

AS LIMITAÇÕES INTRÍNSECAS DOS TESTES PADRONIZADOS DE APRENDIZAGEM

Remo Moreira Brito Bastos
Doutorando em Educação - UFC
GRUPETPE - Grupo de Pesquisas em Trabalho, Práxis e Educação

RESUMO

Com o advento das reformas educacionais promovidas por governos suscetíveis à pressão das grandes corporações empresariais, um aumento desproporcional das horas destinadas a testes, o estreitamento de conteúdos curriculares e as avaliações verticalizantes, focando responsabilização de professores e diretores, passaram a conformar a agenda educacional dos países onde tal fenômeno se manifesta. O presente trabalho, desenvolvido por meio de pesquisa bibliográfica e documental, questiona as premissas que fundamentam essas políticas, fortemente estruturadas em torno dos testes padronizados de larga escala, os quais constituem a “ferramenta por excelência” da *accountability* subjacente às mencionadas reformas. Ademais das limitações intrínsecas de tais avaliações com vistas a aferir a efetiva aprendizagem dos educandos e a mensurar suas habilidades não cognitivas interpessoais e individuais, salienta-se os deletérios efeitos do estreitamento curricular àquele corpo discente impingidos, que abrangem desde o atrofiamento cognitivo que impede o exercício da criatividade e do senso crítico, fatores essenciais para a formação do trabalhador requerido pelo novo capitalismo, culminando com a insidiosa suposição de que algumas crianças simplesmente não podem aprender.

Palavras-chave: Testes padronizados. Política educacional. Estreitamento Curricular

ABSTRACT

With the advent of educational reforms promoted by governments susceptible to pressure from big business corporations, a disproportionate increase in hours devoted to testing, the narrowing of curricula and the verticalizing reviews, focusing on accountability of teachers and principals, began to shape the educational agenda of countries where this phenomenon manifests itself. This work, developed through bibliographical and documentary research, questions the assumptions underlying these policies, strongly structured around standardized large-scale tests, which constitute the "par excellence" tool of accountability underlying the mentioned reforms. In addition to the intrinsic limitations of such assessments in order to measure the effective learning of students and measure their non interpersonal and individual cognitive abilities, it is noted the deleterious effects of curriculum narrowing impinging upon that student body, which ranges from cognitive stunting which prevents the exercise of creativity and critical thinking, essential factors for worker formation required by the new capitalism, culminating with the insidious assumption that some children simply can not learn.

Keywords: Standardized tests. Educational policy. Curriculum narrowing.

Introdução

No bojo do *No Child Left Behind*, o programa educacional federal que vigorou de 2002 a 2015 nos Estados Unidos, as avaliações estandardizadas, ferramenta “por excelência” desse programa, com vistas à operacionalização de seus fundamentos e ao alcance de seus objetivos, constituem elemento nuclear na estruturação da *accountability* educacional que vai se consolidando na administração pública daquele país, marcada pela obsessão em aferir o desempenho acadêmico dos alunos e responsabilizar os professores pelos resultados daqueles por meio da utilização de testes padronizados de larga escala, de aplicação censitária. Adiantamos que nossas pesquisas têm sinalizado que a implantação desse modelo educacional imediatista, mercantil e adestrador fracassa em lograr (com efeito, em muitos casos agrava a situação) os declarados objetivos de recuperar o nível de desempenho educacional daquela nação e de eliminar a diferença nesse quesito (*achievement gap*) que desfavorece os negros e os hispânicos em relação aos brancos.

Nesse paradigma educacional, correntemente dominante no Ocidente, e notabilizado por seu caráter dual, iníquo, segregador, concentrador e reproduzidor das desigualdades constitutivamente imanentes ao modo de produção capitalista, os testes padronizados têm subtraído parcela substancial do tempo em sala de aula destinado à própria instrução, conforme relatório do prestigioso *FairTest - National Center for Fair & Open Testing* (FairTest - Centro Nacional para Testagem Aberta e Justa), nos Estados Unidos:

