Revista de Estudos Curriculares, vol 9, nº 2, 2018

A "FABRICAÇÃO" DAS NOTAS ESCOLARES: EQUIDADE, SUCESSO E AVALIAÇÃO

Eusébio André Machado[1]

UNIVERSIDADE PORTUCALENSE, INSTITUTO PORTUCALENSE PARA O DESENVOLVIMENTO HUMANO - INPP

RESUMO

Neste artigo, a partir de um acervo significativo, sistemático e abrangente de investigação, realizado ao longo de mais cinquenta anos e em vários domínios (sociologia, psicologia, ciências da educação, etc.), o qual pode ser subsumido pela metáfora da "fabricação" das notas escolares, discute-se a relação entre as práticas de avaliação escolar e a equidade e a promoção do "sucesso" escolar. Neste sentido, apresenta-se uma modelização das práticas de avaliação que considera quatro tipos de critérios: os critérios internos conscientes (Tipo 1), os critérios externos conscientes (Tipo2), os critérios internos inconscientes (Tipo 3) e os critérios externos inconscientes (Tipo 4). Por último, e face à modelização proposta, sugerimos algumas pistas, que recuperam várias sugestões com inscrição reconhecida na literatura, tendo como propósito final sistemas educativos e escolas eficazmente equitativos.

Palavras-chave: equidade; escola; avaliação; práticas; modelização.

Introdução: equidade, sucesso e avaliação

A "equidade" é uma das marcas retóricas (Machado, 2016) - e, em muitos casos, consubstanciada em medidas e práticas efetivas - mais fortes das recentes gerações de reformas educativas realizadas em muitos países[2], no contexto mais alargado não só de consolidação de processos de democratização caracterizados por dinâmicas de "inclusão" e "justiça social", mas também de esforços para "usar os recursos mais eficientemente, e aumentar a oferta de conhecimento e competências que alimenta o crescimento económico e promove a coesão social" (Schleicher, 2018, p. 141).

---

[1] eusebio@upt.pt

[2] No caso da União Europeia, veja-se o quadro estratégico para a cooperação europeia no domínio da educação e formação até 2020 – EF 2020.

Não obstante, as avaliações em larga escala da proficiência dos sistemas educativos, como é o caso do PISA, continuam a evidenciar que o propósito da construção de uma escola equitativa e democrática está ainda muito longe de se considerar alcançado, sendo certo que prevalecem, em muitos casos, efeitos segregativos resultantes, designadamente, do índice socioeconómico dos pais, do género ou da origem étnico-geográfica. Num relatório publicado recentemente (*Equity in Education -Breaking Down Barriers to Social Mobility*), a OCDE enfatiza que, apesar de uma tendência geral de melhoria,

> em todos os países e economias que participaram no PISA 2015, o estatuto socioeconómico tem uma significativa influência no desempenho dos alunos em ciências, leitura e matemática. Por exemplo, a média dos resultados em ciências dos alunos desfavorecidos foi 88 pontos mais baixa do que a média dos resultados dos alunos favorecidos, em média entre os países da OCDE. Este fosso é equivalente a três anos completos de escolaridade. (OCDE, 2018, p. 13)

Na concretização das políticas de "equidade" em contexto escolar, as práticas de avaliação assumem um papel central, crítico e, em muitos sentidos, determinante: a avaliação é, ao mesmo tempo, o mais poderoso instrumento de equidade e de injustiça, de inclusão e de exclusão, de mobilidade social e de "reprodução" e, enfim, de promoção da "coesão social" ou de acentuação das desigualdades. As práticas de avaliação escolar, por isso mesmo, não podem deixar de se situar no cerne do escrutínio sobre a equidade de qualquer sistema educativo, sendo certo que, em todo o caso, haverá outros fatores a montante e jusante que concorrerão para promoção ou despromoção de uma escola efetivamente equitativa.

Ora, nas práticas de avaliação escolar, a atribuição de "notas" adquiriu uma tal importância que, rapidamente, se confundiu com o próprio ato de avaliar, tentando operar um efeito epistemológico que procurou conferir a um processo intrinsecamente complexo um carácter "objetivo". Com efeito, a introdução das "notas" nas práticas de avaliação, um processo historicamente situado[3], insere-se no âmbito mais alargado de quantificação da realidade que as ciências humanas prosseguiram na afirmação do seu estatuto disciplinar e científico (Foucault, 1966). Por outro lado, a introdução de um "regime de medida" através de "notas" por referência a uma escala arbitrária, particularmente evidente na engenharia avaliativa dos "exames", respondeu a uma necessidade profunda de conformidade dos sistemas educativos com o cunho cada vez mais meritocrático das sociedades modernas: classificar, hierarquizar e selecionar "cientificamente" os sujeitos.

[3] Segundo Merle (2018), no sistema educativo francês, o recurso às notas começa a ser frequente a partir da segunda metade do século XIX, representando uma rutura com as práticas anteriores: o "regime de sentenças" é substituído pelo "regime de medida". Mais concretamente, é a partir de 1852 que a Escola Politécnica introduz a escala de 0-20 nos concursos de acesso e, posteriormente, se assiste à sua difusão nos outros níveis de ensino.

