

Avaliação de larga escala e o (pseudo) foco na aprendizagem

Thomaz Edson Veloso da Silva

Graduação em licenciatura em física - UFC/BR

Mestre em engenharia de teleinformática - UFC/BR

Doutor em engenharia de teleinformática - UFC/BR

PhD in science - UCPH/DK

Pós-doutorando em análise de dados educacionais - UFC/BR

Coordenador de pós-graduação e inovação pedagógica - FIED

Ex-coordenador de análise de dados, pesquisa e avaliação educacional - Esfapege/Sobral-CE✉

thomazveloso@virtual.ufc.br

Recebido em 24 de outubro de 2018

Aceito em 15 de novembro de 2018

Qual o real propósito de fazer uso de uma avaliação de larga escala?

É uma questão de *accountability*. O gestor educacional, seja ele de qualquer esfera, tem a necessidade de prestar contas com a sociedade sobre o trabalho que vem sendo desenvolvido em sua unidade ou rede de ensino. A utilização de testes padronizados, com forte suporte analítico estatístico, dá a possibilidade de comparação entre séries e unidades escolares. Dessa forma, pode-se perceber, de uma forma geral e sem entrar em especificidades, se uma política educacional está dando certo ou não, com base na compreensão do desempenho acadêmico do alunado que esses tipos de testes procuram medir.

Torna-se importante salientar que nos últimos anos, o marketing associado às grandes escolas do Brasil está associado a números obtidos em olimpíadas, vestibulares e, mais recentemente, no Exame Nacional do Ensino Médio (Enem). Em se tratando de mercado, a livre concorrência permite que os produtos oferecidos à sociedade sejam suficientemente expostos para atrair clientes. Mas parece que o foco exacerbado em testes vem tirando o foco do que realmente importa: a aprendizagem dos alunos.

No ensino público, falar dos números e indicadores das escolas e das redes de ensino é uma realidade. A ampla maioria daqueles que compõem a educação pública brasileira sabem o que significa Prova Brasil, Ana, Pisa, Ideb e etc. Torna-se importante destacar que as avaliações de larga escala, que alguns conhecem pelo nome de "avaliação externa", foram criadas com dois objetivos principais:

- Possibilitar aos gestores do processo educacional uma visão macro de como os alunos estão aprendendo; e
- Conseguir parametrizar seus resultados de modo que modelos estatísticos comparativos possam ser utilizados para facilitar a análise de desempenho dos alunos.

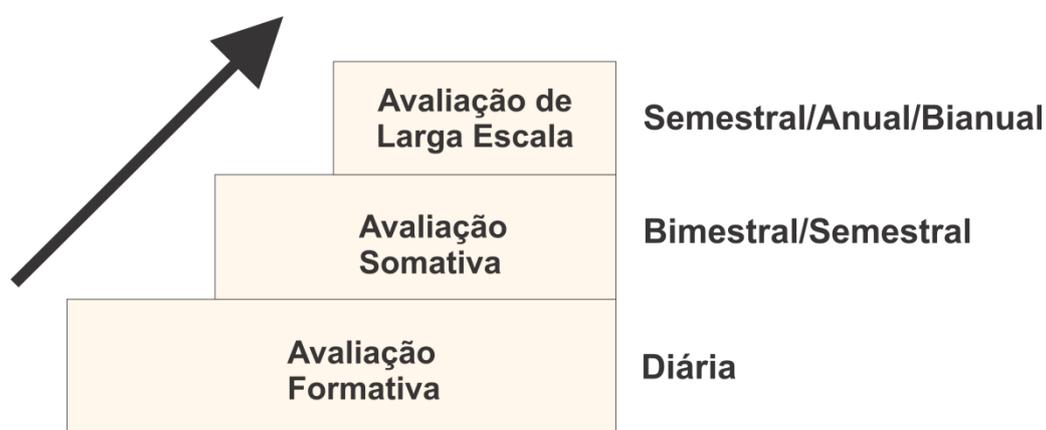
Bom, isso realmente acontece e não é ponto de discussão aqui se o gestor faz bom ou mal-uso desses dados obtidos. O problema é quando se tenta a qualquer custo atingir metas e cumprir agendas que não estejam diretamente associadas ao desenvolvimento da aprendizagem dos alunos. O que se sabe é que o uso desses dados pode servir para: fins comparativos e de benchmarking; melhorar o sistema educacional geral do país por meio de políticas diretas; melhorar o acesso e a equidade; melhorar a prática de ensino e aprendizagem; reformas curriculares; e utilizando estratégias e indicadores para monitorar e avaliar processos educacionais.

