

A importância da avaliação: o que suas provas podem te explicar

Vinício dos Santos

Manual do Professor Iniciante,
manualdoprofessoriniciante.com.br

Material sob licença Creative Commons CC BY-ND 4.0¹

Suas memórias escolares devem estar impregnadas de momentos que envolvem provas: notas boas inesperadas, notas baixas que te traumatizaram, noites estudando, estudar faltando dez minutos para a aula, professores carrascos e professores bonzinhos. Pensar em escola é, automaticamente, pensar naquilo que parece ser o propósito dela: fazer provas. Assim, ao avançar nos anos e retornar à escola, desta vez como docente, é difícil se libertar deste imaginário. Neste texto, vamos defender que devemos sim falar e pensar em provas o tempo inteiro, mas em uma perspectiva bem diferente daquela a que nos acostumamos – e entender como as avaliações são fundamentais para promover a aprendizagem dos nossos alunos.

O que as provas têm a nos ensinar

Tradicionalmente, para alunos e professores, a prova é um momento de encerramento. Depois de algumas semanas de aula trabalhando um objeto de conhecimento, habilidade ou procedimento, o docente aplica uma prova que irá medir o que os estudantes aprenderam e, em seguida, classificar os resultados como sucessos ou fracassos, supostamente conectados à conduta dos alunos nas semanas anteriores.

Por conta disso, a prova assume uma função dupla, tanto de elemento de classificação, por definir quais alunos passam ou não passam de ano, como de coação, pelas ameaças do que o fracasso pode causar. Com tamanha importância e peso, logo a prova deixa de ser o final, para virar a finalidade da aula: estudantes assistem às aulas para fazer

¹ Este texto foi licenciado sob a [Licença Creative Commons CC BY-NC 4.0](https://creativecommons.org/licenses/by-nd/4.0/), que permite aos usuários copiar e distribuir o material sob qualquer formato, com direito a alterá-lo como quiser. Entretanto, é necessário dar crédito ao autor original, indicar a fonte do trabalho e se houve modificações no conteúdo original. Não é permitida qualquer exploração comercial do material.

prova, e o professor, para aplicar prova, precisa dar uma sequência de aulas de preparação. As provas são vistas pelos estudantes como um adversário a ser superado, e se ele tiver exigências baixas, ou for possível vencê-lo com alguma trapaça, não tem nenhum problema. Ao mesmo tempo, para o professor, a prova e seus resultados são o argumento de força para mostrar aos estudantes que seus comprometimentos serão determinantes, repetindo o mantra de que “quem estudou foi bem, que não estudou foi mal”.

Esta visão está cristalizada na cabeça dos alunos e dos docentes. Já discutimos antes como nossas práticas de professores ecoam e repetem nossas experiências como alunos e aqui não seria diferente – mas, desta vez, os resultados são muito mais negativos do que positivos. Quando tomamos as avaliações apenas como acertos de contas, esquecemos sua natureza básica: avaliar é executar uma análise rigorosa, a partir de evidências de qualidade, que nos permitam emitir um juízo ponderado sobre uma situação.

Ao submeter um estudante a uma prova, avaliamos se ele adquiriu as habilidades e competências desejadas e trabalhadas durante um ciclo. Estes resultados têm caráter duplo, pois falam sobre o aluno, mas também sobre a eficiência do trabalho do professor. Esta segunda parte, porém, não é compreensível facilmente, já que a separação entre o trabalho docente e como o aluno se apropriou dele é sutil, e mantém relações estreitas com a construção da avaliação.

Na obra *A prática educativa: como ensinar*, publicada pela primeira vez em 1995 pelo pedagogo catalão Antoni Zabala, o autor traça uma comparação entre a Medicina e a Educação, contrapondo a relação paciente-médico com a relação aluno-professor: “Quando vamos ao médico, não tentamos lhe esconder os sintomas nem o resultado do tratamento, porque consideramos que seus objetivos são os mesmos que os nossos, que o que ele quer é nos ajudar” (p.209). Por outro lado, na avaliação escolar, “o objetivo básico do aluno não seja dar a conhecer suas deficiências para que o professor ou a professora ajudem-no, mas, ao contrário, demonstrar ou aparentar que sabe muito mais [...] A filosofia da prova é a do engano, a do caçador e da caça e, portanto, não promove a cumplicidade necessária entre professor e aluno” (p.209).

Se nos mantivermos na analogia, veremos também que os exames estão para Medicina como as provas estão para a Educação. Na primeira, eles traçam quadros da saúde

do paciente, que permitem identificar doenças ou averiguar a eficácia de um tratamento. Fazer um exame de sangue, por exemplo, é o mesmo que produzir uma descrição dele, que será apreciada por um profissional. Assim deveríamos encarar as provas escolares: ao respondê-las, os alunos apresentam quadros robustos do que conseguem ou não conseguem fazer, e o docente avalia, a partir de parâmetros estáveis, se aquelas características são satisfatórias ou não. Novamente na Medicina, veremos que é ridículo pensar se “passamos ou não” em um exame de urina. O que obtemos é uma avaliação que aponta se o corpo, para os critérios dados, está saudável ou não e, se não estiver, a partir do resultado, um tratamento é proposto – exatamente o que devemos buscar com as provas escolares.

