 6. Visão global da pesquisa

A análise de uma avaliação, ou seja, a avaliação da avaliação é chamada de meta-avaliação (FIRME; LETICHEVSKY, 2002). Segundo Hertz (2006, p. 736), a meta-avaliação

[...] pode ser conceituada operacionalmente como um processo de descrição e julgamento de um estudo ou qualquer procedimento de avaliação, fundamentada em padrões e critérios propostos e validados pelas associações profissionais e outros organismos auditores.

Hertz (2006) ainda complementa afirmando que essa fundamentação deve ser feita em relação a padrões internacionais e/ou governamentais. Conforme abordado no referencial teórico, na falta de padrões nacionais definidos para se analisar a qualidade de uma avaliação, as referências nacionais pesquisadas – Vianna (2000); Firme e Letichevsky (2002); Dannemann, Firme e Letichevsky (2005);

Ribeiro e Gusmão (2010) – são unânimes em adotar os padrões de qualidade da avaliação norte-americanos, definidos pelo JCSEE.

A análise do SAEM se deu a partir de cada uma das quatro faces definidas para um sistema de avaliação. Em sua face Teórica, foram explorados os aspectos teóricos que dão fundamentação para que o SAEM seja reconhecido como um sistema de avaliação. Desse modo, analisou-se o SAEM, a partir do que se entende por meta-avaliação, ou seja, verificou-se a adequação da avaliação aos padrões de qualidade definidos pelo JCSEE para as avaliações de estudantes, conforme está apresentado no Quadro 2 (p. 56). Além disso, elaborou-se um quadro comparativo entre as principais características do SAEM e da Prova Brasil, usado como referência de sistema de avaliação.

Já em sua face Desenvolvimento, foram apresentados os passos que constituem a produção de uma avaliação temática, desde a definição do tema, passando pela elaboração das questões, até a impressão e organização logística de todo o material, além da verificação da validade de conteúdo, pertinente a essa etapa do processo.

A face Aplicação, por sua vez, procurou descrever, inicialmente, como é o funcionamento da aplicação de uma avaliação do SAEM. Num segundo momento, a análise se deteve a analisar os resultados de uma avaliação e como seus relatórios são usados pelos interessados.

Finalmente, a face Expansão buscou identificar melhorias que podem ser implementadas e novos usos e aplicações para os resultados que, por ventura, não estejam sendo explorados até o presente momento.

Foi considerado que a análise de um sistema de avaliação a partir da definição de suas faces é uma contribuição deste trabalho para o estudo de sistemas de avaliação.

3.2.1. FACE TEÓRICA

A primeira etapa desta pesquisa consistiu em fazer uma descrição detalhada do modelo de avaliação proposto para o SAEM a partir de um levantamento da referencias bibliográficas existentes que abordam o tema avaliação educacional e, mais especificamente, avaliação em larga escala.

Conforme explanado no referencial teórico, as referências sobre a qualidade da avaliação convergem para os padrões estabelecidos pelo JCSEE. Esse comitê definiu uma série de padrões para se garantir a qualidade de uma avaliação, com sua primeira edição publicada em 1981 (VIANNA, 2000; RODRIGUES, 2009). Os padrões estão agrupados em quatro categorias: 1) Adequação Ética; 2) Utilidade; 3) Exequibilidade (viabilidade); e 4) Rigor/Exatidão (precisão). Assim, o resultado dessa primeira etapa foi a verificação da adequação do SAEM aos padrões do JCSEE.

Ao aprofundar o conhecimento sobre a teoria que sustenta os métodos adotados por alguns desses sistemas de avaliação, como a Prova Brasil, foram buscadas as referências que conceituam a metodologia utilizada para a elaboração, montagem e aferição dos resultados gerados por esses sistemas. Em relação à teoria dos testes, Pasquali (2003) apresenta os dois modelos: a TCT (Teoria Clássica dos Testes) e a TRI (Teoria de Resposta ao Item), que podem ser adotados para a elaboração das questões de uma avaliação. Pretendeu-se analisar as características dos dois modelos. Já em relação à montagem dos cadernos de avaliação, alguns sistemas de avaliação utilizam a metodologia denominada Blocos Incompletos Balanceados (BIB) (INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA, 2001). Em relação à apresentação dos resultados, grande parte dos sistemas de avaliação, seguindo o SAEB, trabalha com a construção de uma escala de proficiência que associa o desempenho dos alunos a uma matriz de habilidades e competências. Já o SAEM apresenta os resultados dos alunos por desempenho proporcional ao número de questões certas em relação ao total de questões propostas. Pretendeu-se analisar as vantagens e desvantagens dos dois métodos.

Conforme pode ser observado a partir de alguns autores já citados (BARLOW, 2006; FERNANDES, 2009; HADJI, 2001; VIANNA, 1989), as avaliações recebem várias classificações – avaliação somativa e formativa; avaliação interna e externa, avaliação responsiva, avaliação autêntica, etc., conforme o ponto de vista do qual se deseja destacar. O SAEM usa a denominação de avaliação “temática” para ressaltar a sua característica de uma avaliação na qual todas as questões são referentes a um tema único. Esse termo, avaliação temática, não é encontrado na bibliografia como um modelo já definido de avaliação. Assim, foi feita a devida caracterização do que é uma avaliação temática, relacionando o modelo aos modelos existentes.

Ao fim desta etapa, objetivou-se ter uma visão detalhada do objeto de estudo, o SAEM, em sua face teórica. Nesse sentido, esperou-se atingir o objetivo específico **(a)** referente à apresentação do modelo de avaliação proposto pelo SAEM, comparando-o com um sistema de avaliação de qualidade, além da comprovação da proposição P1, relativa aos padrões de qualidade da avaliação.

3.2.2. FACE DESENVOLVIMENTO

A produção de uma avaliação para o SAEM passa por várias etapas, desde a escolha do tema até a entrega dos resultados à Secretaria de Educação e aos diretores escolares, que os repassarão aos alunos e seus pais. Nessa etapa, pretendeu-se apresentar todos os passos dessa produção, o que representa a face prática do SAEM.

Para tanto, foram coletados dados a partir dos documentos produzidos a cada avaliação pelo Colégio Cenecista Dr. José Ferreira, responsável pela produção das avaliações. Vale lembrar que, segundo Yin (2010), a informação documental pode tomar várias formas, tais como correspondência eletrônica, cartas, minutas de reuniões, anotações e outros registros internos. Além disso, foram realizadas entrevistas com os profissionais, tanto do Colégio como da Secretaria de Educação

da Prefeitura Municipal de Uberaba, envolvidos nessa etapa, objetivando dirimir possíveis dúvidas não esclarecidas através da documentação que será pesquisada.

Ainda dentro desta etapa, pretendeu-se proceder a análise da validade de conteúdo, visto que essa análise é pertinente a essa etapa do trabalho. Assim, foram analisados o atendimento aos três grandes pontos levantados por Pasquali (2003) para a validação de conteúdo: 1) definição do conteúdo, 2) explicitação dos processos psicológicos (os objetivos) a serem avaliados e 3) determinação da proporção relativa de representação no teste de cada tópico do conteúdo, a partir dos passos definidos na teoria.

Ao final dessa etapa, buscou-se a geração de informação suficiente que oriente o desenvolvimento de uma avaliação temática do SAEM por qualquer outra pessoa. Assim, esperou-se ter atingido parcialmente o objetivo específico **(b)**, referente à descrição dos passos de produção de uma avaliação e o objetivo **(c)**, no que se refere à verificação de validade de conteúdo da avaliação.

3.2.3. FACE APLICAÇÃO

Para essa etapa, pretendeu-se descrever, em complemento à etapa anterior, como o SAEM é efetivamente aplicado levando-se em conta um caso específico, a rede municipal de ensino de Uberaba, buscando identificar de que forma todos os envolvidos (secretaria, diretores de escolas, alunos e pais de alunos) interpretam e usam os resultados de uma avaliação.

As avaliações do SAEM são aplicadas aos alunos em suas escolas de origem sob a responsabilidade da própria Secretaria Municipal de Educação devido, principalmente, aos elevados custos dessa etapa. Nesse sentido, inicialmente, foram entrevistados os responsáveis da Secretaria Municipal de Educação de Uberaba para, em complemento à pesquisa de documentos, conhecer-se de que forma a Secretaria atua na organização da efetiva aplicação do SAEM. Um exemplo específico que pode ser citado é o uso de pais de alunos como fiscais voluntários no dia de aplicação do SAEM. Em complemento, foi usada a observação direta como

fonte de coleta de dados, uma vez que o pesquisador acompanhou o processo de aplicação da 2ª avaliação do SAEM 2010, realizado no dia 24 de novembro de 2010, e a 1ª avaliação do SAEM 2011, realizada no dia 18 de maio de 2011.

O passo seguinte foi descrever como a comunidade escolar recebe, interpreta e usa os resultados de uma avaliação do SAEM. Para tanto, foram escolhidas 3 escolas com características diferentes – desempenho, número de alunos, localização (urbana ou rural) – para fazer parte da pesquisa. Em cada escola, foi entrevistado o diretor, o coordenador pedagógico, dois professores, dois alunos e dois pais de alunos, buscando evidências acerca do uso dos resultados de uma avaliação. Foi importante o uso da observação direta, ou seja, a ida do pesquisador às escolas para observar como eram ministradas – e se eram ministradas – as aulas de reforço para alunos com baixo desempenho no SAEM.

Dentro da etapa de aplicação, também foram feitas as análises estatísticas relativas à validade da avaliação, em complemento às verificações iniciais realizadas na fase de desenvolvimento, uma vez que essas análises dependem do resultado da avaliação. Após a aplicação de uma avaliação do SAEM pôde ser feita a verificação da validade dos itens. Para tanto, conforme já relatado, foram usados os dados da 1ª avaliação do SAEM 2010. Essa escolha se justifica devido ao fato de os dados dessa avaliação serem os mais recentes quando da verificação da validade de construto.

Segundo Pasquali (2003), para testes na TCT, a validade do item pode ser obtida pela determinação da correlação entre o item e o critério, que pode ser o desempenho em um teste paralelo. Assim, foi verificada a validade das questões que compõe a avaliação em análise.

Em relação à validade do teste, a validade de construto só pode ser garantida quando obtida a partir de mais de uma técnica (PASQUALI, 2003). Conforme a adequação dos testes que usam a TCT, foram aplicadas as técnicas: Correlação com outros testes de mesmo construto e validação convergente-discriminante.

Já para verificar a validade de critério, pode-se usar um teste comprovadamente válido para ser o parâmetro contra o qual será feita a validade, nesse caso do tipo concorrente. A avaliação da Prova Brasil, promovida pelo Governo Federal e já analisada na face Teórica, devido à sua comprovada validade

em medir conhecimento, será usada como critério para testar a validade de critério da avaliação do SAEM.

A confiabilidade de um teste é expressa, dentro da Teoria Clássica dos Testes (TCT), pelo coeficiente de fidedignidade, que é obtido pela razão entre a variância verdadeira e total do teste (PASQUALI, 2003). Dos três procedimentos experimentais de coleta da informação para testar a confiabilidade: a) um mesmo teste em uma única ocasião; b) dois testes em uma única ocasião; e c) um mesmo teste em duas ocasiões, conhecido como teste-reteste. Visto que foram trabalhados com dados secundários, já coletados, o único delineamento possível nesse estudo foi o uso do mesmo teste em uma única ocasião. Ainda segundo Pasquali (2003), há duas técnicas principais para a estimação do coeficiente de fidedignidade: a correlação simples e as técnicas alfa. A primeira usa a correlação enquanto que a segunda a variância. A obtenção do coeficiente de fidedignidade por correlação quando se tem apenas um teste, exige que o mesmo seja dividido em duas metades, devendo-se tomar cuidado para a forma da seleção dos itens que comporão cada uma das metades, para que os “dois” testes sejam equivalentes. Por essa técnica, o coeficiente de fidedignidade é obtido pela correlação medida entre os dois conjuntos de testes. Já a técnica alfa mais usada é o coeficiente alfa de Cronbach, que visa verificar a consistência interna do teste através da consistência interna dos itens. Empregaram-se as duas técnicas citadas para a determinação do coeficiente de fidedignidade e, assim, testar a confiabilidade do SAEM.

Terminada essa etapa, buscou-se ter conhecimento sobre as particularidades da aplicação de um caso do objeto de estudo do presente trabalho, o SAEM Uberaba, atingindo, assim, a totalidade do objetivo específico **(b)** acerca da descrição da aplicação do sistema de avaliação e o objetivo específico **(c)**, relativo à validade e confiabilidade do mesmo. Além disso, esperou-se aceitar ou rejeitar a hipótese H1, relativa à inexistência de diferenças significativas de desempenho dos alunos na Prova Brasil e no SAEM.

3.2.4. FACE EXPANSÃO

Passadas todas as etapas propostas anteriormente, acumulam-se subsídios suficientes para a identificação de melhorias na metodologia atual do SAEM a fim de que seus resultados sejam mais representativos da realidade educacional que se pretende aferir. Nesse sentido, criou-se um quadro comparativo entre as características do SAEM e de outros sistemas de avaliação, identificando-se quais são as características ainda não presentes no SAEM que podem ser acrescentadas ao sistema para o seu aperfeiçoamento.

Por outro lado, buscou-se identificar também novos usos dos dados resultantes da avaliação como, por exemplo, a definição de metas de desempenho das escolas com contagem de pontuação para a avaliação de desempenho dos professores. Nesse sentido, procurou-se orientar a Secretaria da Educação para a efetiva implantação de um sistema de metas para a rede municipal.

Ainda sobre a identificação novos usos do SAEM, foram feitas entrevistas com pessoas envolvidas em todas as etapas de produção e aplicação do sistema de avaliação, além de seus usuários, como os alunos e seus pais, objetivando a coleta de algumas ideias de aplicação do sistema que ainda não haviam sido implementadas.

Finalmente, ao término desta última etapa, procurou-se alcançar o objetivo específico **(d)** que buscava identificar pontos de melhoria e oportunidades de usos ainda não exploradas, além de atingir o objetivo geral proposto: analisar o SAEM em todas as suas faces.

3.3. VISÃO GERAL E ANÁLISE DOS RESULTADOS

Descrita as metodologias que foram usadas nesse trabalho, o Quadro 7 a seguir apresenta uma visão geral das mesmas e que formam o delineamento dessa pesquisa.

Etapa	Objetivo	Instrumento de coleta de dados	Resultados Esperados
Face Teórica	<ul style="list-style-type: none"> • Descrição teórica do modelo de avaliação do SAEM, segundo a teoria sobre qualidade da avaliação e os sistemas de avaliação existentes; 	<ul style="list-style-type: none"> • Pesquisa bibliográfica; • Documentação existente; • Dados do SAEM. 	<ul style="list-style-type: none"> • Verificação da adequação do SAEM aos padrões do JCSEE; • Elaboração de quadro comparativo entre as características do SAEM e da Prova Brasil; • Verificação da proposição P1, relativa à qualidade da avaliação.
Face Desenvolvimento	<ul style="list-style-type: none"> • Conhecer sobre todas as etapas de desenvolvimento de uma avaliação do SAEM; • Verificar se os critérios de validade de um teste são seguidos. 	<ul style="list-style-type: none"> • Documentação existente; • Entrevistas com funcionários do Colégio e da Secretaria de Educação. 	<ul style="list-style-type: none"> • Descrição das etapas de produção de uma avaliação do SAEM; • Testar a validade de conteúdo do SAEM.
Face Aplicação	<ul style="list-style-type: none"> • Conhecer sobre a aplicação do SAEM na rede municipal de Uberaba; • Identificar os atuais usos dos resultados de uma avaliação do SAEM. 	<ul style="list-style-type: none"> • Documentação existente; • Entrevistas com funcionários da Secretaria de Educação, diretores, coordenadores pedagógicos, pais e alunos. 	<ul style="list-style-type: none"> • Complementação da descrição das etapas de produção de uma avaliação do SAEM; • Aplicar a validade dos itens, a validade de construto e de critério e a confiabilidade do SAEM; • Aceitação ou rejeição da hipótese da indiferença no desempenho dos alunos na Prova Brasil e no SAEM.
Face Expansão	<ul style="list-style-type: none"> • Identificar possíveis melhorias na metodologia atual do SAEM; • Conhecer possíveis novos usos dos resultados do SAEM. 	<ul style="list-style-type: none"> • Documentação existente; • Entrevistas com funcionários da Secretaria de Educação, diretores, coordenadores pedagógicos, pais e alunos. 	<ul style="list-style-type: none"> • Elaboração de quadro comparativo mostrando as melhorias que podem ser incorporadas ao SAEM em relação à teoria de avaliação e a os outros sistemas de avaliação; • Descrição de novos usos dos resultados do SAEM

Quadro 7. Resumo do delineamento da pesquisa

O planejamento de como será feita a análise dos resultados, conforme cita Yin (2010) é uma etapa importante de todo projeto de pesquisa, mesmo reconhecendo a sua dificuldade de construção devido à variação das formas de coleta de dados. Gil (2002) destaca a dificuldade, dentro de estudos de caso, da falsa sensação de certeza que o pesquisador pode ter a cerca de suas conclusões. Este fato foi particularmente importante nesse trabalho, uma vez que o mestrando é parte integrante da equipe de produção do SAEM, sendo de suma importância que o mesmo não se deixasse influenciar pelas possíveis certezas sobre as conclusões a que imaginava chegar.

Yin (2010) expõe alguns critérios que devem ser levados em conta para a análise dos dados. Segundo o autor, as análises estatísticas oferecem alguns critérios explícitos para as interpretações. Entretanto, ainda segundo Yin (2010), muitas análises não permitirão o uso da estatística, devendo se considerar outras possibilidades. O mais importante é que haja uma estratégia analítica que permita contar uma “história” que engloba seus dados, com princípio, meio e fim. Yin (2010) define quatro estratégias gerais, não mutuamente exclusivas, para análise dos dados: 1) contando com proposições teóricas; 2) desenvolvimento da descrição do caso; 3) uso de dados quantitativos e qualitativos; e 4) pensando em explicações rivais. A estratégia que foi usada para esse trabalho foi a do uso de dados quantitativos e qualitativos, seguindo uma orientação do autor relativa à estratégia escolhida (YIN, 2010, p.161):

Determinados estudos de caso podem incluir quantidades substanciais de dados quantitativos. Se esses dados forem submetidos à análises estatísticas ao mesmo tempo em que os dados qualitativos permanecem, apesar disso, centrais a todo estudo de caso, você terá conseguido, com sucesso, uma forte estratégia analítica.

Além da análise dos dados quantitativos com o uso de ferramentas auxiliadas pelo computador, os dados qualitativos, oriundos das entrevistas, foram analisados conforme as dimensões que estão definidas no Quadro 8, apresentado a seguir.

Dimensões	Face	Roteiro B1: Funcionários do Colégio	Roteiro B2: Servidores da Secretaria Municipal	Roteiro B3: Diretores e Coordenado- res	Roteiro B4: Alunos e pais de alunos
Papel	Desenvolvi- mento	Questão 1	Questões 1 e 4	Questões 1 e 6	Questões 1 e 4
Etapas	Desenvolvi- mento	Questões 2 a 5	Questões 2, 9, 10 e 11		
Replicabilidade	Desenvolvi- mento	Questão 6	Questão 12	Questão 16	
Documenta- ção	Desenvolvi- mento	Questão 7	Questão 13	Questão 17	
Melhorias	Desenvolvi- mento	Questões 8 a 10	Questão 14		
Organização da aplicação	Aplicação		Questões 3 e 5	Questões 2 a 4, 10, 14 e 15	Questões 2 e 7
Preparação dos alunos	Aplicação			Questões 7 e 8	Questões 5 e 6
Uso dos resultados	Aplicação		Questões 6 a 8	Questões 9 a 13	Questões 8 a 11
Melhorias	Aplicação		Questões 14 e 15	Questão 5, 18 e 19	
Novos usos	Expansão		Questão 16	Questão 20	Questão 12

Quadro 8. Dimensões de análise dos dados qualitativos

4. RESULTADOS

O objeto de estudo deste trabalho é o SAEM, analisado a partir de suas características e identificando suas vantagens e limitações como instrumento de medida do desempenho acadêmico de alunos. Para tanto, no capítulo de Metodologia, foram definidas abordagens e técnicas para o desenvolvimento dessa análise. Os resultados da aplicação dessas técnicas foram organizados conforme mostra a Figura 7.

Figura 7. Organização dos Resultados

4.1. FACE TEÓRICA

O objetivo da face Teórica é realizar uma análise detalhada do modelo de avaliação adotado pelo SAEM, à luz das teorias sobre qualidade da avaliação em larga escala de redes de ensino e em relação ao principal sistema de avaliação da educação básica existente no Brasil, a Prova Brasil/SAEB.

Assim, a primeira etapa consiste em verificar a adequação do SAEM aos padrões de qualidade da avaliação definidos pelo JCSEE. Para tanto, foi criado o Quadro 9, que apresenta a relação de *Standards*, por categoria, as descrições sobre o que cada *Standard* aborda e as justificativas da adequação ou não do SAEM em

relação a cada *Standard*. Finalmente, foi necessária a criação de uma escala de três níveis, denominada Adequado, para julgar o nível de adequação do SAEM ao referido padrão. Os níveis da escala são: sim, parcialmente e não. A última coluna do quadro 9 apresenta o resultado do julgamento do item desenvolvido pelo pesquisador com base nas informações coletadas pelas entrevistas realizadas.

<i>Standards</i>	Descrição	Justificativa	Adequado
<i>Standards de Adequação Ética</i>			
Serviço aos estudantes	A avaliação deve promover os princípios sólidos de educação e o estudo eficaz dos alunos	Conforme as entrevistas coletadas, o SAEM estimula aos alunos a se dedicarem mais aos estudos.	SIM
Políticas e Procedimentos Adequados	Os procedimentos por escrito devem ser desenvolvidos, implementados e disponibilizados, de modo que a avaliação seja imparcial e justa.	A avaliação é elaborada por equipe de professores fora da rede municipal e aborda conteúdos divulgados antecipadamente	SIM
Acesso à informação avaliativa	O acesso à informação do aluno deve ser permitido, mas somente a ele e a pessoas autorizadas, de modo a proteger a confidencialidade das informações	O desempenho do aluno é divulgado somente para os pais, através do Boletim Individual de Desempenho, e para a gestão escolar	SIM
Tratamento dos estudantes	Os estudantes devem ser tratados com respeito em todos os aspectos da avaliação	Os estudantes recebem lanche antes da avaliação, são respeitados em suas dificuldades de tempo para a realização da prova e necessidades de ir ao banheiro, por exemplo.	SIM
Direito dos estudantes	As avaliações devem estar em acordo com as leis e com os direitos humanos	A avaliação segue todas as legislações, e os alunos são livres para discutirem as questões da prova e questionarem as respostas, uma vez que eles ficam com o caderno de questões.	SIM
Avaliação equilibrada	As avaliações devem fornecer informações que identificam os pontos fortes e fracos dos alunos	O relatório de desempenho por questão apresenta as porcentagens de acerto e de erro de cada um dos descritores cobrados na prova, com dados para cada turma, cada série e para a rede toda	SIM

Standards	Descrição	Justificativa	Adequado
Standards de Utilidade			
Orientação construtiva	As avaliações devem ser construtivas, de modo que as decisões educacionais visem aos interesses dos alunos.	O SAEM é uma avaliação temática, ou seja, contextualizada e interdisciplinar, que busca a construção do conhecimento para que os alunos tenham sucesso na resolução das questões propostas	SIM
Usos e utilizadores definidos	Os usuários e os usos da avaliação precisam estar definidos para que a avaliação contribua com o processo de aprendizagem	O SAEM tem seus usuários (Secretaria de Educação, Direção Escolar, alunos e pais) e seus usos (conhecer a realidade educacional; apoiar estratégias de recuperação, informar a família) bem definidos.	SIM
Alcance da informação	As informações recolhidas devem ser abrangentes de modo que as respostas esperadas sejam respondidas	O SAEM avalia todos os conteúdos do ciclo comum (Ciências, Geografia, História, Língua Portuguesa e Matemática), satisfazendo às necessidades dos usuários quanto ao conhecimento do desempenho do aluno	SIM
Competência do avaliador	Os avaliadores devem ter o conhecimento e as habilidades necessárias para que a avaliação seja conduzida com competência e os resultados sejam confiáveis.	Os responsáveis pelo SAEM têm formação na área de educação. A equipe de elaboradores tem se mantido praticamente a mesma desde a primeira avaliação. Os processos de apuração dos resultados – correção dos cartões de respostas – foram melhorados com a introdução do código de barras	SIM
Valores explícitos	Os professores que avaliam alunos devem identificar e justificar os valores usados para avaliar o desempenho dos alunos na avaliação	O SAEM utiliza a contagem de acertos nas questões, o que é facilmente identificável pelos usuários dos resultados	SIM
Relatórios eficazes	Os relatórios devem ser claros, precisos e relevantes para serem úteis aos usuários	O SAEM tem relatórios voltados para todos os seus usuários. As informações transmitidas variam conforme a relevância ao usuário	SIM
Follow-up (acompanhamento)	A avaliação deve ter procedimentos de acompanhamento do	Não foram identificadas ações específicas de acompanhamento do entendimento. Entretanto, os	PARCIAL-MENTE

Standards	Descrição	Justificativa	Adequado
	entendimento das informações pelos usuários	gestores atendem às dúvidas via e-mail e os pais recebem os Boletins Individuais em reunião com a direção e professores, onde recebem orientações quanto à análise dos resultados.	
Standards de Viabilidade			
Orientação prática	Os procedimentos da avaliação devem ser práticos, produzindo as informações necessárias de modo eficiente	Os procedimentos para o desenvolvimento, a aplicação e a emissão dos relatórios atingiram um estágio de maturidade ao longo dos cinco anos ininterruptos de existência que os tornam extremamente eficientes	SIM
Viabilidade política	A avaliação deve ser planejada e conduzida de modo a responder, antecipadamente, às dúvidas de seus usuários.	Foi identificado que há atenção em disponibilizar a maior quantidade possível de informação aos alunos, professores, diretores, etc.	SIM
Apoios à avaliação	O tempo e os recursos adequados devem ser investidos na avaliação, para que elas sejam efetivamente planejadas e implementadas, com ampla comunicação dos resultados	O SAEM é uma das principais políticas públicas da Secretaria Municipal de Educação, não faltando apoio político e financeiro para a sua realização ininterrupta durante esses cinco anos.	SIM
Standards de Rigor/Exatidão			
Orientação para a validade	A avaliação deve ser desenvolvida e implementada de modo que as interpretações a cerca do desempenho dos alunos seja válida	A validade do SAEM foi comprovada na face Desenvolvimento, item 4.2.9, e na face Aplicação, item 4.3.4.	SIM
Expectativas definidas para os alunos	As expectativas de desempenho dos alunos devem estar claras, de modo que os resultados sejam defensáveis e significativos	Os conteúdos e os descritores abordados nas duas avaliações anuais do SAEM ano são divulgadas no início do ano letivo.	SIM
Análise do contexto	As variáveis contextuais devem ser identificadas e levadas em conta na análise de desempenho	No SAEM, todos os alunos são submetidos à avaliação e recebem o seu boletim de desempenho. Entretanto, os desempenhos dos	PARCIAL- MENTE

Standards	Descrição	Justificativa	Adequado
	dos alunos	alunos com deficiência de aprendizagem (atestado por laudo médico) não são considerados para o cálculo do desempenho da turma e da escola. Por outro lado, os alunos das escolas na zona rural são classificados com os mesmos critérios que os alunos das escolas urbanas	
Procedimentos Documentados	Os procedimentos para execução da avaliação devem estar descritos, de modo a facilitar a sua explicação e justificativa	Notou-se que as etapas de elaboração e produção da avaliação são pouco documentadas, o que não ocorre na face Aplicação.	PARCIAL-MENTE
Informação defensável	A coleta dos dados deve ser adequada, facilitando decisões corretas e possibilitando sua defesa e justificativa	A coleta dos dados é feita através da aplicação da prova do SAEM a todos os alunos simultaneamente. Conforme foi conferido durante a realização das entrevistas, o processo ocorre de modo a não comprometer a coleta dos dados.	SIM
Informação confiável	Os procedimentos da avaliação devem ser escolhidos ou desenvolvidos e implantados, para que forneçam informações confiáveis sobre o desempenho dos alunos	A confiabilidade do SAEM foi confirmada nos resultados da face Aplicação, item 4.3.4.	SIM
Identificação e Gestão de tendências/de-turpações	A avaliação deve ser livre de preconceitos (cultural, social, físico, etc.), para que as conclusões sejam justas	Os resultados gerados pelo SAEM não levam em conta nenhum tipo de característica dos alunos. Os alunos com dificuldades de aprendizagem (comprovada por laudo médico) fazem as avaliações normalmente, recebem seus resultados, mas seus desempenhos não são contados para o desempenho da turma e o da escola	SIM
Processamento da Informação e Controle da qualidade	As informações coletadas, processadas e relatadas devem ser sistematicamente revistas, corrigidas caso seja necessário, e mantidas em segurança	Segundo relatou o funcionário do Colégio responsável pela apuração dos resultados, todos os dados, desde a primeira edição do SAEM, estão armazenados de duas formas: em estado bruto, nos arquivos gerados pela leitora óptica, e nos relatórios com os resultados do processamento dos dados.	SIM

Standards	Descrição	Justificativa	Adequado
Análise da informação	As informações recolhidas devem ser sistematicamente e precisamente analisadas, para que as finalidades da avaliação sejam alcançadas	Há rotinas de verificação das respostas assinaladas pelos alunos. Há rotinas de conferência dos resultados antes dos mesmos serem divulgados. Todos os procedimentos são realizados por software desenvolvido exclusivamente para essa função	SIM
Conclusões justificadas	As conclusões sobre o desempenho do aluno devem ser justificadas, de modo que seus usuários tenham confiança nelas	Os resultados do SAEM podem ser conferidos diretamente pelos alunos, que, de posse do caderno de questões da avaliação e do gabarito oficial, fazem a sua correção. Os diretores também podem conferir o desempenho de cada turma, ou mesmo da escola, a partir do desempenho dos alunos.	SIM
Meta-avaliação	Os procedimentos da avaliação devem ser periodicamente analisados a partir dessas ou de outras normas, para que possíveis erros sejam prontamente identificados e corrigidos	O presente trabalho é a primeira experiência de meta-avaliação pela qual o SAEM foi submetido.	PARCIAL-MENTE

Quadro 9. Análise da adequação do SAEM aos Standards da JCSEE
Fonte dos dados: Rodrigues (2009); Joint (2011) e resultados da pesquisa

Ao final da análise de adequação do SAEM aos Standards do JCSEE, observa-se que dos 27 padrões estabelecidos o SAEM pode ser considerado adequado em 23 desses. Nos demais quatro padrões, o SAEM foi considerado parcialmente adequado.

