

# AVALIAÇÃO EDUCACIONAL E MODELOS DE VALOR ACRESCENTADO: TÓPICOS DE REFLEXÃO

MARIA EUGÉNIA FERRÃO\*

*RESUMO:* Desde o início do século XXI deparamo-nos com uma onda crescente na implementação de sistemas de avaliação educacional que têm como propósito a responsabilização – prestação de contas (*accountability*) –, com consequências directas ao nível das escolas e dos professores (*high-stakes*). Assistimos a desenvolvimentos muito rápidos, quer em termos da influência da avaliação educacional na definição de políticas públicas, quer na sua aplicação para a avaliação destas políticas. Neste contexto, o uso do conceito de valor acrescentado em educação é consensualmente aceite. O artigo apresenta algumas reflexões sobre aspectos da sua operacionalização e uso, trazendo à liça temas tais como a escolha do modelo, resultados escolares, caminhos para a apropriação pela comunidade educativa. Adicionalmente, descrevem-se exemplos onde o uso inapropriado de indicadores de desempenho teve consequências imprevistas e adversas para a educação.

*Palavras-chave:* Prestação de contas. Modelo de valor acrescentado. Avaliação educacional.

## EDUCATIONAL EVALUATION AND MODELS OF VALUE ADDED: TOPICS FOR REFLECTION

*ABSTRACT:* Since the beginning of the 21st century that we face a growing wave towards the implementation of educational evaluation systems with high-stake accountability purposes at the level of teachers and schools. The development and implementation of such systems is quite fast, either in terms of the influence of the educational evaluation on public policies, or in terms of the public policy evaluation. In this context, the use of the value-added concept is consensually accepted. This paper presents some reflections on its implementation and use, emerging themes such as the model choice, school and student outcomes, and its appropriateness by practitioners. In addition, some examples are given where the misuse of performance indicators had adverse consequences to education.

*Key words:* Accountability. Value-added model. Educational evaluation.

---

\* Doutora em Teoria de Controle e Estatística e professora da Universidade da Beira Interior (Portugal). *E-mail:* meferrao@ubi.pt

EVALUATION EDUCATIONNELLE ET MODELES DE VALEUR AJOUTEE:  
THEMES DE REFLEXION

**RÉSUMÉ:** Depuis le début du XXI<sup>ème</sup>, siècle nous nous rencontrons face à un mouvement croissant dans la mise en oeuvre de systèmes d'évaluation éducationnelle qui a pour but la responsabilisation – reddition de comptes (accountability) –, ayant des conséquences directes au niveau des écoles et des professeurs (high-stakes). Nous assistons à des développements très rapides, soit en termes de l'influence de l'évaluation éducationnelle dans la définition des politiques publiques, soit dans l'application pour l'évaluation de celles-ci. Dans ce contexte, l'utilisation du concept de valeur ajoutée en éducation est consensuellement acceptée. L'article présente quelques réflexions sur des aspects de sa mise en oeuvre et utilisation, mettant en débat des thèmes comme le choix du modèle, les résultats scolaires, les chemins pour l'appropriation par la communauté éducative. En outre, on y décrit des exemples où l'utilisation inappropriée des indicateurs de performance ont eu des conséquences imprévues et adverses pour l'éducation.

**Mots-clés:** Reddition de comptes. Modèle de valeur ajoutée. Evaluation éducationnelle.

## Introdução

**H**á algumas décadas que se verifica, em diversos países, grande insatisfação quanto aos resultados dos sistemas educativos face à despesa pública feita em educação. Em 1988, a reforma educativa realizada em Inglaterra, País de Gales e Irlanda do Norte constituiu um marco temporal para o advento da responsabilização e prestação de contas (*accountability*) como função principal da avaliação educacional. A tendência veio a acentuar-se com a publicação da Lei *No Child Left Behind* nos Estados Unidos da América, em 2001. Na Europa, no mesmo ano, a responsabilização e prestação de contas foi um dos princípios enunciados para a boa governança das instituições, a par de outros quatro, a saber: abertura, participação, eficácia e coerência (Comissão das Comunidades Europeias, 2001). Uma das faces visíveis das mudanças na educação foi a publicação periódica dos resultados escolares sob a forma de *rankings* como instrumento para a comparação do desempenho institucional. Deste modo, a responsabilidade pelo sucesso ou insucesso educativo foi centrada na escola e no seu pessoal docente. Aumentou quer o número de estudos que visam a comparação do desempenho, quer o volume de investimento realizado nesses estudos. Assim, o tema ganhou rapidamente a atenção dos meios de comunicação social e da classe política, cujo discurso associa a necessidade de avaliação de desempenho à melhoria da qualidade da educação.