Os testes e a preparação para testes têm se proliferado - a quantidade de tempo gasto em testes em algumas escolas dobrou. Um estudo para o Congresso feito pelo Government Accountability Office (Escritório Governamental de Accountability) estimou que os estados teriam que criar mais de 433 testes (a um custo de entre US \$ 1,9 bilhões e 5,3 bilhões entre 2002 e 2008) para satisfazer as exigências da lei NCLB. Isso tornou-se apenas a ponta do iceberg de um aumento maciço no uso de testes. Não é incomum de 20 a 60 dias letivos por ano serem gastos em preparação, além dos dias gastos nos testes em si, que são consideráveis. [...] Um prejuízo considerável é o desvio do tempo de ensino destinado aos alunos que mais necessitam. A Wisconsin Association for Supervision and Curriculum Development (Associação Wisconsin para Supervisão e Desenvolvimento Curricular) tentou quantificar o tempo de

aprendizagem perdido com os testes em geral e para os alunos com necessidades especiais. Eles descobriram que os professores de Wisconsin gastaram uma média por distrito de 976 horas [anuais] administrando testes. Isso foi particularmente prejudicial para alunos com necessidades especiais. (GUISBOND; NEILL; SCHAEFFER, 2012, p. 4-5)

Não obstante, vige no senso comum o entendimento de que os testes padronizados de aprendizagem são instrumentos científicos confiáveis tão objetivos e precisos como um termômetro ou qualquer outro dispositivo utilizado para medições no âmbito da ciência. Nada mais ilusório. Examinemos algumas das limitações desse tipo de avaliação discente.

1 Visão geral e argumentação central da problemática

A primeira limitação refere-se à abrangência do conteúdo testado. Tais tipos de testes conseguem atingir apenas uma pequena amostra de um amplo campo cognitivo ou de uma determinada disciplina ministrada em uma série escolar específica. Por essa razão, a nota do educando em um teste dessa espécie não consegue espelhar o seu nível de aprendizagem de todo o conteúdo ministrado. Explica-se: segundo Popham (1999), o intuito dos profissionais que elaboram a grande maioria dos testes padronizados de larga escala é criar um instrumento de avaliação **referenciado à norma**, que, com apenas um punhado de itens, produza interpretações válidas do domínio do aluno a respeito de uma parte substancial do conteúdo **em comparação com o grupo normativo**¹.

Referenciado à norma, termo cunhado por Glaser (1963), contrapõe-se a **referenciado ao critério** (*norm-referenced* e *criterion-referenced*, respectivamente). Segundo Brito (2013, p. 136-137):

[...] [b]asicamente, há duas formas de analisar dados sobre o rendimento dos alunos. A primeira seria uma interpretação dos resultados tomando-se por base

¹ “[...] uma amostra nacional de estudantes da mesma idade ou nível escolar” (POPHAM, 1999, p. 9).

um referencial de desempenho, representado pelos objetivos de ensino e respectivos critérios de alcance. Esse tipo de interpretação descreve quais os objetivos que o aluno atingiu dentre aqueles que lhe foram propostos, sem relação a qualquer outro desempenho, seja individual ou de turma, série ou grupo. [...] Esse tipo de medida ‘referenciada ao critério’ é tido como uma medida de padrão absoluto, cuja base é o desempenho individual. Em contraposição, a medida ‘referenciada à norma’ baseia-se no desempenho do grupo, seguindo um padrão relativo. Por essa forma de análise, pode-se determinar os desempenhos, discriminando-os entre si. Na construção dos instrumentos, deve-se atentar para que os itens maximizem as diferenças e promovam uma grande variabilidade de resultados.

Os itens que fazem melhor esse trabalho de **diferenciar os alunos** são aqueles respondidos corretamente por cerca de metade deles. Evitam-se itens que possam ser respondidos corretamente por excessivamente muitos ou poucos discentes, o que não contribuiria para **espalhar os resultados dos testes dos estudantes**, que é precisamente o propósito desse tipo de avaliação².

