

# Uma nota crítica sobre a avaliação estandardizada: a perda de qualidade e a segmentação social

JUAN CASASSUS

jasassus@vtr.net

Universidade Metropolitana de Ciências da Educação, Chile

## RESUMO:

A qualidade da educação e as concepções e práticas de avaliação são as preocupações centrais do autor deste artigo. O seu objectivo é mostrar como as provas estandardizadas referidas à norma, como são as dos sistemas de avaliação centralizados, não servem os propósitos que lhe são atribuídos socialmente, uma vez que não dão informações acerca da qualidade da educação. O autor chama a atenção para o poder desmesurado que é dado aos instrumentos estandardizados de medição na avaliação e gestão do sistema educativo, argumentado que estes, ao contrário do que se espera, promovem a desigualdade e diminuem a qualidade da educação.

## PALAVRAS-CHAVE:

Avaliação e qualidade da educação, Provas estandardizadas de medição, Gestão educacional, Segmentação social, Perda de qualidade.

## INTRODUÇÃO

Nos inícios dos anos 80 assistiu-se a uma viragem nas políticas de educação a nível mundial, que depressa chegou à América Latina. Esta alteração consistiu na mudança do foco das políticas, que deixou de ser a expansão do sistema para se concentrar no que ocorria dentro do sistema. Na altura dizia-se que estávamos a passar de um enfoque centrado na quantidade, para outro, cujo centro seria a qualidade. Para tal, a nova política centrou-se em mudar a regulação do sistema, apoiando-se na disciplina que acabava de nascer: a Gestão Educacional. Lamentavelmente, as esperanças que muitos de nós tivemos nessa altura foram frustradas e temos que reconhecer que as políticas concebidas fracassaram nos seus objectivos. Uma das razões primordiais deste fracasso foi nunca se ter dito o que se entendia por qualidade da educação. De facto, o que era a *qualidade* nunca foi debatido e, portanto, nem sequer houve uma tentativa de consenso. É surpreendente que nunca se tenha dado conteúdo à palavra qualidade. Como no princípio do século XX, quando se discutiu sobre o que era a inteligência e sobre a dificuldade de a definir e se concluiu que inteligência seria o que fosse medido pelos testes de inteligência, com a qualidade da educação usou-se um processo semelhante: a qualidade foi interpretada como sendo equivalente a uma pontuação numa prova estandardizada. Em mais do que um sentido, os testes de inteligência foram os precursores dos sistemas de medição da qualidade da educação.

A ausência de conteúdo e de consenso sobre a qualidade pregou-nos uma grande partida. Por algum motivo, ainda por esclarecer, subentendeu-se que toda a gente sabia o que era a qualidade da educação. De tal maneira que, desde essa altura, quando se formulavam políticas educativas bastava que se dissesse que eram para “melhorar a qualidade da educação” para que as autoridades as aceitassem. No entanto, nunca se trataram de políticas de educação, mas sim de políticas de gestão. As políticas educativas centraram-se na gestão do sistema e nunca da educação.

Só para dar alguns exemplos, recordemos algumas das políticas destinadas a melhorar a qualidade da educação: mais horas de aulas, prolongamento dos calendários escolares, descentralização, medição (sim, muita medição estandardizada e centralizada), medidas centradas nos directores e na sua gestão, concorrência entre escolas, privatização, etc. Nenhuma destas políticas, tão valorizadas pelas autoridades nacionais e internacionais, tiveram algo a ver com a educação. Tiveram, sim, a ver com a gestão do sistema. Não foram geradas pelas disciplinas que se ocupam da educação e suas finalidades. Não têm origem na filosofia, nem na psicologia, nem na pedagogia, nem na sociologia, mas apenas na economia.

Se observarmos os critérios estabelecidos nestas políticas, podemos afirmar que as que têm sido colocadas em prática, desde há 20 anos, fracassaram: os resultados divulgados pelos sistemas de avaliação não mostram progressos. Pelo contrário, em muitos países mostram retrocesso (as pontuações desceram)

e a segmentação social e desigualdade social saíram fortalecidas.

