

# Modelação hierárquica ou multinível. Uma metodologia estatística e um instrumento útil de pensamento na investigação em Ciências do Desporto

José A. Maia<sup>1</sup>, Vítor P. Lopes<sup>2</sup>, Rui G. da Silva<sup>1</sup>, André Seabra<sup>1</sup>, João V. Ferreira<sup>3</sup>, Manuel V. Cardoso<sup>1</sup>

<sup>1</sup> Faculdade de Ciências do Desporto e de Educação Física, Universidade do Porto, Portugal; <sup>2</sup> Escola Superior de Educação, Instituto Politécnico de Bragança, Portugal; <sup>3</sup> Escola Superior de Educação, Instituto Politécnico de Viseu, Portugal

## RESUMO

O propósito deste trabalho é apresentar aspectos fundamentais da modelação hierárquica ou multinível aos investigadores das Ciências do Desporto dos países de língua oficial portuguesa. Acima de tudo, salienta a necessidade, cada vez maior, de considerar a natureza hierárquica da informação contida num número elevado de pesquisas. Mostra, também, que a não consideração desta estrutura enviesará fortemente as conclusões dos estudos. É efectuada uma apresentação fortemente didáctica do tipo tutorial, que percorre as ideias de base da metodologia, apresenta um exemplo ilustrativo de alguns aspectos da análise e sugere algumas pistas fundamentais para se realizar pesquisa neste domínio.

*Palavras-chave:* modelação, hierarquia, multinível, Ciências do Desporto.

## ABSTRACT

*Hierarchical or multilevel modeling. A statistical methodology and a usefull thinking tool of research in sport sciences.*

*The purpose of this paper is to present some fundamental aspects of multilevel or hierarchical modeling to researchers of Sport Sciences within the Portuguese speaking community. It is stressed the relevance of a major concern for considering the multilevel data characteristics presented in most of the research. It shows also the bias present in many conclusions for not considering the clustered nature of data sets researchers have at hand. We then present a didactic example that follows major issues of hierarchical modeling, and suggest some ideas for conducting research in this methodology.*

*Key Words:* modeling, hierarchical, multilevel, sport sciences.

## 1. INTRODUÇÃO

É nossa intenção reforçar a ideia, bem disseminada aliás na investigação em Ciências Sociais e Humanas, que a maior parte da informação recolhida na pesquisa empírica possui um padrão hierárquico ou contextual inequívoco, de que os exemplos seguintes são uma mostra bem reduzida: os alunos estão agrupados em classes, as classes em diferentes escolas, as escolas em áreas geográficas distintas; trabalhadores estão hierarquicamente dependentes de sectores, sectores em áreas distintas das empresas, e estas em diferentes localidades. De um modo equivalente, os atletas estão dependentes de diferentes treinadores, que pertencem a clubes diferenciados; os professores estagiários estão associados hierarquicamente a orientadores distintos, que leccionam em diferentes escolas, cuja localização e características são bem diversas. Outros exemplos podem provir da pesquisa em agregação familiar nos hábitos de actividade física (ainda que nunca realizados no contexto desta metodologia), em que filhos estão agrupados em famílias e estas em locais distintos do ponto de vista sócio-económico e demográfico. Se considerarmos com alguma atenção e cuidado qualquer estrutura de dados recolhidos no seio de uma qualquer pesquisa, facilmente “veremos” padrões hierárquicos ou multiníveis (12), que Heck e Thomas (9) designam, genericamente, de estrutura organizacional da informação. A ausência de consideração desta estrutura hierárquica, que salienta uma interligação ou dependência forte da unidade de análise mais baixa na hierarquia (i.e., alunos, trabalhadores, atletas, professores estagiários, etc.) da mais elevada (e que podem ser classes, sectores, treinadores, supervisores, etc.), acarreta uma leitura demasiado parcelar e truncada da informação disponível. Daqui que Plewis (17) tenha referido, justamente, que qualquer investigador que ignore o padrão hierárquico dos seus dados será confrontado com uma perspectiva altamente enviesada dos seus resultados e conclusões, proporcionando uma visão distorcida e fragmentada daquilo a que pretendia dar uma resposta mais esclarecida e abrangente. A história da investigação multidisciplinar nas Ciências do Desporto tem sido fecunda em ilustrar, um sem número de vezes, a confusão estabelecida entre unidade observacional e unidade experimental,

ou entre micro e macro aspectos da informação disponível. A esta evidência associa-se, inapelavelmente, o uso inadequado de métodos de análise de dados que retiram a estrutura hierárquica saliente na investigação. Assume-se, pois, e esquece-se de seguida, a interdependência da informação que se cruza em diferentes planos do próprio delineamento. A história deste desencontro foi salientada, pela primeira vez, por Lindquist em 1940 no contexto das Ciências da Educação, e somente em 1997 por Weimo Zhu na prestigiada revista *Research Quarterly for Exercise and Sport* no vasto domínio polifacetado das Ciências do Desporto. Um resumo suficientemente esclarecedor do percurso da Modelação Hierárquica ou Multinível (MHMN) é encontrado, por exemplo, em Kreft, Leeuw (12). Apesar das diferentes propostas para solucionar de modo adequado o problema interpretativo de dados com estrutura ou padrão em diferentes níveis, é somente nos anos 80 que estatísticos e metodólogos ingleses (por exemplo, Harvey Goldstein) e americanos (por exemplo, Stephen Raudenbush) solucionam, de modo adequado, os enormes problemas levantados à análise deste tipo de informação, propondo *software* de fácil manuseamento (*HLM* ou *MLwiN*) e com enormes possibilidades de modelação. Estava, pois, aberta a porta, não só à colocação de maiores interrogações aos dados disponíveis, como também à possibilidade da sua resposta, assumindo um delineamento cuidadoso da pesquisa e um conhecimento adequado da metodologia. Parece ser inquestionável a necessidade em atribuir uma maior importância a um pensamento e a uma estrutura centrada na MHMN. Por exemplo, num espaçamento de 11 anos quatro prestigiadíssimas revistas internacionais dedicaram números inteiros ao tratamento de matérias que vão desde aspectos de natureza didáctico-metodológica do uso da MHMN, das suas enormes potencialidades de interrogação da informação, das características relevantes da sua flexibilidade, às aplicações mais diversas no domínio substantivo (ver *International Journal of Education Research*, 1990; *Journal of Education and Behavioral Statistics*, 1995; *Counseling Psychologist*, 1999; *Multivariate Behavioral Research*, 2001). Uma busca exclusiva no *Medline* relativa aos anos de 1999-2001 inventariou 53 trabalhos nos mais variados

domínios – das ciências sociais e humanas, aos assuntos mais “hard” das ciências biológicas. Os livros de texto sobre o lato território da MHMN são já em número “substancial”. Eis uma amostra suficientemente esclarecedora e actual:

- *Introducing multilevel modeling* (12).
- *An introduction to multilevel modeling techniques* (9).
- *Multilevel modeling. Techniques and applications* (11).
- *Multilevel analysis: an introduction to basic and advanced multilevel modeling* (24).
- *Hierarchical linear models. Application and data analysis methods* (19).

Contudo, e apesar da forte explosão no uso de um pensamento hierárquico ou multinível, o facto é que nas Ciências do Desporto a sua utilização é ainda tímida, pelo menos a fazer fé na produção disponível e que é extremamente exígua. O primeiro texto didáctico que introduz o assunto a partir de um exemplo bem esclarecedor é devido a Zhu (25) na prestigiada revista *Research Quarterly*, se bem que Baxter-Jones et al. tenham utilizado a MHMN em 1993 no *Journal of Applied Physiology* numa pesquisa sobre a modelação do desenvolvimento da potência aeróbica em jovens atletas.

## 2. ASPECTOS DO PROBLEMA DA ANÁLISE CENTRADA EM UNIDADES MACRO OU MICRO

Antes de avançarmos mais no texto gostaríamos de salientar, uma vez mais, as insuficiências que surgem, necessariamente, quando se efectua uma qualquer análise exclusivamente no seu nível mais baixo, desagregando a informação (i.e., por exemplo, ao nível exclusivo dos alunos) ignorando, inadvertidamente, o padrão hierárquico que a informação contém<sup>1</sup>.