A necessidade de diferenciação das notas dos discentes, para atender aos requisitos desse tipo de avaliação, elimina itens em que os estudantes teriam um bom desempenho. No entanto, esses itens geralmente cobrem precisamente o conteúdo que, devido à sua importância, os professores priorizam. Assim, por mais que os docentes privilegiem esses conhecimentos e habilidades importantes, o menos provável é que questões que cubram esse conteúdo sejam contempladas em um teste de desempenho referenciado à norma. Dessa forma, **avaliar a eficácia do ensino dos professores utilizando instrumentos de avaliação de seus alunos**³ que evitam deliberadamente conteúdos importantes é fundamentalmente insensato.

² A maioria das empresas de testes padronizados de larga escala utilizam as avaliações **referenciadas à norma**, em contraposição às **referenciadas ao critério**. Nos últimos anos, vem crescendo o número de empresas que vêm adotando a metodologia da **teoria de resposta ao item**, sendo a Pearson, a maior do mundo no segmento, uma delas. No Brasil, o Exame Nacional do Ensino Médio (Enem) vem utilizando essa metodologia em suas últimas edições, a qual, segundo o pesquisador estadunidense Walter Stroup, igualmente padece das mesmas limitações da **referenciada à norma** (SMITH, 2012).

³ Essa é a premissa fundamental do *Value-added Model* (Modelo de valor agregado). Para aprofundamento, *vide* Adler (2014), American Statistical Association (2014), Amrein-Beardsley (2008), Darling-Hammond et al. (2011), Hewitt (2015), Kupermintz (2003), Ravitch (2013) e Rothstein (2011).

Observe-se, ademais, outro efeito desfavorável da adoção desse tipo de avaliação: com vistas a espalhar as notas dos alunos e premidos pela limitação do número de itens, os elaboradores desses testes têm uma predileção toda especial por questões que exijam do educando uma habilidade intelectual inata, não facilmente adquirida na escola⁴. Todavia, itens que essencialmente medem as diferenças de capacidades intelectuais inatas dos educandos, ou do que eles aprenderam fora da escola, obviamente, não contribuem para inferências válidas sobre como as crianças foram ensinadas **na escola**, assim não poderiam ser utilizados para avaliar nem a escola, nem o professor. Perceba-se, igualmente, o viés social e de classe embutido na questão: indubitavelmente, os discentes oriundos de famílias de maior capital intelectual e social têm probabilidade bem maior do que os das classes inferiores de acertar esse tipo de item.

Kohn (2000, p. 77, grifos nossos) esclarece ainda que:

Não importa quantos alunos façam o teste, não importa quão bem ou mal eles sejam ensinados, não importa quão difícil sejam as perguntas, o padrão de resultados é garantido para ser o mesmo: exatamente 10% daqueles que fazem o teste pontuarão entre os 10% com maior escore, e a outra metade vai sempre cair abaixo da mediana. Isso não é porque nossas escolas estão falhando; isso é por causa do que significa ‘mediana’. Uma boa pontuação em um teste referenciado à norma significa ‘melhor do que outras pessoas’, mas não sabemos mesmo **quanto** melhor. Pode ser que as pontuações reais de cada um sejam todas muito semelhantes, caso em que as distinções entre elas são sem sentido – um pouco como dizer que eu sou a pessoa mais alta no meu quarteirão, embora eu seja apenas meio centímetro mais alto do que a pessoa mais baixa no meu quarteirão.

Tal característica das avaliações referenciadas à norma leva a outra séria deficiência, que consiste na questão de que, como bem salienta o autor supra, pelo fato de que toda distribuição de notas sempre conterà um patamar mais baixo, vai aparecer sempre que algumas crianças estão indo muito mal. Isso, por sua vez, reforça a sensação de que as escolas estão falhando. Pior, isso

⁴ A título de ilustração, segue o exemplo do item citado por Popham (1999, p. 13), uma questão de um teste de estudos sociais aplicado a uma turma da 6ª série: “Se alguém quer realmente conservar os recursos, uma boa maneira de fazer isso é: a) deixar luzes acesas mesmo se elas não forem necessárias; b) lavar pequenas quantidades de roupa em vez de grandes quantidades em uma máquina de lavar; c) escrever em ambos os lados de uma folha de papel; d) colocar jornais utilizados no lixo” (a resposta certa é a letra C).

contribui para a insidiosa suposição de que algumas crianças simplesmente não podem aprender – especialmente se as mesmas crianças sempre parecem ficar abaixo da mediana (KOHN, 2000).

