

# PROVAS DE AVALIAÇÃO AFERIDA EM EDUCAÇÃO FÍSICA

João Comédias

Esc. Sec. D. Luísa de Gusmão

Nos últimos quinze anos, participámos numa mudança importante na realidade e no Projecto de Educação Física Escolar em Portugal.

No início deste período, em 1988, o processo dos Programas Nacionais de Educação Física, estabeleceu um novo quadro de referência para o “desenvolvimento curricular baseado na escola” e para as orientações curriculares de âmbito nacional.

As inovações que o modelo do programa de E.F. integra e a adequação à escola, conjugando «desenvolvimento» e «exequibilidade» (Bom et al, 1990) tornaram-no numa excelente referência para a desejável articulação entre o currículo nacional e aquilo que são as responsabilidades das “estruturas pedagógicas” da escola e a deliberação de cada professor.

No domínio da compreensão da importância da Educação Física (E.F.) e dos seus valores educacionais, os Programas Nacionais iniciaram uma mudança, a que não é alheia uma pedagogia social praticada pelos representantes do movimento associativo e também pelos professores de E.F., em muitas escolas, no contacto sistemático com restantes colegas e encarregados de educação.

No entanto, um importantíssimo factor de desenvolvimento não tem acompanhado estas interessantes mudanças — a área da Avaliação em Educação Física.

O argumento que tem sido avançado para a discriminação e a desvalorização da Avaliação em E.F., na Avaliação de Frequência e nas Provas Globais, é esclarecedor — a grande heterogeneidade de recursos materiais para as aulas.

No entanto, face ao considerável investimento na construção de instalações para a Educação Física, o argumento da “diversidade de condições das escolas” se mostra, cada vez mais, incapaz de justificar a anulação da E.F. na avaliação dos alunos e do sistema de educação escolar.

De facto, cada vez mais, a construção de instalações para a Educação Física nas escolas tem sido uma área de considerável investimento. Apesar das carências qualitativas, quer das instalações cobertas, quer dos espaços de aula exteriores além da degradação de muitos desses recursos, o parque escolar tem evidenciado uma mudança quantitativa.

Também no lançamento da Avaliação Aferida a E.F. não tem sido contemplada. Continuamos, assim, sem informações preciosas sobre como vai a Educação Física no país, sobre o aproveitamento que os alunos fazem numa área/disciplina que faz parte do currículo do 1º ano ao 12º ano.

No caso concreto da Educação Física, essa responsabilização pela qualidade do serviço prestado emerge como uma necessidade actual. Saber se está a fazer diferença ter boas instalações, tempo para aprender e correctamente distribuído e professores qualificados é algo que não pode ser ignorado. Trata-se de informação necessária para que se possam tomar as medidas adequadas ao desenvolvimento e ao bom funcionamento do sistema público de educação. «*De facto, são as aprendizagens dos alunos que provam as vantagens das instalações e da formação dos professores*» (Carta Aberta do CNAPEF, 2002).

Estamos perante uma situação em que a ausência de dados sobre as aprendizagens e os progressos reais dos alunos, impede uma caracterização rigorosa e objectiva da população escolar, limitação que as Provas Aferidas podem resolver ou atenuar.

O objecto deste estudo situa-se, precisamente, na área da Avaliação em Educação Física — pretende-se com este trabalho chegar a um modelo de Provas Aferidas, que permita construir as situações de prova mais adequadas para avaliar as aprendizagens dos alunos.

*«Importa contribuir para que a avaliação externa (aferrida) venha a constituir um instrumento de avaliação do currículo, da qualidade de ensino e da aprendizagem dos estudantes assim como um veículo que pode induzir, nas escolas, novas formas de ensinar e de avaliar»* (Fernandes, 1994).



No actual plano de organização curricular do ensino básico, e do ensino secundário, o estatuto da Educação Física é equivalente ao das outras áreas disciplinares. Isso não significa que as suas especificidades possam ser ignoradas ou, pelo contrário, justificar medidas de gestão que desvalorizem, na prática, esta área.

Nessa estrutura de avaliação, embora equivalente, tem que se considerar a especificidade da E.F., que a valoriza no conjunto das áreas e disciplinas curriculares.

Neste momento, em que as mudanças estruturais no Sistema Educativo se dão no âmbito do compromisso político sobre o currículo e a equidade na educação básica, torna-se cada vez mais relevante avaliar as competências em todas as áreas do currículo.

É na clarificação desta questão que reside, no essencial, a ideia chave deste estudo, que se orienta no sentido de responder à seguinte questão:

Como organizar as Provas Aferidas em Educação Física?

Esta questão especifica-se em duas outras:

Elenco — Como definir o elenco das matérias das provas de Avaliação Aferida da Educação Física, por Ciclo (referência aos Objectivos gerais), de âmbito nacional?

Segundo o modelo curricular de EF, cada área tem objectivos gerais a realizar em cada ciclo, segundo as matérias nucleares e os níveis de especificação indicados.

O problema que importa resolver é o da metodologia de definição dessa matriz, especificamente a definição do conjunto de áreas e matérias a avaliar.

Trata-se de um problema de selecção, visto que não é possível avaliar todo o espectro de matérias.

Jogos Colectivos — Como organizar as Provas de Avaliação Aferida dos Jogos Desportivos Colectivos, matérias que se caracterizam pela performance colectiva e pela incerteza das situações de jogo?

De que modo estas diferenças e especificidades poderão determinar que os Jogos Desportivos Colectivos se constituam como um referencial seguro para a construção de provas aferidas em Educação Física? Este é um problema tanto mais importante quanto sabemos que os JDC constituem a principal área dos Programas de E.F.

Conseguir determinar aquilo que é crítico para avaliar pode não ser suficiente para ter sucesso. Muitas vezes, os contextos em que se pretende avaliar não proporcionam essa oportunidade. Alguns comportamentos podem não ser observados pela qualidade superior das oposições ou pelas diferenças significativas entre colegas de equipa.

Para complicar isto o jogo nunca se repete. As situações de jogo são sempre diferentes, mesmo que apresentem semelhanças e as acções a observar podem não ocorrer com a frequência necessária ou nas situações que mais nitidamente as revelam.

Poderemos então constatar, que para além dos mesmos problemas de observação e recolha de informação, que as matérias com habilidades “fechadas” têm, os Jogos Desportivos Colectivos têm outros bem mais difíceis de resolver e que as outras matérias não têm.

Em termos de objectivo para a nossa pesquisa, interessa-nos, na resposta às questões acima, definir as linhas metodológicas fundamentais para organizar as Provas Globais e as Provas de Avaliação Aferida.

Assim, constituímos um painel de especialistas para responder em três fases, às questões da necessidade e organização das Provas Aferidas em Educação Física.

Depois procedemos a um estudo quasi-experimental, de observação vídeo, em que investigámos a capacidade de classificar situações de jogo por professores especialistas e professores estagiários.

## **INQUÉRITOS POR QUESTIONÁRIO A PAINEL DE ESPECIALISTAS EM EDUCAÇÃO FÍSICA**

Objectivos:

1— Construir um entendimento colectivo sobre a necessidade de realização das Provas Aferidas em Educação Física.

