

A ideia de Aptidão Física. Conceito, operacionalização e implicações

José Maia*

A importância do estudo da aptidão física é inquestionável. O seu interesse é por demais evidente em áreas distintas, extravasando o domínio das Ciências do Desporto para se associar ao âmbito médico, por exemplo, à Epidemiologia e a Saúde Pública. Este novo fôlego na pesquisa decorre das posições oficiais de organismos de utilidade reconhecida como são, por exemplo, a Associação Americana de Cardiologia, O Centro de Controlo de Doenças, a Organização Mundial de Saúde ou a Associação Portuguesa de Cardiologia.

Permitam que refira, desde já, e no início desta nossa viagem, quatro exemplos suficientemente esclarecedores da revisitação contínua a este tema, aparentemente simples, mas que teima em ser fonte inesgotável para pesquisadores de diferentes quadros de interesses e formação académica:

1. Em 1987, a prestigiada revista *Research Quarterly*, no seu número 4 do volume 58 dedicava-lhe uma parte substancial do seu conteúdo ao lidar com a temática a partir de enquadramentos pedagógicos e de saúde pública. Os autores convidados para apresentarem as suas posições são bem conhecidos - Sallis, Seefeldt e Vogel, Bar-Or, Corbin e Cureton.
2. Em 1988, o prestigiado Colégio Americano de Medicina Desportiva emitia a sua posição oficial em torno da relevância do estudo e implicações da expressão da Aptidão Física de crianças e jovens.
3. Em 1992, uma vez mais, a revista *Research Quarterly*, no seu número 2 do volume 63 retoma o tema a partir da questão - será que crianças e jovens americanos possuem níveis adequados de aptidão? As implicações decorrentes desta questão, bem como as das respostas avançadas são mais que evidentes. Os autores convidados para se pronunciarem sobre os resultados de pesquisas efectuadas foram, entre outros, Blair, Corbin e Pangrazi, Freedson e Rowland, e Updyke.
4. Em 1995, a revista *Quest*, no seu número 3 do volume 47 avança uma questão altamente relevante que pretende ver resolvida - a da relação entre actividade física, aptidão física e saúde para que convida autores consagrados, como são, por exemplo Blair, Pate, Shephard e Lohman.

Conforme anteriormente referi, em 1988, num documento breve, síntese esclarecida do estado do conhecimento sobre a matéria da aptidão física, bem como das perspectivas futuras da pesquisa neste domínio, o Colégio Americano de Medicina Desportiva emitiu o seu parecer oficial a propósito da importância da expressão da aptidão física em crianças e jovens. No coração deste parecer é evidente o seguinte pulsar: **é inquestionavelmente aceite a importância do desenvolvimento de programas de actividade física e aptidão física em crianças e jovens com o propósito de encorajar a adopção de hábitos apropriados de exercício que se**

* Professor Associado da Faculdade de Ciências do Desporto e de Educação Física da Universidade do Porto.

prolonguem pela vida adulta, e cujos propósitos centrais são a manutenção da capacidade funcional, bem-estar e melhoria da saúde.

Este pulsar, de importância inigualável, tem percorrido o pensamento pedagógico dos profissionais da Educação Física de inúmeros países, e encontra-se consignado nos programas oficiais da disciplina. Portugal não é excepção. Basta uma leitura breve aos documentos programáticos dos vários ciclos de ensino para constatar tal presença. Daqui que um dos pressupostos centrais de toda a acção pedagógica em torno da actividade física e do movimento da Aptidão Física, a designada educação para a aptidão (Safrit, 1995), bem como o que actualmente se designa de educação para novos estilos de vida, possa ser apresentada do seguinte modo:

Atitudes, hábitos e comportamentos de actividade física e aptidão física que se promovam e adquiram cedo transferem-se em forte capital acumulado de um estilo de vida saudável no adulto.

Contudo, permitam que coloque a seguinte questão: onde estão os estudos e os dados que dão força a este forte sentimento que percorre uma parte nuclear da nossa acção educativa?

Tentarei responder, ainda que muito tangencialmente, à questão levantada. A meu ver a resposta tem três níveis de aproximação e que nos conduzem ao coração dos resultados das pesquisas nesta matéria (uma abordagem mais profunda deste assunto é referida em Maia et al., 1998a):

- O **primeiro** situa-se ao nível dos estudos da equipa do Prof. Gaston Beunen da Universidade Católica de Lovaina, na Bélgica, de onde têm saído o maior número de pesquisas neste domínio a nível europeu. Apresentarei quatro trabalhos bem esclarecedores do propósito anteriormente traçado. Importa salientar que esta equipa dirige um vasto projecto de implicações diversas, e que começou quando os sujeitos da amostra tinham 10 anos de idade, e que neste momento possuem 45 anos. Este projecto é conhecido por Estudo Longitudinal de Lovaina relativo ao estilo de vida, aptidão e saúde. Trata-se, pois, de uma pesquisa longitudinal quase que única em todo o mundo. Vejamos, então, e a título de exemplo, quatro trabalhos. Beunen et al. (1990) mostraram o baixo a moderado valor de estabilidade ($r < 0.60$) das diferentes componentes da Aptidão Física num intervalo de idade situado entre os 13 e os 30 anos de idade. Em duas outras pesquisas (Lefevre, 1994; Lefevre et al., 1997; Vanreusel et al.,) relativas ao comportamento da actividade física e aptidão física de sujeitos em pontos extremos da escala verificou-se que os valores de estabilidade eram baixos, $r = 0.34$ a $r = 0.44$. Uma das pesquisas conduzidas por Lefevre et al. (1997) refere que: (1) os hábitos de actividade física em crianças e adolescentes não estavam altamente associadas com o seu estado de aptidão física e nível da saúde na vida adulta; (2) e os hábitos de actividade física na vida adulta eram mais benéficos para afectar o nível de aptidão física e saúde nessa mesma idade. Finalmente, o quarto estudo (Beunen et al., 1992) refere não ser possível prever, com precisão, os valores de AF aos 30 anos a partir dos valores conhecidos na infância.
- O **segundo** nível da resposta encontra-se nas extensas revisões sobre esta temática efectuadas pelo Prof. Robert Malina da Universidade de Michigan, nos EUA (Malina, 1990; 1996). Estes estudos são percorridos por dois elementos vertebradores comuns: (1) a consciência precisa da reduzidíssima pesquisa acerca da estabilidade da AF associada à saúde, e o mesmo se constata no tocante à Actividade Física e que se situem na infância, adolescência e estado adulto; (2) os valores de estabilidade encontrados nas pesquisas disponíveis são baixos-a-moderados, sem grande significado substantivo.