Zhu (25), Raudenbush e Bryk (19) e Curran (8) inventariaram as insuficiências seguintes:

### — Heterogeneidade das rectas de regressão

Espera-se que haja uma tendência linear negativa quando se estuda, por exemplo, a relação entre o desempenho motor na prova da milha e a idade cronológica (i.e., quanto maior for a idade dos alunos, tanto menor será o tempo necessário para cobrir a distância da prova). Tal facto é bem conhecido e

documentado. Está associado ao aumento da potência cárdio-respiratória de crianças e jovens em função do incremento da sua idade cronológica (sobre esta matéria ver (21)). Contudo, é também de esperar que o desempenho médio seja diferente entre escolas (cada escola terá a sua recta de regressão, distintas que são umas das outras), dado que em cada escola actua, de modo distinto, um conjunto variado de factores que contribuem, também, para explicar as diferenças encontradas. Ignorar esta fonte de variabilidade não parece ser o mais adequado em qualquer tipo de análise.

### — Ausência de independência das observações

Face à circunstância de grupos de alunos pertencerem a escolas diferentes, cada uma das quais com as suas particularidades bem próprias, é de esperar que os alunos no seio de cada escola sejam relativamente homogéneos entre si (i.e., as observações ou registos dos seus valores de aptidão física, por exemplo, não são independentes, verificando-se alguma correlação entre sujeitos da mesma escola). Os alunos da escola A, de nível sócio-económico médio-elevado, que têm aulas de Educação Física duas vezes por semana com um professor especialista, com material didáctico suficiente e infra-estruturas adequadas são relativamente mais homogéneos nos seus níveis de aptidão físico-motora, mas suficientemente distintos de outros que não têm aulas de Educação Física, ou outros ainda que só têm uma aula de Educação Física por semana, não possuem infra-estruturas gímnico-desportivas e não têm um professor especialista para conduzir as aulas.

Torna-se imperioso, pois, que qualquer procedimento de análise considere, em simultâneo, as diferenças interindividuais dos alunos (nível 1 ou nível micro da informação) e as características diversificadas das escolas (nível 2 ou nível macro da informação).

### — Agregação

O problema da agregação ocorre quando, em estudos de natureza diferencial, os dados são agrupados ao nível das escolas (ignorando a variação interindividual dos alunos), ou somente ao nível das diferenças entre sujeitos, como ocorre em estudos de regressão linear simples ou múltipla) ignorando os efeitos da variação encontrada ao nível das próprias escolas.

Ainda que corramos o risco de repetição, nunca será de mais salientar a urgência do recurso a modelos com estrutura hierárquica ou multinível, que considerem, numa única estrutura de análise, a informação contida nos dois níveis da hierarquia – alunos e escolas.

### 3. ESTRUTURA ESSENCIAL DO TEXTO

Depois de lançada e estabelecida a importância da pesquisa em MHMN, é imperioso que se refira agora a estrutura deste trabalho, que pretende ser uma primeira viagem de natureza didáctico-metodológica ao território fascinante da identificação de estruturas ou padrões em dados de natureza multivariada.

Assim, referiremos em primeiro lugar alguns dos aspectos da pesquisa em Ciências do Desporto que recorreram a esta forma de pensar e analisar a informação disponível nos diferentes estudos. De seguida concentrar-nos-emos em aspectos fundamentais da MHMN, concretamente na sua estrutura fundamental, nos métodos de estimação, na precisão do cálculo das estimativas dos parâmetros mais importantes dos diferentes modelos a serem testados, bem como da extensão da metodologia a dados discretos e/ou binários. Será referido, também, o *software* disponível, bem como os *sites* na *Internet* onde pode ser obtida informação mais detalhada sobre cada um deles. Num terceiro momento apresentaremos um exemplo substantivo que permitirá, de uma forma didáctica, seguir algumas das ideias e vantagens da MHMN. Neste particular, seguiremos as propostas metodológicas sugeridas por Raudenbush e Bryk (19), Heck e Thomas (9) e Hox (10, 11).

Abordaremos finalmente, ainda que de um modo muito breve, algumas sugestões relativas ao delimitamento de uma pesquisa hierárquica ou multinível.

### 4. ALGUNS EXEMPLOS DE PESQUISA COM MODELAÇÃO HIERÁRQUICA

Ao contrário do que acontece no lato universo das Ciências da Educação, a disponibilidade informacional da aplicação da MHMN ao território das Ciências do Desporto não é tão extensa quanto seria de esperar. Essencialmente, a investigação disponível remete-nos para o coração da interpretação das diferenças interindividuais (situada quer ao nível micro, quer ao macro) no que ao desempenho motor ou performance desportivo-motora diz respeito. Nesta maté-

ria, um dos textos mais substanciais e que coloca a tónica na vasta problemática da alometria relativa à performance diferencial é o de Nevill e Holder (16)). Uma pesquisa às bases de dados internacionais permitiu localizar um número muito restrito de trabalhos os quais, em síntese, se poderiam agrupar em quatro olhares inquisitivos ao vasto território da performance diferencial:

1. Os trabalhos relativos ao recurso à MHMN no domínio estrito da fisiologia remetem-nos, na sua essência, para o estudo das mudanças no consumo máximo de  $O_2$  em crianças e jovens circum-pubertários. Trata-se, não somente de modelar as mudanças intraindividuais no consumo máximo de  $O_2$  proveniente de informação longitudinal (1, 3, 5), mas também de interpretar tal desenvolvimento em função da perspectiva alométrica (6), ou a relação do crescimento somático e a maturação biológica com a potência mecânica média (2).

2. Um outro território de aplicação da MHMN é oriundo de um cruzamento da Auxologia com a Fisiologia (20). Numa pesquisa longitudinal com rapazes e raparigas durante a adolescência foi estudada a influência de factores hormonais (concretamente da testosterona e IGF1) na produção de força, quer nas curvas da distância, quer nas da velocidade alinhadas pela idade em que ocorre o pico de velocidade da altura.

3. Um texto substancial no domínio da MHMN é o de Zhu (25) que lança um olhar fortemente didáctico e metodológico do seu uso a partir da análise de factores relativos às características das escolas e dos professores e que estariam associados às diferenças interindividuais nos valores de aptidão física associados à saúde.

4. Um outro trabalho (talvez o primeiro em língua portuguesa no lato universo das Ciências do Desporto) que abre uma janela sobre a MHMN é o de Maia et al. (15) sobre a modelação do desempenho motor na coordenação motora e nos valores de aptidão física associada à saúde de crianças dos 6 aos 10 anos de idade da Região Autónoma dos Açores.

Em suma, a informação disponibilizada nestes trabalhos é, em primeira mão, a matéria-prima dos interessados pela MHMN que encontrarão nela, não só

território de aprendizagem do pensamento e da técnica da modelação hierárquica, mas também um conjunto diversificado de aplicações de natureza substantiva para apreciar a sua relevância.

## 5. IDEIAS FUNDAMENTAIS DO MODELO HIERÁRQUICO OU MULTINÍVEL

### 5.1. Estrutura fundamental

Estudos de natureza contextual, hierárquica ou multinível implicam, necessariamente, a especificação de duas equações, uma para cada um dos níveis em estudo, alunos (micro nível) e escolas (macro nível), por exemplo. Para tornar “mais fácil” o entendimento deste sistema algo complexo de equações, imaginemos uma qualquer prova de desempenho dos alunos, e que consideramos, também, que a sua variabilidade intraindividual é explicada por um conjunto de preditores dos próprios alunos (nível 1) e das características das escolas (nível 2).

A equação ao nível dos alunos modela as relações entre diferentes características (i.e. preditores das diferenças interindividuais, ou variáveis consideradas relevantes e que irão ajudar a interpretar as diferenças encontradas nos desempenhos dos alunos) que se situam ao nível 1,

$$\text{Desempenho}_{ij} = \text{Valor na ordenada}_{0j} + \sum \text{Declive}_{qj} (\text{preditores dos alunos})_{qij} + \text{Erro}_{ij},$$

em que  $i$ =aluno,  $j$ =escola a que pertence,  $q$ =variável preditora e  $\text{Erro}_{ij}$ =erro aleatório ao nível dos alunos.