2– Apurar qual o entendimento, percepção e tendências dos especialistas sobre as Provas Aferidas em Educação Física, particularmente no que respeita ao seguinte:

- Pertinência da avaliação aferida em Educação Física.
- Selecção das matérias das Provas Aferidas, por ciclo.
- Organização das Provas Aferidas dos Jogos Desportivos Colectivos.

Na constituição do painel de especialistas não procurámos, assumidamente, satisfazer os requisitos de uma amostra representativa, mas somente adequar a amostra aos objectivos do estudo.

A constituição do Painel de especialistas obedeceu, principalmente, a três critérios:

- Qualidade profissional dos inquiridos, segundo a nossa opinião pessoal.
- Ocupar posição de responsabilidade na área da especialidade.
- Facilidade de obtenção de contactos para rápido envio de correio.

O painel de especialistas tinha a composição que passamos a apresentar:

Professores da U.L.H.T.	12
Professores da F.M.H.	2
Orientadores de Estágio da U.L.H.T.	6
Orientadores de estágio da F.M.H.	4
Coordenadores de Departamento	2
Presidente de Conselho Executivo	1
Membros dos Departamentos dos Ensinos Básico e Secundário	2
TOTAL	29



Todos os inquiridos exercem a sua actividade na região de Lisboa.

Um dado relevante prende-se com o facto de quatro destes especialistas serem autores dos Programas Nacionais de Educação Física.

Existe uma quantidade apreciável de professores universitários (12) e existe também uma quantidade apreciável de orientadores de estágio (10).

Este conjunto de inquéritos constitui-se num estudo exploratório sobre as principais tendências dos especialistas sobre uma matéria em que não existe nada publicado na nossa área de especialidade – As provas Aferidas em Educação Física.

Na construção dos inquéritos optamos por perguntas abertas, que permitem uma maior recolha de informação, mais rica e detalhada, apesar de poder funcionar o elemento surpresa com o surgimento de informação inesperada.

Procurámos realizá-los directivamente. As próprias técnicas de inquirição dependem da maior ou menor variabilidade deste factor.

Nalguns casos reduzimos as possibilidades de variação pedindo duas ou três razões, dois ou três conselhos.

Enfatizou-se a importância de se projectarem tendências futuras sobre algo que nunca aconteceu e de se chegar a uma ideia de consenso, a uma convergência de respostas sobre uma matéria em que há um desconhecimento absoluto sobre o que a especialidade pensa.

Partimos da questão geral sobre a importância da Avaliação Aferida em Educação Física para a resposta chave do estudo: Como organizar as Provas Aferidas em Educação Física?

A especificação desta questão em duas outras fundamentais esteve presente ao longo dos inquéritos:

Como definir o elenco das matérias das provas de Avaliação Aferida em Educação Física?

Como organizar as provas nos Jogos Desportivos Colectivos?

Os questionários foram elaborados em sequência, dependendo o número de rodadas do grau de convergência das respostas dadas a cada grupo de questões. A construção dos 2º e 3º questionários foi função das respostas aos anteriores.

A lógica da sequência prende-se com a necessidade de ir construindo o entendimento colectivo de que fizemos referência anteriormente.

Alguns elementos do painel foram contactados previamente pelo telefone para consentirem participar.

Outros elementos, afectivamente mais ligados a nós, receberam, sem qualquer contacto prévio, uma carta que foi enviada para todos os potenciais inquiridos, explicando o objectivo do estudo e as regras de funcionamento do painel. O 1º inquérito foi também enviado em conjunto com a carta.

Nesses contactos explicámos as regras de funcionamento do painel, particularmente a do anonimato (não comunicar com ninguém, nem mesmo com o investigador) e a existência de várias rodadas de inquéritos.

Com os dados do 1º questionário (29 respostas), através dum procedimento estatístico sumário e descritivo, debruçámo-nos sobre a análise de conteúdo das questões.

As conclusões dessa análise sumária (tendências principais dos especialistas) foram enviadas para o painel juntamente com o 2º questionário que repetia algumas e acrescentava novas questões.

A este segundo questionário responderam 27 dos especialistas.

O grau de especificidade das respostas e o consenso sobre as principais questões em análise permitiram construir, logo, o 3º questionário, que foi enviado para o painel.

A este terceiro questionário responderam os mesmos 27 especialistas.

Na posse das respostas a todos os inquéritos codificámos essas mesmas respostas procedendo a uma análise de conteúdo das mesmas e organizámos categorias de acordo com cada resposta.

Procurámos não recorrer a aproximações demasiado interpretativas para decidir em que categoria colocar uma resposta.

Decidimos que cada resposta poderia ser atribuída a mais que uma categoria, já que correspondia a uma possibilidade frequente.

Determinadas questões e preocupações de investigação, como no nosso caso, deram origem a determinadas categorias.

Definimos regras de categorização das respostas, para que as respostas inclassificáveis fossem reduzidas.

Após clarificação das respostas e de acordo com as categorias definidas procedeu-se ao seu tratamento através de frequências relativas e absolutas.

### *1.º Questionário*

---

#### **A Avaliação Aferida deve-se fazer em Educação Física? Porquê? Em que anos? Porquê? Como fazer?**

---

Todos os professores inquiridos (29) consideram que se deve fazer Avaliação Aferida em Educação Física.

A maioria dos professores inquiridos coloca a Avaliação Aferida ao nível da avaliação do Sistema Educativo.

Saber como vai a Educação Física no país, quais os resultados alcançados (20), qual o impacto da aplicação dos programas (7), são as grandes preocupações evidenciadas.

Não se ficando, apenas, pela constatação da realidade, alguns professores também consideram que a análise dos resultados pode permitir a quem decide definir prioridades e tomar decisões sobre que investimentos fazer, ao nível dos recursos e da formação de professores (6).

Outras preocupações que merecem algum destaque são as da responsabilização das escolas e dos professores (7) e a da avaliação aferida poder contribuir para aumentar o estatuto da Educação Física (8).

No entanto, um número apreciável (10) de respostas/razões centra a avaliação aferida na escola e nos problemas gerais da aferição. As respostas podem ter sido dadas tomando, apenas, em consideração a necessidade de melhorar e aperfeiçoar a aferição de critérios entre os professores e não a possibilidade de realizar uma Prova Aferida, propriamente dita.

Os anos de escolaridade mais considerados são todos os anos finais de ciclo (17). Estas respostas mantêm-se no 2º questionário (20). Tal escolha é argumentada com o facto de serem anos que marcam o final de ciclo, segundo a lógica do próprio Sistema Educativo e dos Programas de Educação Física (1º questionário – 35 razões - e no 2º questionário – 16 respostas.

Aliás, os Programas constituem-se como uma grande referência para os professores inquiridos. Isto pode ser constatado ao longo dos três questionários nas respostas a variadíssimas perguntas.

Há, no entanto, respostas (7), que consideram que a escolha dos anos se prende com decisões que a Educação Física tem de tomar no quadro de cada escola. Novamente, não nos parece que

estas respostas estejam centradas na possibilidade de haver uma prova aferida, ao nível do sistema educativo.

Um número apreciável de inquiridos considera a possibilidade da Avaliação Aferida se fazer nos inícios de ciclo (10). Este número de respostas reduz-se no 2º questionário, quando as escolas estão mais fechadas (6).

Com destaque, podemos considerar o facto de nove (9) inquiridos não considerar nas respostas, o 1º ciclo de escolaridade.

