

EDUCARE

Ano X - Nº ESPECIAL - 2004

ISSN nº 0873-0504

EDUCERE



Aprendizagens e Avaliação Nas Áreas Artísticas



REVISTA DA ESCOLA SUPERIOR DE EDUCAÇÃO
DE CASTELO BRANCO

EDIÇÃO E PROPRIEDADE
Escola Superior de Educação de Castelo Branco

DIRECTORA
Helena Menezes

COORDENAÇÃO
Helena Menezes
Luísa Reis
Madalena Leitão
Mário Fonseca

CONSELHO EDITORIAL

António Camilo, António Rafael, António Faustino, Clotilde Agostinho, Cristina Pereira, Dolores Correia, Domingos Santos, Eduarda Santos, Ernesto Martins, Fátima Boliqueime, Fátima Matela, Fátima Paixão, Fátima Regina, Gabriela Nunes, Helena Menezes, Helena Mesquita, Helena Tomás, Henrique Gil, João Petrica, João Ruivo, João Serrano, José Pires, José Rafael, Luis Costa, Luisa Reis, Lurdes Barata, Lurdes Cardoso, Madalena Leitão, Manuela Abrantes, Margarida Afonso, Margarida Morgado, Maria João Moreira, Natividade Pires, Paulo Afonso, Sofia Pissarra, Teresa Gonçalves, Valter Lemos e Virginia Brunheta

APOIO À EDIÇÃO
Centro de Recursos da Escola Superior de Educação

PRODUÇÃO
Instituto Politécnico de Castelo Branco

CAPA
Rosário Quelhas

CONCEPÇÃO GRÁFICA
José Castela

REVISÃO DE TEXTO

Fátima Boliqueime
Teresa Gonçalves

EXECUÇÃO GRÁFICA
Serviços Gráficos IPCB

IMPRESSÃO
IPCB

ASSINATURAS
Escola Superior de Educação de Castelo Branco
Rua Prof. Dr. Faria de Vasconcelos
6000-266 Castelo Branco, Portugal,
Telef. 272339100, Fax. 272343477
Email: esecb@mail.es.ipcb.pt

TIRAGEM
500 exemplares

PERIODICIDADE
A revista EDUCARE / EDUCERE é editada semestralmente (2 números/ano)

PREÇO E ASSINATURA
Assinatura anual (2 números): Portugal - 7,50Euro + 2,00 Euro de portes de correio;
Outros países - 7,50 Euro + portes de correio; Avulso 4,00 Euro.

ÍNDICE

Editorial

5

I AVALIAÇÃO

Estudo de validade do teste *Intermediate Measures of Music Audiation* de Edwin Gordon para uma situação de desempenho musical de crianças do 1º ano de escolaridade.

11

Helena Rodrigues

Avaliação externa nas Artes Visuais em Portugal e Inglaterra

35

Teresa Eça

O Portfolio como instrumento Avaliativo na Formação Inicial de Professores de Artes Visuais

49

Leonardo Charréu

Avaliação da expressão plástica no contexto da Educação de Infância

67

Mário Fonseca

Artes Visuais e Minorias Étnicas

Um Estudo de caso junto da Comunidade Cigana de Évora:
A Proposta Pedagógica

"O caminho pelas Artes - Desenhar a Escola"

83

*Leonardo Charréu*II ASPECTOS CURRICULARES
E APRENDIZAGENS

Porquê a Arte na Escola?

- Notas sobre Educação em Arte

127

Margarida Rocha

Sobre Educação Artística e Programas Artísticos Integrados:
O caso paradigmático das aprendizagens e ensino das
artes visuais na educação especial

137

Maria do Pilar de Andrade Falcão

- Abrir o Sentido da Criatividade:**
A música na actual reorganização curricular do Ensino Básico
na Perspectiva dos Professores. 149
António Ângelo de Vasconcelos
Sandra Barbosa

- A Dança e o seu Valor na Educação Rítmica e Social**
da Criança/Jovem 179
Luísa dos Santos Reis

- Drama e Desenvolvimento Moral:**
Perspectivas, modalidades e alguns equívocos 189
Maria Madalena Leitão

III PROJECTOS EDUCATIVOS

- Projecto Literatura & Literacia**
- Um Percurso do Pré-escolar ao final do 2º Ciclo 211
Maria da Conceição Rolo

- MUS-E na Escola Básica EB1 de Algés** 215
Carmelinda Pereira

ESTUDO DE VALIDADE DO TESTE INTERMEDIATE MEASURES OF MUSIC AUDITION DE EDWIN GORDON PARA UMA SITUACAO DE DESEMPENHO MUSICAL DE CRIANCAS DO 1º ANO DE ESCOLARIDADE

Helena Rodrigues*

Resumo

Foi realizado um estudo de validade de uma metodologia que consiste em aplicar o teste Intermediate Measures of Music Audition (IMMA) de Edwin Gordon, tendo como contexto de realização o estudo da validade preditiva após instrução. O IMMA foi aplicado antes e 27 crianças do 1º ano do 1º Ciclo de escolaridade, antes e depois de um período de instrução musical. Durante este período, duas melodias foram apresentadas às crianças. No termo do período de instrução musical foi gravada a cantabilidade cantada, por cada criança individualmente, de ambas as melodias. A validade do desempenho musical foi classificada através de análises de cantabilidade por três juízes independentes, tendo sido atribuídos elevados coeficientes de fiabilidade.

Os resultados do IMMA antes e após o período de instrução foram correlacionados com os resultados no teste de desempenho musical. Foram encontradas relações elevadas, sendo os coeficientes de validade concomitante ligeiramente superiores aos da validade preditiva.

Pode assim concluir-se que os resultados do IMMA, para quando será dos estudos futuros, poderão ser utilizados para outras situações de desempenho musical.

AVALIAÇÃO

Abstract

The aim of this study was to evaluate the extent developmental music aptitude test Intermediate Measures of Music Audition (IMMA) for a Portuguese population.

The test was administered to 27 first grade school children before and after a six month period of musical instruction. During this period, two songs were taught to the children. After the musical instruction period was IMMA applied, which required singing both songs. These performances were evaluated by three independent judges using Gordon's reliability scale. The interjudge reliability and the intrajudge reliability were quite high.

