

Texto de Apoio 8

A classificação ao serviço da  
melhoria das aprendizagens –  
contributos da investigação

Fátima Braga

Eusébio André Machado

PROJETO DE  
MONITORIZAÇÃO  
ACOMPANHAMENTO  
E INVESTIGAÇÃO  
EM AVALIAÇÃO  
PEDAGÓGICA





## Texto de Apoio 8

# A classificação ao serviço da melhoria das aprendizagens – contributos da investigação

Fátima Braga

Eusébio André Machado



### **Ficha Técnica**

**Título:** *A classificação ao serviço da melhoria das aprendizagens – contributos da investigação*

Texto de apoio à formação - Projeto de Monitorização Acompanhamento e Investigação em Avaliação Pedagógica (MAIA)

**Autor:** Fátima Braga e Eusébio André Machado

**Editor:** Ministério da Educação/Direção-Geral da Educação

**ISBN:** 978-972-742-536-5

**Data:** 2023

Por favor, cite esta publicação como:

Braga, F., & Machado, E. A. (2023). *A classificação ao serviço da melhoria das aprendizagens – contributos da investigação*. Projeto de Monitorização Acompanhamento e Investigação em Avaliação Pedagógica (MAIA). Ministério da Educação/Direção-Geral da Educação.



## Índice

<b>Introdução: em torno do conceito de classificação</b>	4
<b>Para que é que deve servir a classificação?</b>	5
<b>Desafios para uma classificação ao serviço da melhoria das aprendizagens</b>	7
<b>Conclusões</b>	11
<b>Referências</b>	12



“O que é que a investigação nos diz sobre as práticas de classificação e avaliação que as escolas podem usar?”

(Brookhart, 2020, p. 1)

### Introdução: em torno do conceito de classificação

A **classificação (grading)** é uma forma de avaliação (*assessment*) que consiste na associação de símbolos a um trabalho dos alunos ou a um conjunto de desempenhos dos alunos (Brookhart *et al.*, 2016; Brookhart, 2017). No entanto, a avaliação e a classificação obedecem a intenções distintas: a primeira faz uma apreciação do desempenho dos alunos, enquanto a segunda comunica, através de notas, os resultados dessa apreciação (Link & Guskey, 2022). Salva-se, no entanto, que a avaliação e a classificação têm um propósito comum: contribuir para que os alunos aprendam melhor (Fernandes, 2021).

As **notas (grades ou marks)** são símbolos que se colocam nos trabalhos dos alunos, ou que se escrevem nos boletins de final de um período de aprendizagem, traduzindo balanços do desempenho dos discentes (Guskey & Brookhart, 2019a; Wiliam *et al.* 2020). Podem ser letras, números, figuras ou um conjunto de descritores; no entanto, remetem sempre para diferentes níveis de desempenho e indicam o quão bem os alunos alcançaram objetivos de aprendizagem (Guskey, 2020a; Wiliam *et al.* 2020).

Por isso, recomendam estes autores que a classificação se faça por **referência a critérios** e com base em *standards* (expectativas para a qualidade do trabalho, que aparecem explicitados no nível superior da descrição de cada critério e remetem para os resultados desejados da aprendizagem - *learning outcomes*). Neste contexto, importa ainda esclarecer o conceito de **pontuação - score** -, que explicita a organização dos números numa categoria ordenada (escalas de classificação ou rubricas) (Brookhart, 2017).

A este propósito, Brookhart *et al.* (2019) referem que os resultados da aprendizagem (*learning outcomes*) são a base de qualquer sistema de avaliação abrangente e equilibrado e a referência relativamente à qual a informação recolhida pelos professores sobre as aprendizagens deve ser interpretada, do mesmo modo que alertam para a necessidade de acautelar a qualidade da avaliação, nomeadamente no que diz respeito ao processo usado para suscitar, reunir, interpretar e usar a informação, e ainda à garantia de intervenção dos diferentes participantes no processo, nomeadamente os alunos.

Os diferentes autores convocados neste Texto de Apoio defendem também que é necessário distinguir a informação que é usada com intenção formativa, na avaliação para



melhorar a aprendizagem, e a informação que se destina a ser usada sumativamente, como balanço, na avaliação das aprendizagens e na certificação dos alunos.

