



Educação & Sociedade

ISSN: 0101-7330

ISSN: 1678-4626

Centro de Estudos Educação e Sociedade - Cedes

Ferrão, Maria Eugénia

**ESTATÍSTICA EDUCACIONAL E POLÍTICA PÚBLICA: A  
PROPÓSITO DOS MODELOS DE VALOR ACRESCENTADO\***

Educação & Sociedade, vol. 39, núm. 142, 2018, Janeiro-Março, pp. 19-38

Centro de Estudos Educação e Sociedade - Cedes

DOI: 10.1590/ES0101-73302017176230

Disponível em: <http://www.redalyc.org/articulo.oa?id=87358190014>

- Como citar este artigo
- Número completo
- Mais informações do artigo
- Site da revista em [redalyc.org](http://www.redalyc.org)

UABM [redalyc.org](http://www.redalyc.org)

Sistema de Informação Científica Redalyc

Rede de Revistas Científicas da América Latina e do Caribe, Espanha e Portugal

Sem fins lucrativos acadêmica projeto, desenvolvido no âmbito da iniciativa  
acesso aberto

# ESTATÍSTICA EDUCACIONAL E POLÍTICA PÚBLICA: A PROPÓSITO DOS MODELOS DE VALOR ACRESCENTADO\*

MARIA EUGÉNIA FERRÃO<sup>1</sup>

*RESUMO:* O artigo apresenta uma revisão sobre modelos de valor acrescentado e ligação à política pública em educação e ensino superior, omitindo os aspectos técnico-metodológicos que são por natureza complexos. Assim, o artigo destina-se, prioritariamente, a pesquisadores e atores de política educativa. Também procede à revisão sistemática da literatura científica publicada em periódicos do Brasil e de Portugal registados na plataforma SciELO, donde resulta a constatação de uma escassa produção face à agenda científica imposta pelas potencialidades inerentes à melhoria da educação e do ensino superior e, consequentemente, à prosperidade da sociedade. Por fim, são enunciadas três etapas principais para o desenvolvimento de um sistema de avaliação que inclua o indicador de valor acrescentado.

*Palavras-chave:* Valor acrescentado. Avaliação. Responsabilização. Indicadores de desempenho.

## EDUCATIONAL STATISTICS AND PUBLIC POLICY: ON VALUE ADDED MODELS

*ABSTRACT:* The article presents a review on value added models and connection to public policy in education and higher education, omitting the technical and methodological aspects that are by nature complex. Thus, the article is primarily aimed at researchers and actors of educational policy. It also proceeds to the systematic review of the scientific literature published in Brazilian and Portuguese journals registered in the SciELO platform, which results in the scarce production in view of the scientific agenda imposed by the potentialities inherent to the improvement of education and higher education and, consequently, to the prosperity of society. Finally, three main steps are set out for the development of an evaluation system that includes the value added indicator.

*Keywords:* Value added. Evaluation. Accountability. Performance indicators.

---

\*Artigo desenvolvido no âmbito do projeto de pesquisa *Statistical modelling in education: multilevel models applied to large and complex data*.

<sup>1</sup>Universidade da Beira Interior – Covilhã, Portugal. Centro de Matemática Aplicada à Previsão e Decisão Económica (CEMAPRE) – Lisboa, Portugal. E-mail: meferrao@ubi.pt

DOI: 10.1590/ES0101-73302017176230

## STATISTIQUES DE L'ÉDUCATION ET LES POLITIQUES PUBLIQUES: LE BUT DES MODÈLES À VALEUR AJOUTÉE

**RÉSUMÉ:** L'article présente un aperçu des modèles à valeur ajoutée et la connexion à la politique publique dans l'éducation et l'enseignement supérieur, en omettant les aspects techniques et méthodologiques qui sont complexes par nature. Ainsi, l'article est destiné principalement aux chercheurs et acteurs de la politique éducative. Réalise également l'examen systématique de la littérature scientifique publiée dans des revues du Brésil et du Portugal enregistrés dans la plate-forme SciELO, d'où la faible production en raison de l'agenda scientifique imposé par le potentiel inhérent à l'amélioration de l'éducation et de l'enseignement supérieur et, par conséquent, à la prospérité de la société. Enfin, ils ont mis trois étapes principales dans le développement d'un système d'évaluation qui comprend l'indicateur de la valeur ajoutée.

**Mots-clés:** Valeur ajoutée. Évaluation. Responsabilité. Indicateurs de performance.

## Introdução

A vertente transdisciplinar da estatística tem se desenvolvido ao longo dos últimos anos com novos temas e desafios. As últimas duas décadas foram portadoras de metodologia que nos permite desvendar a complexidade do mundo real (GOLDSTEIN, 1998), em especial na subárea estatística educacional, que é fundamental ao desenvolvimento disciplinar e científico da educação. Os avanços tecnológicos e computacionais, notadamente no que se refere aos métodos computacionais e estatísticos para lidar com *large and complex surveys* e dados longitudinais, potencializam um salto gigantesco na produção de conhecimento nesse campo. Ora, a declaração da *American Statistical Association* (ASA) sobre modelos de valor acrescentado (VA) suscita uma nova agenda de investigação científica em estatística educacional, mas também em política educativa (MORGANSTEIN & WASSERSTEIN, 2014). Nesse caso, uma das principais exigências é produzir indicadores úteis para os atores políticos nos vários níveis de tomada de decisão e, simultaneamente, demonstrar que a qualidade de tais indicadores é objetivamente superior e mais abrangente do que a daqueles resultantes de outras abordagens. De acordo com a própria declaração (MORGANSTEIN & WASSERSTEIN, 2014), ela é parte integrante da missão da ASA, enunciada da seguinte forma: “*Part of the ASA’s mission is promoting sound statistical practice to improve public policy and improve human welfare*” [Parte da missão da ASA é promover robustez na aplicação e prática da estatística para melhorar a política pública e melhorar o bem-estar humano]. Tal como se mostrará adiante, a definição de “valor acrescentado” ou “valor agregado” é muito intuitiva.