IMMA pre-instruction results and IMMA post-instruction results were correlated with the vocal achievement rating scale results. High values were found for both situations, concomitant validity being slightly higher than predictive validity.

We may conclude that results in IMMA not only predict the success of children singing. Future studies will investigate IMMA productivity for other music achievement situations.

A avaliação da aptidão musical é um tema clássico no âmbito da Psicologia da Música. O nascimento desta disciplina está, aliás, associado às contribuições de Carl Seashore, a quem se deve a criação da primeira

Foramadas de Ciências Sociais e Humanas - Universidade Nova de Lisboa

ESTUDO DE VALIDADE DO TESTE *INTERMEDIATE MEASURES OF MUSIC AUDIATION* DE EDWIN GORDON PARA UMA SITUAÇÃO DE DESEMPENHO MUSICAL DE CRIANÇAS DO 1º ANO DE ESCOLARIDADE

Helena Rodrigues*

Resumo

Este estudo integra-se no âmbito de uma investigação que conduziu à adaptação portuguesa do teste *Intermediate Measures of Music Audiation (IMMA)* de Edwin Gordon. Neste contexto foi realizado o estudo de validade preditiva aqui relatado.

O IMMA foi administrado a 37 crianças do 1º ano do 1º Ciclo de escolaridade, antes e depois de um período de instrução musical. Durante este período, duas melodias foram ensinadas às crianças. No termo do período de instrução musical foi gravada a execução cantada, por cada criança individualmente, de ambas as melodias. A avaliação do desempenho musical foi classificada através da utilização de *rating-scales*, por três juizes independentes, tendo sido encontrados elevados coeficientes de fiabilidade.

Os resultados no IMMA antes e após o período de instrução foram correlacionados com os resultados na prova de desempenho vocal. Foram encontrados valores elevados, sendo os coeficientes de validade concorrente ligeiramente superiores aos da validade preditiva.

Pode assim concluir-se que os resultados no IMMA predizem parcialmente o desempenho vocal das crianças. Estudos futuros deverão investigar a preditividade do IMMA para outras situações de desempenho musical.

Abstract

The aim of this study was to validate Gordon's developmental music aptitude test *Intermediate Measures of Music Audiation (IMMA)* for a portuguese population.

This test was administrated to 37 first grade school children before and after a six month period of musical instruction. During this period, two songs were taught to the children. After the musical instruction period each child was individually recorded singing both songs. These performances were evaluated by three independent judges using a Tonal and a Rhythm rating scale. The interjudge reliability and the intrajudge reliability were quite high.

IMMA pre-instruction results and IMMA post-instruction results were correlated with the vocal achievement rating scale marks. High values were found for both situations, concurrent validity being slightly higher than predictive validity.

We may conclude that results in IMMA partialy predict the success of children singing. Future studies will investigate IMMA predictivity for other music achievement situations.

A avaliação da aptidão musical é um tema clássico no âmbito da Psicologia da Música. O nascimento desta disciplina está, aliás, associado às contribuições de Carl Seashore, a quem se deve a criação da primeira

* Faculdade de Ciências Sociais e Humanas - Universidade Nova de Lisboa.

bateria de testes no domínio e a obra *The Psychology of Musical Talent*, publicada em 1919.

Embora a temática do talento musical seja uma velha questão, é sobretudo a partir de Seashore que se desencadeiam debates mais acesos e surgem diversos instrumentos de avaliação destinados a medir a aptidão musical. Na actualidade, o contributo de Edwin Gordon é, neste domínio, absolutamente ímpar.

Edwin Gordon define a aptidão musical como o potencial para aprender. Distingue claramente entre aptidão e realização musical (aquilo que foi alvo de aprendizagem), embora ambas se situem num mesmo *continuum*, não sendo mutuamente exclusivas. Não é possível construir um teste puro para avaliar a aptidão musical da mesma forma que também não é possível construir um teste puro para avaliar a realização musical.

A base da aptidão musical é a audição, termo este cunhado pelo autor e que designa a capacidade para ouvir e compreender musicalmente quando o som não está fisicamente presente. Ou seja, a audição está para a Música assim como o pensamento está para a linguagem.

Edwin Gordon acredita que no momento do nascimento estão lançadas as bases da aptidão musical de um indivíduo, sendo a altura em que esta apresenta o seu maior nível. Até aos nove anos de idade é possível restabelecer este nível mas, posteriormente, o que se pode melhorar diz respeito apenas à realização musical. Pode assim falar-se em dois tipos de aptidão: em desenvolvimento e estabilizada.

Gordon defende ainda que a aptidão musical se distribui segundo a curva normal, à semelhança de outras aptidões. Ou seja, não há dúvida de que todas as pessoas são capazes de aprender música. A questão está em saber em que extensão, até que nível. Todas as pessoas são capazes de aprender música, embora não aproveitem da mesma maneira essa instrução. Finalmente, a multidimensionalidade é outra característica da aptidão musical, tendo o autor, ao longo do seu trabalho, identificado mais de vinte dimensões. Todas estão interrelacionadas, embora possuam também a sua unicidade própria. Normalmente um sujeito tem diferentes níveis de aptidão nas várias dimensões.

Visando avaliar a aptidão musical em desenvolvimento, Gordon começou por procurar adaptar o teste *Musical Aptitude Profile* (MAP) destinado a estudantes a frequentar entre o 4º e o 12º anos de escolaridade. O autor desenvolveu outras folhas de resposta, experimentou que fossem adultos a escrever as respostas das crianças, etc. Acabou por concluir, no entanto, que para obter um teste de aptidão musical válido para crianças pequenas, tudo deveria ser modificado: desde as direcções do teste, às folhas de resposta, até ao próprio conteúdo do teste.

É assim que entre 1971 e 1978 Gordon desenvolve um conjunto de três investigações, envolvendo 1114 padrões tonais e 486 padrões rítmicos estudados junto de 10 000 crianças, que o hão-de conduzir à elaboração de uma taxonomia de padrões tonais e rítmicos. A elaboração de uma taxonomia prendia-se com um princípio básico da filosofia dos testes criados por Gordon que é a de o ouvinte encontrar significado musical naquilo que ouve, em termos tonais e rítmicos. Isto é, ao autor não interessava saber se o ouvinte conseguia discriminar entre dois sons de altura diferente ou de duração diferente - o que apenas revelaria uma capacidade de âmbito acústico e não requereria pensamento musical, ou categorização em termos de audição - mas antes fazer com que o sujeito respondesse de forma imediata e intuitiva.

