

MATRIZ DE ESPECIFICAÇÕES: suporte operacional de uma avaliação

SÉRGIO FRANCISCO COSTA*

A PROVA COMO INSTRUMENTO DE MEDIDA

Por que se fazem **provas**? Que é que uma prova pretende **medir**? Começamos dizendo que uma prova, quando bem concebida, bem montada, é um **instrumento de coleta** e ao mesmo tempo de **medida**. Assim, da correta aplicação desse instrumento, resulta, em geral, um **número que exprime uma relação quantitativa** entre o que queremos medir e a própria **escala de medida**. Por seu turno, a **escala de medida** é uma espécie de “**linguagem do instrumento**”.

Mas por que o educador precisa desse número? Precisa dele por várias razões, dentre as quais:

a) estabelecer comparações entre as competências individuais dos alunos;

b) ter uma visão geral do andamento e desenvolvimento dos alunos com relação aos objetivos gerais e específicos propostos no início de cada período escolar;

c) calibrar (ou recalibrar) a sua ação docente em função da qualidade das respostas dos alunos.

Não percamos de vista que

**medir é atribuir números a uma grandeza,
de conformidade com um critério.**

* Fundação Carlos Chagas e Universidade do Grande ABC - São Caetano do Sul

Então, o importante é que esse critério contenha algumas características desejáveis: **sólido**, no sentido de **não-ambíguo**, **simples**, ou seja, **pouco ou nada dependente de sofisticações operacionais**, e **interpretável**, o que significa **não deixar dúvidas sobre as relações quantitativas ou posicionais por ele sugeridas**.

Essas características desejáveis consolidam-se numa escala de medida que, sendo a **linguagem do instrumento**, anuncia, através de um número, a **quantidade** que a magnitude medida tem de **determinado traço**.

Assim, uma prova, enquanto instrumento de medida, "pergunta" ao aluno o **que e quanto** ele sabe sobre **determinado assunto** e não há como conseguir que ela cumpra esse desiderato se, na base de sua concepção, não existirem **objetivos claramente definidos**.

Embora soe redundante, **uma avaliação é sempre maior do que a prova que lhe dá origem**. O avaliador (professor), tendo atribuído, com base numa prova, por exemplo, certa nota a um aluno, cumpriu apenas uma **parte da tarefa avaliativa**; a avaliação se completa a partir do momento em que são feitas **reflexões** sobre alguns aspectos bem como tomadas algumas **providências** tendentes a produzir **resultados educacionais melhores**. Por exemplo, e seguramente sem esgotar a lista:

- o aluno não está aprendendo ou eu (professor) não estou tendo sucesso em fazê-lo aprender?;
- o aluno está próximo ou distante da média da classe?;
- os objetivos propostos (para a unidade submetida à prova) foram efetivamente cumpridos?;
- os materiais didáticos utilizados contribuíram efetivamente para a obtenção dos resultados?;
- o meu discurso (de professor) leva em conta a maturidade do aluno e os seus recursos de comunicação e compreensão?

Além disso, essa avaliação, para ser boa, deve caracterizar-se por **justiça, adequação e coerência**:

- **Justiça** - o aluno tem o direito de ser avaliado não só com vistas a uma promoção senão também para situar-se relativamente ao conteúdo e ao nível da classe. Além disso, ele não deve ter a sensação de que há um componente paternalista no processo avaliativo segundo o qual as notas poderiam não estar relacionadas com o seu desempenho. Por isso, a comunicação do critério de correção deve ser anterior à aplicação da prova, de modo que, ao concluí-la, o aluno possa ter uma razoável idéia da nota que irá tirar.

- **Adequação** - uma prova escolar é um instrumento de medida como qualquer outro. Por essa razão, sua característica fundamental deve ser a **validade**. Se a prova não medir o traço ou o conjunto de traços que o avaliador tem em mente medir, sua validade estará seriamente comprometida. A adequação também diz respeito à relação entre o **tamanho da prova** e o **tempo disponível** para realizá-la. Submeter um aluno a um conjunto de 100 questões de certa dificuldade em **tempo insuficiente** é correr o risco de estar medindo,

além do **desempenho, resistência física** (ou psíquica). Ora, para medir resistência física existem instrumentos bem melhores do que uma prova escolar: corrida, levantamento de pesos, vigília etc...!

- **Coerência** - a coerência diz respeito a como o professor usa a avaliação relativamente aos seus acordos com a classe. Em outras palavras, **não deixar de cobrar**, mas não cobrar **além** do combinado, não cobrar **alguém** do combinado.