Contudo, a complexidade envolvida na especificação dos modelos, no procedimento de estimação e em outros aspectos técnico-metodológicos afasta a atenção dos mentores e atores de política educativa do benefício que tais modelos e respectivos indicadores representam para programas de melhoria da educação e do ensino superior. As barreiras entre a pesquisa de base quantitativa e o processo político podem ser mitigadas se os pesquisadores, avaliadores, formuladores e/ou analistas de política educativa tirarem o máximo proveito da interdisciplinaridade inerente ao objeto de estudo. A declaração também refere que a comunidade estatística norte-americana está envolvida na compreensão e melhoria dos modelos estatísticos que podem ser usados em educação e que a comunidade tem um papel fundamental na colaboração com as organizações educativas que desejam usar dados para melhorar a educação. De fato, o debate em torno dos modelos de VA emerge da onda crescente na implementação de sistemas de avaliação educacional que têm como principal propósito responsabilização — ou prestação de contas (*accountability*) — com consequências de alto risco para indivíduos e instituições. Cabe aqui uma nota para referir que o termo “alto risco” constitui a tradução livre de “*high stakes*” (e.g. FERRÃO & KLEIN, 2013), geralmente usado em avaliações que têm muita influência na vida de estudantes, de professores e/ou nos planos de atividades e de ação das instituições. Alguns autores brasileiros têm preferido a tradução “alto impacto”. Por exemplo, a responsabilização foi um dos princípios enunciados para a boa governança das organizações da União Europeia, a par de outros quatro: abertura, participação, eficácia e coerência (COMMISSION OF THE EUROPEAN COMMUNITIES, 2001) — princípios vertidos na legislação dos Estados-membros.

Por todo o mundo, tem-se verificado a implementação, em ritmo acelerado, de sistemas de avaliação que visam à comparação do desempenho profissional e institucional, apesar de ainda nos depararmos com desafios metodológicos e técnicos consideráveis para que os resultados da avaliação contribuam efetivamente para a melhoria da educação e para a prosperidade da sociedade (FERRÃO, 2016). Tal como Bird *et al.* (2005) acautelam, a comparação do desempenho pode ser bastante produtiva para os envolvidos quando é bem feita. Quando é mal feita, pode ser muito custosa e não apenas inútil, mas prejudicial e até mesmo destrutiva. De acordo com Foley & Goldstein (2012), nos últimos anos o ensino superior também teve um rápido crescimento de rankings institucionais, cuja elaboração foi formulada em termos de um “mercado” crescente, haja em vista o aumento drástico no número de universidades e na introdução e/ou elevação das taxas de frequência. Os autores procedem à revisão que estabelece as características no monitoramento do desempenho e dos rankings no ensino superior, bem como dão recomendações para melhorar o seu uso e abordar os aspectos mais problemáticos desse sistema de avaliação.

Contribuições recentes têm mostrado que é possível obter estimativas da eficácia educacional (eficácia escolar e eficácia docente) na promoção das apren-

dizagens. Apesar das limitações de tais abordagens, alguns mentores de política educativa têm sido excessivamente entusiastas na incorporação dessas estimativas como indicadores de desempenho nos sistemas de avaliação educacional dos seus países. As propriedades associadas a esses indicadores são cruciais para a legitimidade desses sistemas, particularmente quando são usados com consequências de alto risco. A maioria dos pesquisadores em estatística está consciente quanto às implicações da assunção de determinados pressupostos, ou da sua violação, nos resultados obtidos; dependendo do uso dado àqueles indicadores, assim serão os efeitos nas pessoas e instituições envolvidas.

Este artigo tem como objetivo geral contribuir para a divulgação do conhecimento sobre modelos de VA e respectivos indicadores junto de uma audiência não interessada na complexidade dos aspectos técnico-metodológicos. Principalmente uma audiência formada por pesquisadores, mentores e decisores de política educativa. O artigo está estruturado em quatro seções, além desta introdução. Na seção um, abordo alguns aspectos sobre a ligação entre a estatística e a política pública em educação e ensino superior na perspectiva da concepção e implementação de políticas orientadas pelo conhecimento de base científica. Na seção dois apresento, ao nível introdutório, uma revisão sobre VA em educação, considerando: a análise de diferentes definições registradas na literatura especializada; características com implicação de ordem metodológica; e finalidade. Na seção três analiso o tema a partir da literatura científica publicada no Brasil ou em Portugal; e, na seção quatro, apresento algumas considerações finais.

## Estatística e política pública em educação e ensino superior

O conhecimento de que dispomos sobre VA em educação é o resultado de, pelo menos, 45 anos de investigação científica. O caminho percorrido mostra avanços e recuos. Especialmente nas Ciências Sociais, o conhecimento científico assenta em cânones que o caracterizam como sistemático, replicável, generalizável, entre outros. Obedecer a esses cânones requer tempo. É a dinâmica da ciência. Uma dinâmica lenta face à premência de soluções com a cientificidade desejável. Entre os marcos do conhecimento científico que traçaram o sentido irreversível do progresso científico-disciplinar da educação, encontra-se aquele que deu lugar ao livro *School matters* (MORTIMORE et al., 1988) e que está intrinsecamente associado ao desenvolvimento metodológico de VA, em especial ao desenvolvimento dos modelos de regressão multinível (FERRÃO, 2011a; MORTIMORE et al., 1994; SAUNDERS, 1999).

Os modelos de VA são caracterizados por elevada complexidade, podendo comprometer a sua interpretabilidade tão desejável à tomada de decisão política. Para ser confiável e amplamente aceito, o modelo deve ser parcimonioso, transparente e interpretável. O caminho a percorrer faz-se de redes multidisciplinares, estabelecendo *pipelines* de comunicação e ação, que permitam transformar a

sociedade a partir de conhecimento científico. Ou seja, a ligação entre a estatística educacional e a política pública em educação e no ensino superior visa a essa transformação no sentido da melhoria. De acordo com Ferrão & Klein (2013, p. 7-8),

[...] a ideia de melhoria educacional e, de modo mais amplo, melhoria social está implícita a qualquer processo de avaliação. Pois tem-se verificado em ritmo acelerado sistemas de avaliação que visam à comparação de desempenho profissional e institucional, apesar de ainda nos depararmos com desafios metodológicos e técnicos consideráveis para que os resultados efectivos da avaliação contribuam efectivamente para a melhoria da educação e para a prosperidade da sociedade.

O relatório *Getting Value Out of Value-Added* publicado pela *National Academy of Sciences* (BRAUN *et al.*, 2010) reporta o parecer de alguns especialistas sobre as vantagens e o uso dos modelos de VA, tais como:

Adam Gamoran and Robert Gordon, among others, focused on what they saw as the advantages of these models over indicators based on student status. Gordon observed that although many important technical issues still need to be resolved, it is not realistic to think that policy makers will wait 20 years until all of the difficulties are worked out before making use of such methods. Decisions about schools and teachers are being made, and, as Jane Hannaway noted, there is enormous demand from the policy side, to which the testing and research communities need to respond as quickly as possible (BRAUN *et al.*, 2010, p. 55-56)<sup>1</sup>.