Em suma, um sistema de avaliação coerente fornece informação significativa, relevante e suficiente, em quantidade e qualidade, para a utilização a que se destina (avaliação formativa, ou avaliação para a aprendizagem e avaliação sumativa, ou avaliação das aprendizagens), pelo que importa, por isso, insistir na questão da intencionalidade pedagógica da avaliação e da classificação.

Recomenda-se, assim, que é preciso **resistir à tentação de tudo registar, para tudo considerar na tomada de decisões classificatórias**, uma vez que:

- i) os dados de avaliação formativa não se destinam a tomar decisões relacionadas com a atribuição de notas, mas apenas a promover aprendizagens e a construir *feedback*;
- ii) os dados de avaliação sumativa é que podem ser usados para tomar decisões relacionadas com as notas e estes só devem ser recolhidos depois de os alunos terem tido oportunidades de aprender e de perceberem o que sabem e conseguem fazer com o que aprenderam. (Brookhart, 2017)

### Para que é que deve servir a classificação?

Brookhart (2017) afirma que um sistema de classificação ao serviço da aprendizagem tem que se basear em duas premissas:

- 1.ª Todos os alunos podem aprender. Este é um compromisso que se estabelece entre professores e alunos; não aprenderão todos as mesmas coisas, no mesmo tempo e com o mesmo nível de proficiência, mas aprenderão todos;**
- 2.ª As notas devem refletir a qualidade dessas aprendizagens, isto é, devem comunicar, de modo claro, os resultados das aprendizagens dos alunos, traduzindo o nível em que se encontram no momento em que são comunicadas.**

Relativamente à 2.ª premissa, reforça-se que o propósito das notas é comunicar o nível de desempenho, isto é, o estado em que se encontram as aprendizagens dos alunos relativamente aos critérios (William *et al.*, 2020).

É o que Muñoz e Guskey (2015) tinham verbalizado antes, ao enunciarem que o propósito da classificação é mostrar a qualidade com que os objetivos de aprendizagem estão



presentes no desempenho dos alunos, assim cumprindo as quatro funções que a classificação deve desempenhar (Wolvoord & Johnson Anderson, 2010):

- i) avaliação da qualidade do trabalho dos alunos;
- ii) comunicação com os alunos, os pais e outros interessados;
- iii) orientação da aprendizagem dos alunos, do seu foco e do seu envolvimento;
- iv) organização do trabalho dos alunos e dos professores.

Por isso, atualmente, diferentes autores (Brookhart et al., 2016; Guskey & Brookhart, 2019b e c; Wiliam et al., 2020; Brookhart et al., 2020; Guskey, 2022; Pasquini, 2018, 2019 e 2021) são unânimes ao considerarem que **a classificação deve ser pedagógica**, tal como a avaliação que fornece os dados que as notas comunicam. A esta premissa acrescentam Wiliam et al. (2020) que as notas devem emergir como parte de um processo de aprendizagem em que há articulação do ensino, da aprendizagem, da avaliação e da classificação, o que significa, como Link & Guskey (2022) explicitam, que é necessário alterar, simultaneamente, as dinâmicas pedagógicas e a avaliação. Só assim, defendem, se consegue que:

- i) a classificação, baseada em critérios, ajude os professores a direcionarem melhor os seus esforços para remediarem as lacunas de aprendizagem e abordarem as pequenas dificuldades de aprendizagem dos alunos, antes que se tornem grandes problemas de aprendizagem;
- ii) os alunos tenham um *feedback* mais preciso sobre o progresso da sua aprendizagem;
- iii) os professores sejam mais transparentes sobre o que se pretende que os alunos aprendam, sem que estes tenham que adivinhar o que é importante aprender ou o que precisa de ser feito para atender às expectativas dos professores e da escola;
- iv) a diferenciação pedagógica seja efetiva;
- v) os alunos percebam as intenções de aprendizagem de cada aula e se envolvam em interações centradas na aprendizagem;
- vi) os alunos se responsabilizem pela aprendizagem e se sintam intelectualmente seguros.