Datada de 1990, 42 anos passados sobre a Declaração Universal dos Direitos do Homem,<sup>1</sup> a Declaração da Educação para Todos (Unesco, 1990), resultante da Conferência Mundial de Jomtien, reafirma o princípio fundamental do direito de todos à

educação e enuncia explicitamente, no artigo 3º, que, para essa finalidade, os serviços de educação básica de qualidade devem ser expandidos e medidas consistentes devem ser tomadas para reduzir as disparidades. Não obstante o reconhecimento da necessidade de garantir o acesso à educação de qualidade, a concepção de qualidade da educação é tudo, menos consensual (ver, p. ex., Unesco, 2005). A diversidade de interpretações do que seja a qualidade da educação é, talvez, o principal obstáculo à obtenção de convergência de entendimento no que se refere: (1) à explicitação de objectivos e metas educacionais; (2) à identificação e descrição das várias dimensões envolvidas na análise e monitoramento da qualidade da educação; (3) à identificação das variáveis envolvidas e sua mensuração; (4) ao diagnóstico atempado das disparidades geradoras de iniquidade; (5) à concepção de um sistema integrado que vise a melhoria do sistema educativo e, por conseguinte, da qualidade da educação provida à população.

O grande desígnio da qualidade da educação tem, assim, justificado o interesse crescente na implementação de sistemas de avaliação que permitam identificar os pontos fortes e fracos do(s) sistema(s) educativo(s) e, conseqüentemente, monitorizar as mudanças ao longo do tempo. A monitorização é impossível sem indicadores de desempenho. Contudo, reduzir os sistemas de avaliação à produção e disseminação de *rankings* elaborados com base em um ou outro indicador, sem cuidar da informação contida nesse indicador e do quão bem ele representa a realidade para mensuração da qual foi concebido, pode ter efeitos catastróficos na educação. A comparação do desempenho pode ser bastante produtiva para os envolvidos quando é bem feita. Quando mal feita, pode ser muito custosa, e não apenas inútil mas prejudicial e até mesmo destrutiva (Bird et al., 2005).

A par da crítica que surgiu em torno da publicação dos *rankings*, desenvolveram-se alternativas metodológicas para a produção de indicadores de desempenho escolar, entre as quais a abordagem de valor acrescentado. O conhecimento científico associado a tais alternativas ainda se depara com algumas limitações, matéria em pauta para a investigação científica. Dependendo do uso e conseqüências de tais sistemas, a omissão dessas limitações pode representar, na atualidade, forte adversidade à consecussão do grande desígnio. Em total concordância com a afirmação em Braun, Chudowsky e Koenig (2010), não é realista pensar que os mentores e decisores de política educativa esperarão mais 20 anos até que todas as limitações e dificuldades estejam resolvidas para então fazerem uso dos métodos de valor acrescentado. “A co-responsabilidade na melhoria da educação pode e deve existir. (...) A comunidade científica tem a sua quota de responsabilidade, particularmente na transferência do conhecimento científico” (Ferrão, 2009, p. 52). Todavia, importa invocar a ideia de avaliação educacional como uma actividade que acontece no contexto político (e.g. Weiss, 1975). A produção de conhecimento científico que supere aquelas limitações e dificuldades está fortemente dependente da política para a ciência, em especial da que concerne ao incentivo e

financiamento das subáreas em que os avanços determinantes são obstaculizados pela falta de massa crítica. Adicionalmente, levar em conta que as políticas e programas avaliados são criações de decisão política; que a avaliação tem por objectivo último dar *feedback* para a tomada de decisão e, como tal, os seus resultados têm implicações na arena política; que a avaliação é, por natureza, uma actividade política na perspectiva, menos consensual, de que pode ou não legitimar políticas ou programas.

As reflexões apresentadas ao longo do artigo incidem na articulação entre a avaliação educacional, a qualidade da educação e a abordagem de valor acrescentado; no uso e finalidades da abordagem de valor acrescentado; e na sua relação com os critérios da avaliação. Atentaremos com prudência sobre as suas consequências, em particular as que decorrem das limitações ainda não superadas pelo conhecimento científico e tecnológico.

## Avaliação e valor acrescentado

De acordo com uma das primeiras definições, a avaliação educacional trata do processo de determinar em que medida os objectivos educacionais estão efetivamente a ser atingidos (*The process of determining to what extent educational objectives are actually being realized* [Tyler, 1950, p. 69]). A classe de modelos estatísticos que permite estimar a contribuição das escolas para o progresso dos alunos, no que se refere aos objectivos educativos estabelecidos, é designada por modelos de valor acrescentado (e.g. OECD, 2008). Em outros termos, mas com significado semelhante, a *National Academy of Sciences* menciona que a abordagem se refere à quantificação do efeito da escola, professor ou programa no desempenho dos alunos, considerando as diferenças entre os alunos quanto ao conhecimento prévio e também a outras variáveis extra-escolares que influenciam o desempenho dos alunos (Braun, Chudowsky & Koenig, 2010). Braun e Wainer (2006) apresentam-no-los como sendo a família de modelos estatísticos que são usados para inferir sobre a eficácia das unidades educativas, usualmente escolas e/ou professores.