O teste *Intermediate Measures of Music Audiation* (IMMA), criado por Edwin Gordon em 1982, baseia-se naqueles estudos, sendo constituído pelos padrões difíceis da taxonomia. O teste *Primary Measures of Music Audiation* (PMMA), baseado também nos mesmos estudos, é constituído pelos padrões fáceis da taxonomia. Ambos avaliam duas dimensões a que correspondem os testes Tonal e Rítmico.

De referir, no entanto, que nos estudos preliminares Gordon procurou incluir testes que medissem as dimensões timbre e intensidade. Não conseguiu, porém, encontrar índices de fiabilidade aceitáveis dado que os itens eram ou demasiado fáceis ou demasiado difíceis. Chegou ainda à conclusão de que as crianças no estágio da aptidão musical em desenvolvimento podem concentrar-se apenas numa dimensão musical - isto é, as suas decisões são fieis quando o padrão tonal que estão a ouvir não está inserido num contexto rítmico e quando o padrão rítmico não é

apresentado num contexto tonal - sendo por isso necessário apresentar as dimensões tonal e ritmo separadamente. Não foi também possível encontrar medidas fiáveis da preferência musical em termos de expressividade musical.

O autor concluiu ainda ser preferível utilizar um sintetizador na gravação do conteúdo musical deste teste do que um instrumento acústico pois esta situação oferecia índices de fiabilidade mais elevados. Finalmente, a não ser que se fornecesse algum apoio (como, por exemplo, a pulsação definidora do tempo do padrão em simultâneo ou previamente à sua definição) as crianças evidenciavam dificuldades na organização dos padrões rítmicos. Por esta razão, a gravação do teste Rítmico apresenta uns "cliques" subjacentes à audição dos padrões.

O IMMA compõe-se, pois, de dois testes, um destinado a avaliar a aptidão tonal (teste Tonal) e o outro a aptidão rítmica (teste Rítmico). Cada teste compõe-se de alguns exemplos demonstrativos e de 40 itens em que a criança ouve um par de padrões e tem de detectar se são iguais ou diferentes. O teste Tonal é composto por um par de padrões tonais (um padrão tonal é uma série de sons de altura diferente mas com a mesma duração). Cada frase do teste Rítmico é também composta por um par de padrões rítmicos (um padrão rítmico é uma série de sons com a mesma altura mas com durações diferentes).

Para responder não é necessário saber ler ou escrever música: cada criança dispõe de uma folha de respostas em que, para cada item, há um desenho de um par de faces iguais e um desenho de um par de faces diferentes, devendo a criança assinalar com um círculo um dos pares, conforme julgue que o par de frases musicais ouvido é, também, igual ou diferente. Cada um dos quarenta itens de cada teste está identificado na cassete com o nome de um objecto (ex.: carro, colher) a que corresponde na folha de respostas o mesmo objecto. Depois de ouvir anunciar esse objecto, a criança ouve a palavra "primeiro" seguida da primeira frase do par e depois a palavra "segundo" seguida da segunda frase do par. Entre cada item há cinco segundos para a criança marcar a sua resposta: se as duas frases soarem iguais a criança deverá desenhar, na folha de respostas, um círculo à volta do desenho de duas faces iguais; se as duas

frases soarem diferentes deverá desenhar um círculo à volta do desenho contendo duas faces diferentes. Ou seja, também não é necessário dominar a linguagem escrita. As direcções verbais que acompanham o teste são padronizadas.

De acordo com a reputada publicação no domínio da Psicometria *The Mental Measurements YearBook*, os testes de aptidão musical de Edwin Gordon apresentam requisitos técnicos excelentes, nomeadamente no que diz respeito a questões de fiabilidade e validade. A validade do IMMA é o objecto de interesse neste estudo.

Segundo a definição tradicional, a validade de um teste refere-se ao facto de o teste medir ou não aquilo que se propõe. Esta definição tem sido posta em causa quer pela circularidade do princípio (Karma, 1982) como também pela observação de que testes elaborados por diferentes autores, com diferentes concepções de aptidão musical possuem, naturalmente, diferentes validades de construto. Assim, à definição tradicional opõe-se hoje a noção de que um teste é válido ou não em função do fim a que se destina. Ou seja, o mesmo teste pode ser válido ou inválido consoante a aplicação que se lhe destine.

A respeito desta questão, recorde-se que Gronlund e Linn fazem notar que é mais correcto falar da validade da interpretação dos resultados de um teste que da "validade" de um teste - *"They maintain that validity is a unitary concept, for which a test maker provides various kinds of evidence. Depending on the purpose for which the test is to be used, the test maker should provide evidence that is content related, criterion related, or construct related"* (in Boyle, 1992, p. 252). Portanto, não se pode declarar liminarmente se um teste é ou não é válido; o que tem é de se demonstrar a sua validade para um determinado uso ou interpretação. Em vez de se afirmar que um teste é válido, questione-se a validação das finalidades para as quais o teste é usado. A validade depende, pois, da finalidade e utilização dos resultados do teste. Resumindo: *"a test is valid to the extent that inferences made from it are appropriate, meaningful and useful"* (Gregory, 1996, p. 633).

Seguindo esta linha de pensamento, a validade preditiva de um instrumento que se proponha medir a aptidão musical é da maior relevância.

Como refere Gordon (1986), a validade estabelecida com base num critério concorrente é um substituto pobre da validade estabelecida com base num critério preditivo, já que não se pode estabelecer uma relação causal entre ambas as variáveis. Pelo contrário, um estudo de validade longitudinal permite relacionar dados obtidos através dos instrumentos de avaliação com comportamentos musicais posteriores dos indivíduos, permitindo, por um lado, inferir causalidade e, por outro, escapar ao "círculo epistemológico" a que anteriormente se aludiu. Por estas razões, a realização de estudos longitudinais de validade preditiva é crucial no estabelecimento da validade de um teste de aptidão musical. Esse é o fulcro do estudo que aqui se apresenta.