A política educativa é uma arte difícil, porque a educação não é uma ciência exacta. Apenas se podem fazer aproximações, com precaução e senso comum, mas também se tem que estar disposto a introduzir mudanças. Chegou o momento de mudar. Chegou o momento de usar o senso comum e enunciar algumas questões, tais como:

1. *A finalidade da educação (o que eu entendo que seja a qualidade de educação) é que as pessoas sejam melhores e, com isso, que a sociedade seja melhor.* Esta posição implica fazer uma reflexão aprofundada acerca das finalidades da educação no século XXI e qual o tipo de pessoa e sociedade que se quer construir, o que obriga a que se acabe com o pensamento simplista acerca da educação, que evita o debate de fundo e se fica pelo manejo de ferramentas.
2. *Colocar a educação ao serviço das crianças e não ao serviço das autoridades.* O que implica que, em vez de ter uma educação cuja finalidade é a submissão dos alunos ao juízo moral dos adultos, se tenda a gerar uma educação que lhes ofereça a experiência do respeito e da convivência harmoniosa, que fortaleça a sua autonomia, que permita desenvolver a sua capacidade para reflectir e pensar, que permita despertar os sentimentos de empatia e solidariedade e, sobretudo, que não os exproprie do seu poder e dignidade pessoais.
3. *Trabalhar com os docentes e não contra eles.* O que quer dizer que é necessário apoiar o trabalho dos docentes, ouvi-los, confiar neles, solicitar-lhes que pensem em maneiras de melhorar as coisas, dar-lhes tempo e, sobretudo, não agir por meio de constantes pressões, cujo efeito é deprimi-los e fazer com que se sintam mal.
4. *Considerar que são os processos que modulam as formas de convivência que as pessoas desenvolvem consigo próprias e com os outros.* Por isso é realmente importante trabalhar ao nível dos processos que ocorrem na escola (interacções e ambientes) e não pensar que o trabalho educativo consiste em medir resultados.

O propósito desta *Nota Crítica* é fazer um exame crítico à política actual e discutir as razões que nos

levam a pensar desta forma. Estas razões fundamentam-se na evidência empírica de que estas políticas fracassaram, que a ênfase em apenas um determinado tipo de avaliação, a medição, aprisiona o sistema numa dinâmica pernicioso na procura da melhoria da qualidade, que tem como consequência o efeito contrário ao esperado. Ou seja, diminui a qualidade porque diminui a sua abrangência e, como instrumento de gestão, retira aos docentes a sua vontade de melhorar a educação. Os sistemas centralizados de medição afastam os sistemas educativos das finalidades para os quais foram criados.

## AS POLÍTICAS EDUCATIVAS DOMINANTES FRACASSARAM?

A política educativa dos anos 90 e as reformas educativas que proliferaram na América Latina (a par com outras partes do mundo) constituíram um grande esforço de mobilização das energias, deprimidas pelo custo que significou o ajustamento estrutural no mundo da educação na década de 80. Desde o início dos anos 90 que os governos têm aumentado os salários dos professores, construído novas escolas, entregue muitos mais livros e fomentado novas leis. Ao mesmo tempo, auto-proclamou-se que o progresso em educação se reflectiria num aumento de pontuações nas provas de escolha múltipla.

Apesar de todos os esforços encetados, há já 20 anos, a nível mundial, as pontuações não sobem em nenhum dos países em que rendimento escolar é medido. As provas dos estudos comparativos internacionais (TIMSS, LLECE, PISA), bem como as medições nacionais, que se baseiam em avaliações pontuais, devolvem resultados decepcionantes. Tendo em conta estes critérios, a educação parece estar estagnada. Se aceitamos este princípio, a evidência empírica indica que os esforços encetados não colheram os frutos esperados: em alguns países, como é o caso do Brasil, Chile, Espanha, França e EUA, as pontuações têm descido e a desigualdade persiste. A resposta que as autoridades têm dado a esta falência não tem sido no sentido de mudar a política. Pelo contrário, a sua atitude tem sido a de aprofundar a política actual, centrando-a em políticas de standards e em mais avaliação quantitativa. Esta é uma opção que não faz mais do que piorar a situação.

O meu posicionamento parte do princípio de que é o sistema de gestão o factor crucial para o fracasso do sistema e, particularmente, no sistema de avaliação e nas teorias a ele subjacentes. O sistema de avaliação oficial — que é o principal instrumento de gestão — constitui um travão para o desenvolvimento da educação, gera involução e contribui para diminuir a qualidade da educação. De seguida iremos elaborar, sucintamente, alguns dos argumentos que permitem fundamentar este juízo.