Dentre esses, o padrão *Follow-up* foi considerado parcialmente adequado devido ao fato de que não foram encontrados procedimentos claros de acompanhamento do entendimento das informações. Entretanto, há histórico de esclarecimento de dúvidas de diretores e de professores pelos gestores da Secretaria e do Colégio. Além disso, conforme relatado pelos diretores e confirmado pelos pais entrevistados, a entrega do Relatório de desempenho individual é feita pessoalmente aos pais, que têm suas dúvidas esclarecidas nesse momento.

A questão relativa à Análise de Contexto foi considerada parcialmente adequada, visto que, apesar de se considerarem alunos com *deficit* de aprendizagem, não são separadas, por exemplo, as escolas rurais das escolas urbanas.

As etapas de desenvolvimento da avaliação e de correção e geração dos relatórios não possuem seus Procedimentos Documentados. Por outro lado, os procedimentos de aplicação da avaliação estão bem documentados.

Em relação ao padrão de Meta-avaliação, não há um trabalho sistematizado de avaliação da avaliação, sendo o presente trabalho o primeiro a ser realizado, o que fez com que o padrão fosse considerado parcialmente adequado.

De um modo geral, pode-se analisar esse resultado para cada uma das quatro categorias dos Standards. O Quadro 10 mostra a distribuição dos resultados de adequação.

Adequação	SIM	PARCIALMENTE	NÃO
Adequação Ética	6	0	0
Utilidade	6	1	0
Viabilidade	3	0	0
Rigor	8	3	0
GERAL	23 (85,18%)	4 (14,82%)	0

Quadro 10. Distribuição dos resultados de adequação

O resultado aponta para um ótimo ajuste do SAEM – 85,18% de itens considerados adequados – aos padrões de qualidade de um sistema de avaliação educacional em larga escala, conforme preconiza a teoria avaliativa. As características consideradas parcialmente adequadas e a considerada inadequada devem ser vistas como pontos de melhoria a serem implementadas no sistema de avaliação.

Em continuação à apresentação dos resultados da Face Teórica, desejou-se estabelecer uma comparação entre as características do SAEM e as de outro sistema de avaliação da educação básica de qualidade reconhecida.

Conforme já explanado no referencial teórico, no início da década de 90, o Governo Federal estabeleceu as bases do Sistema Nacional de Avaliação da Educação Básica, o SAEB, que tem se mantido até os dias atuais. A partir de 2005, o sistema passou a ser formado por duas avaliações: a ANEB (Avaliação Nacional da Educação Básica), que manteve as características do SAEB, e a ANRESC (Avaliação Nacional do Rendimento Escolar), conhecida como Prova Brasil (INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA, 2010).

As características que determinam as diferenças entre o SAEB e a Prova Brasil estão representadas no Quadro 11 a seguir:

Característica	Prova Brasil (ANRESC)	SAEB (ANEB)
Séries/Anos	5º e 9º Anos do Ensino Fundamental	5º e 9º Anos do Ensino Fundamental; 3ª série do Ensino Médio
Rede	Escolas públicas das áreas urbanas e rurais	Escolas públicas e privadas das áreas urbanas e rurais
Abrangência	Quase universal , dentro das séries/anos avaliados	Amostral
Resultados	Desempenho para o Brasil, regiões, unidades federativas, municípios e escolas	Desempenho para o Brasil, regiões e unidades federativas

Quadro 11. Diferenças entre a Prova Brasil e o SAEB.

Fonte dos dados: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (2011)

Por outro lado, o SAEM é um sistema de avaliação cujo objetivo principal é, a partir do desempenho acadêmico dos alunos da rede pública municipal de Uberaba-MG, da Classe de 6 anos/1º ano do Ensino Fundamental até a 8ª série/9º ano do Ensino Fundamental³², aferir a qualidade de educação ofertada e fornecer dados para apoiar as políticas de recuperação de alunos com deficiências de aprendizagem. A avaliação do SAEM é temática, ou seja, todas as questões

³² A Secretaria Municipal de Educação adotou, até 2010, a nomenclatura de Classes para as séries iniciais do Ensino Fundamental – Classe de 10 anos, por exemplo – e a nomenclatura de Séries para as séries finais. No início de 2011, foi alterada essa nomenclatura de classes/séries e adotada a de anos, ficando a descrição atual como sendo do 1º ao 9º ano do Ensino Fundamental.

propostas se referem a um tema principal. Assim, não há a estrutura de um banco de questões. Para cada avaliação, conforme o tema escolhido, novas questões são elaboradas.

A avaliação do SAEM envolve as disciplinas de Ciências, Geografia, História, Língua Portuguesa e Matemática. A distribuição do número de questões para cada classe/série e para cada disciplina varia conforme mostrado no Quadro 12.

Classe/série	TOTAL	CIE	GEO	HIS	POR	MAT
Classe de 6 anos do EF	20	3	3	3	7	7
Classe de 7 anos do EF	20	3	3	3	7	7
Classe de 8 anos do EF	20	3	3	3	7	7
Classe de 9 anos do EF	30	4	4	4	9	9
Classe de 10 anos do EF	30	4	4	4	9	9
5ª série do EF	30	4	4	4	9	9
6ª série do EF	30	4	4	4	9	9
7ª série do EF	30	4	4	4	9	9
8ª série do EF	30	4	4	4	9	9

Quadro 12. Distribuição do número de questões por classe/série e disciplina
Fonte dos dados: Colégio Cenecista Dr. José Ferreira

Vale destacar que as 20 questões elaboradas para as Classes de 6 e de 7 anos são discursivas. Para as demais classes/séries, as questões são objetivas, com cinco alternativas cada.

A partir das informações conhecidas das principais características do modelo de avaliação adotado pelo SAEM e pela Prova Brasil/SAEB, é possível apresentar o seguinte quadro comparativo – Quadro 13 – entre os dois sistemas.

Característica	Prova Brasil/SAEB	SAEM - Uberaba
Período de Aplicação	Uma avaliação a cada dois anos	Duas avaliações por ano
Classes/Séries/Anos	4ª série/5º ano e 8ª série/9º ano do Ensino Fundamental; 3ª série do Ensino Médio (somente o SAEB)	Da Classe de 6 anos/1º ano do Ensino Fundamental até a 8ª série/9º ano do Ensino Fundamental
População	Censitária. Alunos da rede pública urbana e rural, acima de 20 alunos em cada ano.	Censitária. Alunos da rede pública municipal urbana e rural
Instrumento de medida cognitiva	Prova com questões objetivas e tradicionais	Prova com questões objetivas ³³ e temáticas
Disciplinas	Língua Portuguesa e Matemática	Ciências, Geografia, História, Língua Portuguesa e Matemática
Conteúdos	Matriz de referência fixa	Conteúdos definidos no início de cada ano e criação dos descritores
Elaboração das questões	Questões extraídas de um banco de itens, já calibradas e pré-testadas para uso da TRI	Questões temáticas, elaboradas exclusivamente para aquela avaliação
Montagem do Caderno de Questões	Uso dos Blocos Incompletos Balanceados, gerando vários cadernos de questões diferentes	Elaboração de um mesmo caderno de questões para todos os alunos de uma mesma série.
Instrumentos contextuais	Questionário aplicado aos alunos, professores e diretores	Não são aplicados questionários
Modelo estatístico	Teoria da Resposta ao Item (TRI)	Teoria Clássica dos Testes (TCT)
Cálculo do desempenho	Resultados apresentados em uma escala de proficiência, com uso da TRI	Média de acertos de cada aluno na avaliação, variando de 0 a 100
Resultados	Por escola, município, unidade federativa, região e país	Por aluno, turma, série, escola e município
Relatórios	Gráficos de desempenho de cada unidade escolar por disciplina	Relatórios de classificação e desempenho por aluno, unidade escolar e toda a rede de ensino

Quadro 13. Comparação entre o SAEM e a Prova Brasil

Fonte dos dados: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (2011) e Colégio Cenecista Dr. José Ferreira

³³ Os alunos das classes de 6 e 7 anos respondem a questões discursivas.

Algumas características constantes do quadro anterior merecem um aprofundamento em virtude das diferenças entre os sistemas. Quando se observam as características de periodicidade das séries e das disciplinas presentes no SAEM, percebe-se que o mesmo é bem mais completo que a Prova Brasil/SAEB. A periodicidade do SAEM permite que se detectem as deficiências de aprendizagem e, no mesmo ano, que se faça um trabalho específico para sanar o problema.

A extensão da avaliação para todas as classes/séries e o envolvimento de todas as disciplinas traz vantagens para que, mais uma vez, se conheça em profundidade as deficiências apresentadas pelos alunos e, principalmente, para que se evite problemas de maior destaque/importância de uma classe/série ou mesmo disciplina. Esse fato evita problemas de reducionismo de conteúdo, conforme relatado por Fernandes e Gremaud (2009) e Bonamino e Franco (1999).

É necessário que se reconheça que o volume de dados coletados em uma avaliação da Prova Brasil/SAEB justifica a questão da periodicidade, bianual, e da aplicação se limitar somente à Classe de 10 anos/ 5º ano e a 8ª série/ 9º ano do Ensino Fundamental. Entretanto, conforme questionam também Bonamino e Franco (1999), a limitação da avaliação da Prova Brasil/SAEB às disciplinas de Língua Portuguesa e Matemática é prejudicial à avaliação.

Em relação à elaboração das questões, a Prova Brasil/SAEB trabalha com um banco de itens, atualmente mantido pelo INEP, de onde as questões são retiradas para a montagem dos Cadernos de Questões. O uso de questões pré-testadas permite que se use a técnica dos Blocos Incompletos Balanceados (BIB), de modo que alunos respondam a grupos de questões diferentes, mas de mesmo nível de dificuldade. A repetição de blocos entre avaliações de edições anteriores possibilita que, através da aplicação da Teoria da Resposta ao Item, os resultados entre as diferentes edições possam ser comparáveis. A grande vantagem do uso dos BIBs é a maior abrangência de conteúdos inseridos na avaliação toda (INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA, 2001).

Já no SAEM, as questões são elaboradas exclusivamente para cada edição de uma avaliação. O caráter temático da avaliação praticamente inviabiliza a construção de um banco de questões e muito menos a repetição das mesmas

questões em diferentes edições. A avaliação temática permite a contextualização e a interdisciplinaridade dos conteúdos, o que torna a avaliação muito mais próxima e atrativa para os alunos. Vale lembrar que os temas são divulgados com antecedência, permitindo que os professores explorem as relações de sua disciplina com o tema em questão, enriquecendo a aprendizagem dos alunos.

Além dos instrumentos cognitivos – as provas –, a Prova Brasil/SAEB faz uso também de instrumentos contextuais – os questionários –, que são aplicados aos alunos, professores e diretores. O objetivo desses questionários é conhecer os fatores intra e extraescolares que afetam o desempenho dos alunos (BONAMINO;FRANCO, 1999; PESTANA, 1998). O SAEM se limita a avaliar o desempenho cognitivo dos alunos, não fazendo uso de instrumentos contextuais.

Talvez a principal diferença entre os dois sistemas de avaliação esteja no modelo estatístico adotado para cada sistema. A Prova Brasil/SAEB adota a Teoria da Resposta ao Item (TRI), desde 1995. O SAEM adota a Teoria Clássica dos Testes (TCT). Entretanto, é importante ressaltar que Lord³⁴ (1980 apud ANDRIOLA, 2009, p. 322) afirma que a TRI não contradiz a TCT, mas somente apresenta supostos adicionais que respaldam soluções não possíveis com a TCT.

A vantagem maior da TRI é a possibilidade de comparação dos resultados entre edições diferentes da avaliação e, com a ajuda de uma escala de proficiência, a comparação entre diferentes séries numa mesma edição. Certamente, essa é uma característica interessante para uma avaliação educacional aplicada periodicamente.

Em termos de resultados, o SAEM é mais detalhista que a Prova Brasil/SAEB, uma vez que ele apresenta o desempenho dos alunos e das turmas de uma série. Já a Prova Brasil apresenta os resultados por escolas, unidades federativas e o do país.

Uma grande diferença também se apresenta em termos dos relatórios produzidos pelos sistemas. Ao consultar o portal da Prova Brasil, no item Resultados³⁵, tem-se a opção de gerar o relatório com os resultados de uma determinada escola, cidade, unidade federativa, região ou do país. Ao escolher uma

³⁴ LORD, F. M. *Applications of item response theory to practical testing problems*. New Jersey: Lawrence Erlbaum Associates, 1980

³⁵ Resultados Prova Brasil: <http://sistemasprovabrasil2.inep.gov.br/ProvaBrasilResultados/home.seam>

escola, a relatório que surge é o que está apresentado no Anexo A. Nele, há, primeiramente, informações sobre o número de alunos participantes, desde o Brasil todo até a escola participante. Em seguida, para cada uma das duas séries/anos avaliadas, o relatório apresenta a taxa de aprovação, os desempenhos em Língua Portuguesa e Matemática e a nota do IDED para a escola, o município, a unidade federativa e o Brasil. Finalmente, são apresentadas, também para as duas séries/anos e para as redes já citadas, as informações sobre a Média de horas-aula diárias, Porcentagem de docentes com curso superior e taxa de distorção idade-série.

A cada edição, os dados colhidos no SAEM resultam na publicação de vários relatórios, direcionados aos diferentes usuários do sistema. Os relatórios são organizados em função dos seus usuários: a) Alunos e Pais; b) Gestão Escolar e c) Secretaria de Educação. A seguir apresenta-se um detalhamento de cada um desses relatórios.

a) Alunos e pais

Os pais recebem um relatório denominado **Relatório de Desempenho Individual**, conforme pode ser visualizado no Anexo B1. Esse é o único relatório impresso que o Colégio Cenecista Dr. José Ferreira produz, visando exatamente que o mesmo seja entregue às famílias. Conforme pode ser constatado as entrevistas com alguns pais e alunos, as escolas têm tentado entregar os boletins diretamente para os pais, em reuniões específicas para esse fim.

O relatório apresenta, inicialmente, o desempenho do aluno em termos de número de questões corretas e porcentagem de acertos em relação ao total de questões existentes. Em seguida, o relatório mostra a classificação do aluno em relação aos alunos de sua turma, de sua escola e da rede toda. Posteriormente, o relatório apresenta o desempenho médio de todos os alunos de sua turma, de sua escola e da rede toda. Finalmente, para cada uma das disciplinas avaliadas, o relatório mostra o número de questões certas, a média de acertos do aluno e as médias de acertos dos alunos de sua turma, de sua escola e da rede toda.

b) Gestão Escolar

Os relatórios da gestão escolar são entregues na forma de CD-ROM para o diretor da unidade e este os repassa para os pedagogos e os professores. Os relatórios disponíveis são:

- **Relatório de Classificação Geral por Escola:** relação, por classe/série e por turma, de todos os alunos – da escola – que participaram da avaliação, com a classificação em ordem crescente, o nome do aluno e o número de acertos obtidos. O seu modelo pode ser visto no Anexo B2.
- **Relatório de Classificação por Disciplina:** relação, por classe/série, por turma e por disciplina, de todos os alunos – da escola – que participaram da avaliação, com a classificação em ordem crescente, o nome do aluno e o número de acertos obtidos na disciplina, como pode ser visto no Anexo B3.
- **Relatório de Desempenho da Escola:** para cada classe/série, o relatório apresenta uma tabela contendo os desempenhos, por disciplina e total, de cada uma das turmas daquela série. O relatório apresenta, ainda, os desempenhos médios da escola e da rede em cada uma das disciplinas avaliadas. A sua visualização está disponível no Anexo B4. O objetivo desse relatório é auxiliar a gestão escolar a identificar se existem turmas com dificuldade em alguma disciplina específica ou mesmo generalizada, a partir da comparação dos desempenhos das turmas com o desempenho da escola e da rede como um todo.
- **Relatório de Desempenho por Questão:** como pode ser visto no Anexo B5, o relatório apresenta, para cada questão de uma determinada classe/série, a distribuição de alternativas marcadas pelos alunos de cada uma das turmas da escola. Além disso, o relatório informa o desempenho de todos os alunos da escola e da rede em cada uma das questões. Ao informar também a que conteúdo a questão se refere, o relatório auxilia no conhecimento detalhado das deficiências dos alunos nos conteúdos cobrados na avaliação,

possibilitando as intervenções necessárias e de modo pontual – somente nos conteúdos em que foram apresentadas dificuldades.

- **Gráfico de distribuição de acertos:** a partir de 2009, foi acrescentado aos relatórios das escolas o gráfico de distribuição de acertos, por classe/série. O gráfico de barras apresenta o número de alunos que tiraram determinado número de acertos, como pode ser visto no Anexo B6.

c) Secretaria Municipal de Educação

A Secretaria Municipal de Educação recebe um CD-ROM com os relatórios exclusivos da Secretaria além de todos os arquivos gerados para cada uma das escolas. Os relatórios exclusivos da secretaria estão listados a seguir.

- **Relatório de Classificação Geral:** relação, por classe/série, de todos os alunos que participaram da avaliação, com a classificação em ordem crescente, o nome do aluno e o número de acertos obtidos. É semelhante ao Relatório de Classificação da Escola, apresentado no Anexo B2.
- **Relatório de Classificação por Disciplina:** relação, por classe/série e por disciplina, de todos os alunos que participaram da avaliação, com a classificação em ordem crescente, o nome do aluno e o número de acertos obtidos na disciplina. É semelhante ao Relatório de Classificação por Disciplina da Escola, que pode ser visto no Anexo B3.
- **Relatório de Desempenho por Questão:** Apresentação, para cada questão de uma determinada classe/série, da distribuição de alternativas marcadas por todos os alunos que participaram da avaliação. De forma semelhante ao relatório também gerado para as escolas, ele auxilia no conhecimento detalhado das deficiências dos alunos nos conteúdos cobrados na avaliação, auxiliando o trabalho do departamento pedagógico da secretaria, conforme relatado por uma das servidoras entrevistadas. O modelo do relatório é semelhante ao

Relatório de Desempenho por Questão das escolas, disponível no Anexo B5.

- **Gráfico de Distribuição de Acertos:** gráfico de barras que apresenta, para cada classe/série, o número de alunos, dentre todos os participantes, que tiraram determinado número de acertos, semelhante ao que pode ser visto no Anexo B6.
- **Relatório Geral de Desempenho:** relatório que apresenta, para cada classe/série, a relação das escolas participantes, em ordem decrescente de desempenho, com os seus respectivos desempenhos médios, geral e por disciplina, conforme está apresentado no Anexo B7. O relatório também apresenta o desempenho médio da classe/série, que é obtido pela razão entre a soma dos acertos de todos os alunos da classe/série e a soma de todas as questões respondidas, multiplicada por 100.
- **Relatório Geral de Desempenho da Rede Municipal:** relatório que relaciona todas as escolas participantes e apresenta os seus respectivos desempenhos médios em cada uma das classes/séries em que a escola tenha alunos participantes. A última coluna desse relatório apresenta o resultado do desempenho médio da escola, que é obtido pela razão entre a soma dos acertos de todos os alunos da escola e a soma do número total de questões respondidas por esses alunos, multiplicada por 100. Nas últimas linhas, são apresentados os desempenhos médios de cada uma das classes/séries e o desempenho médio da rede municipal. Como pode ser visto no Anexo B8, a relação está ordenada em ordem decrescente de desempenho médio da escola.

O **Relatório Geral de Desempenho da Rede Municipal** passou a ser gerado a partir de 2008 por solicitação da Secretaria Municipal de Educação. O desempenho médio da escola, calculado nesse relatório e usado para classificar as escolas, apresenta distorções, uma vez que o seu cálculo “mistura” os desempenhos de classes/séries diferentes. Basta analisar o exemplo apresentado no Anexo B8. As escolas que possuem somente as classes de 6 e 7 anos, cujos desempenhos são

superiores às demais classes/séries, têm um desempenho médio superior às demais escolas, posicionando-se nos primeiros lugares na relação.

Além disso, o relatório possibilita a comparação e o rankeamento das escolas, surgindo, por inevitável, as “melhores” e as “piores”. É importante ressaltar que a comparação e o rankeamento não fazem parte dos objetivos do SAEM e, aliás, era um dos grandes temores dos gestores durante as primeiras avaliações. Entretanto, acompanhando a entrega dos resultados, tem-se a impressão de que os gestores se adaptaram ao rankeamento e não fazem tantos julgamentos a partir desses resultados, atentando-se mais aos seus próprios resultados. Finalmente, na tentativa de não estimular tais julgamentos, o relatório apresenta a relação das escolas em ordem decrescente, mas não enumera essa ordenação (1º, 2º, 3º etc.), para que não haja pelo menos a memorização dessa classificação (Anexo B8).

4.2. FACE DESENVOLVIMENTO

A face Desenvolvimento tem por objetivo conhecer todas as etapas de produção de uma avaliação do SAEM. Para tanto, a partir da coleta de informações contida nos documentos da avaliação e a das entrevistas com os participantes dessa etapa – Colégio Cenecista Dr. José Ferreira e Secretaria Municipal de Educação e Cultura (SEMEC) da Prefeitura Municipal de Uberaba – foi possível organizar um roteiro descritivo de todas as etapas de desenvolvimento de uma avaliação do SAEM, buscando a possibilidade de reprodução do modelo por qualquer interessado em fazê-lo. Essas etapas são apresentadas, a seguir, conforme sua relação cronológica ou de setores de responsabilidades. A Figura 8 mostra a relação entre essas partes.

Figura 8. Etapas de desenvolvimento de uma avaliação do SAEM

4.2.1. DEFINIÇÕES INICIAIS

A montagem de uma avaliação do SAEM se inicia por uma reunião entre os representantes do Colégio Cenecista Dr. José Ferreira e da Secretaria Municipal de Educação. Nesse momento, são discutidas as datas para as duas avaliações anuais e as dificuldades e problemas ocorridos no ano anterior que, junto às novas propostas, definem as mudanças para o próximo ano. Essa reunião sempre aconteceu no final do ano e no início do ano seguinte. Entretanto, com a mudança da data da 1ª avaliação para o início do ano letivo, a partir de 2010, essa reunião deve ser feita sempre ao final do próprio ano.

A partir das informações recebidas, a equipe do Colégio José Ferreira elabora a proposta orçamentária para fornecer todos os serviços solicitados e a apresenta junto com toda a documentação necessária para a formalização do contrato de prestação dos serviços.

4.2.2. CONTRATO

Um processo importante no desenvolvimento de uma avaliação para uma rede pública é a questão contratual, uma vez que esse processo pode ser demorado e, desse modo, comprometer o cronograma das avaliações.

Conforme consta na Lei das Licitações, nº 8666 (BRASIL, 1993), os órgãos públicos são obrigados a celebrar seus contratos, acima de R\$ 8 mil, através de processo licitatório. Entretanto, em seu artigo 24 – sobre a dispensa da licitação –, inciso XIII, a lei cita:

Artigo 24 – XIII – na contratação de instituição brasileira incumbida regimental ou estatutariamente da pesquisa, do ensino ou do desenvolvimento institucional, ou de instituição dedicada à recuperação social do preso, desde que a contratada detenha inquestionável reputação ético-profissional e não tenha fins lucrativos. (BRASIL, 1993)

A CNEC – Campanha Nacional de Escolas da Comunidade –, mantenedora do Colégio Cenecista Dr. José Ferreira, é uma instituição educacional sem fins lucrativos. Desse modo, os contratos celebrados entre o Colégio e a Prefeitura Municipal de Uberaba têm sido firmados por dispensa de licitação. A grande vantagem desse processo é a possibilidade de planejamento das ações a longo prazo, mesmo que os contratos sejam de periodicidade anual.

4.2.3. ESCOLHA DO TEMA

Como já foi dito na parte da Introdução, o modelo da avaliação do SAEM é temático, o que significa que todas as questões são elaboradas a partir de um único tema. A responsabilidade de escolha dos temas do SAEM é da Secretaria Municipal de Educação e Cultura (SEMEC). A escolha do tema deve levar em conta o seu grau de atualidade, a fim de aumentar a disponibilidade de fontes bibliográficas para a elaboração de questões e a sua abrangência, para que todas as disciplinas explorem os conteúdos trabalhados em sala.

A seguir, a Quadro 14 apresenta uma lista com os temas usados nas avaliações do SAEM desde 2006. Nota-se que todos os temas procuram ser abrangente. Por exemplo, o tema da 1ª avaliação de 2010 – “Esporte: unindo povos, disseminando culturas” – não se limita ao tema Esporte, mas abre espaço também para as ciências humanas trabalharem com essa temática.

Avaliação	Tema
1ª avaliação 2006	Uberaba, nossa terra, nossa gente
2ª avaliação 2006	O trabalho, a cidadania e a escola
1ª avaliação 2007	Jogos Pan-americanos
2ª avaliação 2007	Meio ambiente e qualidade de vida
1ª avaliação 2008	Valores humanos: Pessoas melhores fazem um mundo melhor
2ª avaliação 2008	Folclore: nossa cultura, nossas raízes
1ª avaliação 2009	Água: uso consciente, recurso permanente
2ª avaliação 2009	Direitos Humanos – “A Paz é Fruto da Justiça”
1ª avaliação 2010	Esporte: Unindo povos, disseminando culturas
2ª avaliação 2010	Biodiversidade: Cuidando da vida no planeta terra

Quadro 14. Lista de temas do SAEM

Fonte: Relatórios do SAEM – Colégio Cenecista Dr. José Ferreira

Após escolher os temas, a SEMEC os envia para o Colégio José Ferreira para que o mesmo seja aprovado. Assim, há uma reunião com as equipes de professores que trabalharão na avaliação, onde os temas são apresentados. Caso a

equipe do Colégio José Ferreira não veja problemas de associação do tema proposto com os conteúdos, esses são aprovados e podem ser divulgados para as escolas.