Finalidade e problema

O objectivo do presente estudo é verificar a validade preditiva do teste *Intermediate Measures of Music Audiation* (IMMA) de Edwin Gordon.

Este estudo surge integrado num estudo mais amplo relativo às qualidades psicométricas deste teste em que se procurou derivar dados estatísticos que pudessem ajudar na interpretação dos seus resultados comparando-as com os da amostra americana e fazer a aferição do teste para a área educativa de Lisboa (cf. Manual do IMMA, 2001). Da amostra de 1133 crianças, extraíram-se amostras mais reduzidas com as quais se efectuaram estudos de fiabilidade teste-reteste e de validade preditiva.

Justamente, o estudo aqui relatado pretendeu verificar a preditividade do IMMA para uma situação de desempenho musical. Concretamente, definiu-se e circunscreveu-se a situação de desempenho musical à capacidade de aprender e entoar uma canção. Ou seja, pretende-se responder à questão: de que forma os resultados obtidos no IMMA predizem o desempenho vocal de um aluno?

Sujeitos

A amostra deste estudo é constituída por 37 crianças (13 do sexo masculino e 24 do sexo feminino), com idades entre os seis e os sete anos, frequentando o 1º ano de escolaridade do 1º Ciclo da área educativa de Lisboa. Foi seleccionada de entre as 1133 crianças, assinalando duas turmas de uma das escolas que participaram no estudo de adaptação do IMMA.

Instrumentos

Para a avaliação da aptidão musical foi usada a versão portuguesa do IMMA, traduzida e adaptada previamente pela autora.

Para a avaliação do desempenho musical foram criadas as seguintes melodias, procurando-se que as mesmas tivessem características musicais diferentes (uma em tonalidade menor e métrica binária, a outra em tonalidade Maior e métrica ternária) mas idêntico grau de dificuldade :

Melodia A)



Melodia B)



Para a avaliação da execução cantada destas melodias foram criadas duas *rating scales*, uma para avaliação da dimensão Tonal, outra para avaliação da dimensão Ritmo.

Rating scale para avaliação da dimensão Tonal:

1 - termina na tónica eventualmente, mas a afinação é muito fraca

2 - começa na altura certa ou termina na tónica, percebe-se se é maior ou menor

3 - entoa evidenciando a função da tónica, com afinação razoável

4 - entoa evidenciando a função da dominante e tónica; entoa com boa afinação

5 - entoa tudo correctamente, com bastante boa afinação

Rating scale para avaliação da dimensão Ritmo:

1- pode começar no tempo certo, pode haver sentido métrico mas não o correcto

2- o tempo mantém-se correcto durante toda a performance ou há ligeiros ajustamentos

3 - há sentido de pulsação e de métrica

4 - o(s) padrão(ões) assinalados com * são executados correctamente e há um bom sentido rítmico

5 - pulsação e métrica correctas; os padrões rítmicos assinalados com ** são bem executados

Para o registo das classificações atribuídas pelos professores foi elaborada uma ficha de avaliação.

Metodologia

Começou por se contactar o director da respectiva escola, solicitando a colaboração da mesma, tendo-se explicado pessoalmente quais os objectivos do estudo e entregue um resumo escrito do mesmo. Depois do director ter contactado os professores das duas turmas assinaladas, foram marcadas as datas de administração do IMMA seguindo as indicações do manual.

Os testes Tonal e Ritmo foram administrados em dias diferentes, dentro da mesma semana, no início do ano lectivo, antes de se iniciar o período de instrução musical. Durante a aplicação do teste, os professores de cada classe permaneceram na sala de aula tendo colaborado na sua administração.

Ao longo de um período de seis meses, foi facultada instrução musical a cada uma das turmas, tendo a investigadora leccionado cada turma por um período de 40 minutos uma vez por semana. Os professores de ambas as classes revelaram-se extremamente interessados e colaboradores.

A instrução musical ministrada incluía prática vocal e o ensino das duas melodias acima descritas. Estas eram ensinadas em sílaba neutra, alternadamente em cada semana.

No termo do período de instrução musical foi gravada a execução cantada de ambas as melodias. A gravação foi efectuada individualmente com cada criança, depois da investigadora entoar cada uma das melodias, tendo sido deixado algum espaço de tempo livre entre a gravação de ambas as melodias.

Este procedimento foi repetido no dia seguinte. Ou seja, a execução cantada de ambas as melodias foi gravada duas vezes para cada criança, em dois dias seguidos.

O IMMA foi readministrado também no termo do período de instrução.

Para a avaliação do desempenho musical foram contactados quatro professores de Educação Musical com experiência de trabalho com crianças desta faixa etária mas sem contacto com estas crianças. Após um pré-estudo relativo à fiabilidade das suas avaliações seleccionaram-se os três que ofereciam melhores garantias de consistência interna.

Depois de se discutir as *rating-scales* com estes juizes, entregaram-se cópias das duas gravações das melodias cantadas pelas crianças, pedindo-lhes que atribuísem classificações entre 1 e 5 de acordo com aquelas *rating-scales*. No caso de a criança não fazer nada do que se lhe pedia, seria atribuída a classificação de 0. A cada juiz foram facultadas as informações necessárias tendo, cada um, feito a sua avaliação em casa, sem contacto com qualquer dos outros juizes.

Análise dos resultados

Os resultados da média e do desvio-padrão obtidos no IMMA por estas duas classes foram, respectivamente, de 29.7 e 4.46 para o teste Tonal e de 26.5 e 4.32 para o teste Rítmico. Ou seja, são valores semelhantes aos encontrados para o total de sujeitos do 1º ano de escolaridade, o que sugere tratar-se de uma boa amostra da amostra total utilizada na adaptação do teste.

Relativamente às classificações atribuídas pelos juizes verificou-se que as médias das classificações atribuídas são mais elevadas na dimensão tonal que na rítmica, que, de uma forma geral, a melodia no modo Maior e com divisão ternária obteve pontuações ligeiramente mais elevadas, e que na repetição do procedimento as crianças obtiveram resultados ligeiramente superiores.

Foi calculada a fiabilidade dos instrumentos utilizados.