Estas definições são convergentes num conjunto de características sobre as quais vale a pena refletir. Vale a pena refletir sobre aspectos relacionados com a noção de progresso, com a escolha do modelo de valor acrescentado, com a ligação entre resultados escolares e a missão da escola.

### A noção de progresso

O progresso está associado a algo cujo estado muda ao longo do tempo. Para quantificar essa mudança é necessário recolher dados repetidas vezes no tempo – dados longitudinais –, de tal modo que seja possível captar a mudança decorrente da

aprendizagem realizada de permeio. Ou seja, torna-se necessário recolher dados, repetidamente ao longo do tempo, sobre o desempenho do aluno. O desempenho do aluno refere-se à demonstração duma habilidade, aptidão, capacidade ou conhecimento num determinado momento do tempo. Tal é realizado usualmente através de testes padronizados, por conseguinte a aferição das aprendizagens e do progresso é baseada em testes padronizados. Por um lado, a infraestrutura técnica e operacional para a realização de estudos longitudinais é exigente. Por outro, a construção da métrica que permita quantificar o progresso não é trivial. Não se trata portanto de uma dificuldade específica dos modelos de valor acrescentado, mas antes de uma dificuldade inerente ao processo de aferição das aprendizagens e da sua mudança ao longo do tempo. Por este motivo, alguns investigadores propõem a designação *comparação ajustada* por traduzir melhor a transposição do conceito de valor acrescentado da economia para a educação (Goldstein & Spiegelhalter, 1996). Numa outra perspectiva, Linn (2000, p. 14) alerta para o facto da finalidade da aplicação dos testes poder afetar não só o processo como também os resultados obtidos:

Eu sou levado a concluir que, na maioria dos casos, os instrumentos e tecnologia têm sido usados erradamente para os propósitos de responsabilização – prestação de contas (*accountability*) com consequências directas de forte impacto (*high-stakes*). Os sistemas de aferição, que são meios de monitorização úteis, perdem a sua confiabilidade e credibilidade quando são usados com consequências directas de forte impacto sobre os profissionais ou instituições. Os efeitos negativos inesperados decorrentes de tais usos frequentemente ultrapassam os efeitos positivos esperados. (Tradução livre da autora)

O aspecto do *uso* dos resultados da avaliação e das suas consequências no processo de avaliação será retomado adiante. Por ora, vale mencionar os novos desafios de investigação científica e de desenvolvimento metodológico na classe de modelos de valor acrescentado que derivam dessas dificuldades, tais como a linha de investigação que visa o ajuste ou a correcção do erro da medida nos modelos de valor acrescentado (e.g. Ferrão & Goldstein, 2009; Goldstein, Kounali & Robinson, 2008; Ladd & Walsh, 2002; Woodhouse, Yang & Goldstein, 1996).

## A escolha do modelo de valor acrescentado

Outra característica é a existência duma família ou classe de modelos. Assim, o “modelo de valor acrescentado” denota uma escolha entre diversos modelos estatísticos possíveis. A escolha deve ser realizada em concordância com o critério de comparação estabelecido pelo *designer* do sistema de avaliação, de modo adequado ao contexto educativo, social, económico e cultural em que o modelo vai ser aplicado, articuladamente com as opções políticas e/ou programáticas. Este aspecto está relacionado com os cuidados a ter para tornar comparável a acção educativa das escolas, já que cada uma delas serve à população discente com características diferentes. Para

estimar a “contribuição da escola” ou a “eficácia das unidades educativas”, é absolutamente necessário ter em mente que a população discente que frequenta a rede pública de escolas é muito heterogênea e que essa heterogeneidade também se verifica entre escolas. Logo, é preciso que o modelo contemple variáveis extra-escolares que têm influência no desempenho dos alunos e, como tal, garanta que a estimativa da “contribuição da escola” esteja livre dessa influência. Por exemplo, a variável que representa o nível socioeconómico do aluno deve ser incluída no modelo e, eventualmente, a variável que capta a composição da população discente também deve ser contemplada.