TIPOS DE PROVA

Fazendo uma supersimplificação, é possível classificar as provas em dois grandes grupos: **não-objetivas e objetivas**.

As provas **não-objetivas** possuem algumas características importantes:

- fáceis de montar, difíceis de corrigir;
- possuem embutido forte elemento emocional (pessoal);
- impedem a comunicação rápida dos resultados;
- sofrem fortemente a influência de fatores psicológicos da parte de quem as corrige.

As provas **objetivas**, por seu turno, possuem as seguintes características importantes:

- difíceis de montar, fáceis de corrigir (com base em gabaritos de correção);
- elementos emocionais embutidos praticamente inexistentes;
- favorecem a comunicação rápida dos resultados;
- não sofrem a influência de fatores psicológicos da parte de quem as corrige, isto é, duas pessoas distintas logram atribuir a **mesma nota** a uma **mesma prova**;
- **o mesmo avaliador**, ao recorrer a **mesma prova** algum tempo depois, atribui-lhe a **mesma nota**.

NÍVEIS DE CONHECIMENTO

Existem, grosso modo, **três níveis de conhecimento: reconhecimento, evocação e domínio** (ou **maestria**). Se queremos que um aluno **reconheça** algo que lhe foi ensinado, é fundamental que lhe mostremos a **coisa entre outras** para que ele possa **identificá-la**. Ou ele **discrimina** ou **não discrimina**. Se queremos que ele **demonstre** domínio, mandamos que **faça algo**, estimulamo-lo a **exibir comportamentos** que sugiram sua maestria no assunto. A prova de **múltipla escolha** presta-se muito bem a **medir** os dois níveis de conhecimento retromencionados. A prova **dissertativa (não-objetiva)**, por seu turno, é muito boa para dar ao aluno a oportunidade de demonstrar **evocação**; mas, como isso está estreitamente ligado à memória, a oportunidade para demonstrar o que sabe pode não ser equivalente à oportunidade que o professor quis dar ao aluno para que demonstrasse saber o que de fato era julgado importante!

O MAPA SUGERE O CAMINHO

Supondo aceita a hipótese de que **provas objetivas** possam realmente oferecer melhor serviço à causa da avaliação escolar, o foco de nossas considerações, daqui por diante, estará centrado sobre elas.

Relembremos que uma **prova**, entendida como um **conjunto de itens (questões)**, constitui uma **amostra** dentre as **infinitas possibilidades de outras provas** a que os alunos poderiam ser submetidos. Então, para que elas sirvam bem ao seu propósito, é preciso que os itens sejam construídos em sintonia com **objetivos específicos**.

Parece inegável que viajar munido de mapas torna a viagem mais segura e possivelmente mais agradável. O mapa possibilita enxergar, por antecipação, algumas peculiaridades geográficas bem como, se for o caso, evitar outras. Em Educação não é diferente! Os objetivos servem de referencial para a concepção de um curso, para a eleição de disciplinas, para a escolha do docente ideal ou para a construção de provas que forneçam as evidências necessárias à já mencionada avaliação. Por isso, o **ponto de partida** para a construção de itens deve ser sempre o conjunto de objetivos de aprendizagem estabelecidos para a unidade estudada.

Os objetivos de aprendizagem "iluminam" a trajetória do educando, dando sentido às ações a que o processo educacional o submete.

Tecnicamente, o "mapa", isto é, os objetivos de aprendizagem podem ser reunidos num quadro plano que relaciona **traços de interesse (categorias)** com **conteúdos**. O resultado final é uma **matriz ou tabela de especificações**, que informa o autor de itens do que ele deverá questionar (conteúdo), com que objetivos (traços de interesse) e quantidade ideal de itens para a mensuração do cumprimento dos objetivos de aprendizagem estabelecidos.

Naturalmente, essa matriz pode variar em função de diversas especificações: disciplina, grau escolar, tempo disponível (para a prova) e objetivos específicos (por exemplo, nível de dificuldade). De qualquer modo, se a **prova é mais do que uma simples coleção de questões, se o seu objetivo é coletar informações a respeito do que o aluno sabe e quanto sabe sobre determinado assunto**, deve ela representar, através das questões, uma espécie de "sonda" que capte a desejável **relação entre o que elas perguntam e os objetivos que instruem cada uma delas**.

MODELO TRIDIMENSIONAL DA MATRIZ DE ESPECIFICAÇÕES

A tarefa de construir em larga escala bons itens (questões) depende do cumprimento de importantes fases, que se sucedem em ordem mais ou menos

fixa. A primeira fase consiste em decidir que disciplina será objeto de exame (Matemática, Português, História etc.) e em que nível escolar ela se situa (por exemplo, 1º grau - 3ª série; 2º grau - 1ª série etc.). Essas duas informações contextualizam a matéria em alguma proposta curricular.

