

AVALIAÇÃO DO DESEMPENHO DOCENTE: fiabilidade do instrumento de registro utilizado na observação de aulas

Rui Manuel Fernandes Ferreira¹
Ágata Cristina Marques Aranha²

Resumo:

Integrado numa investigação bem mais abrangente efetuada no âmbito da avaliação do desempenho do professor de Educação Física na escola, o presente trabalho visou testar a fiabilidade do instrumento de registro da observação de aulas vigente, pretendendo que o mesmo se constituísse a partir de uma situação a mais aproximada possível da realidade. Aos 4 observadores que integraram a amostra do estudo foi disponibilizado um registro audiovisual de uma aula de Educação Física para que, cada um deles e afastados entre si, observassem a mesma e lhe fosse atribuída uma classificação de acordo com o estabelecido na legislação vigente, a atribuição de um valor numa escala de 1 (insuficiente) a 10 (excelente) para cada um dos quatro parâmetros que integram o instrumento de registro, dois referentes à dimensão científica e os outros dois relativos à dimensão pedagógica. Através da técnica estatística Kappa de Cohen aplicada na análise inter e intra observadores, os resultados evidenciaram valores de fiabilidade baixa e extremamente baixa, considerando, portanto que o instrumento de observação de aulas utilizado na Avaliação do Desempenho Docente não se reveste de um dos requisitos fundamentais exigidos, a sua fiabilidade, ou seja, os resultados alcançados através deste instrumento de observação são inconsistentes, subjetivos e não replicáveis no lugar e no tempo.

Palavras-chave:

Observação de aulas. Fiabilidade. Avaliação do desempenho docente.

EVALUATION OF TEACHER PERFORMANCE: reliability of the recording instrument used in class observation

Abstract:

Integrated into a much broader research performed in the evaluation of the performance of a physical education teacher at the school, the present study aimed to test the reliability of the instrument for registration of classroom observation force, claiming that it is constituted from a situation closer to reality as possible. At 4 observers who joined the study sample was made available a record of an audiovisual Physical Education class for each of them and apart and observe the same, and you were assigned a rating according to the provisions of the current

¹ Professor no Departamento de Desporto, Exercício e Saúde, da Escola Ciências da Vida e do Ambiente na Universidade de Trás-os-Montes e Alto Douro (Portugal). Professor do Quadro do Agrupamento de Escolas Dr. José Leite de Vasconcelos – Tarouca, Coordenador da Área Disciplinar de Educação Física.

² Professora Agregação no Departamento de Desporto, Exercício e Saúde, da Escola Ciências da Vida e do Ambiente, da Universidade de Trás os Montes e Alto Douro (Portugal). Diretora do Departamento de Desporto, Exercício e Saúde, da Escola Ciências da Vida e do Ambiente. E-mail: aranha@utad.pt.

legislation, assigning a value on a scale of 1 (poor) to 10 (excellent) for each of the four components instrument comprising the registration regarding the two scientific dimension and the other two on the teaching dimension. Through the technique of Cohen's Kappa statistic applied in analyzing inter and intra observers, the results showed low values of reliability and extremely low considering so that the classroom observation instrument used in the Teacher Performance Assessment does not put on one of the fundamental requirements required, its reliability, seems that results obtained through this method of observation is inconsistent subjective and non-replicable in place and time.

Keywords: Classroom observation. Reliability. Evaluation of teaching performance.

Introdução

Como se encontra explícito no Joint Committee on Standards for Educational Evaluation (1981), todo o empreendimento humano é passível e desejável ser avaliado. É um facto hoje verificar o desenvolvimento da avaliação nas mais variadíssimas áreas sociais, na economia, nas finanças, na justiça, na segurança social, na saúde, no trabalho e, por conseguinte, também na educação.

