

Elaboração de instrumentos para avaliação de Desempenho Escolar: processo colaborativo de criação de itens.

Developing tools for students' performance assessment: collaborative process of items creation.

Juliana Batista Pereira dos Santos

Mestranda do Programa de Pós-graduação em Educação em Ciências: Química da
vida e Saúde
Universidade Federal de Santa Maria
juhbpereira@gmail.com

Marcela Santos de Almeida

Mestranda no Programa de Pós-graduação em Ensino de Ciências e Matemática
Universidade Federal de Sergipe
cecelabioalmeida@hotmail.com

Luiz Caldeira Brant de Tolentino-Neto

Docente do Departamento de Metodologia do Ensino e do Programa de Pós-
graduação em Educação em Ciências: Química da vida e Saúde
Universidade Federal de Santa Maria
lcaldeira@gmail.com

Resumo

O presente trabalho apresenta e discute a criação de um instrumento que colabore para suprir a carência de avaliações de caráter nacional (1) voltadas ao desempenho escolar dos alunos da Educação Básica também na disciplina de Ciências e (2) que possam ser realizadas com uso do computador. Relatamos as estratégias desenvolvidas por uma rede multidisciplinar e multicultural vinculada ao Projeto Observatório da Educação, na elaboração de um banco de itens para as disciplinas de Ciências, Língua Portuguesa e Matemática para o 2º, 5º e 9º ano do Ensino Fundamental. Os resultados mostram a viabilidade da criação de itens de forma colaborativa e à distância, e reforçam a importância de avaliações que incluam o monitoramento das Ciências a fim de ampliar o auxílio ao professor na identificação das principais dificuldades de aprendizagem dos alunos.

Palavras chave: Desempenho Escolar, Avaliação de Desempenho, Trabalho Colaborativo, Criação de Questões.

Abstract

This paper presents and discusses the development of a tool that can help to address the lack of assessment of students' performance in a nationwide basis (1) focused on the academic

performance in Science and (2) computer based. We describe the strategies developed by a multidisciplinary and multicultural network linked to the Project Observatory of Education, creating a data bank of items for disciplines of Science, Portuguese Language and Mathematics for the 2nd, 5th and 9th grade of elementary school. Results show the feasibility of developing items in remote and collaborative ways, and reinforce the importance of assessments for the monitoring of Science in order to help teachers to identify the main students' learning difficulties.

Key words: School Performance, Performance Evaluation, Collaborative Work, Items Development.

Contexto da Pesquisa

O programa computacional Ensinando o Cérebro (ENSCER), fomentado inicialmente pela FAPESP no ano de 2004, tem como resultado um software cuja proposta é avaliar o desenvolvimento cognitivo de alunos mediante a resolução de um conjunto de atividades educativas. No estudo inicial realizado por Rocha (2009, p. 90), as atividades foram construídas com base nos estudos da neurociência sobre o aprendizado e a maneira como foram processadas as informações recolhidas pelo instrumento dão indícios de que, com seus resultados, é possível entender a organização neural dos alunos.

Em função do desempenho acadêmico do aluno nas atividades educativas, o software possibilita ao professor obter informações sobre as dificuldades e facilidades de cada aluno. Essas informações podem ser utilizadas como referencial para a elaboração de estratégias de ensino ou de um plano de intervenção educacional adequado a cada caso, permitindo uma aproximação entre a área da saúde e da educação. (ROCHA, 2009, p. 91; PEREIRA, 2012, p. 39).

Atualmente o projeto ENSCER mantém um vínculo com o Projeto de Pesquisa em Rede, “Desempenho Escolar Inclusivo na Perspectiva Multidisciplinar” (Edital 038/2010/CAPES/INEP – Observatório da Educação) iniciado em 2011. O projeto é desenvolvido em cinco núcleos, sediados pelas Universidade de São Paulo (USP), Universidade Federal de São Paulo (Unifesp), Universidade Federal de Santa Maria (UFSM), Universidade Federal de Mato Grosso (UFMT) e Universidade Federal de Sergipe (UFS).