Face a estes resultados parece ser evidente que o primado da acção pedagógica em torno da promoção da actividade física e aptidão física portadores de promessas do seu efeito a longo prazo carece de um forte fundamento empírico. Os valores dos estudos não são nada encorajadores. Parece que ha que rephrasear de forma mais cuidadosa o nosso discurso pedagógico em torno do tema, pelo menos a fazer fé nos resultados da investigação disponível.

- O **terceiro** nível da resposta ao problema (Maia, 1997; 1998a; 1998b; 1998c) mostra um quadro distinto do anteriormente traçado. Uma análise mais severa à estrutura metodológica de alguns dos trabalhos mais importantes nesta àrea mostrou um conjunto variado de insuficiências que parecem limitar o alcance das suas conclusões, pelo menos naqueles que recorreram a procedimntos de auto-correlação. Dentre os problemas mais sérios destacamos: abuso nas interpretações das auto-correlações, restrição na amplitude dos dados, reduzidas dimensões das amostras, efeitos socráticos, valores de r não corrigidos pelas estimativas de fiabilidade, violação da noção de aptidão física e ausência de consideração da estrutura simplex dos dados.

O recurso a procedimentos estatísticos mais poderosos oriundos da Moderna Teoria da Estrutura de Covariâncias, concretamente o uso de estudos de painel auto-regressivos, permitiu evidenciar um quadro algo distinto do anteriormente traçado. O prof. Gaston Beunen permitiu que tivesse acesso aos dados originais da sua pesquisa, e ao reanalizalos, foi possível efectuar uma leitura distinta do problema em análise. Eis pois, uma amostra dos resultados de trabalhos que conjuntamente efectuamos:

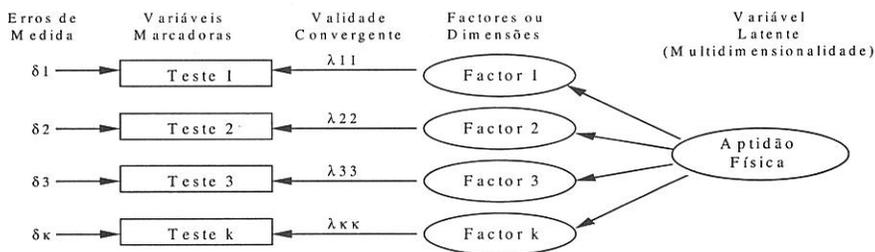
1. A aptidão física, enquanto entidade multidimensional do sujeito é altamente estável entre os 12.75 e os 17.73 anos (Maia et al., 1997). Os valores dos coeficientes de estabilidade da estrutura da AF são de 0.87 (dos 12 para os 14) e 0.84 (dos 14 para os 17 anos). Importa salientar que o valor máximo teórico de estabilidade é de 1. Estes resultados não possuem paralelo na literatura dado que nunca ninguém efectuou este tipo de análise - considerar a estabilidade da totalidade da noção da aptidão física.
2. Um estudo de painel em 100 sujeitos seguidos longitudinalmente dos 11 aos 30 anos de idade e medidos em 7 pontos distintos do tempo (12.7, 13.7, 14.6, 15.6, 16.7, 17.7 e 30 anos) para 6 indicadores de AF, considerados agora individualmente, mostrou a sua elevada estabilidade, situando-se, os valores entre 0.63 e 0.98 (Maia et al., 1998). Estes valores são superiores a todos quantos foram referenciados na literatura.
3. Uma outra pesquisa nos dados da actividade física destes jovens, medidos em 8 pontos do tempo, dos 13 aos 35 anos, revelou, ao contrário do referido na literatura, a sua elevada estabilidade, entre 0.70 e 0.88.
4. Finalmente, numa linha de pensamento semelhante são os resultados apresentados por Lefevre et al. (1997) ao pesquisar o comportamento de grupos extremos. Cerca de 50% dos sujeitos que se localizaram nos quartis extremos quando possuíam 30 anos de idade, mantinham-se nos mesmos quartis aos 35 anos. Mais, os mesmos sujeitos que possuíam uma maior actividade física aos 18 anos de idade, dispendiam significativamente mais tempo em actividades desportivo-motoras aos 30 anos. Significa, pois, que os anos correspondentes à adolescência podem ser considerados como uma espécie de período sensível de impacto substancial na aquisição e manutenção de um estilo de vida activo e saudável.

Em suma, parece ser evidente, pelo menos, a partir desta nova análise dos dados, um quadro distinto do anteriormente traçado. Afinal, o pressuposto inicialmente apresentado, parece possuir

validade, se bem que seja importante referir que se esteja a lidar com dois traços dos sujeitos passíveis de fortes modificações em qualquer ponto do tempo.

Permitam que relembre, uma vez mais, o conceito e a estrutura operativa da aptidão física (Figura 1).