Apesar da vantagem dos modelos de VA face a indicadores de *status* ou de modelos de resultados contextualizados, o seu uso com o propósito de responsabilização/prestação de contas, com consequências de alto risco para indivíduos e instituições, ainda deve ser encarado com cautela. Se 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 desses modelos, urge que a academia e pesquisadores respondam a tais desafios com a rapidez necessária, sendo, desse modo, crucial reforçar a comunicação entre os dois grupos profissionais — isto é, entre a academia/pesquisadores e os mentores/decisores de política educativa.

## Valor Acrescentado em Educação

Esta seção apresenta diversas definições de modelo de VA em educação/ensino superior, aspectos comuns que representam potencialidades, mas também limitações, e debate as principais finalidades do seu uso por meio de exemplos da literatura científica.

## Definições

Na literatura científica publicada em língua portuguesa, encontram-se as designações sinônimas “valor acrescentado” e “valor agregado”. Apesar de considerar relevante a análise da terminologia usada numa perspectiva filológica, ela ultrapassa o domínio deste artigo. Por uma questão de simplicidade, doravante usaremos o termo “valor acrescentado” (VA), por ter sido o primeiro a ser utilizado na literatura científica publicada em língua portuguesa.

Na atualidade, há consenso em relação aos requisitos metodológicos mínimos que devem ser satisfeitos para apurar o VA em educação. O relatório *Measuring improvements in learning outcomes, best practices to assess the value-added of schools* (OECD, 2008) define modelo de VA como pertencente à classe de modelos estatísticos que permitem estimar a contribuição das escolas para o progresso dos alunos. Isso assume particular relevância no que se refere aos objetivos educativos estabelecidos, sendo o progresso medido, pelo menos, em duas ocasiões no tempo. Noutros termos, mas com significado semelhante, a *National Academy of Sciences* (EUA), no relatório *Getting Value Out of Value-Added* (BRAUN *et al.*, 2010), menciona que a abordagem se refere à quantificação do efeito da instituição, do professor ou do programa no desempenho dos estudantes, considerando as diferenças entre os alunos no que se refere ao conhecimento prévio e também a outras variáveis extraescolares que influenciam essa performance.

O volume 26 do *Handbook of Statistics*, no capítulo dedicado à modelação VA (BRAUN & WAINER, 2007), apresenta-os como sendo a família de modelos estatísticos que são usados para inferir sobre a eficácia das unidades educativas, usualmente instituições e/ou professores. Tais definições são convergentes num conjunto de características sobre as quais se refletirá nas próximas secções. Todavia, parece-nos importante registrar que alguns desses conceitos, tais como o de *efeito escola*, têm evoluído ao longo do tempo. Também no que se refere à operacionalização e quantificação de alguns deles, têm-se verificado desenvolvimentos, muitos deles decorrentes das possibilidades metodológicas ao dispor dos pesquisadores.

Artigo sobre os resultados da pesquisa de eficácia escolar em países de língua portuguesa, especificamente Brasil e Portugal (FERRÃO, 2014), apresenta uma breve síntese de definições em eficácia escolar e efeito escola. Entre outros autores, cita Teddlie & Reynolds (2000) para referir que nos últimos 50 anos surgiram pelo menos seis definições de “efeitos escolares”: efeito absoluto da escolarização (grupo experimental versus controle); média não ajustada da proficiência dos alunos por escola; impacto da escolarização na média da proficiência dos alunos por escola, ajustada por variáveis de antecedentes familiares e/ou conhecimento prévio; quantificação da proporção da variação “entre escolas” dos resultados escolares face à variação total; medida do efeito exclusivo de cada escola sobre os resultados dos alunos; e mensuração do impacto das escolas sobre o desempenho dos

alunos ao longo do tempo. Outros detalhes podem ser encontrados em Teddlie & Reynolds (2000, p. 65-69). As repercussões da escolha de uma ou outra definição e/ou operacionalização no âmbito de um sistema de avaliação educacional são, em parte, debatidas em Ferrão (2012), do seguinte modo:

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 (reflecte 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 (FERRÃO, 2012, p. 460).

No que se refere ao ensino superior, ao longo das últimas três décadas verificou-se o consenso internacional sobre a necessidade de medição da qualidade académica, o que se reflectiu no aumento da produção de rankings institucionais e de cursos (DILL & SOO, 2005). De acordo com Foley & Goldstein (2012), os rankings de desempenho destinam-se a dois propósitos principais: o primeiro pode ser descrito como responsabilização/prestação de contas; e o segundo é fornecer informação aos usuários dos serviços públicos de educação e ensino superior. De acordo com Hazelkorn (2011), após a crise financeira de 2008, os rankings das universidades passaram a simbolizar o impulso para maior responsabilização e transparência, fornecendo a evidência subjacente à agenda de modernização e à adoção de novas práticas de gestão pública.



Foley & Goldstein (2012) recomendam que qualquer ranking deve ser contextualizado, por meio da abordagem de VA. Adicionalmente, Goldstein (2014) lembra que, apesar do uso de modelos de VA estar razoavelmente bem estabelecido na educação básica, tal não acontece no ensino superior porque é difícil obter variáveis de controle mensuradas na unidade estatística aluno e ter os estudantes associados às respectivas instituições.

No que concerne à aplicação de modelos de VA no ensino superior brasileiro, alguns pesquisadores (e.g. Soares *et al.*, 2001, p. 365) defendem-na como sendo a abordagem metodológica mais adequada, uma vez que “a efetividade de uma instituição de ensino deve ser medida pelo que acrescenta ao conhecimento do aluno durante sua permanência nela, e não pela nota final obtida por ele em determinado exame”.

### **Características, potencialidades e limitações**

As definições de VA anteriormente apresentadas têm três características em comum. A primeira é a da existência de uma família ou classe de modelos. Assim, a expressão “o modelo de valor acrescentado” denota uma escolha entre diversos modelos estatísticos possíveis. Na recente revisão sistemática elaborada por Everson (2016) sobre modelos de VA para efeito de responsabilização/prestação de contas, a autora refere o seguinte: “A literatura selecionada para esta revisão mostra a diversidade nas escolhas de modelagem que são feitas, refletindo a complexidade e grande variedade de possibilidades na criação de modelos VA” (EVERSON, 2016, p. 48). Assim, tal como defendi em 2012, a escolha do modelo estatístico deve ser realizada em concordância com o critério de comparação estabelecido pelo *designer* do sistema de avaliação e de modo adequado ao contexto educativo, social, econômico e cultural em que o modelo vai ser aplicado (FERRÃO, 2012).