A divulgação antecipada dos temas é de grande importância, pois permite que a Secretaria, as escolas e os professores possam definir projetos de estudo. Os próprios alunos podem se preparar, de forma autônoma, para a avaliação buscando fontes de leitura (livros, revistas, filmes, sites etc.) sobre a temática escolhida. Em tempo, estas são algumas das principais características de um sistema de avaliação como o SAEM: o estímulo à leitura e a capacidade de aproximar os conteúdos das disciplinas ao cotidiano dos alunos.

4.2.4. DEFINIÇÃO DOS CONTEÚDOS

A definição dos conteúdos que serão abordados pelas avaliações do SAEM é também de responsabilidade da SEMEC. No primeiro ano, a SEMEC solicitou às escolas o envio dos programas das disciplinas, por classe/série, para que se pudessem estudar as interseções desses programas que seriam inseridas na avaliação. O resultado desse trabalho foi a descoberta de que, enquanto em algumas disciplinas de algumas séries não havia conteúdos em comum – chegando ao extremo de haver professor ministrando História do Brasil em uma turma da 6ª série e outro professor ministrando História Geral para a outra turma da 6ª série, na mesma escola – em outras apresentava-se um programa que necessitaria de pelo menos dois anos letivos para ser cumprido.

Diante da situação encontrada e de posse de uma compilação dos programas recebidos das escolas, os professores do Colégio definiram os conteúdos que estariam presentes nas avaliações do SAEM em 2006. De 2007 a 2010, no início de cada ano, a SEMEC sempre revisou os conteúdos, levando em conta as sugestões recebidas dos professores da rede, e enviou-os ao Colégio para que os professores pudessem também dar suas sugestões. Ao final desse processo, o documento com

os conteúdos a serem aferidos nas duas avaliações do SAEM era enviado às escolas.

Para as avaliações do SAEM 2011, conforme reunião ocorrida ao final de 2010, a 1ª avaliação passou a ser realizada no início do ano letivo – a 1ª avaliação do SAEM 2011 ocorreu dia 13 de abril de 2011. Assim, a avaliação não teve a divulgação de conteúdos específicos de cada disciplina em cada classe/série. Houve somente a divulgação dos descritores (serão abordados no item 4.2.5), que serão os mesmos para as duas avaliações do ano. Os conteúdos foram tomados como sendo os pré-requisitos necessários para que o aluno esteja em uma determinada classe/série, formando, assim, uma avaliação diagnóstica.

4.2.5. ELABORAÇÃO DOS DESCRITORES

A partir dos conteúdos, a cada ano também são definidos os descritores que serão abordados em cada uma das questões. Araújo e Luzio (2005, p. 15), ao descreverem a construção das Matrizes de referência do SAEB, apresentam definição de descritores

Os conteúdos, associados às habilidades a serem desenvolvidas em cada série e disciplina, foram subdivididos em partes menores, cada uma especificando o que os itens das provas do SAEB deveriam medir. Essas unidades são denominadas “descritores”.

Assim, os descritores são uma associação entre os conteúdos curriculares e as operações mentais desenvolvidas pelos alunos (INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA, 2001).

Até o ano de 2010, os professores do Colégio José Ferreira, responsáveis pela elaboração das questões, após receberem a versão final dos conteúdos, elaboravam uma matriz de descritores associados a esses conteúdos. Esse documento é revisado pela equipe da SEMEC que poderia dar sugestões de alterações. Após o documento estar concluído, ele, juntamente com o documento contendo os conteúdos e os temas, é enviado às escolas para que as mesmas possam repassá-los a seus professores e alunos.

Em 2011, com as alterações citadas anteriormente, foram enviados às escolas somente o documento com os descritores selecionados para a 1ª avaliação e que serão os mesmos também para a 2ª avaliação.

4.2.6. ELABORAÇÃO DAS QUESTÕES

Na reunião com os professores da equipe do Colégio José Ferreira, são distribuídas as informações necessárias, tais como o tema, o número de questões e os conteúdos, por classe/série, para que se inicie o processo de elaboração das questões. Desde 2008, as avaliações passaram a ter o apoio de uma Atividade de Revisão, elaborada nos mesmos moldes da avaliação. O objetivo dessa Atividade é o de capacitar os alunos da rede municipal ao modelo da avaliação, não utilizada com frequência nas avaliações tradicionais das escolas. Junto à Atividade de Revisão, é enviada uma cópia do cartão de resposta³⁶ para que o aluno seja capacitado no correto preenchimento do mesmo. A Atividade de Revisão possui um número menor de questões que a avaliação – a metade – e é disponibilizada com pelo menos duas semanas de antecedência do dia da avaliação do SAEM. Além disso, não há correção formal dessa Atividade, o Colégio somente disponibiliza o gabarito das questões às escolas.

Desse modo, os professores elaboram as questões para a Atividade de Revisão e para o SAEM. Vale ressaltar que não há repetição de questões entre a Atividade e o SAEM, mas há a cobrança dos mesmos descritores.

As questões são entregues ao departamento de avaliações do Colégio. Além da questão propriamente dita, o professor deve anexar o texto que por ventura tenha usado de suporte à questão e deve mencionar o descritor a que a mesma se refere. Cada questão deve estar em uma folha separada, a fim de auxiliar o processo de montagem da avaliação, explicado no item 4.2.7. O Colégio, devido às suas necessidades internas e aos crescentes serviços de elaboração de Processos

³⁶ Os alunos das Classes de 6 e 7 anos não usam os cartões de respostas. Essas avaliações possuem somente questões discursivas.

Seletivos, mantém uma equipe de profissionais responsáveis pela digitação e pela diagramação dessas avaliações. O mestrando, além de professor, atua também como coordenador do departamento de avaliações do Colégio.

Conforme relatado pela funcionária A em sua entrevista, a equipe conta com nove profissionais. Como a avaliação do SAEM tem nove provas diferentes – do 1º ao 9º ano do Ensino Fundamental –, cada um fica responsável por uma avaliação.

4.2.7. MONTAGEM DA AVALIAÇÃO

Após o recebimento de todas as questões de uma determinada classe/série, elas são enviadas ao coordenador que fará a montagem dos cadernos de questões, buscando a melhor ordenação possível. O modelo de avaliação temática, proposto para o SAEM, pressupõe que as questões não estejam ordenadas por disciplinas. O coordenador deve agrupar as questões conforme a sua ligação (questões que se referem a um mesmo texto) ou mesmo conforme a sua proximidade. Mesmo tendo um só tema, sempre são identificados alguns sub-temas associados a ele. Por exemplo, o tema da 2ª avaliação de 2010 do SAEM foi: Biodiversidade: cuidando da vida no planeta Terra. Como pode se perceber, essa temática engloba alguns sub-temas, como: Amazônia, diversidade da flora, diversidade da fauna; as relações do homem com o meio ambiente, dentre outros. Assim, o coordenador agrupa as questões relacionadas a cada sub-tema que vai surgindo e, finalmente, reúne todos os sub-temas, de modo que se tente criar uma história do início ao fim da avaliação. Pela qualidade de textos apresentados, é comum que a avaliação se inicie com a Língua Portuguesa.

Após a montagem da avaliação, é feita a digitação e a diagramação da mesma. Terminada essa etapa, a avaliação passa por uma correção de padronização e de digitação, em que são observados os aspectos de atendimento aos padrões de qualidade de imagens, fontes de letras, erros de digitação do texto original e a ordenação numérica das questões. Após essa etapa, a avaliação é encaminhada para a equipe de profissionais de Língua Portuguesa que verifica a

adequação linguística e gramatical das questões, além da adequação do vocabulário à idade da criança, principalmente nos textos usados como contextualização.

Após todas as etapas de correção, os professores de cada disciplina fazem uma releitura das questões. Geralmente, as equipes de cada disciplina têm pelo menos dois professores responsáveis por cada classe/série que se alternam na elaboração e na verificação final das questões. Finalmente, o coordenador faz uma leitura completa da avaliação, buscando eliminar alguns últimos problemas, tais como: a) o uso excessivo de textos suporte às questões, o que pode tornar a avaliação muito extensa e cansativa, b) a coerência e a clareza entre o enunciado, o comando e as alternativas das questões, c) a clareza e a qualidade das imagens usadas nos textos, d) a existência de questões que ultrapassam a página ou que o texto a que se referem se encontra distante da mesma e e) a organização das alternativas de cada questão de modo que não permitam que o aluno encontre a alternativa correta por outros meios – eliminação, extensão da alternativa etc. – senão pela aplicação dos seus conhecimentos.

Após todo esse processo, as avaliações de cada classe/série estão prontas para serem impressas na Edigraf, gráfica da CNEC, também situada em Uberaba-MG, que é responsável pela produção de todo Material Didático do Sistema de Ensino CNEC.

Nota-se que não é feita a calibragem das questões e não há a necessidade de construção de um banco de questões prontas devido à avaliação ser temática. O Anexo D apresenta um exemplo desse modelo de avaliação.

4.2.8. ORGANIZAÇÃO E DISTRIBUIÇÃO DAS AVALIAÇÕES

A Secretaria de Educação é responsável pelo envio dos dados dos alunos para o Colégio. As informações enviadas são: nome da escola, nome do aluno, classe/série e turma. O funcionário B, em sua entrevista, declarou que o processo de envio dos dados – arquivos em formato de planilha eletrônica – está sujeito a falhas, pois depende de que pessoas na secretaria façam o filtro dos dados e os enviem ao

Colégio. O funcionário relata, ainda, que já houve problema de troca de dados, já foram enviados dados de uma série do ano anterior ao invés dos nomes dos alunos do ano corrente.

Os dados dos alunos, acrescidos da data de aplicação do SAEM, são usados para a personalização do Cartão de Respostas e para a impressão da lista de presença de cada turma. Além disso, há a impressão dos dados de cada turma em envelopes brancos, os quais são usados para o lacre e para a devolução dos Cartões de Respostas, após a avaliação.

Por conterem somente questões discursivas, as Classes de 6 e 7 anos recolhem os cadernos de questões para a correção. Nas demais séries, recolhe-se somente o Cartão de Respostas. Todo esse material (cartões, listas e envelopes) é levado ao departamento de expedição da Edigraf que, de posse dos cadernos de questões já devidamente impressos, organiza os pacotes de cada turma e os lacra com o uso de uma seladora plástica. É feito um pacote com as sobras de cadernos de questões e de Cartões de Respostas em branco para uso caso seja necessário.

Todos os pacotes das turmas de cada escola, por turno (matutino, vespertino e noturno), são novamente embalados. Na tarde do dia anterior à aplicação do SAEM, as diretoras das escolas retiram os pacotes referentes à sua escola e os guarda até o início da avaliação no dia seguinte.

4.2.9. VALIDADE DE CONTEÚDO

Encerrada a primeira etapa da Face Desenvolvimento: descrição do desenvolvimento de uma avaliação do SAEM, parte-se para a segunda etapa dessa Face, que é a de testar a validade de conteúdo do SAEM.

Conforme citado no referencial teórico, Pasquali (2003) afirma que a validade de conteúdo é garantida pela técnica da construção do teste. Daí que essa validade se enquadra na Face Desenvolvimento.

Seguindo as orientações descritas por Pasquali (2003, p. 190), analisou-se o enquadramento do SAEM a cada um dos passos propostos. Para tanto, foi

construído o Quadro 15, que apresenta a relação dos passos, uma descrição de seu objetivo e a justificativa do enquadramento ou não do SAEM à condição imposta. De modo semelhante à análise de adequação aos padrões de qualidade da avaliação, já apresentada nesse trabalho, foi criada uma escala de três níveis, denominada Válido, para mensurar o enquadramento do SAEM àquela etapa.

Passo	Descrição	Justificativa	Válido
1. Definição do domínio cognitivo	Definir os objetivos que se quer avaliar. Em termos educacionais, significa a construção de uma Matriz de Referência.	Conforme descrito, uma das primeiras etapas do SAEM é a definição dos descritores que serão objeto de avaliação.	SIM
2. Definição do universo do conteúdo	Definir e delimitar o universo do conteúdo em termos de divisões e subdivisões.	Antes da definição dos descritores, há a delimitação dos conteúdos que serão cobrados nas avaliações do SAEM.	SIM
3. Definição da representatividade do conteúdo	Definir a proporção com que cada conteúdo será representado na avaliação.	O número de descritores é, geralmente, igual ao número de questões, ficando a proporção de ocorrência definida pela quantidade de questões de cada disciplina (ver Quadro 12, p. 87).	SIM
4. Elaboração da tabela de especificação	Construir uma tabela de especificação, na qual serão relacionados os conteúdos com os processos cognitivos.	No desenvolvimento do SAEM, não há a construção dessa tabela de especificação.	NÃO
5. Construção do teste	A elaboração das questões deve levar em conta as técnicas existentes para tal fim.	A equipe de elaboradores de questões do SAEM tem treinamento em construção de itens e vem acumulando experiência nessa área.	SIM
6. Análise teórica dos itens	Verificar a compreensão das tarefas (comandos) propostas nas questões e adequação do item ao conteúdo.	Nas várias etapas descritas de correção, as questões passam pela análise semântica, feita pelo coordenador, e pela verificação de adequação da questão, realizadas pela equipe de avaliação.	SIM
7. Análise empírica dos itens	Realizar a validade estatística das questões.	A validade dos itens foi comprovada nos resultados da Face Aplicação desse trabalho.	SIM

Quadro 15. Análise da validade de conteúdo do SAEM

Fonte dos dados: Pasquali (2003)

Como pode ser visto no Quadro 15, para os sete passos descritos para a validade de conteúdo, o SAEM foi considerado não válido em somente um item. Mesmo assim, foi garantida a validade de conteúdo para o SAEM.

Detalhadas as etapas de desenvolvimento de uma avaliação e comprovada a validade de conteúdo do SAEM, passa-se agora para a descrição das ações necessárias para a aplicação do SAEM, que envolve, além da aplicação da avaliação propriamente dita, a etapa de apuração dos resultados e os atuais usos desses resultados.

4.3. FACE APLICAÇÃO

Os resultados da face Aplicação estão apresentados em quatro etapas. Inicialmente, procurou-se descrever os procedimentos para a aplicação do SAEM na Rede Municipal de Ensino de Uberaba. Em seguida, apresentou-se a descrição dos procedimentos de correção das provas, geração dos relatórios e entrega dos mesmos. Numa terceira etapa, foi feita uma apresentação dos atuais usos dos relatórios gerados com a aplicação de uma avaliação do SAEM. Finalmente, a partir dos resultados de uma avaliação, procurou-se aplicar as técnicas estatísticas para a determinação da validade, tanto dos itens quanto da avaliação, bem como a sua confiabilidade.

4.3.1. APLICAÇÃO DAS PROVAS

A partir da retirada dos pacotes de provas da Gráfica do Colégio até a devolução dos envelopes com os cartões de respostas e dos cadernos de provas para correção e apuração dos resultados, a responsabilidade da gestão do SAEM é da Secretaria Municipal de Educação.

Desse modo, foram realizadas entrevistas com alguns integrantes da Secretaria, além de três diretores, três coordenadores e seis professores de escolas

participantes do SAEM com distintos perfis, a fim de se conhecer os procedimentos para a aplicação da avaliação. As entrevistas também realizadas com seis alunos e seis pais tinham como objetivo confirmar as informações passadas pelos gestores e professores.

É importante destacar que a seleção das escolas e das pessoas dessas escolas para serem entrevistadas não teve a intenção de selecionar uma amostra representativa da população de pessoas envolvidas com o SAEM. O objetivo foi o de levantar as informações necessárias e, na medida do possível, confirmar a veracidade das mesmas ao diversificar o perfil dos entrevistados. Assim, as informações coletadas não foram usadas para fazer inferências sobre a população, mas para descrever e analisar o processo.

Conforme coletado nas entrevistas, os diretores das escolas municipais são os responsáveis pela retirada do material da avaliação, um dia antes da realização da mesma. A princípio, tem-se a impressão de que há um grave problema de confiabilidade nessa ação, pois as provas passam ao poder dos diretores um dia antes da sua realização. Entretanto, conforme informado, os pacotes são abertos somente no dia seguinte, minutos antes do início da avaliação e com a presença de um representante da Secretaria de Educação e de alguns pais que são convidados a participar da aplicação.

Durante as entrevistas, pode-se perceber que as escolas se preparam antecipadamente para o dia da avaliação. Algumas ações relatadas foram: 1) visita do Diretor às salas de aula, reforçando a data e incentivando a presença dos alunos; 2) a organização logística para alocação dos professores que vão aplicar as provas, de modo que eles não estejam em salas que já atuem como professores da turma; 3) a organização da refeição que os alunos fazem antes do início da avaliação; 4) reunião com os professores para informar sobre o tema da avaliação e definir projetos interdisciplinares que o envolvam e 5) reunião para treinamento dos professores na aplicação das provas.

Quanto a essas pré-ações, pode-se destacar o esforço feito para que os alunos não faltem. Houve depoimentos e até sugestão de mudança de agendamento de consulta médica, para que o aluno não faltasse à aplicação do SAEM.

Sobre o treinamento dos professores para a aplicação da avaliação, a Secretaria de Educação elaborou um Manual do Professor Aplicador, o qual é enviado às escolas e as mesmas os reproduzem para os seus professores. A questão da preparação dos professores para a aplicação da prova foi unânime entre alunos e pais, eles demonstram estarem preparados para a aplicação da avaliação e, mesmo entre os professores, a demonstrações de segurança para essa etapa da aplicação.

Ainda sobre as ações prévias para a realização da aplicação das provas do SAEM, é importante relatar a estrutura existente na Secretaria de Educação que está envolvida no processo. Conforme relatado pelas servidoras da Secretaria que foram entrevistadas e de acordo com o documento **Estrutura Orgânica SEMEC**³⁷, o SAEM está sob a responsabilidade do Departamento Pedagógico da Secretaria.

A Figura 9 mostra parte do organograma da Secretaria Municipal de Educação de Uberaba, apresentando somente os cargos que têm relação direta com o SAEM.

Figura 9. Organograma (parte) da Secretaria Municipal de Educação de Uberaba.
 Fonte dos dados: Adaptado de Organograma³⁸ da SEMEC

³⁷ Arquivo eletrônico fornecido por uma das servidoras, que contém a relação hierárquica entre os departamentos, suas finalidades e competências.

³⁸ Arquivo eletrônico fornecido por uma das servidoras, que contém o organograma completo da Secretaria Municipal de Educação e Cultura

Conforme os depoimentos colhidos, o SAEM é uma das principais políticas públicas da Secretaria de Educação. De um modo geral, a distribuição das responsabilidades é organizada pelo Departamento Pedagógico, que é o responsável pelas ações relativas à realização do SAEM, como, por exemplo, a definição dos temas e do cronograma, contato com o Colégio, organização dos servidores que trabalharão no dia da avaliação, etc. A Diretoria de Gestão faz o planejamento das ações e acompanha o seu desenvolvimento, não só na secretaria mas também junto às escolas. À Subsecretaria cabe resolver as questões de celebração de contratos e o acompanhamento do seu cumprimento. O Secretário define, junto ao prefeito, a sua equipe e aos diretores, as metas e as estratégias que serão usadas a partir dos resultados obtidos.

O Departamento Pedagógico da Secretaria de Educação trabalha com o assessoramento pedagógico, através de equipe de pedagogas que fazem visitas periódicas às escolas. Essa equipe é responsável por acompanhar o andamento das atividades das escolas, em relação dos alunos para a realização do SAEM.

Das atividades preparatórias para o SAEM, talvez a mais efetiva em termos de efeitos do SAEM sobre o sistema educacional – chamada de validade consequential por Pasquali (2003) – seja a revisão que os professores realizam com os alunos. Apesar da questão do reducionismo do currículo ser um dos problemas das avaliações em larga escala, conforme relatado por Fernandes e Gremaud (2009) e Oliveira (2011), os alunos entrevistados destacam que a grande maioria dos professores fazem aulas de revisão, alguns preparam materiais complementares, que, na opinião deles, são importantes. Nesse contexto, destaca a frase de uma aluna, ao comentar sobre as mudanças causadas pelo SAEM: “A forma dos professores ensinar (sic) mudou”, diz a aluna completando que, a partir do SAEM, os professores estão mais interessados na aprendizagem dos alunos.

Outra atividade que merece destaque é a Atividade de Revisão, elaborada nos mesmos moldes da avaliação – mesmo tema, mesmos descritores –, mas com um número de questões reduzido à metade. O objetivo principal da Atividade de Revisão é o de eliminar possíveis erros dos alunos por falta de hábito na resolução de questões temáticas e no preenchimento do cartão de respostas. Entretanto, a

Atividade de Revisão, como pôde ser constatado, tem servido para a revisão dos conteúdos envolvidos na avaliação.

No dia da aplicação das provas, o Departamento Pedagógico organiza para que em cada escola haja a presença de um representante da Secretaria, apoiando a direção na condução da aplicação e fiscalizando o cumprimento das determinações da Secretaria, principalmente quanto à lisura do processo.

A transparência da aplicação também é destacada pela participação de alguns pais de alunos no dia da avaliação. Essa iniciativa é digna de todo o louvor, já que, além de auxiliar na lisura do processo, ela busca aumentar a participação dos pais na vida escolar de seus filhos. Nas três escolas entrevistadas, os diretores relataram que não têm tido dificuldades em convidar pais para ajudarem no dia da aplicação das provas.

Os pais têm algumas funções pré-definidas. A primeira é a de conferir a inviolabilidade dos pacotes de provas que serão enviados às salas de aula, considerando que eles foram retirados na gráfica do Colégio no dia anterior. Posteriormente, auxiliam na fiscalização do andamento da aplicação e ajudam a acompanhar alunos ao banheiro. Ao final da avaliação, os pais ajudam a conferir os envelopes com os cartões de respostas, a lacrá-los e, ao final, eles os assinam.

Sobre especificamente a aplicação das provas, alguns alunos declararam desconforto somente com a questão de que todos os alunos da escola, após o término de suas provas, têm que aguardar, sentados em seus lugares, até que nenhum aluno da escola esteja fazendo a prova. Uma das alunas disse que já houve casos de que a escola toda ficou aguardando por vários minutos um último aluno terminar a sua prova.

Um problema registrado por vários entrevistados foi em relação às alterações da data de aplicação do SAEM. Em virtude de problemas diversos, tal como a demora na formalização do contrato entre o Colégio e a Prefeitura Municipal, já houve algumas mudanças de data. Os alunos e professores entrevistados reclamaram que, após terem feito as revisões e estarem aptos a fazer a avaliação, a data é alterada e todo o trabalho fica prejudicado.

4.3.2. CORREÇÃO E GERAÇÃO DOS RELATÓRIOS

Após a aplicação das provas, os diretores encaminham os envelopes lacrados, contendo os cartões de respostas dos alunos da classe de 8 anos (3º ano) à 8ª série (9º ano), para a correção no Colégio Cenecista Dr. José Ferreira. Além dos envelopes, são enviados também os Cadernos de Questões dos alunos das classes de 6 e 7 anos, visto que para essas classes as questões são discursivas.

O gabarito com as respostas é enviado pelo Colégio à Secretaria de Educação somente após a aplicação, e só depois é repassado às escolas. Conforme pôde ser percebido nas entrevistas, é comum, dentre as escolas pesquisadas, que os professores façam a correção da avaliação diretamente do caderno de questões dos alunos e deem o resultado a eles bem antes dos resultados “oficiais”. A discussão entre professores e alunos sobre as questões do SAEM é salutar e faz parte do processo de aprendizagem.

A Secretaria tem uma importante função durante a etapa de correção das provas. Ela recebe os questionamentos quanto a possíveis erros nas questões e os encaminha para o Colégio, junto com uma análise realizada pelos docentes da própria Secretaria. Os questionamentos são analisados e, caso seja necessário, a questão pode ter sua resposta alterada ou mesmo pode ser anulada. Nos casos de anulação das questões, o procedimento é que todos os alunos da classe/série recebam o acerto relativo à questão anulada.

Historicamente, nesses cinco anos de aplicação do SAEM, houve a ocorrência de uma a duas questões, dentre 250 questões de cada edição, que apresentaram problemas e tiveram de ser anuladas.

O processo de correção se inicia no Colégio com uma equipe de professores que trabalham com as séries iniciais realizando a correção das questões discursivas. Conforme relatado por um dos funcionários entrevistados, as professoras trabalham em duplas, fazendo a correção e a revisão da correção. Após o término dos trabalhos, o número de acertos dos alunos é lançado no banco de dados daquela edição do SAEM, a partir de um aplicativo específico para tal função. Esse software possui a relação dos nomes dos alunos por turma e ainda permite a inclusão de

novos alunos, cujos nomes não constavam na relação inicial, fato um tanto comum de ocorrer na rede municipal, devido à rotatividade de alunos.

Por outro lado, os cartões de respostas dos alunos, a partir da classe de 8 anos, são submetidos a um equipamento de leitura óptica para apuração das alternativas assinaladas pelos alunos. Conforme o funcionário explicou, o equipamento gera um arquivo eletrônico com as respostas dos alunos. É interessante destacar que, após a leitura automática feita pela máquina, é feita uma revisão manual dessa leitura, a partir de um aplicativo que identifica, por exemplo, as questões que não tiveram resposta identificada pelo equipamento. Esse procedimento busca eliminar as falhas de mau preenchimento das alternativas no cartão de respostas.

Corrigidos todos os possíveis problemas, é feita a apuração do desempenho dos alunos, a partir da comparação das respostas assinaladas com o gabarito das provas, gerando o desempenho individual do aluno. É a partir da informação do desempenho dos alunos que todos os outros dados (desempenho médio da turma, da série, da escola, da rede, por questão, etc) são calculados.

Vale destacar a informação dada pelo funcionário em relação ao armazenamento e à segurança dos dados. O armazenamento principal que é feito é somente dos arquivos de respostas dos alunos e dos desempenhos individuais, possibilitando uma nova correção em caso de, por exemplo, haver erro no gabarito de respostas informado. Sobre a segurança, a informação dada é de que as cópias dos dados são guardadas em três ambientes físicos diferentes, dentro do Colégio e até fora dele.

Passadas todas essas etapas, os relatórios para as escolas e para a Secretaria são gerados em formato PDF, padrão comum de leitura de arquivos eletrônicos atualmente. Como já foi apresentado nos resultados da face Teórica, é gravado um CD-ROM para cada uma das escolas e outro para a Secretaria, que contém diretórios com todas as informações de cada uma das escolas, além dos relatórios específicos da Secretaria, com as informações relativas à rede municipal. Entretanto, o Relatório de Desempenho Individual é entregue impresso, para que o mesmo seja repassado aos pais de alunos, sem que haja a dependência da escola imprimir tais relatórios.

Desde a realização da primeira edição do SAEM, a entrega dos resultados é feita em uma reunião em que os representantes do Colégio apresentam os resultados para a equipe da Secretaria Municipal de Educação e Cultura e para os diretores das escolas. Além dos relatórios constantes no CD, o Colégio disponibiliza um Relatório da Avaliação, entregue de forma impressa durante a reunião. Esse relatório contém um resumo das informações gerais relativas àquela edição do SAEM, tais como taxa de abstenção e desempenho médio da Rede por classe/série. Além disso, o relatório apresenta as particularidades relativas àquela edição, como problemas nas questões, falhas de envio de dados, etc. Assim, aproveita-se o momento de reunião para colher informações sobre fatos ocorridos e discutir sugestões de melhorias para as próximas edições.

4.3.3. USO DOS RESULTADOS

A pesquisa pelos usos que são feitos dos resultados do SAEM é de extrema importância para este trabalho, na medida em que apontou também as necessidades de melhorias a serem implementadas no SAEM. De acordo com os documentos disponíveis e com os depoimentos colhidos nas entrevistas, foram levantados os principais usos dos resultados.