Para isso, a classificação não deve ser pensada apenas no final do processo (Yerly, 2022), antes deve moldá-lo desde o primeiro momento em que as aulas são planificadas, para se garantir que as tarefas de avaliação, cujos resultados vão servir para tomar decisões relativas às notas, avaliam o que deve ser aprendido. Pensar no início do processo o modo



como as aprendizagens serão avaliadas e classificadas facilita o trabalho do professor na gestão pedagógica do processo e dos alunos na autorregulação das aprendizagens (Wolvoord & Johnson Anderson, 2010).

### **Desafios para uma classificação ao serviço da melhoria das aprendizagens**

Os princípios acima referidos implicam três grandes desafios em termos de processos de decisão e operacionalização: o alinhamento curricular, a atribuição de notas aos desempenhos dos alunos, nas tarefas de avaliação sumativa, e a decisão sobre as notas a atribuir no final de um período de aprendizagem.

#### **Desafio 1: Relativamente ao alinhamento curricular**

Consideram os autores que temos vindo a cotejar que as práticas de classificação devem ser parte de um sistema de avaliação rigoroso, que promova e certifique a aprendizagem dos alunos, sendo o elo entre o ensino, a aprendizagem e o bem-estar do aprendente.

Trata-se de conseguir o alinhamento do que está escrito no currículo, do que é ensinado e aprendido e do que é avaliado e classificado, como bem mostram Guskey & Brookhart (2019b e 2019c), Townsley & Buckmiller (2020), Wiliam *et al.* (2020), Brookhart *et al.* (2020), Guskey (2022) e Pasquini (2018, 2019 e 2021).

Estes autores afirmam a necessidade de que a classificação se inscreva numa forte relação entre o ensino e a aprendizagem, a avaliação formativa e a avaliação sumativa, pressupondo que essa relação deve ter origem na análise do currículo e ser feita com base nos objetivos de aprendizagem e nos critérios, que devem ser comuns aos três grupos de atividade pedagógica (ensino, aprendizagem e avaliação).

Os referidos autores são, ainda, unânimes ao considerarem que, na gestão curricular que o professor é convocado a realizar, as notas devem emergir como elementos que integram o processo de aprendizagem, aparecendo como resultado e balanço do que foi aprendido, materializado no que os alunos conseguem fazer com os conhecimentos que adquiriram.

É nesta linha de pensamento, e para esclarecer esta articulação necessária, que surge o modelo de alinhamento curricular alargado (Pasquini, 2018, 2019 e 2021).



## **Desafio 2: Relativamente à atribuição de notas aos desempenhos dos alunos nas tarefas de avaliação sumativa**

Brookhart & Nitko (2008) e Guskey & Brookhart (2019b) recomendam que, para atribuição de notas aos desempenhos dos alunos, nas tarefas de avaliação sumativa:

- i) se usem ferramentas que organizem os critérios e que sejam guias para a atribuição da pontuação (rubricas, listas de verificação ou escalas de classificação);
- ii) se usem escalas curtas, que se fundem com os critérios, uma vez que cada nível é caracterizado e descrito através do domínio de conteúdos e de competências referentes às orientações curriculares;
- iii) se tenha um modelo de resposta como referência para as expectativas;
- iv) se classifiquem as respostas de todos os alunos à mesma questão/proposta de trabalho (classificação analítica) em vez da classificação de todas as respostas do mesmo aluno (classificação global);
- v) se ignorem os autores das respostas quando estamos a classificá-las;
- vi) se trabalhe colaborativamente, trocando alguns dos produtos com colegas e aferindo a aplicação dos critérios usados.

Walvoord & Johnson Anderson (2010) acrescentam um alerta importante, que consiste na necessidade de tornar a classificação dos trabalhos dos alunos mais eficiente em termos de tempo. Para isso, esclarecem que não é preciso escrever comentários que não se orientem para a revisão e melhoria dos trabalhos, que é desejável usar escalas com poucos níveis e recorrer às novas tecnologias, que permitem atingir níveis mais altos de desempenho dos alunos, remetendo para o pensamento complexo (por exemplo, resolução de problemas, pensamento crítico e criativo, ...) e facilitam a correção e a classificação.

