Trabalhando por dentro da caixa preta: avaliação para a aprendizagem na sala de aula¹

Paul Black*

Christine Harrison**

Clare Susan Lee***

Bethan Marshall****

Dylan Wiliam*****

*(Department of Education and Professional Studies, King's College London, Londres, Reino Unido)

** (Department of Education and Professional Studies, King's College London, Londres, Reino Unido)

*** (School of Education, Childhood, Youth, and Sports, Open University, Londres, Reino Unido)

**** (Department of Education and Professional Studies, King's College London, Londres, Reino Unido)

***** (Institute of Education, University College London, Londres, Reino Unido)

Resumo: Em seu amplamente divulgado artigo “Por dentro da caixa-preta”, Black e Wiliam demonstraram que melhorar a qualidade da avaliação formativa eleva o desempenho dos estudantes. Agora, eles e seus colegas reportam os resultados de um projeto que se seguiu ao primeiro com o intuito de ajudar professores a mudarem suas práticas e os estudantes a mudarem seu comportamento para que todos compartilhem a responsabilidade pelo processo de aprendizagem.

Palavras-chave: Avaliação formativa. Devolutivas. Ensino e aprendizagem.

¹ BLACK, Paul; HARRISON, Christine; LEE, Clare; MARSHAL, Bethan; WILIAM, Dylan. Working Inside the Black Box: assessment for learning in the classroom. Phi Delta Kappan, p. 9-21, set. 2004. Direitos autorais reservados para Phi Delta Kappan. Traduzido e publicado com autorização. Tradução e revisão técnica: Bárbara Barbosa Born.

Em 1998, o artigo que precedeu esse trabalho, “Por dentro da caixa preta”², apareceu nessa mesma revista (Phi Delta Kappan). Desde então aprendemos muito sobre as ações práticas que são necessárias para alcançar os propósitos expressos no subtítulo daquele artigo: “elevando os padrões de desempenho por meio da avaliação na sala de aula”.

Na primeira parte de “Por dentro da caixa preta”, respondemos a três questões. A primeira era: existe evidência de que melhorar a avaliação formativa eleva os padrões de desempenho? A resposta foi um inequívoco “sim”, conclusão essa baseada na revisão de evidências publicadas em mais de 250 artigos de pesquisadores de diferentes países³. Poucas iniciativas em educação tiveram um corpo de evidências tão robusto para apoiar seu argumento sobre como elevar os padrões de desempenho.

Essa resposta positiva nos levou naturalmente à segunda pergunta: existem indícios de que ainda há espaço para melhoria? Aqui, novamente, a evidência disponível deu uma resposta clara e positiva, apresentando um quadro detalhado que identificou três problemas centrais: 1) os métodos de avaliação que os professores utilizam não são efetivos para promover uma aprendizagem de qualidade; 2) as práticas de atribuição de notas tendem a enfatizar competição em vez de crescimento pessoal; 3) as devolutivas⁴ das avaliações normalmente têm um impacto negativo, particularmente entre os estudantes com baixo desempenho, que são levados a acreditar que eles não possuem as “capacidades” requeridas e que, portanto, são incapazes de aprender.

No entanto, para a terceira questão — existe evidência sobre como melhorar a avaliação formativa? —, a resposta não estava tão clara. Enquanto as evidências proviam diversas ideias para a melhoria, elas careciam de detalhes que possibilitariam aos professores implementar tais ideias nas salas de aula. Nosso argumento foi que os professores precisariam de uma variedade de exemplos concretos sobre a implementação dessas ideias.

² BLACK, Paul; WILLIAM, Dylan. Inside the Black Box: raising standards through classroom assessment. Phi Delta Kappan, out. 1998.

³ BLACK, P.; WILLIAM, D. Assessment and Classroom Learning. Assessment in Education, p. 81-94, mar. 1998.

⁴ No original, os autores utilizam o termo “feedback”, cuja tradução literal seria “retroalimentação”. Tem o sentido de alimentar os estudantes com comentários que os façam continuar aprendendo, daí a escolha pelo termo “devolutiva” em português. [N. do T.]

A JORNADA: APRENDENDO COM OS PROFESSORES

Desde 1998, planejamos e implementamos diversos programas nos quais grupos de professores da Inglaterra receberam apoio para desenvolver práticas inovadoras em suas salas de aula, partindo das ideias do artigo original. Enquanto esse esforço confirmou amplamente a proposta original, também adicionou uma série de novos achados que são simultaneamente práticos e autênticos. Por esse motivo, estamos agora confiantes de que podemos fazer recomendações fundamentadas para a melhoria das avaliações de sala de aula.

O PROJETO KMOFAP

Para conduzir o trabalho exploratório necessário, precisávamos da colaboração de um grupo de professores dispostos a correr alguns riscos e ter algum trabalho extra, bem como de um apoio seguro de suas respectivas escolas e redes de ensino. O financiamento para o projeto foi provido por meio da generosidade da Fundação Nuffield, e fomos afortunados em encontrar duas redes de ensino no sul da Inglaterra — Oxfordshire e Medway — cujos gestores compreenderam as questões que estávamos propondo e se mostraram dispostos a trabalhar conosco. Cada rede selecionou três escolas de anos finais do Ensino Fundamental e do Ensino Médio⁵: Oxfordshire escolheu três escolas mistas e Medway escolheu uma escola mista, uma escola de garotos e uma escola de garotas. Cada escola selecionou dois professores de ciências e dois de matemática. Discutimos os planos com os diretores de todas as escolas, e então convocamos a primeira reunião com os 24 professores. Assim, em janeiro de 1999 nasceu o Projeto de Avaliação Formativa Kings’s-Medway-Oxfordshire (Kings’s-Medway-Oxfordshire Formative Assessment Project – Kmofap).

Todos os detalhes do projeto podem ser encontrados em nosso livro *Assessment for Learning: Putting It into Practice*⁶ [em tradução livre, *Avaliação para a aprendizagem: colocando-a em prática*]. Para o propósito deste artigo, são os resultados do projeto que importam. Os achados apresentados aqui são baseados em observações e gravações das visitas às salas de aula pela equipe do King’s College, gravações dos encontros com o grupo de professores, entrevistas com professores e análise de seus registros escritos,

⁵ No original, “secondary schools”. [N. do T.]

⁶ BLACK, P.; HARRISON, C.; LEE, C.; MARSHAL, B.; WILIAM, D. *Assessment for Learning: Putting It into Practice*. Buckingham, RU: Open University Press, 2003.

e algumas discussões com grupos de estudantes. Inicialmente, trabalhamos com professores de matemática e ciências, porém recentemente nosso trabalho se expandiu para atender professores de língua inglesa das mesmas escolas e professores de outras disciplinas de outras escolas.

DIVULGANDO A PALAVRA

Ao longo do desenvolvimento do projeto, atendemos a inúmeros convites para falar para outros grupos de professores e seus formadores. Eles eram de diferentes áreas do conhecimento, tanto do Ensino Fundamental quanto do Ensino Médio. Adicionalmente, temos desenvolvido um trabalho sistemático com algumas escolas de anos iniciais do Ensino Fundamental⁷. Todas essas experiências têm nos dado a confiança de que nossos achados gerais terão valor para todos, embora alguns importantes detalhes talvez difiram entre grupos etários ou disciplinas distintas. Além disso, um grupo da Universidade de Stanford obteve um financiamento da Fundação Nacional de Ciências⁸ dos Estados Unidos para desenvolver um projeto similar em escolas californianas em parceria conosco. Esse financiamento possibilitou a ampliação do nosso trabalho. Também reconhecemos o apoio de indivíduos em diversas agências governamentais que participaram da direção do projeto, oferecendo orientação e conselhos, e que nos ajudaram a assegurar que a avaliação para a aprendizagem se tornasse um tema central da política educacional da Inglaterra e da Escócia (ver “Avaliação para a aprendizagem”, abaixo).

Avaliação para a aprendizagem

Avaliação para a aprendizagem é qualquer processo de avaliação no qual a prioridade tanto no seu planejamento quanto na sua implementação seja servir ao propósito de promover as aprendizagens dos estudantes. Nesse sentido, ela difere de avaliações cujo foco primário é servir aos propósitos de responsabilização, ranqueamento ou para certificar competências.

Uma atividade avaliativa pode ajudar os estudantes se ela prover informações que os próprios estudantes e os professores possam utilizar enquanto devolutivas para avaliarem a si próprios e uns aos outros, e que aja na modificação das atividades de ensino e de aprendizagem nas quais ambos estão engajados. Ela se torna uma “avaliação formativa” quando as evidências coletadas são efetivamente utilizadas para adaptar o trabalho do professor de modo a atender às necessidades dos estudantes.

⁷ No original, “primary schools”. [N. do T.]