Nota em avaliação regular

Foi detectado que as escolas têm usado o resultado dos alunos no SAEM como uma das notas regulares dos alunos. Conforme informado pelos diretores e confirmado pelos alunos entrevistados, essa foi uma solução para a falta de motivação dos mesmos em terem um bom desempenho na avaliação, problema comum em avaliações externas.

Assessoramento pedagógico das escolas

O principal uso encontrado para os resultados do SAEM está no apoio ao trabalho de assessoramento pedagógico prestado pela equipe do Departamento Pedagógico da Secretaria às escolas municipais. De posse dos resultados, conforme relato de uma das servidoras da Secretaria, são agendadas visitas às escolas, nas quais os seus resultados são analisados, série por série, turma a turma, com o objetivo de detectar quais foram os problemas que geraram os possíveis maus desempenhos apresentados e de discutir propostas de intervenção para resolver esses problemas.

Seleção para atividades de reforço

Uma das principais funções do SAEM quando o sistema foi elaborado era a de, a partir da avaliação do desempenho dos alunos, apoiar as ações de encaminhamento de alunos de baixo desempenho para atividades de reforço. Entretanto, o que foi detectado nas entrevistas é que o SAEM não tem sido usado para esse fim, uma vez que o resultado da 1ª avaliação de cada ano só é divulgado no início do segundo semestre³⁹. Os professores informaram que a seleção dos alunos para frequentarem as atividades de reforço é feita a partir das avaliações regulares que eles mesmos aplicam aos alunos. Eles também fazem uso do resultado informal que eles geram ao corrigirem o SAEM com os alunos. Desse modo, o resultado “oficial” do SAEM é usado somente como confirmação da manutenção do aluno frequentando as aulas, caso ele ainda não tenha recuperado a sua nota.

Em termos de atividade de reforço, a Rede Municipal de Ensino tem basicamente dois projetos: o SOS Pedagógico e o Tempo Integral. Segundo informado pelas servidoras, o projeto SOS Pedagógico é um projeto de recuperação paralela que ocorre no contra-turno (aluno do matutino tem aulas no turno vespertino

³⁹ A partir de 2011, a primeira avaliação passou a ser diagnóstica, aplicada no início do ano letivo. Mesmo assim, a avaliação desse ano, inicialmente prevista para 13 de abril, foi aplicada somente em 18 de maio. A entrega dos resultados ocorreu no dia 06 de julho último.

e vice-versa) ou na extensão do turno regular de aulas (aula-extra após o término das aulas regulares), sendo a segunda opção a mais frequente, sobretudo na zona rural, porque, assim, o aluno não precisa retornar à escola. Nesse projeto, os alunos recebem aulas de Matemática, com ênfase em raciocínio lógico, e de Língua portuguesa, com foco em leitura e interpretação de textos.

Na Rede Municipal, há 21 escolas denominadas CIEM (Centro Integrado de Educação Municipal) com oferta de atividades em tempo integral para alguns alunos, fundamentalmente das classes de 6 a 10 anos. Esses alunos também recebem aulas de reforço em Língua Portuguesa e em Matemática, além de serem acompanhados na realização das tarefas propostas para casa.

Acompanhamento do trabalho docente

Os coordenadores entrevistados relataram que usam os resultados do SAEM como uma das ferramentas de acompanhamento do trabalho dos professores em sala de aula. Conforme relatado, eles, nas reuniões com todos os docentes para a discussão dos resultados gerais da escola e em reuniões particulares, conversam com os professores sobre rendimento de suas classes/turmas.

Comunicação aos pais

O Relatório de Desempenho Individual é destinado aos pais dos alunos ou a seus responsáveis. Dentre as escolas pesquisadas, todas afirmaram que entregam esse relatório em reunião de pais, aproveitando para conversar com eles sobre o desempenho dos seus filhos, tanto na avaliação quanto em sala de aula. Nesse sentido, é importante ressaltar que o SAEM cumpre um importante papel social ao compartilhar os resultados não só do aluno como também de sua turma e de sua escola com as famílias numa clara função de *accountability*, conforme destacam Fernandes e Gremaud (2009).

Entrega de certificados

O SAEM tem como um dos seus objetivos, além de avaliar o desempenho dos alunos da Rede Municipal, auxiliar na melhora desse desempenho. Assim, é interessante notar que um dos usos da avaliação é premiar os alunos que obtêm o maior desempenho em cada uma das classes/séries. Em um dos anfiteatros da cidade, é realizada a cerimônia de entrega dos certificados aos alunos premiados, que conta com a presença de familiares, de gestores das escolas e de professores. Essa ação tem um objetivo motivacional importante para manutenção do desempenho dos alunos.

Durante as entrevistas, buscou-se conhecer se as escolas também realizavam ações que valorizassem os alunos com alto desempenho. As respostas mostraram que, quando existentes, as ações são isoladas, ficando dependentes muito mais da postura do professor em reconhecer esses alunos e parabenizá-los em sala do que de uma ação planejada.

Contrato de Gestão

As servidoras da Secretaria Municipal de Educação informaram também que os resultados do SAEM são usados para a definição de metas a serem atingidas pelas escolas nos próximos anos, contidas num documento denominado **Contrato de Gestão**⁴⁰. Esse documento define 35 metas, tanto da Rede Municipal quanto da escola. As metas se subdividem em metas quantitativas e qualitativas, tanto para a Educação Infantil quanto para o Ensino Fundamental. As metas quantitativas são: a) desempenho na Prova Brasil; b) desempenho no Pró-Alfa⁴¹; c) desempenho em Língua Portuguesa e em Matemática no SAEM; d) taxa de retenção por classe/série e e) desempenho no IDEB.

Vale destacar que os valores das metas quantitativas citadas, conforme relatado, são obtidos por estimativa. É a equipe gestora da Secretaria que determina

⁴⁰ Modelo de documento fornecido pelas servidoras da Secretaria.

⁴¹ Avaliação externa coordenada pela Secretaria Estadual de Educação, aplicada aos alunos da classe de 8 anos (3º ano) do Ensino Fundamental e que avalia o nível de alfabetização dos alunos dessa faixa etária

as metas da Rede Municipal, enquanto os gestores das escolas (diretor, vice-diretor e coordenador) definem as metas da escola.

Avaliação de desempenho dos professores

Conforme relatado nas entrevistas com os servidores da Secretaria e com alguns professores, as metas de desempenho das escolas no SAEM, definidas no Contrato de Gestão, contam pontos na avaliação de desempenho dos professores. Assim, dentre vários outros aspectos, tais como a assiduidade e a presença em cursos de formação, o bom desempenho dos alunos no SAEM contribui para que os mesmos tenham direito a um 14º salário.

4.3.4. VALIDADE E CONFIABILIDADE

A última etapa da face Aplicação aplicou os procedimentos previstos no delineamento dessa etapa, visando determinar a validade dos itens, as validades de construto e de critério do SAEM, além da sua confiabilidade/fidedignidade, a partir dos dados disponíveis da 2ª edição do SAEM 2007 e da 1ª edição do SAEM 2010.

Conforme define Vianna (1978, p. 182), “a fidedignidade é apenas um aspecto da validade. Se um teste não for fidedigno, não poderá ser válido” Assim, é necessário que se comprove primeiramente a confiabilidade do SAEM para somente depois, caso os resultados se apresentem positivos, comprovar a validade do teste.

Confiabilidade

Dentre os três delineamentos possíveis para a verificação da confiabilidade de uma avaliação, uma única possível de ser utilizada em dados secundários é a que considera uma amostra de sujeitos realizando um mesmo teste em uma única ocasião. Para essa situação, segundo Pasquali (2003), existem duas técnicas

estatísticas possíveis: 1) análise das duas metades e 2) análise pelas técnicas alfa. A principal técnica de coeficiente alfa é a do alfa de Cronbach.

Assim, as duas técnicas citadas foram aplicadas nos dados do SAEM. Para tanto, foram escolhidas as avaliações da 2ª edição do SAEM 2007, usadas nos testes de validade de critério, e as avaliações da 1ª edição de 2010, a mais recente à época. Em cada uma dessas edições, foram utilizados os dados das avaliações a partir das classes de 9 anos, uma vez que, a partir dessa classe, há 30 questões para cada classe/série. Em relação aos sujeitos, optou-se por considerar os desempenhos das turmas de alunos e não o dos alunos individualmente. Essa opção objetivou trabalhar com dados de variabilidade maior do que somente 0 e 1, que ocorre ao se trabalhar com o desempenho individual do aluno em cada questão.

A análise das duas metades, como o próprio nome mostra, necessita que se dividam as questões em duas partes equivalentes. Assim, para cada classe/série, foram impressos pequenos papéis com o número da questão e a disciplina envolvida na mesma. Agruparam-se os papéis de cada disciplina, viraram-nos e foram sorteadas as questões que iriam compor cada um dos dois testes. Desse modo, cada teste ficou com a mesma quantidade de questões de Ciências, Geografia e História. As disciplinas de Língua Portuguesa e Matemática têm 9 questões por avaliação. Assim, para o sorteio, guardou-se o último papel de cada uma delas e fez-se o sorteio dentre esses dois papéis, definindo qual das duas metades ficaria com uma questão a mais de Língua Portuguesa ou de Matemática.

Feita a separação das questões em duas metades, procedeu-se o cálculo da correlação entre os desempenhos médios de cada turma nos dois testes. Conforme explica Pasquali (2003), para a técnica da análise das duas metades, há a necessidade de se aplicar a correção de Spearman-Brown, pois a correlação se baseia somente na metade da escala.

A segunda técnica para análise de confiabilidade é através do cálculo do coeficiente alfa. Para tanto, foi usado o coeficiente alfa de Cronbach. Essa técnica visa analisar a consistência interna do teste a partir da consistência interna dos itens. Segundo Pasquali (2003, p. 205), “a lógica dessa suposição é a de que quanto menos variabilidade um mesmo item produz numa amostra de sujeitos, menos erros

ele provoca”. Os resultados dos cálculos do coeficiente de alfa de Cronbach para todas as avaliações analisadas estão apresentados na Tabela 1.

A análise da confiabilidade do SAEM, a partir da determinação do coeficiente de confiabilidade, trouxe resultados próximos para as duas técnicas aplicadas. Os resultados para as duas avaliações analisadas, uma de 2007 e outra de 2010, também apresentaram resultados muito próximos. Entretanto, em relação às diferentes classes/séries analisadas, nota-se um decréscimo dos coeficientes à medida que a seriação aumenta. Esse fenômeno ocorreu nas duas avaliações e nas duas técnicas utilizadas.

Tabela 1. Resultados do cálculo do coeficiente de confiabilidade

Classe/Série	Nº de turmas	Coeficiente de confiabilidade	
		Correlação (teste das duas metades)	Coeficiente alfa (Cronbach)
2ª avaliação do SAEM 2007			
Classe de 9 anos	63	0,9486	0,9431
Classe de 10 anos	79	0,9415	0,9219
5ª série	78	0,8733	0,8568
6ª série	66	0,8323	0,7644
7ª série	52	0,7261	0,7333
8ª série	54	0,7049	0,6575
1ª avaliação do SAEM 2010			
Classe de 9 anos	66	0,960091	0,9569
Classe de 10 anos	84	0,956217	0,9434
5ª série	85	0,857446	0,8149
6ª série	69	0,815723	0,8111
7ª série	60	0,687033	0,7025
8ª série	44	0,606842	0,6661

A verificação da confiabilidade da avaliação passa pela determinação do valor esperado do Coeficiente de confiabilidade. Segundo Vianna (1978), a fidedignidade esperada de um teste depende do número de itens e do número de alternativas que compõe cada item. A expressão que calcula o coeficiente esperado está apresentada na equação 1 (VIANNA, 1978, p. 160).

$$r_{xx} = \frac{k}{k-1} \cdot \left[1 - \frac{9 \cdot (N+1)}{k \cdot (N-1)} \right]$$

Equação 1. Coeficiente de confiabilidade esperado

Sendo: r_{xx} = coeficiente de confiabilidade;
k = número de itens do teste;
N = número de alternativas dos itens;

Para o caso do SAEM, as avaliações analisadas possuíam sempre 30 questões de 4 alternativas cada. Assim, aplicados à equação 1, chega-se a um valor de **Coeficiente de Fidedignidade Esperado de 0,5172**.

Como pôde ser verificado nos valores apresentados na Tabela 1, todas as avaliações estudadas apresentaram coeficientes de confiabilidade/fidedignidade superiores ao valor esperado.

Em complemento, é interessante observar a queda do coeficiente de confiabilidade com o crescimento das classes/séries. Um dos motivos que pode ser relacionado a essa queda de desempenho é a mudança de envolvimento dos alunos mais velhos com o seu desempenho na avaliação, conforme cita Vianna (1978), esse é um dos fatores que impactam na confiabilidade. Os diretores e os professores entrevistados também afirmaram que há um declínio do envolvimento dos alunos em relação ao seu rendimento na avaliação, uma vez que o seu desempenho tem pouca influência na sua aprovação escolar.

Assim, conforme os resultados do coeficiente de confiabilidade calculados, foi comprovada a confiabilidade do SAEM, devendo-se buscar estratégias para que esse coeficiente melhore para as séries finais do Ensino Fundamental.

Validade dos itens

Conforme define Pasquali (2003), validade do item se refere à relação do mesmo com aquilo que se pretende medir. Assim, um item é válido quando tem alta correlação com um critério⁴². Pasquali (2003) afirma que a validade do item pode ser expressa pelo coeficiente de discriminação do item em relação ao teste, uma vez que, quanto maior for esse coeficiente, mais o item foi capaz de discriminar os sujeitos que dominam o construto avaliado. Uma das formas de se obter o coeficiente de discriminação é a partir da análise da diferença entre as médias dos grupos superior e inferior formados pela separação dos sujeitos em quartis. A técnica estatística que compara as médias de dois grupos de dados é chamada de teste T (ANDERSON; SWEENEY; WILLIAMS, 2007).

Assim, a validade dos itens do SAEM foi verificada a partir dos dados da 2ª avaliação de 2007 e a 1ª avaliação de 2010, como já foi usado na verificação da confiabilidade. Em relação às classes/séries, foram usados os dados da 4ª e 8ª série nas duas edições citadas. Foram considerados como sujeitos todos os alunos de cada classe/série analisada e não as turmas, pois a discriminação fica mais bem representada quanto mais sujeitos houver.

O teste T foi aplicado com o auxílio do software estatístico SPSS. Após a importação dos dados dos alunos de uma das classes/séries e de seus desempenhos em cada uma das questões, foi criada uma variável de nome *Total* para armazenar o desempenho total do aluno na avaliação. O software oferece a opção de cálculo dos quartis da variável *Total*. Assim, conhecidos os valores dos quartis, criou-se uma nova variável, *Grupo*, para armazenar quais sujeitos pertenciam ao grupo superior e ao grupo inferior.

Definidos os grupos, procedeu-se a comparação das médias de cada uma das questões nos dois grupos-critério, ou seja, o teste T. A validade do item pode ser comprovada pelo nível de significância dos resultados. Como todos os cálculos foram realizados com um nível de confiança de 95%, os itens foram considerados válidos caso o nível de significância (p-valor) fosse inferior a 5% (0,05).

⁴² Segundo Pasquali (2003), no caso da Teoria Clássica dos Testes (TCT) o critério é um teste paralelo ao teste em que o item analisado está inserido.

Foram analisados de 116 itens relativos aos desempenhos dos alunos da 4ª e da 8ª série da 2ª avaliação de 2007 e da 1ª avaliação de 2010. Desses, somente dois itens – a sexta questão de Língua Portuguesa e a nona questão de Matemática da 4ª série, na 2ª avaliação de 2007– não apresentaram diferença significativa entre os dois grupos, considerando-se um nível de significância inferior a 0,05. A questão P6 apresentou um p-valor de 0,115 e a questão M9 apresentou um p-valor de 0,174. Todas as outras 114 questões apresentaram um p-valor inferior a 0,05. Dessa forma, devida a alta proporção de itens válidos, foi comprovada a validade dos itens das avaliações do SAEM.

Validade de Construto

Vianna (1978) e Pasquali (2003) afirmam que uma das formas de comprovar se uma avaliação mede efetivamente o que se propõe a medir, ou seja, a sua validade de construto é correlacionar os seus resultados com os de outro teste que também meça o mesmo construto. A Prova Brasil é um teste em que, da mesma forma que o SAEM, se propõe medir o desempenho acadêmico dos alunos da rede pública. Desse modo, ao se determinar a correlação entre os resultados dos alunos nos dois testes com a comprovação da validade de critério que será apresentada a seguir, comprovou-se, também, a validade de construto do SAEM.

Entretanto, devido a deficiências apresentadas em todas as técnicas de comprovação da validade de construto (PASQUALI, 2003), é importante que a mesma seja comprovada por mais de uma delas.

Dentre as outras técnicas de validação do construto, está a validade convergente-discriminante. Nessa técnica, tem-se que a validade é comprovada a partir da determinação de dois aspectos: 1) o teste deve correlacionar com variáveis as quais o construto deveria medir (validade convergente) e 2) o teste de **não** se correlacionar com variáveis as quais ele deveria diferir (validade discriminante) (PASQUALI, 2003).

Para tanto, usando-se mais uma vez o conceito dos grupos-critério, ou seja, de que a validade pode ser obtida ao se verificar se o teste é capaz de discriminar

grupos-critério em termos do construto medido, foi estudada a capacidade de o SAEM avaliar os conhecimentos dos alunos em relação às disciplinas envolvidas na avaliação: Ciências, Geografia, História, Língua Portuguesa e Matemática.

Como nos testes anteriores, os dados utilizados foram os acertos de todos os alunos da 4ª e da 8ª série na 2ª avaliação de 2007 e na 1ª avaliação de 2010. Contou-se, mais uma vez, com o auxílio do software estatístico SPSS. Importados os dados, foram criadas variáveis para armazenar as somas dos acertos totais e de cada uma das disciplinas. Além disso, para gerar os grupos-critério sem a influência da própria disciplina, foram criadas outras variáveis. Essas variáveis armazenaram a soma dos acertos dos alunos no teste, excluindo-se as questões da própria disciplina. A variável *Grupo* foi usada, mais uma vez, para armazenar quais sujeitos pertenciam ao grupo superior e inferior na separação por quartis.

Definidas as variáveis que foram usadas, partiu-se para a aplicação dos testes. Sobre a validade convergente, para cada uma das cinco disciplinas, foi feita a definição dos grupos superior e inferior, a partir do campo da soma dos acertos sem a disciplina, e aplicada uma comparação das médias – teste T – entre esses dois grupos. A hipótese de pesquisa era a de que, se as médias apresentassem diferenças significativas, o teste seria capaz de avaliar o construto daquela disciplina, ou seja, o teste era capaz de aferir os conhecimentos dos alunos na disciplina testada.

Foram feitos testes nas classes/séries selecionadas, e, em todos os casos, obteve-se um nível de significância menor que 0,05. Isso representa que as médias dos dois grupos – superior e inferior – são diferentes significativamente. Assim, ficou comprovada a validade convergente da avaliação.

Em complemento, foi feita a validade discriminante do teste, o que significa determinar se os construtos, nesse caso as disciplinas, possuem baixa correlação entre si. Assim, a partir das variáveis dos desempenhos das disciplinas, calculou-se o coeficiente de correlação desses campos entre si. Os resultados dos coeficientes de correlação das classes/séries estudados estão apresentados nas tabelas 2, 3, 4 e 5.

Tabela 2. Coeficientes de correlação entre os desempenhos por disciplina da Classe de 10 anos na 2ª avaliação de 2007

	Sport	Smat	Scien	Sgeo	Shis
Sport	1	0,378	0,365	0,446	0,340
Smat	0,378	1	0,249	0,325	0,220
Scien	0,365	0,249	1	0,319	0,207
Sgeo	0,446	0,325	0,319	1	0,260
Shis	0,340	0,220	0,207	0,260	1

Fonte dos dados: Relatório da 2ª avaliação do SAEM 2007 - Colégio Cenecista Dr. José Ferreira

Tabela 3. Coeficientes de correlação entre os desempenhos por disciplina da 8ª série na 2ª avaliação de 2007

	Sport	Smat	Scien	Sgeo	Shis
Sport	1	0,060	0,117	0,107	0,122
Smat	0,060	1	0,136	0,102	0,081
Scien	0,117	0,136	1	0,052	0,026
Sgeo	0,107	0,102	0,052	1	0,058
Shis	0,122	0,081	0,026	0,058	1

Fonte dos dados: Relatório da 2ª avaliação do SAEM 2007 - Colégio Cenecista Dr. José Ferreira

Tabela 4. Coeficientes de correlação entre os desempenhos por disciplina da Classe de 10 anos na 1ª avaliação de 2010

	Sport	Smat	Scien	Sgeo	Shis
Sport	1	0,380	0,324	0,348	0,342
Smat	0,380	1	0,346	0,497	0,508
Scien	0,324	0,346	1	0,339	0,321
Sgeo	0,348	0,497	0,339	1	0,479
Shis	0,342	0,508	0,321	0,479	1

Fonte dos dados: Relatório da 1ª avaliação do SAEM 2010 - Colégio Cenecista Dr. José Ferreira

Tabela 5. Coeficientes de correlação entre os desempenhos por disciplina da 8ª série na 1ª avaliação de 2010

	Sport	Smat	Scien	Sgeo	Shis
Sport	1	0,156	0,136	0,242	0,148
Smat	0,156	1	0,117	0,048	0,063
Scien	0,136	0,117	1	0,063	0,064
Sgeo	0,242	0,048	0,063	1	0,131
Shis	0,148	0,063	0,064	0,131	1

Fonte dos dados: Relatório da 1ª avaliação do SAEM 2010 - Colégio Cenecista Dr. José Ferreira

O valor esperado do coeficiente de correlação para que haja validade discriminante deve ser menor do que 0,7. Tal parâmetro se estabelece a partir do conceito de coeficiente de determinação, que é o quadrado do coeficiente de correlação. Assim, uma correlação de 0,7 apresenta uma determinação de 49%, ou seja, há uma correspondência de 49% entre os dados correlacionados.

Aplicada a correlação entre os desempenhos nas disciplinas dos alunos das classes/séries selecionadas, todos os coeficientes de correlação encontrados foram inferiores a 0,7. Observou-se, dessa forma, baixa correlação entre os desempenhos nas disciplinas, o que significa que elas representam construtos diferentes dentro da avaliação, comprovando a validade discriminante.

Com a aplicação da técnica de validade convergente-discriminante, além da comprovação por correlação com outros testes, confirmou-se a validade de construto do SAEM.

Validade de Critério

O terceiro tipo de validade a ser testada estatisticamente é a validade de critério. Quando existem testes comprovadamente válidos, estes podem ser usados como referência de critério. Assim, o desempenho das escolas da rede municipal de Uberaba na Prova Brasil pôde ser usado como critério para validade dos resultados dessas escolas no SAEM.

Para tanto, foi medida a correlação dos desempenhos das escolas nessas duas avaliações. Conforme explana Anderson, Sweeney e Williams (2007), o coeficiente de correlação de Pearson entre populações – e não amostrais – é obtido pelo quociente entre a covariância das populações e os desvios padrão de cada população. A Equação 2 apresenta a lei matemática para essa operação.

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \cdot \sigma_y}$$

Equação 2. Coeficiente de correlação de Pearson

Sendo: ρ_{xy} = coeficiente de correlação da população;
 σ_{xy} = covariância populacional;
 σ_x = desvio padrão da população para x;
 σ_y = desvio padrão da população para y;

A covariância entre as populações é obtida pelo quociente entre o somatório dos produtos das diferenças entre cada medida e sua média, de cada população e o número de elementos da população, cuja formulação matemática está descrita na equação 3.

$$\sigma_{xy} = \frac{\sum (x_i - \mu_x) \cdot (y_i - \mu_y)}{N}$$

Equação 3. Covariância populacional

Sendo: σ_{xy} = covariância populacional;
 μ_x = média da população para x;
 μ_y = média da população para y;
 N = número de elementos da população;

Os dados para a medida da correlação foram obtidos a partir dos desempenhos das escolas na Prova Brasil em 2007, coletados do site do INEP (www.inep.gov.br) e dos dados do desempenho médio das mesmas escolas no SAEM de novembro de 2007, ambos referentes à 4ª série (5º ano) e à 8ª série (9º ano) do Ensino Fundamental. Assim, como algumas escolas não possuem

desempenho na Prova Brasil e não são todas que possuem as duas séries estudadas, das 27 escolas de Ensino Fundamental existentes no município de Uberaba-MG, fizeram parte do estudo 22 escolas da 4ª série e 21 escolas da 8ª série do Ensino Fundamental.

O banco de dados com os desempenhos na Prova Brasil apresenta os resultados de todas as escolas participantes da avaliação de 2007. Filtradas as escolas da rede municipal de Uberaba, usou-se, dentre os vários desempenhos que são apresentados, somente a nota média padronizada de cada escola, que possui uma escala de 0 a 10.

Tabela 6. Desempenho das escolas na Prova Brasil em 2007. 4ª série do Ensino Fundamental

	Nome da Escola	Nota Média Padronizada (N)
1	EM PROFA ESTER LIMIRIO BRIGAGAO	4,45
2	EM PROF ANISIO TEIXEIRA	4,00
3	EM MONTEIRO LOBATO	4,74
4	EM MARIA LOURENCINA PALMEIRO	4,71
5	EM PROF JOSE MACCIOTTI	4,88
6	EM JOUBERT DE CARVALHO	4,87
7	EM ADOLFO BEZERRA DE MENEZES	4,96
8	EM REIS JUNIOR	5,01
9	EM PROF JOSE GERALDO GUIMARAES	4,74
10	EM ARTUR DE MELO TEIXEIRA	4,95
11	EM SEBASTIAO ANTONIO LEAL	4,80
12	EM PADRE EDDIE BERNARDES	5,05
13	EM PROFA OLGA DE OLIVEIRA	5,31
14	EM PROFA GENI CHAVES	5,00
15	EM SANTA MARIA	5,42
16	EM PROFA NIZA MARQUEZ GUARITA	5,65
17	EM MADRE MARIA GEORGINA	5,16
18	EM GASTAO MESQUITA FILHO	5,16
19	EM NORMA SUELY BORGES	5,24
20	EM BOA VISTA	5,81
21	EM UBERABA	5,52
22	EM FREI EUGENIO	5,81

Fonte dos dados: “Divulgação_4serie_Escolas.xls” disponível em:
http://portalideb.inep.gov.br/index.php?option=com_content&view=article&id=43&Itemid=7

Tabela 7. Desempenho das escolas na Prova Brasil em 2007. 8ª série do Ensino Fundamental

	Nome da Escola	Nota Média Padronizada (N)
1	EM PROF JOSE MACCIOTTI	4,10
2	EM SEBASTIAO ANTONIO LEAL	3,99
3	EM JOUBERT DE CARVALHO	4,70
4	EM ARTUR DE MELO TEIXEIRA	4,51
5	EM PROF JOSE GERALDO GUIMARAES	4,73
6	EM PADRE EDDIE BERNARDES	4,34
7	EM MARIA LOURENCINA PALMEIRO	4,66
8	EM PROFA ESTER LIMIRIO BRIGAGAO	4,16
9	EM PROFA STELA CHAVES	4,96
10	EM NORMA SUELY BORGES	4,67
11	EM PROF ANISIO TEIXEIRA	4,60
12	EM PROFA NIZA MARQUEZ GUARITA	5,25
13	EM MONTEIRO LOBATO	5,07
14	EM GASTAO MESQUITA FILHO	4,90
15	EM SANTA MARIA	5,26
16	EM ADOLFO BEZERRA DE MENEZES	5,00
17	EM BOA VISTA	5,47
18	EM PROFA GENI CHAVES	5,29
19	EM MADRE MARIA GEORGINA	5,15
20	EM UBERABA	5,93
21	EM FREI EUGENIO	6,13

Fonte dos dados: “Divulgação_8serie_Escolas.xls” disponível em:
http://portalideb.inep.gov.br/index.php?option=com_content&view=article&id=43&Itemid=7

Em 2007, foram realizadas duas avaliações do SAEM. Para a comparação com os dados da Prova Brasil, foi calculada a média do desempenho das escolas nas duas avaliações desse ano. Os resultados são percentuais, variam de 0 a 100 e estão expressos nas Tabelas 08 e 09.