Naturalmente, verificam-se diferenças entre os indicadores VA resultantes da aplicação de cada um dos modelos possíveis e, à luz dessas diferenças, deverá ser realizada a respectiva interpretação dos resultados obtidos (FERRÃO, 2010). O princípio da transparência recomenda que qualquer que seja a abordagem de VA usada, os métodos estatísticos e o *software* devem ser completamente descritos, testados e documentados (MORGANSTEIN & WASSERSTEIN, 2014). Embora os estudos de consistência e da correlação forte entre as estimativas obtidas por modelos com diferentes especificações sejam encorajadores quanto ao seu uso como critério de avaliação educacional (FERRÃO, 2010; FERRÃO & GOLDSTEIN, 2009), é necessário ter em mente que a classificação da escola (ou do professor) por meio da aplicação do modelo pode depender da escolha do modelo. De acordo com Everson (2016), a inclusão ou a exclusão de covariáveis no modelo deve depender antes de tudo da resposta a algumas questões de natureza

filosófica, pois na comparação inerente ao processo de avaliação, nem sempre se sabe com quem um professor ou os alunos de uma escola devem ser comparados. Sem responder a tais questões, as decisões podem continuar a ser tomadas com alguma arbitrariedade e apenas baseadas nos dados disponíveis ao invés de levar em consideração os propósitos ou finalidades do sistema de avaliação. Na mesma linha de pensamento, Ferrão (2012, p. 460) considera que:

[...] 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.

E, portanto, “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” (FERRÃO, 2012, p. 461).

A segunda característica está relacionada com os cuidados a ter para tornar comparável a ação educativa e formativa das instituições envolvidas, já que cada uma delas serve a uma 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 é muito heterogênea, e que essa heterogeneidade também se verifica entre as escolas. Algo similar ocorre no ensino superior, em que a necessidade de contextualização é referida pela maioria dos pesquisadores (e.g. FOLEY & GOLDSSTEIN, 2012; GOLDSTEIN, 2014; SOARES *et al.*, 2001). Logo, é preciso que o modelo contemple variáveis extraescolares, ou de contexto, que têm influência no desempenho dos alunos e, assim, garanta que a estimativa da “contribuição da escola” ou “contribuição da universidade” esteja livre dessa influência.

A terceira característica envolve a noção de “progresso”, que está associada a um atributo que varia no tempo; para quantificar essa mudança e, por conseguinte, o progresso, são necessários dados longitudinais — de tal maneira que seja possível captar as alterações decorrentes da aprendizagem realizada de permeio. Por um lado, a infraestrutura técnica para a realização de estudos longitudinais é exigente. Por outro, a construção da métrica que permite quantificar o progresso do aluno não é trivial. Nesse caso, não se trata de uma

dificuldade específica ou limitação dos modelos de VA, mas antes do processo de aferição das aprendizagens. Geralmente, usa-se a classificação dos alunos em testes padronizados como medida dos resultados escolares. A tal respeito, alguns pesquisadores norte-americanos (LOCKWOOD *et al.*, 2007) mostraram que a escolha dos instrumentos de aferição (testes) causaram maior instabilidade nas estimativas de VA do que a escolha do modelo estatístico, sugerindo que o problema detectado poderia ser ainda maior se tivessem sido usados instrumentos desenvolvidos por diferentes produtores de testes. Problema similar é também tratado por McCaffrey *et al.* (2009).

Alguma resistência à abordagem VA radica na crítica à (suposta) unidimensionalidade do indicador. As metodologias estatísticas permitem considerar múltiplas dimensões dos resultados escolares, quer elas sejam aferidas através de variáveis contínuas ou discretas. Encontram-se estudos que consideram também o comportamento disciplinar dos alunos, a sua assiduidade, a autoestima, etc. Aque-la crítica está, portanto, desfocada. Levanta-se, então, a questão: que resultados escolares devem ser considerados nos modelos de VA? O propósito da educação é assegurar que crianças e jovens se tornem cidadãos aptos nas dimensões pessoal, social, política e econômica. Nessas dimensões há aprendizagens que são da estrita esfera da escola enquanto outras não o são. A valia da abordagem VA e o seu uso ao serviço da melhoria da qualidade da educação depende, assim, da definição do papel da escola e dos objetivos educativos que são da sua inteira responsabilidade e quais são os da responsabilidade dos demais intervenientes. Depende, também, da garantia de níveis aceitáveis de validade, fiabilidade e comparabilidade das medidas de resultados escolares. Só então valerá a pena elaborar sobre a finalidade ou o uso do VA em educação (FERRÃO, 2011b).

### **Finalidade: Responsabilização/prestação de contas, escolha da escola ou melhoria da educação e ensino superior**

No sentido do avanço gradual da autonomia das escolas, espera-se cada vez menos regulação nos *inputs* e nos processos e cada vez mais monitorização e avaliação do sistema educativo e do desempenho das escolas. Isso quer dizer que aspectos organizacionais, de liderança e de natureza pedagógica, por exemplo, devem ser cada vez mais da responsabilidade da escola (ao nível da qual a autoavaliação ocupa um papel determinante). Sabendo que o indicador VA representa a contribuição da escola na aprendizagem realizada durante um certo período de tempo pelos seus alunos, defendemos que ele é indispensável a um sistema de monitorização e avaliação que tenha como desígnio a melhoria da qualidade da educação, incluindo o ensino superior. Aliás, algo consensual em termos internacionais, pois diversos países têm vindo a contemplá-lo nos respectivos sistemas, nem sempre coincidindo quanto à finalidade da sua utilização. De um modo geral, as principais são as seguintes:

1. Incorporar a lógica da responsabilização e da prestação de contas nas instituições de educação/formação;
2. Fornecer informação a pais e encarregados de educação para a escolha da instituição de educação/formação;
3. Fomentar a implementação de programas de melhoria.

Estas três 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 em que se insere o programa de desenvolvimento. A prestação de contas deve ser globalmente entendida nas suas dimensões política, pública, de gestão e administração, profissional e pessoal. A responsabilização — e, por conseguinte, a respetiva prestação de contas — atribuída ao governo, departamentos centrais e regionais/federais, escolas/instituições de ensino superior, encarregados de educação e alunos, deve ser tal que, articulada e cumulativamente, sejam criados os meios, o contexto e o apoio para que as escolas possam maximizar as aprendizagens efetuadas por cada um dos seus alunos. Assim, o indicador de VA ajuda a rastrear as escolas que pertencem sistematicamente a grupos extremos de desempenho (FERRÃO & COUTO, 2014). Servindo como critério objetivo para estabelecer prioridades de intervenção nas escolas com VA deficitário, a sua principal finalidade pode ser a de identificar, a tempo, unidades educativas que carecem de acompanhamento externo no processo de mudança ou de melhoria. Ou seja, no pormenor das práticas educativas em todos os domínios da prestação do serviço educativo visando ao desenvolvimento de um programa específico para a superação das disfunções detectadas.