Neste artigo, convocamos um acervo significativo, sistemático e abrangente de investigação, realizado ao longo de mais cinquenta anos e em vários domínios (sociologia, psicologia, ciências da educação, etc.), o qual pode ser subsumido pela metáfora da "fabricação". Utilizada originalmente por Perrenoud (1984) no âmbito de uma "sociologia da avaliação", esta metáfora corresponde ao programa epistemológico mais vasto da chamada "construção social da realidade" (Berger & Luckmann, 1999). No caso das práticas de avaliação escolar, a "fabricação" procura mostrar que os processos de atribuição de "notas" e de "classificação" dos alunos e, de um modo geral, de construção da "excelência escolar", mesmo tendo em conta os esforços de objetividade e cientificidade reconhecidos, assenta num conjunto de "arranjos" e enviesamentos sociais, económicos, culturais e psicológicos, de natureza inconsciente e/ou explícita, que comprometem, em larga medida, "equidade" e o "sucesso" dos sistemas educativos, elementos constitutivos da narrativa emancipatória da modernidade que legitimou o projeto de escolarização obrigatória e universal (Lyotard, 1976).

Do nosso ponto de vista, importa convocar este *corpus* no atual contexto para o desenho de respostas que coloquem a avaliação ao serviço das aprendizagens, da promoção do "sucesso" escolar e, em última instância, de sociedades mais coesas e equitativas. Na esteira do movimento da "avaliação formativa" (Abrecht, 1994; Allal, Cardinet & Perrenoud, 1986; Barreira, Boavida & Araújo; Scriven, 1967), trata-se de reforçar a necessidade de uma mudança nas práticas de avaliação escolar: sendo importante e até fundamental o esforço realizado de tornar rigoroso e científico o processo de avaliação "das" aprendizagens, é indispensável, de forma concomitante, reconhecer e operacionalizar práticas de avaliação "para" as aprendizagens. A preocupação com a "cientificidade" da avaliação, sendo algo no qual se deve continuar a investir, não pode ser um fim em si mesmo, mas um instrumento ao serviço das aprendizagens dos alunos.

Neste sentido, e face à modelização que apresentaremos de seguida das práticas de avaliação escolar, sugeriremos algumas pistas, que recuperam várias sugestões com inscrição reconhecida na literatura, tendo como propósito final sistemas educativos e escolas eficazmente equitativos[4].

---

[4] Na limitada economia argumentativa do presente artigo não é possível abordar a *vexata quaestio*: é preciso acabar com as notas? Não obstante, e com risco de uma abordagem demasiado esquemática e simplista, consideramos pertinente projetar esta problemática no contexto português, sobre a qual a produção empírica é manifestamente escassa. Recorde-se que em diversos contextos, em particular o de expressão francófona, a questão das notas ultrapassou os muros da academia, dando origem a um envolvimento mais alargado e público "contra as notas". Objeto de uma abundante literatura, embora ainda sem uma "avaliação científica" (Merle, 2018, p. 274), o argumentário da defesa do "fim das notas" salienta problemas como o impacte negativo nas aprendizagens e na motivação dos alunos, a classificação dos indivíduos segundo uma lógica meritocrática, a intoxicação e adição às notas, a mercadorização da escola com um espaço de transações comerciais, etc. (Cf. Pasquini, 2018, p. 42). Do nosso ponto de vista, para além de relativamente obsoleta, esta questão incide sobre um ponto que escamoteia o que está antes e depois da atribuição das notas, bem como o próprio processo em si de

1. Uma modelização das práticas de avaliação escolar

Seguindo a proposta de Merle (2018, p. 208), a qual se encontra sustentada numa extensa revisão de literatura, apresentaremos uma modelização que conjuga dois pares de categorias de análise dos critérios utilizados nas práticas de avaliação escolar: por um lado, o par categorial interno/externo e, por outro, o par categorial consciente/inconsciente.

O primeiro par categorial – interno/externo – traduz a constatação de que, nas práticas de avaliação escolar, estão envolvidos tanto critérios internos diretamente relacionados com as características intrínsecas do que é o "objeto" de avaliação (neste caso, estamos a considerar o que vulgarmente são chamados os "critérios de correção" de uso particularmente conhecido nos exames nacionais), quanto critérios externos que não têm qualquer relação direta com o "objeto" de avaliação (por exemplo, o nível da turma, as relações afetivas, o efeito de género, etc.), embora produzam também efeitos diretos e determinantes nas "notas" atribuídas e no modo como os alunos são avaliados.

Por sua vez, o segundo par categorial – consciente/inconsciente – corresponde à existência quer de critérios explícitos, intencional e deliberadamente construídos, muitas vezes por obrigação legal, dos quais os diversos intervenientes na avaliação têm consciência, quer de critérios implícitos, resultantes de processos sociais, psicológicos, culturais e outros, que funcionam como um conjunto de lógicas implícitas que escapam ao controlo consciente e deliberativo dos diversos intervenientes na avaliação dos alunos.

A conjugação dos dois pares categoriais permite identificar quatro tipos de critérios (cf. Figura 1): os critérios internos conscientes (Tipo 1), os critérios externos conscientes (Tipo2), os critérios internos inconscientes (Tipo 3) e os critérios externos inconscientes (Tipo 4). Convirá ressalvar, no entanto, que esta tipologia se situa no plano concetual e abstrato de uma modelização e que as práticas de avaliação escolar, por natureza, seguem lógicas híbridas e contaminadas, nas quais não se encontram os tipos puros e estanques a cuja concetualização procederemos de seguida, embora possuam um valor heurístico que pretendemos enfatizar.