Juntos, os núcleos trabalham em torno de quatro eixos norteadores: (1) Estimativa da influência dos Distúrbios de Aprendizagem nos índices de Desempenho Escolar em Língua, Matemática e Iniciação à Ciência; (2) Estudo da influência da tecnologia em testes de avaliação padronizados de desempenho em Língua, Matemática e Iniciação à Ciência; (3) Estudo da relação entre o Desempenho Escolar em Língua, Matemática e Iniciação à Ciência e Livro Didático e (4) Produção de materiais formativos e instrucionais que contemplem o enfrentamento das Dificuldades de Aprendizagem no processo de formação do professor pesquisador.

Entre as diversas ações desenvolvidas pelo projeto, destaca-se a possibilidade de estudos realizados em conjunto entre os diferentes professores da educação básica, alunos de iniciação científica e pós-graduação, bem como, coordenadores dos distintos núcleos, distribuídos entre diferentes áreas do conhecimento (Língua Portuguesa, Matemática, Biologia, Física, Química, Pedagogia, Psicologia, Estatística e Educação Física). Os participantes interagem por meio eletrônico (lista de e-mail e Ambiente Virtual de Aprendizagem Moodle) e presencial, em grandes reuniões anuais.

As primeiras atividades realizadas pelo grupo de pesquisa focaram na elaboração de questões para compor o banco de atividades que ampliarão a abrangência e profundidade do software ENSCER. Como previsto pelo projeto original do Observatório, as disciplinas envolvidas na atividade foram Língua Portuguesa, Matemática e Ciências, para o 2º, 5º e 9º anos do Ensino Fundamental.

Diante desse contexto, pode-se perceber estamos diante de um trabalho essencialmente colaborativo, conforme Damiani (2008). A autora comenta que nesse tipo de trabalho os membros possuem objetivos comuns, compartilham as decisões que são tomadas pelo coletivo e estabelecem relações de não-hierarquização. Assim, ocorre trabalho em conjunto que visa a responsabilidade mútua pelas ações.

O processo colaborativo de elaboração de itens

No Brasil, a elaboração dos itens das avaliações em larga escala – de abrangência nacional - é gerenciada pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) que possui um banco de dados que reúne as questões usadas em todos os exames realizados pelo Ministério da Educação (MEC). A metodologia que é utilizada para a seleção das questões dos testes e análise de resultados é a Teoria de Resposta ao Item (TRI) (BRASIL, 2008; INEP, 2009) que permite a comparabilidade entre os resultados das avaliações ao longo do tempo.

Com o intuito de possibilitar inferências adequadas sobre o desempenho dos alunos nestas avaliações, algumas questões se repetem ao longo dos anos de avaliação, bem como, podem ser utilizadas em diferentes anos escolares. Esta característica permite calibrar a qualidade das questões quanto ao público alvo, linguagem e conteúdo. (ARAÚJO; LUZIO, 2005, p.18)

O ENSCER, em sua versão inicial, dispõe de itens para avaliar os estudantes áreas de Língua Portuguesa e Matemática. O intuito da parceria entre as instituições é ampliar e qualificar este banco de itens, acrescentando questões de Ciências, testando-as quanto à compreensão, qualidade das respostas, nível de dificuldade e também estabelecendo uma análise comparativa entre versões em papel e digital do instrumento.

O processo de elaboração iniciou com reuniões dos coordenadores dos núcleos para a preparação das ações futuras para a construção das questões do teste de desempenho. Os anos escolares foram escolhidos em função das séries avaliadas pelo já consagrado Sistema Nacional de Avaliação da Educação Básica (SAEB) e pela Provinha Brasil, que além de outras estatísticas, fornece dados sobre o desempenho de escolares na realização de atividades educativas em lápis e papel, porém apenas para as disciplinas de Língua Portuguesa e Matemática.

Desta maneira, tratando-se da elaboração das questões de Língua Portuguesa e Matemática, foram utilizadas como referência as avaliações padronizadas já existentes, como a Prova Brasil e a Provinha Brasil, tomando como principal base as questões já existentes e os seus temas e descritores, para a elaboração de novas.

Em Ciências, não há no Brasil um monitoramento da qualidade de seu ensino-aprendizagem desde 1999, quando a disciplina foi excluída do SAEB. O que há são ações isoladas, de Estados e Municípios, de avaliação de seus sistemas educacionais (GAME/UFMG, 2011) entre eles, destacam-se o AvaliaBH (de Belo Horizonte/MG) e o Programa de Avaliação do Sistema Educacional do Paraná (AVA), utilizados como referências em outras pesquisas deste Grupo. O primeiro, avalia alunos de Ensino Fundamental desde 2008 para Língua Portuguesa

e Matemática, e incluiu itens de Ciências a partir de 2010. Já o AVA monitorou tal desempenho entre 1996 e 2002 com alunos de todo o Ensino Fundamental.