Neste alinhamento, entendemos, pois que o XVII Governo Constitucional de Portugal, no período da sua legislatura que decorreu de 2005 a 2009, teve não só o direito, mas, e acima de tudo, o dever de implementar na classe docente um modelo de avaliação que, ao contrário do anterior, viesse diferenciar as boas das más práticas, por outras palavras, os bons dos maus professores.

O referido modelo de avaliação encontra na atual legislação a observação de aulas (Artigo 18º do Decreto Regulamentar n.º 26/2012, de 21 de fevereiro, 2012) como instrumento de registro para avaliar a competência científica e pedagógica do professor. Integrada na avaliação externa, o instrumento de observação obedece ao estipulado no guião de observação da dimensão científica e pedagógica, na classificação da observação de aulas e nos parâmetros científicos e pedagógicos e níveis de desempenho, respetivamente anexo I, II e III do Despacho n.º 13981/2012, de 26 de outubro (2012).

Como refere Estrela e Nóvoa (1999, p. 9), logo na nota de apresentação da obra, cito, “a avaliação é, atualmente, uma área de enorme complexidade técnica e científica, o que dificulta a construção de modelos pertinentes de intervenção”, fim de citação. O Conselho Científico para a Avaliação dos Professores (2008), alerta igualmente para a complexa e morosa tarefa que é a produção de instrumentos de aferição fiáveis e de reconhecida credibilidade científica, sugerindo a sua elaboração por instâncias competentes e alheias ao

processo de avaliação docente. Para além da complexidade enunciada, o treino dos observadores é visto como uma das variáveis mais importantes a ter em conta se pretendemos rigor e credibilidade em todo o processo (Estrela, 1994) (Peck & Tucker, 1973). Por último, o modelo e respectivos instrumentos de avaliação que venham a ser adotados devem salvaguardar as especificidades das áreas disciplinares dos professores (Ferreira & Azevedo, 2010).

1. Metodologia

1.1 Caracterização da amostra

A amostra foi constituída por 4 observadores do grupo de recrutamento 620, Educação Física do terceiro ciclo do ensino básico e do secundário, todos eles do gênero masculino, 2 deles com catorze anos de serviço docente, 1 com quinze e 1 com treze, com uma média de idades de 38,5 anos de idade (ver tabela 1).

Tabela 1 – Caracterização da amostra

	Género	Idade	Anos de Serviço	Experiência em Observação
Observador 1	Masculino	46	14	Sim
Observador 2	Masculino	37	15	Não
Observador 3	Masculino	36	14	Não
Observador 4	Masculino	35	13	Não

1.2 Procedimentos

Aos observadores foi-lhes solicitado a leitura e posterior interpretação dos articulados legais (Decreto Regulamentar n.º 26/2012, de 21 de fevereiro, 2012) (Despacho n.º 13981/2012, de 26 de outubro, 2012) que, no momento, regiam e orientavam o instrumento de observação referente à observação de aulas, não tendo havido lugar para qualquer discussão ao nível da concordância intra e inter observadores ou treino de observação.

A base que sustenta esta decisão prendeu-se com a prática institucionalizada, ou seja, a ausência de formação específica em observação de aulas por parte daqueles que poderiam assumir a responsabilidade de avaliar a componente científica e pedagógica dos seus pares, tentando assim ajustar a situação projetada à situação real.

Seguidamente, foi disponibilizado aos quatro observadores um registro audiovisual de uma aula de Educação Física para que, cada um deles e afastados entre si, observassem a mesma e lhe fosse atribuída uma classificação de acordo com o estabelecido na legislação vigente, particularmente o previsto no anexo III do Despacho n.º 13981/2012, de 26 de outubro (2012), que obriga a utilização de uma escala de classificação baseada em níveis de desempenho de 1 a 10, quantitativos e qualitativos (1 a 4,9 Insuficiente; 5 a 6,4 Regular; 6,5 a 7,9 Bom; 8 a 8,9 Muito Bom; 9 a 10 Excelente), a preencher em cada uma das dimensões, científica e pedagógica, cada uma delas com dois parâmetros, num total de quatro categorias que o observador tem de classificar. Muito embora o instrumento de registro (Anexo I) integre uma componente qualitativa que enquadra e fundamenta a sua componente quantitativa, só esta última foi tida em conta neste estudo.