---

notificação. Com efeito, os impactes negativos apontados às notas são os problemas das práticas de avaliação escolar considerados em toda a sua amplitude: fará sentido acabar com as notas e continuar uma lógica de avaliação sumativa? Os enviesamentos sinalizados pela literatura desaparecerão se as notas forem suprimidas? A ausência de notas na escola acabará com as desigualdades sociais que a própria escola internaliza e produz? Que avanços se conseguirão em prol da equidade se avaliarmos com testes sem notas? Se aceitarmos o princípio de que a avaliação está relacionada com a permanentemente prestação de informação aos alunos, o que se ganha com a eliminação das notas? Trata-se de um urgente trabalho empírico a realizar sob pena de eventuais passos facialmente bondosos, mas inconsequentes e, sobretudo, contraproducentes.

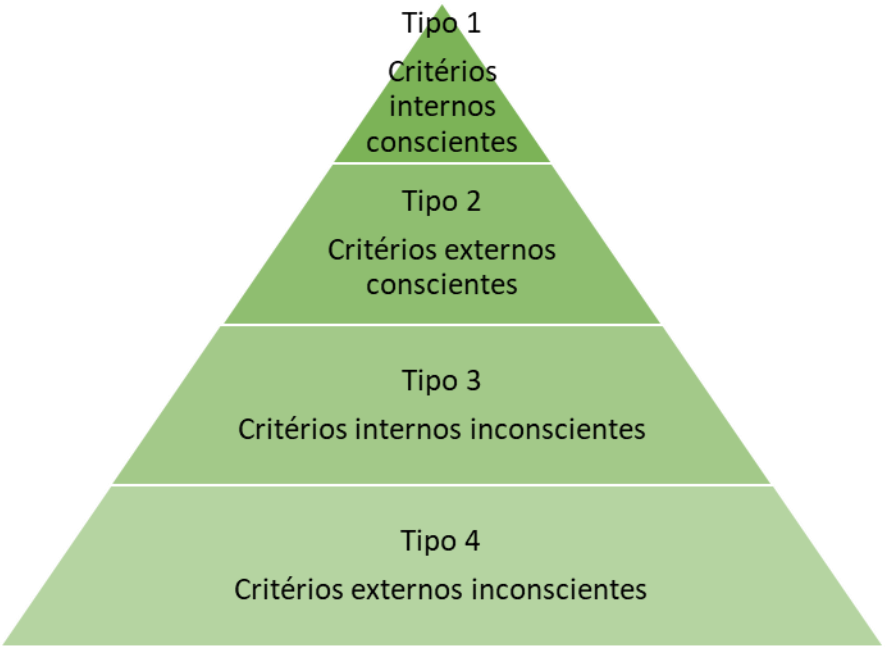

Figura 1 – Modelização das práticas de avaliação escolar[5]

a. Critérios internos conscientes (Tipo 1):

Os critérios tipo 1 resultam de um processo de elaboração realizado pelos professores de acordo com as orientações normativas e em função dos programas das disciplinas, sendo objeto de uma aprovação formal (obrigatória em Portugal desde 2001[6]), normalmente pelos órgãos pedagógicos da escola, e de uma divulgação pública, cujo objetivo principal é garantir o conhecimento formal dos principais intervenientes na avaliação, sobretudo alunos e professores. Estes critérios revestem-se de particular importância, uma vez que sustentam, formal e informalmente, as principais tomadas de decisão dos professores e da escola: as classificações a atribuir (avaliação sumativa) e a retenção/progressão dos alunos.

No caso de Portugal, este tipo de critérios apresenta-se sob duas modalidades distintas cujas diferenças principais assentam no grau de formalização e de publicitação: por um lado, há os critérios que, por imperativo legal, se configuram como "referenciais comuns na escola" e se apresentam com um elevado grau de formalização e de publicitação, sendo assumidos transversalmente por todas as disciplinas ou áreas curriculares; por outro lado, de natureza mais informal e nem sempre publicitados, há os critérios que ocasionalmente são definidos para a correção dos instrumentos de avaliação, em particular os testes, cuja informação é dada aos

[5] Para apresentar a nossa modelização, recorremos à imagem gráfica da pirâmide. Não há dados empíricos que nos permitam estabelecer o peso relativo de cada um dos tipos de critérios nas práticas de avaliação dos professores. Mas a nossa hipótese, meramente heurística, é a de que, pelo menos, os critérios inconscientes são tão determinantes como os critérios conscientes.

[6] Despacho Normativo n.º 30/2001, de 19 de Julho.

alunos por iniciativa dos professores, embora não correspondam a uma prática universal e legalmente compulsória.

Os critérios tipo 1 correspondem, pois, ao protocolo explícito das práticas de avaliação, às "regras do jogo" conscientemente assumidas pelos diversos intervenientes, através das quais a escola e os professores desenvolvem os processos de legitimação técnica e social das notas atribuídas, embora decorram de uma "fabricação" cheia de "segredos" (Merle, 2007) resultante de processos implícitos e tácitos que constituem a dimensão recalcada da avaliação escolar. Em todo caso, do ponto de vista formal, apenas estes critérios relevam para a litigiosidade que própria avaliação gera, circunscrevendo o espaço de escrutínio aberto, transparente e público da equidade das práticas de avaliação escolar.

b. Critérios externos conscientes (Tipo 2):

Os critérios tipo 2 dizem respeito a aspetos/fatores que não estão relacionados com o desempenho dos alunos nas suas produções de avaliação (e.g.: testes), mas que certos professores[7] assumem deliberadamente que têm um peso na correção das provas e na avaliação periódica ou final dos alunos. Correspondem também a contratos informais, para além dos critérios internos explícitos, com os quais os professores regulam as suas relações afetivas e de poder com os alunos e as turmas, dando origem a formas de controlo indireto que se refletem na correção dos instrumentos de avaliação e na avaliação sumativa no final de cada período.