Os defensores dos testes padronizados de alto impacto ignoram o fato de que cada criança tem um potencial único de aprendizado chamado por Vygotsky de *Zona de Desenvolvimento Proximal* (doravante ZDP). De forma sucinta, o conceito significa a distância entre o nível de desenvolvimento cognitivo atual da criança, determinado pela capacidade de resolver um problema sem ajuda, e a gama de potencialidades atingíveis sob a orientação de um adulto ou em colaboração com outro colega. Importa assinalar que a determinação de um certo objetivo de aprendizagem desafiador para o aluno somente resultará em melhoria se o objetivo for estipulado levando em conta sua ZDP, do contrário o efeito será o inverso do desejado. Precisamente o atributo de *padronizado* compromete inexoravelmente o êxito desses testes, por não se levar em consideração que cada criança tem uma ZDP distinta (CHAIKLIN, 2003).

Outra fragilidade científica desses exames consiste em ignorar os estágios de desenvolvimento da criança, conforme teorizado por Piaget:

De acordo com a concepção piagetiana, o desenvolvimento cognitivo compreende quatro estágios ou períodos: o sensório-motor (do nascimento aos 2 anos), o pré-operacional (2 a 7 anos); o estágio das operações concretas (7 a 12 anos) e, por último, o estágio das operações formais, que corresponde ao período da adolescência (dos 12 anos em diante). Cada período define um momento do desenvolvimento como um todo, ao longo do qual a criança constrói determinadas estruturas cognitivas. (PALANGANA, 1994, p. 16)

Dessa forma, pela ignorância ou pela simples desconsideração da existência dessas etapas cognitivas do aluno, testes dessa natureza exigem que a criança resolva questões de níveis de abstração tais que estão à frente de seu estágio de desenvolvimento normal.

Por fim, observamos que os testes padronizados não conseguem mensurar as habilidades não cognitivas interpessoais⁵, tampouco as individuais, tais como as capacidades de escutar e

⁵ As capacidades interpessoais referem-se à forma como nos relacionamos com os outros em um ambiente social e inclui aspectos como colaboração, respeito, amizade, liderança, comunicação, solidariedade e partilha, entre muitas outras características (LEVIN, 2013).

de refletir, o controle dos impulsos, a paciência, a persistência, a empatia, a curiosidade, a criatividade, o pensamento crítico, a resiliência, a autodisciplina, a desenvoltura, as responsabilidades social e ambiental, a dimensão cívica e outras de mesma índole e importância social. Ambos os tipos de capacidades não cognitivas, as interpessoais e as individuais, interagem com as habilidades cognitivas para o desenvolvimento integral do ser humano.

A seção a seguir tratará do estreitamento curricular resultante da adoção de políticas focadas na instrumentalização dessas avaliações estandardizadas como esteio para legitimar a segregação educacional, ocupacional e socioeconômica dos estudantes, do início ao final de sua vida escolar.

2 Adestramento para testes padronizados e estreitamento curricular: a diferença entre treinar e educar

Um dos efeitos mais daninhos da instituição dos testes padronizados de larga escala como ferramenta fulcral das políticas educacionais de viés empresarial tem sido a subtração do tempo anteriormente destinado à instrução com o fito de intensificar a preparação para aqueles testes, os quais se transmutaram de meio de aferição (com todas as suas limitações) do conteúdo efetivamente ministrado em objetivo último de todo o processo de ensino-aprendizagem, seu zênite supremo. Para atingi-lo, não se medem as consequências do escanteamento das disciplinas do campo das humanidades e mesmo das ciências naturais, causando um estreitamento curricular que tem engendrado um corpo discente hábil na aplicação de técnicas e “bizus” com vistas a se sair bem naqueles testes, todavia ignorante de conhecimentos básicos daquelas disciplinas menosprezadas, não obstante essenciais para a conformação de um adulto minimamente instruído para o convívio nas sociedades contemporâneas. Ravitch (2010, p. 79, grifos nossos), a partir de Montefinise (2007), ilustra a situação:

Na cidade de Nova Iorque, professores disseram a uma jornalista que eles **eliminaram estudos sociais, arte e ciências por um mês antes dos testes estaduais de leitura e matemática para se concentrarem nas atividades de preparação para tais exames.** Uma professora disse que **seus alunos não sabem quem era o presidente durante a Guerra Civil, mas eles podem te dizer como eliminar respostas em um teste de múltipla escolha.** ‘E, enquanto nossos resultados dos testes estiverem altos, todos estarão felizes’. Seu diretor ordenou-lhe que ‘esquecesse de tudo, exceto da preparação para o teste’. Outra professora disse que os diretores são parcialmente avaliados de acordo com os resultados dos testes, por isso ‘naturalmente, eles querem as notas altas, [e] essa é a nossa prioridade. Educação verdadeira é secundário’.

O episódio é paradigmático da forma como as escolas vêm comprometendo o ensino de disciplinas indispensáveis a uma sólida formação na desesperada tentativa de apresentar notas elevadas em testes⁶. Trata-se, com efeito, de uma distorção do próprio propósito original dos testes, qual seja, avaliar a efetiva aprendizagem, e não tentar “a ferro e fogo” produzir notas altas, independentemente da real e consistente assimilação prévia do conteúdo por parte do aluno.

O aspecto possivelmente irônico ou paradoxal da questão consistiria na constatação de que o foco estreito apenas nas habilidades básicas no idioma nacional e em matemática é incompatível com o desenvolvimento de um “capital humano” criativo, preparado para o desenho de novas tecnologias, dotado de habilidades de interação interpessoal para trabalhar em equipe e outras da mesma natureza⁷, características essenciais na conformação do trabalhador (“colaborador”, no sádico e cruel jargão empresarial) de perfil adequado ao que Sennet (2006) cunha de **novo capitalismo**. Tal incongruência minaria ou dificultaria o suprimento da força de trabalho apta a laborar na decantada “economia do conhecimento”. Nossa hipótese, a qual exploramos em pesquisa em andamento, é a de que inexistente tal paradoxo, porquanto são outras as

⁶ Para mais exemplos desse *modus operandi*, que constitui a regra, e não a exceção, reportar-se ilustrativamente a Fox e Hacker (2008), McMurrer (2008), McMurrer e Kober (2007) e Perlstein (2007). Para acesso a evidências consistentes e abundantes dos efeitos do estreitamento curricular no sistema educacional estadunidense, por exemplo, remetemos a Au (2007), Meyer (2005) e Morton e Dalton (2007), dentre outros autores.

⁷ “Pessoas inovadoras não podem vir de escolas que forcem os alunos a memorizar respostas corretas em testes padronizados ou recompensam aqueles que se destacam em regurgitar conhecimento mastigado” (ZHAO, 2009, p. vii). Na mesma perspectiva, assevera o autor, em outra obra, que “[...] o que gera ótimas notas em testes pode arruinar as qualidades empreendedoras. Os testes padronizados e o foco em memorização mecânica, por exemplo, são talvez os maiores inimigos da capacidade empreendedora.” (ZHAO, 2012, s/n)

necessidades do capital na presente etapa histórica de seu ciclo de reprodução no que se refere à exploração do **trabalho vivo**, na acepção marxiana.

Ravitch (2010, 2013), lembrando que o propósito essencial da escola pública, a razão pela qual ela recebe fundos públicos, é a formação dos jovens para o exercício da cidadania, advoga que toda escola deve ter um currículo rico e balanceado, incluindo artes, ciências, história, literatura, civismo, geografia, línguas estrangeiras e educação física⁸. Observa a pesquisadora:

Hoje, os formuladores de políticas pensam em educação apenas em termos dos seus propósitos secundários. Eles falam das crianças como futuros concorrentes mundiais. Eles, às vezes, referem-se às crianças, em termos bastante infelizes, como ‘ativos humanos’, esquecendo-se que elas são pessoas únicas e que não são fungíveis. Eles querem que todos os alunos estejam ‘prontos para a faculdade e para a carreira’. Eles tendem a falar apenas de preparação para o mercado de trabalho, e não da educação para a cidadania. Mas isso é um engodo. A formação da força de trabalho pode ocorrer nas escolas ou no local de trabalho. Isso não é sem importância. Também não é a preparação para a faculdade sem importância. Mas **preparar para a faculdade não é o objetivo central da educação. Nem o é o treinamento da força de trabalho. O objetivo central da educação é preparar todos para assumirem os direitos e responsabilidades da cidadania em uma democracia.** (RAVITCH, 2013, p. 182, grifo nosso).