⁸ National Science Foundation. [N. do T.]

OS GANHOS DE APRENDIZAGEM

A partir da revisão da literatura internacional que havíamos feito, estávamos convencidos de que a melhoria na avaliação formativa produziria ganhos nas aprendizagens dos estudantes, mesmo quando essas fossem expressas pelas medidas limitadas dos testes padronizados. Desde o início tínhamos clareza da importância de ter algum indicador sobre ganhos que poderiam ser alcançados em salas de aula reais e ao longo de um longo recorte temporal. Uma vez que cada professor nesse projeto estava livre para escolher a sala de aula com a qual gostaria de trabalhar essas ideias, discutimos individualmente com cada um que tipo de dado estava disponível em cada escola e desenhamos um “miniexperimento” para cada um dos professores.

Cada professor decidiu, assim, qual seria a medida de rendimento de sua turma. Para as turmas do 10º ano, essa medida normalmente era a nota que o aluno tirava no exame de conclusão do Ensino Médio que os alunos fazem aos 16 anos (a Certificação Geral da Educação Secundária⁹, em tradução livre, ou GCSE, na sigla em inglês). Para as turmas de 8º ano, essa medida era normalmente a nota ou nível alcançado no exame nacional que é administrado para todos os alunos com 14 anos. Para as outras turmas, uma variedade de medidas foi utilizada, incluindo notas nos testes de final de módulo e nos exames escolares de final de ano.

Para cada turma que estava no projeto os professores tiveram que encontrar uma turma de comparação. Em alguns casos era uma turma paralela para a qual o mesmo professor dera aula em anos anteriores (e, em um caso, no mesmo ano). Em outros casos, utilizamos turmas paralelas que tinham o mesmo professor ou um professor diferente, e quando isso não era possível, selecionamos uma turma não paralela ensinada pelo mesmo professor ou por algum outro professor. Quando as turmas de controle não eram estritamente paralelas às do projeto, controlamos por possíveis diferenças no desempenho pregresso dos alunos com base em medidas de “entrada”, tais como testes escolares de anos anteriores ou outras medidas de aptidão.

Essa abordagem significou que o tamanho da medida de melhoria foi mensurada de forma diferente para cada um dos professores. Por exemplo, uma turma de 10º ano do projeto pode ter um desempenho melhor que a turma de controle que equivalha à metade dos pontos possíveis no GCSE, enquanto outra turma de 8º ano pode ter um desempenho que seja 7% maior que o controle no teste de final de ano da escola. Para que pudéssemos agregar

⁹ General Certificate of Secondary Education. [N. do T.]

os resultados, portanto, adotamos uma estratégia de padronização bastante comum chamada “tamanho de efeito padronizado” (standardized effect size), calculado ao tirar a diferença entre os resultados do grupo experimental e de controle e dividir pelo número de desvios-padrão (que é uma medida de distribuição das notas do grupo).

Para os 19 professores para os quais nós temos os dados completos, o efeito médio foi de 0,3 desvio-padrão. Se essa mesma melhoria fosse produzida em toda a escola, ela teria o potencial de elevar uma escola do quartil mais baixo na performance nacional para muito acima da média. Com isso fica claro que, muito além de ter de escolher entre ensinar direito ou conseguir um bom desempenho nas provas, os professores podem efetivamente melhorar o desempenho dos estudantes ao trabalhar com as ideias que apresentamos aqui.

COMO A MUDANÇA PODE ACONTECER?

Organizamos nossos achados sobre o trabalho na sala de aula em torno de quatro eixos: questionamentos, devolutivas por meio de notas, autoavaliação e avaliação dos pares, e o uso formativo de instrumentos somativos. A maioria das citações diretas que utilizamos nas próximas páginas foi retirada das reflexões dos professores participantes. Os nomes dos professores e das escolas são pseudônimos, mantendo nossa política de anonimato.

QUESTIONAMENTOS

Muitos professores não planejam nem conduzem diálogos em sala de aula de forma que possam auxiliar as aprendizagens dos estudantes. A pesquisa tem demonstrado que após realizar diversas perguntas, os professores esperam menos de um segundo pela resposta e, se ninguém se habilita a responder, fazem outras perguntas ou respondem à pergunta eles mesmos¹⁰. A consequência de um “tempo de espera” tão pequeno é que o único tipo de pergunta que “funciona” é o daquelas que podem ser respondidas rapidamente, sem refletir — ou seja, questões que mobilizam apenas memorização de fatos. Consequentemente, o diálogo acontece apenas no nível superficial. Como um professor sugeriu:

¹⁰ ROWE, M. B. Wait time and rewards as instructional variables, their influence on language, logic, and fate control. *Journal of Research in Science Teaching*, v. 11, p. 81-94, 1974.

Eu estava insatisfeito com o estilo de perguntas e respostas fechadas no qual meu ensino pouco reflexivo havia se transformado, e como consequência me tornei preguiçoso com relação à aceitação das respostas corretas e algumas vezes era mesmo cúmplice da turma ao assegurar que nenhum de nós tivesse que trabalhar muito duro... Quando percebíamos que as perguntas e respostas não estavam fluindo suavemente, eu mudava a pergunta, respondia eu mesmo, ou apenas buscava as respostas dos alunos mais “brilhantes”. É possível que tenha havido circunstâncias (ainda existem?) nas quais um observador externo visse minha aula como um pequeno grupo de discussão rodeado por diversos espectadores sonolentos – James, Escola Two Bishops.

A chave para transformar essa situação é permitir um tempo maior de espera. Todavia, muitos professores acham isso bastante difícil, porque requer um rompimento com os hábitos que já estão estabelecidos. Uma vez que eles mudam, as expectativas dos estudantes são desafiadas:

Aumentar o tempo de espera após fazer perguntas mostrou-se incrivelmente desafiador em decorrência do meu desejo habitual de acrescentar alguma coisa imediatamente após fazer a pergunta original. Essa pausa após o questionamento era muitas vezes dolorosa. Me parecia antinatural ter esse “período morto”, mas perseverei. Ao dar mais tempo para pensar, os estudantes perceberam que eu estava esperando uma resposta mais elaborada. Agora, após vários meses de mudança no meu estilo de fazer perguntas, percebi que a maioria dos estudantes dá uma resposta e uma explicação (quando necessário) sem que eu precise dar um comando adicional. – Derek, Escola Century Island.

Outra professora sintetizou os efeitos gerais de seu esforço para melhorar o uso de diálogos baseados em perguntas e respostas na sala de aula da seguinte forma:

Questionamentos

Todo o meu estilo de ensino se tornou mais interativo. Em vez de mostrar como encontrar soluções, uma pergunta é feita e os alunos recebem um tempo para explorá-la conjuntamente. Minha turma de 7º ano que participa do projeto já está acostumada com essa forma de trabalhar. Eu também percebi que estou usando esse método cada vez mais com outros grupos.

Sem mãos

A menos que especificamente requisitados, os alunos sabem que não devem levantar as mãos para responder a uma pergunta. Todos sabem que devem estar preparados para responder a qualquer momento, ainda que seja um “eu não sei”.

Clima de apoio

Os alunos estão confortáveis para dar respostas incorretas. Eles sabem que elas podem ser tão valiosas quanto as corretas. Eles ficam felizes quando seus colegas os ajudam a explorar suas respostas incorretas na sequência. — Nancy, Escola Riverside.

Aumentar o tempo de espera pode ajudar mais estudantes a se engajar nas discussões e aumentar a profundidade de suas respostas. Outra forma de ampliar a participação é solicitar aos estudantes que façam uma tempestade de ideias, em duplas, por 2 ou 3 minutos, antes que o professor demande contribuições de toda a sala. De um modo geral, a consequência de tal mudança é que os professores aprendem mais sobre os conhecimentos prévios dos alunos, bem como potenciais aspectos que eles desconhecem ou compreenderam incorretamente sobre aquele tópico, de tal modo que as estratégias seguintes se direcionam para suprir as necessidades reais dos alunos.

