

**ESTRUTURAS DE AVALIAÇÃO ESCOLAR PARA MAPEAR HABILIDADES QUE
TOMA COMO BASE AS TAXONOMIAS DE BLOOM EM QUESTÕES DE
MÚLTIPLA ESCOLHA**

Autor: Henrique Araken Martins

Instituição: Universidade Federal do ABC (UFABC)

E-mail: prof.h.araken@gmail.com

Resumo:

Na educação, desde sempre existe a necessidade de melhorar os resultados do desempenho de aprendizado de nossos alunos em avaliações internas e externas (SARESP). Sendo assim, com grande frequência surgem novas propostas com intenção de aperfeiçoar o todo, ou parte do processo educativo. O trabalho realizado iniciou sua configuração atual no ano de 2014. Na prática, desenvolvemos uma planilha de correção on-line que dinamizou o processo de correção, fornecendo-nos indicadores de processo instantâneos, permitindo-nos refletir, discutir e propor intervenções do nosso ensino, através dos resultados de aprendizado de nossos alunos representados em um gráfico de cores. Contudo, definimos como hipótese: *As informações no gráfico de cores são confiáveis? Ou, são meramente ilustrativas?* Essa hipótese foi discutida no início de 2016, na UFABC, sendo aprovada por banca a utilização do gráfico de cores em prol do ensino e aprendizado. Observando também, comprovação de ganhos nos resultados da avaliação do SARESP.

Palavras-chave: Planilhas de autocorreção, Taxonomia de Bloom, Avaliação e Educação Matemática.

Introdução

O trabalho “*Estruturas de Avaliação escolar para mapear habilidades tomando como base a Taxonomia de Bloom em questões de múltipla escolha*”, foi desenvolvido com o paralelismo de aspectos da Educação Matemática observando o Currículo/Avaliação Bimestral e por métodos Estatísticos, garantindo um alto nível de confiabilidade dos resultados apresentados. Tentamos demonstrar através de testes estatísticos a validade das técnicas de mapeamento sugeridos na formatação, aplicação e devolutiva das avaliações feitas na unidade, Escola Estadual Prefeito Nestor de Camargo - Barueri. O prazo da devolutiva dos resultados foi também priorizado, pois nada adiantaria um resultado com alto nível de confiabilidade sendo entregue após um ou dois meses da realização de Avaliações Bimestrais. De modo concomitante, vimos essas técnicas de mapeamento, discutindo a dependência de habilidades na formatação das questões dos temas dados, utilizando conceitos da Taxionomia de Bloom. Nessas duas frentes, nos preocupamos a responder a interrogação de nossa pesquisa: *As informações no gráfico de cores são confiáveis? Ou, são meramente ilustrativas?*

Existem diversos tipos de Avaliações, e nesta pesquisa analisamos apenas avaliações objetivas (Questões de Múltipla Escolha - QME), descrevendo a sua elaboração com as mesmas orientações para se construir itens em Teoria de Resposta ao Item (TRI), sugeridos em Rabelo (2013). Mantivemos grande foco na aplicação digital, na correção instantânea e principalmente, na devolutiva dos resultados para todos os envolvidos. A decorrente busca de resultados em avaliações externas nos faz refletir e tentar correlacionar junto às Avaliações Internas meios de influenciar beneficemente o aprendizado de nossos alunos. Hoje as avaliações externas de larga escala como: PISA, SAEB, Prova Brasil e o ENEM em nível Nacional e o SARESP em nível Estadual, tornaram-se referências em devolutiva com um alto padrão de confiabilidade, sendo realizadas com a metodologia da Teoria de Resposta ao Item. Porém, essas avaliações são realizadas apenas uma vez por ano (ou, a cada dois ou três anos, sendo que às vezes é observada apenas uma amostra), e as devolutivas, devido a suas características de correção, são entregues em longo prazo. Além disso, suas intervenções propostas para melhora desses resultados não atingem 100% da população dos avaliados e não mostram as defasagens individuais. Portanto, seria de suma importância equiparar as avaliações bimestrais com as avaliações externas, construindo itens para as provas bimestrais com as mesmas normatizações das provas externas. As habilidades presentes nessas avaliações externas, após serem feitos todos os processos de correção e análise da TRI, são classificadas