A fiabilidade do IMMA estudada na população portuguesa apresentou dados aceitáveis ainda que inferiores aos encontrados na população americana. Concretamente, para os resultados Compósitos encontraram-se valores entre .67 e .85 no cálculo teste-reteste e .80 e .82 no método das duas metades corrigido pela fórmula de Spearman-Brown (cf. Manual do IMMA).

A fiabilidade da medida de desempenho musical (medida usada como critério no estudo de validade) foi também objecto de análise, quer através do cálculo da correlação entre as classificações obtidas em ambas as gravações por cada criança, como do estabelecimento de correlações entre as classificações atribuídas pelos três juizes. Estes valores estimam a fiabilidade inter-juizes, enquanto que os primeiros oferecem uma estimativa da fiabilidade teste-reteste (dando a ideia da estabilidade de ambas as execuções das crianças) e, também, uma estimativa da fiabilidade intra-juiz.

Assim, como estimativa da fiabilidade intra-juiz e da estabilidade das execuções da canção por parte das crianças foram calculadas as

correlações entre a 1ª e a 2ª gravação relativamente às classificações atribuídas por cada juiz a cada uma das canções, avaliando-as nos parâmetros Tonal e Ritmo. Determinaram-se também as correlações entre a 1ª e a 2ª gravação dos compósitos (soma do Tonal e do Ritmo) das classificações atribuídas por cada juiz bem como da soma das classificações dos três juizes.

Quadro 1 - Correlações entre as classificações atribuídas por cada juiz à 1ª e 2ª gravações (aspectos Tonal, Rítmico e Compósito da execução cantada das melodias)

	Juiz 1	Juiz 2	Juiz 3	3 Juizes
Melodia M / Tonal	.80	.81	.77	.83
Melodia m / Tonal	.84	.82	.52	.78
Melodia M / Ritmo	.72	.68	.66	.76
Melodia m / Ritmo	.74	.76	.70	.81
Melodia M / Compósito	.85	.82	.82	.87
Melodia m / Compósito	.85	.85	.75	.86

Em face destes resultados verifica-se que, quando considerados individualmente, os juizes 1 e 2 são mais fieis internamente, apresentando valores quase sempre acima de .70. O juiz 3, na avaliação do parâmetro Tonal da melodia no modo menor apresenta um valor menos satisfatório de .52 mas, de uma forma geral, os valores encontrados são encorajadores.

Quando se correlacionam as classificações obtidas em ambas as gravações e que foram atribuídas ao parâmetro Tonal (somando as classificações no aspecto Tonal obtidas nas duas canções) e as atribuídas ao parâmetro Ritmo (adoptando o mesmo procedimento) os valores sobem, como seria de esperar, respectivamente para .87 e .84. Quando se considera a correlação entre o total das classificações compósitas obtidas na 1ª e na 2ª gravação obtém-se o valor de .89.

Os dados referentes à fiabilidade inter-juizes revelaram-se também bastante satisfatórios. No quadro 2 as correlações apresentadas mostram o acordo entre os juizes na avaliação das duas canções nos parâmetros Tonal e Ritmo, na 1ª e na 2ª gravação.

Quadro 2 - Correlações entre as classificações atribuídas pelos três juizes (aspectos Tonal, Ritmo e Compósito da execução cantada das melodias

Classif. na 1ª gravação

	Juiz 1 / 2	Juiz 1 / 3	Juiz 2 / 3
Melodia M / Tonal	.85	.91	.86
Melodia m / Tonal	.83	.73	.72
Melodia M / Ritmo	.90	.74	.81
Melodia m / Ritmo	.87	.65	.65
Melodia M / Comp	.92	.89	.91
Melodia m / Comp	.90	.77	.79

Classif. na 2ª gravação

	Juiz 1 / 2	Juiz 1 / 3	Juiz 2 / 3
Melodia M / Tonal	.93	.95	.89
Melodia m / Tonal	.89	.75	.76
Melodia M / Ritmo	.86	.82	.79
Melodia m / Ritmo	.85	.75	.74
Melodia M / Comp	.94	.93	.90
Melodia m / Comp	.92	.90	.88

Quando estas correlações são calculadas considerando em conjunto todas as tarefas em que o aspecto Tonal está a ser avaliado ou todas as tarefas em que o aspecto Ritmo está a ser avaliado, estes coeficientes, como seria esperado, sobem e situam-se à volta de .90.

Perante os dados reportados nos Quadros 1 e 2 poder-se-á concluir que as condições de fiabilidade ligadas à avaliação do desempenho musical das canções efectuada pelos juizes são satisfatórias, pelo que se passa à análise dos resultados da validade.

Para o cálculo da validade preditiva determinaram-se as correlações entre os valores obtidos no IMMA antes de iniciado o período de instrução musical e as classificações atribuídas pelos três juizes a cada uma das situações de avaliação do desempenho musical anteriormente definidas. Como é óbvio, as correlações estabelecem-se entre as avaliações do mesmo aspecto: isto é, os resultados no teste Tonal foram correlacionados com as classificações atribuídas pelos juizes ao aspecto Tonal; os resultados no teste Rítmico foram correlacionados com as classificações atribuídas pelos juizes ao aspecto Ritmo e os resultados no Compósito com o total das classificações atribuídas pelos três juizes. Os valores encontrados são bastante satisfatórios, como pode ser verificado consultando o quadro 3.

Quadro 3 - Correlações entre as classificações atribuídas pelos três juizes e os resultados obtidos no IMMA antes do período de instrução

Avaliação na 1ª gravação

	Teste T	Teste R	Compósito
Melodia M.	.51**	.42**	.56**
Melodia m	.60**	.38*	.57**
Total Tonal	.59**		
Total Ritmo		.46**	
Total Classif.			.61**

Avaliação na 2ª gravação

	Teste T	Teste R	Compósito
Melodia M	.44**	.54**	.66**
Melodia m	.57**	.55**	.80**
Total Tonal	.52**		
Total Ritmo		.55**	
Total Classif.			.74**

(* $p < .05$; ** $p < .001$)

Como se pode verificar, todos os valores são estatisticamente significativos. Os valores encontrados, para os resultados Compósitos, foram de .61 e .74.