Figura 1 - Operacionalização genérica da ideia de aptidão física



Convém precisar, contudo, que tal conceito multidimensionalmente expresso, e tal como uma moeda, contém duas faces aparentemente divergentes nas suas preocupações e alcances - a que se dirige para o quadro da *performance* desportivo-motora, e a que procura orientar-se para o domínio da saúde. Este carácter bi-fronte da noção de aptidão física expressa, fundamentalmente, uma preocupação didáctica, se bem que procure marcar, tão clara quanto possível, uma fronteira nos horizontes que pretendem alcançar - o da capacidade em realizar trabalho de modo eficiente em condições de exigências energético-funcionais precisas, e o de cunho epidemiológico que deseja colar-se à preocupação emergente do nosso tempo - o da saúde das populações numa óptica essencialmente preventiva. Passemos pois a apresentar cada uma das faces da moeda:

1. No quadro da *performance* desportivo-motora, a aptidão física é entendida como a capacidade funcional múltipla de um dado indivíduo para realizar actividades que exijam empenhamento muscular, ou também, a aptidão individual demonstrada em competições desportivas, sobretudo na capacidade em realizar trabalho (Bouchard e Shephard, 1992; Fleishman, 1964; Safrit, 1990).

Nesta face da moeda é possível encontrar um conjunto alargado de propostas de operacionalização que se encontram formalizadas nas múltiplas e variadas baterias de testes, de que as baterias da AAHPERD (1988), *Eurofit* (1988) ou *Facdex* (1991) são exemplos bem ilustrativos. Esta última bateria, a do FACDEX (Quadro 1) é composta por 9 testes, procura marcar cinco componentes da aptidão física, e tem sido muito utilizada em Portugal, sobretudo no contexto de pesquisas utilizadas por alunos de Mestrado da FCDEF-UP na zona norte do continente e na ilha da Madeira (Freitas, Marques e Maia, 1997).

2. A outra face da moeda, a que procura redimensionar a ideia de aptidão física na busca da sua eventual associação à saúde, é normalmente apresentada do seguinte modo - estado caracterizado por uma aptidão para realizar actividades físicas com vigor, bem como pela demonstração de traços e características que estão intimamente associadas a um risco reduzido de desenvolvimento de doenças designadas de hipocinéticas (Bouchard e Shephard 1992; Morrow, Jackson, Dish e Mood, 1995; Pate, 1988; Safrit, 1990).

Quadro 1 - Estrutura operativa da bateria de testes FACDEX (adaptado de Marques et al., 1991)

Componentes (Dimensões)	Testes
Resistência	Corrida de 12 minutos
Flexibilidade	Sit and reach
Força Superior	Arremesso de um peso de 2 kg Lançamento da bola de hóquei Dinamometria manual
Média Inferior	Sit-up's (60 segundos) Salto em comprimento sem corrida preparatória
Velocidade	Corrida de 50 metros
Coordenação/Agilidade	Corrida vai-vem 10x5 metros

A representação da multidimensionalidade associada a esta noção é bem evidente, por exemplo, em dois trabalhos recentes:

1º O de Corbin (1991) que atribui a este domínio uma hierarquia condicionada pela ideia de desenvolvimento, procurando evidenciar a horizontalidade e verticalidade da expressão da aptidão física com o desenvolvimento de habilidades, integrando-as num modelo coerente e lógico. Contudo, tal proposta, face à sua latitude, afigura-se de difícil operacionalização em estudos de larga escala, dada a ausência de uma bateria de testes que marque cada um dos aspectos das diferentes componentes propostas.

2º No Quadro 2 é apresentada a proposta de Bouchard e Shephard (1992), inquestionavelmente mais complexa e multidimensional onde, por exemplo, não se vislumbram as noções de horizontalidade e verticalidade. Cinco grandes componentes são marcadas por 22 factores.

Ressalta deste quadro o carácter altamente complexo e fortemente inter-dependente das componentes e factores. Este esforço de síntese é deveras louvável pela riqueza do seu alcance. Contudo, torna-se mais que evidente que operacionalizar esta macro-estrutura, no respeito estrito da noção fulcral de parcimónia e do princípio de economia, se torna uma tarefa impossível.

Daqui que a orientação dos autores se tenha concentrado, partindo do respeito estrito das noções de parcimónia e economia, em baterias simples, de aplicação universal, e que pudessem ser utilizadas em estudos essencialmente pedagógicos e epidemiológicos, cobrindo milhares de sujeitos. Exemplos bem sucedidos desta aventura, percorrendo somente algumas das componentes da macro-estrutura anteriormente apresentada, são as baterias *Physical Best da AAHPERD*, a *Chrysler-Fund*, a *Fit Youth Today* e a *Fitnessgram* (Safrit, 1990). Por exemplo, a *Physical Best* é constituída por 4 componentes que são marcadas por cinco itens ou testes (Quadro 3).

Quadro 2 - Componentes e factores da aptidão física associada à saúde (adaptado de Bouchard e Shephard, 1992)

Componentes	Factores
Componente morfológica	Índice ponderal Composição corporal Distribuição da adiposidade subcutânea Gordura visceral abdominal
Componente muscular	Densidade óssea Flexibilidade Potência Força Resistência
Componente motora	Agilidade Equilíbrio Coordenação Velocidade de movimento
Componente cardíaco-respiratória	Capacidade sub-máxima de exercício Potência aeróbia máxima Função cardíaca Função pulmonar Tensão arterial
Componente metabólica	Tolerância à glucose Sensibilidade à insulina Metabolismo lipídico e lipoproteico Características de oxidação de substratos

Quadro 3 - Estrutura operativa da bateria Physical Best (adaptado de Safrit, 1990)

Teste (item da bateria)	Componente da aptidão
Marcha/corrida de 1 milha (ou corrida de 6 minutos)	Capacidade aeróbia
Soma de duas pregas de adiposidade (tricipital e geminal, tricipital e subescapular ou índice de massa corporal)	Composição corporal
<i>Sit and reach</i>	Flexibilidade
<i>Sit-up's</i> modificados	Força e resistência musculares
Elevações na barra	Força e resistência musculares

A avaliação da aptidão física é norteada por duas avenidas de pensamento inequivocamente distintas e de alcances diversos - a normativa e a criterial.