No âmbito das normas éticas e morais essenciais que devem nortear os trabalhos de investigação, nomeadamente acerca do consentimento informado do sujeito alvo de investigação, os autores deste trabalho solicitaram com êxito os seguintes consentimentos: (i) dos encarregados de educação dos alunos que integraram a turma alvo de registro audiovisual; (ii) do professor que lecionou a aula nessa mesma turma; (iii) do Diretor do Agrupamento de Escolas onde a aula se desenrolou; (iv) dos quatro observadores que integraram a amostra do estudo.

1.3 Limitações

O estudo efetuado apresenta algumas limitações que, em nosso entender, se prendem com questões tecnológicas e metodológicas, são elas:

- a. A capacidade limitada que o registro audiovisual comporta na percetibilidade total da realidade observada (ANGUERA, 1998);
- b. Os fenômenos de reatividade junto dos alunos da turma observada e respectivo professor que lecionou a aula, bem como, de expectância relativamente aos observadores que integraram a amostra (ANGUERA, BLANCO, & LOSADA, 2001);

c. As ameaças à validade interna, externa e de constructo que se encontram associadas à elaboração e aplicação dos instrumentos de registro (PIÉRON, 1996);

d. A natureza complexa, dinâmica e imprevisível que é reconhecida ao fenômeno do processo de ensino e aprendizagem (HADJI, 1995).

2. Resultados e discussão

Uma vez que o estudo utiliza a técnica de determinação da fiabilidade através do índice de Kappa de Cohen, entendemos de primordial importância esclarecer os seus conceitos. A fiabilidade é definida como a razão entre a variabilidade inter-individual e a variância total das pontuações, expressa por um número entre 0 e 1, não existe fiabilidade e fiabilidade perfeita, respetivamente (BLANCO & ANGUERA, 2003).

Relativamente ao índice estatístico utilizado, o Kappa de Cohen é definido por Pestana e Gageiro (2000) como a proporção de concordâncias observadas, corrigida explicitamente pela proporção de concordâncias aleatórias, sendo que os valores obtidos neste índice confirmam uma excelente taxa de fiabilidade quando se verificam acima do percentil 0,75 ($\square.75$).

Conforme se pode observar através da tabela 2, o estudo contou com a participação de quatro observadores na tentativa de garantir a robustez dos resultados, originando do seu cruzamento 10 valores do Kappa de Cohen, sendo 6 deles referentes à análise inter-observadores e os outros 4 (a cinza) referentes à análise intra-observadores.

Tabela 2 – Matriz dos Resultados (valores do Kappa de Cohen)

Observadores	1	2	3	4
1	.2792			
2	.5760	.6706		
3	.4503	.6374	.6695	
4	.3346	.1129	.2010	.4103

Sem exceção, os resultados evidenciam valores de fiabilidade baixa e extremamente baixa, considerando, portanto, que o instrumento de observação de aulas utilizado na Avaliação do Desempenho Docente não se reveste de um dos requisitos fundamentais exigidos, a sua fiabilidade, ou seja, os resultados alcançados através deste instrumento de observação são inconsistentes, subjetivos e não replicáveis.

