

BALANÇO DA PRODUÇÃO TEÓRICA SOBRE AVALIAÇÃO DE SISTEMAS EDUCACIONAIS NO BRASIL: 1988 A 2011

Adriana Bauer – FCC

Adriana Teixeira Reis – FCC

No final do século XX, a avaliação de sistemas passou a ser um ponto de destaque nas propostas de políticas públicas em vários países¹. No Brasil, essas avaliações começaram a ser concretizadas no final da década de 1980, mais especificamente em 1988, com a proposição do Sistema Nacional de Avaliação do Ensino Público de 1º grau (SAEP) que lançou as bases para uma política de avaliação federal abrangente. (Bonamino, 2002; Freitas, 2004)

A partir da proposição do SAEB (1990), e mais recentemente da Prova Brasil (2005), diversos estados e municípios² federados consolidaram seus próprios sistemas de avaliação e monitoramento da educação. Levantamento realizado por Brooke, Cunha e Faleiros (2011)³ mostra a existência, em 2011, de sistemas de avaliação em quinze estados. Além desses, outros quatro estados tiveram a experiência de implantar políticas de avaliação, descontinuadas posteriormente.⁴

Apesar da consolidação e expansão de sistemas de avaliação, nota-se que mesmo após quase vinte e cinco anos da realização da primeira experiência, ainda há muita controvérsia em torno deles. Inclusive, tais controvérsias ganham maior relevo quando, a partir de 2007, o Instituto Nacional de Estudos e Pesquisas Educacionais (INEP) criou o Índice de Desenvolvimento da Educação Básica (IDEB), pretendo indicador de qualidade do ensino.

Sabe-se que muito já se produziu em termos de conhecimento acadêmico, em diversas áreas, sobre os sistemas de avaliação da educação e as avaliações em larga escala. Essa produção, por suas características multidisciplinares, sua diversidade e

¹ Por motivo de espaço, as questões contextuais que levaram a avaliação de sistemas a assumir um papel importante nas políticas públicas educacionais não serão exploradas nesse artigo. Para informações sobre o tema, consultar Afonso (2000).

² Até o momento não há levantamento de quantos dos municípios brasileiros já possuem seus sistemas próprios de avaliação. Contudo, experiências municipais começam a aparecer, não apenas em grandes centros, como São Paulo, Campinas e Rio de Janeiro, mas em municípios de médio e pequeno porte, como Vitória (ES), Castro (PR), Angra dos Reis (RJ), João Monlevade (MG)

³ Trata-se da pesquisa intitulado “A avaliação externa como instrumento de gestão educacional nos estados”. Para mais informações consultar Brooke, Cunha e Faleiros (2011).

⁴ São eles: Sergipe (EXAEB, em 2004, 2005 e 2006), Bahia (PAE, entre 2001 e 2004), Paraná (AVA, de 1995 a 1998, com nova aplicação no ano 2000) e Rio Grande do Sul (SAERS, de 1996 a 1998, continuamente e, posteriormente, com aplicações em 2005, 2007, 2009 e 2010). Outros três estados planejavam, em 2011, implantar ou retomar a implantação de sistemas de avaliação: Goiás (que já contava com a experiência do SAEGO em 2001, 2002 e 2004), Piauí (que implantou a Prova Piauí em 2008) e Rondônia.

pelos interesses que envolvem a temática nem sempre é facilmente identificada. Também há uma lacuna no conhecimento sobre a existência de referenciais teórico-metodológicos comuns para compreender a expansão e consolidação das avaliações de redes de ensino e as eventuais contribuições de cada um ao campo da avaliação.

Para contribuir com essas questões, a pesquisa da qual deriva o presente texto em sua primeira fase buscou mapear a produção acadêmica sobre os sistemas de avaliação no Brasil no período de 1988 a 2011, tendo como produto um banco de dados sobre os trabalhos localizados e a análise de suas características gerais⁵.

Outros trabalhos já realizaram o inventário sistemático das produções sobre avaliação educacional, mas não com o foco definido para esse estudo nem com sua abrangência temporal⁶.

Partindo desses trabalhos, procurou-se completar o mapeamento da produção acadêmica sobre o tema por meio do estudo de teses e dissertações, com o objetivo de contribuir para a reflexão sobre essa produção. As questões que nortearam a análise da produção acadêmica na primeira fase do estudo foram: quais são os temas e subtemas mais frequentes nos estudos sobre avaliação de sistemas educacionais? Quais as temáticas que emergem ao longo do tempo e quais são esquecidas? Em quais autores e referenciais os discentes dos programas de pós-graduação se apoiaram para fundamentar suas pesquisas? Quais as metodologias utilizadas nesses estudos? Que características demarcam a produção científica brasileira em avaliação de sistemas educacionais entre os anos de 1988 e 2011?

Para a constituição do corpus de análise foram selecionados os resumos das dissertações e teses, disponíveis no Banco de Dados da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). Este texto focaliza os achados resultantes da análise do levantamento bibliográfico realizado, trazendo algumas considerações sobre possibilidades e desafios de se trabalhar com a base Capes.

Análise da produção acadêmica: limites e contribuições (???)

⁵ Em uma segunda fase, o estudo se centrará em aspectos relativos à constituição desse campo de estudos no Brasil. Neste trabalho, focaliza-se nos resultados obtidos na fase do levantamento bibliográfico e da constituição do banco de dados supramencionado.

⁶ Ver Barretto e Pinto (2001), Borges e Calderón (2011), Poltronieri e Calderón (2012), Martins e Sousa (2011) e Sousa, Pimenta e Machado (2012).