Há ainda as avaliações internacionais que incluem Ciências, dentre elas o PISA (Programme for International Student Assessment) que é uma avaliação trienal, aplicada a estudantes de 15 anos, envolvendo as áreas de Leitura, Matemática e Ciências. No entanto o uso destes instrumentos criados no exterior, para medições de desempenho de estudantes brasileiros, deve ser feito com cuidado para evitar descontextualizações, exageros e, sobretudo, comparações descabidas.

Cabe destacar que, segundo o Ministro da Educação Aloizio Mercadante¹, a disciplina de Ciências será incluída nas avaliações do SAEB, que contemplam o 5º e 9º anos do Ensino Fundamental e o 3º ano do Ensino Médio. Se realizado, este será um grande passo para o aprimoramento das avaliações nacionais, uma vez que estas estão englobando conteúdos necessários à formação discente, porém, o público alvo ainda não contempla todos os anos escolares, já que a entrevista não cita a inclusão da disciplina na Provinha Brasil.

Todos os núcleos participaram do processo de construção, tendo como principal foco as questões de Ciências, inexistentes até então. Como inspiração para a construção da maioria das questões, utilizou-se o tema Sítio, explorando as situações do dia-a-dia rural e suas relações com a cidade. Para complemento, os elaboradores poderiam optar pelo uso de imagens, já que tínhamos disponível um banco de imagens construídas especialmente para uso no software .

Dessa maneira, cada núcleo gerenciou seu processo alicerçado em um Roteiro comum, que norteou a elaboração das questões, estruturado como um modelo de identificação e descrição das atividades construídas. Dado o grande número de envolvidos no processo e a necessidade da rigorosidade do rigor para garantir a confiabilidade dos itens construídos (ARAÚJO e LUZIO 2005, p.18), o Roteiro apontou a necessidade de um momento em comum a todos para o compartilhamento das atividades realizadas e troca de experiências.

Em Sergipe, trios compostos por um aluno de iniciação científica, um de pós-graduação e um professor de educação básica se formaram para a tarefa de pesquisa/produção de itens. Cada pequeno grupo ficou responsável de elaborar questões de Português, Matemática e Ciências para um ano específico. Os pequenos grupos para a elaboração dos itens de Ciência se dividiu da seguinte maneira: para as questões do 2º Ano de Ciências dois professores ficaram responsáveis, sendo um de Biologia e o outro de Química; já para as atividades do 5º ano foram duas professoras de Química; e para o 9º ano foram duas professoras de Física. A partir das diferentes visões dos componentes, buscou-se a possibilidade de ocorrer a troca de informações, sugestões e reflexões para a elaboração das questões.

De maneira similar, o núcleo Santa Maria dividiu seus bolsistas em pequenos grupos, sendo dois deles responsáveis por criar itens de Língua Portuguesa e Matemática e um terceiro grupo que colaborava com orientações psico-pedagógicas, cada qual composto por profissionais das áreas específicas, acompanhado de um profissional das Ciências.

O processo de construção dos itens seguiu o seguinte roteiro:

(1) criação das questões;(2) encaminhamento das atividades para os pequenos grupos do próprio núcleo para análises e sugestões; (3) retorno das sugestões e reformulações necessárias por emails; (4) apresentação das questões pré-finalizadas para os demais integrantes do núcleo e discussões das atividades com as adaptações eventualmente sugeridas;

¹Entrevista publicada pelo jornal O Estado de São Paulo em fevereiro de 2013. Pode ser encontrada em <<http://www.estadao.com.br/noticias/geral,prova-brasil-avaliara-ciencias-a-partir-deste-ano,993703.0.htm>>

(5) compartilhamento via Moodle com os demais núcleos (momento em comum); (6) leitura e análise das questões dos demais núcleos e realização de sugestões se necessário; (7) adequação das sugestões; (8) envio das questões finalizadas para os responsáveis pela inserção no software (concentrada no núcleo Sergipe).