Mas uma pergunta muito importante pode vir à tona: **Será que os resultados gerados pela avaliação são realmente fidedignos ao sistema educacional avaliado?** Na verdade, é imprescindível ressaltar que as avaliações de larga escala contam com um robusto formalismo estatístico no tratamento dos seus dados. O Funcionamento Diferencial dos Itens (do inglês *Differential Item Functioning* - DIF) refere-se a itens elaborados para cada disciplina avaliada, classificando-os relativamente quanto à abordagem de conteúdos mais fáceis ou mais difíceis para diferentes grupos de pessoas. Claramente, o DIF pode ameaçar a validade e a confiabilidade dos resultados da pesquisa, pois está aberto à manipulação na seleção de itens que favorecem certos grupos e precisa-se ter um cuidado redobrado para evitar fraudes. Em pesquisas internacionais, está se tornando claro até que ponto os itens funcionam de maneira diferente em diferentes países e grupos culturais.

Um exemplo muito claro de como o DIF funciona está associado à nota atribuída ao candidato que presta o Enem. Exemplificando, se temos 40 questões em uma prova e dois candidatos, um acerta as primeiras 20 questões, tidas como mais fáceis, e o outro acerta as 20 últimas questões, tida como as mais difíceis. O nome DIF compreende que não existe lógica num padrão de acerto de todas as questões mais difíceis sem ter acertado nenhuma questão fácil. Logo, embora ambos os candidatos tenham acertado o mesmo quantitativo de questões, aquele que acertou as mais fáceis vai receber uma nota maior do que aquele que acertou as mais difíceis.

Devemos reconhecer que a avaliação de larga escala não substitui a avaliação do professor, também conhecida como avaliação interna. Referimo-nos a essas avaliações como "internas" porque são projetadas, selecionadas e usadas pelos professores de acordo com as necessidades de avaliação dos seus alunos, sendo utilizada para subsidiar a prática pedagógica do professor no dia a dia de sala de aula, além de fomentar a visão macro do coordenador pedagógico ao suporte à formação docente. Ainda dentro do conceito de avaliação interna, essa pode ser definida como formativa ou somativa.

Figura 1 – Avaliações associadas à aprendizagem do aluno e suas relações de complementaridade.



Fonte: Própria

Na Figura 1, destaca-se uma escada evolutiva onde se têm como base a avaliação formativa. O objetivo da avaliação formativa é monitorar a aprendizagem do aluno para fornecer *feedback* contínuo que pode ser usado pelos professores para melhorar seu ensino e pelos alunos para melhorar seu aprendizado. Já no degrau subsequente, temos a avaliação somática, que tem como objetivo avaliar a aprendizagem do aluno no final de uma unidade instrucional, comparando-a com algum padrão ou benchmark. Pode-se observar que ambas estão associadas ao desenvolvimento da aprendizagem dos alunos, ou seja, tem forte influência no último estágio da escada, a avaliação de larga escala. Ainda pode ser observada na Figura 1, a costumeira periodicidade de aplicação de cada tipo de avaliação.

Existe um corpo substancial de literatura sobre o valor da educação e seus benefícios para um indivíduo e para a sociedade. Pesquisas empíricas de todo o mundo demonstram o papel crítico da educação em ajudar as pessoas a sair da pobreza, melhorar sua qualidade de

vida, fortalecer sua saúde e a de sua família, aumentando suas oportunidades de emprego e contribuindo para o desenvolvimento econômico de seu país. Mas isso efetivamente acontece, quando existe um cuidado em sala de aula com o aluno, o preparando para algo muito além do que ser bem sucedido em um teste.

CONSIDERAÇÃO FINAL

Temos que focar todos os esforços na aprendizagem e no desenvolvimento socioemocional dos nossos alunos. O bom desempenho em avaliações, de larga escala ou internas, é fruto de um trabalho pedagógico diário. Dessa forma, encaramos a avaliação como uma etapa natural do processo educacional, tendo sua importância ao longo da trajetória do aluno, mas sem precisar ser o foco de todo processo. No mais, não precisamos de criança que sabem macetes e "bizús" para resolverem questões padronizadas, precisamos de cidadãos pensantes e críticos!

LEITURAS RECOMENDADAS

Petra Lietz & Mollie Tobin. The impact of large-scale assessments in education on education policy: evidence from around the world. **Research Papers in Education**, Vol. 31, n. 5, pp. 499-501, 2016. DOI: 10.1080/02671522.2016.1225918

Phillip Perrenoud. Avaliação: da Excelência à Regulação das Aprendizagens - entre duas lógicas. Trad. Patrícia Chittoni Ramos – Porto Alegre: Artes Médicas Sul, 1999.

Heraldo Marelim Vianna. Fundamentos de um Programa de Avaliação Educacional. Brasília: Liber Livro Editora, 2005.

OECD. PISA 2015 Results (Volume II): Policies and Practices for Successful Schools, PISA, OECD Publishing, Paris, 2016.



Este trabalho está licenciado com uma Licença [Creative Commons - Atribuição 4.0 Internacional](https://creativecommons.org/licenses/by/4.0/).