Por que realizar a análise de produção acadêmica se os trabalhos já teriam sido foco de avaliação em exame para concessão de título ao pesquisador que o realizou?

Diversos autores têm enfatizado a necessidade de refletir sobre o processo de pesquisa acadêmica, realizando uma meta-análise da qualidade dessa produção (Larocca, Rosso, Souza; 2005). Com o aumento recente da produção científico-acadêmica brasileira, faz-se necessário atentar para seu processo de elaboração, sem desconsiderar questões relativas ao rigor e qualidade da pesquisa (André, 2001), discutindo, inclusive, a formulação dos objetivos de pesquisa na pós-graduação (Larossa; Rosso; Souza, 2005), e questionando os instrumentos teórico-metodológicos utilizados e os parâmetros de avaliação da qualidade da pesquisa científica existentes. (Gatti, 2001).

Cabe destacar que já no início da década de 1990, alguns autores chamavam a atenção para a necessidade de atentar para a fragilidade da revisão da bibliografia nas teses e dissertações (Alves, 1992) e para o caráter pragmático de alguns trabalhos (Warde, 1990). Tais preocupações ilustram a necessidade de análise constante da produção acadêmica.

Metodologia da pesquisa

Durante o processo de produção do banco de dados foi possível notar que não há unicidade, entre esses trabalhos, na nomenclatura utilizada para referenciar as avaliações de sistemas educacionais. Ora chamada de avaliação de sistemas educacionais, ora de avaliação externa, ou avaliação de rendimento de alunos, dentre outros conceitos que aparecem como sinônimos nos resumos (embora não o sendo), a localização e busca pelos trabalhos criou um desafio à parte para o levantamento bibliográfico proposto. Essa multiplicidade foi notada logo de início, quando se buscou conhecer o mecanismo de localização dos trabalhos criado pela Capes ao disponibilizar o serviço do Banco de Teses ao público em geral.

Primeiramente, notou-se uma incongruência entre a descrição das possibilidades trazidas pela ferramenta de busca existente no sítio da Capes e seus recursos reais. Na página inicial do Banco de Teses, é informada a possibilidade de realizar pesquisas por autor, título e palavras-chave. No entanto, os campos de busca existentes são autor, assunto e instituição, sendo possível associar essas informações ao ano de produção e ao

nível de pós-graduação desejado (mestrado ou doutorado). Fica a dúvida, então, do que deve ser colocado no campo “assunto”, pois a depender da escolha realizada, assunto, palavra-chave do trabalho ou seu título, os trabalhos listados pelo programa são diferentes.

De todo modo, ao perceber essa limitação no Banco, optou-se por realizar a pesquisa por alguns descritores previamente definidos no campo “assunto”. Para a escolha dos descritores de busca⁷ foi procurado o suporte biblioteconômico de técnicos da Biblioteca Ana Maria Poppovic (BAMP), da Fundação Carlos Chagas. Nesse momento, percebeu-se a inexistência de uma padronização no que se refere a esse elemento, já que a catalogação dos trabalhos é realizada a partir de descritores definidos nas próprias bibliotecas. A BAMP utiliza o *Thesaurus* Brasileiro da Educação e o *Thesaurus* da UNESCO como base para a elaboração de seu catálogo de descritores, mas não há uma obrigatoriedade no uso dessas referências. A partir da pesquisa nestes dois documentos, foram encontrados vocábulos gerais, como avaliação educacional, avaliação escolar e avaliação de alunos.

Tendo em vista a necessidade de realizar a pesquisa a partir de diversos descritores mais específicos, optou-se por um primeiro levantamento de termos comuns em diversos artigos científicos sobre os sistemas de avaliação da educação conhecidos previamente pelas autoras, tendo assim a seleção inicial uma base empírica. Assim, foi elaborada uma lista com catorze expressões, quais sejam: avaliação de/em larga escala, avaliação educacional/ avaliação da educação, avaliação de sistemas educacionais, avaliação de monitoramento educacional, avaliação do rendimento escolar, sistemas de avaliação educacional, sistemas de avaliação da educação, avaliação externa, avaliação de desempenho de alunos, SAEB, Prova Brasil e Provinha Brasil.

Também foram acrescentados aos descritores de busca o nome dos sistemas estaduais de avaliação existentes. Para essa pesquisa, tomaram-se como base os sistemas identificados por Nigel Brooke, em pesquisa realizada em 2010/2011. Foram eles:

- AC – Avaliação Escolas Públicas do Acre (1999 e 2003)
- SEAPE – Sistema Estadual de Avaliação da Aprendizagem Escolar (2009)
- AM – Avaliação Escolas Públicas do Amazonas (2003)
- SADEAM – Sistema de Avaliação de Desempenho Educacional do Amazonas
- AEP – Avaliação Escolas Públicas do Tocantins
- PAM – Projeto de Avaliação do Maranhão (SADa partir de 2010)

⁷ Nesse trabalho, descritores estão sendo entendidos como “vocabulário controlado que reúne termos e conceitos relacionados entre si a partir de uma estrutura conceitual da área” (BRASIL, 2011).