em níveis de complexidade de acordo com sua régua de proficiência, sendo eles: Abaixo do Básico, Básico, Adequado e Avançado. Para adquirirmos maturidade para construção de nossas avaliações, analisamos os resultados fornecidos no relatório pedagógico de matemática dos SARESPs realizados nos anos de 2008 a 2013. Neles observamos que os exercícios (itens) com características de reconhecimento se inseriam nos níveis Abaixo do Básico/Básico e os exercícios com características de compreensão, em sua maioria, se inseriam nos níveis Adequado/Avançado. Analisamos também antigas avaliações internas da área de Ciências da Natureza e Matemática dos anos de 2012, 2013 e primeiro Bimestre de 2014, e observamos que as questões tinham características de resolução mecânica com muitos termos e aspectos *livrísticos*.

A configuração atual do trabalho contempla: Investigação, estudo e experiência com uso de procedimento de instrumentos de avaliação on-line. Privilegiando, a avaliação da aprendizagem em Matemática e nas disciplinas relacionadas à área de Ciências da Natureza, em três níveis de complexidade, na modalidade de ensino médio. Os objetivos desse trabalho são: propor uma metodologia de elaboração de questões seguindo os patamares da hierarquia de aprendizado proposto por Bloom; aplicação on-line de avaliações de múltiplas escolhas; priorizar a devolutiva dos resultados com a finalidade de mapear por meio de um gráfico de cores as habilidades dos alunos, das turmas, das séries, e da escola; obter indicadores do ensino dos professores e obter indicadores dos conteúdos envolvidos.

Cronologia das estruturas de avaliação

Será apresentada uma síntese do desenvolvimento: O trabalho se iniciou como uma boa prática escolar, e evoluiu com o decorrer do tempo. Dividimos esta evolução em três etapas da pesquisa, que são apresentadas a seguir.

Etapa I

Na Escola Estadual Prefeito Nestor de Camargo, desde o início de 2012, aplicávamos a avaliação bimestral contendo quarenta questões divididas igualmente nas quatro disciplinas de Ciências da Natureza e Matemática de modo convencional, imprimíamos as avaliações e após todos os alunos responderem corrigíamos manualmente.

Os recursos tecnológicos foram essenciais para realização dessa empreitada. No final de 2012 recebemos na Escola 120 Net books, para uso diverso voltado para educação e aprendizado.

No segundo bimestre de 2013, construímos a avaliação bimestral completamente on-line utilizando como ferramenta os formulários do Google. Após o aluno enviar a avaliação, as informações se direcionam para uma planilha Google. Então, copiávamos as informações da turma e colávamos em um gabarito preparado no Excel, que fornecia o resultado geral. Devido à quantidade de computadores disponíveis, realizávamos a avaliação em três turmas. A correção eletrônica tinha o objetivo de mostrar a quantidade de acertos da Avaliação de modo instantâneo. Conseguimos efetuar todo o processo de correção, com a média de cinco minutos. Portanto, as nossas expectativas estavam satisfeitas momentaneamente.

Etapa II

A aplicação de Avaliações com resultados quase que instantâneos trouxe-nos um dinamismo no tempo de correção. Porém, percebemos que poderíamos explorar ainda mais os recursos que já manuseávamos. Melhoramos a planilha de correção, automatizando todas as informações necessárias com recursos de *aplicativos Google*, e começamos a analisar as alternativas, observando o percentual de distribuição entre as respostas por questão. Com isso obtivemos ganho no momento da correção, e na devolutiva da Avaliação aos alunos. Conseguimos identificar rapidamente a alternativa que a maioria optou. Caso ela fosse a errada, tínhamos discussões de grande valia com os alunos e principalmente com a equipe de elaboração

das questões. Neste momento nosso objetivo não se restringia em ter a nota do aluno, e quantas questões ele acertou por disciplina. Tínhamos a ambição de retirar mais informações da avaliação. Assim, modificamos completamente o modelo da prova; de Avaliação Somatória passamos a elaborar Avaliação Formativa, pensando na formação do aluno e na mensuração do que ele assimilou no Bimestre. Nossa avaliação passou a ter trinta e seis questões, sendo nove por disciplina. Cada disciplina fechava todo conteúdo em três *temas/assunto* trabalhado no Bimestre. De cada tema, o professor fornecia três questões com níveis de interpretação e resolução distintas, sendo uma questão de cada nível. As informações de complexidade contidas em cada nível são descritas a seguir:

Nível I - Identificação / Observação – Uma questão mais simples possível, podendo ser: teórica, informativa ou de um fato cotidiano. O aluno teria somente a necessidade de ter prestado atenção na explicação do professor, ter a habilidade de leitura e interpretação ou simplesmente conhecer brevemente sobre do que o assunto tratava. Em linguagem técnica, essa questão se insere no primeiro patamar da Taxonomia de Bloom: Lembrar.

Nível II - Resolução / Desenvolvimento – Uma questão que aborda principalmente o processo mecânico. O aluno deve saber como desenvolver o cálculo ou desenvolver as ideias presentes em seu contexto, mesmo que seja de modo mecânico. Seria o segundo e o terceiro patamar da taxonomia de Bloom: Entender e Aplicar.

Nível III - Compreensão / Aplicação – Uma questão teoricamente mais complexa, podendo ser uma situação problema ou uma atividade retirada de Avaliação Externa, ou de Vestibulares. O aluno precisa reconhecer a parte teórica e desenvolver conhecimentos para resolver a atividade. Atingindo os patamares Aplicar, Analisar e Sintetizar.

Observe que os *Patamares da Taxonomia de Bloom*¹ podem sobrepor-se na alocação dos níveis. Com essas normalizações conseguimos comparar alunos de turmas, e séries diferentes. Para facilitar na comparação, relacionamos essas três questões, construindo um gráfico de cores, que nos traz várias informações que nos direcionam a comentar e intervir no processo. Apresentando a configuração completa da planilha de respostas, temos: todas as alternativas por aluno; a tabulação por pergunta; separação de notas por disciplina; total de acertos; tabulação de competência obtida; gráfico de mapeamento; tabulação de conteúdos aplicados. Observemos a imagem na página a seguir, (foram ocultadas algumas linhas e colunas):

¹ Será explicado no capítulo seguinte

Etapa III

No final de 2014, eu já cursava o quarto semestre do curso de Mestrado Profissional (PROFMAT) no campus da UFABC. Nesta etapa, cogitamos utilizar o trabalho realizado na escola como tema para dissertação para obtenção de Título. Em torno de alguns professores da escola, já existia o questionamento referente ao gráfico de cores. Contudo, junto ao orientador da dissertação (Prof. Dr Valdecir Marvulle) reafirmamos como hipótese as seguintes questões: Esse gráfico de mapeamento (colorido) é confiável ou não? Ele realmente representa uma devolutiva de aprendizagem efetiva do aluno? Assim, decidimos analisar com um pouco mais de detalhes o referido gráfico. Tomando como base, a parte pedagógica, e um estudo estatístico. Como base pedagógica, utilizamos Taxonomias de Bloom, no estudo estatístico, utilizamos Teste de Friedman e o Teste de Wilcoxon para dados emparelhados.

Será apresentado, breve resumo das bases teóricas utilizadas para responder a interrogação e validação do gráfico de cores.

Taxonomia de Bloom nas Estruturas de Avaliação para Mapeamento de Habilidades

A Classificação de Metas Educativas, proposta e direcionada por Benjamin Samuel Bloom foi executada por ele e uma comissão de pesquisadores em universidades americanas. Iniciou-se em meados do ano de 1950 projetando a classificação hierárquica de domínio psicométricos dependentes entre si, divididas em três grandes grupos: O cognitivo, o afetivo e o psicomotor.

Essa taxonomia será uma das principais ferramentas pedagógicas e educativas contidas neste Trabalho. Toda a base pedagógica se formou a partir das informações do domínio cognitivo, observando a hierarquia taxonômica de complexidade em aprendizagem intelectual em construção de questões objetivas, e principalmente para validação do gráfico de cores que expõe a devolutiva da assimilação de alunos no período estudado.