Os efeitos do estreitamento curricular, todavia, não atingem igualmente os alunos de todas as classes sociais, sendo os oriundos de famílias de menor *status* socioeconômico e de etnias minoritárias os mais prejudicados, como explicam Finn e Ravitch (2007, s.p.):

Precisamos de escolas que preparam as crianças para se destacar e competir não apenas no mercado de trabalho global, mas também como participantes de pleno direito em nossa sociedade, nossa cultura, nossa política e nossa economia. [...] Abandonar as disciplinas de humanidades em nome das ciências duras [STEM⁹] sozinhas também tende a alargar as divisões sociais e aprofundar as desigualdades internas. Os ricos que entendem o valor da

⁸ Argumenta a historiadora, na primeira dessas duas obras, que, não obstante o currículo não seja a “varinha de condão” que resolverá, por um passe de mágica, todos os problemas educacionais, não tê-lo indica uma inabilidade ou mesmo uma falta de vontade de definir o que se está tentando realizar, tendo como efeito o movimento de andar em círculos e não chegar a lugar nenhum. Mas o pior de todos os efeitos, salienta a autora, é deixar que as editoras de livros didáticos decidam o que as crianças devem aprender.

⁹ Acrônimo para *Science, Technology, Engineering and Math* (Ciência, Tecnologia, Engenharia e Matemática).

aprendizagem das disciplinas de humanidades podem ser os únicos capazes de pagá-las para seus filhos. Escolas privadas de elite e alguns sistemas escolares dos subúrbios ficarão com a educação em sentido lato, assim como as faculdades de elite. Crianças ricas vão estudar filosofia e arte, música e história, enquanto seus colegas pobres preencherão bolinhas em folhas de teste. Os poucos sortudos darão origem à próxima geração de magnatas, líderes políticos, inventores, autores, artistas e empresários. As massas menos sortudas terão oportunidades mais estreitas. Alguns não encontrarão oportunidade nenhuma, e a frustração vai tentá-los a saquear os mais afortunados, que, por sua vez, irão trancar-se em condomínios fechados, clubes exclusivos e ‘privado isso – privado aquilo’, alargando assim as clivagens internas e piorando as nossas perspectivas para a coesão social e a civilidade.

Trazendo a questão para a realidade brasileira, Freitas (2012, p. 389-390, grifos nossos) contextualiza o dano que tal processo inflige à nossa juventude pobre:

O argumento para justificar a limitação ao básico é que os outros aspectos mais complexos dependem de se saber o básico primeiro. Um argumento muito conhecido no âmbito do sistema capitalista e que significa postergar para algum futuro não próximo a real formação da juventude, retirando dela elementos de análise crítica da realidade e substituindo-se por um ‘conhecimento básico’, um corpo de habilidades básicas de vida, suficiente para atender aos interesses das corporações e limitado a algumas áreas de aprendizagem restritas (usualmente leitura, matemática e ciências). A consequência é o **estreitamento curricular** focado nas disciplinas testadas e o esquecimento das demais áreas de formação do jovem, em nome de uma **promessa futura: domine o básico e, no futuro, você poderá avançar para outros patamares de formação. Todos sabemos que a juventude mais pobre depende fundamentalmente da escola para aprender; e, se a sua passagem pela escola for limitada às habilidades básicas, nisso se resumirá sua formação.**

Conclusão

A resistência das comunidades educacionais dos países afetados por essa obsessão pelos testes padronizados vem crescendo, e já mostra resultados significativos em alguns deles.