Este tipo de critérios estão consubstanciados em aspetos como o nível escolar da turma, os afetos e o comportamento individual e coletivo dos alunos, embora os mesmos também possam surgir como critérios tipo 4 dado o carácter fluido e aberto dos critérios de avaliação, como referimos anteriormente. Assim, por exemplo, há certos professores que indicam explicitamente ter em conta nos seus critérios de correção o comportamento individual dos alunos ou os comportamentos coletivos da turma, funcionando a avaliação como estratégia de controlo. Do mesmo modo, há também professores que "reconhecem que a atração e a repulsão em relação aos alunos são suscetíveis de influenciar as suas notas" (Merle, 2018, p. 207).

Neste aspeto, Portugal é um caso sui generis e ímpar ao nível internacional, como, de resto, a própria OCDE reconheceu (2012, pp. 59-60). Com efeito, no âmbito dos chamados "critérios de avaliação", definidos ao nível de cada escola como "referenciais comuns" e cuja

[7] Há outros professores que afirmam que estes critérios não têm qualquer influência explícita, intencional e assumida nos seus critérios de avaliação, embora a literatura aponte para efeitos inconscientes de fatores como o comportamento, a atratividade física, o género, etc. Neste sentido, estes critérios também podem ser inconscientes.

principal finalidade é determinar a forma de ponderação da classificação da avaliação sumativa, constata-se uma prática consuetudinária de consideração formal de aspetos como "atitudes", "comportamento", "dimensão sócioafetiva", etc., cujo peso em determinadas disciplinas, segundo relatório recente da Inspeção-Geral de Ensino e Ciência, pode assumir variações significativas (2018, p. 16).

Como critério externo consciente, o "comportamento" não deixa de suscitar vários e sérios problemas, cuja dilucidação não cabe no âmbito deste artigo. Mas a dificuldade em distinguir, nas práticas de avaliação escolares, entre um juízo técnico, sustentado em instrumentos de avaliação e de acordo com "comportamentos" observáveis, e um juízo moral decorrente apreciações e/ou inferências que subsumem sentimentos, impressões e valores constitui uma fonte de enviesamentos na avaliação dos alunos que urge questionar e, sobretudo, mitigar, senão mesmo eliminar[8].

c. Critérios internos inconscientes (Tipo 3):

Os critérios tipo 3, tal como os critérios tipo 4, correspondem por excelência ao campo da investigação feita no âmbito das práticas de avaliação escolar, designadamente na malograda "docimologia" (De Landsheere, 1976), a "ciência dos exames", cujo propósito central de tornar a avaliação "objetiva" acabou por redundar, paradoxalmente, na demonstração da "subjetividade" da avaliação. Trata-se do domínio de uma espécie de "avaliação invisível" (ou "avaliação oculta") que remete também para um conjunto de experiências rotineiras e naturalizadas pelo tempo que constituem um dos aspetos fundamentais da "vida nas salas de aula" (Jackson, 2010). Naturalmente, é neste tipo de avaliação baseada em critérios inconscientes, invisíveis ou ocultos que se inscreve, com mais pregnância e acuidade, a problemática da equidade e, por inerência, do "sucesso" (ou "insucesso") em contexto escolar.

No caso dos critérios tipo 3, estamos a considerar os fatores/aspetos que estão relacionados com as características internas e, sobretudo, formais das produções dos alunos em contexto de avaliação (e.g., testes), as quais exercem um efeito direto na correção e na avaliação, embora o professor não tenha consciência, pelo que não consegue eliminar ou mitigar os diversos enviesamentos que produzem. Dois exemplos sobejamente conhecidos e relatados pela literatura de critérios tipo 3 são o "efeito de halo" e o "efeito de contaminação". Ambos os "efeitos" atuam em função de processos psicológicos, como as "impressões", "crenças" ou perceções", que determinam a correção/avaliação do "conteúdo" de acordo com os aspetos formais inerentes às produções escritas (e até certo ponto, orais) dos alunos.

[8] Segundo Merle, "as modalidades mais contestáveis de arranjo das notas foram suprimidas (em França) por uma circular de julho de 2001, proibindo o uso do zero por "má conduta", bem como uma descida das notas para sancionar uma agitação na turma. Desde então, a avaliação das competências escolares dos alunos não pode regulamentarmente ter em conta os comportamentos escolares desviantes dos alunos" (2018, p. 166).