No posicionamento normativo, aquilo que se procura realizar é o seguinte (Maia, 1996): (1) elaborar um conjunto de valores de referência que caracterizam um dado estrato populacional de que as normas percentílicas são um exemplo acabado; (2) traçar o perfil multidimensional de aptidão de um dado sujeito; (3) posicioná-lo no seio do seu grupo de referência; (4) e identificar as diferenças entre sujeitos. O que aqui está presente não é mais do que uma ideia diferencialista bem simples - diferenciar as respostas dos sujeitos num dado conjunto de testes e posicioná-los no seu grupo de referência.

Este tem sido, sem sombra de qualquer dúvida, o percurso da investigação em Portugal. Ainda que timidamente, e às vezes sem grande consistência metodológica, as tabelas percentílicas ou decílicas propostas, não têm servido praticamente para nada, dada a ausência da sua utilização, e quando são utilizadas incorre-se em erros de natureza interpretativa, como seja, por exemplo, o da tentativa de situar médias de grupos em tabelas percentílicas e ajuizar do valor do grupo em termos da sua aptidão por contraste a outros grupos populacionais.

Mas os problemas que gostaria de apresentar relativamente à perspectiva normativa, essencialmente as interpretações e decisões relativas à avaliação a partir do seu posicionamento na tabela, exigem, em meu entender, um extremo cuidado e muita parcimónia, ou então de uma forma mais drástica, abandonar a sua utilização quando for perspectivada no domínio da avaliação associada à saúde. Os motivos que avanço, e que terão oportunidade de ajuizar do seu valor, são os seguintes:

1. As tabelas percentílicas devem provir de amostras representativas, em termos de quantidade e qualidade dos sujeitos amostrados, do universo estudado. Esta exigência tem sido violada face às amostras de conveniência da quase totalidade dos estudos realizados em Portugal, o que enviesza os seus valores, e condiciona a sua interpretação.
2. Deseja-se que as tabelas sejam actualizadas, o que não se faz. Pretende-se, também, que os testes sejam relevantes. Neste último aspecto é forte a discussão entre os autores, dado que a maior parte das propostas de baterias, as construídas em Portugal, e também nos USA, são propostas *had hoc*, baseadas em consensos de peritos cuja validade pode e deve ser questionada. Ao invés desta situação, saliento dois exemplos altamente esclarecedores da processologia de construção e validação de baterias de testes, recorrendo aos procedimentos mais exigentes, e que são as propostas de Simons et al. (1969) na Bélgica e Marsh (1993) na Austrália. Nestas duas pérolas da literatura da avaliação é bem evidente um quadro de vastas implicações que os investigadores portugueses e americanos parece terem esquecido.

A título de mero exemplo, da insuficiência de invariância estrutural de baterias de testes, vamos referir as baterias utilizadas nos Açores (Sobral, 1989) e a do FACDEX (Marques et al., 1991), para avaliar a aptidão física de crianças e jovens. Dado que são baterias *had hoc*, de validade de conteúdo outorgada por painéis de peritos, não é claro, porque nunca foi mostrado, que medissem a mesma estrutura de aptidão física ao longo dos diferentes valores de idade e nos dois sexos. Contudo, não têm deixado de revelar utilidade no contexto de alguns estudos.

3. A ideia de aptidão de um sujeito é sempre representada por um perfil multidimensional, se se quiser manter intacta a própria noção de aptidão. Um teste isolado não possui qualquer representação da totalidade da ideia de aptidão. Daqui que a interpretação do perfil de aptidão de cada indivíduo reclame, primeiro, uma análise exaustiva da sua globalidade e, só depois, se efectue uma pesquisa detalhada teste a teste. Ora este percurso em duas etapas nem sempre tem sido cumprido pela maioria dos autores portugueses.
4. As propostas de baterias de testes devem ser culturalmente condicionadas e a interpretação do resultado organizado em função de indicadores do estatuto socio-económico dos sujeitos, bem como do seu nível de actividade física, o que não tem sido efectuado de forma exaustiva.
5. Afinal o que é a norma do grupo? O que significa um valor médio? Convém não esquecer que jamais a ideia de normalidade é expressa por um ponto, o P_{50} . É preciso irradicar, de uma vez por todas, a ideia que o normal é representado por um único valor, um ponto numa recta. A normalidade é expressa por uma enorme variação de resultados que pode percorrer o intervalo do P_{10} ao P_{90} . Norma é variância e não a média. A norma não possui qualquer valor

em si mesma. A única coisa que reflecte, se tiver sido bem construída, é o quadro ilustrativo da posição de determinados valores e nada mais.

6. Como é que determina que um sujeito é apto? Convém não esquecer que a aptidão física é expressa de forma multidimensional e que dois sujeitos podem ser considerados de aptos e, no entanto, possuírem perfis de aptidão completamente distintos, mesmo diametralmente opostos. Esta matéria nunca foi abordada de forma a esclarecer este problema.
7. A interpretação do resultado obtido por uma dada criança ou jovem num qualquer teste de aptidão reclama uma análise judiciosa e esclarecida, sobretudo se o seu valor for fortemente condicionado pelo factor dimensional. Ora este assunto, só recentemente tem sido revisitado e de forma altamente esclarecedora do seu alcance interpretativo.
8. Afinal, apto para quê? Qual é a relevância substantiva de um dado valor na vida de uma criança e jovem, a que se encontra adstrita esta qualificação. Não sei responder de forma suficientemente concreta e esclarecida a esta questão.