Dado que o valor na ordenada e o declive (i.e. os coeficientes de regressão) variam entre escolas, sendo portanto variáveis aleatórias com uma dada distribuição, a variação na sua distribuição pode ser, também, função de um conjunto distinto de preditores ao nível da escola, ou nível 2, tal que

$$\text{Coeficientes de regressão}_{qj} = \text{Valor na ordenada}_{q0} + \sum \text{Declive} (\text{preditores ao nível da escola})_{sj} + \text{Erro}_{qj},$$

em que  $s$ =preditores ao nível da escola,  $j$ =escola e  $\text{Erro}_{qj}$ =erro aleatório ao nível da escola.

As etapas da análise multinível são pensadas de acordo com uma estratégia de complexidade crescente sugerida por Raudenbush e Bryk (19):

— **Em primeiro lugar** realiza-se uma análise de variância com efeitos aleatórios (do inglês *random effects anova*), de modo a providenciar informação acerca de quanta variação observada no desempenho existe no seio de cada escola (i.e. ao nível dos alunos – nível 1) e entre escolas (i.e. ao nível 2).

Ao nível dos alunos ( $i$ ) de uma dada escola ( $j$ ), o desempenho numa dada prova ( $Y_{ij}$ ) é função da média da sua escola ( $\beta_{0j}$ ) mais um dado erro aleatório ( $r_{ij}$ ), tal que,

$$Y_{ij} = \beta_{0j} + r_{ij}.$$

Ao nível das escolas ( $j$ ), a média de cada escola ( $\beta_{0j}$ ) é função da grande média ( $\gamma_{00}$ ) mais um erro aleatório ( $u_{0j}$ ), tal que,

$$\beta_{0j} = \gamma_{00} + u_{0j}.$$

Juntando estas duas equações, temos pois que  $Y_{ij} = [\gamma_{00}] + [u_{0j} + r_{ij}]$ , em que  $\gamma_{00}$  é a média do desempenho motor de todos os alunos de todas as escolas numa dada prova, e  $r_{ij}$  uma componente aleatória. A variância de  $Y_{ij}$  é igual à variância entre escolas ( $\tau_{00}$ ) mais a variância entre sujeitos ( $\sigma^2$ ) e possuem uma grande importância, conforme veremos mais adiante na apresentação dos resultados do exemplo que mostraremos.

As questões que aqui podem ser colocadas são as seguintes: (1) haverá ou não variação suficiente entre alunos no seu desempenho que exige interpretação adequada, desde que sejam identificados os seus preditores? (2) quanta variação observada no desempenho é devida à circunstância das crianças pertencerem a escolas diferentes? (a resposta a esta questão, considerada fundamental na modelação hierárquica, é dada pela magnitude do coeficiente de correlação intraclassa); (3) existirá, ou não, variação suficiente ao nível do desempenho médio das escolas, que reclama uma explicação circunstanciada?

— **Em segundo lugar** especifica-se um modelo de coeficientes aleatórios (do inglês *random coefficient*

model) para examinar as equações de regressão no seio de cada escola e entre escolas. Aquilo que se deseja ver respondido é o seguinte:

— Quais são os valores médios dos coeficientes de regressão das escolas, incluindo valores na ordenada e declives? Trata-se, tão somente, de apresentar o perfil normativo médio de desempenho em função da idade, se a idade fosse a variável fundamental “contra” a qual se contrastam os valores do desempenho dos alunos.

— Qual é a magnitude da variação dos coeficientes de regressão entre escolas? Espera-se, nesta situação, identificar aspectos relativos às diferenças no desempenho entre escolas – diferenças nos valores de partida, bem como nos incrementos médios do desempenho em função da idade. A Figura 1 ilustra estes pontos.

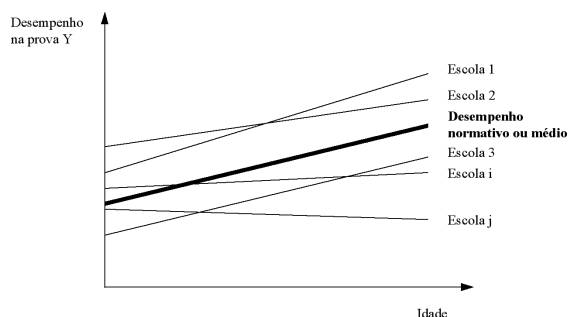


Figura 1: Perfis normativo e individual de cada escola para o desempenho numa dada prova em função da idade.

— Quanta variação presente no desempenho motor é explicada, por exemplo, pelas diferenças de idade e género sexual dos alunos? A Figura 2 pretende ilustrar esta situação.

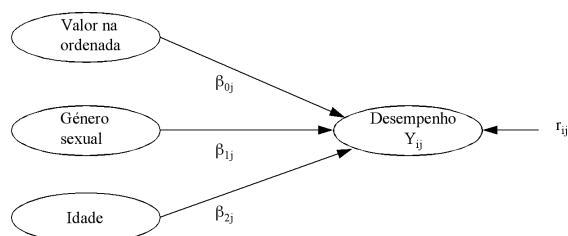


Figura 2: Representação esquemática da regressão ao nível 1, tendo como preditores a idade dos alunos e o seu género sexual.

Da Figura emerge a seguinte equação,

$$Y_{ij} = \beta_{0j} + \beta_{1j} (\text{género sexual}) + \beta_{2j} (\text{idade}_{ij} - \text{idade média}_j) + r_{ij}$$

Que ao nível hierárquico superior conduz a três novas equações,

$$\beta_{0j} = \gamma_{00} + u_{0j},$$

$$\beta_{1j} = \gamma_{10} + u_{1j},$$

$$\beta_{2j} = \gamma_{20} + u_{2j},$$

em que  $\gamma_{00}$  é a média das médias das escolas,  $\gamma_{10}$  e  $\gamma_{20}$  são a média dos declives da idade e género sexual entre escolas.

— **Em terceiro lugar**, assumindo que os coeficientes de regressão são diferentes entre escolas e que uma “reduzida” quantidade de variância pode ser explicada ao nível dos alunos, deve ser utilizado um modelo mais complexo para determinar o porquê de determinadas escolas possuírem médias mais elevadas no desempenho, bem como associações mais fortes entre idade, género sexual e desempenho motor numa dada prova.

## 5.2. Métodos de estimação

Face ao carácter extremamente técnico deste e dos dois pontos seguintes, tentaremos abordá-los de um modo mais substantivo e “ligeiro”. Convidamos os mais interessados neste território, a percorrer as páginas dos manuais dos programas HLM 5.0 e MLwiN. Uma outra leitura altamente proveitosa é o texto de Hox (11).

Antes de nos lançarmos na apresentação dos métodos de estimação, convém esclarecer, de um modo muito tangencial, três termos que se nos afiguram importantes – modelo, técnica estatística e algoritmo. No seu sentido mais amplo, um modelo é uma representação simplificada da realidade. No caso concreto de um modelo estatístico, enquanto instrumento de apreciação sempre simples de uma realidade mais complexa, é composto por um dado número de equações que descrevem as relações entre quantidades aleatórias. É importante salientar que apesar da MHMN possuir preditores fixos nos seus diferentes níveis hierárquicos, o modelo contém, sempre, um termo estocástico e/ou aleatório. Os modelos pos-

suem, na sua generalidade, um conjunto de parâmetros desconhecidos e que são utilizados para descrever aspectos que se consideram fundamentais no modelo. Uma técnica estatística é uma função ou programa, que considera os dados como *input* e produz valores para os parâmetros desconhecidos. Na maior parte dos casos, a técnica estatística provém da aplicação de um princípio estatístico ao modelo e que pode ser, por exemplo, a máxima verosimilhança ou os mínimos quadrados.