Se as provas são momentos de obter informações precisas e verossímeis sobre as habilidades dos alunos, quanto mais forem realizadas, melhor. A literatura escolar já tem uma definição para esta abordagem: são as avaliações formativas. Em *Avaliação das aprendizagens, para as aprendizagens e como aprendizagem* (2022), a pedagoga e doutora em Educação Benigna Villas Boas informa que a avaliação formativa “promove as aprendizagens usando evidências do que os estudantes aprenderam, em relação aos objetivos que se pretende alcançar, para que sejam planejados os próximos passos e a maneira de atingi-los [...] [e] propicia a identificação das necessidades, facilitando o trabalho docente, munindo o professor de condições de apresentar aos pais a situação de aprendizagem de seus filhos”.

A noção de avaliação formativa surge como contraponto à tradição da avaliação somativa que, nas palavras de Villas Boas, “tem hora para acontecer; faz um balanço da conquista das aprendizagens em certos componentes curriculares ou em todos, durante determinado período; costuma ser realizada por intermédio de provas”. Este “balanço das conquistas” é traduzido para o sistema numérico de notas e, posteriormente, será utilizado para embasar a decisão do corpo docente da escola se o aluno pode ou não ser promovido a etapa seguinte do ciclo.

Nos discursos educacionais atuais, há uma tendência a opor avaliação formativa e somativa, em uma dicotomia superficial em que a primeira é desejável e moderna, en-

quanto a segunda é prejudicial e antiquada. Aqui, defendemos que estas duas abordagens devem coexistir, inclusive na mesma prova: é perfeitamente possível avaliar os estudantes o tempo todo, para compreender como ocorrem suas aprendizagens, e fornecer a eles feedbacks sobre seus avanços e pontos de melhoria e, ao mesmo tempo, registrar com notas essas avaliações para, ao final de um período, embasado nestes resultados, julgar se o aluno aprendeu de maneira suficiente ou não.

Tipos de conteúdo

Voltando aos apontamentos do professor Antoni Zabala, encontramos a definição de quatro tipos de conhecimentos que podem ser avaliados em provas: os conhecimentos factuais, os conceituais, os procedimentais e os atitudinais.

Os conhecimentos factuais são as informações isoladas que um estudante pode guardar na sua memória – como o nome do escrivão da esquadra de Pedro Álvares Cabral, os títulos dos livros de Guimarães Rosa, ou a letra que representa o tungstênio na tabela periódica. Para Zabala, esses conhecimentos não têm qualquer valor se estiverem dissociados de conceitos, mas, ainda assim, à escola, às vezes, “interessa saber, neste momento, se [os alunos] são capazes de lembrar os nomes, os dados, os títulos, as datas, etc.” (p.203). Se retomarmos as categorias cognitivas de Bloom, veremos que os conhecimentos factuais são aqueles ligados à habilidade de lembrança, e nada mais.

O segundo tipo, o conhecimento conceitual, implica que o aluno consiga explicar um conceito – aquilo que a taxonomia de Bloom descreveria como a habilidade de “entender”. Zabala faz a ressalva de que avaliar os conhecimentos conceituais não é simples como os factuais (que estão certos ou errados): “Teremos que falar de graus ou níveis de profundidade e compreensão, algo que implica a necessidade de propor atividades em que os alunos possam demonstrar que entenderam, assim como sua capacidade para utilizar convenientemente os conceitos aprendidos”. (p.204).

Já o terceiro tipo – o conhecimento procedimental – pode ser compreendido como saber o que fazer com conceitos e fatos. Voltando à taxonomia de Bloom, a relação é com as habilidades de “aplicar”, “analisar”, “avaliar” e “criar”. Para Zabala, “O que define sua aprendizagem não é o conhecimento que se tem dele, mas o domínio ao transferi-lo

para a prática” (p.207) – e isto também é delicado. Enquanto alguns procedimentos são realizados no papel – como cálculos ou redações – outros só podem ser vistos na prática, como uma titulação ácido-base, em que de nada adiantaria se o aluno puder descrever o procedimento, se não for capaz de realizá-lo em laboratório.