O teste foi ainda readministrado no termo do período de instrução musical, tendo-se encontrado valores superiores aos que tinham sido encontrados na primeira administração. A diferença entre as médias dos resultados obtidos antes e após o período de instrução revelou-se significativa ao nível de .05.

Como estudo de validade concorrente foram determinadas as correlações entre estes resultados (valores obtidos no IMMA no termo do período de instrução musical) e as classificações atribuídas pelos juizes, tendo-se encontrado valores significativos ao nível de .001. (cf. quadro 4).

Quadro 4 - Correlações entre as classificações atribuídas pelos três juizes e os resultados obtidos no IMMA após o período de instrução

Avaliação na 1ª gravação

	Teste T	Teste R	Compósito
Melodia M	.48**	.55**	.60**
Melodia m	.61**	.38*	.52**
Total Tonal	.57**		
Total Ritmo		.52**	
Total Classif.			.61**

Avaliação na 2ª gravação

	Teste T	Teste R	Compósito
Melodia M	.64**	.67**	.75**
Melodia m	.68**	.68**	.76**
Total Tonal	.69**		
Total Ritmo		.68**	
Total Classif.			.76**

(* $p < .05$; ** $p < .001$)

À semelhança do que acontecia para o quadro anterior, todos os valores são estatisticamente significativos.

Interpretação e discussão dos resultados

Os valores de .61 e de .74 (coeficientes de correlação entre os compósitos de ambas as provas encontrados no estudo de validade

predictiva) mostram que o nível de aptidão musical tal como é medido pelo IMMA explica, respectivamente, 37% e 55% das razões do sucesso das crianças na execução das tarefas musicais com que foram avaliadas pelos juizes. Trata-se de valores da mesma ordem de grandeza que os encontrados por Gordon (1986) que, em estudos semelhantes, obteve coeficientes de .67 e .70.

É evidente que se pode argumentar que a manifestação das competências musicais de uma criança se pode revelar doutros modos diferentes daqueles para os quais se procurou aqui testar se o teste era ou não predictivo. Mas ninguém pode negar que cantar uma melodia é uma das manifestações possíveis do comportamento musical. Os dados aqui relatados mostraram a predictividade do teste para esta situação concreta.

A diferença significativa encontrada entre os resultados obtidos no IMMA antes e depois do período de instrução musical pode ser interpretada como manifestação do carácter desenvolvimental da aptidão musical. Fica, no entanto, por esclarecer se esse aumento se deve ao facto de as crianças terem tido instrução musical durante este período de tempo, (e se se deve especificamente ao tipo de instrução musical recebida), ou se mesmo não tendo recebido instrução musical, se teriam verificado diferenças da mesma ordem de grandeza. Recorde-se, entretanto, que estes resultados são coerentes com os reportados noutros trabalhos (Flohr, 1981; Schleuter e Schleuter, 1989) em que se verificaram também aumentos significativos nos resultados do PMMA ao fim de períodos de instrução musical de três e seis meses.

Os valores encontrados no estudo de validade concorrente foram, de um modo geral, ligeiramente superiores aos de validade predictiva, o que é natural, dado o facto de os dois tipos de dados (pontuações obtidas no IMMA e classificações obtidas nas *rating -scales*) terem sido recolhidos aproximadamente na mesma altura. O facto de as crianças terem agora maior tempo de escolarização e, portanto, estarem mais familiarizadas e motivadas para o tipo de tarefas exigidas em ambas as provas pode ter aumentado a fiabilidade do IMMA, o que tem como consequência directa o aumento da validade. É certo que este facto pode também ser interpretado como uma demonstração de que o IMMA é, ele

próprio, um teste de desempenho musical, ideia esta que o autor tem rejeitado. Por outro lado, a constelação dos factos (o IMMA revelou-se um teste preditivo do desempenho vocal; os valores obtidos no IMMA são superiores depois de um período de instrução musical) é igualmente coerente com a ideia defendida por Gordon de que o IMMA mede a aptidão musical em desenvolvimento.

Conclusão

O estudo longitudinal apresentou resultados muito satisfatórios em termos de validade preditiva, evidenciando também o carácter desenvolvimental da aptidão musical. Os resultados obtidos neste estudo confirmam a validade do IMMA para a situação de desempenho musical estudada e permite colocar à disposição da comunidade científica portuguesa um teste de aptidão musical para o qual passou a haver dados de comparação com a amostra americana. No que concerne à avaliação da aptidão musical é a primeira vez que esta questão é abordada - até ao momento não havia qualquer teste de aptidão musical que tivesse sido estudado para a população portuguesa.

Os factos encontrados no estudo longitudinal apresentam grande coerência. Na sua base estarão aliadas as condições criadas para a apreciação das tarefas de realização musical em que, recorde-se, os juizes limitaram a avaliação do desempenho musical à avaliação da execução gravada de duas canções. Deste modo, foi possível operacionalizar e objectivar melhor os critérios de avaliação.

De modo geral, observa-se que neste caso a consistência inter-juizes apresentou melhores resultados que a consistência interna de cada juiz.

Estas reflexões permitem indirectamente questionar os procedimentos tradicionalmente utilizados em escolas do ensino vocacional de Música relativamente à avaliação da aptidão musical. Torna-se assim mais relevante o facto de, a partir de agora, se poder contar com um instrumento de avaliação que pode desafiar expectativas e complementar

as tradicionais fontes de informação subjectiva. Fica, no entanto, em aberto a necessidade de no futuro se levarem a cabo estudos referentes à avaliação do desempenho musical efectuadas por professores e juizes independentes, estendendo-a a outras situações para além do desempenho vocal.

Note-se ainda que a amostra utilizada é proveniente de um meio urbano do litoral. Embora estudos efectuados previamente pelo autor tenham demonstrado a irrelevância de factores como a localização geográfica e o estatuto sócio-económico na determinação dos resultados num teste de aptidão musical, em termos de aptidão musical em desenvolvimento poderão ser relevantes variáveis relacionadas com a estimulação musical do meio familiar, eventualmente definidas em termos mais globais como "estilo de vida". Neste sentido, seria desejável a replicação deste estudo numa localidade de meio rural e interior, em que o "estilo de vida" e o ambiente musical familiar associado são muito provavelmente diferentes do vivenciado pelas crianças da região de Lisboa.