SIMADE – Sistema Maranhense de Avaliação de Desempenho, PROVA PIAUÍ (2008)
SPAECE – Sistema Permanente de Avaliação da Educação Básica do Ceará
SAEPE – Sistema de Avaliação Educacional de Pernambuco
SAVEAL – Sistema de Avaliação de Alagoas
EXAEB-SE – Exame de Avaliação da Educação Básica do Estado de Sergipe (2004-2006)
PAE – Projeto de Avaliação Externa da Bahia
NAE – Núcleo de Avaliação Educacional do Mato Grosso do Sul
SAEMS – Sistema de Avaliação da Educação Básica do Mato Grosso do Sul (2003)
SAEGO – Sistema de Avaliação da Educação de Goiás
SIADE – Sistema de Avaliação de Desempenho das Instituições Educacionais do Distrito Federal
SIMAVE – Sistema Mineiro de Avaliação (SIMAVE – PROEB desde 2000)
PAEBES – Programa de Avaliação da Educação Básica do Espírito Santo
PNE (2000) – Programa Nova Escola
SAERJ – Sistema de Avaliação da Educação do Estado do Rio de Janeiro ou SAERJINHO (2011)
SARESP – Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo
AVA – Programa de Avaliação do Sistema Educacional do Paraná
SAERS – Sistema de Avaliação do Rendimento Escolar do Rio Grande do Sul

Finalmente, foi acrescido na busca o sistema municipal de avaliação da cidade de São Paulo, a Prova da Cidade.

Durante a fase inicial, chamou a atenção a falta de padronização de descritores entre as instituições e ao caráter generalista do vocabulário controlado da UNESCO e do INEP, fatores que podem ter dificultado a localização de trabalhos que seriam da área de interesse da pesquisa, mas que não foram localizados por meio da base Capes, constituindo, assim, uma limitação do trabalho.

A Tabela 1 traz a quantidade de trabalhos encontrados, por descritor pesquisado, na etapa inicial de levantamento de dados. Cabe ressaltar que dos 37 descritores utilizados para a busca do tipo “por expressão exata”⁸, 17 não retornaram resultados, sendo que se estabeleceu que os trabalhos seriam selecionados por meio de 20 descritores.

⁸A ferramenta de busca da Capes permite que a busca seja realizada “por expressão exata”, “todas as palavras” ou “qualquer uma das palavras” digitadas no campo de interesse do pesquisador.

Inicialmente foram encontrados 854 estudos. Em uma primeira análise, foi possível excluir trabalhos repetidos e que não eram se referiam ao tema de interesse da pesquisa. Também se procedeu a um refinamento da pesquisa, a partir da leitura dos títulos e resumos de todos os trabalhos. Após essa triagem, os estudos selecionados para compor o banco de dados da pesquisa resultaram em 294. Procedeu-se, então, à impressão da ficha do trabalho disponível no sítio da Capes e à localização dos trabalhos, visando ao preenchimento do banco de dados.

Nota-se, pela comparação dos resultados da Tabela 1 com os descritores originalmente pesquisados, que alguns sistemas estaduais têm sido foco sistemático de análise na academia. São também os sistemas dos estados das regiões sul e sudeste que têm sido foco sistemático de pesquisa, talvez pela concentração de grupos de pesquisa sobre avaliação e gestão educacional existentes nas universidades dessa região. À exceção do Spaece do Ceará, e do Saepe, de Pernambuco, parece que os sistemas de avaliação dos estados no Norte, Nordeste e Centro-Oeste não têm sido foco do estudo acadêmico. Ou, se o são, os autores não identificam seus estudos de modo que possam ser facilmente associados com a temática de avaliação de sistemas educacionais.

Tabela 1 - Resultados da pesquisa por descritor

Descritor	N
Avaliação educacional	16
Avaliação de sistemas educacionais	10
Sistemas de avaliação educacional	3
Sistemas de avaliação da educação	17
Avaliação em larga escala/ Avaliação de larga escala	33
Avaliação do rendimento escolar	36
Avaliação de monitoramento educacional	3
Avaliação externa	29
Avaliação de desempenho de alunos	2
SAEB – Sistema de avaliação da Educação Básica	48
AVA – Programa de Avaliação do Sistema Educacional do Paraná	1
PNE – Programa Nova Escola -	5
Prova Brasil	36
Prova da Cidade	1
Provinha Brasil	9
SAEPE – Sistema de Avaliação Educacional de Pernambuco	7
SAERS – Sistema de Avaliação do Rendimento Escolar do Rio Grande do Sul	2
SARESP	25
SIMAVE/ PROEB	8

SPAECE	3
TOTAL	294

Os descritores de busca que aparecem mais relacionados aos trabalhos de avaliação de sistemas educacionais são as expressões “avaliação do rendimento escolar” e “avaliação em larga escala”. No entanto, vários outros termos são utilizados pelos autores dos trabalhos, sem que se note a preocupação de explicar o sentido em que o conceito é entendido, como nos casos dos trabalhos pesquisados sob os descritores “avaliação de monitoramento educacional”, “avaliação de rendimento escolar” e “avaliação de desempenho de alunos”.

A diversidade e confusão no uso da terminologia relativa à avaliação educacional e sua conseqüente imprecisão, já foi objeto de reflexão de Popham, há quase trinta anos (Popham, 1983). À época, o autor imputava a displicência no uso de algumas expressões ao caráter instável do campo da avaliação educacional e a pouca tradição dos pesquisadores em trocar entre si os significados da terminologia utilizada. Apesar de estar preocupado com o uso de outros conceitos (mensuração, classificação, avaliação, apreciação, prestação de contas, estimativa, etc.), a reflexão realizada pelo autor parece atual quando se pensa nas expressões utilizadas em trabalhos sobre avaliações de sistemas educacionais.

Durante a busca efetiva dos trabalhos selecionados, algumas dificuldades foram recorrentes: a não disponibilização dos trabalhos completos em meio virtual, a ausência de informação sobre os trabalhos, o conflito de informações entre os dados da ficha Capes e os trabalhos, a precariedade de ferramentas de busca oferecida pelas instituições, quer pela lentidão no processo de busca ou pela impossibilidade de acesso ao trabalho pelo público em geral.