Tomada essa decisão, convidam-se a seguir professores experientes da disciplina escolhida e estudiosos de *curricula* e pede-se-lhes que, com base no material didático disponível no mercado, na sua experiência pessoal, no exame que tem sido feito das estruturas lógicas da disciplina e em suas convicções enquanto educadores, construam uma lista que arrole os *conteúdos* (da disciplina em causa) considerados relevantes. Essa lista, depois de exaustivamente analisada, é hierarquizada (do geral para o específico) e cada **tópico** é desdobrado em seus aspectos **essenciais**. Por exemplo, na disciplina Matemática, o tópico **números** pode ser desdobrado em **inteiros, frações, porcentagens e decimais**. Assim, **tópicos** definem **áreas de conteúdo** e desdobramentos definem **conteúdos específicos**.

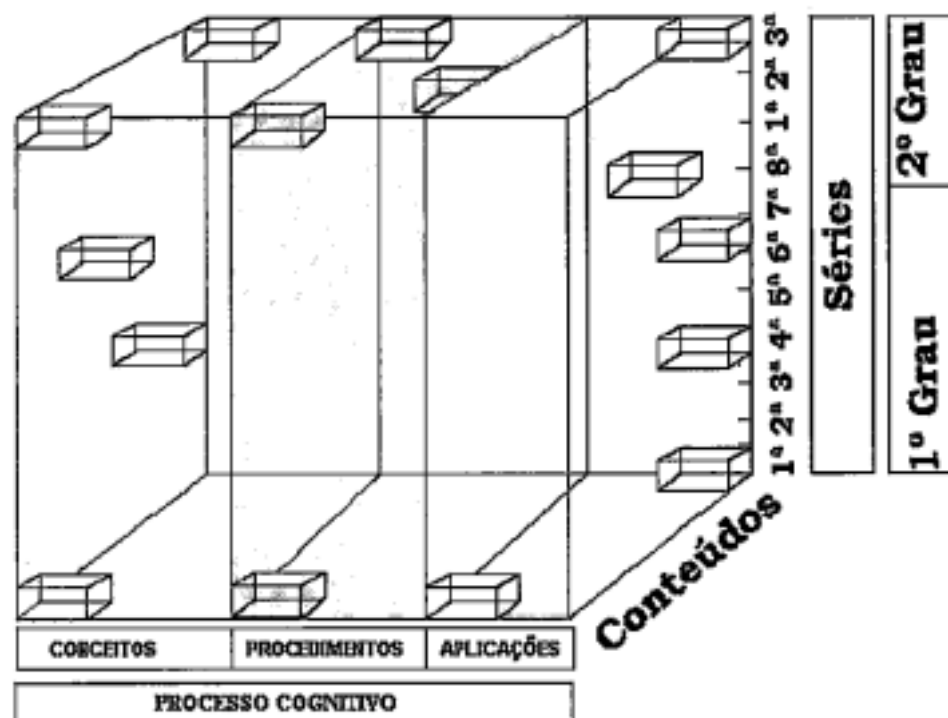
A tarefa, a seguir, consiste em estabelecer, relativamente a cada área de conteúdo e a cada conteúdo específico, os **processos cognitivos** implicados na aquisição e no desenvolvimento do seu conhecimento. Isso, naturalmente, implica ter, na base de todas as decisões, uma taxonomia; e embora existam várias disponíveis, é sempre possível buscar uma que resulte da própria matéria sob exame, a exemplo da que emergiu em trabalho recente realizado na Fundação Carlos Chagas, a propósito de uma pesquisa solicitada pelo Sistema Nacional de Avaliação da Educação Básica (SAEB) realizado em 1995 em continuação de etapas anteriores (1990 e 1993). Assim, relativamente à **dimensão cognitiva**, ficou estabelecido trabalhar, em Matemática, com **conceitos**, com **procedimentos** e com **aplicações**, reunidos sob a denominação genérica de **processos cognitivos**. Nessa linha de raciocínio, ficam, por conta de conceitos, termos, definições, princípios; por conta de procedimentos ficam **cálculos, reformulações, construções, estimações, execução de algoritmos**; finalmente, sob **aplicações** reúnem-se **provas, inferências, solução de problemas expressos em linguagem natural e aplicações no mundo real***.