Este quadro de fortes insuficiências no domínio da avaliação normativa conduziram à minha adesão ao pensamento da avaliação referenciada a critérios que penso ser mais promissor e de horizontes mais vastos. Não é, contudo, isenta de problemas que representam um forte desafio aos investigadores, e que mais adiante referenciarei. Apesar de já ter apresentado e escrito acerca da essência (Maia, 1996), estrutura operativa e alcances deste “novo” entendimento da avaliação da aptidão física, bem como da vasta bibliografia disponível, ainda que em inglês, reitero a minha perplexidade diante do desconhecimento do seu valor, da sua falta de aplicação e utilização mais generalizada no nosso país.

A proposta mais interessante e desafiadora deste novo posicionamento é tentar responder à eterna questão - *how fit is fit enough?*

Antes de apresentar algumas das ideias, princípios e procedimentos da avaliação referenciada a critérios, impõe-se uma explicação prévia. Este tipo de avaliação, tanto quanto julgo saber, não é utilizada pelos profissionais de Educação Física. Deste modo, a abordagem usada será tão didáctica quanto possível, sem me concentrar em detalhes que não interessam no momento.

A essência da avaliação referenciada a critérios é devida ao trabalho clássico de Glasser e Klaus (1962) cuja essência se dirigia para a avaliação de comportamentos de aprendizagem, i.e. para a obtenção de valores que providenciassem informação acerca do comportamento alcançado, que designou de comportamento critério.

A ideia de base da avaliação criterial pode ser apresentada do seguinte modo (Glasser e Nitko, 1971):

Um teste referido ao critério é aquele que foi deliberadamente construído para produzir uma medida que é directamente interpretada em termos de um padrão determinado de performance. Os padrões de performance são especificados a partir da definição de uma classe ou domínio de tarefas que devem ser realizadas por um indivíduo. As medidas são realizadas em amostras representativas de tarefas do domínio em causa, e tais tarefas são referidas directamente ao domínio de cada sujeito.

O que aqui está em causa é, acima de tudo, a interpretação da *performance* do sujeito relativamente a um conjunto bem definido de competências e que necessita:

1. Que se descreva a *performance* do sujeito numa base estritamente individual. Dito de outro modo, não se pretende em nenhuma circunstância, tal como se faz na avaliação normativa, comparar a *performance* do sujeito com a do seu grupo de referência (Looney, 1989).

2. Que se atribua graus de proficiência numa medida critério. Neste caso o resultado é interpretado de forma puramente dicotómica, zero e um. Zero se falha, um se passa. Ou seja, que se classifiquem os sujeitos em função do sucesso ou insucesso na realização de determinada tarefa perfeitamente definida, ou que se obtenha um dado valor num teste cuja validade é conhecida. É também possível aumentar a latitude do grau de mestria para incluir níveis diferenciados - não-proficiência, proficiência mínima, proficiência óptima.

Tal como na avaliação normativa, também aqui se elaboram tabelas com valores de referência. Só que estes são valores que servem, exclusivamente, para classificar sujeitos em termos de alcance, ou não, de uma meta perfeitamente definida. Vejamos um exemplo simples desta noção de valor barreira, conhecido por valor de *cut-off*. Dada a presença de forte variabilidade de estatura dos alunos da turma A, por questões de equidade na formação das equipas de basquetebol, o professor de Educação Física decidiu dividi-los em dois grupos. Pretendia que os alunos altos jogassem com os altos e os baixos com os baixos. Assim, os que tivessem mais de 175 cm de estatura pertenciam às equipas dos altos. Os que tivessem menos de 175 cm jogavam com os baixos. Cento e setenta e cinco centímetros é o valor que classifica os alunos, categorizando-os, sem contudo os comparar. Este valor é conhecido por valor de corte, ou valor *cut-off*.

No domínio da aptidão física associada à saúde, a mudança radical de perspectiva residiu no estabelecimento de valores a serem alcançados pelos sujeitos, em cada teste, e que se pensa estarem associados a níveis desejados de saúde (Plowman, 1992). O que aqui está implícito é a resposta, do modo mais preciso possível à eterna questão *do how fit is fit enough* - nível de aptidão quanto basta. O valor de *cut-off* é aqui interpretado como o nível desejado de aptidão associado à saúde. O valor representa, em certo sentido, uma medida reduzida de risco.

A primeira versão de uma bateria de testes referenciados a critérios, a *South Carolina Physical Fitness Test*, foi desenvolvida em 1978 e era constituída pelos testes da corrida da milha, *sit-up*, *sit and reach* e pregas de adiposidade subcutânea, e foi desenvolvida para crianças e jovens dos 9 aos 16 anos de idade (Cureton e Warren, 1990).

Pela primeira vez na história da avaliação da aptidão física associada à saúde foram apresentados standards mínimos que se pensava estarem associados à saúde. Curiosamente, o valor de *cut-off* foi situado no percentil 50 em cada um dos testes. Os ajustamentos a estes valores eram efectuados sempre que fossem justificados por factores culturais.

A esta bateria seguiram-se, em 1986 a *Fit Youth Today*, a *Prudential Fitnessgram* em 1987, e a *Physical Best* da AAHPERD em 1988. Actualmente a bateria mais utilizada é a *Prudential Fitnessgram*, adoptada pela própria AAHPERD em substituição da sua própria proposta.