Explorar tais mudanças significa se distanciar da prática rotineira de perguntas focadas em conhecimentos factuais, reposicionando a atenção quanto à qualidade e as diferentes funções que perguntas possuem para a aprendizagem. Considere, por exemplo, as “grandes questões”: uma pergunta de resposta aberta ou uma solução de problemas que pode definir o cenário para uma aula, evocar uma discussão mais ampla ou ainda fornecer a base para uma discussão em pequenos grupos. Se os professores esperam que essa estratégia seja produtiva, tanto as respostas que essa tarefa podem evocar quanto as formas de dar sequência a essas respostas devem ser antecipadas. A colaboração entre os professores para trocar ideias e experiências sobre a elaboração de boas questões é muito valiosa. As próprias questões, dessa forma, se tornam uma parte significativa do ensino, quando a atenção está voltada a como elas podem ser construídas e utilizadas para desenvolver a aprendizagem dos alunos. Isto é o que uma professora pensa a respeito:

Escolhi uma turma de 8º ano, com desempenho intermediário, e realmente comecei a pensar sobre que tipo de perguntas eu estava fazendo — se elas levavam a respostas instantâneas de uma palavra —, o que essas perguntas estavam testando — se conhecimento ou compreensão —, se eu estava dando à turma tempo suficiente para responder à questão, se estava aceitando rapidamente a resposta correta, se estava solicitando às meninas que explicassem seu raciocínio, como eu estava lidando com as respostas incorretas. Quando realmente parei para pensar, percebi que eu poderia fazer uma grande diferença na aprendizagem das garotas ao utilizar todas as suas respostas para ditar o ritmo e o conteúdo das aulas. — Gwen, Escola Waterford.

Questionamento efetivo é também um importante aspecto das intervenções improvisadas que os professores conduzem uma vez que os alunos iniciam suas atividades. Elaborar questões simples, tais como “por que você pensa isto?” ou “como você expressaria essa ideia?” pode se tornar parte da dinâmica de interação da aula e oferecer oportunidades valiosas para ampliar o pensamento dos alunos por meio de devolutivas imediatas sobre seu trabalho.

De um modo geral, as principais sugestões para a ação nesse eixo que emergiram das experiências dos professores são:

- Mais esforço deve ser colocado na elaboração de perguntas que valem a pena ser realizadas, isto é, questões que explorem aspectos que são críticos para o desenvolvimento da compreensão pelos estudantes.
- O tempo de espera deve ser aumentado em vários segundos para assegurar que os estudantes tenham tempo para pensar, e a expectativa deve ser que todos tenham uma resposta e possam contribuir para a discussão. Com isso, todas as respostas, certas ou erradas, podem ser utilizadas para desenvolver a compreensão. O objetivo é que ocorra um aprendizado refletido, e não que se chegue à resposta correta na primeira tentativa.
- Atividades de seguimento devem ser ricas, de tal modo que possam criar oportunidades para ampliar a compreensão dos estudantes.

De forma simples, o único sentido em fazer perguntas é levantar questões sobre as quais os professores precisam de informações ou sobre as quais os alunos precisam refletir. Quando essas mudanças foram realizadas, a experiência demonstra que os estudantes se tornaram participantes mais ativos e perceberam que a aprendizagem depende menos da capacidade que eles têm de dar a resposta correta e mais da prontidão em expressar e discutir seus próprios entendimentos. O papel dos professores também se modifica, passando de apresentadores de conteúdos para líderes na exploração e desenvolvimento de ideias nas quais os estudantes estão envolvidos.

DEVOLUTIVAS POR MEIO DE NOTAS

Ao dar devolutivas aos trabalhos orais e escritos dos estudantes, o aspecto mais crítico é a natureza dos comentários, e não a quantidade. Alguns experimentos demonstraram que, ao passo que a aprendizagem pode ser melhorada por meio de devolutivas na forma de comentários, a atribuição de notas (tanto numéricas quanto na forma de conceitos) tem um efeito

negativo, porque os estudantes ignoram os comentários quando as notas são dadas em conjunto¹¹. Esses resultados no geral surpreendem os professores, mas aqueles que abandonaram a atribuição de notas descobriram que suas experiências estavam de acordo com os achados da pesquisa: os estudantes se engajaram mais efetivamente na melhoria de seus trabalhos.

Muitos professores se preocupam com os efeitos de entregar os trabalhos para seus alunos apenas com os comentários, mas não com as notas. Potencialmente pode haver conflitos com as políticas de avaliação das escolas:

Meu sistema de correção se desenvolveu de um que utilizava comentários acompanhados dos objetivos de aprendizagem e da nota, que é a política da escola, para comentários e objetivos de aprendizagem apenas. Os estudantes trabalham no alcance dos objetivos e fazem correções mais produtivamente quando as notas não são dadas. Clare (Lee) observou em diversas ocasiões quão pouco tempo os alunos gastavam lendo os meus comentários quando havia notas acompanhando. Minha rotina agora na sala do projeto é: a) não dar notas, apenas comentários; b) oferecer comentários que destacam o que foi feito corretamente e o que ainda necessita de trabalho; e c) esclarecer aos alunos qual é o mínimo de trabalho necessário e esperado para que completem a tarefa da próxima vez que eu olhar os seus cadernos. — Nancy, Escola Riverside.

Os medos iniciais sobre como os estudantes reagiriam mostraram-se injustificados, e nem os pais nem os inspetores das escolas reagiram de forma adversa. De fato, a oferta de comentários ajudava os pais dos alunos a focar os problemas de aprendizagem em vez de tentar interpretar a nota que havia sido dada. Hoje acreditamos que o esforço que muitos professores devotam para dar notas nas lições de casa talvez seja mal direcionado. Uma nota, seja ela numérica ou na forma de conceito, não diz ao aluno como melhorar o seu trabalho, de modo que a oportunidade de melhoria é perdida.

Um comprometimento com a melhoria dos comentários requer mais trabalho em um primeiro momento, dado que os professores precisam atentar para a qualidade das devolutivas que escrevem em cada trabalho de seus alunos. A colaboração entre professores ao compartilharem exemplos de comentários efetivos pode ser de grande valia, e a experiência levará à fluência. Todavia, há mais coisas envolvidas, porque os comentários somente serão devolutivas úteis se os estudantes os utilizarem para guiar seus trabalhos futuros, de modo que outros procedimentos são necessários.

¹¹ BUTLER, R. Enhancing and undermining intrinsic motivation: the effects of task-involving and ego-involving evaluation on interest and performance. *British Journal of Educational Psychology*, v. 58, p. 1-14, 1998.

Após a primeira formação em serviço, eu estava ansioso para tentar uma forma diferente de corrigir as tarefas e dar devolutivas mais construtivas para meus alunos. Estava ansioso por tentar e encontrar um método mais fácil de monitorar as respostas dos estudantes aos meus comentários sem que eu tivesse que rastrear seus livros cada vez para descobrir se eles haviam reagido aos meus comentários. Implementei uma folha de comentários no verso dos livros dos meus alunos de 8º ano. Trata-se de um papel A4, em que a metade esquerda contém um espaço para os meus comentários e a metade direita possui um espaço para os estudantes se referenciarem à página do livro onde eu posso encontrar a evidência de que eles fizeram o trabalho esperado. Os comentários se tornaram mais significativos ao longo do tempo, e continuo gastando uma hora para corrigir os livros. — Sian, Escola Estadual Cornbury.

Encontramos uma variedade de formas de acomodar a nova ênfase nos comentários. Alguns professores deixaram de atribuir notas completamente, outros colocavam as notas apenas em seus diários de classe, mas não nas atividades dos alunos, e outros ainda apenas davam a nota após o aluno responder aos seus comentários. Alguns professores gastavam mais tempo em algumas tarefas específicas para assegurar que elas recebessem uma devolutiva mais completa e, para ganhar tempo para essas tarefas, ou não olhavam diretamente todos os trabalhos, ou apenas corrigiam um terço das lições de casa que os alunos faziam em seus livros por semana, ou envolviam os próprios estudantes na checagem de suas tarefas.

Um método particularmente valioso é dedicar algum tempo em aula para que os alunos possam reescrever algumas peças selecionadas de seu trabalho, de modo que a ênfase possa ser colocada na devolutiva para as melhorias dentro de um ambiente de apoio. Essa prática pode mudar as expectativas dos alunos quanto aos propósitos do trabalho em sala e para casa.

Enquanto tentavam criar comentários que se constituíssem em devolutivas úteis, muitos dos professores participantes no projeto perceberam que reavaliaram as próprias tarefas que haviam demandado aos estudantes. Eles perceberam que algumas atividades eram úteis para revelar compreensões e incompreensões dos estudantes, ao passo que outras estavam focadas apenas em informações factuais. Com isso, algumas atividades foram eliminadas, outras modificadas, e novas e melhores tarefas emergiram ativamente desse processo.

De um modo geral, as principais ideias para a melhoria de devolutivas podem ser sintetizadas como se segue:

- As tarefas escritas, em conjunto com questionamentos orais, devem encorajar os alunos a desenvolver e demonstrar a compreensão de

características centrais sobre o assunto que estão aprendendo.

- Os comentários devem identificar o que foi feito adequadamente e o que ainda precisa ser melhorado e oferecer alguma forma de orientação sobre o que fazer para melhorar.
- As oportunidades para que os estudantes respondam às devolutivas devem ser planejadas como parte do processo geral de aprendizagem.