Do mesmo modo, o critério ajuda a identificar escolas de referência, das quais podem ser extraídas lições sobre boas práticas generalizáveis, por meio de programas de melhoria. Nesses termos, o seu uso para determinar as prioridades de atuação e a elaboração, a implementação e o acompanhamento de programas de melhoria encerra simultaneamente dimensões de política, de gestão e de administração — profissional e pessoal — da prestação de contas. Em alguns países, tais dimensões estão apostas ao caráter profissional, por meio da atribuição de prêmios de desempenho, de mérito, ou de outros incentivos individuais. Esses mecanismos de responsabilização ganharam muita popularidade entre governantes e mentores de política. Os argumentos a favor são principalmente de ordem racional, não havendo, até a data, demonstração da existência de uma relação causa-efeito entre a prática de incentivos do tipo *performance-related pay* e a melhoria das aprendizagens.

Nessa linha, mesmo supondo que os incentivos individuais se baseiam em múltiplas medidas do desempenho ou do mérito, 100% justas e 100% fiáveis, convém levar em consideração a natureza da profissão e o efeito de competitividade gerada entre professores pelo mecanismo de incentivos individuais.

Isso porque a competitividade, que é salutar em outros setores de atividade, pode ser um obstáculo ao trabalho colaborativo na educação, sendo esse considerado fator determinante do sucesso educativo. Note-se que o processo de criação de valor, produção e transmissão de conhecimento é acentuatadamente marcado por externalidades positivas resultantes da cooperação e difusão de conhecimento e práticas de aprendizagem coletiva e não da retenção desse conhecimento como estratégia de reforço do posicionamento individual dos profissionais perante avaliações futuras. Adicionalmente, a avaliação de desempenho de professores com consequências de alto risco deve ser feita com muito cuidado para evitar situações tais como a relatada por Anderson (2013). De acordo com o relato, um profissional foi despedido na sequência da avaliação de desempenho com consequência de alto risco, tendo-se posteriormente concluído que havia um erro no apuramento. A pontuação resultante da abordagem de VA representava 35% na avaliação de desempenho dos professores. Os argumentos de ordem racional radicam essencialmente na necessidade de valorização da profissão docente, considerando que a qualidade da prática é muito importante para as aprendizagens dos alunos. Logo, a valorização da profissão contribuirá para atrair e reter os melhores para a docência.

A esse propósito, 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 a ele, as escolas da Carolina do Norte — cuja população discente se caracterizava por maior proporção de alunos com défice educativo (ou menor desempenho) — passaram a ter mais dificuldade em atrair e reter “bons” professores. Mesmo admitindo que tal tivesse ocorrido ao acaso, verificou-se a diminuição da estabilidade do corpo docente nessas escolas, o que, por si só, já tem um custo educativo a ser considerado. A esse respeito, estudos sobre a estabilidade do corpo docente mostram que é mais provável um professor bem preparado (aferição pelo *National Teacher Exam* ou instrumento similar) abandonar a profissão do que um mal preparado.

Goldstein (2014) refere-se à utilidade dos rankings, quando os dados adequados à sua produção estão disponíveis. O autor alerta para a incerteza que lhes está associada e que pode ser representada através de intervalos de confiança. Acrescenta, ainda, que talvez o uso mais eficaz dos rankings institucionais seja como instrumento de rastreamento que possa sugerir onde os problemas estão ocorrendo, ao invés de diagnósticos de quais são os problemas.

## Valor acrescentado na literatura científica de língua portuguesa

A pesquisa dos artigos de educação publicados em língua portuguesa foi realizada na base de dados SciELO considerando, de acordo com o teor da plataforma, o método de pesquisa integrada, na coleção de periódicos do Brasil e de Portugal, com as palavras “valor acrescentado” ou “valor agregado” em todos os índices, com filtro

WOS e as áreas temáticas: educação e pesquisa educacional. A pesquisa foi efetuada no dia 13 de fevereiro de 2017, e teve como resultado oito artigos. Na primeira fase de análise foram excluídos dois deles por estarem fora do âmbito deste artigo. Um deles, publicado num periódico de ensino da física, trata de um modelo de simulação em processos de polimerização. O outro, publicado no periódico *Educação & Sociedade*, trata do VA, na perspectiva profissional, no mercado das escolas de estilismo. Entre os seis artigos restantes, o critério de inclusão estabelecido foi o artigo apresentar explicitamente a definição e operacionalização do conceito de VA. Os cinco artigos incluídos na análise encontram-se resumidos no Quadro 1, que simplificadamente enuncia a contribuição do artigo, o propósito ou finalidade, bem como o enquadramento ou referencial teórico em que o artigo se desenvolve, através das principais referências bibliográficas usadas para suportar metodologicamente o respectivo artigo.

### Quadro 1

Aspectos tratados nos artigos sobre valor acrescentado.

Artigo	Contribuição	Propósito ou finalidade	Enquadramento valor acrescentado
(FERRÃO, 2012)	Revisão e reflexão sobre modelo de valor acrescentado: definição, características, especificações e requisitos metodológicos, uso e finalidades, limitações	Ligação entre valor acrescentado e sistema de avaliação educacional	(BRAUN <i>et al.</i> , 2010; BRAUN & WAINER, 2007; GOLDSTEIN & SPIEGELHALTER, 1996; OECD, 2008)
(BRAUN, 2013)	Revisão e reflexão sobre modelo de valor acrescentado e política de prestação de contas-responsabilização com consequências de alto risco	Resumir os principais resultados de pesquisa nos Estados Unidos; sugerir modos de uso para a melhoria do ensino	(BRAUN, 2005; BRAUN <i>et al.</i> , 2010; REARDON & RAUDENBUSH, 2009)
(FERRÃO & COUTO, 2013)	Debate e comparação de modelo de resultados contextualizados ( <i>status</i> ) e de valor acrescentado	Mostrar evidência comparativa do indicador de <i>status</i> e de valor acrescentado, estudo das propriedades de consistência e estabilidade	(BRAUN & WAINER, 2007; OECD, 2008; SAMMONS <i>et al.</i> , 1997; SAUNDERS, 1999)
(BROOKE <i>et al.</i> , 2014)	Comparação de duas abordagens de modelação do valor acrescentado, considerando variáveis de alunos, turmas, escolas	Identificar fatores que influenciam a aprendizagem ano a ano e os padrões de crescimento ao longo dos anos	Não consta
(SOARES <i>et al.</i> , 2017)	Comparação de modelo de resultados contextualizados ( <i>status</i> ) com modelo de valor acrescentado	Verificar a possibilidade de quantificar a eficácia escolar através do modelo de <i>status</i> , comparando com o de valor acrescentado. Estudo de estabilidade	(BRAUN & WAINER, 2007; FERRÃO, 2014; FERRÃO & COUTO, 2013; REARDON & RAUDENBUSH, 2009; RUBIN <i>et al.</i> , 2004)