O “efeito de halo” é um tipo de enviesamento que resulta de indícios gerais de qualidade[9] de qualquer produção dos alunos, sobretudo escrita, como a apresentação “limpa”, o aspeto “estético”, a legibilidade da escrita, a ausência de rasuras, etc., bem como indícios específicos de cada disciplina ou de cada professor (os alunos mais diligentes no desempenho do seu ofício detetam rapidamente estes indícios idiossincráticos dos professores), tais como a diversidade de referências, o rigor dos gráficos, o domínio da língua escrita, a estrutura formal da argumentação, etc. (Merle, 2018, pp. 123-124). As produções dos alunos criam uma impressão imediata (um “halo”) que determina uma atitude positiva e negativa em relação ao seu conteúdo: um “efeito de halo” positivo favorece uma “avaliação aditiva”, ao passo que um “efeito de halo” negativo provoca uma “avaliação subtrativa”.

Por sua vez, o “efeito de contaminação” ocorre quando, perante uma determinada produção escrita ou até oral, o professor não consegue separar os diferentes critérios utilizados para avaliar os objetivos pretendidos, originando, desta forma, um processo de contaminação. Trata-se, por exemplo, da situação em que certos professores subavaliam a riqueza de conhecimentos devido a uma fraca organização formal de teste, de uma resposta ou de trabalho realizado pelos alunos (o “efeito de contaminação” também pode ocorrer ao contrário: uma elevada qualidade formal pode originar uma sobreavaliação de uma produção escrita pouco rica no que respeita aos conhecimentos demonstrados). A “má” ou “boa” apresentação de um trabalho pode funcionar, ao mesmo tempo, como “efeito de halo” que contamina todos os outros critérios definidos, dando origem a um enviesamento que tanto pode originar uma “avaliação aditiva” ou uma “avaliação subtrativa”. Neste aspeto, como demonstra o estudo Simler (cit. por Merle, 2018, pp. 127-130), realizado a partir da análise da correção de provas do “baccalauréat” em ciências económicas e sociais em França, é fundamental a ordem de aplicação dos critérios: o primeiro critério a ser aplicado acaba por exercer um “efeito de halo” nos critérios ulteriores.

d. Critérios externos inconscientes (Tipo 4):

Os critérios externos inconscientes correspondem, por seu turno, a normas externas (leia-se: externas às produções dos alunos) que atuam de modo inconsciente e não controlado nos processos, enviesamentos e “arranjos” da avaliação, evidenciando que as desigualdades sociais do sucesso escolar são construídas dentro das próprias instituições educativas (Merle, 1998). Estes critérios relacionam-se, assim, com as normas específicas e implícitas de cada disciplina, com os enviesamentos individuais e sociais (características escolares, nível escolar dos alunos, nível escolar da turma, origem social, origem étnica, género, etc.) e com os aspetos afetivos e traduzem-se em vários “efeitos” (“efeito de âncora”, “efeito de ordem”, “efeito de

[9] Note-se que alguns destes indícios também são considerados como critérios explícitos por certos professores, pelo que assumem um carácter consciente e deliberado.

reputação", etc.) que funcionam de forma implícita na correção e avaliação de produções escritas dos alunos, como os testes. Emblemático deste tipo de critérios é a lei de Phostumus[10] ou a "constante macabra" (Antibi, 2003) que funciona como uma espécie de *a priori* universal da avaliação e cujo principal pressuposto é a naturalização do "insucesso".

A lei de Phostumus ou a "constante macabra" assentam no princípio de que, em qualquer produção dos alunos e em qualquer nível de ensino, a correção e avaliação deve dar origem a uma distribuição "normal" dos resultados, baseada numa alegada lei de Gauss, segundo a qual deve haver, necessariamente e de forma constante[11], um número reduzido de alunos excelentes e fracos e um número elevado de alunos de nível médio (é a curva de Gauss)[12]. Esta norma externa de notação, mesmo variando conforme os níveis de ensino (é menos frequente no 1.º ciclo e mais frequente no ensino secundário) e de acordo com as disciplinas (é mais frequente nas disciplinas do *core curriculum* e menos frequente nas outras disciplinas), condiciona os "arranjos" das classificações dos alunos, tendo por base a norma da turma, em função de expetativas pessoais, profissionais e sociais. Como refere Merle, "a conceção dominante e institucional das práticas de notação torna suspeito, em todo o caso não conforme e anormal, um teste ou uma média trimestral na qual todos os alunos tivessem a nota média" (2018, p. 121).

A "constante macabra" traduz-se, desde logo, na forma como os testes e os exames nacionais são "calibrados" em termos de dificuldades, sendo evidente a preocupação de evitar a todo custo o "facilitismo", uma aberração face à lei de distribuição normal dos resultados. Neste sentido, e porque um teste ou exame bem elaborados devem conduzir sempre a resultados "negativos", a construção dos instrumentos de avaliação tende a privilegiar a escolha de questões difíceis, e até com "armadilhas", que garantam o mínimo normal de total insucesso e o mínimo normal de pleno sucesso. Como refere De Landsheere,

> O fim em vista é classificar cada um, atribuir-lhe um lugar que lhe pertence num grupo numeroso. Numa palavra, trata-se de organizar uma espécie de concurso, em que o melhor virá a ocupar necessariamente o primeiro lugar.
> É por isso que muitos testes de aptidões ou inventários de conhecimentos são ótimos *instrumentos de seleção*. (1976, p. 245)

---

[10] Phostumus "é o nome do professor holandês a quem é atribuída a paternidade de uma lei relativa à distribuição das notas dadas pelos professores" (Merle, 2018, p. 118) e confirmada por várias investigações ulteriores sob a forma da "curva de Gauss" ou "constante macabra" (Antibi, 2003).