Segundo a Taxonomia de Bloom, os objetivos educacionais no domínio cognitivo são classificados em uma hierarquia de seis níveis: Conhecimento (lembrar), Compreensão (entender), Aplicação, Análise, Síntese e Criação. Esses seis níveis são ilustrados nas seguintes etapas de demonstração de aquisição e devolução cognitiva.

Taxonomias de Bloom como alicerces da avaliação, esperavam certo comportamento na análise dos resultados. Lembrando, temos três habilidades que compõem o conjunto de uma competência, dadas em respectivamente três questões. Essas três questões chamaremos de: Q1 a questão de identificação, Q2 a questão de processo mecânico e Q3a questão de compreensão.

Fonte: Autor (2016)

Vejamos um exemplo de tabela que representou os resultados:

Figura 3: Modelo de tabela analisada

Bimestre	Questão	Média (DP)	IC 95%	Mediana (Mín-Máx)	p*
2º/2014	1	0,570 (0,280)	(0,505; 0,634)	0,670 (0,000 – 1,000)	0,001
	2	0,502 (0,267)	(0,441;0,563)	0,330 (0,000 – 1,000)	
	3	0,351 (0,301)	(0,281; 0,420)	0,330 (0,000 - 1,000)	
3º/2014	1	0,757 (0,259)	(0,697; 0,816)	0,670 (0,000 – 1,000)	0,001
	2	0,333 (0,280)	(0,268;0,397)	0,330 (0,000 – 1,000)	
	3	0,551 (0,256)	(0,492; 0,610)	0,670 (0,000 - 1,000)	
1º/2015	1	0,493 (0,327)	(0,418; 0,568)	0,330 (0,000 – 1,000)	0,135
	2	0,480 (0,271)	(0,417;0,542)	0,330 (0,000 – 1,000)	
	3	0,417 (0,259)	(0,357; 0,476)	0,670 (0,000 - 1,000)	

* *Teste de Friedman*

Na tabela acima, foi verificado e comparado de modo combinado a proporção de acertos de cada questão: Q1, Q2 e Q3, analisando se, dentre elas, existe diferença significativa (com $p < 0,001$) (teste de Wilcoxon com correção de Bonferroni). Verificando, se a movimentação teve o aspecto esperado $Q1 > Q2 > Q3$. Tirando como conclusão se a amostra em estudo diferenciou o grau de dificuldade dentre as questões.

Considerações finais

O foco dessa análise, não foi comprovar a evolução desses alunos na realização deste tipo de avaliação, e sim, discutir os fatos ocorridos com essa amostra. Quanto maior a discussão, mais informações teremos sobre o aluno e sobre o processo, conseqüentemente construiremos mais proposta de intervenção que beneficiariam futuras avaliações e o aprendizado de nossos alunos.

Chegamos á conclusão que, o gráfico de cores realmente nos trás informações pertinentes que se aproximam muito da realidade dos conteúdos assimilados pelo aluno. Em relação à construção da avaliação, sua correção computadorizada e as regras de formação das

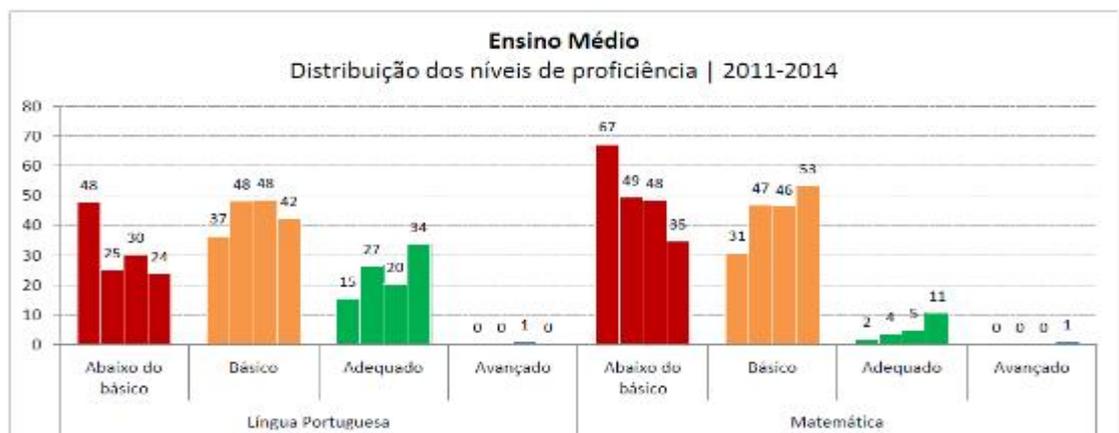
cores, nossos estudos mostraram que toda a estrutura pedagógica tem bases em estudos psicométricos; os itens devem ser construídos com todas as normas especificadas para sua formatação abrangendo os conteúdos do bimestre. Vimos também o comportamento dos alunos nos resultados das avaliações, onde percebemos que no decorrer das aplicações, estatisticamente, o resultado converge para o padrão esperado. Desse modo, obtemos mutuamente as duas frentes pedagógica e estatística se validando.