Por fim, o conteúdo atitudinal quebra a sequência, por se debruçar sobre temas imateriais: a dedicação do aluno em aprender, sua organização, sua proatividade, sua autonomia, sua maturidade, seu respeito com os colegas e professores, sua forma de dar opiniões etc. Zabala define que “a fonte de informação para conhecer os avanços nas aprendizagens de conteúdos atitudinais será a observação sistemática de opiniões e das atuações nas atividades grupais, nos debates das assembleias, nas manifestações dentro e fora da aula, nas visitas, passeios e excursões, na distribuição das tarefas e responsabilidades, durante o recreio, nas atividades esportivas, etc” (p.209). Aqui, não há paralelo a ser feito com a taxonomia de Bloom, e também somos levados a pensar que os conteúdos atitudinais se relacionam aos anteriores de maneira indireta, e nem sempre precisa: ao adotar atitudes desejáveis na escola, supõe-se que o aluno aprenda mais, ainda que não haja garantias de que a adoção ou não de determinadas posturas seja determinante para aprendizagem – basta pensarmos que há estudantes que passam as aulas fazendo anotações e cópias da lousa e, mesmo assim, não alcançam bons resultados, enquanto outros tiram boas notas, mesmo com cadernos vazios.

O zero, o dez e a média

Considerando o equilíbrio necessário entre avaliação formativa e somativa, a melhor proposta seria oferecer aos alunos relatórios detalhados sobre seus avanços e problemas, após cada avaliação e, de maneira objetiva, uma nota em uma escala. Não temos tradição em produzir tais relatórios, mas a escala é conhecida: ela vai de zero (completo desconhecimento) até dez (completo domínio), com um valor intermediário, conhecido como média.

Talvez na escola onde você trabalhe a média seja 5, 6 ou 7, mas estes números não têm valor nenhum, e são configurados a partir da noção de desempenho mínimo. Seguindo a trajetória do planejamento, o professor e a escola determinam qual habilidade os alunos devem dominar; depois, quais ações devem acontecer na sala de aula para que

haja o domínio dessa habilidade; neste ponto, também se determina o desempenho mínimo, ou seja, quanto dessa habilidade será considerado minimamente satisfatório e sinônimo de ensino e aprendizagem bem-sucedidas. É a este mínimo que damos o nome de média e, depois, atribuímos a ele um valor entre 0 e 10.

Atingir a média, portanto, é atingir o mínimo – e ficar abaixo dela é sequer ter conseguido isso. O que nos leva a pensar nos extremos: o que significa tirar 0 ou 10? O primeiro ponto é a completa ausência da habilidade. Uma prova com zero é aquela em que o aluno não demonstra sequer um resquício do que deveria ter aprendido. Já a prova com 10 é a habilidade dominada por completo para o nível de exigência estabelecido, sem nenhuma falha. Se damos uma nota 10 a uma prova com problemas, perdemos a noção da excelência – como se houvesse duas notas 10, mas uma fosse mais 10 do que a outra.

A transformação de qualquer diagnóstico que tenhamos sobre nossos alunos em um valor numérico é uma simplificação grosseira – e, ao mesmo tempo, necessária. De modo a conseguir analisar o desempenho do aluno ao longo do tempo, ou de comparar sua aprendizagem com colegas de sala, os números serão infinitamente mais práticos do que descrições textuais, o que não significa que o docente não possa manter os dois registros. Contudo, a maneira como os números representarão visões sobre o aluno irá mudar largamente de um docente para o outro, por conta das decisões sobre o que é o mínimo necessário, e também porque cada docente terá escolhas pessoais na hora de montar sua “tabela de conversão”. Esta diferença está na base de todos os conflitos de conselhos de classe na hora de decidir se um aluno deve ser aprovado ou não, já que os números de notas comparados ali são formados por critérios díspares e, portanto, não poderiam ser colocados lado a lado.

Como a escola, mesmo com o avanço da noção de avaliação formativa, não vai abandonar o sistema de notas, o melhor que os professores podem fazer é conseguir torná-lo objetivo e criterioso, para que, como sistema de classificação, ele possa ser válido e estável. O oposto disso, porém, acontece com frequência, especialmente quando o professor adota as práticas que descreveremos a seguir.

Avaliação e confusão

Ao adotar critérios estáveis apenas de maneira parcial, o docente termina o ano letivo com uma quantidade enorme de notas e parâmetros que não pode ser comparada nem reunida – mas, mesmo assim, usa desses valores para tomar decisões sobre a vida do aluno ou, ao menos, afirmar o que pensa sobre o seu aprendizado. A seguir, listamos cinco problemas comuns que acontecem com as avaliações.

#1 misturar avaliação de conteúdos diferentes na mesma nota

Antoni Zabala, ao tratar das avaliações, defende que “na análise e avaliação das aprendizagens é indispensável diferenciar os conteúdos que são de natureza diferente e não situá-los num mesmo indicador. Não podemos resolver a valoração de um aluno numa determina área com um único dado que se refira às aprendizagens de conteúdos conceituais, procedimentais e atitudinais ao mesmo tempo” (p.214).