Estudiosos em avaliação educacional são unânimes em reconhecer a importância de dispôr de vários indicadores para formar juízo de valor relativamente ao(s) critério(s) formulado(s) no sistema de avaliação e, assim, suportar adequadamente decisões de política educativa. Não se espera que as “melhores” escolas de acordo com um indicador sejam também as “melhores” de acordo com outro indicador. Cada um deles pode revelar uma faceta diferente do mesmo objecto. Por exemplo, se o objectivo é demonstrar que os alunos atingem determinado nível de desempenho estabelecido como meta, então o indicador de *status* (reflete o desempenho, resultados escolares, num determinado momento do tempo) é o mais apropriado. Convém, todavia, levar em conta que algumas das escolas, onde os alunos atingem níveis de desempenho elevado, pouco contribuem para esse desempenho. Geralmente, são escolas frequentadas por alunos com elevado nível de desempenho à entrada (escolas selectivas). Essas crianças e jovens são normalmente provenientes de grupos sociais favorecidos. Em contrapartida, as escolas que recebem alunos com baixo desempenho e onde, no final do período sob avaliação, o desempenho se mantém aquém da meta, podem, apesar disso, ter contribuído fortemente para alavancar a aprendizagem desses alunos e, portanto, apresentam elevado valor acrescentado. A sua contribuição não está reflectida no indicador de *status*, mas sim no indicador de valor acrescentado. Este é o caso típico das escolas cuja população discente é extremamente desfavorecida do ponto de vista socioeconómico. Um sistema de incentivos orientado pelo indicador de *status* tende a premiar escolas com reduzido valor acrescentado e a desvalorizar o serviço educativo prestado por aquelas que trabalham junto de populações onde a escola marca a diferença.

Em síntese, as decisões e escolhas inerentes à especificação funcional do modelo estatístico devem consubstanciar, tanto quanto possível, as opções de natureza política e/ou programática estabelecidas pelos governos. Assim, de modo estruturado e formal, o sistema de avaliação educacional, concebido nas dez dimensões – definição; função; objecto; tipo de informação a recolher face a cada objecto; critério para formar juízo de valor; a quem se destina a avaliação (*stakeholders*); processo; método; quem é avaliador e constituição das equipas de avaliação; padrões que regem uma avaliação (Nevo, 1983; 2006) –, acomoda diversas e complementares

vertentes de avaliação, tais como a avaliação das aprendizagens, de métodos de ensino, de materiais, do desempenho de pessoal docente, da avaliação institucional, da auto-avaliação, da avaliação externa, da avaliação de programas e de políticas. Instanciar cada uma dessas dimensões depende da vertente de avaliação em apreço. Como exemplo, o critério no âmbito da auto-avaliação da escola ou da avaliação de materiais pode não coincidir com o critério no âmbito da avaliação institucional. Também pode não coincidir face aos potenciais interessados no uso da avaliação (e.g. Raudenbush & Willms, 1995). Isto é, ainda no âmbito da avaliação institucional, o critério com interesse para uso de pais e encarregados de educação na escolha da escola não é o mesmo critério a aplicar por governantes, por directores de escola ou por professores. Os pais desejam que a escola produza o maior progresso possível, independentemente de como ela o consegue produzir – quer seja através de professores altamente qualificados, do método de ensino praticado, da influência da comunidade na qual a escola está inserida, dos recursos materiais disponíveis, etc. Assim, interessa-lhes saber qual é o desempenho escolar esperado, condicionado às características individuais dos seus educandos. Já o critério que represente o efeito da prática docente e escolar na aprendizagem, “isolado” o contexto da escola, deve ser o critério de interesse para todos aqueles que têm a responsabilidade de promover a aprendizagem de todos os alunos, independentemente das suas características individuais e antecedentes (Ferrão, 2008). Note que o contexto da escola inclui fatores que, apesar de serem exógenos a ela e ao esforço educativo levado a cabo pelos seus profissionais, podem ser catalizadores ou inibidores dos ambientes propícios à aprendizagem. Nestes termos, a escolha do modelo de valor acrescentado do qual resulte(m) o(s) indicador(es) a ser(em) usado(s) como critério no sistema de avaliação educacional deve atender à complexidade inerente ao sistema educativo.

## Resultados escolares

Alguma resistência à abordagem de valor acrescentado assenta na crítica à (suposta) unidimensionalidade do indicador face a múltiplas dimensões do desenvolvimento do aluno. As metodologias estatísticas permitem considerar múltiplas dimensões dos resultados escolares, sejam elas aferidas através de variáveis contínuas ou discretas. Geralmente, para o cômputo usa-se a classificação dos alunos em testes padronizados como medida dos resultados escolares. Adicionalmente, há estudos que consideram também o comportamento disciplinar dos alunos, a sua assiduidade, a auto-estima, etc. Aquela crítica está, portanto, desfocada. Levanta-se então a questão: que resultados escolares? O propósito da educação é assegurar que crianças e jovens se tornem cidadãos aptos nas dimensões pessoal, social, política e económica. Nestas dimensões há aprendizagens que são da estrita esfera da escola enquanto outras não o são. A valia da abordagem de valor acrescentado e o seu uso ao serviço da melhoria da

qualidade da educação dependem, assim, da definição do papel da escola e dos objectivos educativos que são da sua inteira responsabilidade (e.g. Biesta, 2009). Depende, também, da criação da infraestrutura necessária para garantir a qualidade dos dados e níveis aceitáveis de validade, fiabilidade e comparabilidade das medidas de resultados escolares. Erros amostrais e não-amostrais podem ser minimizados, mas nunca a situação ideal é alcançável em estudos de avaliação educacional ou outros estudos em ciências sociais. Portanto, é preciso acautelar as limitações deles decorrentes. A este propósito, Cronbach e Suppes (1969, p. 15) dizem o seguinte:

Qualquer que seja a natureza do estudo, se é rigoroso, o investigador antecipadamente estabeleceu as questões pertinentes, instituiu controles nas etapas de recolha de informação e o raciocínio associado a essas questões, de modo a evitar possíveis fontes de erro. Se os erros não puderem ser evitados, o investigador deve levá-los em consideração, discutindo as limitações das suas conclusões. (Tradução livre da autora)

## Uso e finalidades

A finalidade da avaliação não é necessariamente a criação de conhecimento (Alkin, 2003). A avaliação educacional é orientada à tomada de decisão, estabelecendo implicitamente os seus potenciais utilizadores – aqueles para quem a avaliação é relevante no que respeita à fundamentação das decisões (*stake-holders*). Da revisão do tema *Evaluation use revisited* elaborada por Hofstetter e Alkin (2003) é possível identificar quatro usos da avaliação educacional: o instrumental, o conceptual, o de substituição e o de legitimação. O primeiro uso está associado à acção; o segundo, ao melhor entendimento dos fenómenos e dos sistemas; o terceiro está associado à sinalização de que algo está a ser feito perante um problema, através da avaliação, mas as acções/intervenções necessárias são adiadas ou ignoradas; finalmente, o mais comum, o de legitimação, quando os dados ou resultados são usados selectivamente para fundamentar publicamente uma decisão tomada. Estes últimos usos têm paralelo nos modelos tático e político, respectivamente, da tipologia proposta por Weiss (1977, apud Hofstetter & Alkin, 2003) para o uso dos achados da investigação em ciências sociais na decisão política. É possível que o uso da avaliação educacional se tenha complicado na atualidade como consequência do aumento das expectativas por parte dos múltiplos interessados nos resultados da avaliação. A não antecipação dessas expectativas na fase de planeamento do sistema de avaliação, nos termos em que anteriormente aludimos, citando Cronbach e Suppes (1969), poderá levar ao uso inapropriado dos resultados da avaliação. O uso incorrecto remete para questões de ética de todos os potenciais interessados.

Existe a tendência para considerar o uso incorrecto a par do uso e do não-uso. O uso e uso incorrecto operam em diferentes dimensões. A primeira é um continuum

do não-uso para o uso da avaliação; a segunda é um continuum desde o não-uso incorrecto para o uso incorrecto. Ou seja, o uso incorrecto não está no sentido oposto do uso da avaliação – é um conceito relacionado com a forma do uso (Hofstetter & Alkin, 2003; Patton, 1988). O uso incorrecto dos resultados da avaliação educacional pode ficar a dever-se à omissão do sistema de avaliação em atender a todos os seus potenciais interessados.

Nos termos ora definidos, o indicador de valor acrescentado ajuda a responder a perguntas tais como: “Qual é a contribuição da escola X (ou do professor X) para a aprendizagem (ganho ou progresso) dos alunos, comparativamente com a contribuição das restantes escolas?”, “Quais são as escolas que sistematicamente mais contribuem para a aprendizagem dos alunos?”. Algo consensual em termos internacionais, diversos países têm vindo a contemplar o valor acrescentado nos respectivos sistemas de avaliação educacional, nem sempre coincidindo quanto ao uso e finalidade. De um modo geral, as principais finalidades são: (1) incorporar a lógica da responsabilização e da prestação de contas nas escolas; (2) fornecer informação para a escolha da escola; (3) fomentar a implementação de programas de melhoria da escola.