Das baterias disponíveis só mencionarei com algum detalhe a *Prudential Fitnessgram*, provavelmente a mais conhecida e estudada. Esta bateria foi desenvolvida no Instituto Aeróbico de Dallas sob a direcção do eminente epidemiologista da actividade física, Steven Blair. Os valores de *cut-off* foram determinados a partir da investigação disponível em estudos epidemiológicos, clínicos e dados observados em pesquisas normativas. O próprio Steven Blair referiu que os valores de corte são propostas de trabalho. Enquanto tal, reclamam da comunidade científica e dos profissionais de Educação Física um exame meticuloso e contínuo por forma a estabelecer uma maior intimidade com os valores critério das medidas de saúde que lhes estão adstritos. As componentes mais solidamente pesquisadas e validadas são as que se referem à capacidade aeróbia e à composição corporal, para as quais os valores de corte são consensuais.

Nesta bateria são incluídas, fundamentalmente, três componentes essenciais da aptidão associada à saúde: (1) a capacidade aeróbia, (2) a composição corporal, (3) a força muscular, resistência e flexibilidade.

A bateria *Prudential Fitnessgram* dirige-se ao intervalo etário dos 5 aos 17 anos. O pilar da sua construção assenta nas componentes anteriormente referidas que procuram ligar-se ao estado de saúde do sujeito e sua funcionalidade para responder de modo adequado às tarefas do dia-a-dia.

Os resultados obtidos nos testes são submetidos a um programa informático devidamente actualizado que não só classifica o sujeito em cada item da aptidão, como também faz sugestões interessantes do ponto de vista da melhoria ou manutenção de cada componente de aptidão.

Quadro 4 - Estrutura operativa da bateria Fitnessgram

Componentes/itens da bateria	Sub-componentes da aptidão
CAPACIDADE AERÓBIA	
Corrida/marcha de 1 milha	Capacidade aeróbia
FORÇA MUSCULAR, RESISTÊNCIA E FLEXIBILIDADE	
<i>Curl-up</i>	Força abdominal
<i>Push-up</i> a 90 graus	Força da parte superior do tronco
Tempo de suspensão na barra	tronco
<i>Pull-up</i> modificado	
Elevações do tronco	Força e flexibilidade dos extensores do tronco
<i>Back saver sit and reach</i>	Flexibilidade
<i>Shoulder stretch</i>	
COMPOSIÇÃO CORPORAL	
Pregas de adiposidade subcutânea	Composição corporal
Índice de massa corporal	

Quadro 5 - Limites do intervalo para classificação do nível de aptidão (só se apresentam os valores relativos aos testes da corrida da milha, dos *curl-up*'s, das elevações do tronco e da % de gordura para o sexo masculino)

Idade	corrida milha (min:seg)		Porcentagem de gordura		<i>Curl up</i> 's		Elevações do tronco	
5	completar		25	10	2	10	6	12
6	completar		25	10	2	10	6	12
7	completar		25	10	4	14	6	12
8	completar		25	10	6	20	6	12
9	completar		25	10	9	24	6	12
10	11.30	9.00	25	10	12	24	9	12
11	11.00	8.30	25	10	15	28	9	12
12	10.30	8.00	25	10	18	36	9	12
13	10.00	7.30	25	10	21	40	9	12
14	9.30	7.00	25	10	24	45	9	12
15	9.00	7.00	25	10	24	47	9	12
16	8.30	7.00	25	10	24	47	9	12
17	8.30	7.00	25	10	24	47	9	12
17+	8.30	7.00	25	10	24	47	9	12

O Quadro 5 mostra os valores critério para cada teste para o sexo masculino. São mencionados em cada intervalo de idade, os limites inferiores e superiores para as zonas de aptidão associadas à saúde. O limite inferior reflecte o mínimo aceitável, e o limite superior serve para motivar os alunos a alcançar níveis mais elevados de aptidão - este limite não é mais do que um desafio.

O problema da validade é fulcral em qualquer estudo da avaliação da aptidão física, seja ela normativa ou criterial, dado que se aplica, sempre, a mesma estrutura conceptual e substantiva. A diferença situa-se sobretudo no domínio operativo, dada a especificidade da avaliação referida ao critério.

Na avaliação criterial identificam-se duas categorias de estudos de validade (Safrit, 1989; 1990):

- A validade referida ao domínio
- A precisão da decisão

Quando um teste é utilizado para identificar a extensão de um domínio de tarefas que um sujeito consegue realizar, diz-se que possui validade referida ao domínio, i.e. o teste mede os objectivos identificados no domínio em causa. Esta noção repousa nos pressupostos de verdade e valor. O que se pretende, afinal, tal como na avaliação normativa, é que o teste evidencie validade lógica na sua construção.

A grande fatia do estudo da validade na avaliação criterial vai, sem sombra de dúvida, para o estudo da precisão da decisão. O que aqui está em causa é classificar, da forma mais inequívoca possível, dois grupos de sujeitos em duas categorias - proficientes e não proficientes. Este modelo operativo, também designado por testagem de proficiência, é baseado no modelo binomial do erro, e é composto por 3 etapas:

- A primeira refere-se à utilização de valores dicotómicos, 0 e 1, para classificar os sujeitos em proficientes e não-proficientes.
- A segunda pretende identificar um valor de *cut-off* para separar os dois grupos. Mais à frente veremos a complexidade deste problema, e acima de tudo, a enorme controvérsia em torno deste famoso valor, da sua interpretação substantiva e das implicações que arrasta.
- A terceira centra-se na identificação dos erros de classificação, i.e. a determinação de quantos são falsos positivos e falsos negativos, bem como pretende conhecer o limite da validade do teste em função da extensão destes erros.