Tabela 8. Desempenho médio das escolas no SAEM em 2007. 4ª série do Ensino Fundamental

	Nome da Escola	Nota Média
1	EM PROFA ESTER LIMIRIO BRIGAGAO	34,11
2	EM PROF ANISIO TEIXEIRA	34,19
3	EM MONTEIRO LOBATO	36,14
4	EM MARIA LOURENCINA PALMEIRO	37,16
5	EM PROF JOSE MACCIOTTI	37,32
6	EM JOUBERT DE CARVALHO	38,34
7	EM ADOLFO BEZERRA DE MENEZES	38,70
8	EM REIS JUNIOR	39,06

	Nome da Escola	Nota Média
9	EM PROF JOSE GERALDO GUIMARAES	39,63
10	EM ARTUR DE MELO TEIXEIRA	40,85
11	EM SEBASTIAO ANTONIO LEAL	42,72
12	EM PADRE EDDIE BERNARDES	42,82
13	EM PROFA OLGA DE OLIVEIRA	43,10
14	EM PROFA GENI CHAVES	43,36
15	EM SANTA MARIA	44,20
16	EM PROFA NIZA MARQUEZ GUARITA	44,79
17	EM MADRE MARIA GEORGINA	45,87
18	EM GASTAO MESQUITA FILHO	46,48
19	EM NORMA SUELY BORGES	47,02
20	EM BOA VISTA	48,37
21	EM UBERABA	50,80
22	EM FREI EUGENIO	52,99

Fonte dos dados: “Relatório da 1ª avaliação do SAEM 2007” e “Relatório da 2ª avaliação do SAEM 2007”, fornecidos pelo Colégio Cenecista Dr. José Ferreira

Tabela 9. Desempenho médio das escolas no SAEM em 2007. 8ª série do Ensino Fundamental

	Nome da Escola	Nota Média
1	EM PROF JOSE MACCIOTTI	36,60
2	EM SEBASTIAO ANTONIO LEAL	37,87
3	EM JOUBERT DE CARVALHO	38,60
4	EM ARTUR DE MELO TEIXEIRA	39,65
5	EM PROF JOSE GERALDO GUIMARAES	40,68
6	EM PADRE EDDIE BERNARDES	40,90
7	EM MARIA LOURENCINA PALMEIRO	40,90
8	EM PROFA ESTER LIMIRIO BRIGAGAO	41,22
9	EM PROFA STELA CHAVES	41,25
10	EM NORMA SUELY BORGES	41,43
11	EM PROF ANISIO TEIXEIRA	41,77
12	EM PROFA NIZA MARQUEZ GUARITA	42,26
13	EM MONTEIRO LOBATO	42,68
14	EM GASTAO MESQUITA FILHO	42,75
15	EM SANTA MARIA	42,91
16	EM ADOLFO BEZERRA DE MENEZES	43,78
17	EM BOA VISTA	44,13
18	EM PROFA GENI CHAVES	44,40
19	EM MADRE MARIA GEORGINA	46,59
20	EM UBERABA	50,70
21	EM FREI EUGENIO	51,08

Fonte dos dados: “Relatório da 1ª avaliação do SAEM 2007” e “Relatório da 2ª avaliação do SAEM 2007”, fornecidos pelo Colégio Cenecista Dr. José Ferreira

Coletados os dados e com o auxílio da planilha de dados Microsoft Excel, foi possível calcular as médias e os desvios padrão de cada população, possibilitando, dessa forma, o cálculo da covariância populacional e o coeficiente de correlação. Os resultados obtidos serão apresentados nas Tabelas 10 e 11.

Após a aplicação dos cálculos citados anteriormente, obteve-se o coeficiente de correlação de Pearson para as duas séries estudadas, conforme está ilustrado nas tabelas a seguir, sendo uma para cada série estudada.

Tabela 10. Resultados do cálculo da correlação de Pearson para a 4ª série do Ensino Fundamental

4ª série do Ensino Fundamental	Prova Brasil	SAEM
Média	5,06	42,18
Desvio Padrão	0,42	5,04
Covariância Populacional	1,850	
Coeficiente de correlação	0,872	

Tabela 11. Resultados do cálculo da correlação de Pearson para a 8ª série do Ensino Fundamental

8ª série do Ensino Fundamental	Prova Brasil	SAEM
Média	4,90	42,48
Desvio Padrão	0,54	3,51
Covariância Populacional	1,546	
Coeficiente de correlação	0,815	

Os desempenhos das escolas em cada uma das avaliações foram usados para a construção dos gráficos de dispersão mostrados na Figuras 10 e 11 a seguir.

Figura 10. Gráfico de dispersão – 4ª série do Ensino Fundamental

Figura 11. Gráfico de dispersão – 8ª série do Ensino Fundamental

Esses gráficos de dispersão mostram que os dois coeficientes de correlação foram positivos, determinando uma relação diretamente proporcional entre os dois desempenhos. Pode-se notar, também, que os desempenhos das escolas estão

próximos e não existem *outliers*, o que significaria discrepância entre o desempenho de alguma escola em um dos exames.

Apesar de possuírem diferenças metodológicas, observou-se, através da comparação de desempenhos por escola nos dois exames, alta correlação obtida pelo método de Pearson. Os valores de correlação obtidos foram de **0,872** para a 4ª série do Ensino Fundamental e de **0,815** para a 8ª série do Ensino Fundamental.

Conforme cita Pasquali (2003, p. 187)

[...] se houver um teste comprovadamente válido para a medida de algum traço latente, ela certamente pode servir de critério para a validação de um novo teste. Espera-se, nesse caso, que a correlação do novo teste seja elevada, de pelo menos 0,75"

Desse modo, conclui-se que foi comprovada a validade de critério do sistema da avaliação do SAEM.

Encerrada a exposição dos resultados da face Aplicação, pretendeu-se, nessa etapa, apresentar os procedimentos para a aplicação, correção e entrega dos resultados de uma avaliação do SAEM na Rede Municipal de Ensino de Uberaba, atingindo-se, assim, o objetivo específico **(b)** deste trabalho. Ao validar o SAEM quanto ao seu construto e seu critério, além de comprovar a confiabilidade dos seus resultados, atendeu-se ao proposto no objetivo **(c)** deste trabalho.

4.4. FACE EXPANSÃO

A face Expansão se constitui como a face “futuro” do trabalho, porque apresentou os caminhos para a implementação de melhorias nos procedimentos do SAEM e relacionou possíveis novos usos dos resultados das avaliações, a partir da pesquisa e das entrevistas realizadas.

4.4.1. PROPOSTAS DE MUDANÇAS

Seguindo a ordem dada neste capítulo de resultados, a apresentação das propostas de mudanças se iniciará pela análise do SAEM em relação aos padrões de qualidade da JCSEE. Em seguida, serão expostas as propostas de inserção de algumas das características encontradas na Prova Brasil/SAEB e, finalmente, as propostas de mudanças coletadas durante as entrevistas.

4.4.1.1. PROPOSTAS DE ADEQUAÇÃO AOS PADRÕES DO JCSEE

A análise realizada no item 4.1 apontou alguns itens em que os procedimentos do SAEM foram considerados parcialmente adequados em relação aos padrões de qualidade de uma avaliação, segundo o *Joint Committee Standards for Educational Evaluation* (JOINT, 2011). Portanto, é importante que se destaquem esses pontos, como está no Quadro 16, que apresenta a relação de padrões/Standards não atendidos ou parcialmente atendidos, as suas descrições e uma proposta para a adequação ao mesmo.

Standards	Descrição	Proposta
<i>Standards de Utilidade</i>		
<i>Follow-up</i> (acompanha- mento)	A avaliação deve ter procedimentos de acompanhamento do entendimento das informações pelos usuários.	Oficialmente, não existem ações previstas para o acompanhamento do entendimento das informações, conforme preconiza o item. Entretanto, tanto a Secretaria quanto os diretores têm opções de comunicação com os representantes do Colégio para esclarecimentos de quaisquer dúvidas. Mesmo assim, é importante formalizar esse acompanhamento do entendimento, seja pela criação de formulários ou mesmo a elaboração de um material (Manual) explicativo para os pais entenderem os resultados.

Standards	Descrição	Proposta
Standards de Rigor/Exatidão		
Análise do contexto	As variáveis contextuais devem ser identificadas e levadas em conta na análise de desempenho dos alunos.	O padrão foi considerado parcialmente adequado, uma vez que há a consideração somente dos alunos com deficiência de aprendizagem, cujos desempenhos não são considerados nos cálculos dos desempenhos da turma, da classe/série e da Rede. A proposta é que se repense a utilidade dos relatórios que fazem as comparações entre as escolas, classificando-as em função de um desempenho da escola que é calculado “misturando-se” todos os desempenhos de todas as classes/séries.
Procedimentos Documentados	Os procedimentos para execução da avaliação devem estar descritos, de modo a facilitar a sua explicação e justificativa.	Foi detectado que algumas etapas possuem documentação apropriada. Contudo, existem várias outras etapas que não estão devidamente documentadas. Assim, deve-se pensar em produzir um único documento que contenha as descrições de todas as etapas, relatando as experiências vividas e as alterações implementadas ao longo do tempo, possibilitando que novos gestores possam desenvolver eficientemente o SAEM.
Meta-avaliação	Os procedimentos da avaliação devem ser periodicamente analisados a partir dessas ou outras normas, para que possíveis erros sejam prontamente identificados e corrigidos.	Em quatro anos da aplicação ininterrupta de avaliações externas à Rede Municipal de Ensino de Uberaba, o presente trabalho é a primeira meta-avaliação realizada no SAEM. A proposta é que se estimule a inserção de novos pesquisadores interessados em contribuir com a construção de um sistema de ensino cada vez mais eficiente.

Quadro 16. Propostas de adequação do SAEM aos Standards

O atendimento às propostas apresentadas no Quadro 16 possibilitará um grau de maturidade importante ao SAEM, principalmente visando a sua continuidade após as inevitáveis trocas de equipes gestoras e a criação de um processo contínuo de identificação de melhorias.

4.4.1.2. CARACTERÍSTICAS DA PROVA BRASIL/SAEB PARA SEREM IMPLANTADAS NO SAEM

Nos resultados da face Teórica, foi feita também uma comparação entre as características do SAEM e da Prova Brasil/SAEB. O objetivo era o de identificar características da Prova Brasil/SAEB que podem ser implantadas pelo SAEM. As características consideradas importantes a sua implantação estão apresentadas a seguir.

Matriz de Referência

Apesar de o SAEM possuir uma relação de descritores dos conteúdos, elaborados ao início de cada ano letivo, é importante que se tenha uma Matriz de Referência efetiva, para se manter constante por um bom período de tempo, a fim de possibilitar as efetivas comparações de desempenho entre distintas avaliações. Para tanto, como relatado por Pestana (1998) na criação das Matrizes do SAEM, é importante que haja a participação de professores, coordenadores pedagógicos e da equipe do departamento pedagógico na elaboração desses descritores, de modo que o seu resultado seja consensual e reflita a realidade local.

Uso da Teoria de Resposta ao Item (TRI)

A princípio, crê-se que a maior diferença entre os dois modelos de avaliação – e, por isso, deverá ser a mais analisada – é a teoria psicométrica adotada pelos sistemas. Enquanto o SAEM usa a Teoria Clássica dos Testes (TCT), a Prova Brasil adota a Teoria de Resposta ao Item (TRI). Há várias vantagens em passar o SAEM para a TRI. Inicialmente, somente com a aplicação da TRI permite a comparação entre resultados de distintas edições, característica muito solicitada pela equipe da Secretaria.

Em relação à construção de um banco de itens, a característica temática do SAEM não permite tal fato. Entretanto, o método mais seguro para a comparação dos resultados é a ancoragem de itens, ou seja, a repetição de algumas questões entre os dois testes (PASQUALI, 2003). A solução deve ser mais bem equacionada, mas um caminho é fazer a ancoragem não com questões iguais, mas sim com questões que tenham as mesmas características (índice de dificuldade e índice de discriminação) e referentes a um mesmo descritor. Deve-se estudar bem essa situação, para que os pressupostos da TRI sejam compatíveis com a avaliação temática.

Outro desafio para a implantação da TRI talvez seja absorção dos custos provenientes da implantação. Nesse sentido, deve-se trabalhar na expansão do SAEM para outras redes de ensino.

Escala de Proficiência

No SAEM, o desempenho do aluno é obtido pela razão entre o número de questões certas em relação ao total de questões propostas, o que vale também para o cálculo do desempenho da escola. Já na Prova Brasil, o resultado das escolas é apresentado em uma escala de proficiência. A diferença de escalas de resultados atrapalha a comparação entre as avaliações, conforme relatado por uma das servidoras entrevistadas. A adoção de uma escala de proficiência permite a comparação de desempenhos entre distintas séries. Assim, a fim de possibilitar tais comparações, é importante que o SAEM passe a apresentar os seus resultados na mesma escala usada pela Prova Brasil. Entretanto, deve-se lembrar que a escala da Prova Brasil está definida apenas para a 4ª série (5º ano) e a 8ª série (9º ano) do Ensino Fundamental e somente para as disciplinas de Língua Portuguesa e Matemática. Assim, será necessário ampliar a escala para adaptá-la às outras séries e disciplinas do SAEM.

Aplicação de questionários contextuais

Como já detectado na análise dos padrões de qualidade, essa é mais uma característica da Prova Brasil/SAEB que deve ser incorporada ao SAEM, pois o levantamento dos fatores intra e extraescolares que afetam o desempenho escolar deve ser do conhecimento da Secretaria Municipal para pautar a condução das políticas de sua responsabilidade.

4.4.1.3. PROPOSTAS COLETADAS NAS ENTREVISTAS

Encerrada a relação de características da Prova Brasil/SAEB passíveis de serem transplantadas para o SAEM, passa-se a apresentar algumas sugestões coletadas durante as entrevistas realizadas para esse trabalho, relatadas a seguir.

Sistematização do valor da pontuação

Foi notável, apesar de as informações colhidas pelas entrevistas não representarem a opinião da população, a valorização dada pelos alunos ao SAEM pela avaliação valer como uma das avaliações regulares. Assim, é recomendável que a Secretaria formalize essa prática, normatizando a quantidade de pontos distribuídos para que a avaliação tenha um valor efetivo e não simbólico.

Inclusão dos níveis de desempenho/Novos relatórios

Até 2010, conforme apresentados no Anexo B, os relatórios do SAEM apresentavam o desempenho do aluno, de sua turma, da escola e da rede a partir razão entre o número de acertos e o número de questões existentes, o que gera um

cálculo de média. Matematicamente, o cálculo da média oculta a informação da distribuição dos desempenhos. Por exemplo, em uma turma de alunos de uma determinada série, se todos os alunos tiverem um desempenho de 50%, o desempenho médio da turma será de 50%. Por outro lado, se em outra turma da mesma série a metade dos alunos tirar zero e a outra metade tirar 100%, o desempenho médio da turma será os mesmos 50%. Apesar de terem o mesmo desempenho médio, as turmas têm distribuição de desempenhos muito diferentes.

A equipe da Diretoria Pedagógica, durante as entrevistas para este trabalho, externou a vontade de se melhorar a informação dos desempenhos calculados para as turmas, as séries, as escolas e a Rede, complementando-os com algum outro dado que represente essa variabilidade.

A solução encontrada foi criar uma variável denominada **nível de desempenho**, que é uma escala de cinco níveis, em que o desempenho do aluno é classificado em um dos cinco níveis, conforme o seu valor. Assim, o primeiro passo foi criar uma tabela contendo as faixas de desempenho para cada uma das séries/anos, como pode ser visto na Tabela 12.

Tabela 12. Faixas de desempenho do SAEM 2011

	Muito Crítico	Crítico	Básico	Adequado	Avançado
1º ano	Abaixo de 40%	A partir de 40% até menor que 50%	A partir de 50% até menor que 70%	A partir de 70% até menor que 95%	A partir de 95%
2º ano					
3º ano					
4º ano	Abaixo de 30%	A partir de 30% até menor que 40%	A partir de 40% até menor que 60%	A partir de 60% até menor que 80%	A partir de 80%
5º ano					
6º ano	Abaixo de 25%	A partir de 25% até menor que 35%	A partir de 35% até menor que 50%	A partir de 50% até menor que 75%	A partir de 75%
7º ano					
8º ano					
9º ano					

Fonte dos dados: “Relatório da 1ª avaliação do SAEM 2011” – Colégio Cenecista Dr. José Ferreira

A definição das faixas de valores para cada uma das séries/anos levou em conta as distribuições de desempenhos dos alunos nas edições anteriores do SAEM.

Desse modo, além do desempenho do aluno, o SAEM 2011 passou a informar o seu nível de desempenho. A partir desse novo dado, tornou possível calcular a distribuição porcentual de alunos em função da quantidade de alunos classificados em cada um dos níveis, atendendo ao que havia sido solicitado pela equipe da Secretaria.

Consequentemente, houve a necessidade de se reformularem vários dos relatórios do SAEM. Além disso, foi possível criar novos relatórios apresentando a distribuição dos níveis de desempenho das turmas, da série, da escola e da Rede. A seguir, são apresentados os relatórios do SAEM que sofreram modificações e os novos relatórios.

Relatórios remodelados:

- **Relatório de Desempenho Individual:** conforme pode ser visualizado no Anexo C1, na região em que é apresentado o desempenho do aluno, foi criado um campo para informar o nível de desempenho alcançado por ele.
- **Relatórios de Classificação:** todos os relatórios de classificação, da rede ou da escola, geral ou por disciplina, apresentavam somente o número de acertos de cada um dos alunos. Assim, foram incluídas as colunas que informam o desempenho do aluno e o nível de desempenho. Um exemplo de um dos novos relatórios de classificação está disponível no Anexo C2.

Relatórios novos:

- **Relatório de Distribuição de Níveis da escola:** para cada série/ano de ensino, as escolas recebem um relatório da distribuição porcentual dos níveis de desempenho, por turma e da série/ano e da Rede, possibilitando a análise do desempenho das turmas, em complemento ao Relatório de Desempenho da Escola. Um exemplo do novo relatório está disponível no Anexo C3.
- **Relatório de Distribuição de Níveis Geral:** o relatório apresenta as duas informações de desempenho: desempenho médio e distribuição de níveis de desempenho da Rede Municipal para cada uma das séries/anos

avaliados. Anteriormente, junto aos desempenhos de cada série/ano, era calculado e apresentado o desempenho médio da Rede, envolvendo e “misturando”, em seu cálculo, todas as séries/anos. A partir desse novo relatório, o desempenho da Rede foi separado em dois índices: um para as séries iniciais (1º ao 5º ano) e outro para as séries finais (6º ao 9º ano), como pode ser visto no Anexo C4.

Assim, concluem-se a apresentação das propostas de melhorias e as mudanças já efetivadas no SAEM. Passa-se, neste momento, à apresentação de possíveis novos usos dos resultados do SAEM, alguns inclusive já implementados a partir da edição de 2011.

4.4.2. NOVOS USOS DOS RESULTADOS DO SAEM

Esta última etapa contém os resultados relativos aos novos usos dos resultados do SAEM, obtidos, principalmente, a partir das pesquisas relativas a este trabalho.

Algumas das propostas apresentadas, na verdade, já estão sendo colocadas em prática, a partir dos resultados da 1ª avaliação do SAEM 2011, como é o caso da implantação do sistema de metas. A seguir, estão destacados os novos usos propostos.

Sistema de Metas para o SAEM

Conforme citam Fernandes e Gremaud (2009), os programas de *accountability* têm como uma de suas características a presença de um sistema de metas. Como já foi discorrido no Referencial Teórico deste trabalho, os resultados das escolas na Prova Brasil, a partir de 2005, são usados para calcular o IDEB (Índice de Desenvolvimento da Educação Básica) das escolas públicas brasileiras. O

IDEB leva em conta o desempenho dos alunos na Prova Brasil e a taxa de aprovação das escolas (FERNANDES e GREMAUD, 2009).

Com base nos resultados obtidos em 2005 e na definição de uma meta para que, em 2021, o índice nacional atinja o valor 6,0 (para o 5º ano do EF), foram estabelecidas as metas intermediárias para o IDEB de cada escola, a cada dois anos, até 2021 (FERNANDES, 2011).

Segundo Fernandes (2011), a variação dos valores das metas intermediárias é regida por uma função logística com 4 variáveis: o valor inicial do IDEB, o valor final do IDEB, o tempo entre os dois IDEBs e uma variável denominada esforço, que garante o alcance do IDEB final dentro do tempo lançado.

No caso do IDEB, foi estabelecido que a meta seria que o IDEB nacional fosse 6,0 em 2021⁴³. Assim, calculou-se o esforço necessário para atingir essa meta. Obtido esse esforço, ele foi usado para se calcular após quantos anos o IDEB nacional atingirá o valor de 9,9, uma vez que o valor do IDEB não pode ser maior do que 10. Estabelecida essa data, que para o caso do IDEB nacional foi em 2096, pode-se calcular o esforço que cada um dos municípios e estados devem empreender para o cumprimento da meta. Finalmente, tendo o valor inicial do IDEB de um município e o seu esforço, podem-se estabelecer os valores intermediários dos IDEBs para cada edição da Prova Brasil.

Conforme esta pesquisa detectou, o Contrato de Gestão que as escolas assinam com a Secretaria de Educação estabelece diversas metas a serem cumpridas, sendo uma delas o desempenho das escolas no SAEM. A pesquisa detectou, também, que essas metas eram estipuladas pelas escolas, sem nenhum critério pré-definido.

Nesse contexto, baseado no modelo usado para o estabelecimento de metas do SAEM, explanado em um documento (FERNANDES, 2011) do então presidente do INEP, Prof. Reynaldo Fernandes, foram criadas as metas de desempenho para as próximas edições do SAEM.

Para o caso do SAEM, inicialmente, foi necessário estabelecer a meta a ser atingida em 2021, haja vista que já existe a meta do IDEB para a Rede Municipal

⁴³ Válido somente para o 5º ano do EF. A meta para o 9º ano é atingir 6,0 em 2025 (FERNANDES, 2011)

nesse ano. Assim, a partir dos IDEBs do município em 2007 e 2021 e do desempenho da Rede no SAEM em 2007, foi calculada a meta do SAEM para 2021, tanto para as séries iniciais quanto para as séries finais. As Tabelas 13 e 14 apresentam os valores usados e o valor calculado para a meta do SAEM em 2021.

Tabela 13 Projeção do IDEB e do SAEM para 2021 – séries iniciais

SÉRIES INICIAIS	2007	2021
SAEM	60,80	83,00
IDEB	4,5	6,2

Fonte dos dados: Portal IDEB (<http://sistemasideb.inep.gov.br/resultado/>); Relatório da 2ª avaliação do SAEM 2007 e Relatório da 1ª avaliação do SAEM 2011 – Colégio Cenecista Dr. José Ferreira

Tabela 14. Projeção do IDEB e do SAEM para 2021 – séries finais

SÉRIES FINAIS	2007	2021
SAEM	41,27	64,7
IDEB	3,8	5,5

Fonte dos dados: Portal IDEB (<http://sistemasideb.inep.gov.br/resultado/>); Relatório da 2ª avaliação do SAEM 2007 e Relatório da 1ª avaliação do SAEM 2011 – Colégio Cenecista Dr. José Ferreira

De posse das metas do SAEM em 2021, das séries iniciais e das séries finais, foi possível calcular a variável esforço, usando também os valores do SAEM em 2011 e o tempo para atingir a meta.

Finalmente, mantendo-se constante o esforço, foram calculadas as metas intermediárias do SAEM até o ano de 2021, tanto para as séries iniciais quanto para as finais, de Rede Municipal de Ensino de Uberaba, apresentada na Tabela 15.

Tabela 15. Metas intermediárias do SAEM

	1º SAEM 2011	2º SAEM 2011	1º SAEM 2012	2º SAEM 2012	SAEM 2021
Séries Iniciais	66,35	67,35	68,34	69,32	83,00
Séries Finais	51,50	52,18	52,86	53,54	64,70

Fonte dos dados: Relatório da 1ª avaliação do SAEM 2011 – Colégio Cenecista Dr. José Ferreira

Deve-se observar que não foram apresentados os valores entre 2013 e 2020, uma vez que está prevista a revisão dos valores a cada dois anos, objetivando não incorrer num problema do IDEB, que é o que ocorre quando a escola atinge a meta de alguns ciclos futuros, antecipadamente, ficando “sem meta” por vários anos.

Aproveitando a profundidade maior do SAEM, já que ele apresenta resultados por aluno, e buscando o envolvimento mais próximo dos alunos com o sistema de metas, foram criadas também as metas para cada uma das turmas das escolas. Para tanto, usou-se o desempenho da turma na 1ª avaliação do SAEM 2011 e o mesmo esforço definido para a rede, diferenciado somente entre séries iniciais e séries finais, dependendo da série a que a turma pertence.

Reforçando a questão do envolvimento dos alunos com as metas, foi criado um novo relatório, que contém o desempenho da turma na 1ª avaliação de 2011 e a meta para a 2ª avaliação de 2011, que ocorrerá em novembro próximo. Há também um relatório com a meta da escola. Esses relatórios têm o formato de um cartaz para ser impresso e afixado nas salas de aulas (com a meta daquela turma) e nas dependências da escola (com a meta da escola). Um exemplo do modelo desse relatório pode se visto no Anexo C5.

Sistema de bônus para servidores das escolas

Associado à definição do sistema de metas para o SAEM, a Secretaria Municipal de Educação estabeleceu o pagamento de um bônus salarial, ao final de cada ano letivo, a todos os funcionários das escolas que atingirem as metas estabelecidas para a escola.

Uso do desempenho dos alunos em processos seletivos

Na medida em que se comprovou que os resultados do SAEM são válidos e confiáveis, o histórico de desempenho dos alunos no SAEM pode ser usado como

critério de acesso a programas de incentivo e apoio a alunos das escolas públicas ou mesmo para seleção de bolsistas em outras instituições.

Nesse sentido, o Colégio Cenecista Dr. José Ferreira firmou um acordo com a Secretaria Municipal de Educação em relação à seleção de bolsistas para o ano de 2012. Vale lembrar que, a partir de 2011, todos os alunos de Rede Municipal usam o Material Didático desenvolvido pelo Colégio. Como a Rede Municipal oferece cursos regulares até o final do Ensino Fundamental, não trabalhando com o Ensino Médio, esses alunos terão oportunidade de se manterem estudando dentro do mesmo sistema de ensino.

Assim, foram selecionados os 45 alunos com os maiores desempenhos no 1º SAEM de 2011 (nota de corte maior ou igual a 80%) para terem aulas complementares de Língua Portuguesa e de Matemática durante o segundo semestre de 2011 nas dependências do Colégio e com os seus professores. Ao final do ano letivo, de posse do desempenho desses alunos na 2ª avaliação do SAEM e de acordo com a avaliação da postura, do envolvimento e da dedicação desses alunos durante as aulas, alguns desses alunos receberão bolsas de estudo para cursarem a 1ª série do Ensino Médio no Colégio. Vale ressaltar, conforme divulgado pela direção do Colégio, que o número de bolsas que serão disponibilizadas depende de algumas variáveis, como o número de alunos matriculados na 1ª série em 2012, e segue as regras definidas pela mantenedora (CNEC) para a concessão de bolsas de estudo, principalmente quanto à comprovação da carência em relação ao pagamento das mensalidades.

Apoio à Formação Continuada

A Secretaria Municipal de Educação mantém um departamento responsável pela formação continuada de seus professores e gestores. Anualmente, são investidos recursos em cursos de capacitação e é realizado um congresso educacional que já está na sua décima sétima edição. Os servidores que cumprirem uma determinada carga horária anual em cursos têm direito a um décimo quarto salário ao final do ano. Entretanto, conforme informações colhidas nas entrevistas,

há a impressão de que os resultados das formações realizadas não atingem a sala de aula. Uma das causas levantadas é a oferta de cursos a todos os servidores.

Por outro lado, desde a sua criação, o SAEM trabalha com o conceito de concentrar esforços somente nos problemas, ou seja, identificar os alunos com maiores dificuldades e trabalhar as deficiências desses alunos.