## **Desafio 3: Relativamente à decisão sobre as notas a atribuir no final de um período de aprendizagem**

Refletindo sobre este processo, Wiliam *et al.* (2020), Brookhart *et al.* (2020) e Guskey (2022) consideram três condições para que as notas sejam úteis aos alunos: devem ser atribuídas aos desempenhos, não aos alunos; não se reportarem às capacidades dos alunos, nem ao seu potencial de aprendizagem; devem fornecer indicação sobre quão perto ou longe os alunos estão de alcançar as intenções de aprendizagem (o que é que quero/preciso de aprender?).



Por isso, é importante basear as notas num conjunto de evidências, que permitam tirar conclusões claras (sem enviesamentos) sobre o que os alunos são capazes de dizer ou fazer, as quais devem ser recolhidas em diferentes momentos e agregadas, a partir de processos diversificados e adequados, tendo em conta o princípio da triangulação (Fernandes, 2021).

Acresce que as notas devem ser assumidas como temporárias, uma vez que o nível de desempenho dos alunos nunca é permanente e as aprendizagens têm um caráter evolutivo. Neste sentido, ter uma nota baixa apenas significa que o seu desempenho ainda não conseguiu atingir a qualidade necessária e que há trabalho a fazer, uma vez que, à medida que estudam e exercitam, a compreensão dos alunos cresce e o seu desempenho melhora.

Emerge aqui o que no Texto de Apoio n.º 7 (Braga, F., Machado, E. A. & Candeias, F., 2022) dissemos sobre o conceito de critério: uma qualidade que, servindo de base a uma apreciação, está relacionada com a competência, é geral e, se possível, comum a diferentes competências e a várias disciplinas; o ponto de vista sobre o qual nos colocamos para avaliar e tomar uma decisão, que é contextualizado e operacionalizado por indicadores, estes da ordem da situação, observáveis e precisos (Roegiers, 2020).

Assim, Pasquini (2021) apresenta uma organização do que, na literatura da especialidade tem surgido, para sintetizar as diferentes formas de utilização dos critérios (Brookhart, 2013 e 2018):

- Orientados para a tarefa (nas *task-specific rubrics*) destinadas a monitorizar um desempenho específico;
- Orientados para a análise das aprendizagens (nas *analytic rubrics*), que desagregam os resultados evidenciados no desempenho dos alunos nas suas diferentes componentes, o que exige a utilização de diferentes critérios (três ou quatro), articulados entre si, para não se perder a perspetiva global da aprendizagem;
- Utilizados de forma holística (nas *holistic rubrics*), que fazem apelo ao juízo profissional do professor, pois os diferentes critérios são usados simultaneamente, englobando as diferentes dimensões do produto esperado ou da aprendizagem avaliada.

Daqui resulta que, atualmente, os investigadores aqui referidos são unânimes ao considerarem que a classificação resulta de um processo assente em dez princípios, tal como aparecem identificados por Brookhart (2020):