## O USO DAS PROVAS ESTANDARDIZADAS E AS SUAS LIMITAÇÕES COMO INSTRUMENTO DE AVALIAÇÃO DA QUALIDADE DA EDUCAÇÃO

As provas de escolha múltipla, como as que se utilizam nos sistemas nacionais de avaliação (como o SAEB no Brasil ou o SIMCE, no Chile), chamam-se “provas referidas à norma”. O propósito destas provas consiste em determinar posições e estabelecer *rankings* entre os indivíduos examinados. Este *ranking* é construído em relação a uma média estandardizada, invariante, de 250 pontos. Robert Glaser, o criador destas provas, disse muito claramente que estas não medem o que sabem os alunos, nem o que eles sabem fazer. Para avaliar o que os alunos sabem ou sabem fazer é necessário recorrer a outro tipo de avaliações, tais como os *portfolios*, a provas de desempenho ou a provas de carácter construtivista. Segundo Glaser (1963, 1997), a única competência medida por estas provas referidas à norma é a capacidade de recordar procedimentos (para a matemática) ou reconhecer um resultado quando se dão múltiplas escolhas. É um erro de natureza conceptual dizer que este tipo de provas mede o que sabem e/ou o que sabem fazer os alunos das escolas. É um erro ainda mais grave equiparar a pontuação que se obtém com o objectivo de conseguir uma educação de qualidade. Pontuação não é qualidade.

Qualidade não é pontuação. Uma educação de qualidade é o que se pode chamar de boa educação. Uma educação de qualidade tem a ver, essencialmente, com a capacidade que a instituição escolar tem para facilitar que as pessoas se transformem em melhores pessoas, que a sociedade se transforme em melhor sociedade. É uma actividade de conhecimento transformador. O nível de qualidade de uma escola

é proporcional à profundidade das análises que se coloquem à disposição dos alunos, do tipo de perguntas que eles possam colocar, do tipo de projectos em que se possam envolver e do tipo de problemas que são capazes de resolver. A qualidade educativa não é uma actividade centrada em obter pontuações altas. São âmbitos distintos e é um erro confundi-los, pois tem consequências negativas.

## DE QUE FORMA O PODER DESMESURADO DADO AOS INSTRUMENTOS DE MEDIÇÃO EM AVALIAÇÃO AFECTA NEGATIVAMENTE A QUALIDADE DA EDUCAÇÃO

De uma forma voluntária ou involuntária tem vindo a ser dado, a este tipo de avaliação, a legitimidade e o poder de ser considerada o eixo estrutural da política educativa. É das provas de avaliação que resultam as políticas de *accountability*, de incentivos e de castigos, de afectação de recursos, de orientação das competências, de uso da informação para a tomada de decisões, etc. Isto é um contra-senso se o comparamos com as intenções de melhoria da qualidade da educação. Se analisarmos o que aconteceu quando estes sistemas foram colocados em prática, verificamos que são mecanismos que na realidade servem para piorar e não para melhorar a qualidade da educação. Convém também salientar que é um erro assinalar que o êxito ou fracasso de uma política se mede pelo subir de pontos, numa medição que não mede o que se lhe atribui e que o sistema, tal como está concebido, não pode modificar.

Poderíamos fazer a seguinte pergunta: *Por que é que se afirma que este tipo de avaliação, instalada como meio de gestão, se converteu num instrumento que diminui a qualidade, quando foi desenhada para aumentar a qualidade na educação?* A resposta é que não só se converteu num travão ao desenvolvimento da educação, como, ao contrário do que se afirma em vários meios, está a fazer involucionar o sistema. Esta situação pode ser explicada a partir de diferentes ângulos. A título de exemplo podemos destacar que:

1. *As pontuações descem.* Como já se assinalou, a mera observação das estatísticas comparáveis mostra que, apesar de todos os esforços, as

pontuações continuam a descer. Perante esta situação, a tentativa tem sido a de aprofundar o exercício de uma política de pressão sobre os docentes, em vez de fazer uma revisão do sistema.