Cabe mencionar que do total de trabalhos que compõem o banco de dados, foram obtidos 249 trabalhos em meio eletrônico (sendo 216 dissertações e 36 teses) e seis em meio impresso, somando 255 trabalhos.

Focos de interesse dos estudos

No que se refere aos focos de interesse dos estudos, a partir da leitura e análise dos resumos e títulos dos trabalhos foi possível identificar quatro grupos distintos. O Grupo 1 incorpora trabalhos que focalizam a discussão das políticas de avaliação e de

sua inserção nas reformas educacionais. Nesse grupo foram considerados os estudos que têm como objeto a avaliação de sistemas educacionais propriamente dita, propondo reflexões acerca das políticas de avaliação, seus condicionantes, motivações, pressupostos que as baseiam, delineamentos, desenhos metodológicos, planejamento e execução, etc. Também foram considerados, nesse grupo, os estudos sobre o papel da avaliação de sistemas educacionais na política educacional mais ampla ou na reforma educacional.

O Grupo 2 refere-se aos estudos sobre as implicações dessas avaliações no sistema educacional e na escola. É composto pelos estudos que discutem os usos (e desusos) dos resultados das avaliações na gestão, no planejamento escolar, implicações das avaliações no currículo, na dinâmica escolar, na formação e trabalho dos professores ou, ainda, que usam os dados das avaliações para discutir os fatores explicativos dos resultados obtidos pelos alunos. Análises de efeitos e impactos dessas avaliações fazem parte desse agrupamento.

O Grupo 3 foi engloba os estudos que fazem uma discussão centrada no desenho ou na metodologia da avaliação de sistemas educacionais, dos instrumentos de avaliação e o processo de elaboração e análise de conteúdo dos itens, das matrizes de referência, das escalas de proficiência. Estudos cujo foco é a discussão dos procedimentos estatísticos utilizados para processar os dados e/ou a transformação desses dados em indicadores para a análise e monitoramento dos resultados educacionais também foram incorporados nesse agrupamento.

Finalmente, o Grupo 4 é aquele dos estudos que se apropriam das informações geradas pelas avaliações para fazer discussões e análises a partir de outros objetos de estudo, ou seja, o interesse não está na discussão da avaliação de sistemas educacionais, mas em objetos de estudo específicos da Matemática (aprendizado de Geometria, por exemplo), Língua Portuguesa (aquisição da linguagem), Economia (uso de determinado modelo econômico para aferir eficiência na educação), Demografia, Planejamento Urbano e Regional, etc. Também foram identificados estudos cuja preocupação central é discutir as diferenças nos resultados devido a diversidade de gênero, raça e etnia, utilizando as bases de dados das avaliações que, por não se aterem às discussões de avaliação, foram computados nesse grupo.

Para cada categoria foram elaboradas catorze subcategorias, a partir das quais houve a tentativa de classificar os trabalhos segundo seu foco. As subcategorias são apresentadas no Quadro 1.

Quadro 1 – Subcategorias de análise dos trabalhos

Grupo 1 – Discussão das políticas de avaliação e de sua inserção nas reformas educacionais	1.1 – Foco na Avaliação de Sistemas Educacionais	Análise acerca das políticas de avaliação, seus condicionantes, motivações, pressupostos que as baseiam, delineamentos, etc.
	1.2 – Foco na Política Educacional	Estudos que discutem a política ou a reforma educacional e destacam a avaliação de sistemas nesse contexto
	1.3 – Foco na discussão sobre o desenho metodológico da avaliação	Estudos que discutem o planejamento das avaliações, seus descritores, adequação às políticas mais amplas, etc.
Grupo 2 Implicações das avaliações no sistema educacional e na escola	2.1 – Apropriações, usos (e desusos) dos resultados das avaliações	Estudos que buscam analisar a apropriação e usos dos resultados das avaliações na gestão, no planejamento escolar, na formulação de políticas, na reformulação de currículo, no trabalho dos resultados, etc.
	2.2 – Impactos e efeitos das avaliações	Estudos que focalizam nas implicações, impactos e efeitos das avaliações no currículo, na dinâmica escolar, na avaliação escolar, na formação e trabalho dos professores, etc.
	2.3 – Fatores associados aos resultados	Estudos sobre fatores associados/ explicativos dos resultados, influência de fatores intraescolares, de perfil de professor, etc. usando os dados das avaliações
	2.4 – Avaliação interna x avaliação externa	Comparação/ confronto entre os resultados de avaliações internas e externas
Grupo 3 – Discussão sobre a metodologia e as formas de tratamento dos dados das avaliações	3.1 – Desenho avaliativo	Foco na análise das matrizes, descritores, escalas, itens, habilidades avaliadas, questionário, etc.
	3.2 – Procedimentos de análise	Foco na discussão dos procedimentos estatísticos utilizados para processar e analisar os dados
	3.3 – Indicadores de monitoramento	Foco na discussão sobre a produção de indicadores produzidos a partir dos resultados ou produzidos para a análise e monitoramento dos resultados
	3.4 – Aplicação	Discussão da metodologia de aplicação dos testes, instrumentos de avaliação, processo de elaboração dos itens, etc.
Grupo 4 – Apropriação das informações das avaliações em outros objetos de estudo	4.1 – Apropriação dos resultados das avaliações por outras áreas (ensino de) e para orientar reflexões sobre outros objetos de estudo	Estudos de áreas com temáticas específicas da Matemática, Língua Portuguesa, Ed. Física, História, Geografia, etc.
		Estudos sobre perfil e formação de professores, violência, currículo, progressão continuada, Bolsa Família, diferenças nos resultados devido a diferenças de gênero, raça, etnia, etc.