Ao analisarmos a tabela 2, os resultados alcançados variam entre o valor .1129 (muito baixo) e .6706 (baixo), respectivamente, obtido na análise inter-observadores relativa aos observadores 2 e 4 e o obtido na análise intra-observador referente ao observador 2. Estes resultados parecem dar razão às preocupações manifestadas pelos diferentes autores com metodologias credenciadas nesta área do saber, a destacar:

a. A aplicação do novo modelo de avaliação veio a confirmar-se como um autêntico fracasso devido, fundamentalmente, por não se encontrar assegurado os requisitos de construção, nomeadamente a sua precisão, credibilidade, validade e fiabilidade, ao nível dos instrumentos de observação (CCAP, 2008).

b. Partindo da ideia que os instrumentos de registro são a base empírica para fundamentar as apreciações relativas ao desempenho docente, a produção destes instrumentos de aferição que sejam, simultaneamente, fiáveis e de reconhecida credibilidade científica é uma tarefa complexa e morosa, a desenvolver por instâncias competentes alheias ao processo de avaliação do desempenho (CCAP, 2008).

c. Posteriormente à recolha dos dados, urge a necessidade de garantir a sua qualidade. Segundo Anguera, Blanco e Losada (2001), esta garantia é conseguida através da

utilização de um dos mais elementares requisitos de controlo, a fiabilidade do instrumento de registo.

d. A avaliação da prestação docente deve utilizar instrumentos válidos e fiáveis, de referencial qualitativo ou quantitativo, onde a discussão acerca do rigor se encontra intrínseco ao próprio processo de avaliação, marcadamente específico e subjetivo (Blanco, Pacheco & Silva, 1988).

e. Segundo Brito (1998), a observação pode assumir um carácter científico desde que, seja formulado um plano de pesquisa, tenha uma estrutura sistemática, apresente um registo de dados metódico e tenha validade e fiabilidade, permitindo desta forma tornar o ato de observar numa ação consciente, intencional, previsível, controlável e eficaz.

f. A avaliação dos profissionais da educação é uma tarefa extremamente difícil pelo que, entre outras razões, deve responder às seguintes questões: que avaliar? como avaliar? quando avaliar? onde avaliar? quem avalia? porquê e para quê avaliar? (Posada, 2009). Relativamente à questão como avaliar? referindo-se aos problemas técnicos, mais concretamente em relação à fiabilidade, os instrumentos de registo devem garantir que se obtêm os mesmos resultados em diferentes circunstâncias de lugar, tempo e características socioculturais.³

g. No âmbito da avaliação do desempenho docente, Fernandes (2009) refere que um dos princípios que a deve reger é o uso de instrumentação válida, fiável e contextualizada.

h. A necessidade dos instrumentos de registo incorporarem standards de rigor ou exatidão, referindo-se às dimensões e parâmetros de avaliação dos professores, prendem-se com a produção de uma informação fiável, representativa e abrangente, que permita interpretações válidas e conclusões justificadas e apropriadas (RODRIGUES, 2009).

i. Enquanto instrumento classificatório, Vieira e Moreira (2011) salientam que a observação de aulas encontra-se sujeita a diversos constrangimentos que podem colocar em causa a sua fiabilidade, destacando, o mito da objetividade na recolha da informação, a influência da disposição do observador no modo como observa, a sua incompetência para observar e as consequências da distância física e psicológica entre observador e observado na construção de perspectivas sobre a aula.

³ Este foi o juízo de valor difundido pelo Conselho Científico para a Avaliação de Professores (2008, p. 1), enquanto órgão consultivo do governo da república, cuja missão passava pela fundamentação e legitimação das suas decisões políticas, cito, “... levar ao conhecimento dos decisores políticos alguns aspectos ligados à aplicação do modelo de avaliação instituído que não se revelam adequados ou exequíveis, sugerindo possíveis soluções”, fim de citação.

j. Referindo-se aos diferentes tipos de instrumentos de recolha de dados através da observação de aulas, Reis (2011) refere que as escalas de classificação, tal como sucede no instrumento de registro atualmente em vigor na avaliação do desempenho docente e usada neste estudo, encerram um grau elevado de subjetividade.

k. Outra das fragilidades apontadas ao modelo de avaliação vigente prende-se com a inadequação dos instrumentos de registro, onde se inclui o instrumento de observação de aulas, face às especificidades das áreas disciplinares que integram o currículo, quer no ensino básico, quer no ensino secundário (Ferreira & Azevedo, 2010). Sabendo de antemão as particularidades referidas à disciplina de Educação Física, principalmente pela integração prioritária do desenvolvimento de capacidades e competências no âmbito do domínio psicomotor, o seu ensino admite, naturalmente, particularidades significativas que não podem passar indiferentes ao processo de avaliação do professor da disciplina.