É possível encontrar maneiras de equilibrar, na mesma prova, a análise de conteúdos factuais, conceituais e procedimentais, de forma que a nota reflita quais o estudante consegue dominar – como em uma sequência de perguntas que parte do factual e chega ao procedimental, com notas de maior peso conforme a complexidade aumenta. Contudo, quando os conteúdos atitudinais se tornam parte da nota, aumentando-a ou diminuindo-a, a avaliação passou a contemplar duas esferas muito diferentes: a capacidade cognitiva e a postura do estudante. Isso acontece, por exemplo, quando o estudante tem uma dedução na pontuação por entregar um trabalho fora do prazo, ou quando a nota é composta parte pela avaliação, e parte pela soma de vistos no caderno por copiar a lousa. Em um cenário assim, dois estudantes poderiam atingir nota 6 ao final do ano letivo, sendo que o primeiro dominou os conteúdos e habilidades de modo parcial, e o segundo de maneira ainda mais frágil, mas, por ser diligente e cuidadoso, recebeu pontos pela sua postura em aula.

O incentivo em forma de pontuação para práticas valorizadas pela escola - chegar no horário, manter um caderno de anotações organizado, fazer perguntas, entregar trabalhos dentro do prazo etc – pode parecer uma proposta inofensiva, mas não é. Se damos

pontos para algumas atitudes dos estudantes, estamos deixando de fora dessa oportunidade aqueles alunos que não sentem a necessidade de adotar tais práticas para seu aprendizado. Há uma quantidade razoável de estudantes em sala de aula que aprende ouvindo o professor, sem precisar traçar uma linha de anotação.

Por fim, há mais um problema, que acerta em cheio a natureza do conteúdo atitudinal: é impossível avaliá-lo de maneira objetiva, então sobra aos professores apenas basear a nota na diferença entre fez / não fez. Como avaliação, este formato é insatisfatório, porque agrupa respostas muito diferentes dentro da mesma notação. Em resumo, o melhor a se fazer é não pontuar estes conteúdos, e tentar compreender, em que medida, eles favoreceram ou prejudicaram os estudantes partindo do desempenho deles nos outros conteúdos.

#2 comparar notas de tipos diferentes de avaliação

Instrumentos de avaliação diferentes – como provas com consulta, ou em dupla, ou teste, ou escrita – não são apenas variações da mesma tarefa, mas ferramentas que avaliam habilidades distintas. Quando os resultados parciais de um estudante (as notas de bimestre ou trimestre) são construídas com tipos diferentes de provas, pouco pode ser falado sobre a evolução do aluno, já que estamos comparando itens de natureza diferente. Se um estudante tira 5 ao escrever uma redação, e 8 ao responder uma prova teste, no mês seguinte, é incorreto afirmar que houve qualquer avanço no desempenho dele.

Para retomar a analogia com a medicina, se um paciente dá entrada no hospital depois de uma queda e tem uma fratura no braço constatada via radiografia, a cicatrização do osso será atestada, dali algumas semanas, com uma nova radiografia, e não com um exame de vista. O rigor da avaliação passa, portanto, pela manutenção das condições de avaliação, que nos permitirá dizer que a mudança no quadro não aconteceu por fatores externos, mas por um avanço na característica que importa.

#3 avaliar usando questões de vestibular

Esse é um expediente mais comum no ensino médio, porque resolve dois problemas de uma só vez: proporciona uma avaliação de qualidade, porque as questões vêm de

fonte confiável, e submete os estudantes a um treino para os processos seletivos. Entretanto, para a avaliação de aprendizagem, esta ideia é controversa. Quando incorporamos questões-teste de vestibular a uma de nossas provas, estamos adicionando a elas uma pergunta produzida com um propósito diferente das demais. Enquanto as questões originais são pensadas para estabelecer um diagnóstico do que nossos alunos aprenderam, a lógica das questões de vestibular é classificar estudantes e fazer com que eles tenham que competir entre si pelo maior número de acertos, o que inclui formular questões com distorções propositais, vocabulário complicador e, às vezes, o envolvimento de outros temas, ainda não trabalhados na sua aula. A questão de vestibular é excelente como desafio e deve constar do trabalho diário em sala, mas como instrumento de avaliação não é recomendável, pois não foi pensada para sua aula, e nem para oferecer informações de aprendizagem.

#4 Avaliar uma habilidade diferente da trabalhada em aula

Já mencionamos este problema no texto sobre os tipos de aula: pode acontecer que uma determinada habilidade tenha sido trabalhada em aula mas, na hora da avaliação, o professor exige outra habilidade para aquele mesmo conteúdo, mais sofisticada. Isso acontece quando a aula de matemática se concentra na execução de algoritmos, mas a prova está repleta de situações-problema, que vão exigir a transformação do texto em contas; ou quando aulas trabalham exclusivamente com leitura e interpretação de texto, mas as avaliações versam sobre análise e produção. A consequência dessa discrepância entre aula e avaliação serão resultados baixos dos alunos e um sentimento de incompetência deles e do professor, que terá em mãos uma avaliação sem qualquer valor, pois testa algo que jamais foi trabalhado.