[11] No Japão, até à reforma das práticas de avaliação em 2001, os resultados dos testes, numa escala de 1 a 5, deveriam respeitar obrigatoriamente a curva gaussiana (Merle, 2018, p. 119).

[12] A distribuição gaussiana é a seguinte: "aproximadamente 70% de médios, 13% de bons, 13% de medíocres, 2% de muito bons, 2% de muito maus" (De Landsheere, 1976, p. 245).

Ora, a questão fundamental que os critérios tipo 4 colocam é, pois, a congruência entre a narrativa emancipatória, assente em propósitos democráticos e equitativos, e as práticas de avaliação escolar que, deliberadamente ou não, se inscrevem em lógicas seletivas e meritocráticas[13]. Como vimos, a instituição escolar internaliza as desigualdades sociais e, em certo sentido, reprodu-las, acentua-as e reconfigura-as, sob a forma de diversos enviesamentos e "efeitos" que inviabilizam a matriz axiológica que fundou o projeto iluminista da escolarização universal (Machado, 2016). Por isso, se a equidade e o sucesso forem, de facto, a ambição dos sistemas educativos, das escolas e dos professores, importa reforçar a pregnância das práticas de avaliação, sendo certo que, pelo menos, duas hipóteses de trabalho devem ser tidas em consideração:

1.ª hipótese: toda a prática de avaliação escolar implica o uso de critérios, sejam eles implícitos ou explícitos, internos ou externos, conscientes ou inconscientes, pelo que é desejável que sejam adotados procedimentos de avaliação que promovam a equidade e eliminem, tanto quanto for possível, os enviesamentos indutores de "desigualdades";

2.ª hipótese: a existência de critérios de avaliação explícitos conscientes não elimina a existência de critérios de avaliação implícitos e inconscientes, embora, como veremos adiante, a definição de critérios de avaliação é fundamental para a concretização da avaliação formativa, do feedback útil e da competência de autoavaliação dos alunos.

2. Uma modelização das práticas de avaliação escolar

Em larga medida, a avaliação não muda se a escola não mudar (Perrenoud, 1993), mas a ausência de mudança da e na escola não deve servir para a ausência de qualquer mudança na avaliação, até porque a avaliação pode ser a alavanca para a "metamorfose" da própria instituição escolar. Por outro lado, mudar a avaliação implica também mudar os processos de ensino e de aprendizagem, sobretudo a chamada "gramática escolar" prevalecente (Illich, 1970; Jackson, 2010; Tyack & Tobin) assente no aprender para ser medido, comparado e avaliado (Flores, Machado & Alves, 2017). Neste sentido, a principal mudança que se impõe, com levado grau de urgência e pertinência, é criar uma "cultura de avaliação para apoiar as aprendizagens dos alunos" (Toom, 2017, p. 75), uma avaliação para as aprendizagens ("assessment for learning") que cumpra eficazmente o propósito de uma escola equitativa. Que pistas podem ser consideradas?

1 - Reverter o modelo psicométrico de avaliar, legitimado num propósito ilusório de "justiça", segundo o qual qualquer momento de avaliação é um ato seletivo, pelo que os alunos são submetidos artificialmente às mesmas "provas", ao mesmo tempo e sob as mesmas

[13] Esta tensão é vivida diária e intensamente nas nossas salas de aula, sobretudo no subsetor do Ensino Secundário, cuja função seletiva se sobrepôs a todas as outras devido à importância que adquiriu no ingresso no Ensino Superior.

condições. A realização de testes, no contexto de uma avaliação interna e "contínua, reveste-se de uma ritualização que, cada vez mais, emula os exames nacionais, uma forma de avaliação externa e pontual que constitui a concretização do modelo psicométrico por excelência. Ora, uma escola atenta à diversidade e comprometida com a equidade, não pode deixar de considerar uma espécie de "desenho universal de avaliação", suficientemente capaz de flexibilizar diversificar tempos, modos e processos através dos quais os alunos são avaliados.

2 - Usar critérios de avaliação explícitos e claros que possibilitem ao professor e, sobretudo, ao aluno desenvolver e praticar uma "avaliação para as aprendizagens" (avaliação formativa) baseada no feedback eficaz e construtivo e no desenvolvimento de competências de autoavaliação. A avaliação criterial é o principal instrumento de combate aos critérios externos inconscientes, como a "constante macabra", que naturalizam o insucesso e produzam enviesamentos artificiais à luz de determinadas expetativas e crenças profissionais e sociais. Como refere Toom a propósito do sistema educativo finlandês, na avaliação baseada em critérios, "os alunos não são comparados entre si quanto ao seu desempenho, antes, o seu desempenho é comparado com os critérios de avaliação previamente definidos" (2017, p. 76). Trata-se também de uma forma de evitar "comparações sociais forçadas" que tendem a acentuar efeitos na progressão dos alunos no âmbito de uma turma, dando origem a fenómenos como as "profecias autorrealizadoras negativas" (Bressoux, 2003).