Considerações finais

Conscientes das limitações e constrangimentos que enformam o registro da observação de aulas (VIEIRA & MOREIRA, 2011), bem como, da dificuldade em definir o objeto de avaliação dos professores que, segundo Hadji (1995), a sua problemática resulta mais da natureza complexa e incerta onde o processo de ensino e aprendizagem se desenvolve do que de problemas técnicos, que para o autor são sempre subalternos e secundários, o presente estudo pretende contribuir para um melhor esclarecimento do assunto tratado e apontar possíveis soluções.

Uma vez conhecidos os valores obtidos através do índice estatístico Kappa de Cohen, os resultados obtidos neste estudo apontam para uma falta de fiabilidade do instrumento de registro utilizado na observação de aulas que atualmente se encontra em vigor, sugerindo para além da necessária e desejável correção técnica, o seu uso para efeitos de uma avaliação formativa, num processo que se pretende ao serviço dos objetivos para o qual foi desenhado, o desenvolvimento profissional e a melhoria da qualidade das aprendizagens, em detrimento do seu uso marcadamente sumativo, com repercussões diretas na gestão dos professores e das suas carreiras profissionais.

No âmbito da correção técnica, a falta de formação e treino dos observadores e o tipo de instrumento que a lei obriga (anexo II do Despacho n.º 13981/2012, de 26 de outubro, 2012) parecem ter funcionado como os fatores principais para a falta de fiabilidade registrada. Relativamente à falta de formação e treino, os requisitos evocados na lei vigente para o exercício da função de avaliador externo, por si só, não garantem a qualidade necessária do avaliador, primeiro porque a avaliação de professores não é a mesma coisa que avaliação dos alunos e, em segundo, porque são muito poucos os professores com características pessoais e profissionais capazes de garantir a máxima imparcialidade possível no ato de avaliar (RUIVO, 2009).

Em relação ao segundo fator, a adoção de uma escala de classificação como instrumento de registro cujas fragilidades são conhecidas por Reis (2011), a saber, o elevado grau de subjetividade que enformam, fundamentalmente por representarem juízos de valor (ex.: adequado, correto, exequível) sem a sua representação objetiva ao nível dos critérios de êxito (FERREIRA & AZEVEDO, 2010) (ex.: uma ocorrência mínima de determinado descritor não é necessariamente negativo, nem será necessariamente positiva uma ocorrência elevado de outro descritor).

Considerando um assunto a tratar na próxima etapa do estudo, a correção técnica aqui evocada deverá ser norteada pelo princípio do pragmatismo e da objetividade possível, incidir sobre o desempenho espontâneo do professor na intervenção pedagógica junto dos seus alunos em contexto natural, salvaguardar a regularidade e o uso com rigor da metodologia na implementação da observação de aulas e proceder à formação e treino dos observadores.

Referências

ANGUERA, M. T. (1998). **Observación en la escuela. Barcelona: Graó.**

ANGUERA, M. T., Blanco, A., & Losada, J. L. (2001). **Disenos observacionales, cuestión clave en el processo de la Metodologia Observacional.** Metodologia de las Ciencias del Comportamiento, 135-160.

BLANCO, A., & Anguera, M. T. (2003). **Calidad de los datos registrados en el ámbito deportivo.** Em A. Hernández Mendo, Psicología del Deporte (Vol. 2). Metodología (pp. 35-73). Buenos Aires: Efdeportes (www.efdeportes.com).