Estas finalidades diferem quanto ao foco e quanto aos programas de desenvolvimento. Todavia, elas podem não ser mutuamente exclusivas, dependendo, entre outros aspectos, do enquadramento conceptual onde se insere o programa. A responsabilização – prestação de contas – deve ser globalmente entendida nas suas dimensões política, pública, de gestão e administração, profissional e pessoal (e.g. Sinclair, 1995). A responsabilização e, por conseguinte, a respectiva prestação de contas, atribuídas ao governo, departamentos centrais e regionais, encarregados de educação e alunos, devem ser tais que, articuladas e cumulativamente, sejam criados os meios, o contexto e o apoio para que as escolas possam maximizar as aprendizagens efectuadas por cada um dos seus alunos. Com este fim em mente, o âmbito da responsabilização – prestação de contas – em cada nível de intervenção deve ser estabelecido coerentemente, mobilizando todos os intervenientes à participação. A afirmação de Sinclair (op. cit., p. 219), “Accountability is subjectively constructed and changes with context”, é actual, apesar das entrevistas a dirigentes terem sido realizadas em 1990, portanto há mais de duas décadas. Em alguns países, as dimensões política e de gestão e administração estão apostas à dimensão profissional através da atribuição de prémios de desempenho/mérito ou outros incentivos individuais ou institucionais. Estes mecanismos de prestação de contas ganharam muita popularidade entre governantes e mentores de políticas públicas. Os argumentos a favor são principalmente de ordem racional, não havendo, até à data, demonstração da existência de uma relação causa-efeito entre a prática de incentivos, por exemplo do tipo *performance-related pay*, e a melhoria das aprendizagens. Mesmo supondo que os incentivos individuais ou institucionais se baseiam em múltiplas medidas

do desempenho ou do mérito, 100% justas e 100% fiáveis, convém levar em consideração a natureza da profissão docente e da missão da escola, bem como o efeito da competitividade gerada entre profissionais pelo mecanismo de incentivos individuais e/ou institucionais. Ora, a competitividade, que é salutar em outros sectores de actividade, na educação pode ser um obstáculo ao trabalho colaborativo, sendo este considerado factor determinante do sucesso educativo. Note-se que o processo de criação de valor, produção e transmissão de conhecimento é acentuadamente marcado por externalidades positivas, resultantes da cooperação e difusão de conhecimento e práticas de aprendizagem colectiva, e não da retenção desse conhecimento como estratégia de reforço do posicionamento individual dos profissionais perante avaliações futuras.

A título de exemplo, o estado federado da Carolina do Norte foi um dos precursores a implementar um programa de prestação de contas com incentivos individuais, alicerçado num modelo de ganho do desempenho dos alunos (modelo semelhante ao valor acrescentado, mas não inclui o controle de variáveis extraescolares do tipo socioeconómico). O relatório oficial publicado em 2002 relata que, entre 1992/93 e 2000/01, aumentou a percentagem dos alunos com classificação (em leitura e matemática) igual ou superior à esperada para o respectivo ano de escolaridade e que o aumento foi mais pronunciado após a entrada em vigor do sistema estadual de prestação de contas. No entanto, Clotfelter et al. (2004) apresentam evidência de que, comparando o período posterior à entrada em vigor do sistema de prestação de contas com o período anterior, as escolas cuja população discente se caracterizava por maior proporção de alunos com défice educativo (ou menor desempenho) passaram a ter maior dificuldade em atrair e reter “bons” professores. A diminuição da estabilidade do corpo docente nessas escolas constitui, por si só, um custo educativo a ser considerado.

Na mesma época, na Inglaterra, foi publicado o relatório *Who wins and who loses from school accountability? The distribution of educational gain in English secondary schools*, de S. Burgess, C. Propper, H. Slater e D. Wilson. O modelo de financiamento das escolas públicas, baseado numa fórmula que relaciona o orçamento anual com o número de alunos e custo *per capita*, conciliado com a possibilidade da escolha da escola, fomentou a concorrência entre as escolas pela captação de alunos. Os indicadores do desempenho das escolas, taxas de aprovação nos exames nacionais, publicados anualmente, fundamentavam, tanto quanto possível, a decisão dos pais quanto à escolha da escola para os seus educandos. Os autores mostram que, na prática, as escolas tendiam a desenvolver estratégias para otimizar globalmente essas taxas de desempenho. Ou seja, investiam mais recursos educativos em alunos que se encontrassem no limiar de aprovação, face aos recursos investidos em alunos com trajectória de sucesso (desempenho elevado) ou insucesso (desempenho baixo).

As evidências reportadas sugerem que a percentagem dos alunos aprovados que se encontravam no limiar de aprovação aumentou, mas a análise ao nível dos alunos mostrou padrões de desempenho individual diferenciado consoante se tratasse, ou não, de escolas que competiam entre si. Num sistema de quase-mercado como este, os grandes perdedores são os alunos com défice educativo, geralmente provenientes de famílias socialmente desfavorecidas. Enquanto os alunos com desempenho elevado podem facilmente mudar de escola, caso sintam que o seu potencial não é devidamente aproveitado e desenvolvido na escola que frequentam, o mesmo não acontece com os alunos com défice educativo. Com efeito, a modelação da segregação social nas escolas públicas do ensino secundário mostra que a segregação aumentou no período compreendido entre 1994 e 2004, e que esse aumento foi mais acentuado nas escolas selectivas (Allen & Vignoles, 2007; Goldstein & Noden, 2003).

Estudos como os já descritos apontam consequências inesperadas de mecanismos de incentivo e servem de orientação para evitar cometer os mesmos erros, para prevenir o fracasso ou efeitos perversos, caso mecanismos similares venham a ser implementados em países que dão agora os primeiros passos.