Também não consideramos a possibilidade de haver Avaliação Aferida em EF no 1º ciclo, porque não se pode contar com a formação específica em EF dos potenciais avaliadores, os professores do 1º ciclo.

Nos conselhos que os professores apresentam sobre a Avaliação Aferida estão representadas algumas preocupações, que nós temos, com a organização da Prova Aferida, nomeadamente:

- A necessidade de considerar os programas de Educação Física
- Objectivos
- Níveis de especificação das matérias
- Actividade referente (jogo, concurso, composição, coreografia)
- A possibilidade da Prova Aferida se fazer por amostra
- A definição das áreas de avaliação
- A constituição do júri e as preocupações com as metodologias e procedimentos de observação e avaliação
- Os cuidados no treino da aferição/observação e da formação de professores para as Provas
- A utilização do vídeo – registo em imagens do desempenho dos alunos
- A necessidade de preparação prévia da prova e da experimentação prévia do protocolo de aplicação
- A necessidade de avaliar os resultados e tomar decisões.

## 2.º Questionário

---

**Em que escolas realizar as P.A.? Porquê? Em que anos? Justifica**

**Que alunos escolher em cada escola? Porquê?**

**Em que momento do ano? Porquê?**

**Quais as áreas do programa a considerar? Porquê?**

**Quais as matérias a considerar? Porquê?**

**Há áreas do programa mais difíceis de avaliar? Porquê?**

**Há matérias mais difíceis de avaliar? Porquê?**

**A aplicação das P.A. deve ser realizada por quem? Justifica**

**Dois sugestões para a formação dos professores para a aplicação das P.A.**

---

Quanto à escolha de escolas para a Avaliação Aferida os professores estão muito divididos sobre aquelas em que se deve realizar a Avaliação Aferida. “Todas” reúne 16 respostas e “Algumas” reúne 10 respostas.

Concordando connosco, estão os argumentos que consideram que uma amostra de escolas é uma boa solução (11), porque é caro realizar as Provas em todas (4) e as técnicas estatísticas permitem considerar a população alvo a partir do estudo de uma amostra (1).

O conhecimento do retracto do país (6), das diferenças regionais, de se tomar em consideração a heterogeneidade (6) e da necessidade das escolas terem um feed-back dos seus próprios resultados (4) não parecem ser argumentos que possam impedir a estatística, sobre uma amostra representativa, de os resolver.

Pode, inclusive, haver respostas centradas na suposição de que a Avaliação Aferida decide da progressão, ou retenção, dos alunos, não considerando a possibilidade de utilização de uma amostra, por isso mesmo.

O mesmo argumento de que a amostra resolve parece não ser utilizado, por alguns dos mesmos professores que a consideraram antes, quando respondem em relação a que alunos escolher em cada escola seleccionada. Visão global da escola (8), feed-back e reflexão (4), evitar o efeito de variável (5) são agora argumentos utilizados, como se uma amostra de alunos de cada escola não resolvesse, também, o problema da representatividade.

Os poucos que consideram a possibilidade de uma amostra, em cada escola seleccionada (8), admitem que uma visão da escola (2) é possível na mesma e que a amostra é melhor, também, por uma questão económica (3).

Quanto ao momento do ano para aplicar a prova, o final do ano (23) é o melhor momento, porque o programa e os objectivos já estarão cumpridos (15).

Alguns professores (5) consideram ser o início do ano o melhor momento, apesar de três deles também assinalarem o final de ano, porque pode ser associada a Prova Aferida à avaliação inicial dos alunos (3) e porque se pode ver o impacto do ciclo anterior e quais os níveis de entrada no ciclo em causa(2).

Segundo os inquiridos, todas as áreas do programa devem ser avaliadas. Tenham respondido “Todas” (18) ou “Algumas” (8), a maioria das respostas refere-se às “nucleares”. Dominam os argumentos sobre a lógica dos programas, o respeito pela extensão da Educação Física e pela sua característica ecléctica (11).

Quanto às matérias a avaliar devem ser as nucleares (16) porque são obrigatórias e prioritárias (11) e garantem o eclectismo (4).

Alguns consideram que são as opções de escola que devem ser consideradas (4).

Quer se estejam a referir às áreas mais difíceis (18) ou às matérias mais difíceis (19), os Jogos Desportivos Colectivos (9+13) são os mais assinalados. As suas características de incerteza e abertura, a performance individual estar dependente da prestação colectiva e as dificuldades de aferição/observação são as razões mais apontadas.

Não é de desvalorizar alguma referência feita às Actividades Rítmicas Expressivas (5).

Quanto à formação necessária para aplicação das Provas a maioria dos professores (15) considera dever haver uma formação específica (25 a 50 horas) para uniformizar a aplicação do protocolo de aplicação e avaliação (7) e cuidar dos aspectos relacionados com a observação (3).

Aliás, nas sugestões para a formação de professores para a aplicação da prova, a observação e registo (14) cruza todas as categorias consideradas. Os programas Nacionais de E.F. (4), o protocolo de aplicação (4), plano de formação (5) e a formação inicial têm algum relevo. Significativo é o facto de nas categorias “Formação específica” e “Formação especializada” a sugestão de haver um curso creditado seguido de formação nacional “em cascata” ser muito referido (9).

### 3.º Questionário

---

**Que situações de avaliação considerar nos Jog. Desport. Colect.? Porquê?**  
**Características das situações? Porquê?**  
**Como formar os grupos/equipa em cada escola? Porquê?**  
**Quem desempenha a função de árbitro nas situações de jogo? Porquê?**  
**Quem observa e aprecia as situações de prova?**  
**Essa apreciação deve ser validada? Porquê?**  
**Como se deve validar a apreciação da prova?**  
**Que condições de realização deve ter a prova para ser válida?**  
**Como se deve apreciar o nível de jogo? Porquê?**  
**Quais são os níveis a considerar nessa prova de jogo formal? Porquê?**

---

No 3º questionário, o Jogo Formal (8) ou o Jogo Formal com Outra situação (16), que muitas vezes é considerada ser o jogo reduzido (como se este não pudesse ser formal!), são maioritariamente consideradas as situações de avaliação nos JDC, porque correspondem à actividade referente/objectivos terminais (8), porque o programa refere as duas situações (6), se deve considerar o jogo reduzido (6) ou as situações de análise de pormenor (4).

A Outra situação é considerada facilitadora da avaliação/observação de questões de natureza técnica (13) e deve preferencialmente ser jogo reduzido (10) e percursos/exercícios (7).

Muitos professores (15) consideram que os grupos/equipa devem ser homogéneos/grupos de nível. Os principais argumentos para se justificarem as características que os grupos equipas devem ter são: facilita a observação/avaliação (10) e/ou a participação dos alunos (5).

Esta opinião está em desacordo com o que pensamos porque partimos da hipótese de se poder classificar um nível de jogo independentemente dos jogadores intervenientes. Estes revelam, sempre, um produto que é colectivo.

A formação de grupos de nível, que nós defendemos, se estivermos perante uma Prova Global em que se decide da progressão dos alunos (avaliação de natureza pedagógica), é menos económica.

A formação de grupos de nível também não nos dá a garantia de se evitarem resultados enviesados numa prova, como a Prova Aferida, que deve ser feita por amostra.

Esta é uma questão central da organização da prova nos Jogos Desportivos Colectivos e que merece que a investigação apure resultados que possam ser considerados válidos e fiéis.