O ponto central aqui é que para ser efetiva, a devolutiva precisa levar ao pensamento. A implementação desse tipo de reforma pode mudar as atitudes de alunos e professores com relação ao trabalho escrito: a avaliação dos trabalhos dos estudantes será vista menos como um julgamento somativo e competitivo, e mais como um passo importante no processo de aprendizagem.

AUTOAVALIAÇÃO E AVALIAÇÃO PELOS PARES

Os estudantes podem atingir um objetivo de aprendizagem apenas quando compreendem o que esse objetivo significa e o que devem fazer para alcançá-lo. Assim, a autoavaliação é essencial para a aprendizagem¹². Muitos professores que tentaram desenvolver as habilidades de autoavaliação de seus estudantes perceberam que a tarefa mais difícil é fazê-los refletir sobre seus trabalhos em termos dos objetivos de aprendizagem. À medida que os alunos o fazem, eles começam a desenvolver uma visão mais ampla do trabalho que lhes permite gerenciá-lo e controlá-lo por si mesmos. Em outras palavras, os estudantes estão desenvolvendo a capacidade de trabalhar no nível metacognitivo.

Na prática, a avaliação pelos pares se mostra um importante complemento da autoavaliação. A avaliação por pares é excepcionalmente valiosa porque os alunos podem aceitar melhor as críticas sobre seus trabalhos quando vindas de seus colegas do que aceitariam as vindas do professor. O trabalho com os pares também é valioso porque a troca ocorre em uma linguagem que os estudantes utilizam entre si naturalmente e porque os estudantes aprendem ao adotar a postura de professores e examinar os trabalhos de outros colegas¹³. Uma das professoras compartilhou sua experiência positiva sobre avaliação entre pares:

¹² SADLER, R. Formative assessment and the design of instructional systems. *Instructional Science*, v. 18, p. 119-44, 1989.

¹³ SADLER, R. Formative assessment: revisiting the territory. *Assessment in education*, v. 5, p. 77-84, 1998.

Além de avaliarem e escreverem comentários sobre seus próprios trabalhos (por meio de discussões e com orientações claras), as alunas também avaliam e oferecem comentários nos trabalhos umas das outras. Isso elas fazem de uma forma bastante sensível e madura, o que tem se mostrado uma experiência muito valiosa. As estudantes sabem que a lição de casa será conferida por elas mesmas ou por alguma outra colega da turma no começo da próxima aula. Isso nos levou a uma rotina bem estabelecida e apenas em ocasiões muito raras as estudantes falham em completar os seus trabalhos. Elas se mostram orgulhosas por um trabalho claro e bem apresentado que uma de suas colegas possa pedir para examinar. Qualquer desacordo sobre as respostas é cuidadosa e abertamente discutido até que o consenso seja alcançado. — Alice, Escola Waterford.

A última frase no comentário dessa professora levanta um ponto muito importante: quando os estudantes não compreendem uma explicação, é mais provável que eles interrompam um colega de uma forma que não interromperiam a professora. Para além dessa vantagem, a avaliação pelos pares é valiosa porque coloca o trabalho nas mãos dos estudantes. O professor fica livre para observar e refletir sobre o que está acontecendo, e assim formular intervenções úteis:

Fazemos avaliação por pares com regularidade, e achamos essa prática realmente útil. Diversas incompreensões vêm à tona, e então nós as discutimos enquanto percorremos a lição de casa. Na sequência eu observo os comentários que os pares fizeram e falo individualmente com os alunos enquanto circulo pela sala. — Rose, Escola Brownfields.

Todavia, a avaliação por pares acontecerá apenas se os professores ajudarem seus alunos, particularmente aqueles com maiores dificuldades, a desenvolver essas habilidades. Isso pode demandar tempo e prática:

As crianças não possuem habilidades no que eu estou querendo que elas façam. Acho que o processo é mais efetivo no longo prazo. Se você investir tempo nele, o custo que se paga será compensado, uma vez que esse processo tornará os estudantes mais independentes na forma como eles aprendem e para assumirem mais responsabilidades. — Tom, Escola Riverside.

Uma ideia simples e efetiva é que os estudantes utilizem símbolos como o das luzes do semáforo, marcando seus próprios trabalhos como verde, amarelo ou vermelho se acham que desenvolveram um entendimento bom, parcial ou pequeno. Essas marcas servem como formas simples de comunicar a autoavaliação. Os estudantes podem então ter de justificar seu julgamento em um grupo de pares, e com isso conectarem a avaliação por pares com a autoavaliação. Essas conexões podem ajudá-los a desenvolver as habilidades e o distanciamento necessário para realizar uma autoavaliação efetiva.

Outra abordagem é pedir que os estudantes marquem o seu próprio trabalho com uma das cores do semáforo e então pedir que eles indiquem, ao levantarem as mãos, se colocaram um ícone verde, amarelo ou vermelho. O professor pode então agrupar os verdes e os amarelos para se ajudarem a resolver os problemas que encontraram, e trabalhar diretamente com os alunos do grupo vermelho para que aprofundem os problemas encontrados por eles. Para que esse tipo de atividade em grupo funcione, os estudantes precisam de orientação sobre como trabalhar em grupo, incluindo habilidades de como escutar uns aos outros e se revezar ao falar.

Em alguns conteúdos, também é bastante útil investir algum tempo para ajudar os alunos a compreender as rubricas de avaliação. Eles podem receber versões simplificadas das rubricas utilizadas pelos professores, ou podem ser encorajados a reescrever essas rubricas ou fazer as suas próprias. Novamente, a autoavaliação e a avaliação entre pares estão intimamente ligadas. Observadores em diversas aulas de linguagem viram os alunos aplicando as aprendizagens que obtiveram no trabalho de avaliação por pares em seus próprios trabalhos. Um comentário frequentemente ouvido era “eu também não fiz isso” ou “eu também preciso fazer aquilo”.

As reflexões dos estudantes sobre suas compreensões podem também ser utilizadas como informação para o ensino, e suas devolutivas podem indicar a quais áreas ou conteúdos o professor deve dedicar mais tempo. Uma orientação útil é pedir que os estudantes indiquem por meio das “cores do semáforo” suas avaliações do final da unidade antes do início da próxima unidade: aquilo que for apontado como vermelho ou amarelo pode ser utilizado como prioridade para ajustar o planejamento do professor. Nossa experiência nos leva a oferecer as seguintes recomendações para aprimorar essa prática na sala de aula:

- Os critérios para avaliar o desempenho em qualquer atividade devem ser esclarecidos aos estudantes para possibilitar que eles tenham um claro entendimento tanto dos objetivos do trabalho quanto do que significa fazê-lo com sucesso. Potencialmente esses critérios podem ser abstratos, mas exemplos concretos podem ser utilizados para modelar o desenvolvimento dessa compreensão.
- Os estudantes devem ser ensinados sobre os hábitos e as habilidades necessários para a avaliação por pares, tanto porque isso tem valor intrínseco quanto porque a avaliação por pares pode auxiliar no desenvolvimento da objetividade necessária para a autoavaliação.
- Os estudantes devem ser encorajados a manter o foco nos objetivos

do trabalho e a avaliar o seu progresso à medida que vão ao encontro desses objetivos. Com isso eles serão capazes de orientar sua própria produção e se tornar aprendizes independentes.

O ponto principal aqui é que a avaliação por pares e a autoavaliação proveem contribuições únicas para o desenvolvimento das aprendizagens dos estudantes. De fato, elas asseguram o alcance de objetivos que não podem ser atingidos com outros recursos.

O USO FORMATIVO DE TESTES SOMATIVOS

As práticas de autoavaliação e de avaliação por pares podem ser utilizadas na preparação para os testes, bem como para lidar com o seguinte problema:

[Os estudantes] não mencionaram nenhuma das estratégias de revisão que havíamos discutido em aula. Quando questionados mais de perto, ficou claro que eles gastaram o tempo utilizando técnicas de revisão bastante passivas. Eles apenas leram os seus trabalhos, mas fizeram muito pouco em termos de revisão ativa. Eles não estavam transferindo as estratégias de aprendizagem ativa que utilizamos em sala de aula para trabalharem em casa. — Tom, Escola Riverside.

Para remediar essa situação, pode-se solicitar aos alunos que utilizem as “cores do semáforo” em uma lista de palavras-chave ou tópicos que serão demandados no teste. O objetivo dessa atividade é estimular que os estudantes reflitam sobre os pontos nos quais consideram sua aprendizagem segura, os quais eles marcam como verdes, e onde precisam concentrar os seus esforços, aqueles em amarelo ou vermelho. Essas cores do semáforo formam a base para um plano de revisão. O professor pode então solicitar que os estudantes identifiquem questões que aprofundem as áreas “vermelhas”. Eles podem trabalhar com seus materiais didáticos em pequenos grupos para responder a essas questões com sucesso.