2. *Efeitos perversos das provas.* A importância que se tem dado a esta forma de avaliação faz com que a qualidade da educação se tenha transformado em sinónimo de pontuações em provas. Esta é uma identificação nefasta e, ao mesmo tempo, com vários efeitos:

- a. Faz com que o foco da educação seja a resposta a provas psicométricas e não o ensino, que é o fundamento da profissão
- b. Esconde aspectos que são importantes finalidades da educação, a saber: o desenvolvimento da personalidade, o respeito, a cidadania, a curiosidade, o desenvolvimento de valores, a vontade de descobrir conhecimentos, o compromisso com a sabedoria, etc. Todas estas coisas que não são avaliadas pelas provas psicométricas. Em definitivo, a finalidade da educação é colocada de lado no processo educativo, uma vez que não é avaliada.

3. *Subir nas pontuações cria ilusões.* Identificar qualidade da educação com a subida de pontuações nas provas de medição gera situações ilusórias. Treinar para responder a uma prova estandardizada é o mesmo que, quando uma pessoa tem febre, lhe dermos medicamentos ou arrefecermos-lhe a cabeça, de forma a quando se coloque o termómetro para medir a febre e esta esteja a descer se conclua que o doente se curou. Da mesma forma se podem subir ou desvirtuar as pontuações de muitas maneiras, como por exemplo:

- a. Inflacionando as notas (como demonstram os sistemas de acesso às universidades)
- b. Não incluindo os “maus” alunos quando se tem que dar um teste (quando este conceito de “mau” não tem um conteúdo específico, é apenas uma interpretação)
- c. Dando aos alunos as respostas às questões colocadas nas provas
- d. Treinando os alunos na resposta às provas
- e. Os alunos são livres para responder aleatoriamente

4. *A educação torna-se algo superficial.* Centrar a política nas provas instala uma dinâmica que faz com que a educação se torne superficial.

- a. Porque a habilidade para responder a provas de escolha múltipla não implica reflectir sobre a forma de alcançar uma resposta, é apenas o acto de escolher uma resposta. Aprender a reflectir é importante, porque é assim que se consegue aprofundar um tema e, por conseguinte, elevar o nível de qualidade do domínio do conteúdo
  - b. Porque responder a provas de escolha múltipla não exige a reflexão nem o raciocínio, mas apenas o reconhecimento de uma resposta. Aprender a pensar e raciocinar é importante, porque permite encontrar as estruturas lógicas e interligar os acontecimentos
  - c. Porque responder a provas deste tipo não exige a competência de construir conhecimento, mas apenas a capacidade de “tomar como sua” uma resposta já dada
  - d. “Ensinar para o exame” implica que os professores ocupem o tempo a exercitar os alunos a escolher uma resposta entre as apresentadas. Implica que os professores não dediquem tempo a ensinar ou a [fazer] pensar
  - e. Subir ou descer pontos é uma actividade superficial que não tem nada que ver com uma educação de qualidade. Pensar que um sistema educativo melhora (ou piora) porque sobem (ou descem) as pontuações, é o mesmo que pensar que a saúde é melhor quando se põem termómetros aos doentes depois de lhes ter dado aspirinas ou aplicado compressas de água fria
5. Consequências sociais: aumenta a desigualdade
- a. Nos colégios com baixos recursos, ensinar para o exame “para subir pontos... a todo o custo”, significa que os professores se vêm forçados não a educar, mas sim a treinar os alunos nas respostas de provas de escolha múltipla. Nos sectores mais abastados da sociedade, esta actividade é paga fora das horas das aulas
  - b. Orienta os gastos para as provas e não para o ensino. Na teoria que está subjacente a esta visão, supõe-se que o nível de rendimentos da família determina o resultado da prova. As diferenças nos resultados dessas provas não se devem a factores educativos. Assim, como já se sabe o resultado de antemão, e, por outro



lado, sendo conscientes de que as variações são marginais, em vez de continuar a avaliar (e a pagar para isso), seria mais útil utilizar os recursos da avaliação em actividades educativas mais ricas e produtivas para os alunos

## USAR PROVAS ESTANDARDIZADAS COMO INSTRUMENTO PARA A GESTÃO DO SISTEMA EDUCATIVO É UM ERRO

A avaliação baseada em provas psicométricas referidas à norma está desenhada para seleccionar e não para incluir. Por isso, em contextos segmentados como são os sistemas educativos da América Latina ou do sistema educativo chileno, promove a desigualdade. Esta é uma dimensão obscura que tem dois aspectos: os sistemas de medição são políticos e o uso da informação é político.