#5 Mudar de instrumento de avaliação porque seus alunos têm notas baixas nele

Após uma rodada de resultados ruins, o professor decide trocar o tipo de prova, convencido de que seus alunos não se dão bem com aquele instrumento e, por isso, precisam fazer algo mais simples, que lhes permitirá resultados melhores. Esta não é uma ação desejável. Antes de trocar de tipo de prova por causa de resultados ruins, é fundamental reconhecer que eles são a consequência de um cenário com problemas, e não a

causa. Quando a prova é substituída por outra, as habilidades avaliadas também mudam, e isso pode acabar favorecendo os estudantes a melhorarem seus resultados – sem que, contudo, haja progresso nas suas aprendizagens.

É preciso refletir que a mudança no tipo de prova pode ser feita, desde que não prejudique nossas intenções de diagnóstico. A troca mais sugerida nas escolas é oferecer avaliações orais a estudantes com problemas de se expressar de maneira escrita – o que é válido, se a proposta é observar e analisar habilidades de lembrança e entendimento do aluno, ou seus conhecimentos factuais e conceituais. Se produzir respostas escritas é parte da habilidade em questão – como produzir um texto de maneira organizada e coesa – então nenhuma outra avaliação poderá dar as mesmas impressões e, por isso, não serve. Quando algo assim acontece, acabamos por driblar os problemas dos estudantes, e criamos um cenário falso de aprendizagem. A máxima aqui é de que as provas não devem ficar menos exigentes – os alunos é que devem ficar mais capacitados.

A prova dissertativa

Modelo sinônimo de prova, a avaliação dissertativa (ou aberta) é a mais versátil de todas, porque permite que diferentes tipos de conteúdo e habilidades sejam observados. Ao propor que o aluno escreva as respostas, ela tanto pode centrar as perguntas em conteúdos factuais e na habilidade de lembrança, como pedir que o estudante faça análises de materiais apresentados.

Ao escrever uma prova dissertativa, é necessário tomar alguns cuidados para garantir que o diagnóstico que ela tem a oferecer seja de qualidade. O primeiro é não utilizar nela questões que não são de natureza dissertativa, como testes, questões de verdadeiro ou falso e questões de associação. Estes tipos de pergunta não são recomendados porque suas respostas são binárias – ou estão corretas ou incorretas – e não deixam nenhuma margem para análise do professor sobre o raciocínio e reflexão do estudante.

Outro ponto é uma mistura comum de habilidades e conteúdo. Dificilmente, uma prova dissertativa é pura, no sentido de propor questões do mesmo tipo. Com vários exercícios, estes certamente serão variados: alguns serão conceituais, outros procedi-

mentais, outros factuais, e envolverão mais de uma habilidade. Se todos forem pontuados da mesma maneira – suponhamos, com um ponto para cada um – ao final, haverá vários alunos com as mesmas notas, mas características completamente diferentes, tornando a notação numérica um dado sem validade. Uma alternativa para isso – ainda que nem sempre ao alcance – é construir uma avaliação em que as habilidades e conteúdos se apresentem de forma crescente, e que questões mais complexas só possam ser respondidas por quem é capaz de solucionar as mais simples. Se esta configuração for possível, o valor de cada questão deve ser inversamente proporcional a sua dificuldade: questões mais simples valem mais, e questões mais complicadas valem menos. Com esse quadro, o professor fica mais próximo de garantir que um aluno tirou 5 porque conseguiu apenas responder questões factuais que envolvem memória, enquanto o estudante que tirou 9 apenas não foi capaz de acertar uma questão de análise e que envolvia um procedimento de mais etapas.

A correção da prova dissertativa pode ser mais ou menos demorada – quanto mais questões conceituais e procedimentais, maior o trabalho do professor. Este tipo de questão – e as habilidades envolvidas nela – também vão se mostrar um desafio de estabilidade, pois é preciso que todas sejam corrigidas com o mesmo parâmetro. Pode acontecer que algumas respostas sejam consideradas incorretas, em determinado ponto da correção, e respostas iguais, mais adiante, já recebam pontos do professor. A dificuldade em manter os mesmos critérios pede uma abordagem mais paciente e reflexiva, em que o professor precisa recordar, várias vezes, que leu uma resposta semelhante antes, mas lhe conferiu uma avaliação diferente – e, por isso, precisa mudá-la, para manter a uniformidade da correção.

Ainda falando em correção, outra necessidade das provas dissertativas – e de todas as outras que envolverem apreciação das respostas pelo professor – é corrigi-las sem saber quem é o estudante autor das respostas. Nossa visão sobre os alunos irá tornar a correção enviesada: podemos ser mais complacentes com alunos que temos em boa conta, ou mais rigorosos com alunos indisciplinados, ou desconfiarmos da lisura de uma resposta correta vinda de um aluno com desempenho historicamente baixo. Ao contrário, se não lermos os nomes no topo das folhas de prova, podemos oferecer uma correção bem mais justa aos alunos e eficiente para nossa compreensão da aprendizagem deles.