Será também importante que em futuras investigações se proceda à aplicação individual do teste, comparando-a com resultados provenientes da aplicação colectiva. Outros estudos a efectuar poderão debruçar-se sobre as instruções de administração do teste: testar a utilização dos termos "igual / não igual" ou "mesmo / diferente" em vez do par "igual / diferente".

Ao disponibilizar-se este teste não se podem deixar de sublinhar todos os cuidados com que a administração de um teste deve ser rodeada. É importante recordar o que Gordon escreve no manual do teste: nenhuma criança deve ser afastada de um plano de estudos musicais com base nos resultados do mesmo. O diagnóstico psicológico e educativo fornecido por um instrumento de avaliação tem que estar contextualizado e ser consequente com medidas educativas a tomar. Doutro modo, as potencialidades de um teste convertem-se numa mera burocracia da medida destituída de qualquer sentido.

A utilização do IMMA deve contribuir para o melhoramento da qualidade de instrução musical das nossas crianças, adaptando-a às suas

diferenças individuais. Para além de permitir avaliar normativamente a aptidão musical (isto é, comparar os alunos entre si), de forma a que a educação musical proporcionada tenha em conta as diferenças interindividuais, outra das finalidades do IMMA é analisar ideograficamente cada criança (isto é, comparar os seus resultados em ambos os testes) para que o ensino se possa adaptar às suas necessidades e perfil específico.

Dadas as estreitas relações defendidas por Gordon entre a avaliação da aptidão musical e a teoria da aprendizagem musical, a disponibilização do IMMA é um suporte importante na implementação dos currículos *Jump Right In: The Music Curriculum* e *Jump Right In: The Instrumental Series* formulados pelo autor, visando a adaptação da intervenção educativa às necessidades individuais de cada criança. A teoria de aprendizagem musical interroga-se sobre o como as crianças aprendem; esta interrogação traz implícita a questão das diferenças individuais pois crianças diferentes aprendem de modos diferentes. O IMMA pode ajudar a esclarecer esta diferença. Para além das potencialidades investigativas oferecidas pelo IMMA, grande parte do valor do teste aqui estudado reside, pois, no que venha a estimular em termos de práticas educativas com as nossas crianças numa fase etária fulcral do seu desenvolvimento em termos musicais.

É importante ter presente que um teste como este é um instrumento de avaliação objectiva mas sempre com uma margem de erro associada (mas, apesar de tudo, inferior aos procedimentos tradicionais que pretendem avaliar a aptidão musical), expressa, nomeadamente, através dos coeficientes de validade e fiabilidade apresentados (cujos valores nunca são questionados nos procedimentos tradicionais).

Concretamente, há que ter presente que a aptidão musical, tal como é medida por este teste, é um, entre outros, dos factores responsáveis pelo sucesso das crianças em termos de desempenho musical. (Por exemplo, no estudo de um instrumento, as preferências pessoais de cada criança, o seu desenvolvimento emocional, aspectos motores e o apoio familiar são também importantes. Os resultados no teste *Instrument Timbre Preference* (Gordon, 1984), poderão também ser considerados). Por outro lado, o elevado coeficiente de validade preditiva encontrado foi

estabelecido para a entoação de canções, o que, representa, obviamente, apenas uma parte das possíveis manifestações da realização musical.

Pessoalmente, vejo o comportamento musical como um cristal cheio de arestas e clivagens. As provas de avaliação devem procurar retratar cada superfície desse comportamento da forma mais objectiva possível (ou, talvez melhor, o menos subjectiva possível), mostrando os vários ângulos da mesma realidade.

É importante não esquecer, no entanto, que cada cristal é sempre único.

BIBLIOGRAFIA

- Boyle, J.D. (1982). A study of the comparative validity of three published standardised measures of music preference. *Psychology of Music*, 10, 11-16.
- Boyle, J.D. (1989). Perspective on evaluation. *Musical Educators Journal*, 76, (4), 22-25.
- Boyle, J.D. (1992). *Evaluation of musical ability*. In R. Colwell (Ed.), *Handbook of research on music teaching and learning*. New York: Schirmer Books.
- Flohr, J. (1981). Short term music instruction and young children's developmental music aptitude. *Journal of Research in Music Education*, 29, (3), 219-223.
- Gordon, E. E. (1965). *Musical aptitude profile*. Boston: Houghton Mifflin Company.
- Gordon, E. E. (1967). *A three-year longitudinal predictive validity study of the musical aptitude profile*. Iowa City: University of Iowa Press.
- Gordon, E. E. (1968). A study of the efficacy of general intelligence and music aptitude tests in predicting achievement in music. *Bulletin of the Council for Research in Music Education*, 13, 40-45.
- Gordon, E. E. (1970). *The psychology of music teaching*. Englewood Cliffs: Prentice-Hall.
- Gordon, E. E. (1974). Toward the development of a taxonomy of tonal patterns and rhythms patterns: evidence of difficulty level and growth rate. *Experimental Research in the Psychology of Music: Studies in the Psychology of Music*, 8, 39-232.
- Gordon, E. E. (1976). *Tonal and rhythms patterns: an objective analysis*. Albany: State University of New York Press.
- Gordon, E. E. (1978). *A factor analytic description of tonal and rhythm patterns and objective evidence of pattern difficulty level and growth rate*. Chicago: G.I.A..
- Gordon, E. E. (1979). Developmental music aptitude as measured by the primary measures of music audiation. *Psychology of Music*, 7, (1), 42-49.
- Gordon, E. E. (1980a). Developmental music aptitudes among inner-city primary children. *Bulletin of the Council for Research in Music Education*, 63, 25-30.
- Gordon, E. E. (1980b). The assessment of music aptitudes of very young children. *The Child Quarterly*, 24, 107-111.
- Gordon, E. E. (1981). *The manifestation of developmental music aptitude in the audiation of "same" and "different" as sound in music*. Chicago: G.I. A..
- Gordon, E. E. (1986a). *The nature, description, measurement and evaluation of music aptitudes*. Chicago: G. I. A..
- Gordon, E. E. (1986b). A factor analysis of the MAP, the PMMA and the IMMA. *Bulletin of the Council for Research in Music Education*, 87, 17-25.
- Gordon, E. E. (1986c). Final results of a two-year longitudinal predictive validity study of the Instrument Timbre Preference Test and the Musical Aptitude. *Bulletin of the Council for Research in Music Education*, 89, 7-17.
- Gordon, E. E. (1986d). *Manual for the PMMA and the IMMA - music aptitude tests for kindergarten and first, second, third and fourth grade children*. Chicago: G.I.A..
- Gordon, E. E. e Woods, D. (1986e). *Reference handbook for using learning sequence activities*