As técnicas estatísticas são implementadas em algoritmos que se utilizam para realizar os cálculos, otimizando a solução de um dado modelo, sobretudo no que respeita aos valores dos parâmetros relevantes. Os estimadores normalmente utilizados em MHMN são estimadores de máxima verosimilhança. O cálculo destes estimadores requer um processo iterativo. Normalmente o programa estatístico utilizado (por exemplo o HLM 5.0 ou o MLwiN) gera valores iniciais razoáveis para os diferentes parâmetros, obtidos a partir de estimativas de mínimos quadrados. Na segunda iteração já se obtêm estimativas de mínimos quadrados generalizados (utilizadas para estimar os coeficientes do segundo nível da hierarquia). Quando o processo iterativo converge (i.e., quando se encontra um mínimo de uma dada função de acordo com o método de estimação), as estimativas obtidas dos diferentes parâmetros são designadas de estimativas de máxima verosimilhança (utilizadas para estimar variâncias e covariâncias nos primeiro e segundos níveis da hierarquia). Por exemplo, o *output* do programa estatístico HLM 5.0 providencia não somente as estimativas de mínimos quadrados, como também, e sobretudo, as estimativas de máxima verosimilhança. Tanto numa como noutra são também calculados erros-padrão robustos somente para os efeitos fixos no modelo.

Há duas versões de estimação por máxima verosimilhança comumente utilizadas em MHMN – *full information maximum likelihood* e *restricted maximum likelihood*. As duas versões estão implementadas nos diferentes programas de MHMN.

### 5.3. Precisão das estimativas dos parâmetros

A estrutura conceptual e analítica da MHMN assenta num conjunto de pressupostos<sup>2</sup> que é importante não violar, para não conduzir à obtenção de estimati-

vas enviesadas dos parâmetros do próprio modelo. Contudo, nem sempre se consegue cumprir aquilo que está explicitamente descrito nos pressupostos do modelo. Daqui que diferentes estudos de simulação tendam a referir o que acontece quando os dados não são normalmente distribuídos e a dimensão da amostra, nos dois níveis, é pequena. Na generalidade, pode referir-se que a precisão das estimativas dos parâmetros depende do que se está a estimar (parâmetros fixos e respectivos erros-padrão, ou parâmetros aleatórios e respectivos erros-padrão), e sobretudo da dimensão das amostras. Num brevíssimo sumário pode dizer-se que:

- As estimativas dos parâmetros fixos não são enviesadas qualquer que seja o método utilizado (mínimos quadrados, mínimos quadrados generalizados, máxima verosimilhança).
- Há diferenças na qualidade das estimativas, sobretudo no que se refere à sua eficiência. Aqui, o método de máxima verosimilhança parece comportar-se melhor que os outros.
- Não é ainda inequivocamente conhecida a vantagem de qualquer um dos métodos no que se refere à estimativa da componente de variância do modelo. Contudo, estudos de simulação referem a supremacia do método de máxima verosimilhança.

### 5.4. Extensões do modelo

A MHMN possui um carácter de forte generalização, sendo aplicada, também, a dados binários e/ou categorizados.

Por exemplo, a análise multinível de proporções usa, genericamente, o modelo em que

$$\text{Logit}(\pi_{ij}) = \gamma_{00} + \gamma_{10} X_{ij} + u_{0j}$$

Assume-se que as proporções observadas  $P_{ij}$  têm uma distribuição binomial de variância conhecida, em que

$$\text{Var}(P_{ij}) = (\pi_{ij}(1-\pi_{ij}))/n_{ij}$$

Exemplos detalhados do uso substantivo de dados binários e categorizados podem ser encontrados em Hox (10, 11), Raudenbush et al. (18), Raudenbush e Bryk (19), Snijders e Bosker (24).

Há muitos outros aspectos do modelo, que não consideraremos aqui, e que compreendem, por exemplo, análise em 3 ou mais níveis da hierarquia dos dados, modelação multivariada, análise de dados longitudinais, uso de variáveis latentes, etc. (consultar, por exemplo o manual do HLM 5.0).

### 5.5. Software disponível

Do *software* disponível para ser utilizado em pesquisa de MHMN iremos destacar somente quatro (os leitores interessados podem consultar as referências de Kreft e de Leeuw, 1998; Snijders e Bosker, 2002):

— HLM (*Hierarchical Linear Models*) de Raudenbush et al. (18). Trata-se de um *software* desenvolvido no contexto das Ciências da Educação, sendo pois o seu *software* “oficial”. A última versão (HLM 5.0) é altamente flexível e extremamente didáctica, oferecendo inúmeras possibilidades de análise (ver por exemplo o livro *Hierarchical Linear Models* de Raudenbush e Bryk, 2002). Mais informações podem ser obtidas a partir dos seguintes endereços:

<http://www.ssicentral.com/hlm>

<http://www.gamma.rug.nl>

— MLn/MLwiN (*Multilevel Modeling*) cujo principal investigador e divulgador é Harvey Goldstein. Este *software* foi desenvolvido pelo projecto multinível do Instituto da Educação da Universidade de Londres. É um programa fortemente interactivo, ainda que aspectos da sua especificação sejam diferentes da estrutura do HLM. É altamente versátil, com enormes potencialidades de análise e procedimentos avançados de simulação. Informações detalhadas podem ser obtidas no seguinte endereço:

<http://www.ioe.ac.uk/mlwin/>

— O PROC MIXED é uma parte do programa de estatística SAS. É também altamente flexível na sua especificação e no seu *output*, ainda que seja distinto dos anteriores (não tem uma versão em ambiente *Windows*). Existem vários documentos fortemente didácticos do seu uso e interpretação quer do *input*, quer do *output*. Informação adicional pode ser adquirida no endereço seguinte:

<http://www.sas.com/>

— O MIXED REGRESSION é um módulo do programa estatístico SYSTAT 10.0 desenvolvido por Don Hedeker. Possui várias possibilidades de análise

se para dados contínuos e discretos. A sua implementação no SYSTAT é de fácil manuseamento e acesso, embora não existam documentos didácticos do seu uso para além dos manuais do SYSTAT e da informação disponibilizada no *site* de Don Hedeker. Este autor disponibiliza, gratuitamente, os seus programas e respectivos manuais no seguinte endereço: <http://www.uic.edu/~hedeker/mixdos.html>

## 6. UM EXEMPLO PRÁTICO APLICADO AO DOMÍNIO DO DESENVOLVIMENTO MOTOR

A amostra deste exemplo ilustrativo é composta por 1255 sujeitos dos 10 aos 18 anos de idade, dos dois sexos, provenientes dos concelhos de Vila Real e Viseu. Estes alunos pertencem a 33 escolas aleatoriamente distribuídas pelas duas regiões.

A medida critério de desempenho motor considerada aqui é o tempo necessário para cobrir a distância da prova de corrida/marcha da milha incluída na bateria de testes *Prudential Fitnessgram*. O valor de fiabilidade da performance intraindividual foi estimado de acordo com os procedimentos habituais para diferentes intervalos de idade e sexo e localiza-se, genericamente, em torno de  $0.89 < r_{tt} < 0.97$ .

A actividade física total foi estimada com base nos valores das respostas ao questionário de Baecke et al (4) adaptado para a língua portuguesa. Este instrumento tem evidenciado forte consistência na sua estrutura ao longo de diferentes pesquisas realizadas na Faculdade de Ciências do Desporto e de Educação Física da Universidade do Porto (FCDEF-UP). As estimativas de fiabilidade para os diferentes índices calculados situam-se entre os valores de 0.81 e 0.90. O índice de massa corporal (IMC) foi calculado de acordo com a fórmula habitual ( $IMC = \frac{Peso (Kg)}{Altura^2 (m)}$ ). O estatuto sócio-económico (ESE) das escolas foi dividido numa escala de 1 a 5 em função do número de alunos classificados nos escalões A, B e C e que correspondem ao nível sócio-económico da família. Um procedimento equivalente permitiu classificar o equipamento das escolas, em função da dimensões e qualidade das instalações desportivas, bem como a qualidade e extensão do material didáctico. As principais medidas descritivas dos dois níveis de variáveis estão no Quadro 1.