Os resultados dos testes também podem ser uma ocasião para o trabalho formativo. Oferecer comentários nos testes dos colegas pode ser útil tanto quanto o é em trabalhos durante as aulas, e essa atividade pode ser ainda mais bem-sucedida se os estudantes tiverem que elaborar uma rubrica de avaliação — um exercício que foca a atenção dos alunos nos critérios de qualidade que são relevantes para a sua produção. Após esse processo de avaliação pelos pares, os professores podem reservar um tempo para discutir as questões que se mostraram difíceis para um grupo grande de alunos, enquanto a tutoria entre pares pode dar conta dos problemas encontrados por uma minoria do grupo.

Outro achado que emergiu dessa pesquisa foi que os estudantes que aprenderam a se preparar para os testes formulando e respondendo suas próprias perguntas tiveram uma performance superior àqueles que se preparam das formas convencionais¹⁴. A preparação de questões auxilia os estudantes a desenvolverem uma visão mais ampla do tópico:

Os alunos tiveram que pensar sobre o que torna uma questão boa para um teste e ao fazer isso tiveram um claro entendimento do material. Como um desdobramento desse processo, as melhores questões foram utilizadas no teste da turma. Dessa forma, os estudantes puderam ver que seu trabalho foi valorizado, e eu pude fazer uma avaliação do progresso que tiveram nessas áreas. Ao revisar o teste, é possível fazer uma boa utilização do trabalho em grupo e das discussões entre os estudantes concentrando-se em áreas específicas que merecem mais atenção. — Angela, Escola Estadual Cornbury.

Experiências como essas desafiam as expectativas comuns. Há pessoas que defendem que a avaliação somativa e a formativa são tão distintas em seus propósitos que devem ser mantidas separadas, e esses argumentos são fortalecidos quando alguém experimenta os efeitos negativos que testes somativos restritivos e de alto impacto têm no ensino. Todavia, é fantasioso esperar que professores e estudantes pratiquem tal separação, de modo que o desafio é alcançar uma relação mais positiva entre os dois. Todas as formas que descrevemos para fazê-lo podem ser utilizadas para testes nos quais os professores possuem o controle sobre o planejamento e a atribuição de notas. Talvez sua aplicação seja um pouco mais limitada para testes sobre os quais os professores têm pouco ou nenhum controle.

De um modo geral, as principais possibilidades para aprimorar as práticas de avaliação ao utilizar testes somativos para fins formativos são as seguintes:

- Os estudantes podem se engajar em uma revisão reflexiva do trabalho que fizeram para possibilitar que planejem suas correções com efetividade.
- Os estudantes podem ser encorajados a elaborar perguntas e corrigir respostas para ganhar um entendimento do processo de avaliação e com isso refinar seus esforços para melhorar.
- Os estudantes devem ser encorajados por meio da avaliação por pares e da autoavaliação a aplicarem critérios que os ajudem a compreender como seu trabalho pode ser melhorado. Isso pode incluir oportunidades para os estudantes trabalharem novamente as respostas de seus testes em sala de aula.

¹⁴ Veja, por exemplo, FOOS, P.; MORA, J.; TKACZ, S. Student study techniques and the generation effect. *Journal of Educational Psychology*, v. 86, p. 567-576, 1994; e KING, A. Facilitating elaborative learning through guided student-generated questioning. *Educational Psychologist*, v. 27, p. 111-126, 1992.

A mensagem geral é que os testes somativos devem se tornar uma parte positiva do processo de aprendizagem. Por meio do envolvimento ativo no processo de verificação da aprendizagem, os estudantes têm a possibilidade de perceber que podem ser beneficiários e não vítimas dos testes, uma vez que esses podem auxiliá-los no processo de aprendizagem.

REFLEXÕES: ALGUMAS QUESTÕES SUBJACENTES

As mudanças que são vinculadas à melhoria da avaliação para a aprendizagem levaram à reflexão — nossa e dos professores — sobre questões profundas relativas à aprendizagem e ao ensino.

TEORIA DE APRENDIZAGEM

Uma das coisas mais surpreendentes que ocorreu durante as primeiras sessões de formação foi que os professores participantes nos solicitaram que elaborássemos um encontro sobre a psicologia da aprendizagem. Pensando em retrospecto, talvez não devêssemos estar tão surpresos com esse pedido. Afinal de contas, tínhamos destacado que as devolutivas teriam função formativa apenas se as informações fornecidas aos estudantes fossem utilizadas por eles para melhorar sua performance. Mas ao passo que uma pessoa pode resolver posteriormente se uma devolutiva teve ou não o efeito desejado, o que os professores realmente necessitavam era uma forma de dar uma devolutiva aos estudantes que eles soubessem de antemão que seria útil. Para fazer isso, eles precisariam elaborar modelos sobre como os estudantes aprendem.

Com isso os professores passaram a tomar extremo cuidado ao selecionar tarefas, perguntas e outros comandos de atividades para assegurar que respostas dos estudantes efetivamente auxiliassem o processo de ensino. Esses tipos de resposta podem “colocar sobre a mesa” as ideias que os estudantes mobilizam para determinada tarefa. A chave para uma aprendizagem efetiva é então encontrar formas de ajudar os estudantes a estruturar o seu conhecimento para construir ideias novas e poderosas. Nas salas de aula do projeto, quando os professores passaram a ouvir as respostas dos alunos com mais atenção, começaram a perceber mais profundamente que a aprendizagem não era um processo de recepção passiva do conhecimento, mas um processo no qual os estudantes estão ativamente criando suas próprias compreensões. Posto de maneira simples, não importa quanta pressão exista para atingir uma boa nota nos testes, a aprendizagem só pode ser realizada pelo estudante.

Os estudantes passaram a compreender o que é considerado um bom trabalho por meio de exemplos. Algumas vezes isso era feito com uma discussão com toda a classe focada em um exemplo particular; outras vezes isso era alcançado por meio do uso de um conjunto de critérios utilizados no trabalho com os pares.

Engajar-se na avaliação por pares ou na autoavaliação é muito mais do que apenas checar os erros ou fraquezas. Esse processo envolve tornar explícito o que normalmente é implícito, e conseqüentemente requer que os estudantes sejam ativos na aprendizagem. Como um estudante escreveu:

Depois que um colega corrigiu minha pesquisa, pude reconhecer meus erros mais facilmente. Espero que não apenas eu, mas também o estudante que deu devolutivas à minha pesquisa, tenha aprendido algo com ela. Da próxima vez terei que tornar minhas explicações mais claras, porque ele disse que elas “estavam difíceis de entender” [...] Agora eu explicarei minha equação novamente de modo que ela fique mais clara.

Os estudantes também se tornam muito mais conscientes sobre quando estão aprendendo e quando não estão. Uma turma que tinha aulas com outro professor que não enfatizava a avaliação para a aprendizagem surpreendeu aquele professor ao reclamar: “Veja, nós lhe dissemos que não entendemos esse conteúdo. Por que você está passando para o próximo tópico?” Enquanto estudantes que estão sintonizados com suas aprendizagens podem criar dificuldades para os professores, nós acreditamos que esse seja exatamente o tipo de problemas que queremos ter.

DIFERENÇAS ENTRE AS DISCIPLINAS

Ao ouvir sobre as pesquisas e discutir ideias com seus colegas, os professores construíram um repertório de habilidades genéricas. Eles planejaram suas próprias questões, ofereceram um tempo de espera apropriado e deram devolutivas que eram planejadas para provocar reflexão. Eles asseguraram que os estudantes tivessem tempo suficiente durante as aulas para avaliar seu próprio trabalho e o de seus colegas.

Todavia, ficou claro, depois de certo tempo, que essas estratégias haviam chegado a um limite. Escolher uma boa questão requer um conhecimento detalhado do conteúdo disciplinar, mas não necessariamente os conhecimentos adquiridos nos estudos avançados nesse conteúdo. Um alto nível de qualificação em determinada área do conhecimento é menos importante do que um conhecimento profundo de seus princípios fundamentais, um conhecimento sobre os tipos de dificuldades que os estudantes podem ter potencialmente e a capacidade de elaborar questões que possam estimular

um pensamento produtivo¹⁵. Além disso, tal conhecimento pedagógico do conteúdo é essencial para interpretar as respostas dos estudantes¹⁶. Isto é, aquilo que os estudantes dizem contém pistas sobre o seu pensamento que requerem atenção, mas captar essas pistas requer um conhecimento profundo sobre as dificuldades comuns da aprendizagem de um dado conteúdo. Assim, ao passo que os princípios gerais da avaliação formativa se aplicam às diferentes disciplinas, a forma como ela se manifesta em cada uma delas pode variar. Encontramos essas diferenças ao fazer comparações entre os professores de matemática, ciências e linguagem.