Ora, o indicador de valor acrescentado ajuda a rastrear as escolas que pertencem sistematicamente a grupos extremos de valor acrescentado, ou porque são sistematicamente deficitárias ou sistematicamente excedentárias. Assim, servindo como critério objectivo para estabelecer prioridades de intervenção nas escolas com valor acrescentado deficitário, a sua principal finalidade seria a de identificar atempadamente unidades educativas que carecem de acompanhamento externo no processo de mudança/melhoria; ou seja, no estudo em pormenor das práticas correntes em todos os domínios da prestação do serviço educativo, visando o desenvolvimento de um programa específico para a superação das disfunções detectadas. Do mesmo modo, o critério valor acrescentado permite identificar escolas de referência, das quais podem ser extraídas lições sobre boas práticas generalizáveis e incorporadas através dos programas de melhoria. Nestes termos, o uso do valor acrescentado para determinar prioridades de actuação e a elaboração, implementação e acompanhamento de programas de melhoria, encerra também as dimensões política, de gestão e administração, profissional e pessoal da responsabilização – prestação de contas intrínseca a qualquer sistema de avaliação educacional. Acresce notar que os benefícios resultantes de qualquer sistema de avaliação dependem da aceitação e valorização por todos os intervenientes. A título de exemplo, registo a declaração da directora da Escola Secundária Campos Melo sobre a participação da instituição no projecto de eficácia escolar, em que um dos objectivos era a estimação do valor acrescentado:

Chegados a esta etapa, consideramos de todo o interesse que este projecto seja alargado a um maior universo de escolas e de disciplinas (...), ou mesmo generalizado, uma vez

que poderá constituir-se como uma poderosa ferramenta para auxiliar as escolas na sua auto-avaliação e na promoção da melhoria sustentada dos seus resultados escolares. (Isabel Fael, discussão em Ferrão, 2009, p. 60)

## Nota final

Os grandes desígnios enunciados no artigo 26 da Declaração Universal dos Direitos do Homem e as medidas para os concretizar passam pela concepção e implementação de um sistema de avaliação educacional aceite e valorizado por todos os intervenientes. As vantagens da abordagem de valor acrescentado no sistema de avaliação educacional são amplamente consensuais. Ao longo do artigo apresentaram-se reflexões sobre a sua definição e características, especificações e requisitos metodológicos, uso e finalidades, bem como as limitações à data. Foram dados alguns exemplos ilustrativos das consequências adversas para a qualidade da educação que determinados mecanismos e instrumentos podem ter e como podem ser evitadas. Destas reflexões emerge o enorme potencial inerente ao conhecimento científico, tecnológico e metodológico que tem sido desperdiçado por não estar a ser transferido à sociedade, usado e aplicado.

## Nota

1. Artigo 26: “1. Toda a pessoa tem direito à educação. A educação deve ser gratuita, pelo menos a correspondente ao ensino elementar fundamental. O ensino elementar é obrigatório. O ensino técnico e profissional deve ser generalizado; o acesso aos estudos superiores deve estar aberto a todos em plena igualdade, em função do seu mérito. 2. A educação deve visar à plena expansão da personalidade humana e ao reforço dos direitos do homem e das liberdades fundamentais e deve favorecer a compreensão, a tolerância e a amizade entre todas as nações e todos os grupos raciais ou religiosos, bem como o desenvolvimento das actividades das Nações Unidas para a manutenção da paz. 3. Aos pais pertence a prioridade do direito de escolher o género de educação a dar aos filhos”.

## Referências

ALKIN, M.C. Evaluation utilization: introduction. In: KELLAGHAN, D.L.; STUFFLEBEAM, D.L. (Ed.). *International handbook of educational evaluation*. Dordrecht: Kluwer, 2003. p. 189-196.

ALLEN, R.; VIGNOLES, A. What should an index of school segregation measure? *Oxford Review of Education*, Dorchester on Thames, v. 33, n. 5, p. 643-668, 2007. Disponível em: <<http://dx.doi.org/10.1080/03054980701366306>>

BIESTA, G. Good education in an age of measurement: on the need to reconnect with

the question of purpose in education. *Educational Assessment, Evaluation and Accountability*, v. 21, n. 1, p. 33-46, 2009.

BIRD, S. M. et al. Performance indicators: good, bad, and ugly. *Journal of the Royal Statistical Society: Series A*, London, v. 168, n. 1, p. 1-27, 2005.

BRAUN, H.; CHUDOWSKY, N.; KOENIG, J. *Getting value out of value-added: Social Sciences*. Washington, DC: National Academies, 2010. Disponível em: <<http://216.78.200.159/Documents/RandD/Other/Getting Value out of Value-Added.pdf>>

BRAUN, H.; WAINER, H. 27 Value-Added Modeling. *Handbook of statistics*, Amsterdam, v. 26, p. 867-892, 2006.