Existe divisão entre os professores que fazem referência a se os alunos devem ser da mesma turma (4) ou de turmas diferentes (3).

Sobre quem deve desempenhar a função de árbitro há muitos professores (14) que consideram dever ser os alunos porque é uma competência dos programas (avaliação do árbitro) (4) e é importante avaliar as atitudes de quem está a jogar (8).

Um número apreciável de inquiridos (10) considera que devem ser alunos com conhecimentos ou professores para evitar variáveis que possam impedir avaliar (8). Esta é uma posição com a qual concordamos, mas, que merece mais aprofundamento e reflexão. Importa tentar perceber se é exequível avaliar os árbitros sem percas significativas na qualidade do jogo.

A esmagadora maioria considera que são os professores que devem apreciar as situações de prova (26). É importante referir o facto de oito (8) inquiridos escolherem professores da escola e nove (9) inquiridos professores que devem ter formação.

A esmagadora maioria (23) também considera que essa apreciação deve ser validada e que tem que ser baseada na conferência de avaliadores (10), nas normas e critérios comuns (5), na pré-testagem (4) e no registo em vídeo para posterior validação (5).

Para dezanove (19) dos inquiridos a prova deve ser considerada válida se as condições de realização forem idênticas porque a fidelidade/validade (10) e a aferição de critérios (9) são importantes. Há, no entanto, algumas referências à definição dos avaliadores (3), aos critérios de observação (7) e ao registo em vídeo (4).

A avaliação deve incidir sobre o nível de jogo em presença e os aspectos globais (12) e como o jogo é colectivo e há aspectos de cooperação a avaliar (7) a apreciação deve ser sobre o desempenho conjunto das duas equipas em jogo (14) ou sobre cada equipa (11).

Poucos referem exclusivamente cada jogador e o argumento principal é de que a avaliação é individual (4).

Os três níveis dos Programas (I, E e A) são a grande referência (24) porque para os professores inquiridos os PNEFs são a grande referência.

## OBSERVAÇÕES SISTEMÁTICAS EM VÍDEO

Objectivos:

- 1 – Apurar quais os principais problemas de aferição nos Jogos Desportivos Colectivos.
- 2 – Verificar se é possível discriminar níveis de jogo através da observação e classificação directa.
- 3 – Verificar se os especialistas em JDC se distinguem, ou não dos outros professores na classificação directa do nível de jogo.

A composição dos professores que participaram na observação sistemática em vídeo é a seguinte que passamos a descrever:

- Trinta e seis (36) Professores de Educação Física.
- Sete (7) são Especialistas em Andebol, pertencendo seis (6) ao Gabinete Técnico das Selecções Nacionais.
- Onze (11) são especialistas de outros Jogos Desportivos Colectivos.
- Dezoito (18) são professores sem qualquer especialidade nos JDC.
- Noventa e dois (92) alunos do 4º ano da ULHT, já colocados como professores estagiários para o ano lectivo de 2002/2003 (95% do total de estagiários).
- Vinte e nove (29) são especialistas em JDC, dos quais sete (7) no Andebol.
- Sessenta e três (63) não são especialistas em JDC.

Todos os sujeitos (128) exercem a sua actividade quer como professores quer como alunos na região de Lisboa.

Os professores especialistas de JDC têm no seu conjunto uma qualidade reconhecida. Seis são professores na Universidade e dois são autores dos Programas de Educação Física.

Os Especialistas de Andebol têm, supostamente uma elevadíssima qualidade no conhecimento da matéria, já que ocupam posições de destaque na estrutura técnica nacional na Federação de Andebol de Portugal.

A convicção de que a construção de Provas Aferidas em Educação Física tendo como referência os Programas Nacionais de Educação Física se pode configurar a partir de um modelo tendo como base os Jogos Desportivos Colectivos, levou-nos a seleccionar um jogo de progressão no campo, jogado à mão – o Andebol.

Parte-se do pressuposto que se podem identificar níveis de jogo a partir do produto colectivo que duas equipas evidenciam quando os jogadores de cada equipa cooperam entre si e se opõem aos da outra, independentemente do nível de cada um.

Pretende-se verificar se os especialistas se distinguem dos restantes professores e se os níveis de jogo podem, ou não, ser discriminados.

Utilizámos o método de observação sistemática para categorizar segundo parâmetros ou critérios a concordância entre observadores.

Seleccionámos quatro jogos de 10 minutos cada um, utilizando os seguintes procedimentos: Observámos duas aulas de uma turma do 9º ano da Escola Secundária D. Luísa de Gusmão em que a matéria de Andebol tinha tratamento privilegiado.

Identificámos os alunos que cumpriam e os que não cumpriam o nível Introdução do Programa.

Organizámos, criteriosamente, em conjunto com o professor da turma a constituição de equipas para se poderem obter diferentes produtos de jogo.

Gravámos vários jogos de 10 minutos com árbitros que garantissem que o jogo não era desvirtuado pela variável arbitragem.

Procedemos ao visionamento, catalogação e escolha dos jogos.

Seleccionámos um jogo de nível Introdução. Neste jogo os jogadores, manifestamente, não cumpriam os objectivos deste nível do programa (Jogo 3).

Seleccionámos um jogo de nível Elementar. Neste jogo os jogadores cumpriam os objectivos de nível Introdução (Jogo 1).

Seleccionámos um jogo de difícil apreciação (Jogo 4), com uma composição de jogadores em que metade deles cumpria o nível Introdução e, exactamente, outra metade não cumpria o nível introdução.

Tomámos uma decisão sobre qual o produto desse jogo (elementar), baseados em quatro características principais que ele evidenciava:

- Cooperação entre os jogadores de cada equipa.
- Protagonismo dos melhores não era fruto de acções, marcadamente, individualistas.
- Os jogadores de nível introdução não faziam descer claramente a qualidade do jogo porque dominavam, globalmente, a desmarcação.
- Existência de muitos golos.

Gravámos um jogo (Jogo 2) entre elementos da selecção nacional júnior feminino para termos um produto de nível Avançado. As jogadoras cumpriam os objectivos de nível Elementar e também os de nível Avançado do programa. Esta escolha tem haver com o facto de não encontrarmos alunos em número suficiente, na escola escolhida, para procedermos a uma gravação de um jogo com este produto.

Definimos uma ficha de registo da observação.

Definimos um protocolo de observação:

- Preenchimento do cabeçalho da ficha de registo. Depois de preenchida não mais se podia mexer na questão relativa ao conhecimento do programa.
- Vinte minutos para ler o programa de Andebol. Acabada a leitura recolhia-se o documento.
- Esclarecimento do protocolo de observação, particularmente, o impedimento de consulta de qualquer documento e de haver comunicação entre observadores.
- Visionamento, sem interrupção, de 5 minutos de cada um dos quatro jogos.
- Visionamento de 10 minutos do Jogo 1, seguido de 5 minutos para preenchimento da ficha de registo.
- Igual procedimento nos três jogos seguintes, podendo utilizar os 5 minutos de preenchimento da ficha de registo para completar os jogos anteriores.

Para cada jogo era pedido que se identificasse o nível de jogo em presença e que se assinalasse, caso houvesse, um jogador que pertencesse a outro nível de jogo.

Para cada uma das questões foi pedido, também, que explicassem as razões da escolha efectuada. As observações realizaram-se nos dias 9 e 10 de Julho de 2002.