1. Promover avaliações sumativas que traduzam o que os alunos aprenderam com as orientações e explicações dos professores e as tarefas realizadas;
2. Privilegiar, para o registo e tratamento de dados, instrumentos referidos a critérios;
3. Atribuir notas que correspondam aos desempenhos e às aprendizagens realizadas, que estejam de acordo com os objetivos de aprendizagem e que remetam para o nível de desempenho dos alunos quando a nota é atribuída;
4. Assumir que as notas são temporárias, sujeitas a revisão, sempre que há evidências mais recentes dos progressos realizados nas aprendizagens;
5. Flexibilizar os momentos de recolha de evidências de avaliação sumativa orientada para a classificação, para que os alunos sejam avaliados pelo que sabem fazer e não pelo momento em que o fazem, o que implica:
  - Implementar práticas de avaliação formativa, com diferentes propostas pedagógicas, para proporcionar tempo para que os alunos aprendam e pratiquem tanto quanto necessitam;
  - Promover a reformulação e a melhoria dos trabalhos, valorizando avaliações assentes em *feedback* informal, oral e fornecido na hora;
  - Recolher evidências de avaliação sumativa orientadas para a classificação apenas quando os alunos atingem o nível de desempenho;
  - Partilhar regularmente os resultados com os alunos.
6. Diversificar as formas de recolha de informação (ao nível dos procedimentos, técnicas e instrumentos), de acordo com as características dos diferentes domínios;
7. Definir tarefas que exijam pensamento complexo, com base na reflexão, planeamento e aplicação de conhecimentos ou demonstração de competências;
8. Alinhar a tarefa, os critérios de avaliação e os objetivos de aprendizagem;
9. Definir regras de avaliação, para as dinâmicas de grupo/trabalho de grupo, atendendo ao princípio de que as classificações devem traduzir as aprendizagens individuais;
10. Usar critérios de aprendizagem que indiquem a realização de objetivos de aprendizagem, que sejam adequados aos alunos e úteis para a sua autorregulação, ao longo do trabalho.



## Conclusões

“A classificação é, atualmente, um dos assuntos mais calorosamente debatidos em educação (...). Infelizmente, as práticas de classificação atuais baseiam-se muito mais na tradição do que na investigação.”

(Guskey & Brookhart, 2019a).

Apesar de ainda se saber pouco acerca de como é que os professores tomam as decisões classificatórias e, concomitantemente, sobre o significado das notas atribuídas, Brookhart, Guskey, Bowers, McMillan, Smith, L. Smith, Stevens e Welsh realizaram, em 2016, um estudo sobre o significado das notas na sala de aula, através de uma meta-análise de investigações sobre práticas de classificação nos últimos 100 anos (Brookhart *et al.*, 2016).

Através dos diversos estudos revistos, constatam que:

- as práticas de classificação variam entre turmas e professores e isso é sentido, por alguns professores, como práticas justas de avaliação;
- os professores usam medidas de desempenho, principalmente testes, como os principais determinantes das notas;
- e incluem nas notas fatores não cognitivos, como habilidade, esforço, melhoria, conclusão do trabalho e outros comportamentos dos alunos.

Concluem, assim, afirmando que as práticas de classificação não são consistentes entre os professores, no que diz respeito ao peso dos fatores não cognitivos, pelo que as notas atribuídas refletem tanto as diferenças de crenças e valores dos professores como as aprendizagens dos alunos.

Assumir a classificação como fazendo parte de um processo, tal como na literatura da especialidade aparece fundamentada e este texto apresenta, implica romper com o paradigma tradicional, em que a classificação tem uma referência normativa.

Trata-se de perceções que têm que ser conhecidas e debatidas, para que as práticas de classificação passem a basear-se mais nos dados fornecidos pela investigação e contribuam para garantir a efetividade do ensino e da aprendizagem.



## Referências

- Braga, F., Machado, E. A., & Candeias, F. (2022). *Avaliação referida a critérios: perspectivas de concepção e utilização*. Texto de apoio à formação - Projeto de Monitorização Acompanhamento e Investigação em Avaliação Pedagógica (MAIA). Ministério da Educação/Direção-Geral da Educação.
- Brookhart, S. (2013). *How to create and use rubrics for formative assessment and grading*. ASCD.
- Brookhart, S. (2017). *How to use grading to improve learning*. ASCD.
- Brookhart, S. (2018). Appropriate Criteria: Key to Effective Rubrics. *Frontiers in Education*, 3:22. doi: 10.3389/feduc.2018.00022.  
<https://www.frontiersin.org/articles/10.3389/feduc.2018.00022/full>.
- Brookhart, S. (2020). *What Grading and Assessment Practices Could Schools Use in the Year Ahead?* Center on Education Policy, Equity and Governance.  
[https://edpolicyinca.org/sites/default/files/2020-08/pb\\_brookhart\\_sept20.pdf](https://edpolicyinca.org/sites/default/files/2020-08/pb_brookhart_sept20.pdf)
- Brookhart, S., & Nitko, A. (2008). *Assessment and Grading in Classrooms*. Pearson Prentice Hall.
- Brookhart, S. et al. (2016). A century of grading research: Meaning and value in the most common educational measure. *Review of Educational Research*, 86(4), 803-848.  
<https://doi.org/10.3102/0034654316672069>.
- Brookhart, S. et al. (2019). *The Future of Assessment Practices: Comprehensive and Balanced Assessment Systems*. Recommendations from the National Panel on the Future of Assessment Practices. LSI.
- Brookhart, S. et al. (2020). Eight Essential Principles for Improving Grading. *Educational, School, and Counseling. Psychology Faculty Publications*. 50.  
[https://uknowledge.uky.edu/edp\\_facpub/50](https://uknowledge.uky.edu/edp_facpub/50).
- Fernandes, D. (2021). *Avaliação pedagógica, classificação e notas: perspectivas contemporâneas*. Folha - Projeto de Monitorização Acompanhamento e Investigação em Avaliação Pedagógica (MAIA). Ministério da Educação/Direção-Geral da Educação.
- Guskey, T. (2020a). Grading policies that work against standards ... and how to fix them. *NASSP Bulletin*, 84 (620), 20-29.