Esse mesmo conceito pode muito bem ser aplicado à formação continuada mantida pela Secretaria. Os resultados de uma avaliação do SAEM podem ser usados para identificar professores cujas turmas apresentaram baixo desempenho. Identificados esses professores e as deficiências que suas turmas apresentaram, é possível planejar uma formação que sane tais deficiências, devido, principalmente, ao número reduzido de alunos na turma e ao foco que será dado ao curso.

É válido salientar que a identificação desses professores não deve se basear somente no valor absoluto do desempenho da turma. Devem ser levados em conta as condições sociais e os fatores intra e extraescolares que influenciam no desempenho dos estudantes.

Análise da série histórica de dados

Ao se analisar os resultados da última avaliação do SAEM, percebeu-se que a série histórica de dados já coletados – de 2006 a 2011 – pode ser usada para se estudar a evolução do desempenho dos alunos que participaram de todas as avaliações, visando detectar se há uma classe/série específica em que os alunos passam a ter baixo rendimento ou mesmo pesquisar se há alunos com baixo desempenho que conseguem melhorar de desempenho em algum ano, buscando verificar a eficiência das políticas de recuperação.

Vale lembrar que dentro de Teoria Clássica dos Testes (TCT), não se pode comparar desempenho de sujeitos em diferentes testes, mas pode-se comparar a classificação dos mesmos.

A produção de série histórica de desempenho dos alunos, ano a ano, é uma característica única do SAEM frente aos demais sistemas de avaliação e, assim, deve ser mais bem explorada.

Ao término da apresentação dos resultados da face Expansão, encerra-se também a apresentação dos resultados alcançados neste trabalho. Desse modo, foi atingido o objetivo específico **(d)**, que tratou de levantar pontos de melhoria e de oportunidades de novos usos dos resultados do SAEM.

4.5. VISÃO GERAL DOS RESULTADOS

No capítulo 3, sobre a metodologia do trabalho, a seção 3.3 apresentou, no Quadro 07 (p. 92), uma visão geral dos métodos que se pretendeu utilizar em cada uma das faces analisadas, os instrumentos usados e os resultados esperados. Ao final da apresentação dos resultados obtidos, é importante retomar os dados do quadro para uma comparação entre os resultados esperados e os resultados obtidos, que está apresentada no Quadro 17.

Etapa	Resultados esperados	Resultados obtidos
Face Teórica	<ul style="list-style-type: none"> • Verificação da adequação do SAEM aos padrões do JCSEE; • Elaboração de quadro comparativo entre as características do SAEM e da Prova Brasil; • Verificação da proposição P1, relativa à qualidade da avaliação. 	<ul style="list-style-type: none"> • Quadro 9, com a análise de adequação dos 27 itens de qualidade do JCSEE; • Quadro 13; discussão sobre as diferenças encontradas; • Proposição verdadeira
Face Desenvolvimento	<ul style="list-style-type: none"> • Descrição das etapas de produção de uma avaliação do SAEM; • Testar a validade de conteúdo do SAEM. 	<ul style="list-style-type: none"> • Figura 9, com as etapas de desenvolvimento e as seções 4.2.1 a 4.2.8 • Validade comprovada no Quadro 15;
Face Aplicação	<ul style="list-style-type: none"> • Complementação da descrição das etapas de produção de uma avaliação do SAEM; • Aplicar a validade de construto e de critério e a confiabilidade do SAEM; • Aceitação ou rejeição da hipótese H1 sobre a indiferença no desempenho dos alunos na Prova Brasil e no SAEM. 	<ul style="list-style-type: none"> • Seções 4.3.1 a 4.3.3 contendo a descrição da aplicação do SAEM na rede municipal de Uberaba e os atuais usos dos resultados; • Validade de construto obtida por outros testes e por análise convergente discriminante; validade de critério obtida por correlação entre Prova Brasil e SAEM; • Hipótese H1 comprovada junto à validade de critério;
Face Expansão	<ul style="list-style-type: none"> • Elaboração de quadro comparativo mostrando as melhorias que podem ser incorporadas ao SAEM em relação à teoria de avaliação e a outros sistemas de avaliação; • Descrição de novos usos dos resultados do SAEM 	<ul style="list-style-type: none"> • Quadro 16 com adequações aos padrões do JCSEE; seção 4.4.1 com discussão de mudanças e melhorias possíveis de implantação; • Seção 4.4.2 com exposição de possíveis novos usos, alguns já implantados.

Quadro 17. Comparação entre resultados esperados e obtidos

Desse modo, de acordo com o Quadro 17, tem-se que os resultados esperados, elaborados durante a construção da metodologia que seria utilizada neste trabalho, foram plenamente alcançados.

Na face Teórica, a adequação, não plena mas predominante do SAEM em relação aos padrões de qualidade da avaliação definidos pelo JCSEE, é um bom resultado para um Sistema de Avaliação que foi criado não levando em conta tais padrões. A comparação do SAEM com a Prova Brasil/SAEB evidencia as diferenças entre esses sistemas de avaliação, apontando também quais características da Prova Brasil podem ser absorvidas pelo SAEM, almejando a sua melhoria.

Os resultados da face Desenvolvimento atingiram o objetivo de levantar as etapas de desenvolvimento de uma avaliação com precisão e com riqueza de detalhes, constituindo-se como uma orientação ao trabalho dos envolvidos nessa etapa de produção do SAEM, principalmente por apresentar-se deficiente de documentação relativa às suas etapas. A Validade de Conteúdo foi comprovada também nessa face.

As etapas de aplicação, geração de resultados e identificação dos usos desses resultados foram apresentadas nos resultados da face Aplicação. A hipótese de que não existem diferenças significativas entre os desempenhos das escolas da rede municipal na Prova Brasil e no SAEM foi aceita a partir do cálculo do coeficiente de correlação entre os desempenhos nas avaliações aplicadas em 2007: 0,872 para a 4ª série do Ensino Fundamental e 0,815 para a 8ª série do Ensino Fundamental, comprovando-se, também, a validade de critério do SAEM.

Na face Aplicação, também foram apresentados os resultados da verificação da confiabilidade dos dados gerados pelas avaliações. Os cálculos do coeficiente de confiabilidade a partir de duas técnicas distintas – correlação por análise das duas metades e coeficiente alfa de Cronbach – mostraram-se bastante próximos e dentro dos limites aceitáveis. A validade de construto também foi verificada nessa face, e os resultados obtidos a partir da aplicação da análise convergente-discriminante foram suficientes para demonstrar que as avaliações do SAEM avaliam os construtos de desempenho acadêmico a que se propõe.

Os resultados da face Expansão buscaram apresentar as opções de melhoria dos processos do SAEM, detalhando os passos para que as mesmas sejam

efetivadas. Além disso, foram apresentadas algumas opções de novos usos desses resultados. Vale lembrar que algumas dessas melhorias, como os novos relatórios e o sistema de metas, já foram implementadas na edição de 2011 do SAEM.

Nessa seção, apresentam-se também os resultados da análise dos dados qualitativos coletados a partir das entrevistas realizadas com os usuários do SAEM.

A rede municipal de ensino é composta por 31 escolas, divididas em escolas na zona urbana e na rural. Das escolas da zona urbana, existem escolas localizadas tanto na região central da cidade quanto na região periférica. Há também distinção quanto ao número de alunos, com escolas pequenas, atendendo somente a algumas classes/séries, e escolas com um grande número de alunos.

Diante dessa diversidade de características, foram escolhidas três escolas: uma urbana com bom desempenho no SAEM; uma urbana com baixo desempenho e uma escola rural. Em cada escola, foram entrevistados o diretor, o coordenador pedagógico, dois professores, dois alunos e dois pais de alunos. Foram entrevistados também alguns funcionários do Colégio Cenecista Dr. José Ferreira e algumas servidoras da Secretaria Municipal de Educação (SEMEC).

Assim, a amostra de pessoas selecionadas para as entrevistas é não probabilística, o que não representa um problema, uma vez que os resultados não foram usados para fazer inferências sobre a população, e sim para confirmar as informações colhidas com os demais entrevistados.

As entrevistas foram importantes para o conhecimento e para a descrição das etapas de desenvolvimento e de aplicação, apresentadas anteriormente. Além disso, na seção 3.3 do capítulo sobre a metodologia do trabalho, foi proposta a análise das informações contidas nas entrevistas, a partir das dimensões definidas no Quadro 8 (p. 94).

Assim, o Quadro 18 retoma essas dimensões e apresenta a análise das informações obtidas.

Dimensões	Face	Análise das informações colhidas
Papel	Desenvolvimento	<ul style="list-style-type: none"> Os funcionários do Colégio, os servidores da SEMEC e os diretores têm claro o seu papel para a realização do SAEM; Os professores não reconheceram o papel da Secretaria na condução do SAEM; Os alunos e os pais reconheceram que os diretores exercem a função de incentivadores do processo.
Etapas	Desenvolvimento; Aplicação	<ul style="list-style-type: none"> Os funcionários do Colégio demonstraram conhecer as etapas de desenvolvimento da avaliação; As servidoras da Secretaria de Educação demonstraram conhecer pouco sobre a etapa de desenvolvimento, já que isso não as compete. Entretanto, demonstraram conhecer profundamente as etapas de aplicação.
Replicabilidade	Desenvolvimento; Aplicação	<ul style="list-style-type: none"> O funcionário responsável pelo banco de dados, não tem quem o substitua, comprometendo a replicabilidade; O trabalho da equipe de gestão responsável pelo SAEM é bem coordenado, de modo que, na ausência de qualquer um desses, o processo ocorre normalmente; Os diretores disseram acreditar que o processo ocorra sem a presença deles. Entretanto, eles não deixam de estarem presentes durante a aplicação da avaliação;
Documentação	Desenvolvimento	<ul style="list-style-type: none"> Os funcionários do Colégio não apresentaram documentos que descrevam as suas funções; As servidoras da SEMEC têm suas funções definidas de modo geral no documento apresentado, não há uma descrição das responsabilidades e ações em relação ao SAEM; Os diretores e os professores possuem Manuais para a orientação do seu trabalho, os quais são elaborados pela SEMEC e revisados a cada nova edição.
Melhorias	Desenvolvimento;	<ul style="list-style-type: none"> Os funcionários do Colégio solicitaram que haja um maior tempo para a condução das suas atividades e que fosse revisada a forma de envio dos dados dos alunos; As servidoras da Secretaria da Educação relataram sobre a melhora da elaboração das questões, na medida em que, segundo elas, algumas não correspondem ao descritor a que se propõe; Criação de uma Matriz de referência fixa com os descritores que são cobrados na avaliação.
Organização da aplicação	Aplicação	<ul style="list-style-type: none"> Após os vários anos de aplicação do SAEM, os procedimentos estão consolidados. As ações dos gestores da SEMEC em relação à aplicação da avaliação são claras para todos e bem avaliadas por eles; Os diretores, coordenadores e professores demonstraram que a aplicação é bem organizada, com reuniões de treinamento antes do dia, distribuição dos Manuais para todos e distribuição dos professores por sala; Os alunos e os pais que já auxiliaram na aplicação da avaliação

Dimensões	Face	Análise das informações colhidas
		relataram também que o processo está consolidado. Entretanto, uma das alunas comentou sobre a obrigatoriedade de que todos os alunos devem aguardar para que todos os alunos da escola terminem a prova para irem para o intervalo. Sugere-se que, nesses casos, os alunos que não terminaram a prova a tempo, terminem-na na biblioteca, por exemplo.
Preparação dos alunos	Aplicação	<ul style="list-style-type: none"> Os professores relataram que é uma prática comum a revisão dos conteúdos com os alunos antes da avaliação; Alguns dos alunos declararam que se preparam sozinhos em casa para a avaliação e todos confirmaram que os professores fazem revisão em sala.
Uso dos resultados	Aplicação	<ul style="list-style-type: none"> Professores não aguardam o resultado do SAEM, em julho, para identificarem alunos com necessidade de aulas de reforço; a Secretaria usa os resultados para conferir se o aluno está frequentando as aulas de recuperação. Uso do desempenho dos alunos no SAEM como nota de uma das avaliações regulares dos alunos, no que eles concordam; Os coordenadores usam os dados dos relatórios do SAEM nas reuniões com os professores; um professor afirmou, porém, que essas reuniões não ocorrem; Os diretores disseram entregar o Relatório Individual de Desempenho em reuniões de pais, o que foi confirmado pelos alunos e pelos pais entrevistados; A Secretaria promove a entrega de certificados de honra ao mérito aos alunos em uma cerimônia que conta com a presença do Prefeito e do Secretário Municipal, além dos pais dos alunos; A superação das metas de desempenho da escola no documento <i>Contrato de Gestão</i> gera pontos para os professores da escola para a sua avaliação de desempenho.
Melhorias	Aplicação	<ul style="list-style-type: none"> Não alterar a data da avaliação, o que parece estar ocorrendo com frequência; Uso de escala de proficiência; Comparação entre avaliações de diferentes edições; Aplicação da Teoria da Resposta ao Item (TRI); Resultados por nível de desempenho; Aplicação de questionários contextuais.
Novos usos	Expansão	<ul style="list-style-type: none"> Criação de um Sistema de Metas de desempenho para as escolas, de modo que se alcancem as metas do IDEB para 2021, conforme definido pelo Ministério de Educação. Uso do desempenho dos alunos em um sistema de bonificação de todos os profissionais da escola;

Quadro 18. Análise das informações coletadas nas entrevistas

A análise ora apresentada no Quadro 18 representa um resumo, uma visão geral do que foi apresentado nos resultados das faces Desenvolvimento e Aplicação, itens 4.2 e 4.3. Assim, para um maior aprofundamento dessas análises, sugere-se uma consulta ao conteúdo dos itens citados.

Após a apresentação dessa visão geral dos resultados obtidos por este trabalho, crê-se ter atingido o objetivo geral do mesmo, que era o de analisar as características do SAEM, identificando suas vantagens e suas limitações.

5. CONCLUSÕES

O objeto de estudo desse trabalho foi o Sistema de Avaliação do Ensino Municipal (SAEM), desenvolvido pelo Colégio Cenecista Dr. José Ferreira desde 2006 e aplicado à Rede Municipal de Ensino de Uberaba. O problema de pesquisa definido para esse trabalho foi “Quais são as limitações e as vantagens do SAEM como ferramenta de avaliação do desempenho acadêmico dos alunos de uma rede municipal de educação básica?”. Considerou-se então como abordagem para responder o problema de pesquisa o seguinte objetivo geral: “analisar as características do SAEM, ao identificar suas vantagens e limitações como ferramenta de aferição do desempenho acadêmico de alunos de uma rede de ensino e ao ressaltar pontos de melhoria que possam ser introduzidos ao processo”.

Buscando atingir o objetivo geral utilizou-se como método de desenvolvimento do trabalho a técnica Estudo de Caso cujo objeto de análise foi o Sistema de Avaliação das Escolas Municipais (SAEM). Conforme o trabalho foi conduzido, o mesmo se apresentou como uma Meta-avaliação, uma avaliação da avaliação.

A análise partiu definição das quatro faces do SAEM: Teórica, Desenvolvimento, Aplicação e Expansão. Na face teórica foi comprovada a adequação das características do SAEM aos padrões da JCSEE e feita uma comparação entre as características da Prova Brasil/SAEB e do sistema de avaliação estudado. Na face Desenvolvimento foram descritos os passos de desenvolvimento de uma avaliação e comprovada a validade de conteúdo da mesma, pertinente a essa face. Na face Aplicação, foram descritas as características de aplicação do SAEM à Rede Municipal de Ensino de Uberaba-MG, analisados os relatórios gerados e comprovadas as validades dos itens, do construto e de critério. A face Expansão apresentou propostas de melhorias e novos usos dos resultados do sistema de avaliação.

Os resultados apresentados permitem concluir que o questionamento do problema de pesquisa foi devidamente respondido e que, tanto o objetivo geral quanto os objetivos específicos foram alcançados plenamente.

Destacam-se também as **contribuições acadêmicas**, as **contribuições práticas**, as **limitações** do trabalho e as **perspectivas futuras** que surgiram a partir do mesmo.

Em relação às **contribuições acadêmicas**, têm-se:

- Definição das faces da avaliação: A análise do sistema de avaliação foi feita a partir da definição das quatro faces que o compõem: Teórica, Desenvolvimento, Aplicação e Expansão. As faces foram desenvolvidas pelo autor do trabalho. Essa divisão orientou a condução da análise facilitando o seu desenvolvimento e pode ser aplicada para avaliação de qualquer sistema com características semelhantes.
- Atualização da comparação entre os padrões do JCSEE: O trabalho apresentou uma atualização da tabela comparativa proposta por Rodrigues (2009), que relaciona os padrões de qualidade entre a Avaliação de Estudantes, a Avaliação de Programas e a Avaliação do Pessoal. O trabalho inclui a nova categoria e os novos padrões para Avaliação de Programas, lançados em 2011.
- Protocolo de uma Meta-avaliação: Como este trabalho realizou uma análise da qualidade do SAEM, ele dá uma contribuição de como essa análise pode ser conduzida, principalmente a partir da análise de adequação do sistema de avaliação aos padrões de qualidade do JCSEE, conforme preconiza a teoria sobre os processos de meta-avaliação.

As **contribuições práticas** são:

- Propostas de adequação aos padrões de qualidade: Além de levantar os pontos nos quais o SAEM foi considerado parcialmente adequado e não adequado aos padrões do JCSEE, o trabalho procurou propor soluções para que esses pontos passem a ser também considerados como adequados aos padrões de qualidade.
- Levantamento do pontos de melhoria: A partir da pesquisa bibliográfica e documental e, principalmente, a partir das entrevistas realizadas com

os usuários da avaliação, foi possível levantar algumas características do sistema de avaliação que podem ser incrementadas ou mesmo alteradas visando à melhoria do SAEM.

- Indicação de novos usos do SAEM: Além das propostas de melhorias, o trabalho procurou identificar novos usos dos resultados do SAEM, dentre os quais um deles já foi implantado na 1ª edição de 2011 do exame.

Como um estudo de caso no qual o objeto em análise era o SAEM, o trabalho limitou-se a analisar as características desse sistema de avaliação. Mesmo usando como referência de comparação o Sistema de Avaliação da Educação Básica (SAEB) e as suas avaliações, o trabalho não se propôs a produzir nenhuma análise desse sistema ou de qualquer outro sistema de avaliação.

Em termos de uso dos resultados das avaliações, este trabalho limitou-se a utilizar os dados da 2ª avaliação de 2007, por ser usada na comparação com os resultados da Prova Brasil do mesmo ano, e os dados da 1ª avaliação de 2010, uma das mais recentes. Dentro dessas edições escolhidas, as análises quantitativas se limitaram às avaliações das classes/séries que possuíam pelo menos 30 questões. Assim, não foram analisadas as avaliações das classes de 6, 7 e 8 anos, que possuem somente 20 questões.

Em relação às entrevistas realizadas, o trabalho limitou-se a pesquisar alguns usuários do sistema de avaliação: funcionários, diretores, coordenadores pedagógicos, professores, alunos e pais. A seleção dos usuários dessa amostra não foi probabilística, uma vez que o objetivo era levantar as informações necessárias e confirmar essas informações e não o de fazer inferências sobre a população.

Adicionalmente, vale destacar que a pesquisa teve como foco os aspectos psicométricos da avaliação que determinam a sua qualidade. Assim, não se objetivou mensurar os impactos, positivos e negativos, gerados pela avaliação, tais como o impacto no projeto político pedagógico da escola, na prática pedagógica dos professores e nos hábitos de estudos dos alunos.

No todo, o presente trabalho abre diversas **perspectivas futuras**, tais como:

- Análise de Critério a partir dos resultados da Prova Brasil de 2009: De forma semelhante à análise de critério que foi desenvolvida entre os resultados da Prova Brasil e do SAEM de 2007, é possível realizar a análise nos resultados de 2009. Entretanto, os desempenhos das escolas somente na Prova Brasil ainda não foram divulgados.
- Análise da implantação das melhorias sugeridas: Como continuidade desse trabalho, deve-se acompanhar a implementação das sugestões de melhorias e de novos usos propostos pelo trabalho para que, posteriormente, possam ser analisados os impactos dessas mudanças no sistema de avaliação, principalmente no que se refere à adoção da Teoria da Resposta ao Item (TRI) como modelo de análise dos resultados. Deve-se também acompanhar a evolução da implantação de sistema de metas de desempenho das escolas, com vistas a mensurar o seu impacto no sistema de ensino.
- Análise de outros sistemas de avaliação: Os protocolos de avaliação da qualidade de um sistema de avaliação criados nesse trabalho propiciam que outros sistemas de avaliação possam ser analisados sob a ótica aqui proposta.

Em relação aos resultados obtidos, a pronta adequação, quase integral, do SAEM ao que preconiza a teoria sobre qualidade da avaliação, talvez seja o melhor resultado apresentado por este trabalho. É importante citar ainda que a comparação com as características da Prova Brasil/SAEB trouxe uma rica contribuição, ao identificar pontos de melhoria para o SAEM.

A aplicação das técnicas estatísticas para se comprovar a confiabilidade e a validade do SAEM foi realizada pela primeira vez nos dados desse sistema. Mesmo assim, os resultados se apresentaram positivos, demonstrando que o SAEM é um sistema de avaliação confiável e que seus resultados são válidos, ou seja, são representativos do desempenho acadêmico dos alunos de Rede Municipal de Ensino de Uberaba-MG.

Os resultados apresentados na face Expansão indicam o futuro próximo do SAEM. É interessante destacar que alguns dos novos usos identificados já estão sendo incluídos na edição de 2011. A implantação das mudanças propostas e dos novos usos dos resultados implica a realização de novas análises, de novos processos de meta-avaliação, para que sejam estudados os impactos dessas mudanças na qualidade da avaliação.

Finalmente, pôde-se concluir que o presente trabalho possibilitou ao mestrando, que exerce a função de coordenador de avaliações no Colégio Cenecista Dr. José Ferreira, uma oportunidade única de aprofundamento dos seus conhecimentos sobre a Educação, a legislação educacional, as políticas públicas educacionais e, sobretudo, o papel e a responsabilidade que os sistemas de avaliação possuem na melhoria da qualidade da Educação em nosso país.

6. REFERÊNCIAS⁴⁴

ANDERSON, D. R.; SWEENEY, D. J.; WILLIAMS, T. A. **Estatística Aplicada à Administração e Economia**. 2 ed. São Paulo: Cengage Learning, 2007. 597 p.

ANDRIOLA, W. B. Psicometria Moderna: características e tendências. Estudos em Avaliação Educacional, São Paulo, v. 20, n. 43, p. 319-340, maio/ago. 2009.

ARAÚJO, C. H.; LUZIO, N. Avaliação da Educação Básica: em busca da qualidade e equidade no Brasil. Brasília, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2005. 71 p.

BARLOW, M. **Avaliação escolar**: mitos e realidades. Tradução de Fátima Murad. Porto Alegre: Artmed, 2006. 176 p.

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

BONAMINO, A. C. de. FRANCO, C. Avaliação e política Educacional. **Cadernos de Pesquisa**, São Paulo, n. 108, p. 101-132, nov. 1999.

BONI, V.; QUARESMA S. J. Aprendendo a entrevistar: como fazer entrevistas em ciências sociais. **EmTese: revista eletrônica dos pós-graduandos em Sociologia Política da UFSC**. Florianópolis, v. 2, n.1, p. 68-80, jan./jul. 2005. Acesso em: 12 fev. 2011.

BRASIL. **Constituição (1988)**. Constituição da República Federativa do Brasil. Brasília, DF: Senado, 1988.

_____. Lei nº 8666, de 21 de junho de 1993. Regulamenta o artigo 37, inciso XXI, da Constituição Federal, institui as normas para licitações e contratos da Administração Pública. **Diário Oficial da União**, Brasília, DF, 22 jun. 1993. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/L8666cons.htm> Acesso em: 21 jan.2011.

⁴⁴ De acordo com a Associação Brasileira de Normas Técnicas. NBR 6023.

_____. Lei nº 9394, de 20 de dezembro de 1996. Estabelece as diretrizes e as bases da educação nacional. **Diário Oficial da União**, Brasília, DF, 23 dez. 1996. Disponível em: <http://www.planalto.gov.br/ccivil_03/Leis/L9394.htm> Acesso em: 08 fev. 2010.

_____. Lei nº 10.172, de 9 de janeiro de 2001. Aprova o Plano Nacional de Educação e dá outras providências. **Diário Oficial da União**, Brasília, DF, 9 jan. 2001. Disponível em: <http://www.planalto.gov.br/ccivil03/leis/leis_2001/l10172.htm> Acesso em: 10 abr. 2011.

_____. MINISTÉRIO DA EDUCAÇÃO. Portaria nº 931, de 21 de março de 2005. **Diário Oficial da União**, Brasília, DF, 22 mar. 2005. Seção 1, p. 16/17. Disponível em: <<http://www.sinepe-mg.org.br>>. Acesso em: 11 out. 2010.

_____. INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. Portaria nº 47, de 03 de maio de 2007. **Diário Oficial da União**, Brasília, DF, 07 maio 2007a. Disponível em: <<http://ensinopublico.com/wp-content/uploads/2011/06/Portaria-47.pdf>>. Acesso em: 11 out. 2010.

_____. Decreto nº 6.094, de 24 de abril de 2007. Dispõe sobre a implementação do Plano de Metas Compromisso Todos pela Educação... **Diário Oficial da União**, Brasília, DF, 25 abr. 2007b. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2007-2010/2007/Decreto/D6094.htm> Acesso em: 14 abr. 2011.

_____. MINISTÉRIO DA EDUCAÇÃO. **Plano de Desenvolvimento da Educação: raízes, princípios e programas**. 2007c. Disponível em: <<http://portal.mec.gov.br/arquivos/livro/>>. Acesso em: 16 abr. 2011.

_____. MINISTÉRIO DA EDUCAÇÃO. **Plano de Desenvolvimento da Educação: Prova Brasil: matrizes de referência, tópicos e descritores**. 2007. Brasília: MEC, SEB; INEP, 2008. Disponível em: <http://www.oei.es/salactsi/provabrazil_matriz.pdf>. Acesso em: 14 out. 2010.

_____. INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. Portaria nº 87, de 07 de maio de 2009. **Diário Oficial da União**, Brasília, DF, 08 maio 2009. Disponível em: <<http://www.in.gov.br/imprensa/visualiza/index.jsp?jornal=1&pagina=48&data=08/05/2009>>. Acesso em: 11 out. 2010.

_____. MINISTÉRIO DA EDUCAÇÃO. CONAE 2010. **Construindo o sistema nacional articulado de educação**: o Plano Nacional de Educação, diretrizes e estratégias de ação. Documento Final. Brasília, DF, 2010. Disponível em: <http://conae.mec.gov.br/images/stories/pdf/pdf/documentos/documento_final_sl.pdf> Acesso em: 10 abr. 2011.

_____. INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. Portaria nº 149, de 16 de junho de 2011. **Diário Oficial da União**, Brasília, DF, 17 jun. 2011. Disponível em: <http://download.inep.gov.br/educacao_basica/prova_brasil_saeb/legislacao/2011/portaria_n149_16062011_prova_brasil.pdf>. Acesso em: 12 jul. 2011.

BROKE, N. O futuro das políticas de responsabilização educacional no Brasil. **Cadernos de Pesquisa**, São Paulo, v. 36, n. 128, p. 377-401, maio/ago. 2006.

COELHO, M. I. de M. Vinte anos de avaliação da educação básica no Brasil: aprendizagens e desafios. **Ensaio: avaliação e políticas públicas em educação**, Rio de Janeiro, v. 16, n. 59, p. 229-258, abr./jun.2008.

CONDE, F. N.; LAROS, J. A. A unidimensionalidade e a propriedade da invariância das estimativas da habilidade pela TRI. *Avaliação psicológica*. v. 6, n. 2, p. 205-215, 2007. Disponível em: <<http://pepsic.bvsalud.org/pdf/avp/v6n2/v6n2a10.pdf>>. Acesso em: 16 nov. 2010.