Nos Estados Unidos, por exemplo, a forte resistência dos pais, dos professores e dos alunos¹⁰ levou o Governo Federal a encerrar o programa *No child left behind* (“Nenhuma criança

¹⁰ Notadamente por meio de massivos boicotes (*Opt out*) aos testes estaduais. Cf. Foster (2016).

deixada para trás’, o qual estruturava-se totalmente em torno de testes padronizados de alto impacto), com a promulgação de uma nova lei educacional, *Every Student Succeeds Act* (“ Todo Estudante Vencerá”), que destina aos estados maior protagonismo na condução de sua política de educação, ao permitir que essas unidades federativas estabeleçam múltiplas medidas de aferição do desempenho acadêmico de seus alunos, possibilitando assim que os testes deixem de constituir a única forma de avaliação do aprendizado dos educandos (BROWN, 2015; FOSTER, 2016).

Cabe, agora, àquele corpo social a vigilância com vistas a impedir que a obcecação por esses tipos de testes continue a prejudicar o alunado daquele país, notadamente os vitimados pela desigualdade e pela segregação socioeconômica e étnica.

REFERÊNCIAS

ADLER, Moshe. Review of Measuring the Impacts of Teachers. *Think Twice Review National Education Policy Center Report*, 2014. Disponível em: <http://greatlakescenter.org/docs/Think_Twice/TT_Adler_TeachImpact.pdf>. Acesso em: 22 mar. 2016.

AMERICAN STATISTICAL ASSOCIATION et al. *ASA statement on using value-added models for educational assessment*. Alexandria, VA: Author, 2014.

AMREIN-BEARDSLEY, A. Methodological concerns about the education value-added assessment system. *Educational Researcher*, v. 37, n. 2, p. 65-75, 2008.

AU, W. High-stakes testing and curricular control: a qualitative metasynthesis. *Educational Researcher*, v. 36, n. 5, p. 258-267, 2007.

BRITO, M. S. T. Norma e critério de desempenho como parâmetros da avaliação da aprendizagem. *Estudos em Avaliação Educacional*, São Paulo, n. 15, p. 135-198, 2013.

BROWN, E. How schools would be judged under ‘Every Student Succeeds’, the new No Child Left Behind. *The Washington Post*, Nov. 30, 2015. Disponível em: <<https://www.washingtonpost.com/news/education/wp/2015/11/30/how-schools-would-be-judged-under-every-student-succeeds-the-new-no-child-left-behind/>>. Acesso em: 27 abr. 2016.

CHAIKLIN, Seth. The zone of proximal development in Vygotsky's analysis of learning and instruction. *Vygotsky's educational theory in cultural context*, v. 1, p. 39-64, 2003.

DARLING-HAMMOND, L. et al. Getting teacher evaluation right: a background paper for policy makers (Research Briefing: American Educational Research Association & National Academy of Education). 2011. Disponível em: <<http://files.eric.ed.gov/fulltext/ED533702.pdf>>. Acesso em: 24 mar. 2016.

FINN, C. E.; RAVITCH, D. Not by geeks alone. *The Washington Post*, Aug. 8, 2007. Disponível em: <<http://www.wsj.com/articles/SB118653759532491305>>. Acesso em: 10 abr. 2016.

FOSTER, J. B. The opt out revolt: democracy and education. *Monthly Review*, New York, v. 67, n. 10, p. 1, 2016.

FOX, L.; HACKER, H. K. Dallas-Fort worth students struggle with TAKS' Shortresponse Written Test. *The Dallas Morning News*, July 20, 2008. Disponível em: <<http://www.dallasnews.com/sharedcontent/dws/news/localnews/stories/072008dnmettaksscores.443a381.html>>. Acesso em: 10 abr. 2016.

FREITAS, L. C. Os reformadores empresariais da educação: da desmoralização do magistério à destruição do sistema público de educação. *Educação e Sociedade*, Campinas, v. 33, n. 119, p. 379-404, 2012.

GLASER, R. Instructional technology and the measurement of learning outcomes: some questions. *American Psychologist*, Washington, DC, v. 18, n. 8, p. 519-521, 1963.