O problema central do *standard* ou valor de *cut-off* reside no seu processo de cálculo, na atribuição de significado ao número obtido e na especificação precisa da sua validade ecológica. Para solucionar este problema há quatro abordagens possíveis: a do julgamento dos peritos, a normativa, a empírica e a combinada. A operacionalização da busca deste valor de *cut-off* é efectuada a partir de três possibilidades: o teste critério, os grupos critério e os grupos *borderline*.

Quando se recorre à possibilidade do teste critério, a solução é obtida a partir de um triplo processo: a determinação da probabilidade máxima de decisão correcta, a determinação do valor do coeficiente de validade e a análise da utilidade (os interessados poderão consultar os aspectos técnicos desta matéria num texto publicado por Maia, 1996).

Centremos, finalmente, a nossa atenção no forte problema das limitações desta metodologia de avaliação da aptidão física.

Dentre as várias limitações desta abordagem, apresentaremos a que mais discussão tem levantado, e que é a seguinte:

- Os valores de *cut-off* envolvem sempre um certo julgamento subjectivo, dependente da posição do investigador e do conhecimento disponível. A discordância no valor critério é um

facto. Veja-se, por exemplo, as diferenças de valores critério em diferentes baterias para a prova da milha. Na bateria *Prudential Fitnessgram* o nivelamento não é tão restritivo quanto o é, por exemplo, na *President's National* nos diferentes intervalos de idade.

Quadro 6 - Diferentes valores critério para a prova da milha em rapazes (valores em minutos e segundos).

Idade	<i>Prudential Fitnessgram</i>	<i>Physical Best</i>	<i>President's National</i>
6	15.00	12.36	10.15
7	14.00	11.40	9.22
8	13.00	11.05	8.48
9	12.00	10.30	8.31
10	11.00	9.48	7.57
11	11.00	9.20	7.32

Plowman (1992), num estudo recente sobre o problema da validade dos standards da avaliação criterial, reexaminou três testes, o sit-up, o sit and reach e as elevações na barra. Estes testes são comuns a quatro baterias muito divulgadas nos estados Unidos. São elas, o *Physical Best*, a *Fit Youth Today*, a *Fitnessgram* e a *YMCA Youth Fitness Test*. Partindo de três estratégias diferentes, a dos dados empíricos, a do recurso a grupos extremos e a combinação de dados normativos com o julgamento de peritos, a autora concluiu, entre outros aspectos, o seguinte:

- Que existe discrepância nos valores de *cut-off* das diferentes baterias.
- Que não se possui informação precisa e inequivocamente segura acerca da validade dos *cut-off* propostos.
- E que se desconhece ainda, e apesar da informação disponível, o significado prático do valor de cada *cut-off* relativamente aos valores absolutos desejados de saúde.

Um exemplo esclarecedor da disparidade de *cut-off* e das suas implicações é referido por Morrow et al. (1995). Para o teste de corrida ou marcha da milha possuem-se os valores de *cut-off*. Contudo, Cureton e Warren (1990) propuseram, também, outros valores de VO_2 max para duas baterias, a *Fitnessgram* e a *Physical Best*. Os resultados são altamente sugestivos. A probabilidade máxima de classificações correctas é de 85% na primeira bateria e 61% na segunda. Tanto uma como outra possuem a mesma percentagem de falsos negativos, 4%. O problema são os falsos positivos: 11% na primeira e 35% na segunda, o que significa que há sujeitos considerados aptos numa bateria e considerados inaptos na outra.

Safrit e Looney (1992), ao reanalizarem os dados de Cureton e Warren (1990) referem, uma vez mais, três posições inequívocas sobre o assunto:

- A primeira refere-se à necessidade de um estudo sério e rigoroso da validade do *cut-off* de cada bateria, coisa que ainda não foi realizado.
- A segunda remete para o carácter imperioso da determinação da validade ecológica de cada teste e do seu valor de *cut-off*.
- A terceira, e mais concretamente dirigida aos resultados, é que ainda não se possui um consenso relativo á validade do critério do consumo máximo de O_2 , bem como á percentagem de VO_2 máximo que é utilizado em cada intervalo de idade. Acrescentam a estes factos, a necessidade de ajustar os resultados na prova da corrida/marcha da milha aos valores das pregas de adiposidade de cada sujeito.

É evidente, e apesar da controvérsia instalada, que a avaliação criterial continuará a ser o grande desafio neste final de milénio. As suas perspectivas no contexto da educação para a aptidão, actividade física e saúde são inequivocamente relevantes.

Finalmente é imperioso e urgente repensar de forma muito séria a problemática da avaliação da aptidão física no contexto escolar face à sua importância e relevância social. Os alunos e a sociedade civil merecem-nos este esforço.