O exercício de definição destas subcategorias não pode ser considerado exaustivo e ressalta-se que vários desses estudos poderiam pertencer a mais de um dos subgrupos assinalados.

Descrição e análise dos resultados

Quanto ao nível acadêmico dos trabalhos, pode-se dizer que a maioria deles é de mestrado acadêmico (78,2%) sendo poucos, relativamente ao universo pesquisado, os trabalhos de doutorado (15%).

A maioria dos estudos encontrados (65%) foi produzida em instituições públicas de ensino, com destaque para as universidades federais (41,2%), que concentram os trabalhos de mestrado acadêmico (102 trabalhos) e estaduais (23,5%). As instituições privadas foram responsáveis por 35% dos trabalhos pesquisados, sendo a maioria deles realizada em nível de mestrado acadêmico (75 trabalhos). Os trabalhos de doutorado concentram-se nas universidades estaduais e federais, que produziram 16 e 15 teses, respectivamente.

Os mestrados profissionalizantes concentram-se na esfera privada, ainda que sejam poucos os programas a oferecer essa modalidade de estudos em nível de pós-graduação, enquanto que os acadêmicos foram, em sua maioria, realizados em instituições federais (44%) e privadas (32,5%).

No que se refere à distribuição de bolsas de estudo, percebe-se que os mestrados acadêmicos estão entre os que mais são beneficiados com os fomentos, especialmente pelas bolsas oferecidas pela Capes. Entretanto, a análise dos dados da pesquisa identifica que grande número de fichas da Capes (59,5%) não traz informação sobre se o trabalho obteve ou não algum auxílio. Assim, não é possível fazer inferências sobre a relação entre fomento e qualidade da pesquisa realizada, já que os dados existentes sobre bolsas de estudo ainda são pouco consistentes para permitir análises mais aprofundadas a esse respeito.

Sobre a distribuição temporal dos estudos, a partir da análise das datas de defesa dos trabalhos, é possível perceber que apesar das primeiras iniciativas de avaliação terem ocorrido no final dos anos 1980 e início dos 1990, elas passaram a ser foco de estudo apenas no final dos anos 1990, mas precisamente em 1998, quando foram localizados três estudos: dois sobre o sistema paulista de avaliação, o SARESP, e um

sobre o sistema nacional (SAEB). A produção intensifica-se durante o primeiro decênio dos anos dois mil, atingindo seu auge em 2010, ano em que foram produzidos 56 trabalhos sobre a temática em tela (Gráfico 1).

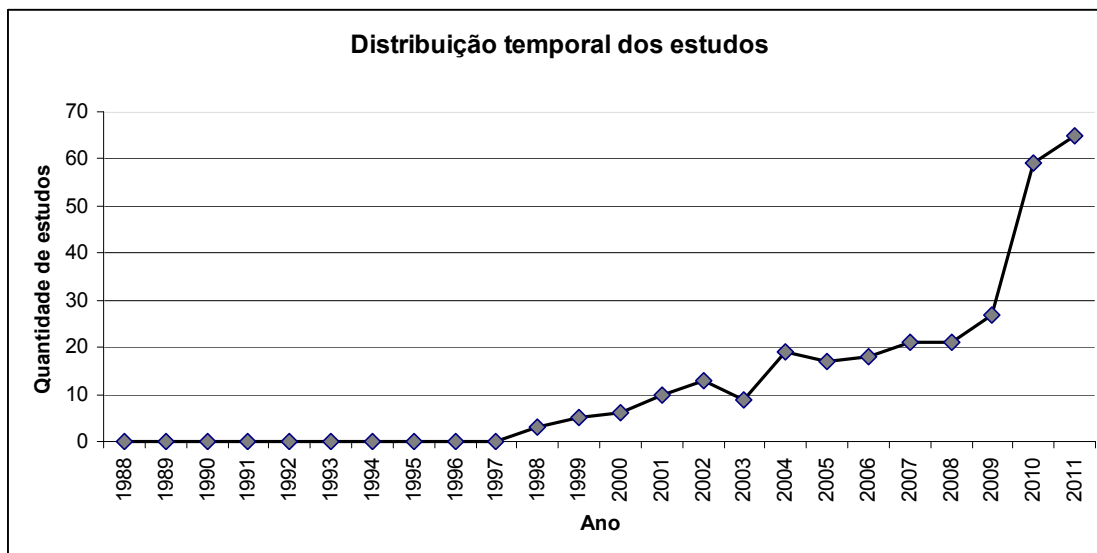


Gráfico 1 - Distribuição dos estudos produzidos ao longo dos anos pesquisados

Observa-se a existência de uma disparidade entre os níveis, ainda que se possa apontar que o primeiro trabalho de doutorado é defendido em 1999, sendo que os estudos do autor na área devem ter se iniciado anteriormente, em meados da década de 1990. Cabe destacar que a produção em nível de doutorado é pontual até 2002 e a expansão dos estudos neste nível começa a partir de 2003. Já os primeiros trabalhos de mestrado acadêmico começam a aparecer em 1998.

A partir desse período, o número de trabalhos tem aumentado, observando-se alguns picos de produção entre 2004 e 2005 e sua acentuação a partir de 2009, provavelmente devido à consolidação do Ideb e da Prova Brasil. É provável que a natureza das informações disponibilizadas de forma censitária por este índice e pela avaliação federal tenha contribuído para o aumento do interesse, no meio acadêmico, por seu estudo.