Prova objetiva (ou a prova de testes)

A prova de testes – em que o estudante apenas assinala a resposta correta em um conjunto de opções – é comumente aplicada em exames de larga escala, como os vestibulares. Isso acontece justamente por causa do seu formato: sua correção é muito mais rápida do que a da prova dissertativa, podendo, inclusive, ser automatizada, e a avaliação é igualitária e livre da subjetividade do professor, não sendo necessário ponderar sobre as respostas – ou elas estão corretas ou não estão. Por isso, entre todos os modelos de avaliação, é o que melhor permite a comparação entre os resultados dos estudantes, porque há a garantia de que todas foram corrigidas do mesmo jeito.

Porém, as vantagens param por aí, e há uma carga considerável de desvantagens na sua aplicação. A primeira é a chance de acerto ao acaso ou acerto por fraude. Dado que o objetivo da avaliação é conseguir um retrato fiel do aluno, qualquer interferência ou informação falsa é prejudicial para o processo de avaliação – e nada pode ser pior para o julgamento do professor do que um relatório de que os estudantes têm habilidades maiores do que realmente têm.

O segundo problema é a natureza da avaliação em si. As questões-teste são de inspiração behaviorista, vertente pedagógica preocupada com resultados, e não com o processo de obtenção deles. Quando um estudante acerta ou erra uma questão teste, esta é a única informação à disposição do professor, e mesmo ela não tem qualquer valor diagnóstico, porque erros e acertos em testes não significam, exatamente, que o estudante desconheça ou conheça o que se pediu a ele.

Questões-teste funcionam com conteúdo factual e a habilidade de lembrança. Quando tentamos produzir perguntas de múltipla escolha para os conteúdos conceituais ou procedimentais, e que exijam do aluno aplicação e análise, acabamos por criar um tipo único de pergunta, que nem sempre é bem-sucedida. Questões de interpretação de texto, por exemplo, não pedem ao aluno que explique uma passagem do texto, mas que julgue qual das explicações oferecidas está correta. Esta é uma habilidade diferente, e de execução menos certa, porque as alternativas acabam escritas quase como charadas a serem decifradas. Já questões que envolvem aplicação de procedimentos – como na Ma-

temática ou na Física – oferecem como alternativas apenas os resultados e, como mencionado antes, o processo de obtenção deles não passa, necessariamente, pelo procedimento correto – o aluno pode testar as respostas, ou simplesmente realizar ações incorretas que, por acaso, chegam ao mesmo lugar.

Em resumo, as provas objetivas têm limitações severas no que diz respeito à avaliação dos estudantes, mas suas vantagens operacionais fazem com que elas continuem o centro de muitos processos seletivos. Não à toa, os vestibulares das principais universidades públicas do país apenas aplicam provas dissertativas quando há um número reduzido de candidatos, nas fases finais.

Prova com consulta

A prova com consulta é uma variação da prova dissertativa, que faz uma mudança de habilidades: sai a capacidade de lembrança, e entram as de aplicação, análise e avaliação. Além disso, este instrumento pode deixar de lado as questões factuais para se concentrar no uso de informações, uma vez que tudo está à disposição do aluno. É possível aplicá-la oferecendo o material de consulta – como o texto base ou um arcabouço de fórmulas – ou permitindo que o aluno use as próprias anotações produzidas ao longo das aulas. Neste cenário, não se garante mais que todos partam das mesmas condições – visto que os alunos farão anotações diferentes – mas há o apelo para que características desejáveis nos estudantes, como a anotação e a organização, façam a diferença no resultado deles.

Prova oral

A prova oral também é uma variação da prova dissertativa, mas com implicações grandes por causa da mudança do suporte da prova. Os objetivos da prova continuam os mesmos: é possível checar a capacidade de memória do estudante, de análise, de criação, e conteúdos factuais, conceituais ou procedimentais. Contudo, como o formato é de um diálogo entre professor e aluno, este instrumento propicia que haja uma modulação da prova enquanto ela acontece: o professor pode oferecer auxílio se sentir que o estudante precisa de algo para seguir adiante (como uma informação) e também alterar

o direcionamento da prova, com questões mais fáceis ou mais difíceis, à medida que percebe o que pode explorar com o aluno.

Também há outra diferença em relação à prova escrita: aqui, interessa muito mais o conteúdo das respostas do aluno do que a organização delas. Além disso, como o professor não pode rever as respostas – o que é possível na prova escrita – a nota vem de anotações ou é dada “a quente”, assim que a prova acaba, o que interfere na isonomia de tratamento. Por fim, há um complicador de procedimento: provas orais são muito extensas e cansativas para o professor, e precisam de muito tempo para aplicação – uma conta rápida que considere 5 minutos por aluno (incluindo o tempo de prova e o tempo de troca entre eles), para uma turma de 30 alunos, pede, no mínimo, 2 horas e meia de realização.