**Quadro 1. Estatísticas descritivas das variáveis nos níveis micro (nível 1) e macro (nível 2).**

| Variáveis                | M±dp       | Mínimo | Máximo |
|--------------------------|------------|--------|--------|
| <b>Nível 1 (n=1255)</b>  |            |        |        |
| Idade (anos)             | 13.98±2.26 | 10.00  | 18.00  |
| Sexo                     |            | 0.00   | 1.00   |
| Milha (minutos)          | 8.51±1.70  | 4.36   | 15.54  |
| ActF (unidades)          | 7.28±2.27  | 2.50   | 16.63  |
| IMC (kg.m <sup>2</sup> ) | 20.37±3.06 | 11.38  | 34.22  |
| <b>Nível 2 (n=33)</b>    |            |        |        |
| ESE escola               | 2.79±0.70  | 2.00   | 4.00   |
| Equipamento              | 3.36±1.08  | 1.00   | 5.00   |

O primeiro passo da análise compreende a determinação da quantidade de variação que está associada ao primeiro (i.e., alunos) e segundo (i.e., escolas) níveis da estrutura dos dados. A especificação do modelo de efeitos aleatórios da Anova (do inglês *random effects Anova*) compreende, pois, um modelo designado de “nulo” que servirá de contraste a

outros modelos mais complexos que especificaremos mais adiante. A informação a sair deste modelo nulo é a seguinte:

- Uma estimativa da média do desempenho na prova da corrida da milha para todos os alunos de todas as escolas ( $\gamma_{00}$ ). Uma interpretação mais adequada dos dados implica que os centremos, i.e., que calculemos as diferenças de cada aluno relativamente à grande média<sup>3</sup>.
- Um fraccionamento da variância total do desempenho em cada teste pelos primeiro ( $\sigma^2$ ) e segundo ( $\tau_{00}$ ) níveis da hierarquia.
- Uma medida de dependência dos resultados ao efeito específico das escolas, que é dada pelo coeficiente de correlação intraclasses ( $\rho$ ).
- Informação acerca da hipótese, a testar posteriormente, das escolas possuírem as mesmas médias no desempenho da prova da corrida.

Os resultados desta análise preliminar estão no Quadro seguinte:

**Quadro 2: Resultados no modelo de Anova de efeitos aleatórios (random effects Anova) para determinar a presença de estrutura hierárquica, ou organizacional no desempenho na prova da corrida/marcha da milha.**

| Corrida da milha           | Efeito fixo             | Coefficiente            | Erro-padrão   | Valor t  | p      |
|----------------------------|-------------------------|-------------------------|---------------|----------|--------|
| Grande média               |                         | 8.50                    | 0.14          | 59.11    | <0.001 |
|                            | <i>Efeito aleatório</i> | Componente de variância | Desvio-padrão | $\chi^2$ | p      |
| Média das escolas          |                         | 0.63                    | 0.79          | 434.05   | <0.001 |
| Efeito ao nível dos alunos |                         | 2.19                    | 1.48          |          |        |

A grande média (i.e., a média populacional de todas as crianças e jovens, independentemente do sexo e da escola a que pertencem) no desempenho da prova de corrida/marcha da milha é de 8.50 minutos. A variância tradutora das diferenças interindividuais é de 2.19. Tal valor exige que se tente a sua modelação para identificar o grau de importância de diferentes preditores na explicação das diferenças entre sujeitos no que diz respeito ao seu desempenho. A variância da performance média entre escolas é de 0.63, e é estatisticamente significativa. Também aqui há que inspeccionar a importância dos preditores do segundo nível da hierarquia. O coeficiente de correlação

intraclasses é substancial, 22% (confirma-se aquilo que foi sugerido a propósito da variância entre escolas); vinte e dois por cento da variância total no desempenho está associada às características distintas das escolas, em “oposição” às características interindividuais das crianças e jovens.

Apesar da magnitude da variância do segundo nível ser de 0.63 e inferior ao do 1º nível que é de 2.19, possui um valor de  $\chi^2$  significativo, rejeitando-se, pois, a hipótese de que as escolas possuem todas o mesmo desempenho médio na prova. Esta é pois a indicação substancial para se recorrer à MHMN para interpretar o significado da variância situada nos

dois níveis da hierarquia sobre o desempenho na prova da corrida/marcha da milha. Se houvesse uma teoria para guiar a modelação do desempenho<sup>4</sup> na prova, a etapa seguinte consistiria em especificar um modelo para o primeiro nível da hierarquia, também designado de modelo não-condicionado (do inglês *unconditional model*). Na sua ausência, uma solução viável é utilizar, de forma exploratória, uma opção do programa estatístico HLM 5.0 para verificar a qualidade (i.e., o significado estatístico) de cada um dos preditores. Uma outra estratégia é recorrer a um lote reduzido de preditores que façam sentido na interpretação das diferenças interindividuais, e testar o seu impacto na variação interindividual da performance dos sujeitos (mas aqui há que fixar os declives, i.e., assumir que o efeito de cada uma destas variáveis é homogénea entre escolas). É esta estratégia que iremos seguir.

De um modo simplificado, as equações fundamentais que iremos utilizar para que o algoritmo optimize as estimativas dos parâmetros do modelo (*level 1 random intercept model*) são as seguintes:

$$\text{Desempenho motor } ij = \beta_{0j} + \beta_{1j}(\text{idade}) + \beta_{2j}(\text{sexo}) + \beta_{3j}(\text{ActF}) + \beta_{4j}(\text{IMC}) + r_{ij},$$

e que no segundo nível da hierarquia, assumindo um efeito homogéneo entre escolas, será,

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40}$$

Os resultados deste modelo estão no Quadro 3.

Quadro 3: Valores dos parâmetros mais importantes do modelo de level 1 random intercept para a corrida/marcha da milha.

| Corrida da milha           | Efeito fixo      | Coefficiente            | Erro-padrão   | Valor t  | p      |
|----------------------------|------------------|-------------------------|---------------|----------|--------|
| Grande média               |                  | 7.97                    | 0.53          | 15.09    | <0.001 |
| Média [idade]              |                  | -0.16                   | 0.02          | -6.49    | <0.001 |
| Média [género sexual]      |                  | -1.30                   | 0.08          | -17.17   | <0.001 |
| Média [ActF]               |                  | -0.08                   | 0.02          | -4.85    | <0.01  |
| Média [IMC]                |                  | 0.08                    | 0.01          | 6.71     | <0.001 |
|                            | Efeito aleatório | Componente de variância | Desvio-padrão | $\chi^2$ | p      |
| Média das escolas          |                  | 0.61                    | 0.78          | 589.66   | <0.001 |
| Efeito ao nível dos alunos |                  | 1.53                    | 1.24          |          |        |

Os comentários indispensáveis que emergem do Quadro são os seguintes:

— Os preditores considerados no nível 1 do modelo para interpretar as diferenças interindividuais no desempenho da prova da milha são todos significativos ( $p < 0.001$ ). Contudo, é importante salientar que, apesar da sua qualidade interpretativa de variância interindividual, há ainda variância residual que exige um esforço suplementar de modelação, dado que  $\sigma^2 = 1.53$ .

— De um modo equivalente, a ideia de homogeneidade das rectas de regressão entre escolas não pare-

ce ser inequívoca, dado que existe variância significativa no desempenho entre escolas,  $\tau_{00} = 0.61$  ( $\chi^2 = 589.66$ ,  $p < 0.001$ ). Mais adiante modelaremos esta heterogeneidade entre escolas.

— A grande média do desempenho é agora de 7.97 minutos (a diferença de valor relativamente ao primeiro modelo, e que era de 8.50 minutos, reside na entrada de variáveis significativas no modelo, i.e., a nova média é uma média ajustada, dado que a média de cada escola é ajustada às diferenças existentes nas variáveis do primeiro nível da hierarquia).

— Em média o desempenho motor na prova da milha melhora com o incremento da idade (-0.16

minutos), é superior no sexo masculino (-1.30 minutos), depende dos valores da actividade física dos sujeitos (-0.08 minutos) e penaliza aqueles e aquelas que têm valores mais elevados de IMC (0.08 minutos).

— Cerca de 26% da variância no desempenho ao nível dos alunos é devida aos preditores considerados no modelo.