Em matemática, os estudantes precisam aprender a utilizar procedimentos válidos e conceitos que os sustentam. As dificuldades podem surgir quando os estudantes aprendem estratégias que se aplicam apenas a contextos limitados e não percebem que elas são inadequadas em outras situações. As questões de debate devem então ser elaboradas para trazer essas estratégias para a discussão e para explorar problemas na compreensão dos conceitos de forma que os estudantes possam captar a necessidade de mudar a forma de pensar. Nesse tipo de aprendizagem, normalmente há um resultado correto bem definido. Em exercícios mais abertos, como uma investigação da aplicação de ideias matemáticas em problemas cotidianos, pode haver uma variedade de boas soluções. Nesse sentido, a compreensão dos critérios de qualidade é mais difícil de alcançar e talvez requeira uma discussão conjunta de exemplos e dos critérios abstratos que eles exemplificam.

Em ciências, a situação é bastante similar. Há vários aspectos do mundo natural para os quais as ciências proveem um modelo “correto” de explicação. Todavia, fora da escola, muitos estudantes adquirem diferentes ideias. Por exemplo, muitos estudantes podem acreditar que os animais são seres vivos porque se movem, mas que flores e vegetais não o são porque são estáticos. Ou os estudantes podem pensar que os astronautas parecem não ter peso algum na

¹⁵ Veja, por exemplo, ASKEW, M. et al., *Effective teachers of numeracy: final report*. Londres: King’s College London, School of Education, 1997. Nesse estudo, não foi encontrada correlação entre o progresso feito pelos alunos de anos iniciais do Ensino Fundamental em aritmética e o nível mais alto de matemática estudada pelo professor. De fato, havia uma relação negativa (porém estatisticamente insignificante) entre ambos. Os estudantes que fizeram mais progresso foram ensinados por professores que não detinham altos níveis de conhecimento da disciplina, mas que enfatizavam as conexões entre os conceitos matemáticos.

¹⁶ Sobre a ideia de conhecimento pedagógico do conteúdo, ver: SHULMAN, L. *Conhecimento e ensino: fundamentos para a nova reforma*. Cadernos Cenpec | Nova Série, v. 4, n. 2. Disponível em: <<http://cadernos.cenpec.org.br/cadernos/index.php/cadernos/article/view/293>>. Acesso em: 16 mar. 2019. [N. de T.]

Lua porque lá não existe ar. Muitas dessas “concepções alternativas” podem ser antecipadas porque estão bem documentadas. O que também está bem documentado é que meramente apresentar a visão “correta” não é efetivo. A tarefa nesses casos é propor uma discussão aberta de tais ideias e oferecer devolutivas que desafiem tais concepções ao introduzir novas evidências e argumentos que apoiem o modelo científico.

Existem outros aspectos para os quais um resultado fechado é menos definido. Tal como em matemática, investigações abertas demandam uma abordagem diferente para a avaliação formativa. Ainda mais abertas são as questões sobre as implicações sociais ou éticas dos avanços científicos, porque para elas não existe uma “resposta correta”. Assim, esse trabalho tem de ser “aberto” de uma maneira mais fundamental. Para tanto, a prioridade é fornecer devolutivas que desafiem os estudantes a provar seus pressupostos e que os ajudem a ser críticos acerca da qualidade de qualquer argumento.

A avaliação por pares e a autoavaliação possuem uma longa história na área de linguagens. Tanto a natureza dessa disciplina quanto o caráter mais aberto das produções que lhe são características fazem desse tipo de prática fundamental para um dos muitos objetivos da área, que é ampliar o julgamento crítico dos alunos.

Uma segunda função importante da avaliação por pares e da autoavaliação foi introduzida por Royce Sadler, que argumentou que os critérios isoladamente eram insuficientes para julgar a qualidade de determinado trabalho ou orientar a progressão nele, porque sempre haveria muitas variáveis¹⁷. A chave está em como interpretar os critérios para um caso específico, o que envolve “conhecimento cooperativo¹⁸”. Os professores adquirem esse conhecimento por meio da avaliação que fazem dos estudantes, e é esse processo que lhes permite diferenciar entre níveis de proficiência e desenvolver um senso sobre como o progresso está sendo alcançado. A avaliação por pares e a autoavaliação provocam uma oportunidade semelhante para os alunos para que eles sejam aprendizes dentro da “cooperativa”, desde que os critérios de qualidade sejam claramente comunicados.

Em linguagens, tal como em matemática ou ciências, é preciso prestar atenção às atividades centrais. As mais bem-sucedidas são aquelas que

¹⁷ SADLER, R., 1989

¹⁸ No original, guild knowledge. Guild significa corporação, associação, grêmio. A opção por traduzir como “conhecimento cooperativo” se deu pelo entendimento global de que se trata de um conhecimento que se desenvolve no coletivo, no grupo, e não individualmente. [N. do T.]

proveem aos estudantes a oportunidade tanto de ampliar sua compreensão de um contexto em determinado texto quanto de estruturar progressivamente suas ideias antes de escreverem. Caracteristicamente, isso inclui trabalho em pequenos grupos ou em duplas, com os resultados normalmente sendo discutidos em conversas com toda a classe. Novamente, esse tipo de trabalho não é incomum em aulas de linguagem, e requer a habilidade de estruturar a tarefa o suficiente para provocar a construção progressiva da aprendizagem, mas sem deixá-la tão controlada que limite o pensamento. Essas atividades não apenas oferecem aos estudantes uma oportunidade para desenvolver seus pensamentos por meio da fala, mas também oferecem aos professores uma oportunidade de darem devolutivas ao longo de uma aula por meio de questionamentos e orientações que emergem da discussão. Quanto maior a qualidade da tarefa, maior a qualidade das intervenções.

As diferenças entre as tarefas de aprendizagem podem ser compreendidas em termos de um espectro. Em uma ponta estão as tarefas fechadas com uma única resolução bem definida; na outra ponta estão as tarefas abertas com uma ampla gama de resoluções aceitáveis. As tarefas em aulas de linguagens — por exemplo, escrever um poema — estão concentradas principalmente no grupo das tarefas abertas. Mas há elementos fechados mesmo para esse tipo de tarefas — por exemplo, a observância de regras gramaticais ou convenções do gênero. As tarefas em matemática, por sua vez, são normalmente mais fechadas, mas a aplicação da matemática em problemas cotidianos requer avaliações mais abertas. Assim, em diferentes medidas, será necessária orientação para esses dois tipos de atividades de aprendizagem em ambas as disciplinas.

A despeito das diferenças, a experiência tem demonstrado que o desenvolvimento de habilidades mais genéricas se aplica entre as diferentes disciplinas. Um dos professores de ciências do projeto deu uma formação para os outros professores de sua escola sobre sua experiência na utilização de algumas dessas habilidades mais genéricas que vimos discutindo e descobriu posteriormente que tais práticas haviam se espalhado por várias disciplinas.

Os professores de arte e encenação o fazem o tempo todo, bem como os professores de tecnologia (alguma coisa relacionada com atividades abertas, projetos de longo prazo e talvez um currículo menos limitado?). Mas um professor de inglês veio até mim hoje e disse: “Ontem à tarde foi fantástico. Tentei hoje com minha turma de 8º ano e funcionou. Sem mãos erguidas e dando tempo para eles pensarem nas respostas. Tive um retorno fantástico de algumas crianças que mal falaram nas aulas durante todo o ano. Todas elas queriam dizer alguma coisa, e a qualidade das respostas foi brilhante. Essa é a primeira vez em anos em que aprendi algo novo que realmente fará diferença no meu ensino.” — James, Escola Two Bishops.

MOTIVAÇÃO E AUTOESTIMA

A aprendizagem não é apenas um exercício cognitivo: ela envolve a pessoa como um todo. A necessidade para motivar os estudantes é evidente, mas normalmente se assume que recompensas extrínsecas como notas, estrelas douradas ou prêmios são a melhor forma de fazê-lo. Todavia, há inúmeras evidências que desafiam esse pressuposto.

Os estudantes vão investir seu esforço em uma tarefa apenas se acreditarem que podem alcançar alguma coisa. Se uma tarefa de aprendizagem é vista como uma competição, então todos estarão cientes de que haverá vencedores e perdedores, e aqueles que possuem um histórico de “perdedores” verão pouco sentido em participar da atividade. Assim, o problema é motivar a todos, mesmo que alguns acabem alcançando menos que os outros. Para lidar com esse problema, o tipo de devolutiva oferecido é muito importante. Muitos estudos apoiam essa ideia. Aqui estão alguns exemplos.