3 - Reforçar as práticas de avaliação formativa, na base de um diálogo constante com os alunos através de um feedback útil. Deste ponto de vista, face à quantidade de informação recolhida ao longo ano, para a qual é dado apenas um uso "sumativo" e, muitas vezes, seletivo, é importante que o uso formativo seja preponderante e se verifique uma "avaliação para as aprendizagens", integrando a avaliação nos processos de aprendizagem e respeitando o princípio do "alinhamento curricular" (Pasquini, 2018). Uma "avaliação para as aprendizagens", de natureza eminentemente formativa, aparece associada à definição de critérios, à flexibilidade nos modos de avaliar, à diversificação dos instrumentos de avaliação e a uma cultura profissional que combate a "naturalização" das desigualdades sociais e as normas externas de sucesso e/ou insucesso.

4 - Diversificar os instrumentos de avaliação, uma vez que constitui a maior garantia de uma avaliação para as aprendizagens, equitativa e democrática face à própria diversidade dos alunos e dos seus modos de aprender. Alguns dos enviesamentos estudados, cuja evidenciação empírica, ao fim de mais de cinquenta anos de investigação, está significativamente consolidada, poderão ser evitados e suprimidos com o recurso a diferentes tipos de instrumentos. Os efeitos de halo, contaminação, ordem, contraste, âncora, etc. podem ser eliminados através do uso de instrumentos de avaliação como o portefólio que permitem um perspetiva mais holística, integrada, multirreferencial e contínua do desempenho dos alunos. Concentrar as práticas de avaliação escolar num único instrumento (seja ele qual for) reduz a

amplitude e a pregnância da informação recolhida e diminui drasticamente a sensibilidade à singularidade e à diferença.

5 - Preservar o anonimato social e escolar do aluno (Merle, 2014, 2018), o que implica pôr em causa coisas como as fichas de informação preenchidas pelos alunos no início do ano letivo, as informações veiculadas sobre o percurso escolar, as reuniões de "rotulagem" prévias ao início das aulas. Implica também questionar as chamadas práticas de "avaliação diagnóstica" que funcionam, muitas vezes, numa lógica de pré-teste de seleção que é quase sempre confirmado pelos resultados ulteriores, seguindo a dinâmica psicológica das profecias que se autorrealizam. Trata-se, pois, da necessidade de controlar os processos de estigmatização e de cristalização de expetativas que acarretam efeitos reputacionais profundos e longitudinais dos quais resultam os principais enviesamentos sociais e psicológicos nas práticas de avaliação escolar.

6 - Incrementar o uso da conetividade digital e dos múltiplos dispositivos tecnológicos para a construção de processos de avaliação mais autónomos, mais motivadores e mais formativos (o *feedback* oportuno e pertinente). Por outro lado, a apresentação das produções dos alunos em suporte digital pode mitigar enviesamentos resultantes do "efeito de halo", do "efeito de ordem" e do "efeito de contaminação", uma vez que, até certo ponto, uniformiza fatores idiossincráticos e aleatórios que influenciam a forma como os alunos são avaliados. Além disso, a partilha e a colaboração ficam significativamente potenciadas, permitindo criar espaços de aprendizagem mútua e formas de *feedback* (mesmo quantificado) mais anónimas e personalizadas.

7 - Relacionar a avaliação dos alunos e a avaliação das escolas (Merle, 2014), sendo certo que há um efeito de reputação da escola nos enviesamentos sociais registados pela literatura. Os mecanismos de comparação entre escolas, designadamente através dos *rankings* elaborados a partir dos exames nacionais, produzem vários impactes negativos à luz de um referencial político de equidade, inclusão e democracia. Neste sentido, como acontece no caso finlandês, não se trata de diminuir ou até acabar com a avaliação (interna ou externa) das escolas, mas, pelo contrário, de incrementá-la na deteção de problemas, na monitorização do desempenho e no apoio construtivo. Dito de outro modo, o que está em causa é também um modelo formativo (e não normativo ou comparativo) de avaliação das escolas com o objetivo de "estudar as dificuldades específicas detetadas pelos professores em certas escolas e de encontrar soluções para melhorar o nível médio dos alunos" (Merle, 2014, s/p.). Nesta linha, é também fundamental reconfigurar o modelo de avaliação externa que funciona quase exclusivamente como processo de controlo de aquisição de conteúdos e instrumento seletivo (sobretudo, ensino secundário), acentuando o seu carácter formativo junto dos alunos, pais e professores (Flores, Machado & Alves, 2017).

8 - Promover momentos de não-avaliação, de tal modo que seja possível reverter o princípio totalitário da "omniavaliação" (Machado, 2012) que se tornou a Némesis na vida dos alunos, dos professores e das escolas. O discurso da preparação para a "vida real" (é a ideologia da escola como "vida simulada" e espaço idílico sem negatividade), que exige que a escola coloque os alunos face aos problemas reais com os quais se confrontarão na vida pós-escolar, legitimou a expansão da avaliação a todas as esferas e interstícios da vida escolar. Não se trata da avaliação enquanto instrumento de apoio e de ajuda, mas da avaliação enquanto "competição", "concurso" e quase "prova de vida" a que os alunos estão permanentemente sujeitos. É preciso que os alunos descubram também a alegria da aprendizagem como um fim em si mesmo e não com o fim instrumental da avaliação, o hegemónico "telos" existencial que o "currículo oculto" transmite aos alunos.

Referências

Abrecht, R. (1994). *A avaliação formativa*. Rio Tinto: Edições Asa.

Allal, L., Cardinet, J., & Perrenoud, P. (1986). *A avaliação formativa num ensino diferenciado. Coimbra: Livraria Almedina.*

Antibi, A. (2003). *La constante macabre ou comment a-t-on découragé des générations d'élèves?* Paris: Math'Adore.