Consideremos agora um novo esforço de modelação com um modelo mais complexo e que se designa por *level-1 random-intercept and random-slope model*. Um dos problemas nucleares na especificação de modelos hierárquicos reside, precisamente, na consideração de que os preditores do primeiro nível poderão implicar um comportamento distinto entre escolas, sobretudo na possibilidade de serem representadas por um declive distinto. O problema situa-se ao

nível da escolha das variáveis que melhor reflectirão esta possibilidade. A título de mero exemplo foi considerada a variável actividade física (o IMC não mostrou distinção significativa nos declives entre escolas). As equações consideradas são iguais às do modelo anterior, com a diferença única das equações do segundo nível que são agora,

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + u_{0j} \\ \beta_{1j} &= \gamma_{10} \\ \beta_{2j} &= \gamma_{20} \\ \beta_{3j} &= \gamma_{30} + u_{3j} \\ \beta_{4j} &= \gamma_{40} \end{aligned}$$

Os resultados estão no Quadro 4 e são facilmente interpretáveis pelos leitores com base nas explicações anteriores.

Quadro 4: Valores dos parâmetros mais importantes do modelo de level 1 random intercept and random-slope model para a corrida/marcha da milha.

| Corrida da milha           | Efeito fixo      | Coefficiente            | Erro-padrão   | Valor t  | p      |
|----------------------------|------------------|-------------------------|---------------|----------|--------|
| Grande média               |                  | 7.87                    | 0.49          | 16.08    | <0.001 |
| Média (idade)              |                  | -0.16                   | 0.03          | -5.13    | <0.001 |
| Média (género sexual)      |                  | -1.30                   | 0.10          | -12.67   | <0.001 |
| Média (ActF)               |                  | -0.08                   | 0.02          | -4.53    | <0.01  |
| Média (IMC)                |                  | 0.09                    | 0.02          | 4.84     | <0.001 |
|                            | Efeito aleatório | Componente de variância | Desvio-padrão | $\chi^2$ | p      |
| Média das escolas          |                  | 1.06                    | 1.03          | 117.094  | <0.001 |
| Efeito do ActF (declive)   |                  | 0.006                   | 0.08          | 51.28    | 0.017  |
| Efeito ao nível dos alunos |                  | 1.49                    | 1.22          |          |        |

Um dos esforços mais complexos da MHMN é considerar um dos seus modelos mais interessantes e que se designa por *level-2 intercept-and slopes-as-outcomes model*. Esta nova tentativa de modelação mais complexa radica na circunstância de se ter verificado que existe variação significativa entre escolas em termos de médias do desempenho ( $\tau_{00}=1.06$ ;  $\chi^2=117.09$ ,  $p<0.001$ ), e o mesmo ocorre nos níveis de actividade física ( $\tau_{30}=0.006$ ;  $\chi^2=51.28$ ,  $p=0.017$ ). Estes resultados reflectem, necessariamente, a rejeição da hipótese nula de que o desempenho motor na prova da milha é constante entre escolas. O passo a dar,

agora, é tentar perceber o porquê da diferença de desempenho entre escolas. As variáveis que explicam esta diferença podem ser diversificadas, desde as características dos professores, qualidade dos planeamentos escolares e sua implementação, aspectos da organização didáctico-metodológica das aulas, equipamento disponível, características sócio-económicas do local da escola, etc. Para ilustrar melhor a flexibilidade da estrutura conceptual e analítica da MHMN, iremos testar as hipóteses das diferenças de desempenho médio entre escolas estarem associadas às características do seu

estatuto sócio- económico (ESE), qualidade das instalações e material disponível para as aulas de Educação Física (QIMEF).

Nesta situação, o modelo ao nível 1 permanece tal como formulado anteriormente,

Desempenho motor  $ij = \beta_{0j} + \beta_{1j}(\text{idade}) + \beta_{2j}(\text{sexo}) + \beta_{3j}(\text{ActF}) + \beta_{4j}(\text{IMC}) + r_{ij}$

A alteração substancial está agora no segundo nível da hierarquia,

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{ESE}) + \gamma_{02}(\text{QIMEF}) + u_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30} + \gamma_{31}(\text{ESE}) + \gamma_{32}(\text{QIMEF}) + u_{3j}$$

$$\beta_{4j} = \gamma_{40}$$

O *output* resultante da optimização da solução para este sistema algo complexo de equações referentes a este modelo está no Quadro 5.

Quadro 5: Valores dos parâmetros mais importantes do modelo de level-2 intercept-and slopes-as-outcomes model para a corrida/marcha da milha.

| Corrida da milha           | Efeito fixo      | Coefficiente            | Erro-padrão   | Valor t  | p      |
|----------------------------|------------------|-------------------------|---------------|----------|--------|
| Grande média               |                  | 9.09                    | 0.97          | 9.41     | <0.001 |
| ESE                        |                  | -0.48                   | 0.22          | -2.09    | 0.045  |
| QIMEF                      |                  | 0.03                    | 0.19          | 0.15     | 0.880  |
| Média [idade]              |                  | -0.16                   | 0.03          | -5.08    | <0.001 |
| Média [género sexual]      |                  | -1.30                   | 0.10          | -12.56   | <0.001 |
| Média [ActF]               |                  | -0.07                   | 0.08          | -0.83    | 0.415  |
| ESE                        |                  | 0.02                    | 0.02          | 0.75     | 0.460  |
| QIMEF                      |                  | -0.02                   | 0.02          | 0.96     | 0.343  |
| Média [IMC]                |                  | 0.09                    | 0.02          | 4.81     | <0.001 |
|                            | Efeito aleatório | Componente de variância | Desvio-padrão | $\chi^2$ | p      |
| Média das escolas          |                  | 1.03                    | 1.02          | 104.62   | <0.001 |
| Efeito do ActF (declive)   |                  | 0.016                   | 0.08          | 48.29    | 0.018  |
| Efeito ao nível dos alunos |                  | 1.49                    | 1.22          |          |        |

Centremos a nossa atenção em informação relevante e “nova” deste modelo. É claro para o leitor, neste ponto da abordagem do assunto da MHMN que a grande média pode variar em função, não só dos preditores do 1º nível da hierarquia, mas também da inclusão de preditores do 2º nível. Daqui que o seu valor seja ligeiramente diferente. Esta discrepância será solucionada a partir do momento em que se encontre o modelo que melhor se ajusta aos dados (este assunto será abordado de seguida quando labormos sobre a noção de *Deviance*).

O ESE (i.e., as escolas de estatuto sócio-económico mais elevado) possui um efeito marginalmente significativo ( $p=0.045$ ) na redução do tempo de prova (-0.48 minutos). Já a QIMEF não possui qualquer influência na heterogeneidade das médias entre escolas.

A inclusão de preditores associados à actividade física total não revelou, inexplicavelmente, qualquer influência significativa na heterogeneidade das médias entre escolas. A variância residual ( $\tau_{30}=0.016$ ), ainda que significativa ( $p=0.018$ ), é baixa. Contudo, não deixa de reclamar alguma atenção para interpretar o porquê deste valor e da possibilidade de se incluírem outros preditores que não o ESE e o QIMEF. Salientamos ainda que a variância na média das escolas ( $\tau_{00}=1.03$ ) é estatisticamente significativa ( $p<0.001$ ) e exigiria um novo esforço de modelação.

A estatística *Deviance* é utilizada enquanto tradutora da qualidade do modelo. Quanto maior for o seu valor, tanto pior será o ajustamento do modelo. Uma solução consensual entre os autores, quando se lida

com modelos hierarquicamente contidos uns nos outros (do inglês *nested models*), é comparar o valor da *Deviance* entre modelos. As diferenças entre *Deviances* possuem uma distribuição de qui-quadrado cujos graus de liberdade são iguais às diferenças entre o número de parâmetros testados em cada modelo. Dos resultados do Quadro 6, constata-se uma redução substancial no valor da *Deviance* do modelo nulo ( $M_0$ ) para o primeiro modelo ( $M_1$  – resultados no Quadro 3). Deste modelo ainda se verifica uma redução na *Deviance* para o segundo modelo  $M_2$  (resultados no Quadro 4). Contudo, o último modelo,  $M_3$  (Quadro 5) ao aumentar a *Deviance* sugere que não será a melhor hipótese de modelação.