Em relação à distribuição dos estudos pelos grupos de análise ao longo do período que constitui o recorte da pesquisa, é possível inferir que o foco principal do interesse acadêmico parece ter sido as implicações dos sistemas de avaliação na escola (professores, alunos, gestão) e no próprio sistema educacional, temas que constituem o foco do Grupo 2. Esses trabalhos representam cerca de 35% do total dos estudos

selecionados (102 estudos), aumentando quantitativamente a partir de 2009. Dentre as muitas dimensões deste debate, encontra-se a influência da divulgação dos resultados sobre o seu uso e a relação entre os objetivos da avaliação sistêmica em relação às informações por ela produzidas, bem como a discussão política e ideológica do uso dos resultados.

Os primeiros trabalhos produzidos pertencem ao Grupo 1. Proporcionalmente, os estudos com esse foco de interesse contemplam 19% do total de estudos selecionados (57 trabalhos). Destaca-se que o interesse nessa temática parece diminuir no início dos anos 2000, mas, em alguns momentos, como no ano de 2005 a análise da política de avaliação volta a se destacar enquanto objeto de estudo. No período de 2010 e 2011 há um aumento de trabalhos no Grupo 1, que passam a explorar a Provinha Brasil e a introdução de sistemas de avaliação em pequenos municípios, ou se propõem a uma reflexão mais aprofundada sobre a institucionalização da política de avaliação no Brasil ou em sistemas estaduais de avaliação mais consolidados, como Minas Gerais, por exemplo.

Os trabalhos que se debruçam sobre aspectos teórico-metodológicos das avaliações (Grupo 3) constituem cerca de 14,0% do total de estudos encontrados (41 trabalhos). É possível notar que só mais recentemente houve uma intensificação dos trabalhos nessa categoria, notadamente a partir de 2008. Os trabalhos anteriores a esse período tinham como foco a discussão da metodologia da prova, das escalas de proficiência, das matrizes de referência e da Teoria da Resposta ao Item propriamente dita. Os estudos mais recentes debruçam-se sobre as metodologias de construção de indicadores para acompanhamento dos resultados e seu monitoramento.

A apropriação das avaliações de sistema como base de estudo de outras áreas disciplinares (Letras e Linguística, Matemática, Geografia e História) e como objeto de interesse de outras áreas que não a Educação também parece digna de nota, atingindo aproximadamente 31,6% dos estudos produzidos (Grupo 4). Três grandes áreas de conhecimento dominam este grupo, a Educação com 34 trabalhos, o que temos denominado de “Ensino de” com 19 trabalhos e a Economia, com 15 trabalhos (36,5%, 20,5% e 16%, respectivamente).

Foram identificados, ainda, estudos que se baseiam nos resultados das avaliações para realizar escolhas metodológicas no direcionamento da pesquisa. Um exemplo desse foco de interesse, por exemplo, são estudos que discutem os efeitos da progressão continuada ou a gestão democrática da escola e utilizam resultados das avaliações para

escolher as unidades escolares nas quais será realizada a pesquisa de campo. Estudos como esses foram desconsiderados por não focalizarem a discussão das avaliações de sistema.

Investigando o caminho destes trabalhos pelos programas de Pós-Graduação das instituições pode-se observar o domínio específico de programas de pós-graduação em Educação. Os 294 trabalhos selecionados estão distribuídos em 71 diferentes programas de pós-graduação em 76 diferentes instituições, sendo que a maioria dos trabalhos concentra-se em programas de Pós-Graduação em Educação (162 trabalhos) ou programas relativos à educação, administração, gestão e ensino em áreas específicas (como Matemática e Ciências), totalizando 25 trabalhos.

Cabe destacar que programas de Pós-Graduação em Economia são responsáveis por 22 dos trabalhos selecionados, enquanto que programas de Pós-Graduação em Psicologia e Psicologia da Educação produziram 13 trabalhos. Programas de Pós-Graduação em Linguística Aplicada, Língua Portuguesa, Letras ou Estudos da Linguagem produziram 19 trabalhos. Foram encontrados, ainda, dez trabalhos em programas diversos de mestrado profissional, seja da área de avaliação, economia, educação (com ênfase em ensino de matemática e ciências) ou administração.

Nota-se que dentre os 294 autores pesquisados foram identificados apenas 10 que deram continuidade à temática da avaliação de sistemas educacionais no mestrado e no doutorado dentro dos limites dos descritores pesquisados.

Além disso, há 217 diferentes orientadores nos 294 trabalhos que constam no banco de dados. Dentre eles, foi possível perceber que apenas 38 orientaram mais de uma produção. Desses 38 orientadores, 23 foram responsáveis pela orientação de apenas dois dos trabalhos pesquisados. A atividade de orientação dos outros 15 pesquisadores contribui para a elaboração de 23,5% do total dos trabalhos selecionados. Devido ao grande número de orientadores encontrados, as análises descritivas que se seguem focalizaram os orientadores com maior quantidade de orientações.

Em relação à inserção profissional daqueles que mais orientaram trabalhos sobre a temática é possível perceber que há concentração dos trabalhos em algumas instituições (UnB, UFJF, UFRJ, PUC-Rio, PUC-SP, FEUSP, UFBA, Unisinos, FEA-USP, FEA-RPUSP, UFRPE, UFPE, UNESP- Marília, Cesgranrio e UFMG).

As linhas de pesquisa que se destacam referem-se à avaliação educacional (teoria e metodologia), currículo, política e gestão educacional e didáticas ou metodologias específicas de ensino.