Qualquer dos cinco artigos traz ao debate as diferentes possibilidades de modelação na produção de indicadores de desempenho institucional e profissional. Dois deles (FERRÃO & COUTO, 2013; SOARES *et al.*, 2017) abordam o tema a partir do modelo de resultados contextualizados e do de VA aplicados aos dados do Estudo Longitudinal da Geração Escolar (GERES) 2005, norteando a utilização com vantagem no apuramento do VA da escola. Todavia, torna-se mais exigente uma vez que obriga a existência de dados longitudinais. O debate em torno do desempenho profissional com consequências de alto risco é tratado em detalhe por Braun (2013), invocando os achados de pesquisa científica norte-americana. Outra ideia partilhada em todos os trabalhos é a de que um bom sistema de avaliação deve conter múltiplos indicadores. Nas palavras de Braun (2013), “multiple indicators are not only necessary for fair evaluations, but also for the accountability system to contribute to the systematic and sustained improvement of teaching and learning”.

A necessidade de prestação de contas através da informação à sociedade também é lembrada por SOARES *et al.* (2017). Os autores referem o seguinte:

Acreditamos que a principal informação a ser prestada à sociedade deve ser os níveis de proficiências alcançados pelos alunos ao final das etapas escolares, explicados por meio de uma interpretação pedagógica e, ainda, contextualizados por meio de medidas como as da condição socioeconômica dos alunos e, também, por meio de medidas relevantes de processos escolares. Mas, para amortizar os conflitos que os resultados brutos trazem, e não ser injusto com o esforço da comunidade escolar, nos mais variados contextos sociais e políticos, uma medida direta da qualidade da escola pode ser uma informação útil, senão para a comunidade escolar se espelhar nos resultados, pelo menos para que os gestores possam formular suas políticas públicas adequadamente (SOARES *et al.*, 2017, p. 83).

Todos os artigos elencam os desafios e limitações que justificam uma agenda de pesquisa muito exigente. Por exemplo, Ferrão & Couto (2013) mencionam que “Os resultados apresentados estão ainda limitados pela ocorrência de erros não amostrais que ultrapassam o âmbito deste artigo mas que é fundamental investigar em trabalhos futuros. Um é o do erro da medida, e o outro é o do erro de representatividade ou viés de selecção devido ao atrito”. Contudo, os autores chamam a atenção para outros fenômenos revelados pela modelação — tais como “a importância da informação dada pelo modelo referente ao desempenho relativo de grupos minoritários” — e que, por serem recorrentes, reclamam a atenção dos mentores de política educativa. Ainda sobre as limitações da investigação reportada, Ferrão (2012) cita Cronbach & Suppes (1969) para lembrar boas práticas na condução de investigação científica, isto é:

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 (FERRÃO, 2012, p. 462, tradução livre da autora).

De fato, o reduzido número de artigos publicados, visíveis na plataforma SciELO, nas coleções de Brasil e Portugal, é a manifestação de uma extensa agenda de pesquisa e avaliação educacional que está por se concretizar. A comunidade científica tem a sua cota de responsabilidade na melhoria da educação básica e do ensino superior, especificamente na transferência do conhecimento científico (FERRÃO, 2009, p. 52). Todavia, importa invocar a ideia de avaliação educacional como uma atividade que acontece no contexto político (WEISS, 1975).

A produção e comunicação de conhecimento científico que supere as limitações e dificuldades enunciadas ao longo deste artigo está fortemente dependente da política para a ciência, em especial da que concerne ao incentivo/financiamento das subáreas em que os avanços determinantes são obstaculizados pela falta de massa crítica (FERRÃO, 2012). O reduzido número de artigos e a falta de massa crítica sobre tais matérias e afins são ainda mais visíveis na perspectiva comparativa internacional. Note-se, como exemplo, que o artigo de revisão sobre modelos de VA publicado em 2016 (EVERSON, 2016) no periódico *Review of Educational Research* — com *Scientific Journal Rankings* (SJR) no 1º percentil superior ao longo dos últimos anos na área de ciências sociais e subárea educação — contém 112 referências bibliográficas. Inclui apenas um dos pesquisadores que constam no Quadro 1 anteriormente apresentado. Importa referir que a capacidade de pesquisa científica instalada está condicionada à disponibilidade de dados longitudinais que satisfaçam os requisitos metodológicos enunciados ao longo do artigo.

## Considerações Finais

A declaração da ASA (MORGANSTEIN & WASSERSTEIN, 2014) sobre modelos de VA suscita uma nova agenda de pesquisa científica em estatística educacional, tal como foi preconizado por Ferrão & Klein (2013), mas também em política educativa. Ao longo deste artigo, apresentou-se uma revisão sobre modelos de VA e respectivos indicadores, destinando-se prioritariamente a pesquisadores e a mentores/decisores de política educativa; isso é, a uma audiência não interessada na complexidade dos aspectos técnico-metodológicos dos modelos. Também se procede à revisão sistemática na literatura científica publicada em periódicos do Brasil e de Portugal, registados na plataforma SciELO, donde resulta a diminuta produção face à agenda científica imposta pelas potencialidades inerentes ao tema.



Finalmente, em jeito de síntese e conclusão, vale a pena registrar, de acordo com a *Organisation for Economic Co-operation and Development* (OECD, 2008), as etapas necessárias para um país dispor de um sistema de monitorização e avaliação que inclua o indicador de VA. Delineiam-se em seguida:

- Decidir quanto ao propósito do indicador de VA e outros na perspectiva dos objetivos de política e de programas de melhoria da educação;
- Estabelecer o referencial passível de monitorização e avaliação, e selecionar o modelo estatístico mais adequado ao contexto social e educacional;
- Assegurar a infraestrutura necessária à implementação do modelo estatístico. Isso inclui um sistema de aferição dos resultados que obedeça técnica e metodologicamente aos padrões recomendados, que permita a recolha de dados longitudinais considerando o estudante como unidade estatística de observação ao longo da sua trajetória; inclui também a capacidade técnica e científica para a análise e modelação desses dados, depois de anonimizados, bem como para a disseminação de resultados que promova atividades *feedback* da avaliação, envolvendo as previstas no âmbito dos programas de melhoria já referidos.