Quadro 6. Resultados da *Deviance* e sua alteração em função dos diferentes modelos considerados no exemplo da prova de corrida/marcha da milha.

| Modelos     | <i>Deviance</i> | $\Delta$ na <i>Deviance</i> |
|-------------|-----------------|-----------------------------|
| Modelo nulo | 4625.96         |                             |
| Modelo 1    | 4199.89         | 426.07                      |
| Modelo 2    | 4189.49         | 10.40                       |
| Modelo 3    | 4199.43         | -9.94                       |

Somos pois de opinião que o melhor modelo é o  $M_2$ , dado que a inclusão de preditores no segundo nível da hierarquia não reduziu o valor da *Deviance*, nem se mostrou estatisticamente significativa<sup>5</sup>.

Se quisermos pensar em termos de proporção de variância explicada no desempenho da prova da milha, o “melhor” modelo explica cerca de 73% da variância total.

Poderemos pensar que os preditores considerados no segundo nível do Modelo 3 não serão os mais adequados para interpretar as diferenças que ocorrem nas médias entre escolas. Daqui que tenhamos referido anteriormente (Maia et al., 2002) que:

- Se houvesse informação sobre a qualidade dos documentos relativos ao planeamento anual das aulas de Educação Física,
- Se houvesse uma qualquer forma de entender a extensão e qualidade dos grandes objectivos das aulas de Educação Física em cada ano do 1º ciclo,
- Se houvesse forma de avaliar a estrutura e delimitação das opções didáctico-metodológicas de cada professor,

- Se tivéssemos acesso ao modo como cada professor avalia e determina o grau de prontidão desportivo-motora e coordenativa de cada criança,
- Se pudessemos quantificar o tempo efectivo das crianças nas tarefas motoras que as aulas encerram,
- Se houvesse informação sobre uma estrutura sólida de avaliação da Aptidão Física dos alunos, respectiva interpretação e *feed-backs* construtivos para as crianças,
- Se houvesse informação coerente sobre o número efectivo de aulas dadas por semana, bem como do seu tempo de duração,
- Se houvesse informação sobre o tempo dedicado a actividades de natureza motora de intensidade moderada e intensa,

então seria possível construir um modelo mais sólido da interpretação das diferenças verificadas nas médias das escolas no que ao desempenho motor dos alunos diz respeito. As respostas às perguntas anteriores foram consideradas no modelo construído por Zhu (25) no seu estudo interpretativo da Aptidão Física de crianças americanas. Os resultados parecem corroborar dois factos que, em nosso entender, são por demais inequívocos: (1) a importância das variáveis que descrevem as características do desempenho do professor e que foram anteriormente listadas nos “se”, e (2) a diversidade e qualidade das instalações disponibilizadas para as aulas de Educação Física.

Em suma, a grande vantagem da modelização hierárquica reside, precisamente, na oportunidade de interpretação das diferenças de desempenho motor que se verificam entre crianças, assumindo que essas diferenças estão dependentes de, pelo menos, dois níveis de influências: (1) as características dos próprios alunos, e (2) as características das escolas e dos professores. Assim haja informação da maior fidedignidade nestas matérias, sobretudo no 2º nível da estrutura hierárquica, que a intervenção dos responsáveis pela condução pedagógica do ensino da Educação Física nas escolas saberá dar-lhe o devido valor e retirar delas lições de um serviço cada vez mais eficaz e dedicado ao desenvolvimento da aptidão física de crianças e jovens.

## 7. DELINEAMENTO DE UMA PESQUISA HIERÁRQUICA OU MULTINÍVEL (ASPECTOS INTRODUTÓRIOS)

Esta parte do trabalho pretende lançar um conjunto necessariamente simples de pistas para a realização de uma pesquisa de natureza hierárquica ou multinível. Detalhes de nível intermédio e avançado podem ser consultados em 10, 11, 19, 24. As sugestões situam-se em três níveis:

— O primeiro radica, obrigatoriamente, na necessidade em recorrer a uma teoria ou corpo conceptual substantivo que contenha, em si mesma(o), uma noção forte de hierarquia ou estrutura organizacional. É também possível que o próprio problema e o modo como é formulado radiquem numa imagem clara de padrão hierárquico da informação a ser obtida. É da maior importância considerar conjuntos de preditores que ajudarão a interpretar as diferenças interindividuais que ocorrem nos dois níveis – alunos e escolas, por exemplo. Em língua portuguesa uma primeira abordagem a esta matéria é apresentada por Maia et al. (15), relativa à pesquisa do crescimento somático, aptidão física, actividade física e coordenação motora de crianças dos 6 aos 10 anos da Região Autónoma dos Açores. Outras janelas sobre a matéria do delineamento hierárquico são abertas por 10, 11, 14, 22, 23.

— O segundo nível exige uma forte consideração do problema da dimensão amostral que se situa, sempre, nos diferentes níveis da hierarquia, i.e., aos níveis micro e macro. As questões relativas à dimensão da amostra remetem qualquer pesquisa para aspectos da potência dos testes estatísticos inferenciais, precisão das estimativas dos parâmetros a que se associam erros-padrão de baixo valor e, sobretudo, para a faceta da generalização dos resultados para contextos equivalentes ou díspares dos da pesquisa em causa. Um texto fundamental neste domínio é o de Cohen (7) – *Statistical Power Analysis for the Behavioral Sciences*. As sugestões que aqui se apresentam, sob a forma de “rules of thumb” para pesquisa hierárquica com dois níveis (alunos agrupados em escolas, atletas em treinadores, trabalhadores em sectores de empresas, etc.) variam entre 30/30, i.e. pelo menos 30 grupos com 30 elementos cada um, a 50/20 se houver interesse substantivo em interacções cruzadas entre níveis. Há ainda a possibilidade

de se pensar em 100/10, i.e. 100 grupos com 10 elementos cada, se o propósito da pesquisa se localizar na parte aleatória do modelo (i.e. variâncias e covariâncias). Uma discussão detalhada destas matérias, sobretudo sob a forma de sumários bem esclarecedores, pode ser encontrada em 12, 13, 14, 24.

— O terceiro nível, ainda que destaque aspectos da dimensão amostral, coloca-se em função de aspectos de delineamento e propósitos da pesquisa: (1) estimar a média da população para a variável em causa; (2) estimar associações entre variáveis que se situem em planos hierárquicos distintos e, (3) explorar a estrutura da variância no segundo nível da hierarquia e que nos remete, não só para a estimação do coeficiente de correlação intraclasse, mas também para os parâmetros da variância nos dois níveis. Aspectos mais técnicos desta matéria específica são abordados de modo didáctico com base em vários exemplos em 12, 24 para onde remetemos os leitores mais interessados nestas matérias da MHMN.

Parece estar bem demonstrada a importância do recurso à MHMN na pesquisa do universo das Ciências do Desporto que possua um padrão hierárquico ou organizacional nos dados. Contudo, importa lançar um forte aviso aos navegadores principiantes destes territórios (e que é extensível a outros companheiros de aventura), o qual pode ser consubstanciado nas sugestões de Kreft e de Leeuw (12):

— É importante considerar que a metodologia proposta pela MHMN, como qualquer outra metodologia estatística, ou não, só funciona se todo o processo de delineamento da pesquisa e a estrutura de recolha da informação tiver sido pensada para responder às questões previamente formuladas. É imperioso considerar que os métodos estatísticos, por mais sofisticados que sejam, não são mais do que ferramentas imperfeitas para interpretar um mundo extremamente complexo e em mudança constante.

— A MHMN é, também, um instrumento de análise que busca identificar padrões na estrutura necessariamente multivariada dos dados. Neste sentido, é pois um olhar um tanto ou quanto fortuito da realidade. E neste sentido, é necessariamente parcelar. Daqui que possa funcionar, ou não!

— Os modelos hierárquicos ou multinível funcionam, e bem, quando os seus pressupostos não são violados! A violação de tais pressupostos implica, pelo menos, duas coisas: (1) que os testes de significância não devam ser considerados como válidos, e o mesmo se refere às estimativas dos parâmetros; (2) o pesquisador pode e deve recorrer a outras técnicas e procedimentos mais adequados às características dos seus dados e respectivas distribuições.