Quanto aos membros de banca, foram encontrados 522 diferentes acadêmicos na composição das bancas dos trabalhos analisados. Destes, apenas 99 tiveram participação em mais de uma banca na temática, sendo que são poucos os que possuem presença em três ou mais bancas. É possível perceber que aqueles que mais orientam trabalhos nas temáticas pesquisadas (presidente da banca) nem sempre são os que mais foram convidados para serem membros de bancas examinadoras. Há apenas nove pessoas com participação constante nessas duas atividades e que concentram 12% das orientações ocorridas no período em estudo.

Comentários sobre as palavras-chaves e os resumos

Considerando as palavras-chave como termos que resumem os temas e idéias principais de um texto, observou-se alguns aspectos problemáticos do registro desse componente na Ficha Capes⁹: incompletude de palavras-chave (indicação de menos de três delas); palavras sem sentido ou digitadas de forma incompleta, já que há espaço limitado de caracteres nos campos destinados a essa informação; divergências das palavras assinaladas na ficha Capes e nos trabalhos.

Uma leitura acurada dessas palavras¹⁰ permitiu organizá-las em algumas categorias, assim definidas: avaliação educacional; ensino e aprendizagem; formação de professores; métodos e técnicas de análise; planejamento, gestão e administração educacional e política educacional. As palavras-chave que não eram consideradas adequadas para nenhuma dessas categorias foram agrupadas sob o termo “genérico”. A tabela 2 traz a distribuição das palavras-chave dos trabalhos selecionados pelas categorias supracitadas, destacando quantas vezes palavras relacionadas a cada categoria foram utilizadas.

Observa-se que muitas das palavras-chave utilizadas nos trabalhos são genéricas demais não auxiliando o pesquisador na compreensão mais acurada da natureza do trabalho e, portanto, a definição desses termos mereceria mais atenção por parte dos autores e orientadores, tendo em vista que fazem parte dos componentes que permitem identificar o trabalho e selecioná-lo.

⁹ Na ficha Capes, tradicionalmente registram-se até três palavras-chaves para cada trabalho.

¹⁰ As análises tomam como base 204 palavras referentes aos trabalhos dos Grupos 1 e 2.

Tabela 2 – Organização das palavras-chave dos trabalhos em categorias de análise

Categorias de análise das palavras-chave	N
Avaliação educacional	78
Ensino e aprendizagem	41
Formação de professores	13
Genérico	89
Métodos e técnicas de análise	30
Planejamento, gestão e administração educacional	43
Política educacional	156
Total	450

Nesse trabalho considerou-se que um resumo informativo não deveria apenas dizer do que trata a pesquisa, mas informar, sinteticamente, as finalidades da pesquisa, a metodologia utilizada e os resultados atingidos. Assim, foi definido um padrão de estrutura mínima esperado, composto de quatro elementos básicos:

- 1) Definição do objeto da pesquisa
- 2) Metodologia utilizada em caso de pesquisa empírica
- 3) Resultados da pesquisa
- 4) Considerações finais

Adicionalmente, buscou-se observar se os resumos traziam indicativos sobre o referencial teórico utilizado, dado que este aspecto é controverso nas indicações que se referem à estrutura deste tipo de texto. Nota-se que muitos resumos dos trabalhos selecionados não contemplam algumas dessas características, dificultando a compreensão do escopo da pesquisa

Quanto ao conjunto dos elementos básicos componentes dos resumos, pode-se perceber que estavam presentes em 99 dos trabalhos (34%)¹¹. Três trabalhos (1%) não contemplaram nenhum dos elementos básicos mencionados acima, não oferecendo suporte ao entendimento do que se trata a pesquisa. Os 190 resumos restantes (65%) estavam incompletos. Observa-se, assim, a necessidade de um maior cuidado, por parte de orientandos e orientadores, para com esse texto.

No que se refere à questão metodológica, procurou-se identificar se os autores definiam a abordagem assumida no trabalho (considerando-se as abordagens mais gerais, ou seja, qualitativa, quantitativa e mista). O caráter hegemônico da abordagem

¹¹ Somente 292 trabalhos foram analisados quanto à estrutura de seus resumos, visto que dois deles não apresentaram esse texto.

qualitativa (base de 124 trabalhos, ou 42,5%) é significativo em pesquisas na área das Ciências Humanas. Em contrapartida, a abordagem quantitativa está anunciada em 18 trabalhos (6%), enquanto que pesquisas que associam abordagens qualitativas e quantitativas (abordagem mista) correspondem a 17 trabalhos (6,0%). É preocupante o número de trabalhos que não anunciam, com clareza, a perspectiva metodológica utilizada (133, ou 45,5%), muitas vezes apresentando uma confusão entre técnicas e instrumentos utilizados e a abordagem metodológica em um sistema mais amplo.

Percebe-se, ainda, que a grande maioria dos resumos não apresentou referenciais teóricos (236 ou 80,8%) ou, quando isso ocorre, os autores tendem a citar fontes de pesquisa ou autores utilizados na construção do quadro teórico e não referenciais teóricos no sentido mais amplo.

Cabe ressaltar que a leitura dos resumos possibilitou observar um uso praticamente indistinto dos termos “avaliação de sistemas educacionais”, “sistemas de avaliação”, “avaliação externa” e “avaliação de rendimento dos alunos”. Esses termos foram tomados como sinônimos por diversos dos autores dos trabalhos pesquisados, sendo possível inferir que não parece haver consenso, entre os que produzem conhecimento acerca dessas avaliações, sobre significado conceitual de cada um deles.