Seguidamente procedeu-se ao tratamento estatístico das respostas.

Fez-se uma análise de conteúdo relativa ás razões das respostas.

Codificaram-se as razões e organizaram-se por categorias, após definirmos regras de atribuição das respostas a uma categoria.

Validaram-se as respostas confrontando-se as opiniões dos especialistas com a nossa própria opinião utilizando o procedimento de evocação da memória, colocando-os perante casos de falsos positivos e falsos negativos.

Após clarificação das razões e de acordo com as categorias definidas procedeu-se ao seu tratamento através de frequências relativas e absolutas.

#### Conhecimento do Programa de Andebol

		Especialistas Andebol	Especialistas JDC	Outros	Total
<b>PROFESSORES</b>	Muito Bem	1	-	-	1
	Bem	2	3	6	11
	+/-	3	5	9	17
	Não	1	3	3	7
	Total	7	11	18	36
<b>ESTAGIÁRIOS</b>	Muito Bem	-	-	-	-
	Bem	1	-	1	2
	+/-	6	16	37	59
	Não	-	4	23	27
	Não respondeu	-	2	2	4
	Total	7	22	63	92

De acordo com os próprios os professores parecem conhecer melhor os programas de Andebol que os estagiários.

Não existe diferença entre os professores especialistas em JDC e os outros professores, apesar dos especialistas em Andebol afirmarem que conhecem um pouco melhor os programas.

Também os estagiários especialistas em JDC afirmam conhecer melhor os programas que os outros estagiários.

#### Erros na Apreciação do Nível de Jogo

	N	JOGO 1	JOGO 2	JOGO 3	JOGO 4
Professores Especialistas	18	0	0	0	2
Professores	18	2	1	0	9
Estagiários Especialistas	29	4	0	0	14
Estagiários	63	4	1	0	29

Há uma elevada percentagem de correcção nas identificações dos níveis nos três primeiros jogos observados:

Jogo 1 – (92,2%) - 10 erros observados

Jogo 2 – (98,5%) - 2 erros observados

Jogo 3 – (100%) - 0 erros observados

Os professores especialistas em Andebol identificam correctamente todos os níveis de jogo observados (100%).

Os outros professores especialistas em JDC também têm uma percentagem elevada de correcção na identificação dos níveis (88,9%). Apenas dois professores não o fazem no Jogo 4, tendo um deles errado não na observação, mas por ter utilizado critérios errados que não estavam de acordo com o protocolo de observação.

Já os restantes professores erram mais, havendo dois que erram no Jogo 1 e um, que repete, errando no Jogo 2. No Jogo 4 há 9 a errar.

Os estagiários têm percentagens de erro semelhantes às dos outros professores não especialistas, menos os estagiários especialistas em Andebol que erram muito menos.

O Jogo 4 de difícil solução porque reúne 50% de jogadores de nível introdução e 50% de jogadores de nível elementar tem uma resposta muito positiva da parte dos professores especialistas em JDC. Tal, dá-nos um indicador muito forte da possibilidade de utilizar este método se os professores conhecerem bem os programas.

Os erros dos restantes professores podem estar associados à conjugação de dois factores, pouco conhecimento dos programas (situação semelhante à dos professores especialistas) e menos conhecimento da matéria.

Os professores especialistas em JDC argumentam mais de acordo com os critérios expressos nos objectivos dos programas quando justificam identificações relativas ao nível introdução do que quando se referem ao nível elementar.

Utilizam poucos argumentos tendo como referência os objectivos dos programas quando justificam aspectos relativos ao nível elementar. Na maior parte das vezes não fazem qualquer referência a competências marcantes deste nível como a exploração horizontal e vertical do espaço do defesa ou a marcação de controlo, que é um elemento identificador e revelador da utilização do contacto físico. Isto é particularmente visível quando fazem identificação de jogadores.

São, também, os que mais identificam que no Jogo 2 já estão cumpridos os objectivos de nível avançado.

Para justificar que o nível elementar está cumprido usam mais vezes os critérios do nível avançado que são visíveis, porque o nível está cumprido, do que os de nível elementar.

Os restantes professores, os estagiários especialistas e os restantes estagiários seguem as principais tendências verificadas nos professores especialistas em JDC.

No entanto, a utilização de argumentos que tomem claramente como referência os objectivos dos programas é sucessivamente reduzida se passarmos dos professores (especialistas ou não) para os estagiários especialistas e ainda mais para os restantes estagiários.

No Jogo 1, os estagiários dão mais respostas que ou não contêm argumentação relevante ou são baseadas noutros critérios que não os que identificam o nível.

Na categoria “outros critérios” verifica-se que os estagiários utilizam critérios relevantes do nível elementar, só que deslocados da observação em causa.

Se a utilização de argumentos que tomem claramente como referência os objectivos dos programas é um indicador forte do conhecimento do mesmo, pode-se afirmar que os professores conhecem mal o programa. Só conseguem percentagens acima dos 50% no Jogo 3 quando justificam o não cumprimento do nível introdução (professores especialistas, outros professores e estagiários especialistas).

Quando há referência ao cumprimento de objectivos de nível elementar, os pouquíssimos que utilizam argumentos que tomem claramente como referência os objectivos dos programas, não fazem referências significativas à defesa (por exemplo, marcação de controle).

A maioria dos professores quando distinguem jogadores (abaixo ou acima do nível geral do jogo) não fundamentam essa distinção em critérios expressos nos objectivos dos programas.

O mesmo já não acontece com os estagiários que, inclusivamente, erram menos.

Em muitas respostas dos professores fica a dúvida se o destaque de jogadores se refere a uma avaliação por comparação com os outros jogadores, ou, se esse destaque representa que o jogador pertence a outro nível do programa, como pretendia o protocolo.

## CONCLUSÕES

Há hoje, em alguns países, uma análise sistemática do aproveitamento escolar dos alunos, ou seja, um apuramento de indicadores para basear o planeamento global a médio e longo prazo.

Esta necessidade dos sistemas públicos de educação pode ser aproveitada para, em processos participados que entusiasmem professores, investigadores e instituições, se garantir uma atitude global de mudança no sistema que promova a qualidade de ensino e que possa influenciar a formação de professores.

De igual forma, as novas práticas de ensino, a avaliação baseada no desempenho como principal referência do processo de avaliação dos alunos, as definições claras do que é o nível de sucesso em cada área/disciplina, são referências que podem ser induzidas nas escolas, através do contributo da avaliação aferida.

O desafio de consolidar e desenvolver a Educação Física, área/disciplina que, com a Língua Portuguesa, faz parte do currículo geral do 1º ao 12º ano e de ampliar seriamente os efeitos educativos do próprio currículo escolar não se projecta ignorando quais as aprendizagens e os progressos reais dos alunos e eternizando a ausência de dados sobre a realidade da Educação Física Escolar.

Os especialistas em Educação Física que constituíram o painel de inquiridos reconhecem a importância de pôr a avaliação aferida na ordem do dia e são unânimes na opinião de que se deve fazer.

Ao justificar tal, com a necessidade de saber como vai a Educação Física no país, o que se ensina e o que se aprende, quais as áreas fortes e fracas, qual a distância a que os alunos estão dos objectivos do programa, estão a desejar uma caracterização objectiva e rigorosa da população escolar, uma avaliação do “sistema de Educação Física” na escola.