Esse ciclo foi contínuo, de modo que enquanto as questões elaboradas na etapa (1) encontravam-se já na etapa (5) um novo ciclo de elaboração se iniciava, repetindo-se os mesmos processos, por isso a necessidade de um trabalho – ainda que independente entre os núcleos- organizado com etapas em comum a todos. O processo se realizou ao longo de dois meses e meio, totalizando aproximadamente 150 questões distribuídas entre as disciplinas de Ciências, Língua Portuguesa e Matemática.

A etapa (1) envolvia o preenchimento de um Formulário de elaboração, fazendo com que todas as questões perfizessem o mesmo processo. Assim, ao submeter uma atividade, seus elaboradores deveriam indicar o **objetivo da atividade**, correspondente aos descritores da Prova Brasil, no caso de Linguagem e Matemática, ou a um objetivo equiparado a um descritor, construído pelo grupo no caso de Ciências. Em seguida, fazer uma **descrição** indicando o conhecimento ou a interpretação que o aluno necessitaria para o desenvolvimento da questão.

Além disso, indicar o **grau de complexidade** da questão entre os anos escolares avaliados e uma **justificativa** para todas as tomadas de decisão realizadas ao longo do processo, como a escolha do grau de complexidade, as referências utilizadas para a elaboração, o conteúdo escolhido e o formato adotado.

Ao avaliarmos as questões elaboradas pelos demais núcleos na etapa (6), era necessário o preenchimento de um formulário, contendo itens importantes quanto à qualidade das questões. Dessa forma, o avaliador deveria assinalar **adequado**, **inadequado** ou ainda **não avaliado** para os seguintes itens: (a) correção gramatical; (b) correção conceitual; (c) adequação da atividade ao grau de complexidade proposto; (d) adequação da linguagem à série; (e) adequação da imagem ao texto; (f) atendimento ao objetivo proposto; (g) clareza da justificativa apresentada; (h) pertinência da justificativa apresentada e (i) adequação à diversidade étnica de gênero.

Além dessa avaliação, o grupo de Santa Maria voltado à orientação psico-pedagógica, também elaborou pareceres aos conjuntos de questões elaboradas por todos os núcleos.

Após o processo de formulação e inclusão das questões no software, obteve-se a versão digital das atividades educativas a serem aplicadas aos alunos da educação básica e posteriormente as questões foram transcritas do meio digital para o formato impresso. Dessa maneira obteve-se 3 questionários em papel, um para cada ano escolar, que continham as três disciplinas desejadas. Para essa passagem, o formato e as figuras das questões foram mantidas, para fins de comparação posterior.

Segundo as orientações da Provinha Brasil², o questionário do 2º ano deve ser lido, em sua grande parte, aos alunos, uma vez que nessa fase da escolarização os mesmos podem não ter desenvolvido completamente essa habilidade. Essa indicação implicou na necessidade do software incluir áudios em seu sistema, contendo a parte que deveria ser lida aos alunos, para que a prova pudesse ser realizada individualmente.

² Informações extraídas do Manual do Aplicador do SAEB/2009 elaborado pelo Ministério da Educação (MEC) e Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP).

Analogamente, para a aplicação da versão em papel, foi necessária a presença de um aplicador responsável somente pela leitura das questões, conforme as instruções do manual do INEP.

Como resultado deste processo colaborativo obteve-se três questionários, um para cada ano escolar avaliado, em que o instrumento do 2º ano foi composto por 26 atividades de Ciências, 18 de Língua Portuguesa e 20 de Matemática; o instrumento do 5º ano foi composto de 10 questões de Língua Portuguesa, 28 questões de Matemática e 21 itens de Ciências; e por fim, o questionário do 9º ano trouxe 11 de Língua Portuguesa, 36 de Matemática e 21 de Ciências.

Reflexões sobre o processo colaborativo de construção de questões

O processo de elaboração das questões, como todo novo processo que se inicia, exigiu muitos estudos e discussões entre os envolvidos. Ao elaborar um item muitas incertezas e angústias fizeram-se presentes e, conseqüentemente, muitas indagações surgiram. Questões como o tipo de enunciado, a utilização ou não de imagens, a adequação da questão ao seu objetivo e o número de alternativas de resposta, permearam todo processo. No entanto, entendeu-se que tais levantamentos são conseqüências naturais da proposta diferenciada que trazemos e, além disso, a análise destas questões após a aplicação do Piloto, contribuirá para as criações futuras.