- jump right in: the music curriculum*. Chicago: G. I. A..
- Gordon, E. E. (1987). *The nature, description, measurement and evaluation of music aptitudes*. Chicago: G. I. A..
- Gordon, E. E. (1989). *Predictive validity studies of the Intermediate Measures of Music Audiation and the Instrument Timbre Preference Test*. Chicago: G.I.A..
- Gordon, E. E. (1990). The nature and description of developmental and stabilized music aptitudes: implications for music learning. In R. F. Wilson e L. F. Roehmann (Eds.), *Music and child development*. Missouri: Biology of Making, Inc..
- Gordon, E. E. (1998). *Introduction to research and the psychology of music: measurement and evaluation*. Chicago: G. I. A.
- Gordon, E. (2000). *Teoria da aprendizagem musical - Competências, Conteúdos e Padrões*. Lisboa: Fundação Calouste Gulbenkian.
- Gordon, E. (2000). *Teoria da aprendizagem musical para recém-nascidos e crianças em idade pré-escolar*. Lisboa: Fundação Calouste Gulbenkian.
- Gordon, E. (2001). *Manual do Intermediate Measures of Music Audiation*. Lisboa: CEGOC (Tradução e adaptação portuguesa de Helena Rodrigues).
- Gregory, R. (1996). *Psychological testing*. Boston: Allyn e Bacon.
- Karma, K. (1982). Validating tests of musical aptitude. *Psychology of Music*, 10, (1), 32-35.
- Karma, K. (1984). Musical aptitude as the ability to structure acoustic material. *International Journal of Music Education*, 3, 27-30.
- Karma, K. (1985a). Components of auditive structuring - towards a theory of musical aptitude. *Bulletin of the Council for Research in Music Education*, 82, 1-13.
- Karma, K. (1985b). Selecting students to music instruction. *Bulletin of the Council for Research in Music Education*, 75, 23-32.
- Karma, K. (1986). Item difficulty values in measuring components of musical aptitude. *Bulletin of the Council for Research in Music Education*, 89, 19-31.
- Kingsbury, H. (1988). *Music, talent and performance - a conservatory cultural system*. Philadelphia: Temple University Press.
- Radocy, R. E. e Boyle, J. D. (1987). *Measurement and evaluation of musical experiences*. New York: Schirmer Books.
- Rodrigues, H. (1993). *Aptos, preparados ou esclarecidos - Contribuição para o estudo da aptidão musical*. Tese de Mestrado apresentada à Faculdade de Psicologia e Ciências da Educação da Universidade de Coimbra.
- Rodrigues, H. e Ferreira, J. (1994). Avaliação da aptidão musical: elementos para uma definição. *Psychologica*, 11, 91-100.
- Rodrigues, H. (1996). Thank you Dr Gordon, *Boletim da Associação Portuguesa de Educação Musical*, 88, 8-14.
- Rodrigues, H. (1998a). *Avaliação da Aptidão Musical em Crianças do 1º Ciclo de Escolaridade - Aferição do Teste Intermediate Measures of Music Audiation (I.M.M.A.) para a área educativa de Lisboa*. Tese de Doutoramento apresentada à Faculdade de Psicologia e Ciências da Educação da Universidade de Coimbra.
- Rodrigues, H. (1998b). Pequena crónica sobre notas de rodapé na Educação Musical - reflexões

a propósito da teoria da aprendizagem musical, *Boletim da Associação Portuguesa de Educação Musical*, 99, 15-25.

- Rodrigues, H. (1999). Medidas Intermédias de Avaliação da Aptidão Musical. In Simões, M.; Gonçalves, M. e Almeida, L. (Eds), *Testes e provas psicológicas em Portugal*. Braga: Sistemas Humanos e Organizacionais, Lda.
- Seashore, C. (1919). *The psychology of musical talent*. New York: Silver Burdett.
- Seashore, C. (1938). *The psychology of music*. New York: Silver Burdett.
- Shuter-Dyson, R. (1994). Le problème des interactions entre hérité et milieu dans la formation des aptitudes musicales. In A. Zenatti (Ed.), *Psychologie de la Musique*. Paris: Presses Universitaires de France.
- Shuter-Dyson, R. e Gabriel, C. (1981). *The psychology of musical ability*. London: Methuen.
- Schleuter, S.L. e Schleuter, L.J. (1989). The relationship of rhythm response tasks and PMMA scores with music training, grade level and sex among K-3 students. *Bulletin of the Council for Research in Music Education*, 1, 2-14.

Abstract

This article focuses on aptitude validity, reliability, impact and predictability of measurement in art and design in England and Portugal. The theme is discussed in relation to theoretical framework, applied measurement instruments, procedures of data collection and analysis. The Portuguese situation is highlighted, the importance of music, theatre and other products, but also limitations and should analysis include other work methods. Two different systems of aptitude assessment are analysed, in order to address similarities in England and Portugal at pre-university level. The text concludes with a study.

Introdução e quadro conceptual para análise de sistemas de avaliação

Este artigo descreve as conclusões de três anos de trabalho em Inglaterra e Portugal (2000, 2003) onde se conduziu uma análise sistemática dos instrumentos e processos de avaliação externa em vigor a nível do ensino pré-universitário: 12º ano em Portugal (idades 18 anos) e A/AS levels em Inglaterra (idades 17, 18 anos). A análise foi feita a partir de um quadro conceptual elaborado a partir de uma revisão de literatura centrada nos temas: avaliação; ensino artístico e criatividade. O quadro de referência parte dos conceitos de validade, fiabilidade e

Este trabalho foi desenvolvido na Universidade de Surrey, Inglaterra.