Diz-se que os testes de inteligência (QI) ou as provas dos sistemas nacionais de medição são objectivos, mas, na realidade, não o são, pois estão sujeitos a ênfases e às condições específicas de quem os elabora. Mas enfatizar que são objectivos tem um efeito perverso. Assim sendo, aceita-se que, em sociedade, a inteligência e o desenvolvimento mental se distribuem numa curva normal. Ou seja, é por sorte ou azar que alguns estão no extremo dos “burros” e outros estão no outro extremo, o dos “inteligentes”. Se os testes de inteligência (QI) ou os sistemas nacionais de medição fossem objectivos, então, supunha-se que, dada a normalidade da curva, no extremo dos “burros” tinha que haver tantos ricos como pobres e, no extremo dos “inteligentes”, também deveria haver tantos ricos como pobres. Mas não é o que acontece. Os resultados das provas indicam-nos que no extremo dos “burros” e das baixas pontuações se encontram os pobres, e que os ricos se encontram no extremo dos “inteligentes” e com pontuações altas. É esta a única informação válida e consistente deste tipo de provas: as escolas municipais e os pobres têm más pontuações, porque se verifica, “cientificamente”, que é ali que estão os “burros”, que, por azar, são pobres. Um sistema meritocrático numa democracia formal, ou seja uma democracia que não toma em consideração a desigualdade de condições e os contextos de pobreza, que nos diz que há uma razão de ser para os burros serem pobres: é porque são

burros. Assumir isto é esquecer um ponto crucial: os efeitos negativos que a pobreza tem nas aprendizagens e no desenvolvimento cognitivo dos alunos.

Em democracias formais, como o Chile, os sistemas de medição nos quais a informação se distribui de uma forma aberta, partindo da suposição que isso permitirá situar os alunos em relação à sua geração, faz precisamente isso, situa-os na parte alta ou baixa da curva. A razão para que tal aconteça reside no facto de que o sistema montado é uma estrutura de *controlo social* e o uso que se faz das informações é o mecanismo de manutenção do *controlo social*. Isto faz com que os sectores carenciados se auto-convençam que têm más pontuações porque são *burros*. Consequentemente, é “natural” que estejam onde estão e igualmente “natural” que os ricos estejam onde estão. Que as elites económicas dirigem porque são inteligentes e que os pobres os seguem porque são “burros”. Em última instância implica que se isto é assim, é por razões eugénicas.

Por outro lado, é uma ilusão pensar que serve para melhorar a qualidade da educação. Este tipo de avaliações não são úteis aos professores para que possam dar melhores aulas e, acrescentando-se, não são úteis à gestão. Criam a ilusão de se que está a tomar o pulso ao sistema (ver a metáfora do termómetro acima mencionada), mas esta ideia apenas se move na irrealidade. Usá-las como instrumentos de gestão apenas serve para identificar posições e rankings, ou dito de outra forma: segregar e desintegrar; gerando, em consequência e a todos os níveis do sistema, muito sofrimento inútil e injusto.

Neste texto não falamos da avaliação em geral, uma vez que é conveniente ter algum tipo de avaliação, mas de uma avaliação que tem como consequência não a melhoria da qualidade, mas sim o aumento da pressão sobre os docentes, mediante uma parafernália de prémios e castigos. É muito difícil compreender a lógica subjacente às teorias que levam a pensar que os professores ensinarão melhor se castigados (ou premiados). É bem conhecido na literatura do “management” e nas práticas de gestão de ministérios e secretarias da educação que este sistema não funciona. Tem, por vezes, algum efeito a curtíssimo prazo, mas que rapidamente desaparece.