GUISBOND, Lisa; NEILL, Monty; SCHAEFFER, Bob. *NCLB's lost decade for educational progress: what can we learn from this policy failure?* FairTest: National Center for Fair & Open Testing. 2012. Disponível em: <http://fairtest.org/sites/default/files/NCLB_Report_Final_Layout.pdf>. Acesso em: 30 dez. 2015.

HEWITT, K. K. Educator Evaluation Policy that Incorporates EVAAS Value-Added Measures: undermined intentions and exacerbated inequities. *Education Policy Analysis Archives*, Arizona, v. 23, n. 76, p. 1-49, 2015.

KOHN, A. *The schools our children deserve: moving beyond traditional classrooms and "tougher standards"*. Boston: Houghton Mifflin Harcourt, 2000.

KUPERMINTZ, H. Teacher effects and teacher effectiveness: a validity investigation of the Tennessee Value Added Assessment System. *Educational Evaluation and Policy Analysis*, Washington, DC, v. 25, n. 3, p. 287-298, 2003.

LEVIN, H. M. The limits of test score comparisons. Unesco Office in Santiago - *Newsletter LLEC*. June 13, 2013. Disponível em: <<http://www.unesco.org/new/en/santiago/press->

room/single-
new/news/the_limits_of_test_score_comparisons_henry_m_levin_teachers_college_columbia_un
iversity/#.vutgendw7iu>. Acesso em: 12 mar. 2016.

MCMURRER, J. *Instructional time in elementary schools: a closer look at changes for specific subjects*. Washington, DC: Center on Education Policy, 2008.

MCMURRER, J.; KOBER, N. *Choices, changes, and challenges: curriculum and instruction in the NCLB era*. Washington, DC: Centre on Education Policy, 2007.

MEYER, L. The complete curriculum: ensuring a place for the Arts in America's Schools. *Arts Education Policy Review*, v. 106, n. 3, p. 35, 2005.

MONTEFINISE, A. Lost lessons in test-prep craze. *The New York Post*. Jan. 28, 2007.
Disponível em: <<http://nypost.com/2007/01/28/lost-lessons-in-test-prep-craze/>>. Acesso em: 6 abr. 2016.

MORTON, B. A.; DALTON, B. National Center for Educational Statistics, and Institute of Education Sciences (US). *Changes in Instructional Hours in Four Subjects by Public School Teachers of Grades 1 through 4*. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, 2007.

PALANGANA, I. C. *Desenvolvimento e aprendizagem em Piaget e Vygotsky: a relevância do social*. São Paulo: Plexus, 1994.

PERLSTEIN, L. *Tested: one American school struggles to make the grade*. New York: Henry Holt, 2007.

POPHAM, W. J. Why standardized test scores don't measure educational quality. *Educational Leadership*, v. 56, n. 6, p. 8-15, 1999.

RAVITCH, D. *Reign of error: the hoax of the privatization movement and the danger to America's public schools*. New York: Basic, 2013.

RAVITCH, D. *The death and life of the great American school system: how testing and choice are undermining education*. New York: Basic, 2010.

SENNET, R. *A cultura do novo capitalismo*. Rio de Janeiro: Record, 2006.

SMITH, M. A serious design flaw is suspected in state tests. *The New York Times*. July 28, 2012.
Disponível em: <http://www.nytimes.com/2012/07/29/education/texas-studies-suggest-test-design-flaw-in-taks.html?_r=0>. Acesso em: 14 mar. 2016.

I JOINGG – JORNADA INTERNACIONAL DE ESTUDOS E PESQUISAS EM ANTONIO GRAMSCI
VII JOREGG – JORNADA REGIONAL DE ESTUDOS E PESQUISAS EM ANTONIO GRAMSCI
Práxis, Formação Humana e a Luta por uma Nova Hegemonia
Universidade Federal do Ceará – Faculdade de Educação
23 a 25 de novembro de 2016 – Fortaleza/CE
Anais da Jornada: ISSN 2526-6950

ZHAO, Y. *Catching up or leading the way: American education in the age of globalization*.
Alexandria, VA: ASCD, 2009.