Cruzando-se então **conteúdos** com **processos cognitivos**, numa tabela de dupla entrada, resultam **caselas** que encerram ("aprisionam", enquadram) **objetivos curriculares específicos**. Aí, é construir itens que meçam especificamente cada objetivo curricular, com o cuidado de, na medida do possível, **evitar contaminações entre caselas contíguas**. Mas a construção de itens deve levar em conta mais uma variável: o **nível de escolaridade** pressuposto na sua resolução, o que é dado pelo **grau** e pela **série** curriculares.

Numa tentativa de tornar mais visível a matriz de especificações, procuramos compará-la a um **tabuleiro**, cuja **largura** representasse os **processos cognitivos**;

* Algumas das idéias contidas nesta página foram inspiradas em trabalhos ainda não divulgados de autoria de Phillip R. Fletcher, Pesquisador Sênior da Fundação Carlos Chagas.

o comprimento, os conteúdos; e a altura, a segmentação dos conteúdos e dos processos, em ordem crescente, e em função do grau e da série. Na figura a seguir, o lado esquerdo do cubo abriga conceitos; o lado direito, aplicações; a parte central, procedimentos. Cada cubo miniaturizado, por sua vez, abriga, pelo menos teoricamente, um único objetivo curricular e a totalidade dos cubinhos representa a matriz de especificações completa.



Construída a matriz, e antes de utilizá-la como suporte para a construção de itens, é preciso garantir a sua **validade**. Nesse sentido, é ela submetida a uma **banca de especialistas** que têm a incumbência de pronunciarem-se sobre a sua adequação. Nesse momento, o que se pergunta é o seguinte: **1. Será que a estrutura dos tópicos (e seu detalhamento) corresponde à ordenação lógica da disciplina?** ; **2. Será que a matriz explora, à exaustão, os objetivos curriculares propostos para a série X do grau Y?** Trata-se de trabalho minucioso e árduo, mas que tem o condão de expor eventuais fragilidades na matriz, ensejando alterações importantes.

Completada mais essa etapa, uma equipe de professores experientes é convocada para a elaborar os itens. Algumas regras, de longa data testadas, servem de respaldo para que os itens saiam com qualidade. Por exemplo, dar preferência a textos não-redundantes e em linguagem natural, evitar ordem indireta, não misturar objetivos numa mesma questão, procurar manter o item independente dos demais etc.

A construção de itens não é tarefa fácil. Os professores, mesmo os mais experientes, logo se dão conta de que, depois de terem construído alguns itens, começam a repetir-se, mudando pequenos detalhes, mas mantendo a mesma estrutura. Isso pode ser bom, por um lado, uma vez que contribui para o aumento do acervo de um particular tipo de item; por outro lado, ofuscados pela abundância de itens semelhantes, os professores podem perder a oportunidade de exploração de todas as caselas da matriz de especificações.

Finalmente, construídos os itens, o passo seguinte é submetê-los a uma outra banca de especialistas com o objetivo de assegurar que não haja descompasso entre o que propõe a matriz e o que cobram os itens. Essa fase busca dar **legitimidade** aos itens e impedir que sejam propostas questões inadequadas a algum nível escolar.

FINALIZANDO

Esperamos ter conseguido demonstrar a importância de que se reveste a construção de bons **itens** com vistas numa **avaliação** que produza **crescimento** simultâneo de **alunos** e respectivos **professores**. À semelhança de uma estrada que une dois pontos, uma proposta curricular precisa, além dos pontos de saída e chegada, de facho que iluminem a trajetória de todos os implicados no processo: os **objetivos**.

Os objetivos, por sua vez, devem merecer tratamento adequado, a fim de que especifiquem claramente os **comportamentos de saída, observáveis**, que possam denunciar a **aquisição**, pelo aluno, das **dimensões cognitivas planejadas pelo educador**. Nesse sentido, a **matriz tridimensional de especificações** cumpre o importante papel de "aprisionar", em unidades pequenas, **objetivos reduzidos à sua forma mais simples**, de modo que os itens construídos com o seu suporte possam ser **específicos, não-redundantes e isentos de contaminações de objetivos colaterais**.

A moderna **Teoria da Resposta ao Item**, concebida nos anos 50 e só mais recentemente desenvolvida graças aos avanços da Estatística e da Informática, constitui campo fértil para a utilização de uma matriz tridimensional, como a apresentada, e presta-se, com algumas vantagens sobre as teorias clássicas, à **produção de itens em larga escala**, capazes de, a um só tempo, **avaliar a "saúde" curricular de uma comunidade** bem como a **posição individual de alunos no seu grupo de referência**. Vale a pena tentar!