BLANCO, E., Pacheco, J. A., & Silva, B. (1988). **A avaliação do professor.** Revista Portuguesa de Educação, 89-102.

BRITO, A. (1998). **Observação Direta e Sistemática do Comportamento**. Cruz Quebrada: FMH Serviços de Edições.

Conselho Científico para a Avaliação dos Professores. (2008). **Princípios orientadores sobre a organização do processo de avaliação do desempenho docente**. Recomendações N.º 2 / CCAP / 2008.

Decreto Regulamentar n.º 26/2012, de 21 de fevereiro. (2012). **Diploma que regulamenta o sistema de avaliação do desempenho do pessoal docente**. MEC.

Despacho n.º 13981/2012, de 26 de outubro. (2012). **Parâmetros nacionais de avaliação externa, bem como os modelos de referência para os instrumentos de registo a utilizar na observação de aulas**. Gabinete do Secretário de Estado do Ensino e da Administração Escolar.

Despacho normativo n.º 24/2012, de 26 de outubro. (2012). **Despacho que regulamenta o processo de constituição e funcionamento da bolsa de avaliadores externos**. Secretaria de Estado do Ensino e da Administração Escolar.

ESTRELA, a. (1994). **teoria e prática de observação de classes**. porto: porto editora.

ESTRELA, A., & Nóvoa, A. (1999). **Avaliações em Educação: Novas Perspectivas**. Porto: Porto Editora.

FERNANDES, D. (2009). **Para uma avaliação de professores**. ELO 16 - Avaliação do Desempenho Docente, 19-23.

FERREIRA, P., & Azevedo, I. (2010). **Instrumentos de observação de aulas: da opressão à emancipação? Avaliação e Currículo: Actas do 22º Colóquio Internacional da ADMEE-EUROPE** (pp. 698-725). Braga: Universidade do Minho.

HADJI, C. (1995). **A avaliação dos professores. Linhas directivas para uma metodologia pertinente**. Em A. Estrela, & P. Rodrigues, Para uma Fundamentação da Avaliação em Educação (pp. 27-36). Lisboa: Edições Colibri.

JOINT COMMITTEE ON STANDARDS FOR EDUCATIONAL EVALUATION. (1981). **Standars For Evaluation of Educational Programs, Projects and Materials**. New York: McGraw-Hill Company.

PECK, R., & Tucker, J. (1973). **Research on teacher education**. Em R. Travers, **Handbook of research on teaching** (pp. 940-978). Chicago: Rand McNally.

PESTANA, M. H., & Gageiro, J. N. (2000). **Análise de dados para ciências sociais - A complementaridade do SPSS**. Lisboa: Sílabo.

PIÉRON, M. (1996). **Formação de Professores. Aquisição de Técnicas de Ensino e Supervisão Pedagógicas**. Lisboa: Faculdade de Motricidade Humana. Universidade Técnica de Lisboa.

POSADA, J. A. (2009). **La evaluación de los profesores: Trivializar la evaluación o evaluar lo trivial?** Em J. Ruivo, & A. Trigueiros, Avaliação de Desempenho de Professores (pp. 79-89). RVJ, Editores.

REIS, P. (Junho de 2011). **Observação de Aulas e Avaliação do Desempenho Docente.** Cadernos do CCAP - 2. Conselho Científico para a Avaliação de Professores.

RODRIGUES, P. (2009). Recensões. Sísifo / Revista de Ciências da Educação, 115-118.
Ruivo, J. (2009). **Formação, Avaliação e Autonomia Docente.** Em J. Ruivo, & A. Trigueiros, A Avaliação do Desempenho de Professores (pp. 33-44). RVJ, Editores.

VIEIRA, F., & Moreira, M. A. (Abril de 2011). **Supervisão e avaliação do desempenho docente: Para uma abordagem de orientação transformadora.** Cadernos do CCAP - 1. Conselho Científico para a Avaliação dos Professores.