É normal pensar-se que os prémios e castigos são incentivos para a acção e, por isso, aparecem como uma medida fácil de implantar, mas note-se

que quando falamos de “incentivos” o que se está a dizer é que o sistema utiliza uma motivação externa (extrínseca) para que os professores façam o seu trabalho. Há pouco já se mencionou que as motivações extrínsecas não funcionam como resposta ao que querem as autoridades do sistema. No entanto, é verdade que funcionam para outra coisa. A motivação extrínseca, como instrumento e política de Estado, parte de falsas suposições e tem efeitos perversos:

1. *Diminui a dignidade dos docentes.* Parte-se da suposição que os docentes não estão intrinsecamente (internamente) motivados para fazer o seu trabalho. Esta é uma falsa suposição e tem um efeito perverso. Temos que perguntar: Porque é que se parte da suposição que os docentes não estão internamente motivados para fazer bem o seu trabalho? Supor que não estão, como política de estado, é retirar dignidade ao docente e ao seu trabalho. Esta perda de dignidade não se consegue compensar com um aumento de remuneração.
2. *Mina a motivação intrínseca.* Há também outros efeitos perversos, por exemplo, supor de forma sistemática que os docentes não estão motivados para fazer o seu trabalho, mina de forma igualmente sistemática a motivação intrínseca. Uma boa educação depende da motivação intrínseca dos professores. Ter uma política que nega esta motivação intrínseca é ter uma política que estimula uma educação de má qualidade.
3. *Torna os processos mais rígidos.* Estabelecer standards e aplicar avaliações, não apenas promove a motivação extrínseca, como também faz com que os procedimentos sejam mais rígidos e estandarizados. Por um lado, é uma situação que convém àqueles professores não motivados para responder de forma criativa aos problemas colocados pelo ensino. Mas, por outro lado, inibe os docentes que estão motivados e “corta-lhes as asas” para tentarem encontrar formas mais adequadas para ensinar as crianças, particularmente se se trata de uma política curricular construtivista.
4. *Destrói o clima na sala de aula.* Inimiza professores e alunos e fragiliza a confiança e vínculos entre eles. A possibilidade de seleccionar alunos faz com que se expulsem os “maus” alunos para as escolas municipais, que estão obrigadas a recebê-los. Num sistema são e convivial, um aluno com

dificuldades é um desafio e é a razão do trabalho dos docentes em escolas mais pobres. Mas num sistema que se move sobre a base de juízos relacionados com as pontuações, no qual a sobrevivência de um docente depende das pontuações dos seus alunos, um “mau” aluno não pode deixar de ser visto pelo docente como uma ameaça e como um problema que “oxalá não existisse”.

## CONCLUSÃO

Para terminar esta *Nota Crítica*, só nos podemos perguntar como é que chegámos a este estado de coisas. O que terá acontecido para que se abandonasse a tradição de educar para a formação humana, de nos preocuparmos com os alunos, de educar para melhorar a sociedade? Porque é que se tem obrigado os docentes a abandonar o entusiasmo pelo seu trabalho, fazendo com que entrem neste jogo de pressões e ameaças? São estas as perguntas que coloquei e a procura dolorosa de respostas levou-me a concluir que tal aconteceu porque a educação e, particularmente, a avaliação actual se transformaram num negócio muito lucrativo. Enquanto continuarmos a pensar que a pontuação numa prova psicométrica é um indicador de qualidade, as autoridades continuarão dispostas a comprar sistemas de provas. E fá-lo-ão com o entusiasmo de quem crê que está a fazer um favor aos mais desprotegidos. Acreditarão que quando treinam os alunos num sistema de instruções para que respondam a provas compradas a alto preço, lhes estão a oferecer uma educação de qualidade... E não perceberão, como se afirmou antes, que lhes estão a dar uma educação empobrecida, uma vez que o sistema de standards e provas concentra-se, precisamente, em aspectos, como a memorização de informações, que possam ser medidos pelas referidas provas. Os *standards estandarizados* e as provas psicométricas deixam de lado tudo aquilo que é mais difícil medir: aprender a pensar, aprender a respeitar, aprender a viver com os outros, aprender a fazer perguntas relevantes e a resolvê-las, a procurar a evidência do conhecimento, a determinar o que é importante e válido, a aprender com o contexto. Ou seja, tudo aquilo que parece estar no âmago de uma educação de qualidade.

## REFERÊNCIAS BIBLIOGRÁFICAS

- GLASER, R. (1963). Instructional technology and the measurement of learning outcomes: some questions. *American Psychologist*, 18, pp. 519-521.
- GLASER, R. (1997). Education for all: Access to learning, and achieving usable knowledge. *International Seminar on Educational Evaluation*. Ministry of Education, Rio de Janeiro, November 30 — December 3.

Tradução de Gabriela Lourenço