Considerações finais

Ao produzir um banco de dados com estudos sobre avaliação de sistemas educacionais, pode-se perceber a necessidade de realizar uma meta-análise da produção acadêmica e aprimorar as formas existentes de disseminação dessa produção.

A análise da produção acadêmica dos alunos de pós-graduação do Brasil no período de 1988 a 2011 revelou interesse crescente dos pesquisadores pela temática das avaliações de sistemas educacionais, tanto na área da Educação, quanto em outras áreas do conhecimento. Esse interesse pode ser indicativo de preocupações como 1) identificar e compreender o papel da avaliação de redes de ensino na evolução da qualidade da escola pública; 2) contribuir para o desenvolvimento teórico-metodológico das avaliações em curso no país; 3) compreender as implicações dessas avaliações para a política educacional, para os programas e projetos educacionais existentes e para as escolas e seus agentes. Ao analisar os 294 trabalhos selecionados, identificam-se as instituições e intelectuais que, dialogando com a implantação de uma cultura da

avaliação, buscam compreendê-la em sua relação com outras políticas educacionais, como as de currículo, de formação de professores, de distribuição de materiais didáticos.

No entanto observam-se lacunas que a produção acadêmico-científica tem tido pouca preocupação em preencher: poucos são os estudos que discutem o planejamento das avaliações, seus descritores, adequação às políticas mais amplas, tanto em suas especificações técnicas, quanto no que se refere aos posicionamentos políticos que subjazem as escolhas teórico-metodológicas realizadas. Escassos também são os estudos que procuram iluminar as possíveis relações entre os resultados de avaliações internas e externas e como estes são significados e apropriados nas escolas.

Ao finalizar esse mapeamento, fica evidente a importância das sínteses integrativas da produção para o acompanhamento da constituição de um campo de estudos e para a identificação dos aspectos já consolidados e dos que carecem de mudanças, principalmente no que se refere às redundâncias na produção, ao rigor teórico-metodológico das pesquisas, aos nichos de produção de conhecimento existentes nas instituições de pós-graduação e ao uso de referenciais teórico-metodológicos adequados aos objetivos e problemas de pesquisa propostas. Tais informações podem possibilitar o aperfeiçoamento dos estudos pós-graduados e contribuir para a qualidade da pesquisa acadêmica.

Referências

AFONSO, A. **Avaliação Educacional: regulação e emancipação**. São Paulo: Cortez, 2000.

ANDRÉ, M. Pesquisa e educação: buscando rigor e qualidade. **Cadernos de Pesquisa**, n.113, p.51-64, julho/2001.

ALVES, J.A. A “revisão da bibliografia” em teses e dissertações: meus tipos inescapáveis. **Cadernos de Pesquisa**, São Paulo, n.81, p. 54, maio 1992.

BARRETTO, E. S. S; PINTO, R. P. (Coord.). **Avaliação na educação básica (1990-1998)**. Brasília: MEC/INEP/Comped, 2001. Série Estado do Conhecimento, n.4.

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

BORGES, R. M.; CALDERÓN, A. I. Avaliação na Educação Básica tendências temáticas predominantes na Revista Ensaio: Avaliação e Políticas Públicas em

Educação (1999-2008). In: **Anais do 10º Encontro de Pesquisa da Região Sudeste**, 2011. Disponível em: http://www.fe.uftj.br/anpedinha2011/trabalhos/POSEDUC_PUCCAMPINAS_744.440.852-15_trabalho.doc. Acesso em 19/08/2011.

BROOKE, N. P.; CUNHA, M. A.; FALEIROS, M. A. Avaliação externa como instrumento da gestão educacional nos estados: **Relatório Final**. GAME/ Fundação Victor Civita, agosto/2011. Disponível em: <http://www.fvc.org.br/pdf/relatorio-avaliacoes-externas.pdf>, Acesso em 27/12/2012.

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

FREITAS, D. N. T. Avaliação da educação básica e ação normativa federal. **Cadernos de Pesquisa**, São Paulo: Fundação Carlos Chagas/Autores Associados, v. 34, n. 123, p. 663-689, set/dez 2004.

GATTI, B. Implicações e perspectivas da pesquisa educacional no Brasil contemporâneo. **Cadernos de Pesquisa**, n. 113, p. 65- 81, julho/ 2001.

LAROCCA, P.; ROSSO, A. J.; SOUZA, A. P. A formulação dos objetivos de pesquisa na pós-graduação em Educação: uma discussão necessária. **Revista Brasileira de Pós-Graduação**, Brasília: CAPES, v. 2, n. 3, p. 118-133, mar. 2005.

MARTINS, A.; SOUSA, S.M.Z.L. Avaliação educacional e gestão de sistema e de escolas: estado do Conhecimento (2000-2008). **XXV Simpósio Brasileiro e II Congresso Ibero-Americano de Política e Administração da Educação**. São Paulo, 26 a 30 de abril de 2011.

POLTRONIERI, H.; CALDERÓN, A. I. Avaliação em educação: a revista Estudos em Avaliação Educacional. **Estudos em Avaliação Educacional**, v. 23, n. 53, set/ dez., p. 82- 103, 2012.

SOUSA, S. Z.; PIMENTA, C. O.; MACHADO, C. Avaliação e gestão municipal da educação. **Estudos em Avaliação Educacional**, São Paulo, v. 23, n.53, p. 14-36, set./dez. 2012.

VIANNA, H. **Avaliações nacionais em larga escala: análises e propostas**. São Paulo: Fundação Carlos Chagas, 2003.

WARDE, M. O papel da pesquisa na pós-graduação em educação. **Cadernos de Pesquisa**, n. 73, p. 67- 75, 1990.