— A estrutura matemática subjacente à MHMN, apesar de altamente complexa, sofisticada e flexível, não é mais do que um olhar particularizado da realidade. Enquanto modelo, é tão-somente uma aproximação, às vezes bem grosseira, da realidade. E dado que a realidade é altamente complexa, e os modelos também o são, pequenas alterações nos valores iniciais ou nas iterações nos métodos de estimação podem conduzir a instabilidade nas estimativas dos parâmetros dos modelos, e conseqüentemente, a soluções impróprias ou a não convergência, pura e simples.

— É da maior importância considerar, no uso da MHMN, o recurso a teoria(s) ou a um corpo sólido de hipóteses, ou a justificações substantivas e metodológicas que justifiquem o seu uso. É urgente considerar o percurso: **no início é o problema; depois vêm as hipóteses; de seguida o delineamento cuidadoso da pesquisa e a recolha da informação; só depois é que vem a análise dos dados.**

<sup>1</sup> É evidente que um pensamento semelhante ocorre quando a análise é efectuada ao nível mais elevado da hierarquia, esquecendo completamente o nível mais baixo, agregando a informação.

<sup>2</sup> Os resíduos no nível 1  $e_{ij}$  possuem uma distribuição normal com média zero e variância comum  $\sigma^2$ . No segundo nível, os resíduos  $u_{0j}$  e  $u_{1j}$  são independentes do nível mais baixo ( $e_{ij}$ ), possuindo uma distribuição normal multivariada com médias zero. Os preditores são fixos e as relações são lineares.

<sup>3</sup> O problema da centragem dos dados é da maior relevância para tornar mais interpretável a informação dos alunos e dos preditores dos dois níveis (ver mais detalhes em 11 e 19).

<sup>4</sup> Desconhecemos a existência de uma qualquer teoria para descrever, explicar e prever o desempenho motor na prova da milha. Aliás, é nosso entendimento que há uma ausência notória de uma teoria geral do desempenho desportivo-motor. No caso concreto da prova da milha há, como é evidente, um conjunto diversificado de explicações de natureza fisiológica relativa ao funcionamento do sistema cárdio-respiratório (ver 21). Contudo, não se trata de uma interpretação generalizada e global do desempenho nesta prova ou num outro qualquer desempenho motor.

<sup>5</sup> Se os modelos a serem comparados não fossem hierarquicamente contidos, i.e., se fossem de natureza competitiva, e portanto com uma estrutura conceptual distinta, o princípio que deveria governar a sua qualidade, riqueza substantiva e relevância dos seus parâmetros, seria o da parcimónia. A “navalha de Occam” é aqui a directriz central (teorias e modelos devem ser o mais simples possível, sem deixar de se comprometer, necessariamente, com a natureza sempre complexa dos problemas). Em modelos competitivos não hierárquicos, mas ajustados ao mesmo conjunto de dados, a medida de ajustamento global que é mais citada na literatura (11) é o critério de informação de Akaike (AIC) que penaliza modelos fortemente expansivos e que possuem, necessariamente, menos graus de liberdade face ao maior número de parâmetros. Pretende-se, pois, penalizar excessos de parâmetros a estimar, em que a fórmula é  $AIC = d + 2q$ , e em que  $d = Deviance$  e  $q =$  número de parâmetros.

## CORRESPONDÊNCIA

José A.R. Maia

Faculdade de Ciências do Desporto  
e de Educação Física

Laboratório de Cineantropometria

Rua Dr. Plácido Costa, 91

4200-450 Porto

Portugal

[jmaia@fcdef.up.pt](mailto:jmaia@fcdef.up.pt)

## BIBLIOGRAFIA

- 1 Armstrong N, Welsman JR (2001). Peak oxygen uptake in relation to growth and maturation in 11- to 17-year-old humans. *European Journal of Applied Physiology*, 85:546-551.
- 2 Armstrong N, Welsman JR, Chia MYH (2001). Short term power output in relation to growth and maturation. *British Journal of Sports Medicine*, 35:118-124.
- 3 Armstrong N, Welsman JR, Nevill AM, Kirby BJ (1999). Modeling growth and maturation changes in peak oxygen uptake in 11-13 yr olds. *Journal of Applied Physiology*, 6:2230-2236.
- 4 Baecke JAH, Burema J, Frijters JER (1982). A short questionnaire for the measurement of habitual physical activity in epidemiological studies. *American Journal of Clinical Nutrition*, 36:936-942.
- 5 Baxter-Jones A, Goldstein H, Helm P (1993). The development of aerobic power in young athletes. *Journal of Applied Physiology*, 3:1160-1167.
- 6 Beunen G, Baxter-Jones ADG, Mirwald RL, Thomis M, Lefevre J, Malina RM, Bailey DB (2001). Intraindividual allometric development of aerobic power in 8- to 16-year-old boys. *Medicine and Science in Sport and Exercise*, 3:503-510.
- 7 Cohen J (1988). *Statistical power analysis for behavioral sciences*. Second Edition. Hillsdale: Lawrence Erlbaum Associates, Publishers.
- 8 Curran P (2002). Apontamentos do curso sobre Modelação da Performance – Modelação Hierárquica. Laboratório de Cineantropometria. FCDEF-UP. Porto.
- 9 Heck RH, Thomas SL (2000). *An introduction to multilevel modeling techniques*. Mahwah: Lawrence Erlbaum Associates, Publishers.
- 10 Hox J (1998). Multilevel modeling: When and why. In: Balderjahn I, Mathar R, Schader M (eds.). *Classification, Data Analysis and Data Highways*. New York: Springer Verlag, 147-154.
- 11 Hox J (2002). *Multilevel analysis. Techniques and applications*. Mahwah: Lawrence Erlbaum Associates, Publishers.
- 12 Kreft I, de Leeuw J (1998). *Introducing multilevel modeling*. Thousand Oaks: Sage Publications Ltd.
- 13 Maas CJM, Hox J (2000). Sample sizes for multilevel modeling. Department of Methodology and Statistics. Faculty of Social Sciences, Utrecht University. Documento não publicado.
- 14 MacCallum RC, Kim C (2000). Modeling multivariate change. In: Little TD, Schnabel KU, Baumert J (eds.). *Modeling Longitudinal and Multilevel Data. Practical Issues, Applied Approaches and Specific Examples*. Mahwah: Lawrence Erlbaum Associates, Publishers, 51-68.
- 15 Maia JAR, Lopes VP, Morais FP, Silva RMG, Seabra A (2002). *Estudo do crescimento somático, aptidão física, actividade física e capacidade de coordenação corporal de crianças do 1º ciclo do ensino básico da região autónoma dos Açores*. DREFD, DRCT, FCDEF-UP.
- 16 Nevill AM, Holder RL (2000). Modeling health-related performances indices. *Annals of Human Biology*, 6:543-559.
- 17 Plewis I (1994). Longitudinal multilevel models. In: Dale A, Davies RB (eds.). *Analyzing Social and Political Change*. London: Sage Publications, Ltd, 118-135.
- 18 Raudenbush S, Bryk A, Cheong YF, Congdon R (2001). *HLM 5. Hierarchical linear and nonlinear modeling*. Chicago: Scientific Software International.
- 19 Raudenbush, SW, Bryk AS (2002). *Hierarchical linear models. Applications and data analysis methods*. Second Edition. Thousand Oaks: Sage Publications, Ltd.
- 20 Round JM, Jobes DA, Honour JW, Nevil AM (1999). Hormonal factors in the development of differences in strength between boys and girls during adolescence: a longitudinal study. *Annals of Human Biology*, 26: 49-62.
- 21 Rowland T (1996). *Developmental exercise physiology*. Champaign: Human Kinetics Publishers.
- 22 Singer JD (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics*, 24: 323-355.
- 23 Singer JD (2002). Fitting individual growth models using SAS PROC MIXED. In: Moskowitz DS, Hershberger SL (eds.). *Modeling Intraindividual Variability with Repeated Measures Data*. Mahwah: Lawrence Erlbaum Associates, Publishers, 135-170.
- 24 Snijders T, Bosker R ((2002). *Multilevel analysis. An introduction to basic and advanced multilevel modeling*. London: Sage Publications Ltd.
- 25 Zhu W (1997). A multilevel analysis of schools factors associated with health-related fitness. *Research Quarterly for Exercise and Sport* 2:125-135.