- Alunos que foram informados de que as devolutivas “os ajudariam a aprender” efetivamente aprenderam mais do que aqueles que foram informados de que “como vocês fazem nos diz quão inteligentes vocês são e quais notas irão tirar”. As diferenças foram mais amplas para os alunos de baixo desempenho¹⁹.
- Os estudantes que receberam as devolutivas na forma de notas são mais propensos a compreendê-las como uma maneira de comparar a si mesmos com os outros (envolvimento do ego); aqueles que receberam apenas os comentários os veem como uma forma de ajudá-los a melhorar (envolvimento com a tarefa). Esse último grupo teve um desempenho melhor do que o primeiro²⁰.
- Num sistema competitivo, os alunos de baixo desempenho atribuem sua performance à falta de habilidades, e os alunos de alto desempenho, ao seu esforço. Num sistema orientado por tarefas, todos relacionam a performance ao esforço e a aprendizagem é enriquecida, particularmente entre os alunos de baixo desempenho²¹.
- Uma revisão ampla de estudos sobre devolutivas demonstrou que

¹⁹ NEWMAN, R.; SCHWAGER, M. Students' help-seeking during problem solving: effects of grade, goal, and prior achievement. *American Educational Research Journal*, v. 32, p. 352-376, 1995.

²⁰ BUTLER, R. Task-involving and ego-involving properties of evaluation: effects of different feedback conditions on motivational perceptions, interest, and performance. *Journal of Educational Psychology*, v. 79, p. 474-482, 1987.

²¹ CRAVEN, R.; MARCH, H.; DEBUS, R. Effects of Internally focused feedback on enhancement of academic self-concept. *Journal of Educational Psychology*, v. 83, p. 17-27, 1991.

essa prática melhora a performance de 60% dos alunos. Nos casos em que as devolutivas não foram úteis, descobriu-se que elas acabavam sendo um mero julgamento ou nota com nenhuma indicação sobre como melhorar²².

De um modo geral, as devolutivas dadas na forma de recompensas ou notas aumentam o envolvimento do ego em vez de favorecer o envolvimento com a tarefa. Elas podem focalizar a atenção dos alunos na sua “habilidade” em vez de na importância do esforço, e com isso prejudicar a autoestima dos alunos de baixo desempenho e levar a problemas de “desamparo aprendido²³”. A devolutiva que foca aquilo que precisa ser feito encoraja os alunos a acreditar que eles podem melhorar. Esse tipo de devolutiva pode melhorar a aprendizagem, tanto diretamente, por meio do esforço que ela estimula, quanto indiretamente, ao apoiar a motivação para investir nesse esforço²⁴.

A GRANDE IDEIA: FOCO NA APRENDIZAGEM

Todas as experiências que tivemos nesse projeto apontam para a necessidade de repensar o objetivo central do ensino: ampliar a aprendizagem dos estudantes. O alcance desse objetivo demanda disposição para repensar o planejamento das aulas, em conjunto com uma prontidão para mudar os papéis que tanto os professores quanto os alunos desempenham no apoio do processo de aprendizagem.

UM AMBIENTE DE APRENDIZAGEM: PRINCÍPIOS E PLANOS

Melhorias da aprendizagem na sala de aula requerem antecipações cuidadosas:

Refletir efetivamente sobre o ensino tem significado que eu me tornei capaz de propor ideias e estratégias para lidar com o que quer que emergja na sala de aula, o que contribuiu imensamente para o meu desenvolvimento profissional. Essa influência provocou uma mudança de “O que eu vou ensinar e o que os alunos vão fazer?” para “Como eu vou ensinar e o que os alunos vão aprender?”. — Susan, Escola Waterford.

Um dos propósitos das antecipações dos professores é um planejamento que melhore as ações de ensino. Assim, por exemplo, o planejamento de

²² KLUGER, A.; DENISI, A. The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, v. 119, p. 254-284, 1996.

²³ DWECK, C. Motivational processes affecting learning. *American Psychologist (Special issue: Psychological Science and Education)*, v. 41, p. 1040-1048, 1986.

²⁴ DWECK, C. *Self-theories: their role in motivation, personality, and development*. Filadélfia: Psychology Press, 2000.

questões e atividades deve ser feito em função dos papéis que eles têm para a aprendizagem.

Eu certamente não gastava tempo suficiente elaborando as questões antes do início dessa formação em avaliação formativa... Você só percebe quão pobres são os seus questionamentos quando consegue analisá-los. Percebi que estava utilizando questões para ocupar o tempo e que fazia perguntas que requeriam pouco pensamento dos alunos. Quando falamos com os alunos, especialmente aqueles que estão experimentando dificuldades, é importante fazer perguntas que os levem a pensar sobre o conteúdo e que lhes permitam dar o próximo passo para clarear o entendimento. — Derek, Escola Century Island.

De mesma importância é a preocupação com a qualidade das reações dos professores, seja no diálogo estabelecido com os alunos, seja nas devolutivas nos trabalhos escritos. Devolutivas efetivas devem tornar claro para os alunos o que está envolvido na elaboração de um trabalho de qualidade e quais são os passos necessários para melhorar. Ao mesmo tempo, as devolutivas podem ampliar as habilidades e as estratégias que os estudantes aplicam para obter uma aprendizagem efetiva.

Aqui existe uma questão profunda. O ambiente de aprendizagem deve ser “arquitetado” para envolver os estudantes ativamente em suas aprendizagens. A ênfase deve estar no pensamento dos estudantes e em como torná-lo evidente. Como um professor apontou:

Em algum ponto definitivamente houve uma transição, com o foco saindo daquilo que eu trazia para o processo para aquilo que os estudantes estavam contribuindo. Tornou-se óbvio que uma forma de fazer uma mudança sustentável era fazendo os estudantes pensarem mais. Passei a procurar formas de tornar o processo de aprendizagem mais transparente para os estudantes. De fato, agora passo a maior parte do tempo buscando formas de fazer os estudantes assumirem a responsabilidade por sua própria aprendizagem, ao mesmo tempo fazendo com que aprendam mais colaborativamente. — Tom, Escola Riverside.

A colaboração entre os professores e os estudantes, bem como entre os estudantes e seus pares pode produzir um ambiente de apoio no qual os alunos podem explorar suas próprias ideias, ouvir ideias alternativas na linguagem de seus pares e avaliá-las.

Uma técnica tem sido organizar os alunos em pequenos grupos e dar a cada estudante uma parte da unidade para explicar para seus colegas. Eles têm alguns minutos para preparação, recebem algumas dicas e utilizam seus livros de exercícios. Na sequência, cada estudante explica o conteúdo que escolheu para o restante do grupo. Os estudantes rapidamente apontam elementos como “eu acho que os exemplos que você escolheu são ótimos e eles não estão no livro. Não acho que teria pensado neles”. Ou “eu esperava que você mencionasse mais as partículas quando estivesse explicando as diferenças entre líquidos e gases”. Essas aulas se mostraram de valor inestimável — não apenas para mim, porque me permitiram descobrir o nível de compreensão dos alunos, mas para eles também. — Philip, Escola Century Island.

Uma vantagem adicional desse tipo de ambiente de aprendizagem é que o professor pode trabalhar intensamente com um grupo, desafiando as ideias e pressupostos de seus membros, e saber que o restante da turma também está trabalhando duro.

Assim, as principais ações que devem ser tomadas para organizar um ambiente de aprendizagem efetivo são:

- Planejar atividades de sala de aula que ofereçam aos estudantes oportunidades para expressar seus pensamentos de forma que as devolutivas possam ajudá-los a se desenvolver.
- Formular devolutivas que orientem a melhoria das aprendizagens.
- Utilizar atividades que demandem colaboração de modo que todos estejam incluídos e sejam desafiados, treinando os estudantes para ouvir e respeitar as ideias de todos.
- Assegurar que todos os estudantes sejam participantes ativos das aulas e enfatizar que a aprendizagem depende menos da capacidade que eles têm de dar uma resposta correta e mais da prontidão que possuem para expressar e discutir seus entendimentos.

UM AMBIENTE DE APRENDIZAGEM: PAPÉIS E EXPECTATIVAS

Uma coisa é planejar novos tipos de atividades para as aulas e outra bem diferente é colocá-las em prática de forma que os objetivos de aprendizagem aos quais elas servem sejam alcançados. Aqui não existem receitas a serem seguidas de maneira uniforme. No artigo “Por dentro da caixa preta” nós fomos claros ao afirmar que o desenvolvimento da avaliação formativa ocorreria apenas se “cada professor encontrasse sua própria forma de incorporar as lições e ideias que eram oferecidas em seus próprios padrões de sala de aula”.

Um segundo princípio é que o ambiente de aprendizagem previsto requer uma cultura de sala de aula que talvez seja pouco familiar ou desconcertante tanto para professores quanto para os estudantes. O efeito das inovações implementadas por nossos professores foi a mudança do “contrato de sala de aula” entre eles e os alunos — as regras que governam o comportamento esperado e compreendido como legítimo tanto por professores quanto por alunos.