Também, nas respostas a várias perguntas, é evidente que os Programas Nacionais de Educação Física são hoje uma referência incontornável. A Lógica do programa, as áreas e os objectivos do programa, as normas de sucesso, os níveis de especificação das matérias, são usados recorrentemente para justificar muitas escolhas e opções.

Apesar do estudo estar circunscrito a uma amostra que exerce a sua actividade profissional na região de Lisboa, não nos parece que, tal, constitua uma limitação séria à generalização destas conclusões. Dez orientadores de estágio, doze professores universitários, membros dos Departamentos dos Ensino Básico e Secundário e autores dos Programas Nacionais são um painel, com toda a certeza, a considerar.

Quanto à definição do elenco das matérias das Provas Aferidas, concluímos que o problema geral da selecção das áreas e matérias passa a deslocar-se, para o problema da selecção de matérias por conjunto de alunos.

De facto, a técnica da amostra multimatricial ao permitir aplicar maior número de itens, avaliando o rendimento de grupos, garante que todas as áreas e matérias do programa possam ser consideradas.

Também o aparecimento das “Normas de Referência para o Sucesso em Educação Física”, que constam dos Programas aprovados recentemente na Revisão Curricular, vem permitir que a dedução dos níveis de sucesso das Provas Aferidas se possa fazer directamente delas.

Igualmente, em Provas em que está em causa uma classificação individual, como é o caso das Provas Globais a selecção das matérias pode ser feita por aluno de acordo com as referidas “Normas”.

Assim, parece-nos que a partir da organização de uma Prova Aferida numa escola ou em escolas escolhidas para o efeito, organizando a amostra e utilizando situações de avaliação de acordo com o que problematizámos, verificar quais os problemas reais de utilização de espaços,

definição de horários, enquadramento de alunos e testar procedimentos de observação e validação das classificações, poderá constituir uma pista de continuidade da investigação.

Quanto ao problema da organização das Provas dos Jogos Desportivos Colectivos, matérias que têm características específicas, e à possibilidade de serem uma referência para a construção de Provas Aferidas em Educação Física, o estudo com base na observação sistemática e objectiva em vídeo teve resultados encorajadores.

A elevada percentagem de correcção nas identificações nos níveis, de três dos quatro jogos considerados, feita pelos participantes na amostra e o facto dos professores especialistas em Jogos Desportivos Colectivos manterem uma percentagem elevada de correcção mesmo no quarto jogo (jogo de apreciação complexa) é um indicador muito relevante da possibilidade de se poder utilizar a apreciação do nível de jogo, como elemento fundamental de avaliação.

No entanto, os sujeitos evidenciaram uma reduzida capacidade de argumentar utilizando os critérios expressos nos objectivos dos programas, o que revela pouco conhecimento dos mesmos.

Os problemas específicos da avaliação nos Jogos Desportivos Colectivos derivados da abertura, incerteza e dependência individual do colectivo, que tornam esta área na mais difícil de avaliar, como os próprios participantes no painel de especialistas reconhece, conjugados com o pouco conhecimento dos programas podem ser factores que determinem a dificuldade de recolher informação relevante.

O grau de discriminação nos resultados entre especialistas, professores e estagiários, e a elevada qualidade dos especialistas em Jogos Desportivos Colectivos, particularmente os especialistas em Andebol, permite-nos concluir que se os professores para além do conhecimento da matéria conhecerem bem o programa, estaremos em condições de usar eficazmente este método de avaliação.

Uma linha de continuidade da investigação poderia aprofundar os resultados estendendo a observação sistemática a outros Jogos Desportivos Colectivos com uma amostra treinada no conhecimento efectivo dos programas das matérias a considerar.

Esta possibilidade poderia estar articulada com o estudo que a organização de uma prova aferida numa escola permitiria.

## Bibliografia

- ALONSO, R.** (2001). *Instrumentos para la Autoevaluación de Instituciones educativas*. Santiago do Chile: Santillana.
- BAKER, E.** (1994). *Making Performance Assessment Work: The Road Ahead*. *Educational Leadership*, 51, 6.
- BAYER, C.** (1985). Pour une Pratique Transférable dans L'Enseignement des Sports Collectifs *apud* **GARGANTA, J.** (1998), Para uma Teoria dos Jogos Desportivos Colectivos. In **A. Graça e J. Oliveira** (Eds.) *O Ensino dos Jogos Desportivos* (pp. 11-25). Porto: FCDEF-UP.
- BAYER, C.** (1994). *O Ensino dos Desportos Colectivos*. Lisboa: Dinalivro.
- BOGDAN, R. & BIKLEN, S.** (1994). *Investigação Qualitativa em Educação. Uma Introdução à Teoria e aos Métodos*. Porto: Porto Editora.
- BOM, L.** (1980). Da Actividade Lúdica à Formação Desportiva. In ISEF (Ed.) *Da Actividade Lúdica à Formação Desportiva: Seminário de Metodologia do Basquetebol* (pp. 19-44). Cruz Quebrada: ISEF.
- BOM, L.; PEDREIRA, M.; MIRA, J.; CARVALHO, L.; CRUZ, S.; JACINTO, J.; ROCHA, L.; COSTA, F.** (1990). A Elaboração do Projecto Programas de Educação Física. *Revista Horizonte*, 6, 35, Dossier.
- BRÁS, J.; MONTEIRO, J.** (1998). A Importância do Grupo para o desenvolvimento da Educação Física. *Revista Horizonte*, 15, 86, Dossier.
- CARR, J. & HARRIS, D.** (2001). Succeeding with Standards. Retirado em 11 de Abril de 2002 da World Wide Web: <http://www.ascd.org/readingroom/books/carr01book.html>
- CASTEJÓN, F.** (1996). *Evaluación de Programas en Educación Física*. Madrid: Editorial Gymnos.
- CIMBRICZ, S.** (2002). State-Mandated Testing and Teachers' Beliefs and Practice. *Education Policy Analysis Archives*, 10, 2.
- CIZEK, G.** (1993). Some Thoughts on Educational Testing: Measurement Policy Issues into the Next Millenium. *Educational Measurement: Issues and Practice*, 12, 3, 10-16.
- CNAPEF** (2000). *Moção sobre Formação aprovada no 5º Congresso Nacional de Educação Física*. Lisboa.
- CNAPEF** (2002). *Carta Aberta: Dez Anos Após a Reforma – Perspectivas para a Educação Física e o Desporto Escolar*. Lisboa: Autor.
- COMISSÃO EUROPEIA – EDUCAÇÃO** (2000). *Relatório Europeu sobre a Qualidade do Ensino Básico e Secundário: Dezasseis Indicadores de Qualidade*.
- CONCEIÇÃO, J.** (1994). Avaliação Aferida – para quê? (I). In IIE (Ed), *Pensar Avaliação, Melhorar a Aprendizagem* (Folha F/1). Lisboa: IIE.
- CONCEIÇÃO, J.** (1994). Avaliação Aferida – para quê? (II). In IIE (Ed), *Pensar Avaliação, Melhorar a Aprendizagem* (Folha F/2). Lisboa: IIE.
- CONCEIÇÃO, J.** (1994). Avaliação Aferida: A Experiência de Outros Países (I). In IIE (Ed), *Pensar Avaliação, Melhorar a Aprendizagem* (Folha F/3). Lisboa: IIE.
- CONCEIÇÃO, J.** (1994). Avaliação Aferida: A Experiência de Outros Países (II). In IIE (Ed), *Pensar Avaliação, Melhorar a Aprendizagem* (Folha F/4). Lisboa: IIE.
- DE LANDSHEERE, G.** (1997). *A Pilotagem dos Sistemas de Educação: Como Garantir a Qualidade da Educação?* Lisboa: Edições ASA.
- DE LANDSHEERE, V.** (1994). *Educação e Formação*. Lisboa: Edições ASA.
- D.G.E.B.S. ME** (1991). *Programa de Educação Física - Ensino Secundário*.
- DINIS, A. & MIRA, J.** (1984). Como "Ler" no Jogo as Prioridades da Formação? Observação do Jogo e Avaliação Formativa. In ISEF (Ed.) *Ensinar, Aprender, Gostar do Basquetebol* (pp. C1-C36). Cruz Quebrada: ISEF.
- FARIA, R. & TAVARES, F.** (1992). O Comportamento Estratégico – acerca da Autonomia de Decisão nos Jogadores de Desportos Colectivos *apud* **GARGANTA, J.** (1998), *Para uma Teoria dos Jogos Desportivos Colectivos*. In **A. Graça e J. Oliveira** (Eds.) *O Ensino dos Jogos Desportivos* (pp. 11-25). Porto: FCDEF-UP.
- FERNANDES, D.** (1992). O Tempo da Avaliação. *Noesis*, 23, 18-21.
- FERNANDES, D.** (1994). Contornos de uma Experiência de Avaliação Desenvolvida no Instituto de Inovação Educacional (1990 – 1993). *Boletim da SPEF*, 10/11, 7-32.
- FERNANDES, D., NEVES, A., CAMPOS, C. & LALANDA, J.** (1996). Das Concepções, Práticas e Organização da Avaliação das Aprendizagens à Formação de Professores. *Relatório de Execução Material* (1º Ano).
- FERNANDES, D.** (1997). Avaliação na Escola Básica Obrigatória: Fundamentos para uma Mudança de Práticas. In Pedro da Cunha (Ed.), *Educação em Debate*. Lisboa: Universidade Católica Portuguesa.
- FINN JR, C., PETRILLI, M. & VANOUREK, G.** (1998). The State of State Standards: Four Reasons why Most "Don't Cut the Mustard". Retirado em 22 de Abril de 2002 da World Wide Web: <http://www.edexcellence.net/library/mustard.html>