- Guskey, T. (2022). Can grades be an effective form of feedback? *Phi Delta Kappan*, 104 (3), 36-41. <https://kappanonline.org/grades-feedback-guskey/>.
- Guskey, T., & Brookhart, S. (2019a). Introdução. Guskey, T. R., e Brookhart, S. M. (Eds.) (2019). *What we know about grading: What works, what doesn't, and what's next* (pp 1-12). ASCD.
- Guskey, T., & Brookhart, S. (2019b). Reliability in Grading and Grading Scales (cap. 1). Guskey, T. R., e Brookhart, S. M. (Eds.) (2019). *What we know about grading: What works, what doesn't, and what's next*. ASCD.
- Guskey, T. & Brookhart, S. (2019c). Where do we go from here? Guskey, T. R., e Brookhart, S. M. (Eds.) (2019). *What we know about grading: What works, what doesn't, and what's next* (214-224). ASCD.
- Yerly, G. (2022). Comment la planification peut-elle permettre de passer d'une évaluation de l'apprentissage à une évaluation pour l'apprentissage ? *CNESCO, Conférence de consensus du CNESCO l'évaluation en classe, au service de l'apprentissage des élèves : Notes des experts* (pp. 107-113). Cnesco-Cnam
- Link, L., & Guskey, T. (2022). Is standards-based grading effective? *Theory Into Practice*. DOI: 10.1080/00405841.2022.2107338.
- Muñoz, M., & Guskey, T. (2015). Standards-based grading and reporting will improve education. *Phi Delta Kappan*, 96 (7), 64-68.
- Pasquini, R. (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 : Univ. Genève, 2018, no. FPSE 704. DOI : 10.13097/archive-ouverte/unige :106442 URN : urn:nbn:ch:unige-1064421.
- Pasquini, R. (2019). Élargir conceptuellement le modèle de l'alignement curriculaire pour comprendre la cohérence des pratiques évaluatives sommatives notées des enseignants : enjeux et perspectives. *Mesure et évaluation en éducation*, 42(1), 63–92. <https://doi.org/10.7202/1066598ar>.
- Pasquini, R. (2021). *Quand la note devient constructive: Évaluer pour certifier et soutenir les apprentissages*. Les Presses de Université Laval.
- Roegiers, X. (2020). L'évaluation des étudiants dans une logique de compétences : enjeux et démarches. INSPE Lorraine. Townsley, M. (2022). *Using Grading to support student learning*. Routledge.



Townsley, M., & Buckmiller, T. (2020). Losing As and Fs: What works for schools implementing standards-based grading? *Educational Considerations*: Vol. 46: No. 1. <https://doi.org/10.4148/0146-9282.2204>.

Walvoord, B. E., & Johnson Anderson, V. (2010). *Effective Grading. A Tool for Learning and Assessment*. Jossey Bass.

William, D. et al. (2020). *Grading in a Comprehensive and Balanced Assessment System. Recommendations from the National Panel on the Future of Assessment Practices*. LSI.