Os estudantes tiveram que deixar de agir como receptores passivos do conhecimento oferecido pelo professor para se tornarem aprendentes ativos que assumem responsabilidades por gerir sua própria aprendizagem.

Para os professores, é necessário coragem. Uma das características mais marcantes do projeto é que, nos estágios iniciais, muitos participantes descreveram as abordagens como “assustadoras” porque sentiram que iriam perder o controle de suas turmas. Ao final do projeto, eles falaram não em perder o controle, mas em dividir responsabilidades pela aprendizagem dos estudantes com a turma — exatamente o mesmo processo visto de duas perspectivas completamente diferentes. Em uma perspectiva, professores e estudantes estão em uma relação de entrega/recebimento; na outra, eles são parceiros para perseguir o mesmo objetivo:

O que a avaliação formativa fez por mim foi manter meu foco menos em mim mesmo e mais nas crianças. Desenvolvi a confiança para empoderar as crianças em levar o processo adiante. — Robert, Escola Two Bishops.

O que aconteceu aqui é que as expectativas de todos — ou seja, aquilo que os professores e os estudantes pensavam que eram os requisitos para ser um professor ou um estudante — foram alteradas. Ao passo que pode parecer intimidador compreender tais mudanças, elas não aconteceram de repente. As mudanças com os professores participantes do projeto vieram lentamente e de forma constante, enquanto a experiência se desenvolvia e a confiança crescia no uso das diversas estratégias para enriquecer as devolutivas e as interações. Por exemplo, muitos professores começaram utilizando questionamentos para encorajar o pensamento. Depois eles melhoraram suas devolutivas orais e escritas, o que impulsionou o desenvolvimento do pensamento e abriu o caminho para a avaliação por pares e a autoavaliação.

Para sintetizar, a cultura e as expectativas na sala de aula podem ser transformadas ao:

- transformar o “contrato” de sala de aula de modo que todos esperem que professores e alunos trabalhem conjuntamente para o mesmo fim: melhorar a aprendizagem de todos.
- empoderar os estudantes para se tornarem aprendentes ativos, com isso assumindo a responsabilidade sobre sua própria aprendizagem.
- incorporar as mudanças no papel do professor aos poucos, uma a uma, conforme foram sendo apropriadas.
- manter a atenção e refletir sobre as formas nas quais a avaliação pode apoiar a aprendizagem.

O QUE VOCÊ PODE FAZER

Para incorporar algumas dessas ideias sobre avaliação formativa em sua prática, o primeiro passo é refletir sobre o que você está fazendo agora. Discutir com seus colegas e observar as práticas uns dos outros pode ajudar a desencadear essa reflexão.

Um próximo passo deve ser experimentar as mudanças. Transformações no atacado podem ser arriscadas e demandantes, então normalmente é melhor pensar sobre algo que você se sinta confiante em tentar — podem ser as “luzes do semáforo”, a avaliação por pares, melhorar os seus questionamentos, etc. — e simplesmente colocar em prática. Se você for um professor dos anos iniciais, pode tentar apenas em uma disciplina. Observamos que, conforme os professores vão ganhando confiança no poder de possibilitar que os estudantes digam o que sabem e o que precisam saber, os professores decidiram que deveriam ampliar o alcance da avaliação para a aprendizagem para todas as suas tarefas de ensino.

A realização de novas estratégias leva a novos progressos. Quando vários colegas estão colaborando, cada um pode começar com diferentes estratégias e compartilhar seus resultados. Esse processo deve levar à formulação de um plano de ação explícito que compreenda um conjunto de estratégias a ser utilizadas, de maneira combinada, preferencialmente começando com uma turma no início do ano. A primeira razão para começar no início do ano é que tanto os alunos quanto os professores terão tempo para se acostumar com a nova forma de trabalhar. A segunda é que pode ser muito difícil mudar hábitos e rotinas que já estão estabelecidos no meio do ano. A experiência de um trabalho sustentável ao longo de um ano, apenas com uma turma, preferencialmente em conjunto com esforços similares por parte de outros colegas, pode prover uma base mais consistente para a adoção dessas práticas em uma escala maior.

A colaboração com um grupo tentando implementar inovações similares é quase essencial. Observação mútua e compartilhamento de ideias e experiências sobre o progresso dos planos de ação podem providenciar o suporte necessário com as táticas e no nível estratégico. O apoio dos colegas é particularmente importante para superar a incerteza inicial ao assumir os riscos da mudança de cultura e as expectativas de sala de aula.

Assim como para qualquer inovação, o suporte administrativo é essencial. Uma forma pela qual a gestão da escola pode apoiar mudanças desse tipo é

ajudando o grupo de professores a ter momentos coletivos regulares para se encontrar. Também devem ser oferecidas oportunidades para os professores dividirem suas experiências durante as reuniões de equipe.

O trabalho de qualquer grupo experimentando inovações é um investimento para toda a escola, de modo que o apoio não deve ser tratado como indulgência para práticas idiossincráticas. De fato, esse trabalho deve estar integrado ao plano de melhoria da escola, com a expectativa de disseminação das práticas que frutificarem após a avaliação das experiências do grupo.

Ao mesmo tempo, é provável que seja necessário rever as políticas correntes da escola porque elas potencialmente podem restringir a possibilidade de utilizar avaliações formativas. Um exemplo notável são políticas que, ao demandarem que todas as atividades dos alunos recebam uma nota, constrem a capacidade dos professores de darem devolutivas apenas utilizando comentários. Cinco das seis escolas do projeto Kmofap modificaram suas políticas de avaliação após a participação de seus professores para possibilitar devolutivas apenas com comentários; em duas delas ficou estabelecido que nenhuma lição de casa receberia notas em toda a escola. Em outro exemplo, um sistema focado em resultados a serem alcançados e que requer revisão frequente estava inibindo qualquer mudança nos métodos de aprendizagem que pudessem desacelerar o “progresso” imediato em troca de produzir ganhos em aprendizagens e habilidades em médio e longo prazo. Aqueles que estão engajados nas inovações precisam ser liberados desse tipo de política.

Assim, o apoio e a disseminação da inovação em avaliação para a aprendizagem serão planejados de forma coerente apenas se a responsabilidade pela visão estratégica de seu desenvolvimento for atribuída a uma pessoa que tenha função de liderança na escola. Nossa experiência corrobora a visão de que alcançar a promessa de uma avaliação formativa apenas esperando que uns poucos indivíduos desejosos de implementá-la o façam é injusto com eles, ao mesmo tempo que esperar que isso aconteça por meio da imposição de uma política que requeira mudanças imediatas de todos os professores em seus papéis e estilos seria absurdo.

O que é necessário é um plano, que se estenda por pelo menos três anos, no qual alguns pequenos grupos sejam apoiados para dois anos de exploração. Esses grupos, então, formam núcleos de experiência e proficiência para a disseminação das ideias pela escola e para apoiar seus colegas a fazerem mudanças similares em suas práticas.

Working Inside the Black Box: assessment for learning in the classroom

Abstract: In their widely read article “Inside the Black Box,” Mr. Black and Mr. Wiliam demonstrated that improving formative assessment raises student achievement. Now they and their colleagues report on a follow-up project that has helped teachers change their practice and students change their behavior so that everyone shares responsibility for the students’ learning.

Keywords: Formative assessment. Feedback. Teaching and learning.

Trabajando desde dentro de la caja negra: evaluación para el aprendizaje en el aula

Resumen: En su ampliamente divulgado artículo “Por dentro da caixa-preta”, Black y Wiliam demuestran que mejorar la calidad de la evaluación formativa aumenta el rendimiento de los estudiantes. Ahora, ellos y sus compañeros aluden a los resultados de un proyecto que sucedió al primero con el objetivo de ayudar a los profesores a cambiar sus prácticas y los estudiantes el comportamiento, para que todos compartan la responsabilidad por el proceso de aprendizaje.

Palabras-clave: Evaluación formativa. Devolutivas. Enseñanza y aprendizaje.

SOBRE OS AUTORES:

Paul Black é professor Emérito na cadeira de Science Education no King's College of London. Doutor em Cristalografia pela Cavendish Laboratory in Cambridge (1954).
E-mail: paul.black@kcl.ac.uk

Christine Harrison é docente do departamento de Science Education no Kings College of London. Doutora em educação pelo King's College of London.
E-mail: christine.harrison@kcl.ac.uk

Clare Susan Lee é docente de educação na Open University of London. Doutora em Educação.
E-mail: clare.lee@open.ac.uk

Bethan Marshall é docente sênior do departamento de English Education do King's College of London. Doutora em Educação.
E-mail: bethan.marshall@kcl.ac.uk

Dylan Wiliam é professor emérito da cadeira de avaliação educacional na University College London. É doutor em Educação pela University of London (1993).
E-mail: dylan.wiliam@ucl.ac.uk