- FIRESTONE, W., MAYROWETZ, D. & FAIRMAN, J.** (1998). Performance-Based Assessment and Instructional Change: The Effects of Testing in Mayne and Maryland *apud* **CIMBRICZ, S.** (2002), State-Mandated Testing and Teachers' Beliefs and Practice. *Education Policy Analysis Archives*, 10, 2.
- GARGANTA, J.** (1998). Para uma Teoria dos Jogos Desportivos Colectivos. In **A. Graça e J. Oliveira** (Eds.) *O Ensino dos Jogos Desportivos* (pp. 11-25). Porto: FCDEF-UP.
- GHIGLIONE, R. & MATALON, B.** (1995). *O Inquérito: Teoria e prática*. Oeiras: Celta.
- GIPPS, C.** (1994). *Beyond Testing: Towards a Theory of Educational Assessment*. London: The Falmer Press.
- GIPPS, C.** (1986). A Critique of the APU *apud* CONCEIÇÃO, J. (1994), Avaliação Aferida: A Experiência de Outros Países (II). In IIE (Ed), *Pensar Avaliação, Melhorar a Aprendizagem* (Folha F/4). Lisboa: IIE.
- GRAÇA, A.** (1998). Para uma Teoria dos Jogos Desportivos Colectivos. In **A. Graça e J. Oliveira** (Eds.) *O Ensino dos Jogos Desportivos* (pp. 27-34). Porto: FCDEF-UP.
- GRANT, S.** (2000). Teachers and Tests: Exploring Teachers' Perceptions of Changes in the New York State- mandated Testing Program *apud* **CIMBRICZ, S.** (2002), State-Mandated Testing and Teachers' Beliefs and Practice. *Education Policy Analysis Archives*, 10, 2.
- GRANT, S.** (2001). An Uncertain Lever: Exploring the Influence of State-Level Testing on Teaching Social Studies *apud* **CIMBRICZ, S.** (2002). State-Mandated Testing and Teachers' Beliefs and Practice. *Education Policy Analysis Archives*, 10, 2.
- GRÉHAIGNE, J.** (1992). L' Organisation du Jeu en Football *apud* GARGANTA, J. (1998), Para uma Teoria dos Jogos Desportivos Colectivos. In **A. Graça e J. Oliveira** (Eds.) *O Ensino dos Jogos Desportivos* (pp. 11-25). Porto: FCDEF-UP.
- GUSKEY, T.** (1994). What You Assess May Not Be What You Get. Retirado em 22 de Abril de 2002 da World Wide Web: [http://www.ncacasi.org/documents/other/what\\_you\\_assess.adp](http://www.ncacasi.org/documents/other/what_you_assess.adp)
- HIBBARD, M., WAGENEN, L., LEWBEL, S., WYATT, S., SHAW, S., PELLETIER, K., LARKINS, B., DOOLING, J., ELIA, E., PALMA, S., MAIER, J., JOHNSON, D., HONAN, M., NELSON, D. & WISLOCKI, J.** (1996). A teacher' s Guide to Performance-Based Learning and Assessment. Retirado em 31 de Julho de 2002 da World Wide Web: <http://www.ascd.org/readingroom/books/hibbard96book.html>
- HILL, M. & HILL, A.** (2000). *Investigação por Questionário*. Lisboa: Edições Síbalo, Lda.
- JACINTO, J., COMÉDIAS, J., MIRA, J. & CARVALHO, L.** (2001). *Programas de Educação Física – Ensino Secundário*. Lisboa: Departamento do Ensino Secundário – Ministério da Educação.
- JACINTO, J., COMÉDIAS, J., MIRA, J. & CARVALHO, L.** (2001). Competências Específicas do Ensino Básico: Educação Física. In *Curriculo Nacional do Ensino Básico – Competências Essenciais*. Lisboa: Departamento do Ensino Básico – Ministério da Educação.
- JAMENTZ, K.** (1994). Making Sure that Assesment Improves Performance. Retirado em 22 de Abril de 2002 da World Wide Web: <http://www.ascd.org/readingroom/edlead/9403/jamentz.html>
- KONZAG, I.** (1991). A Formação Técnico-Táctica nos Jogos Desportivos Colectivos *apud* **GARGANTA, J.** (1998), Para uma Teoria dos Jogos Desportivos Colectivos. In **A. Graça e J. Oliveira** (Eds.) *O Ensino dos Jogos Desportivos* (pp. 11-25). Porto: FCDEF-UP.
- KULM, G.** (1991). The Control of Assessment *apud* CONCEIÇÃO, J. (1994). Avaliação Aferida: A Experiência de Outros Países (I). In IIE (Ed), *Pensar Avaliação, Melhorar a Aprendizagem* (Folha F/3). Lisboa: IIE.
- LEMOS, V., NEVES, A., CAMPOS, C., CONCEIÇÃO, J. & ALAIZ, V.** (1992). *A Nova Avaliação da Aprendizagem: O Direito ao Sucesso*. Lisboa: Texto Editora.
- MADAUS, G.** (1988). The Influence of Testing on the Curriculum *apud* **DE LANDSHEERE, G.** (1997), *A Pilotagem dos Sistemas de Educação: Como Garantir a Qualidade da Educação?* Lisboa: Edições ASA.
- MADAUS, G. & KELLAGHAN, T.** (1992). Curriculum Evaluation and Assessment *apud* **DE LANDSHEERE, G.** (1997), *A Pilotagem dos Sistemas de Educação: Como Garantir a Qualidade da Educação?* Lisboa: Edições ASA.
- MAHLO, F.** (1974). *L'acte Tactique en Jeu*. Paris: Vigot.
- MAIZE PUBLIC SCHOOLS USD 266.** (2001). *District Assessment Report Fall*.
- MARZANO, R., PICKERING, D. & MCTIGHE, J.** (1993). Assessing Student Outcomes. Retirado em 31 de Julho de 2002 da World Wide Web: <http://www.ascd.org/readingroom/books/marzano93toc.html>
- MARZANO, R.** (1996). Eight Questions about Iplementing Standard-Based Education. ERIC/AE Digest Series. Retirado em 8 de Julho de 2002 da World Wide Web: [http://www.ed.gov/databases/ERIC\\_Digests/ed410230.html](http://www.ed.gov/databases/ERIC_Digests/ed410230.html)
- MILLER, M. & LEGG, S.** (1993). Alternative Assesments in High- Stakes Enviroment. *Educational Measurement: Issues and Practice*, 12, 2, 9-15.
- MONTEIRO, J.** (1993). Características da Avaliação no Âmbito dos Novos Programas. *Revista Horizonte*, 10, 55, 29-33.
- NASBE** (1997). A Primer on State Accountability and Large-Scale Assessment. *Report of the NASBE Study Group on Statewide Assessment Systems*.