Barreira, C. Boavida, J., & Araújo, N. (2006). Avaliação formativa. Novas formas de ensinar e aprende. *Revista Portuguesa de Pedagogia*, 40-3, 95-133.

Berger, P. L., & Luckmann, T. (1999). *A construção social da realidade – um livro sobre a sociologia do conhecimento*. Lisboa: Dinalivro.

Bressoux, P. (2003). Jugements scolaires et prophéties autoréalisatrices: anciennes questions et nouvelles réponses. DOI: 10.3406/dsedu.2003.1028

De Landsheere, G. (1976). *Avaliação contínua e exames. Noções de docimologia*. Coimbra: Livraria Almedina.

Flores, M. A., Machado, E. A., & Alves, M. P. (2017). *Avaliação das aprendizagens e sucesso escolar. Perspetivas internacionais*. Santo Tirso: De Facto Editores.

Foucault, M. (1966). *Les mots et les choses*. Paris: Gallimard.

IGEC (2018). *Avaliação das Aprendizagens dos Alunos do Ensino Secundário: Indicador do Alinhamento das Notas Internas em doze escolas — Relatório 2017*. Lisboa: Inspeção-Geral de Educação e Ciência. Consultado a 27 de dezembro de 2018 de http://www.ige.min-edu.pt/upload/Relatorios/AAAES_Relatorio_2017.pdf

Illich, I. (1970). *Deschooling society*. Londres: Marion Boyars Publishers Ltd.

Jackson, P. W. (2010). *La vida en las aulas*. Madrid: Morata.

Lyotard, J.-F. (1979). *La Condition postmoderne. Rapport sur le savoir*. Paris: Les Éditions de Minuit.

Machado, E. A. (2013). *Avaliar é ser sujeito ou sujeitar-se. Elementos para uma genealogia da avaliação.* Mangualde: Edições Pedago.

Machado, E. A. (2016). *A educação que vem: entre a performatividade e a esperança*. Santo Tirso: Whitebooks.

Merle, P. (1998). *Sociologie de l'évaluation scolaire*. Paris: PUF.

Merle, P. (2007). *Les notes. Secrets de fabrication*. Paris: Presses Universitaires de France.

Merle, P. (2014). Faut-il en finir avec les notes? *La Vie des idées*. ISSN: 2105-3030. URL: http://www.laviedesidees.fr/Faut-il-en-finir-avec-les-notes.html

Merle, P. (2018). *Las pratiques d´évaluation scolaire. Historique, difficultés, perspectives*. Paris: PUF.

OCDE (2012). *Reviews of evaluation and assessment in education* – Portugal. Paris: OECD Publishing.

OCDE (2018). *Equity in education: breaking down barriers to social mobility.* Paris: OECD Publishing. DOI: https://doi.org/10.1787/9789264073234-2-en

Pasquini, R. G. (2018). *Le modèle de l'alignement curriculaire élargi pour étudier les pratiques évaluatives sommatives d'enseignants de mathématiques et de français du secondaire: enjeux conceptuels et pragmatiques* (Thèse de doctorat). Genève: Université de Genève. No. FPSE 704. DOI: 10.13097/archive-ouverte/unige:106442. URN: urn:nbn:ch:unige-1064421

Perrenoud, Ph. (1984). *La fabrication de l'excellence scolaire*. Paris-Genève: Librairie Droz.

Perrenoud, Ph. (1993). Não mexam na minha avaliação! Para uma abordagem sistémica da mudança pedagógica. In A. Estrela e A. Nóvoa (Eds.), *Avaliações em educação: novas perspectivas* (171-191). Porto: Porto Editora.

Schleicher, A. (2018). *World class. How to build a 21st-century school system?* Paris: OECD Publishing.

Scriven, M. (1967). The methodology of evaluation. *Perspectives of Curriculum Evaluation*, Chicago, RandMcNally, Ronda.

Toom, A. (2017). Princípios e práticas de avaliação para apoiar as aprendizagens no Ensino Básico finlandês. In M. A. Flores, E. A. Machado & M. P. Alves, M. P. (2017). *Avaliação das aprendizagens e sucesso escolar. Perspetivas internacionais* (pp. 71-89). Santo Tirso: De Facto Editores.

Tyack, D., & Tobin, W. (1994). The "Grammar" of Schooling: Why Has It Been So Hard to Change? *American Educational Research Journal*, 31, 3, 453-479.

THE "FABRICATION" OF SCHOOL MARKS: ASSESSEMENT, SUCCESS AND EQUITY

ABSTRACT

In this paper, from a significant, systematic and comprehensive collection of research, carried out over a further fifty years and in various domains (sociology, psychology, education sciences, etc.), which can be subsumed by the metaphor of "fabrication" of the school marks, the relationship between school assessment practices and equity and the promotion of school "success" is discussed. In this sense, we present a modelization of assessment practices that considers four types of criteria: conscious internal criteria (Type 1), conscious external criteria (Type 2), unconscious internal criteria (Type 3) and unconscious external criteria (Type 4). Finally, and considering the proposed modelization, we give some suggestions, which recover several suggestions with inscription recognized in the literature, being the promotion of equity in educational systems and schools the final purpose.

Keywords: equity; school; evaluation; practices; modelization.