Bibliografia

- AAHPERD (1998): Physical best. Washington, DC.
- Beunen, G., Lefevre, J., Claessens, A., Lysens, R., Maes, H., Renson, R., Simons, J., Vanden Ende, B., Vanreusel, B., Van Den Bossche, C. (1990). Tracking and prediction of adult fitness in males between the ages of 13 and 30 years. Paper presented at the Vith European Research Seminar - The Eurofit Tests of Physical Fitness, Izmir. Council of Europe.
- Beunen, G., Lefevre, J., Claessens, A., Lysens, R., Maes, H., Renson, R., Simons, J., Vanden Ende, B., Vanreusel, B., Van Den Bossche, C. (1992). Age specific correlation analysis of longitudinal physical fitness levels in men. *European Journal of Applied Physiology*. Vol. 64: 538-545.
- Bouchard, C., Shephard, R.J. (1992): Physical activity, fitness, and health: the model and key concepts. in Bouchard, C., Shephard, R.J., Stevens, T. (eds), Physical activity, fitness, and health. International proceedings and consensus statement. Human Kinetics Publishers. Champaign. pp: 77-88.
- Corbin, C.B. (1991). A multidimensional hierarchical model of physical fitness: a basis for integration and collaboration. *Quest*. Vol. 43: 296-306.
- Cureton, K.J., Warren, G.L. (1990). Criterion-referenced standards for youth health-related fitness tests: a tutorial. *Research Quarterly for Exercise and Sport*. Vol. 61: 7-19.
- Fleishman, E.A. (1964). The structure and measurement of physical fitness. Prentice-Hall, Inc. Englewood Cliffs.
- Glasser, R., Klaus, D.J. (1962). Proficiency measurement: assessing human performance. in Gagne, R.M. (ed), Psychological principles in systems development. Holt, Rinehart & Winston. New York. pp: 419-474.
- Glasser, R., Nitko, A.J. (1971). Measurement in learning and instruction. in Thorndike, R.L. (ed), Educational measurement. American Council on Education. Washington. pp: 32-47.
- Lefevre, J. (1994). Tracking of physical fitness. Concepts and results. Comunicação apresentada no I seminário internacional sobre o crescimento, maturação, aptidão física e performance. FCDEF-UP. Porto.
- Lefevre, J., Vanreusel, B., Beunen, G., Claessens, A., Lysens, R., Renson, R., Simons, J., Vanden Ende, B. (1997). Physical activity and physical fitness from adolescence through adulthood in Flemish males: a longitudinal study. in Coetsee, M.F., Van Heerden, H.J. (eds), Nutrition and physical activity. Proceedings of the ICPAFR '95. University of Zululand. pp: 135-142.
- Looney, M.A. (1989). Criterion-referenced measurement: reliability. in Safrit, M.J., Wood, T.M. (eds), Measurement concepts in physical education and exercise science. Human Kinetics Books. Champaign. pp: 137-152.
- Maia, J.A.R. (1995). Avaliação da aptidão física. Aspectos metodológicos e analíticos. *Horizonte*. Vol. XI, nº 65: 190-197.
- Maia, J.A.R. (1996). Avaliação da aptidão física. Uma abordagem metodológica. Vol. XIII, nº 73. *Dossier*.
- Maia, J.A.R., Lefevre, J., Claessens, A., Beunen, G. (1997). Stability of physical fitness in Belgian boys. A panel study using the LISREL approach. Comunicação apresentada no XIXth International Seminar on Pediatric Work Physiology. Exeter.
- Maia, J.A.R., Lefevre, J., Beunen, G., Claessens, A. (1998a). A estabilidade da aptidão física. O problema, essência analítica, insuficiências e apresentação de uma proposta metodológica baseada em estudos de painel com variáveis latentes. *Movimento (no prelo)*.
- Maia, J.A.R., Lefevre, J., Beunen, G., Claessens, A., Vanden Eynde, B., Vanreusel, B., Lysens, R., Renson, R. (1998b). Stability in physical fitness. A study on Belgian males followed longitudinally from 12 to 30 years. Comunicação apresentada no ACSM. Orlando.
- Maia, J.A.R., Lefevre, J., Claessens, A., Renson, R., Vanreusel, B., Beunen, G. (1998c). Tracking for physical fitness during adolescence. A panel study in Belgian boys. (em preparação).
- Malina, R.M. (1990). Tracking of physical fitness and performance during growth. in Beunen, G., Ghesquière, J., Reybrouck, T., Claessens, A. (eds), Children and exercise. Ferdinand Enke Verlag. Stuttgart.
- Malina, R.M. (1996). Tracking of physical activity and physical fitness across the lifespan. *Research Quarterly for Exercise and Sport*. Vol. 64: 48-57.

- Marques, A.T., Costa, A., Maia, J.A.R., Oliveira, J., Gomes, P. (1991). Aptidão Física. in Sobral, F., Marques, A.T. (eds), FACDEX, Desenvolvimento somato-motor e factorres de excelência desportiva na população portuguesa. Ministério da Educação. Gabinete Coordenador do Desporto Escolar. Lisboa. pp: 33-53.
- Marsh, H.W. (1993). The multidimensional structure of physical fitness: invariance over gender and age. *Research Quarterly for Exercise and Sport*. Vol. 64: 256-273.
- Morrow, J.R., Jackson, A.W., Disch, J.G., Mood, D.P. (1995). Measurement and evaluation in human performance. Human Kinetics. Champaign.
- Pate, R.R. (1988). The evolving definition of physical fitness. *Quest*. Vol. 40: 174-179.
- Plowman, S.A. (1992). Criterion referenced standards for neuromuscular physical fitness tests. An analysis. *Pediatric Exercise Science*. Vol. 4: 10-19.
- Safrit, M.J. (1989). Criterion-referenced measurement: validity. in Safrit, M.J., Wood, T.M. (eds), Measurement concepts in physical education and exercise science. Human Kinetics Books. Champaign. pp: 119-136.
- Safrit, M.J. (1990). An introduction to measurement in physical education and exercise science. Times Mirror/Mosby. St. Louis.
- Safrit, M.J. (1995). Complete guide to youth fitness testing. Human Kinetics. Champaign.
- Safrit, M.J., Looney, M.A. (1992). Should the punishment fit the crime? A measurement dilemma. *Research Quarterly for Exercise and Sport*. Vol. 63: 124-127.
- Simons, J., Beunen, G., Ostyn, M., renson, R., Swalus, P., Van Gerven, D., Willems, E. (1969). Construction d'une batterie de testes d'aptitude motrice pour garçons de 1 à 19 ans par la methode de l'analyse factorielle. *Kinanthropologie*. Vol. 58: 295-302.
- Sobral, F. (1989). Estudo de crescimento e aptidão física na população escolar dos açores. SREC-DREFD.RAA/ISEF-UTL. Lisboa.