É comum se sugerir que, quando um estudante apresenta dificuldades com provas escritas, ele seja submetido a uma prova oral. Este é um expediente com validade, mas limitado. Se a intenção é descobrir se o suporte escrito da prova impedia o aluno de mostrar aquilo que sabia, então a prova oral funcionará como bom diagnóstico. Por outro lado, se a produção textual da prova é relevante – a maneira como as respostas são apresentadas e organizadas – então a substituição por uma prova oral acaba sendo um dribble nos problemas, que continuam lá. Neste cenário, a prova oral deve ser apenas um passo intermediário, já que a prova escrita é a que dará o diagnóstico mais preciso sobre o que se deseja saber.

Provas em duplas, trios ou grupos

Este não é exatamente um instrumento, mas uma forma de aplicação – é possível ter provas abertas, objetivas, com consulta ou orais em duplas, trios ou grupos maiores. Quando fazemos isso, adicionamos o componente da negociação e da reflexão a todas estas provas, supondo que os alunos participantes terão que discutir as respostas entre si. Para o professor, há também uma vantagem secundária: diminui-se pela metade, ou até mais, a quantidade de provas a serem corrigidas.

Os pontos positivos, porém, não vão além disso. Há algo inescapável nesta correção: é impossível determinar qual parte da dupla (ou trio, ou grupo) respondeu qual parte

da prova, o que invalida o propósito da avaliação. Ao final dela, qualquer diagnóstico oferecido não pode ser generalizado para ambos os estudantes e, portanto, não tem qualquer valor. Em resumo, esta é uma dinâmica de aula muito bem-vinda para o dia a dia escolar, mas terrível como avaliação.

Seminário

Evento constante nas nossas formações universitárias, o seminário é uma forma de avaliação que mobiliza o trabalho em grupo, a identificação e o recorte de informações, a decisão sobre como apresentá-las e a clareza da apresentação, com uso de recursos visuais. O aluno envolvido com um seminário passa pelos mesmos passos que nós, professores, passamos para produzir uma aula.

Os seminários esbarram na avaliação limitada que se pode fazer deles: ainda que haja um trabalho de preparo complexo até a apresentação, ele é avaliado exclusivamente por ela – e, novamente, ou se confere uma nota para o grupo todo, sem clareza sobre quem fez o que, ou se tenta dar notas individuais, o que a princípio é uma atitude satisfatória para sanar o problema anterior, mas que também não se sustenta, porque a dinâmica sobre “quem fala o que” é decidida nos bastidores do trabalho, e alguns alunos sempre acabarão com menos espaço do que os outros.

O maior problema dos seminários, contudo, é sua utilização como ferramenta de ensino e de avaliação ao mesmo tempo. Não é incomum – ao menos no ensino superior – que o professor da disciplina traga parte do conteúdo da aula em forma de seminários apresentados pelos alunos, o que é uma estratégia bastante questionável, porque a qualidade da aula será desigual. Portanto, assim como acontece com as atividades em dupla ou trio, o seminário é uma excelente dinâmica de sala de aula, mas bastante problemática para ser utilizada como avaliação.

Trabalho para casa

Há uma ampla variedade de trabalhos para serem feitos em casa – lista de exercício, resumos, redações etc – e que irão mobilizar a pesquisa, recorte de informações e o trabalho autônomo, uma vez que são tarefas executadas fora da escola, como uma continuação da sala de aula.

Entretanto, este é o tipo de prova mais propenso a fraudes e, portanto, tem pouco valor avaliativo. Quando dois alunos apresentam o mesmo trabalho – ou quando há uma cópia geral na turma – a avaliação fica prejudicada e o professor se vê na posição desconfortável de identificar qual é o trabalho original, na intenção de anular as cópias, o que nem sempre é possível. Assim, esta é uma forma de avaliação não recomendada para conferir notas aos alunos, mas que poderia ser aplicada sem o uso das notas, ainda que seja pouco viável, em um cenário de muito trabalho do professor, usar o tempo para corrigir tarefa e dar devolutiva para os alunos, sem que isso também contribua para a totalização das notas – tornando necessário montar mais avaliações e fazer mais correções.

Avaliação seriada

A avaliação seriada não é um instrumento de avaliação, mas uma forma de aplicá-los com os alunos. Tradicionalmente, as provas acontecem ao final de uma sequência de aulas, e cobrem todo o conteúdo trabalhado neste espaço de tempo. Na proposta de serialização, a prova é repartida e aplicada ao final de cada aula: por exemplo, ao invés de aulas distribuídas por cinco semanas, seguidas de uma prova, há cinco provas, menores, aplicadas ao final das cinco aulas, cada uma sobre o conteúdo trabalhado naquele dia.