- NASBE** (1999). Setting Cut Scores on Large-Scale Assessments. *Issues in Brief*, 19, 2, 1-12.
- NASPE** (1995). *Moving Into The Future: National Standards for Physical Education: A Guide to Content and Assessment*. Reston: Mosby.
- NEVETT, M., ROVEGNO, I., BABIARZ, M. & MCCAUGHTRY, N.** (2001). Changes in Basic Tactics and Motor Skills in an Invasion-Type Game After a 12-Lesson Unit of Instruction. *Journal of Teaching in Physical Education*, 20, 352-369.
- OSLIN, J., MITCHELL, S. & GRIFFIN, L.** (1998). The Game Performance Assessment Instrument (GPAI): Development and Preliminary Validation. *Journal of Teaching in Physical Education*, 17, 231-243.
- PERRENOUD, P.** (1993). Não mexam na minha avaliação! Para uma abordagem sistêmica da mudança pedagógica. In **Estrela, Albano e Nóvoa, António**, *Avaliação em Educação: Novas Perspectivas*. Lisboa: Educa.
- PERRENOUD, P.** (1998). A Avaliação dos Estabelecimentos Escolares: um Novo Avatar da Ilusão Cientificista. *Ideias*, 30, 193-204.
- PERRENOUD, P.** (2000). Réformer l' école sans la briser : de la décision autoritaire au pilotage négocié. *Éducateur*, 8, 40-44.
- PERRENOUD, P.** (2001). Du Pilotage Partagé au Pilotage Négocié. Retirado em 31 de Julho de 2002 da World Wide Web: [http://www.unige.ch/fapse/SSE/teachers/perrenoud/php\\_main/php\\_2001/2001\\_16.html](http://www.unige.ch/fapse/SSE/teachers/perrenoud/php_main/php_2001/2001_16.html)
- PHELPS, R.** (1999). Why Testing Experts Hate Testing. Retirado em 8 de Julho de 2002 da World Wide Web: <http://www.edexcellence.net/library/phelps.htm>
- POPHAM, J.** (2001). Standardized Achievement Tests: Misnamed and Misleading. Retirado em 8 de Julho de 2002 da World Wide Web: <http://www.edweek.org/ew/newstory.cfm?slug=03popham.h21>
- RAVITCH, D.** (1997). National Tests: A Good Idea Gone Wrong. Retirado em 22 de Abril de 2002 da World Wide Web: <http://www.edexcellence.net/library/natttest.html>
- REDDAWAY, J.** (1991). The Mechanics of National Assessment in the U.K. at Age 16 *apud CONCEIÇÃO, J.* (1994). Avaliação Aferida: A Experiência de Outros Países (II). In IIE (Ed), *Pensar Avaliação, Melhorar a Aprendizagem* (Folha F/4). Lisboa: IIE.
- RICHARDS, C.** (1988). Indicators and Three Types of Educational Monitoring Systems *apud DE LANDSHEERE, G.* (1997), *A Pilotagem dos Sistemas de Educação: Como Garantir a Qualidade da Educação?* Lisboa: Edições ASA.
- RINK, J., FRENCH, K. & TJEERDSMA, B.** (1996). Foundations for the Learning and Instruction of Sport and Games. *Journal of Teaching in Physical Education*, 15, 4, 399-417.
- ROVEGNO, I., NEVETT, M. & BABIARZ, M.** (2001). Learning and Teaching Invasion-Game Tactics in 4<sup>th</sup> Grade: Introduction and Theoretical Perspective. *Journal of Teaching in Physical Education*, 20, 4, 341-351.
- ROVEGNO, I., NEVETT, M., BROCK, S. & BABIARZ, M.** (2001). Teaching and Learning Basic Invasion-Game Tactics in 4<sup>th</sup> Grade: A Descriptive Study From Situated and Constraints theoretical Perspectives. *Journal of Teaching in Physical Education*, 20, 370-388.
- STECHER, B.** (2002). Improving District and State Accountability Systems. Retirado em 17 de Junho de 2002 da World Wide Web: <http://www.cse.ucla.edu/CREST/cont00/stecher.ppt>
- TAVARES, F.** (1998). Para uma Teoria dos Jogos Desportivos Colectivos. In **A. Graça e J. Oliveira** (Eds.) *O Ensino dos Jogos Desportivos* (pp. 35-46). Porto: FCDEF-UP.
- TEODORESCU, L.** (1984). *Problemas da teoria e Metodologia nos Jogos Desportivos*. Lisboa: Livros Horizonte.
- THURLER, M.** (1991). And if School Sel Evaluation was already overcome? "State of Art" of School Evaluation in Switzerland *apud PERRENOUD, P.* (1998), *A Avaliação dos Estabelecimentos Escolares: um Novo Avatar da Ilusão Cientificista*. *Ideias*, 30, 193-204.
- VEAL, M.** (1993). The Role of Assessment and Evaluation in Secondary Physical Education: A Pedagogical View *apud OSLIN, J., MITCHELL, S. & GRIFFIN, L.* (1998), *The Game Performance Assessment Instrument (GPAI): Development and Preliminary Validation*. *Journal of Teaching in Physical Education*, 17, 231-243.
- WRIGHT, J. & GIOVINAZZO, R.** (2000). Delphi – Uma Ferramenta de Apoio ao Planeamento Prospectivo. *Caderno de Pesquisas em Administração*, 1, 12, 54-65.