O tema base para elaboração das questões, ainda que presente em todas as regiões do País, pode apresentar palavras ou termos específicos de cada região. Esse fato nos alertou para as diferenças de vocabulário existente, tanto entre os envolvidos no processo de elaboração, quanto entre os próprios alunos que realizaram a atividade. Dessa maneira, ao analisarmos as questões de outros núcleos, procuramos minimizar a diferença.

O processo deslanchou com a elaboração do Roteiro acima citado, pois até então, cada núcleo apresentava uma visão do processo, assim, com a padronização do processo, tornou-se possível a comparação entre questões, a troca de experiência entre os núcleos e o aprimoramento da atividade como um todo.

Observou-se, durante as discussões das questões de Ciências no núcleo de Sergipe, que os professores se basearam em livros didáticos e nos Parâmetros Curriculares Nacionais (PCN) para a elaboração das questões. Esta opção - certamente alvo de muitos problemas - foi uma conseqüência da falta de descritores de Ciências, diferentemente de Língua Portuguesa e Matemática, que foram utilizados, além dos descritores da Prova Brasil, também livros didáticos e experiências do cotidiano.

As figuras utilizadas nas questões também geraram algumas dificuldades no processo de elaboração, isto porque a proposta de avaliar o desempenho dos alunos exigia que as figuras se adequassem ao contexto e as informações da questão em sua totalidade, sem a possibilidade de distorção de conceitos. Além das imagens já construídas e disponibilizadas para uso, ainda era possível que os elaboradores da questão encomendassem ao ilustrador da equipe uma imagem especial.

Após a elaboração das questões, durante as reuniões presenciais, foram debatidos alguns pontos positivos e negativos. Percebeu-se que alguns grupos interagiram mais que outros devido a proximidade dos componentes. Similar a esta dificuldade, em alguns momentos as questões elaboradas por um núcleo não receberam avaliação dos demais, na etapa destinada à isso prevista pelo Roteiro. Esse fato diminui os efeitos de um trabalho coletivo em prol da

troca de ideias e reforça a necessidade de comprometimento entre todos os envolvidos no processo.

Conclusões

O processo de construção de questões, realizado de forma colaborativa entre uma equipe multidisciplinar, pode alcançar resultados diferentes daqueles de uma produção individualizada, restrita entre áreas do conhecimento ou entre profissionais do mesmo nível de estudo e/ ou trabalho. O diálogo entre profissionais de diferentes áreas e níveis de estudo amplia as discussões, favorece a troca de experiências e proporciona uma reflexão sobre os temas e conteúdos a serem trabalhados.

Além destas características do grupo, cabe destacar a pluralidade de culturas entre os envolvidos no processo, que inclui as regiões Sul, Sudeste, Nordeste e Centro-Oeste. Este fato se mostra positivo, visto que o Teste de Desempenho que se propõe criar é voltado ao contexto das regiões de todo País

Nosso teste busca também atender aos alunos com Dificuldade de Aprendizagem, o que se apresenta como uma inovação, pois as avaliações educacionais nacionais buscam por meio de seus dados, em termos gerais, verificar o desempenho dos alunos e diagnosticar as escolas boas e ruins. Nessa perspectiva, o instrumento de avaliação que a equipe elaborou busca considerar as particularidades dos alunos, visto que seus resultado não pretendem criar rótulos, mas identificar qual sua área de conhecimento mais carente do aluno e, com o auxílio do professor, contornar as dificuldades.

A qualidade de ensino não pode ser confundida ou limitada ao desempenho escolar, este deve ser tomado, apenas, como um fator propulsor de mudanças, seja em sala de aula, seja em políticas públicas. Bizzo (2011) ressalva que

a melhoria da qualidade da educação depende da existência, da aplicação eficiente e do uso adequado de mecanismos sensíveis de monitoramento , que incluam – mas não se limitem a – o acompanhamento do desempenho escolar, algo que carecemos no Brasil. (Bizzo, 2011, p. 256)

Nesta mesma perspectiva, Vianna (2000) contribui nessa discussão apontando outras variáveis que podem influenciar nos resultados de desempenho:

A medida da qualidade em educação, entretanto, não pode ficar restrita apenas ao desempenho escolar. Necessita, também, verificar outras variáveis que se associam e condicionam o rendimento escolar. O que as crianças fazem na escola, o que os professores procuram transmitir aos seus alunos e o que os livros didáticos apresentam refletem expectativas culturais e educacionais da sociedade, bem como seus valores e seus objetivos sociais e econômicos. Assim, é impositivo verificar em que medida a interação dessas variáveis contribui para a qualidade da educação. (Vianna, 2000, p. 190)

Próximas etapas

Após a aplicação do estudo piloto do teste, o ano que se segue será dedicado à análise das informações obtidas com sua aplicação e ao aprimoramento do software. Diversas discussões estão sendo abordadas através de pesquisas de iniciação científica, mestrado e doutorado.

Como uma proposta para melhor orientar os procedimentos da elaboração das próximas questões, surge a criação de um guia de elaboração como uma Matriz, a partir da análise dos dados das aplicações do teste piloto.

Ao fim do edital, será disponibilizado uma versão do software de forma online e gratuita para que professores possam fazer uso, contribuindo assim para um acompanhamento da aprendizagem dos alunos.

Agradecimentos

Agradecimentos são devidos à Universidade Federal de Santa Maria/RS, Universidade Federal de São Paulo, Universidade Federal de Sergipe, Universidade Federal do Mato Grosso, Universidade de São Paulo. Ao Programa de Pós-graduação em Educação em Ciências da UFSM. O presente trabalho é um resultado preliminar de projeto de pesquisa com financiamento CAPES/INEP – Observatório da Educação [edital 038/2010].

Referências

- 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.
- BIZZO, N. Exames de desempenho de estudantes e qualidade da educação no Brasil. *In: Brasil do século XXI*. Coordenador Antonio Delfim Netto. São Paulo: Saraiva, 2011.
- BRASIL. MINISTÉRIO DA EDUCAÇÃO. **Manual dos aplicadores**: SAEB (ANEB/Prova Brasil). Brasília: MEC, INEP, 2009.
- CASTRO, M. H. G. Sistemas de avaliação da educação no Brasil: avanços e novos desafios. **São Paulo Perspec.**, São Paulo, v. 23, n. 1, p. 5-18, jan./jun. 2009.
- COSTA, F. K. B.; ALMEIDA, M. M.B.; LIMA, I. B. O novo ENEM e o livro didático de química. *In: VI Colóquio Internacional: “Educação e Contemporaneidade”*. São Cristóvão, Brasil, Sergipe, set. 2012.
- Damiani, M. F. **Entendendo o trabalho colaborativo em educação e revelando seus benefícios**. Editora UFPR. Curitiba, n. 31, p. 213-230, 2008.
- INEP. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Prova Brasil e Saeb. 2009. Disponível: < <http://portal.inep.gov.br/web/prova-brasil-e-saeb/prova-brasil-e-saeb>>. Acesso em: 11/ 10/2012.
- KUHN, L. B. C. Refletindo sobre formas de avaliação: ENEM que avaliação é essa? *Anais do SILEL*. vol. 2, n.2. Uberlândia: EDUFU, 2011.
- MINHOTO, Maria Angélica Pedra. Da disseminação da cultura de avaliação educacional: estudo sobre a institucionalização do ENEM. **Poiésis, Tubarão**, v. 1, n. 1, p. 67-85, jan./abr. 2008.
- PEREIRA, J. B.; **Dificuldades de Aprendizagem**: um estudo sobre a discalculia. 2012. 49f. Monografia de Especialização (Especialização em Educação Matemática). Universidade Federal de Santa Maria, Santa Maria, 2012.
- ROCHA, F. T. **Sistema informatizado para avaliação de crianças com dificuldades de aprendizagem**. 2009. 117f. Tese (Doutorado em Medicina). Universidade de São Paulo, São Paulo, 2009.
- UNIVERSIDADE FEDERAL DE MINAS GERAIS; **A avaliação externa como instrumento da gestão educacional nos estados**. Relatório final. GAME/FAE/UFMG, Agosto/2011
- VIANNA, H.M. Avaliações nacionais em larga escala: análises e propostas. **Textos Fundação Carlos Chagas**, n.23, 2003. 41p.