Há algumas vantagens importantes neste modelo: a primeira é de que, ao aplicar uma prova no final da aula, corrobora-se com a intenção de que “toda aula vale alguma coisa” na cabeça dos alunos, que entendem que não podem ficar desatentos; do ponto de vista de avaliação, esse modelo consegue observar a aprendizagem assim que ela aconteceu, e, por isso, também descobrir se o desenvolvimento está saindo como foi planejado pelo professor, permitindo correções de rota mais ágeis. Há também outro efeito positivo: a tensão em fazer provas diminui consideravelmente, já que elas deixam de ser eventos especiais e passam a ser incorporadas nas aulas. Estas duas consequências colaboram para que a aula ganha um sentido de urgência, e também para que elas se configurem como eventos fechados, de começo, meio e fim.

Contudo, de todos os modelos apresentados aqui, a avaliação seriada talvez seja a mais exigente. Não há, a princípio, um aumento na quantidade de correção a ser realizada pelo professor – uma prova integral de cinco questões se torna cinco provas de uma

questão cada – mas a dinâmica de correção precisa ser bem mais rápida, já que, na aula seguinte, os alunos precisam conhecer seus resultados. Outro ponto tem relação com o planejamento e a execução da aula: se há a necessidade de aplicar uma avaliação ao final dela, estas provas já devem estar previstas e escritas durante a produção das aulas – nada de chegar ao final da aula e decidir o que perguntar – e a divisão do tempo entre a aula em si e a avaliação precisa ser seguida com muito rigor, para que a avaliação possa ser realizada com a quantidade de tempo adequada. Isso é especialmente problemático porque aulas são imprevisíveis e podem tomar rumos inesperados a partir das dúvidas dos estudantes.

Depois da prova

Ao receber uma prova de volta, o primeiro movimento dos estudantes é checar a nota obtida – uma consequência clara da cultura escolar de resultados e punições. Mesmo que o professor tenha chegado àquele valor equilibrando os diferentes tipos de conteúdo, e de maneira a caracterizar os estudantes a partir daqueles números, este mecanismo é opaco aos alunos, que possuem um termômetro nada explicativo de seus desempenhos.

Um trabalho cuidadoso de correção e feedback após as avaliações é completamente necessário, porque de nada vale ao estudante receber um resultado ruim se ele não conseguir entender as causas disso. Retornando aos conteúdos factuais, conceituais e procedimentais explicados por Zabala, a correção do primeiro tipo acontece sem sobressaltos, porque há somente uma resposta certa, e o estudante se recorda ou não da informação solicitada. As coisas ficam mais delicadas com os outros dois tipos. Os conteúdos procedimentais, por serem divisíveis em etapas menores, requerem atenção para que o aluno consiga identificar em que ponto do procedimento sua resposta tomou o caminho errado. Com os conteúdos conceituais, porém, a correção é de compreensão mais difícil para o estudante, porque não é simples demonstrar a ele porque sua explicação para um conceito é incompleta, ou apenas parcialmente equivocada. Quando a correção é feita de maneira geral, para a turma toda, é comum que o aluno se queixe de que “escreveu aquilo, mas o professor deu errado mesmo assim”. Este é um mal-entendido que precisa ser desfeito, sob pena de que os estudantes passem a achar que a prova é corrigida de

maneira arbitrária – e isso só pode acontecer se for possível estabelecer correções objetivas, que especifiquem quais temas, palavras-chave ou encadeamentos deveriam ter aparecido.

Para concluir

Os estudantes apenas poderão desenvolver uma nova visão sobre avaliação e compreender que ela é benéfica e necessária para suas aprendizagens quando os próprios docentes derem o primeiro passo, e oferecerem provas bem construídas que sirvam, em primeiro lugar, como instrumento de análise de aprendizagem. Esta, porém, não é uma tarefa fácil, porque, além de mudança de postura, demanda também uma compreensão mais profunda sobre como uma prova deve ser construída e analisada, considerando os diversos conhecimentos e habilidades. A cada prova aplicada pelo professor, há um mundo de parâmetros envolvidos, e o conhecimento deles é fundamental para que a avaliação seja eficiente no seu papel de obter informações sobre a aprendizagem dos estudantes.

Já decisão sobre que tipo de prova aplicar ressoa as opções de planejamento e dinâmicas de sala de aula: é preciso equilibrar os objetivos da ação com a viabilidade dela. Idealmente, os alunos deveriam ser avaliados diariamente, com feedbacks detalhados sobre o que apresentaram, mas a carga horária elevada e as turmas volumosas são um impedimento severo para isso. Assim, cabe ao professor conseguir, dentro das possibilidades, aplicar o máximo de avaliações possíveis e que, de fato, permitam que ele compreenda o que se passa com a aprendizagem de seus estudantes – e, para isso, como desejamos mostrar, há instrumentos mais eficazes do que outros.

Bibliografia

ANDERSON, L. W. et. al. **A taxonomy for learning, teaching and assessing: a revision of Bloom's Taxonomy of Educational Objectives**. Nova York: Addison Wesley Longman, 2001.

VILLAS BOAS, Benigna. **Avaliação das aprendizagens, para as aprendizagens e como aprendizagem**. Papirus Editora, 2022.

ZABALA, Antoni. **A prática educativa: como ensinar**. Artmed, 1998.