Os professores têm o mapeamento da sala, sabendo até quem colocou as respostas de modo aleatório. O mapeamento da sala é importante também para o professor saber o grau de complexidade que os alunos atingiram e qual conteúdo deve ser repassado. Conseqüentemente, também começam a perceber as diferenças dos três níveis de complexidade das questões ao longo do processo, evitando falhas na elaboração das mesmas como foi verificado em algumas avaliações.

O coordenador pedagógico tem o mapeamento de todo o conteúdo transmitido, observando o gráfico horizontalmente, vemos o aprendizado do aluno e observando o gráfico verticalmente é possível verificar o ensino (desempenho do professor) monitorando o programa bimestral.

A equipe de gestão se preocupa com o todo, avaliações internas e avaliações externas. Temos resultados de avaliação externa (SARESP) dos últimos quatro anos (2011 à 2014) nas disciplinas de Português e Matemática:

Figura 4: Resultados do Saresp – EE Pref. Nestor de Camargo-Barueri



O interessante é que a estrutura de avaliação foi aplicada apenas em Matemática. Como a mesma população realizou as duas avaliações, pudemos observar o comportamento de rendimento dos alunos nesses anos. Nesse período (2011 à 2014) tivemos também avaliações bianuais do SARESP em Ciências da Natureza (2012 e 2014) :

Figura 5: Tabela de ganhos

	2012	2014	Crescimento
Ciên. Natureza	1,85	2,99	62%
Matemática	1,81	2,61	44%
Língua Portuguesa	3,37	3,66	9%
Mat. Estado	1,64	1,67	2%
Port. Estado	3,1	2,94	-5%

Com esses dados notamos que as estruturas de Avaliação influenciaram de modo benéfico nas avaliações externas. Ou seja, concluímos que todo o esforço desenvolvido para modificarmos os paradigmas de avaliação utilizando a Taxonomia de Bloom e os gráficos de cores foram decisivos para a melhoria do desempenho da escola estudada nas avaliações externas.

Referências

[Bl] Bloom, Benjamin S. Taxonomia de Objetivos Educacionais (1956). Avaliação da Aprendizagem escolar: estudos e proposições, Boston, MA. Copyright (c) 1984 por Pearson Education.

[Bl] Bloom, Benjamin S. Todas as nossas crianças Aprendizagem, New York: McGraw-Hill.

[Bl] Bloom, Benjamin S. Desenvolvimento de Talentos nos jovens, New York: Ballantine Books.

[Ho] HOFFMANN, Jussara Maria Lerch; Avaliação: mito e desafio: uma perspectiva construtivista, Porto Alegre: Editora Mediação, 1991.

[Im] IMBERNÓN, Francisco Formação Docente e profissional: formar-se para a mudança e a incerteza. - 5. ed., São Paulo: Cortez, 2005.

[Lu] Luckesi, Cipriano Carlos Avaliação da Aprendizagem escolar: estudos e proposições, São Paulo: Cortez, 2005.

[Pe] PERRENOUD, Phelippe. Avaliação da excelência à regulação das aprendizagens entre duas lógicas, Porto Alegre: Artes Médicas, 1999

[Ra] RABELO, Mauro. Avaliação educacional: fundamentos, metodologias e aplicações no contexto brasileiro, Rio de Janeiro, RJ : Sociedade Brasileira de Matemática, 2013