Embora seja necessário desenvolver trabalho adicional de modelagem, uma necessidade mais urgente é a de encontrar formas para melhorar a qualidade dos dados e, especialmente, evitar que inferências despropositadas sejam extraídas de qualquer apresentação excessiva dos resultados (GOLDSTEIN, 2014).

## Agradecimentos

A autora agradece a Leandro S. Almeida pelas contribuições à versão inicial do artigo, aos dois pareceristas anônimos pelas sugestões de melhoria, e ainda a Nara Dias, Marília Boldorini e Bianca pela revisão editorial. Este trabalho foi parcialmente apoiado pelo projeto CEMAPRE — UID/MULTI/00491/2013 financiado pela Fundação para a Ciência e Tecnologia (FCT/MCTES), por meio de fundos nacionais, e também pelo Banco Santander, no âmbito da bolsa de jovens professores 2016/17 atribuída à autora.

## Notas

1. Tradução livre da autora: “Adam Gamoran e Robert Gordon, entre outros, centraram-se no que viam como vantagens desses modelos [de valor acrescentado] em relação aos indicadores baseados no status do estudante [ou nos modelos de resultados contextualizados]. Gordon observou que, apesar de muitas questões técnicas importantes ainda precisarem ser resolvidas, não é realista

pensar que os formuladores de políticas irão aguardar 20 anos até que todas as dificuldades estejam resolvidas antes de usar esses métodos. Decisões sobre escolas e professores estão sendo tomadas e, como observou Jane Hannaway, há uma demanda enorme do lado da política, à qual as comunidades de pesquisa precisam responder o mais rápido possível.”

## Referências

- ANDERSON, N. D.C. school officials: 44 teachers were given mistaken performance evaluations. *Washington Post*, 2013. Disponível em: [https://www.washingtonpost.com/local/education/dc-school-officials-44-teachers-were-given-mistaken-performance-evaluations/2013/12/23/c5cb9f26-6c0c-11e3-a523-fe73f0ff6b8d\\_story.html?utm\\_term=.f4c9e9b81ffc](https://www.washingtonpost.com/local/education/dc-school-officials-44-teachers-were-given-mistaken-performance-evaluations/2013/12/23/c5cb9f26-6c0c-11e3-a523-fe73f0ff6b8d_story.html?utm_term=.f4c9e9b81ffc). Acesso em: 14 março 2016.
- BIRD, S.M.; COX, D.; FAREWELL, V.T.; GOLDSTEIN, H.; HOLT, T.; SMITH, P.C. Performance indicators: good, bad, and ugly. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, v. 168, n. 1, p. 1-27, 2005. <http://doi.org/10.1111/j.1467-985X.2004.00333.x>
- BRAUN, H. Value-added modeling and the power of magical thinking. *ENSAIO: Avaliação e Políticas Públicas em Educação*, v. 21, n. 78, p. 115-130, 2013.
- \_\_\_\_\_. Value-added modeling: What does due diligence require. In: LISSITZ, R. (Ed.). *Value-added Models in Education: Theory and Applications*. Maple Grove: JAM Press Books, 2005. p. 19-38. Disponível em: [http://www.cgp.upenn.edu/pdf/Braun - VA Modeling What Does Due Diligence Req.pdf](http://www.cgp.upenn.edu/pdf/Braun_-_VA_Modeling_What_Does_Due_Diligence_Req.pdf). Acesso em: 14 março 2016.
- BRAUN, H.; CHUDOWSKY, N.; KOENIG, J. *Getting value out of value-added*. Social Sciences. Washington, D.C.: National Academies Press, 2010. Disponível em: [http://216.78.200.159/Documents/RandD/Other/Getting Value out of Value-Added.pdf](http://216.78.200.159/Documents/RandD/Other/Getting_Value_out_of_Value-Added.pdf). Acesso em: 14 março 2016.
- BRAUN, H.; WAINER, H. Value-added modeling. In: RAO, C.R.; SINHARAY, S. (Eds.). *Handbook of Statistics: Psychometrics*, v. 26, p. 867-892, 2007. Amsterdam: Elsevier, 2006. [http://doi.org/10.1016/S0169-7161\(06\)26027-9](http://doi.org/10.1016/S0169-7161(06)26027-9)
- BROOKE, N.; FERNANDES, N. da S.; MIRANDA, I.P.H.; SOARES, T.M. Modelagem do crescimento da aprendizagem nos anos iniciais com dados longitudinais da pesquisa GERES I. *Educação & Sociedade*, v. 40, n. 1, p. 77-94, 2014.
- CLOTFELTER, C.T.; LADD, H.F.; VIGDOR, J.L.; DIAZ, R.A. 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*, v. 23, n. 2, p. 251-271, 2004. <http://doi.org/10.1002/pam.20003>
- COMMISSION OF THE EUROPEAN COMMUNITIES. *European Governance: a white paper*. Brussels, 2001.
- CRONBACH, L.; SUPPES, P. *Research for tomorrow's schools: disciplined inquiry for evaluation*. New York: MacMillan, 1969.

DILL, D. D.; SOO, M. Academic quality, league tables, and public policy: A cross-national analysis of university ranking systems. *Higher Education*, 49(4), 495–533, 2005. Disponível em: <https://doi.org/10.1007/s10734-004-1746-8>.

EVERSON, K.C. Value-added modeling and educational accountability: are we answering the real questions? *Review of Educational Research*, p. 1-36, 2016. <http://doi.org/10.3102/0034654316637199>

FERRÃO, M.E. Avaliação educacional e modelos de valor acrescentado: tópicos de reflexão. *Educação & Sociedade*, v. 33, n. 119, p. 455-469, 2012.

\_\_\_\_\_. Estatística Educacional: a propósito dos Modelos de Valor Acrescentado. *Boletim Sociedade Portuguesa de Estatística*, v. 201, p. 23-25, 2016.

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

\_\_\_\_\_. School effectiveness research findings in the Portuguese speaking countries: Brazil and Portugal. *Educational Research for Policy and Practice*, v. 13, n. 1, p. 3-24, 2014. <http://doi.org/10.1007/s10671-013-9151-7>

\_\_\_\_\_. Valor Acrescentado em Educação. *Educação/Jornal de Letras*, p. 7-8, 2010.