DANNEMANN, A. C.; FIRME, T. P.; LETICHEVSKY, A. C. Setting up the Brazilian Evaluation Network: a challenging work with no boundaries. **Ensaio: avaliação e políticas públicas em educação**, Rio de Janeiro, v. 13, n. 49, p. 515-533, out./dez.2005.

DEPRESBITERIS, L. **Avaliação da aprendizagem do ponto de vista técnico-científico e filosófico-político**. Fundação para o Desenvolvimento da Educação (FDE), São Paulo, p. 161-172, 1998. (Série Idéias, 8).

_____. Avaliação de programas e avaliação da aprendizagem. **Educação e Seleção**, São Paulo, n. 19. p. 5-31, 1989.

DOURADO, L. F.; OLIVEIRA, J. F.; SANTOS, C. A. **A qualidade da educação**: conceitos e definições. Brasília: Instituto Nacional de Estudos e Pesquisas

Educacionais Anísio Teixeira (INEP), v. 24, n. 22, p. 5-34, 2007. (Série Documental: Textos para Discussão)

DOURADO, L. F.; OLIVEIRA, J. F. A qualidade da educação: perspectivas e desafios. **Cad. CEDES**, Campinas, v. 29, n. 78, p.201-215, maio/ago. 2009.

EDUCAÇÃO de qualidade para todos: uma questão de direitos. 2. ed. Brasília, DF: UNESCO/OREALC, 2008. 108p.

FERNANDES, D. **Avaliar para aprender**: fundamentos, práticas e políticas. São Paulo: Editora Unesp, 2009. 221p.

FERNANDES, R. **Índice de Desenvolvimento da Educação Brasileira (IDEB)**. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), 2007. 26 p.

_____. **Índice de Desenvolvimento da Educação Brasileira (IDEB)**: metas intermediárias para a sua trajetória no brasil, estados, municípios e escolas. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). Disponível em: <http://download.inep.gov.br/educacao_basica/portal_ideb/metodologias/Artigo_projecoes.pdf>. Acesso em: 05 jul 2011.

FERNANDES, R.; GREMAUD, A. P. Qualidade da educação: avaliação, indicadores e metas. In: VELOSO, F. et al.(org.). **Educação Básica no Brasil**: construindo o país do futuro. Rio de Janeiro: Elsevier, 2009. 280 p.

FERREIRO, E. La Internacionalización de la Evaluación de los Aprendizajes en la Educación Básica. **Avance Y Perspectiva**. v. 24, n. 1, p.37-43, mar. 2005.

FIRME, T. P.; LETICHEVSKY, A. C. O desenvolvimento da capacidade de avaliação no século XXI: enfrentando o desafio através da meta-avaliação. **Ensaio: avaliação e políticas públicas em educação**, Rio de Janeiro, v. 10, n. 36, p. 289-300, jul./set. 2002.

FREITAS, D. N. T. de. **A avaliação da educação básica no Brasil**: dimensão normativa, pedagógica e educativa. São Paulo: Autores Associados, 2007. 224 p.

_____. Ação reguladora da União e a qualidade de ensino obrigatório. **Educar em Revista**, Curitiba, n. 31, p. 33-51, 2008.

FREITAS, L. C. et al. **Avaliação Educacional**: caminhando pela contramão. Rio de Janeiro: Vozes, 2009. 80 p.

FUNDAÇÃO CESGRANRIO. Projetos de avaliação em larga escala. Disponível em: <http://www.cesgranrio.org.br/avaliacoes/avaliacoes_larga_escala.aspx>. Acesso em: 10 nov. 2010.

GIL, A. C. **Como elaborar projetos de Pesquisa**. 4 ed. São Paulo: Atlas, 2002. 175 p.

GATTI, B. A. Avaliação Educacional no Brasil: pontuando uma história de ações. **Ecoos Revista Científica**, São Paulo, v. 4, n. 1, p. 17-41, jun. 2002.

HADJI, C. **Avaliação desmistificada**. Tradução Patrícia C. Ramos. Porto Alegre: Artmed, 2001.

HERTZ, Z. M. A. Princípios e padrões em metaavaliação: diretrizes para os programas de saúde. **Ciência & Saúde Coletiva**, Rio de Janeiro, v. 11, n. 3, p. 733-738, 2006. Disponível em: <<http://www.scielo.br/pdf/csc/v11n3/30987.pdf>> Acesso em: 05 jun. 2011.

INDICADORES da Qualidade na Educação. 3 ed. Ação Educativa, Unicef, Pnud, INEP, Seb/MEC (coordenadores). São Paulo: Ação Educativa, 2007. Disponível em: <http://portal.mec.gov.br/seb/arquivos/pdf/Consescol/ce_indqua.pdf> Acesso em: 19 abr. 2011.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **História do INEP**. Disponível em: <<http://www.inep.gov.br/institucional/historia.htm>>. Acesso em: 03 maio 2009.

_____. **História da Prova Brasil e do SAEB**. Disponível em: <<http://portal.inep.gov.br/web/prova-brasil-e-saeb/historico>>. Acesso em: 06 out. 2010.

_____. **Semelhanças e diferenças.** Disponível em: <<http://portal.inep.gov.br/web/prova-brasil-e-saeb/perguntas-frequentes>>. Acesso em: 15 mar. 2011.

_____. **SAEB 2001: novas perspectivas.** Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2001. 106 p.

_____. **Resultados Nacionais – PISA 2006:** Programa Internacional de Avaliação de Alunos (PISA). Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2008. 153 p.

IOSCHPE, G. Pra pobre analfabeto...tae know do! **Revista Veja**, São Paulo, 04 jun. 2011. p. 88-89.

JOINT Committee Standards for Educational Evaluation. Apresenta os *Standards* de Programas, de Alunos e de Pessoal. Disponível em: <<http://www.jcsee.org>>. Acesso em 25 maio 2011.

LIMA, A. O. **Avaliação escolar:** julgamento ou construção? 8 ed. Petrópolis: Vozes, 2002. 168 p.

LUCKESI, C. C. Verificação e avaliação: o que pratica a escola? In: _____. **Avaliação da Aprendizagem Escolar.** São Paulo: Cortez, 1995. 180 p.

MARCUSCHI, L. A. Gêneros Textuais: definição e funcionalidade. In: DIONÍSIO, A. P.; MACHADO, A. R.; BEZERRA, M. A. (Org.). **Gêneros Textuais e Ensino.** 2 ed. Rio de Janeiro, Lucerna, 2003.

MARTINS, G. A. Sobre confiabilidade e validade. *Revista Brasileira de Gestão de Negócios*, São Paulo, v. 8, n. 20, p.1-12, jan./abr. 2006. Disponível em: <<http://200.169.97.104/seer/index.php/RBGN/article/viewFile/51/44>>. Acesso em: 26 maio 2011.

MICHAELIS moderno dicionário da língua portuguesa. São Paulo: Melhoramentos, 2007. Disponível em: <<http://michaelis.uol.com.br/moderno/portugues/index.php>>. Acesso em: 01 maio 2010.

MEC divulga Plano Nacional de Educação 2011-2020. São Paulo: Todos pela Educação, 2010. Disponível em: <<http://www.todospelaeducacao.org.br/comunicacao-e-midia/noticias/12514/mec-divulga-plano-nacional-de-educacao-2011-2020>>. Acesso em 25 abr. 2011.

MURILLO, J. Desafios na avaliação da qualidade da avaliação. In: FERNANDES, D. (Org.) **Avaliação em Educação**: olhares sobre uma prática social incontornável. Pinhais: Editora Mello, 2011.

OECD. *Knowledge and skills for the life: first results for the PISA 2000*. Organization for Economic Cooperation Development, 2001. Disponível em: <<http://www.oecd.org/dataoecd/44/53/33691596.pdf>> Acesso em: 24 abr. 2010.

_____. *Learning for Tomorrow's World: first results from PISA 2003*. Organization for Economic Cooperation Development, 2004. Disponível em: <<http://www.oecd.org/dataoecd/1/60/34002216.pdf>> Acesso em: 24 abr. 2010. v.1: Analysis

_____. *PISA 2006: Science competencies for tomorrow's world*. Organization for Economic Cooperation Development, 2007. Disponível em: <<http://www.oecd.org/dataoecd/30/17/39703267.pdf>> Acesso em: 24 abr. 2010. v.1: Analysis

_____. *PISA 2009 Assessment Framework: keys competencies in Reading, Mathematics and Science*. Organization for Economic Cooperation Development, 2009. Disponível em: <<http://www.oecd.org/dataoecd/11/40/44455820.pdf>> Acesso em: 24 abr. 2010.

_____. *PISA 2009 Results: what students know and can do – student performance in reading, mathematics and science*. Organization for Economic Cooperation Development, 2010. Disponível em: <<http://dx.doi.org/10.1787/9789264091450-en>> Acesso em: 10 mar. 2011. 1 v.

OLIVEIRA, R. P. de. Da universalização do Ensino Fundamental ao desafio da qualidade: uma análise histórica. **Educação e Sociedade**, Campinas, v. 28, n. 100 - Especial, p. 661-690, out. 2007.

_____. A qualidade do ensino como parte do direito à educação: um debate em torno dos indicadores. In: FERNANDES, D. (Org.) **Avaliação em Educação**: olhares sobre uma prática social incontornável. Pinhais: Editora Mello, 2011.

OLIVEIRA, R. P. de; ARAÚJO, G. C. de. Qualidade de ensino: uma nova dimensão da luta pelo direito à educação. **Revista Brasileira de Educação**, Rio de Janeiro, n. 28, p. 5-23, jan./abr. 2005.

PASQUALI, Luiz. **Psicometria**: teoria dos testes na Psicologia e na Educação. 2 ed. São Paulo: Vozes, 2003. 397 p.

PERRENOUD, P. **Avaliação**: da excelência à regulação das aprendizagens – entre duas lógicas. Porto Alegre: Artmed, 1999. 184 p.

RIBEIRO, V. M.; GUSMÃO, J. B. B. Uma leitura dos usos dos indicadores de qualidade na educação. *Cadernos de Pesquisa*. v. 40, n. 141, p. 823-847, set/dez. 2010. Disponível em: <<http://www.scielo.br/pdf/cp/v40n141/v40n141a08.pdf>> Acesso em: 28 maio 2011.

PESTANA, M. I. O Sistema de avaliação brasileiro. **Revista Brasileira de Estudos Pedagógicos**, Brasília, v. 79, n. 191, p. 65-73, jan./abr.1998.

RODRIGUES, P. Resenha das obras: *The Program Evaluation Standards* (1994); *The Student Evaluation Standards* (2003) e *The Personnel evaluation standards* (2009), do *JOINT COMMITTEE ON STANDARDS FOR EDUCATIONAL EVALUATION*. **Sísifo. Revista de Ciências da Educação**. n. 9, p. 115-118, maio/ago. 2009. Disponível em: <<http://sisifo.fpce.ul.pt/pdfs/Revista%209%20PT%20REC.pdf>>. Acesso em: 22 maio 2011.

RICHARDSON, Roberto Jarry. **Pesquisa Social**: Métodos e Técnicas. 2 ed. São Paulo: Atlas, 1989. 288 p.

SAVIANI, D. O Plano de desenvolvimento da educação: análise do projeto do MEC. *Educação e Sociedade*. Campinas, v. 28, n. 100 – especial, p. 1231-1255, out. 2007.

SOARES, J. F. Escolas não aproveitam bem a avaliação de desempenho. **Folha de São Paulo**, São Paulo, 01 set. 2008. Disponível em: <<http://www1.folha.uol.com.br/fsp/cotidian/ff0109200827.htm>> Acesso em: 10 dez. 2008.

SOUZA, S. M. Z. L. 40 anos de Contribuição à Avaliação Educacional. **Estudos em Avaliação Educacional**, São Paulo, v. 16, n. 31, p. 7-35, jan./jun. 2005.

STEVENSON, W. J. **Estatística aplicada à Administração**. Tradução de Alfredo Alves de Farias. São Paulo: Harper & Row do Brasil, 2001.

THESAURUS BRASILEIRO DA EDUCAÇÃO. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Disponível em: <<http://www.inep.gov.br/pesquisa/thesaurus/>>. Acesso em: 15 maio 2010.

UNIVERSIDADE DE SÃO PAULO. **Diretrizes para apresentação de dissertações e teses da USP**: documento eletrônico e impresso Parte I (ABNT). 2 ed rev. ampl. São Paulo: Sistema Integrado de Bibliotecas da USP, 2009. 102 p.

VIANNA, H. M. **Testes em Educação**. 3 ed. São Paulo: IBRASA, 1978. 220p.

_____. **Introdução à Avaliação Educacional**. São Paulo: IBRASA, 1989. 211p.

_____. Avaliação Educacional: uma perspectiva histórica. **Estudos em Avaliação Educacional**, São Paulo, n. 12. p. 7-24, jul./dez. 1995.

_____. Avaliação Educacional: vivência e reflexões. **Estudos em Avaliação Educacional**, São Paulo, n. 18. p. 69-109, 1998.

_____. **Avaliação Educacional**: teoria, planejamento e modelos. São Paulo: IBRASA, 2000. 193 p.

_____. Programas de Avaliação em Larga Escala: algumas considerações. **Estudos em Avaliação Educacional**, São Paulo, n. 23. p. 93-104, 2001.

_____. Avaliações nacionais em larga escala: análises e propostas. **Estudos em Avaliação Educacional**, São Paulo, n. 27. p. 41-76, 2003.

_____. Avaliação de programas: duas questões. **Estudos em Avaliação Educacional**, São Paulo, n. 32. p. 43-56, 2005.

YIN, R. K. **Estudo de caso**: planejamento e métodos. 4. ed. Porto Alegre: Bookman, 2010. 248 p.

APÊNDICE A – Autorizações de uso do nome e dos dados**Autorização A1 - Colégio Cenecista Dr. José Ferreira**

Uberaba, 16 de agosto de 2010.

Colégio Cenecista Dr. José Ferreira
A/C: Prof. Danival Roberto Alves
Rua Felipe dos Santos, 286 – Uberaba-MG

Prezado Professor Danival,

O sistema de avaliação desenvolvido pelo Colégio Cenecista Dr. José Ferreira e denominado SAEM vem sendo aplicado, ininterruptamente, desde 2006, à rede municipal pública de ensino de Uberaba em uma parceria do Colégio com a Secretaria de Educação e Cultura do Município. A necessidade de explorar ainda mais os dados coletados a cada avaliação fez com que eu me inscrevesse no Programa de mestrado em Administração de Organizações da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto (FEARP – USP). Minha dissertação pretende apresentar um estudo de caso sobre o desenvolvimento de um sistema de avaliação, analisando as vantagens e desvantagens do atual modelo utilizado e identificando pontos de melhoria e novos usos que ainda não foram explorados. Um exemplo que pode ser citado é o relacionamento do desempenho dos alunos com o professor regente da sala. Devido ao grande volume de dados e toda a experiência acumulada, pensei em estudar o caso da aplicação do SAEM em Uberaba.

Desse modo, venho por meio desta, solicitar a autorização do Colégio Cenecista Dr. José Ferreira para usar, em minha dissertação, o nome do Colégio como o desenvolvedor do SAEM.

Peço ainda a colaboração do senhor para realizar algumas entrevistas com o pessoal do Colégio envolvido com o SAEM, objetivando coletar informações sobre o processo de produção de uma avaliação e, principalmente, identificar melhorias para o sistema. Ressalto que todos os participantes terão o direito a ter ou não seus nomes publicados e também assinarão um documento de autorização.

Agradecendo a atenção dispensada, comprometo-me a entregar uma cópia do trabalho para arquivamento no Colégio.

Atenciosamente,

Cristiano Pereira Arruda

Se acordo:

Autorização A2 – Secretaria Municipal de Educação e Cultura de Uberaba

SECRETARIA MUNICIPAL DE EDUCAÇÃO E CULTURA
OFÍCIO GAB/SEMEC/Nº 0724

Uberaba, 23 de agosto de 2010

Assunto: Resposta à correspondência datada de 16/06/2010.

Senhor Cristiano

Reportando-nos à correspondência, datada de 16/06/2010, informamos que a Secretaria Municipal de Educação e Cultura autoriza Vossa Senhoria a utilizar os dados obtidos pelas unidades de ensino da Rede Municipal de Ensino, nas avaliações do SAEM, nos anos de 2006 a 2010, assim como realizar entrevista com a equipe gestora, coordenadores pedagógicos, professores e alunos, como estudo de caso para compor o tema de sua dissertação.

Atenciosamente.

José Vandir de Oliveira
Secretário Municipal de Educação e Cultura

Ao Senhor
CRISTIANO PEREIRA ARRUDA
Professor do Colégio Cenecista Dr. José Ferreira
R. Felipe dos Santos, 286 - Bairro Abadia

MBDC/miad

APÊNDICE B – Roteiros para as entrevistas**Roteiro B1 - Funcionários do Colégio Cenecista Dr. José Ferreira**

Nome (opcional): _____

Cargo: _____

Função no SAEM: _____

1. Explique mais detalhadamente qual é o seu papel na produção de uma avaliação do SAEM?
2. Quem são seus fornecedores internos, ou seja, de quem você depende para executar seu trabalho?
3. Quem são seus clientes internos, ou seja, para quem você entrega o seu trabalho?
4. Há quanto tempo você participa da produção de uma avaliação do SAEM?
5. Quais são os cuidados que se deve ter para executar sua(s) tarefa(s)?
6. Elas podem ser realizadas por outra pessoa? Há quem possa lhe substituir?
7. Há algum material que descreva e oriente o seu trabalho?
8. O que você identifica que pode ser melhorado no seu trabalho para o SAEM?
9. Dentro da sua função, há algo que pode ser implementado e ainda não o foi?
10. Há alguma sugestão a mais sobre a produção de uma avaliação do SAEM, mesmo fora da sua atuação?

Roteiro B2 - Funcionários da Secretaria Municipal de Educação e Cultura de Uberaba (SEMEC)

Nome (opcional): _____

Cargo: _____

Função no SAEM: _____

1. Há quanto tempo você está na Secretaria? Há quanto tempo está envolvido com o SAEM?
2. Qual a participação da (SEMEC) na produção de uma avaliação do SAEM?
3. Internamente, como a Secretaria se organiza para a realização do SAEM?
4. Mais especificamente, qual é a sua participação na organização do SAEM?
5. A aplicação do SAEM na rede municipal é de responsabilidade da SEMEC. Quais são os procedimentos para a aplicação do SAEM?
6. Após o recebimento dos resultados de uma avaliação, como a Secretaria trabalha com esses dados? Como eles são divulgados?
7. A Secretaria tem políticas de recuperação dos alunos com baixo desempenho? Como é feito esse trabalho com as escolas? Elas acatam as orientações passadas?
8. Por outro lado, há algum tipo de incentivo aos alunos com bom desempenho no SAEM?
9. Em relação à(s) sua(s) atividade(s), é possível identificar seus clientes internos e externos na realização do SAEM? Em caso positivo, quem são essas pessoas?
10. Como um todo, em sua opinião, quais são as atividades mais críticas para o sucesso da realização de uma avaliação do SAEM?
11. Quais são os cuidados que se deve ter para executar a(s) sua(s) tarefa(s)?
12. Elas podem ser realizadas por outra pessoa? Há quem possa lhe substituir?
13. Há algum material que descreva e oriente o seu trabalho?
14. O que você identifica que pode ser melhorado no seu trabalho para o SAEM?
15. Dentro da sua função, há algo que pode ser implementado e ainda não o foi?
16. Há alguma contribuição, mesmo fora da sua atuação, para melhoria ou mesmo um novo uso do SAEM?

Roteiro B3 – Diretores, Coordenadores Pedagógicos e Professores

Nome (opcional): _____

Cargo: _____

1. Há quanto tempo você trabalha na rede municipal de Uberaba? Há quanto tempo está envolvido com o SAEM?
2. Qual a participação da escola na aplicação de uma avaliação do SAEM?
3. Internamente, como a escola se organiza para a realização do SAEM?
4. As pessoas envolvidas recebem treinamento? Há alguma orientação por escrito das funções de cada um?
5. Há alguma deficiência na aplicação do SAEM que pode ser melhorada?
6. Mais especificamente, qual é a sua participação no SAEM?
7. Como a escola motiva a participação dos alunos no SAEM? Há algum tipo de preparação específica para a avaliação?
8. E a comunidade? Como é o envolvimento dos pais ou responsáveis?
9. Após o recebimento dos resultados de uma avaliação, como a escola trabalha com esses dados? Há alguma divulgação? Como?
10. Como você vê o trabalho da SEMEC na realização do SAEM?
11. A Secretaria tem políticas de recuperação dos alunos com baixo desempenho? Como é feito esse trabalho com as escolas?
12. A escola tem políticas próprias de recuperação dos alunos com baixo desempenho? Como esse trabalho é desenvolvido?
13. Por outro lado, há algum tipo de incentivo aos alunos com bom desempenho?
14. Como um todo, em sua opinião, quais são as atividades mais críticas para o sucesso da realização de uma avaliação do SAEM?
15. Quais são os cuidados que se deve ter para executar a(s) sua(s) tarefa(s)?
16. Elas podem ser realizadas por outra pessoa? Há quem possa lhe substituir?
17. Há algum material que descreva e oriente o seu trabalho?
18. O que você identifica que pode ser melhorado no seu trabalho para o SAEM?
19. Dentro da sua função, há algo que pode ser implementado e ainda não o foi?
20. Há alguma contribuição, mesmo fora da sua atuação, para melhoria ou mesmo um novo uso do SAEM?

Roteiro B4 – Pais de alunos e alunos

Nome (opcional): _____

☐ Aluno ☐ Pais ou responsáveis

Classe/Série: _____

1. Há quanto tempo você (seu filho) estuda na rede municipal de Uberaba? Há quanto tempo você (seu filho) é avaliado pelo SAEM?
2. A escola se prepara devidamente para a aplicação do SAEM? As pessoas envolvidas demonstram estar treinadas para a aplicação da avaliação?
3. Como é a participação dos pais na aplicação do SAEM?
4. Há alguma deficiência na aplicação do SAEM que pode ser melhorada?
5. Você (seu filho), individualmente, se prepara para fazer a avaliação?
6. Como a escola motiva a participação dos alunos no SAEM? Há algum tipo de preparação específica para a avaliação?
7. E a comunidade? Como é o envolvimento dos pais ou responsáveis?
8. Após o recebimento dos resultados de uma avaliação, como a escola trabalha com esses dados? Há alguma divulgação? Como eles são divulgados?
9. A escola oferece aulas ou outras atividades de recuperação dos alunos com baixo desempenho? Como é feito esse trabalho?
10. Por outro lado, há algum tipo de incentivo aos alunos com bom desempenho no SAEM?
11. A família (Você) acompanha o seu desempenho (do seu filho)? Há algum incentivo / punição conforme o resultado? A escola tem abertura para discutir com a família os resultados do aluno?
12. Há alguma para melhoria do SAEM como um todo?

ANEXO B – Relatórios do SAEM

Relatório B1 – Relatório de desempenho individual do aluno

SAEM - Uberaba 2007		Secretaria Municipal de Educação e Cultura		Uberaba uma cidade para todos
Sistema de Avaliação das Escolas Municipais de Uberaba							
Relatório de Desempenho Individual							
IDENTIFICAÇÃO							
ESCOLA:				AVALIAÇÃO:			
ALUNO:				SÉRIE:			
DESEMPENHO DO ALUNO							
TOTAL DE ACERTOS				DESEMPENHO PORCENTUAL			
				0% _____ 100%			
CLASSIFICAÇÃO DO ALUNO							
NA TURMA		NA ESCOLA		DENTRE TODOS OS ALUNOS PARTICIPANTES			
DESEMPENHO MÉDIO DOS ALUNOS							
DA TURMA		DA ESCOLA		DENTRE TODOS OS ALUNOS PARTICIPANTES			
0% _____ 100%		0% _____ 100%		0% _____ 100%			
DESEMPENHO POR DISCIPLINA							
DISCIPLINA	NÚMERO DE ACERTOS DO ALUNO	NÚMERO DE QUESTÕES PROPOSTAS	MÉDIA DOS ALUNOS DA TURMA	MÉDIA DOS ALUNOS DA ESCOLA	MÉDIA DOS ALUNOS DAS ESCOLAS PARTICIPANTES		

C:\CICLO\SAEM\B1\032004

Relatório B2 – Relatório da Classificação geral dos alunos em sua escola e em sua turma

[illegible]

Relatório B3 – Relatório da Classificação dos alunos em sua escola e em sua turma, POR DISCIPLINA

[illegible]

Relatório B4 – Relatório de desempenho da escola

SAEM - Uberaba 2007

Sistema de Avaliação das Escolas Municipais de Uberaba

Secretaria Municipal
de Educação e Cultura **Relatório de Desempenho da Escola****IDENTIFICAÇÃO**

Avaliação: 1ª Avaliação 2007 - 27/06/2007

Escola: 01 - E. M. Modelo de Ensino

Desempenho Por Série**Classe 06 anos do Ciclo Inicial de Alfabetização**

Turmas	Cien	Hist	Geo	Mate	Port	Média
Turma A	86,25	80,00	80,00	78,33	65,71	75,75
Turma B	82,35	88,24	76,47	82,35	60,50	74,41
Turma C	85,53	94,74	68,42	63,16	45,11	63,42
Turma D	94,12	100,00	97,06	80,39	84,87	87,35
Média da Escola	86,99	90,41	80,14	75,80	63,60	74,93
Média de Todas Escolas Participantes	92,47	92,16	87,18	81,49	77,04	83,23

Desempenho Por Série**Classe 07 anos do Ciclo Inicial de Alfabetização**

Turmas	Cien	Geo	Hist	Mate	Port	Média
Turma A	71,43	90,48	45,24	61,22	60,54	63,33
Turma B	48,00	98,00	60,00	81,71	86,86	79,60
Turma C	83,33	85,71	42,86	45,58	54,42	56,19
Turma D	89,29	100,00	92,86	77,55	74,49	81,43
Média da Escola	70,37	93,21	57,41	66,31	69,49	69,63
Média de Todas Escolas Participantes	80,59	88,46	74,23	70,74	78,09	76,42

Desempenho Por Série**Classe 08 anos do Ciclo Inicial de Alfabetização**

Turmas	Cien	Geo	Hist	Mate	Port	Média
Turma A	56,25	41,67	29,17	33,33	52,38	42,71
Turma B	58,00	22,00	28,00	33,71	47,43	39,20
Turma C	52,17	36,96	39,13	39,13	48,45	43,48
Turma D	23,68	23,68	26,32	25,56	37,59	29,47
Turma E	21,05	21,05	21,05	26,32	25,56	24,47
Média da Escola	44,09	29,55	29,09	32,08	43,25	36,64
Média de Todas Escolas Participantes	46,80	35,01	35,27	39,99	46,70	42,05

Relatório B5 – Relatório de desempenho por questão dos alunos, por turma

SAEM - Uberaba 2007

Sistema de Avaliação das Escolas Municipais de Uberaba

Secretaria Municipal de Educação e Cultura

Uberaba

uma cidade para todos

Relatório de Desempenho por Escola / Disciplina / Turma

IDENTIFICAÇÃO

Escola: 01 - E. M. Modelo de Ensino

Avaliação: 1ª Avaliação 2007 - 27/06/2007

Série: 7ª Série do Ensino Fundamental

Turma: B

Disciplina: Matemática

DESEMPENHO POR QUESTÃO

Nº	Assunto	Respostas Dadas										Resposta Correta	Percentual de Acertos da Turma	Percentual de Acertos da Escola	Percentual de Acertos das Escolas Participantes
		A		B		C		D		N/B					
		Nº	%	Nº	%	Nº	%	Nº	%	Nº	%				
3	Localizar o número irracional na reta real	0	0,0	5	33,3	7	46,7	3	20,0	0	0,0	B	33,33%	34,38%	28,92%
4	Resolver problema que envolva porcentagem	4	26,7	6	40,0	4	26,7	1	6,7	0	0,0	C	26,67%	40,63%	24,85%
5	Calcular o valor numérico de uma expressão algébrica	3	20,0	6	40,0	4	26,7	2	13,3	0	0,0	A	20,00%	25,00%	22,14%
18	Problemas que envolva o cálculo de medida de ângulos	8	53,3	3	20,0	0	0,0	4	26,7	0	0,0	C	0,00%	12,50%	13,55%
19	Problemas que envolvam o perímetro e área de figuras planas	2	13,3	4	26,7	7	46,7	2	13,3	0	0,0	D	13,33%	12,50%	18,07%
23	Interpretar e utilizar dados apresentados em tabelas	4	26,7	4	26,7	4	26,7	3	20,0	0	0,0	C	26,67%	25,00%	32,53%
24	Gráfico de setores que melhor representa os dados da tabela	6	40,0	3	20,0	6	40,0	0	0,0	0	0,0	A	40,00%	46,88%	54,59%
29	Efetuar operações com polinômios	2	13,3	8	53,3	3	20,0	1	6,7	1	6,7	C	20,00%	28,13%	38,78%
30	Efetuar operações com polinômios	3	20,0	4	26,7	3	20,0	5	33,3	0	0,0	C	20,00%	25,00%	36,82%