CLOTFELTER, C.T. et al. Do school accountability systems make it more difficult for low-performing schools to attract and retain high-quality teachers? *Journal of Policy Analysis and Management*, New York, v. 23, n. 2, p. 251-271, 2004.

COMISSÃO DAS COMUNIDADES EUROPEIAS. *European governance: a white paper*. Brussels, 2001.

CRONBACH, L.; SUPPES, P. *Research for tomorrow's schools: disciplined inquiry for evaluation*. New York: MacMillan, 1969.

FERRÃO, M.E. Ainda os rankings e a estatística. *Gazeta de Matemática*, Lisboa, n. 155, p. 36-47, 2008.

FERRÃO, M.E. *Modelo de valor acrescentado: algumas evidências a partir do projecto 3EM, com discussão*. Covilhã: Universidade da Beira Interior, 2009.

FERRÃO, M.E.; GOLDSTEIN, H. Adjusting for measurement error in the value added model: evidence from Portugal. *Quality & Quantity*, Amsterdam, v. 43, n. 6, p. 951-963, 2009.

GOLDSTEIN, H.; NODEN, P. Modelling social segregation. *Oxford Review of Education*, Doschester on Thames, v. 29, n. 2, p. 225-237, 2003.

GOLDSTEIN, H.; KOUNALI, D.; ROBINSON, A. Modelling measurement errors and category misclassifications in multilevel models. *Statistical Modelling*, London, v. 8, n. 3, p. 243-261, 2008.

GOLDSTEIN, H.; SPIEGELHALTER, D. League tables and their limitations: statistical issues in comparisons of institutional performance. *Journal of the Royal Statistical Society: Series A*, London, v. 159, p. 385-443, 1996. Disponível em: <<http://www.jstor.org/pss/2983325>>

HOFSTETTER, C.H.; ALKIN, M.C. Evaluation use revisited. In: KELLAGHAN,

T.; STUFFLEBEAM, D.L. (Ed.). *International handbook of educational evaluation*. Dordrecht: Kluwer, 2003. p. 197-222.

LADD, H.F.; WALSH, R.P. Implementing value-added measures of school effectiveness: getting the incentives right. *Economics of Education Review*, Cambridge, Mass., v. 21, n. 1, p. 1-17, 2002.

LINN, R.L. Assessments and accountability. *Educational Researcher*, Washington, DC, n. 29, p. 4-16, 2000.

NEVO, D. The conceptualization of educational evaluation: an analytical review of the literature. *Review of Educational Research*, v. 53, p. 117-128, 1983.

NEVO, D. Evaluation in education. In: SHAW, I.F.; GREENE, J.C.; MARK, M.M. (Ed.). *The Sage handbook of evaluation*. London: Sage Publications, 2006. p. 440-460.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD). *Measuring improvements in learning outcomes*. [S.l.]: OECD, 2008.

PATTON, M.Q. Six honest serving men for evaluation. *Studies in Educational Evaluation*, Emsford, n. 14, p. 301-330, 1988.

RAUDENBUSH, S.W.; WILLMS, J.D. The estimation of school effects. *Journal of Educational and Behavioral Statistics*, Washington, DC, v. 20, n. 4, p. 307, 1995.

SINCLAIR, A. The chameleon of accountability: forms and discourses. *Accounting, Organizations and Society*, Oxford, v. 20, n. 2, p. 219-237, 1995.

TYLER, R.W. *Basic principles of curriculum and instruction*. Chicago: University of Chicago, 1950.

UNESCO. Declaração Mundial sobre Educação para Todos: satisfação das necessidades básicas de aprendizagem. Paris: Unesco, 1990. Disponível em: <<http://unesdoc.unesco.org/images/0008/000862/086291por.pdf>>

UNESCO. *Education for all: the quality imperative*. Education. Paris: Unesco, 2005. Disponível em: <<http://www.unesco.org/new/en/education/themes/leading-the-international-agenda/efareport/reports/2005-quality/>>

WEISS, C.H. Evaluation research in the political context. In: STRUENING, E.; GUTTENTAG, M. (Ed.). *Handbook of evaluation research*. London: Sage, 1975. p. 13-26.

WEISS, C.H. *Using social research in public policy making*. Lexington, MA: Books, 1977.

WOODHOUSE, G.; YANG, M.; GOLDSTEIN, H. Adjusting for measurement error in multilevel analysis. *Journal of the Royal Statistical Society. Series A*, London, v. 159, n. 2, p. 201-212, 1996. Disponível em: <<http://www.jstor.org/stable/10.2307/2983168>>

---

*Recebido em 22 de maio de 2012.*

*Aprovado em 1º de junho de 2012.*