\_\_\_\_\_. Valor Acrescentado em Educação: Finalidade e Uso. *Educação/Jornal de Letras*, p. 4-5, 2011a.

\_\_\_\_\_. Valor Acrescentado em Educação: Marcos da Investigação Científica. *Educação/Jornal de Letras*, p. 8-9, 2011b.

FERRÃO, M.E., COUTO, A. Indicador de valor acrescentado e tópicos sobre consistência e estabilidade: uma aplicação ao Brasil. *Ensaio: Avaliação e Políticas Públicas em Educação*, v. 21, n. 78, p. 131-164, 2013. <http://doi.org/10.1590/S0104-40362013000100008>

FERRÃO, M.E.; COUTO, A.P. The use of a school value-added model for educational improvement: a case study from the Portuguese primary education system. *School Effectiveness and School Improvement*, v. 25, n. 1, p. 174-190, 2014. <http://doi.org/10.1080/09243453.2013.785436>

FERRÃO, M.E.; GOLDSTEIN, H. Adjusting for measurement error in the value added model: evidence from Portugal. *Qual Quant*, v. 43, n. 6, p. 951-963, 2009. <http://doi.org/10.1007/s11135-008-9171-1>

FERRÃO, M.E.; KLEIN, R. Editorial. *ENSAIO: Avaliação e Políticas Públicas em Educação*, v. 78, n. 21, p. 7-9, 2013.

FOLEY, B.; GOLDSTEIN, H. *Measuring success: League tables in the public sector*. Londres: British Academy, 2012.

GOLDSTEIN, H. *Models for reality: New approaches to the understanding of educational processes*. Londres: Institute of Education Publications, 1998.

\_\_\_\_\_. Using league table rankings in public policy formation: Statistical issues. *Annual Review of Statistics and Its Application*, v. 1, n. 1, p. 385-399, 2014. <http://doi.org/10.1146/annurev-statistics-022513-115615>

GOLDSTEIN, H.; SPIEGELHALTER, D. League tables and their limitations: statistical issues in comparisons of institutional performance. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, v. 159, p. 385-443, 1996. Disponível em: <<http://www.jstor.org/pss/2983325>>. Acesso em: 14 março 2016.

HAZELKORN, E. *Rankings and the reshaping of higher education: The battle for world-class excellence*. London: Palgrave Macmillan, 2011.

LOCKWOOD, J.R.; MCCAFFREY, D.F.; HAMILTON, L.S.; STECHER, B.; LE, V.-N.; MARTINEZ, J.F. The sensitivity of value-added teacher effect estimates to different mathematics achievement measures. *Journal of Educational Measurement*, v. 44, n. 1, p. 47-67, 2007. <http://doi.org/10.1111/j.1745-3984.2007.00026.x>

MCCAFFREY, D.F.; SASS, T.R.; LOCKWOOD, J.R.; MIHALY, K. The intertemporal variability of teacher effect estimates. *Education Finance and Policy*, v. 4, n. 4, p. 572-606, 2009. <http://doi.org/10.1162/edfp.2009.4.4.572>

MORGANSTEIN, D.; WASSERSTEIN, R. ASA statement on value-added models. *Statistics and Public Policy*, v. 1, n. 1, p. 108-110, 2014. <http://doi.org/10.1080/2330443X.2014.956906>

MORTIMORE, P.; SAMMONS, P.; STOLL, L.; LEWIS, D.; ECOB, R. *School Matters*. Somerset: Open Books, 1988.

MORTIMORE, P.; SAMMONS, P.; THOMAS, S. School effectiveness and value added measures. *Assessment in Education: Principles, Policy & Practice*, v. 1, n. 3, p. 315-332, 1994. <http://doi.org/10.1080/0969594940010307>

NEVO, D. Evaluation in Education. In: SHAW, I.F.; GREENE, J.C.; MARK, M.M. (Eds.). *The SAGE Handbook of Evaluation*. London: SAGE Publications, 2006. p. 440-460.

\_\_\_\_\_. The conceptualization of educational evaluation: An analytical review of the literature. *Review of Educational Research*, v. 53, p. 117-128, 1983.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD). *Measuring improvements in learning outcomes*. Paris: OECD Publishing, 2008. <http://doi.org/10.1787/9789264050259-en>

REARDON, S.F.; RAUDENBUSH, S.W. Assumptions of value-added models for estimating school effects. *Education Finance and Policy*, v. 4, n. 4, p. 492-519, 2009. Disponível em: <<http://www.mitpressjournals.org/doi/abs/10.1162/edfp.2009.4.4.492>>. Acesso em: 14 março 2016.

RUBIN, D.B.; STUART, E.A.; ZANUTTO, E.L. A potential outcomes view of value-added assessment in education. *Journal of Educational and Behavioral Statistics*, v. 29, n. 1, p. 103-116, 2004. <http://doi.org/10.3102/10769986029001103>

SAMMONS, P.; THOMAS, S.; MORTIMORE, P. *Forging links: Effective schools and effective departments*. Londres: Paul Chapman, 1997.

SAUNDERS, L. A brief history of educational “value added”: How did we get to where we are? *School Effectiveness and School Improvement*, v. 10, n. 2, p. 233-256, 1999. <http://doi.org/10.1076/sesi.10.2.233.3507>

SOARES, J.F.; RIBEIRO, L.M.; CASTRO, C.D.M. Valor agregado de instituições de ensino superior em Minas Gerais para os cursos de Direito, Administração e Engenharia Civil. *Dados*, v. 44, n. 2, p. 363-396, 2001. Disponível em: <[http://www.scielo.br/scielo.php?pid=S0011-52582001000200005&script=sci\\_abstract](http://www.scielo.br/scielo.php?pid=S0011-52582001000200005&script=sci_abstract)>. Acesso em: 14 março 2016.

SOARES, T.M.; BONAMINO, A.; BROOKE, N.; FERNANDES, N. da S. Modelos de valor agregado para medir a eficácia das escolas Geres. *Ensaio: Avaliação e Políticas Públicas em Educação*, v. 25, n. 94, p. 59-89, 2017. <http://doi.org/10.1590/s0104-40362017000100003>

WEISS, C.H. Evaluation research in the political context. In: STRUENING, E.L.; GUTTENTAG, M. (Eds.). *Handbook of Evaluation Research*. Califórnia: SAGE Publications, 1975. p. 13-26.

---

*Recebido em 22 de fevereiro de 2017.*

*Aprovado em 01 de agosto de 2017.*