Relatório B6 – Gráfico de acertos por aluno

Relatório B7 – Relatório Geral de desempenho

SAEM - Uberaba 2009

Sistema de Avaliação das Escolas Municipais de Uberaba

Secretaria Municipal de Educação e Cultura

Uberaba

UMA NOVA NIDADE

Relatório Geral de Desempenho (em %)

IDENTIFICAÇÃO

Avaliação: 1ª Avaliação 2009 - 08/07/2009

Série: 6ª Série do Ensino Fundamental

DESEMPENHO

Escola	Média Geral	Disciplinas								
		Cien	Geo	Hist	Mate	Port				
E. M. Modelo de Ensino 29	53,40	67,89	37,50	40,44	42,27	70,92	-	-	-	-
E. M. Modelo de Ensino 10	52,95	65,70	45,93	50,58	34,88	69,51	-	-	-	-
E. M. Modelo de Ensino 07	48,84	59,88	43,60	47,09	35,92	59,95	-	-	-	-
E. M. Modelo de Ensino 24	45,52	61,89	28,28	23,77	39,34	61,75	-	-	-	-
E. M. Modelo de Ensino 21	44,68	62,90	25,40	24,80	38,53	60,22	-	-	-	-
E. M. Modelo de Ensino 06	44,43	54,42	33,38	33,23	36,25	58,06	-	-	-	-
E. M. Modelo de Ensino 25	44,33	51,67	35,00	36,67	31,11	61,85	-	-	-	-
E. M. Modelo de Ensino 15	43,44	56,54	29,62	33,46	38,46	53,16	-	-	-	-
E. M. Modelo de Ensino 22	43,28	57,31	34,62	32,31	36,75	52,31	-	-	-	-
E. M. Modelo de Ensino 03	43,21	60,66	29,96	24,82	37,50	55,23	-	-	-	-
E. M. Modelo de Ensino 05	42,82	53,85	30,77	23,08	47,86	47,01	-	-	-	-
E. M. Modelo de Ensino 01	42,75	59,38	30,00	26,25	38,89	52,22	-	-	-	-
E. M. Modelo de Ensino 13	42,67	56,36	32,27	33,18	34,75	53,33	-	-	-	-
E. M. Modelo de Ensino 30	42,60	51,22	25,00	28,66	40,11	55,28	-	-	-	-
E. M. Modelo de Ensino 26	42,51	55,24	32,62	29,52	35,66	53,86	-	-	-	-
E. M. Modelo de Ensino 14	42,37	51,61	32,26	33,87	36,20	52,69	-	-	-	-
E. M. Modelo de Ensino 23	42,11	53,68	31,99	33,09	34,97	52,61	-	-	-	-
E. M. Modelo de Ensino 18	41,44	62,08	30,00	27,50	32,04	52,96	-	-	-	-
E. M. Modelo de Ensino 17	41,34	50,89	28,79	31,03	36,31	52,28	-	-	-	-
E. M. Modelo de Ensino 09	41,33	57,19	34,06	30,00	35,14	48,75	-	-	-	-
E. M. Modelo de Ensino 12	41,07	56,50	32,50	25,50	32,67	53,33	-	-	-	-
E. M. Modelo de Ensino 11	40,40	61,00	26,00	29,00	30,22	52,89	-	-	-	-
E. M. Modelo de Ensino 28	40,25	59,26	25,00	23,15	37,45	48,97	-	-	-	-
E. M. Modelo de Ensino 27	40,00	51,79	21,43	19,64	41,27	50,79	-	-	-	-
E. M. Modelo de Ensino 02	39,72	55,83	31,20	27,07	33,92	47,79	-	-	-	-
E. M. Modelo de Ensino 20	38,36	52,50	25,68	28,86	32,63	47,68	-	-	-	-
E. M. Modelo de Ensino 16	37,50	43,75	27,08	34,38	31,48	46,76	-	-	-	-
E. M. Modelo de Ensino 04	36,83	48,81	32,14	23,81	31,75	44,44	-	-	-	-
E. M. Modelo de Ensino 08	36,35	46,88	24,22	30,47	33,68	42,36	-	-	-	-
Média Geral das Escolas Participantes por Disciplina (em %)		56,79	31,36	30,77	36,07	54,51	-	-	-	-
Média Geral das Escolas Participantes (em %)		43,03								

ANEXO C – Relatórios Remodelados/Novos do SAEM

Relatório C1 – Relatório de desempenho individual do aluno - remodelado

SAEM - Uberaba 2011		Secretaria Municipal de Educação e Cultura		PREFEITURA Uberaba
Sistema de Avaliação das Escolas Municipais de Uberaba				CONSTRUINDO UM NOVO FUTURO			
Relatório de Desempenho Individual							
IDENTIFICAÇÃO							
ESCOLA: 08 - E. M. Modelo de Ensino 08				AVALIAÇÃO: 1ª Avaliação 2011 - 18/05/2011			
ALUNO: 14450 - Aluno Modelo 14450				ANO: 3º ano do Ensino Fundamental			
DESEMPENHO DO ALUNO							
TOTAL DE ACERTOS		DESEMPENHO PORCENTUAL		NÍVEL			
20 acertos em 20 questões		0% 100% (100,00%)		Avançado			
CLASSIFICAÇÃO DO ALUNO							
NA TURMA		NA ESCOLA		DENTRE TODOS OS ALUNOS PARTICIPANTES			
1º lugar de 26 alunos participantes		1º lugar de 105 alunos participantes		1º lugar de 1854 alunos participantes			
DESEMPENHO MÉDIO DOS ALUNOS							
DA TURMA		DA ESCOLA		DENTRE TODOS OS ALUNOS PARTICIPANTES			
0% 100% (77,31%)		0% 100% (76,38%)		0% 100% (70,38%)			
DESEMPENHO POR DISCIPLINA							
DISCIPLINA	NÚMERO DE ACERTOS DO ALUNO	NÚMERO DE QUESTÕES PROPOSTAS	MÉDIA DOS ALUNOS DA TURMA	MÉDIA DOS ALUNOS DA ESCOLA	MÉDIA DOS ALUNOS DAS ESCOLAS PARTICIPANTES		
Ciências	2	2	1,38	1,52	1,41		
Geografia	2	2	1,58	1,70	1,59		
História	2	2	1,48	1,60	1,58		
Língua Portuguesa	7	7	5,69	5,19	4,88		
Matemática	7	7	5,35	5,26	4,82		

CNEC/UBS/2011

Relatório C2 – Classificação por escola - remodelado

SAEM - Uberaba 2011

Sistema de Avaliação das Escolas Municipais de Uberaba

Secretaria Municipal de Educação e Cultura

Classificação por Escola / Turma

IDENTIFICAÇÃO

Escola: 06 - E. M. Modelo de Ensino 06

Avaliação: 1ª Avaliação 2011 - 18/05/2011

Ano: 3º ano do Ensino Fundamental

Turma: D

Nº de questões: 20

Classificação		Acertos	Desempenho (%)	Nível	Nome
Na Turma	Na Escola				
1º	1º	20	100	Avançado	Aluno Modelo 14524
		20	100	Avançado	Aluno Modelo 14532
3º	21º	18	90	Adequado	Aluno Modelo 14542
4º	35º	17	85	Adequado	Aluno Modelo 14530
		17	85	Adequado	Aluno Modelo 14548
6º	48º	16	80	Adequado	Aluno Modelo 14537
7º	54º	15	75	Adequado	Aluno Modelo 14550
8º	65º	14	70	Adequado	Aluno Modelo 14522
		14	70	Adequado	Aluno Modelo 14533
		14	70	Adequado	Aluno Modelo 14534
11º	75º	13	65	Básico	Aluno Modelo 14543
		13	65	Básico	Aluno Modelo 14549
13º	82º	12	60	Básico	Aluno Modelo 14523
		12	60	Básico	Aluno Modelo 14526
		12	60	Básico	Aluno Modelo 14528
		12	60	Básico	Aluno Modelo 14536
		12	60	Básico	Aluno Modelo 14547
18º	89º	11	55	Básico	Aluno Modelo 14529
		11	55	Básico	Aluno Modelo 14545
20º	95º	10	50	Básico	Aluno Modelo 14541
21º	98º	9	45	Crítico	Aluno Modelo 14525
		9	45	Crítico	Aluno Modelo 14539
23º	101º	8	40	Crítico	Aluno Modelo 14540
24º	103º	5	25	Muito Crítico	Aluno Modelo 14531

Relatório C3 – Relatório de Distribuição de Níveis da Escola - novo

Relatório de Distribuição de Níveis Por Escola

IDENTIFICAÇÃO

Avaliação: 1ª Avaliação 2011 - 18/05/2011

Escola: 06 - E. M. Modelo de Ensino 06

Distribuição Porcentual de Alunos Por Nível

7º ano do Ensino Fundamental

Turmas	Des. Médio	Muito Crítico	Crítico	Básico	Adequado	Avançado
Turma A	45,40	3,45	13,79	48,28	31,03	3,45
Turma B	54,11	3,33	6,67	33,33	43,33	13,33
Turma C	53,01	0,00	3,23	32,26	58,06	6,45
Turma D	55,98	5,88	11,76	11,76	58,82	11,76
Turma E	48,17	6,45	9,68	35,48	48,39	0,00
Desempenho da Escola	51,48	3,87	9,03	31,61	48,39	7,10
Desempenho da Rede Municipal	49,23	3,26	12,44	34,66	45,39	4,25

Distribuição Porcentual de Alunos Por Nível

8º ano do Ensino Fundamental

Turmas	Des. Médio	Muito Crítico	Crítico	Básico	Adequado	Avançado
Turma A	46,67	0,00	16,67	41,67	38,89	2,78
Turma B	47,74	0,00	16,13	45,16	38,71	0,00
Turma C	51,31	0,00	6,06	33,33	60,61	0,00
Turma D	49,27	0,00	9,38	37,50	53,13	0,00
Turma E	56,88	0,00	6,25	21,88	62,50	9,38
Desempenho da Escola	50,30	0,00	10,98	35,98	50,61	2,44
Desempenho da Rede Municipal	49,36	1,91	10,66	37,06	47,27	3,09

Distribuição Porcentual de Alunos Por Nível

9º ano do Ensino Fundamental

Turmas	Des. Médio	Muito Crítico	Crítico	Básico	Adequado	Avançado
Turma A	60,48	3,57	0,00	7,14	82,14	7,14
Turma B	57,65	0,00	3,70	14,81	77,78	3,70
Turma C	54,17	0,00	12,50	29,17	45,83	12,50
Turma D	46,39	12,50	8,33	29,17	50,00	0,00
Turma E	56,67	0,00	7,41	22,22	51,85	18,52
Turma F	56,79	0,00	3,57	21,43	64,29	10,71
Desempenho da Escola	55,59	2,53	5,70	20,25	62,66	8,86
Desempenho da Rede Municipal	53,30	2,78	8,41	26,07	55,17	7,57

Relatório C4 – Relatório de Distribuição de Níveis Geral

<div> <div> SAEM - Uberaba 2011 Sistema de Avaliação das Escolas Municipais de Uberaba </div> <div> Secretaria Municipal de Educação e Cultura </div> <div> </div> <div> </div> </div>						
Relatório de Distribuição de Níveis Geral						
IDENTIFICAÇÃO						
Avaliação: 1ª Avaliação 2011 - 18/05/2011						
DISTRIBUIÇÃO PORCENTUAL DE ALUNOS POR NÍVEL						
Ano	Desemp. Médio (%)	Muito Crítico	Crítico	Básico	Adequado	Avançado
1º ano do Ensino Fundamental	86.74	1.57	1.79	7.91	58.28	30.45
2º ano do Ensino Fundamental	83.45	1.74	2.41	10.98	48.36	36.50
3º ano do Ensino Fundamental	70.81	7.77	7.26	25.39	42.62	16.95
4º ano do Ensino Fundamental	54.88	6.59	12.87	37.41	31.22	11.91
5º ano do Ensino Fundamental	52.73	5.05	14.08	44.31	31.56	4.99
Distribuição Geral da Rede Municipal - Séries Iniciais	66,35	4,68	7,95	26,14	41,90	19,33
6º ano do Ensino Fundamental	53.90	2.55	8.66	25.37	55.23	8.18
7º ano do Ensino Fundamental	49.23	3.28	12.25	34.60	45.60	4.27
8º ano do Ensino Fundamental	49.36	1.91	10.61	37.11	47.26	3.10
9º ano do Ensino Fundamental	53.30	2.78	8.41	26.07	55.17	7.57
Distribuição Geral da Rede Municipal - Séries Finais	51,50	2,65	10,04	30,64	50,79	5,87

C:\8\Graf 6/02/011

Relatório C5 – Modelo do Cartaz de Metas do SAEM

ANEXO D – Exemplo de questões de uma avaliação Temática

2) Em "A sociedade é implacável" (linha 9) a palavra destacada pode ser substituída, sem perda de sentido, por:

- a) ☐ justa.
b) ☐ inflexível.
c) ☐ perdoadora.
d) ☐ amável.

3) Segundo o autor, "a mais pura manifestação do desejo e da vontade coletiva de tornar o mundo melhor para o nosso próprio bem e para o bem dos nossos filhos" (linhas 18 e 19) é:

- a) ☐ a esperança.
b) ☐ a reflexão.
c) ☐ a solidariedade.
d) ☐ o esporte.

4) Observe este mapa político da América do Sul.

O país que sediará a Copa do Mundo de 2014 e as Olimpíadas de 2016, no mapa anterior, está representado por:

- a) ☐ 1.
b) ☐ 4.
c) ☐ 9.
d) ☐ 7.

SAEM - Sistema de Avaliação das Escolas Municipais de Uberaba

Uberaba
Município

SAEM - Uberaba 2010
Sistema de Avaliação das Escolas Municipais de Uberaba

Secretaria Municipal
de Educação e Cultura

Classe de 10 anos do Ciclo Complementar de Alfabetização

1ª Avaliação Sistemática - 10/09/2010

“Esporte: Unindo povos, disseminando culturas.”

Leia sua avaliação com muita atenção e responda às questões propostas.
Você é capaz! Boa prova!

Texto para as questões de 1 a 6.

A força de uma Olimpíada

1 Nos últimos 100 anos, o esporte foi o único mecanismo capaz de promover a união e a paz entre os povos, a solidariedade e a esperança para um mundo melhor. O esporte arranca lágrimas numa guerra, une povos diferentes, abre mentes e corações em torno do mesmo ideal. Somente o esporte é capaz de proporcionar recordes inimagináveis ao despertar o desejo incontrolável dos atletas em busca da perfeição e da supremacia, sem necessidade do apelo à violência, em plena terra dos aborígenes.

5 E assim deve ser o esporte, algo pacificador, criador de talentos, inibidor da violência e, por fim, um gerador de esperança. Um novo milênio sugere uma mudança de vida e esta, por sua vez, sugere uma profunda mudança nos conceitos de relacionamento pessoal.

10 A sociedade é implacável, seletiva e discriminadora, mas, quando se trata de esporte, o mundo se transforma, as forças se equilibram, os níveis se igualam e as chances se equivalêm.

Em pleno início do século XXI, não vemos alternativa senão a de provocar a reflexão nas pessoas, nas comunidades, nas famílias, nos clubes e nos governantes. A grande missão é unir, conciliar, despertar o interesse coletivo, partilhar a alegria e dividir a dor, fazer da reflexão um ponto de apoio e crescimento em todos os campos, jamais incitar o desânimo e a violência.

15 Vivemos um novo milênio e, por conta dele, há mais problemas que soluções, mais tristeza que esperança. Somos capazes de virar esse jogo e provar o quanto é possível mudar e evoluir através dos nossos atos.

20 O esporte é a mais pura manifestação do desejo e da vontade coletiva de tornar o mundo melhor para o nosso próprio bem e para o bem dos nossos filhos. Graças a Deus, ainda há tempo, muito tempo. O milênio está apenas começando. E através da Copa do Mundo de 2014 e das Olimpíadas de 2016, o Brasil terá a grande chance de ser o protagonista das duas maiores celebrações do esporte mundial. E viva o Brasil! E viva o Esporte! E viva a Olimpíada!

Jerônimo Mendes

<http://jeronimos.wordpress.com/2009/10/02/a-forca-de-uma-olimpiada/> — acessado em 07/05/2010

1) O texto trata, principalmente:

- a) ☐ da violência.
b) ☐ do novo milênio.
c) ☐ da união promovida pelo esporte.
d) ☐ da discriminação.

SAEM - Sistema de Avaliação das Escolas Municipais de Uberaba

Releia este trecho para resolver a questão 5.

Em... "despertar o desejo incontrolável dos atletas em busca da perfeição e da supremacia..." (linhas 4 e 5)

- 5) Para alcançar a perfeição e a supremacia, os atletas precisam superar todas as limitações de seu corpo. Por isso, durante os jogos:
- a) () os movimentos respiratórios e os batimentos cardíacos dos atletas aumentam.
 - b) () a distribuição de nutrientes pelo corpo dos atletas é interrompida.
 - c) () o coração dos atletas desacelera e faz o sangue circular lentamente.
 - d) () os movimentos respiratórios e os batimentos cardíacos dos atletas diminuem.

6) O texto afirma que "...o esporte foi o único mecanismo capaz de promover a união e a paz entre os povos..." (linhas 1 e 2). No Brasil, a união de diferentes culturas deu-se, em grande parte, devido à chegada dos imigrantes no século XX. Observe a seguir o gráfico que mostra a vinda de alguns imigrantes, entre 1900 e 1930, de acordo com o IBGE:

Número de imigrantes no século XX

www.ibge.gov.br — acessado em 17/05/2010

Após a análise desse gráfico, conclui-se que:

- a) () a imigração de italianos foi a menor.
- b) () os portugueses foram os imigrantes que vieram em maior quantidade para o Brasil.
- c) () a maior imigração apresentada é a de espanhóis.
- d) () os alemães foram os imigrantes que vieram em maior número para o Brasil.

7) No Brasil, a união de culturas também se deu por meio de migrações internas. Assim, vários costumes, vindos de diferentes lugares, proporcionaram uma mistura cultural. Observe o mapa a seguir:

www.mestredahistoria.com — acessado em 17/05/2010

Sobre as migrações internas, na década de 1970, conclui-se que:

- a) () os nordestinos migraram para todas as outras regiões do Brasil.
- b) () o Sudeste recebeu a menor quantidade de migrantes.
- c) () grande quantidade de pessoas da Região Sul foi para o Nordeste.
- d) () a Região Centro-Oeste recebeu muitos migrantes do norte.

- 13) No desenho do campo de futebol, você observou vários quadriláteros. Eles recebem o nome de:
- a) () losango.
 - b) () trapézio.
 - c) () quadrado.
 - d) () retângulo.

- 14) O gol deve ser colocado na região central da linha de fundo do campo de futebol. Como o campo mostrado possui 69 metros de linha de fundo, divida 69 por 3 para descobrir a quantos metros do início da linha de fundo o gol deve ser colocado.

O gol deve ser colocado a uma distância de:

- a) () 36 metros.
- b) () 23 metros.
- c) () 34 metros.
- d) () 26 metros.

Texto para as questões 15 e 16

Para informar a população brasileira sobre as cidades-sede em que acontecerão os jogos da Copa no Brasil, foi lançado o seguinte informe publicitário:

<http://guiadacopa.net/2014/images/stories/Infopadacopa2014.png> — acessado em 10/05/2010

- 15) As cidades de Manaus, Natal, Belo Horizonte, Brasília e Curitiba serão algumas das sedes para os jogos da Copa em 2014. Essas cidades pertencem, respectivamente, às seguintes regiões:

- a) () Nordeste, Norte, Sul, Sudeste e Centro-Oeste.
- b) () Norte, Nordeste, Sudeste, Centro-Oeste e Sul.
- c) () Sudeste, Nordeste, Norte, Sul e Centro-Oeste.
- d) () Centro-Oeste, Norte, Nordeste, Sul e Sudeste.

SAEM - Sistema de Avaliação das Escolas Municipais de Uberaba

7

- 16) O Espírito Santo é o único Estado da Região Sudeste que não sediará nenhum jogo na Copa de 2014. Além do Espírito Santo, também fazem parte da Região Sudeste os seguintes Estados:

- a) () Bahia, São Paulo e Paraná.
- b) () Rio de Janeiro, Bahia e Paraná.
- c) () São Paulo, Minas Gerais e Bahia.
- d) () Minas Gerais, São Paulo e Rio de Janeiro.

Texto para as questões 17 e 18.

Brasília, a capital do Brasil, será também uma das cidades-sede da Copa de 2014. Conheça um pouco mais sobre ela.

Brasília 2014

Estádio Copa 2014: Mané Garrincha
— capacidade atual: 53 250
— capacidade em 2014: 90 000
População: 2 400 000 habitantes
Rede hoteleira: 27 000 leitos
Estabelecimentos de saúde: 1 700
Transporte público: 6 600 ônibus

Cidades Copa do Mundo 2014 — Notícias 2014 (13/04/2010)
Disponível em: <http://www.guiadacopa.net/2014/cidades-sede> — acessado em 10/05/2010

- 17) Se, durante a Copa no Brasil, 1 000 pessoas pagarem cada um dos 6 600 ônibus que circulam em Brasília, qual o total de pessoas que circularão de ônibus nesse período?

Para descobrir esse número, multiplique 6 600 por 1 000.

O número encontrado é:

- a) () 86 000 pessoas.
- b) () 66 000 000 pessoas.
- c) () 6 600 000 pessoas.
- d) () 660 000 pessoas.

SAEM - Sistema de Avaliação das Escolas Municipais de Uberaba

8

- 18) A capacidade atual do estádio Mané Garrincha é de 53 250 lugares. Depois da reforma pela qual esse estádio passará para a Copa de 2014, sua capacidade será de 90 000 lugares.

O estádio terá um aumento de:

- a) () 36 750 lugares.
b) () 143 250 lugares.
c) () 43 250 lugares.
d) () 38 650 lugares.

- 19) A tabela a seguir mostra as capacidades atuais dos estádios em que acontecerão os jogos da Copa do Mundo de 2014 de algumas cidades-sede.

Cidade-sede	capacidade atual
Belo Horizonte – MG	76 500
Rio de Janeiro – RJ	92 000
São Paulo – SP	73 501
Brasília – DF	53 250

Após analisar a tabela, é possível afirmar que:

- a) () Belo Horizonte possui um estádio com capacidade maior do que São Paulo.
b) () Brasília possui o estádio que abriga o maior número de pessoas.
c) () o Rio de Janeiro possui um estádio que abriga o menor número de pessoas.
d) () o estádio onde cabe o menor número de pessoas está em São Paulo.

Observe, no gráfico a seguir, a capacidade futura dos estádios das cidades-sede que foram mostradas na tabela da questão anterior.

20) Ao analisar esse gráfico, pode-se concluir que:

- a) () Belo Horizonte possuirá o estádio com o menor número de lugares disponíveis.
b) () o Rio de Janeiro será a cidade com o estádio de maior capacidade.
c) () Brasília possuirá um estádio com mais lugares do que o Rio de Janeiro.
d) () para lotar o estádio de São Paulo deverão comparecer aos jogos 62 800 pessoas.

- 21) Como a capacidade atual dos estádios é de 76 500 pessoas em Belo Horizonte, 92 000 pessoas no Rio de Janeiro e 73 501 em São Paulo, quantas pessoas podem assistir aos jogos, ao mesmo tempo, nesses estádios, se eles estiverem lotados?

- a) () 241 000 pessoas.
b) () 224 001 pessoas.
c) () 142 002 pessoas.
d) () 242 001 pessoas.

- 22) Para a realização da Copa de 2014, muitos estádios de futebol precisarão ser reformados, para atenderem às exigências da FIFA. Essas reformas mobilizarão diversos trabalhadores, como engenheiros, jardineiros, arquitetos, pedreiros, dentre outros.

Segundo dados do site *folha.uol*, um engenheiro tem média salarial de R\$ 6 000,00 e um jardineiro, de R\$ 560,00.

A renda obtida na atividade de trabalho está relacionada com as condições de vida do trabalhador porque:

- a) () em média, o jardineiro possui salário maior que o engenheiro.
b) () a média salarial do engenheiro e a do jardineiro são iguais.
c) () na maioria dos casos, tem melhor qualidade de vida quem possui maiores rendas.
d) () de acordo com a média salarial, a qualidade de vida do jardineiro é melhor que a qualidade de vida do engenheiro.

- 23) Belo Horizonte também será uma das cidades-sede da Copa.

A imagem abaixo mostra a planta dessa cidade feita em 1896, antes de sua construção.

Com base nas informações anteriores, conclui-se que:

- a) () a cidade de Belo Horizonte foi planejada.
b) () Belo Horizonte foi construída sem planejamento.
c) () a planta representa a cidade de São Paulo.
d) () a planta foi feita atualmente.

Texto A

www.ricbit.com/ilafox/olimpiadas_monica.jpg — acessado em 07/05/2010

Texto B

24) Ao compararmos os dois textos, notamos que:

- a) ☐ ambos referem-se a Copa do Mundo.
 b) ☐ ambos pertencem ao mundo real.
 c) ☐ o **texto B** critica uma situação mundial, a desunião entre os povos.
 d) ☐ ambos têm as mesmas informações.

25) No **texto A** em "Hei! Isso não é esporte!", a pontuação sugere que a personagem está:

- a) ☐ contrariada.
 b) ☐ cansada.
 c) ☐ indiferente.
 d) ☐ desinteressada.

26) Quanto ao **texto A**, pode-se afirmar que:

- a) ☐ é um texto sem humor.
- b) ☐ tem como objetivo instruir.
- c) ☐ é direcionado ao público adulto.
- d) ☐ o humor está relacionado às modalidades esportivas.

27) No **texto A**, a Magali está praticando levantamento de melancia. Para ser bem-sucedido nos esportes, os atletas precisam de uma alimentação saudável, pois é dos alimentos que o organismo retira os nutrientes necessários para a produção de energia.

Como é feita a distribuição desses nutrientes pelo corpo?

- a) ☐ Pela urina, com o auxílio do sistema excretor.
- b) ☐ Pelo sangue, que transporta os nutrientes às células do corpo.
- c) ☐ Pelo intestino, que se liga aos demais órgãos do corpo.
- d) ☐ Pelo processo de transpiração.

28) O **texto A** cita uma corrida de 100 metros. Sabendo-se que as pistas de corrida possuem marcas para identificar a posição dos atletas durante a competição, divida 100 por 25 para descobrir de quantos em quantos metros são feitas as marcas na pista.

As marcas, nas pistas de corrida, são feitas a cada:

- a) ☐ 5 metros.
- b) ☐ 4 metros.
- c) ☐ 10 metros.
- d) ☐ 15 metros.

Observe esta imagem e responda às questões 29 e 30.

http://sarauanet.files.wordpress.com/2008/04/nado_livre.jpg?w=480 — acessado em 07/05/2010

29) A finalidade **principal** do texto é:

- a) ☐ criticar.
- b) ☐ descrever.
- c) ☐ informar.
- d) ☐ relatar.

30) O autor desse texto pretendeu chamar a atenção para:

- a) ☐ a importância das Olimpíadas no Brasil.
- b) ☐ a modalidade nado livre.
- c) ☐ um problema comum em algumas cidades.
- d) ☐ a necessidade de se praticarem esportes.