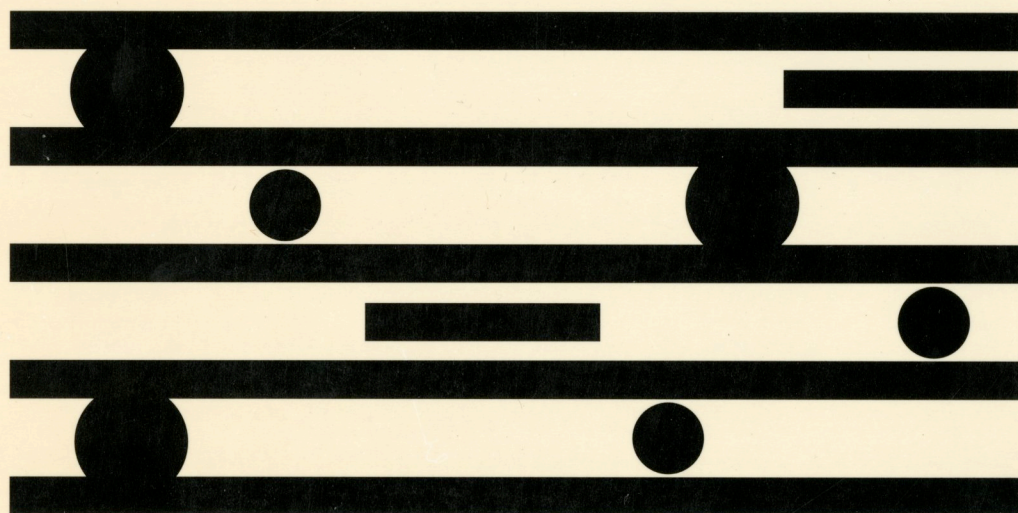


# INVESTIGAÇÃO EM PSICOLOGIA DA MÚSICA

Estudos Críticos

Coordenação  
Helena Rodrigues  
Christopher Johnson



Edições Colibri

---

Centro de Estudos de Sociologia e Estética Musical  
Universidade Nova de Lisboa









**INVESTIGAÇÃO EM PSICOLOGIA DA MÚSICA**

– ESTUDOS CRÍTICOS –



Colecção: **ESTUDOS MUSICOLÓGICOS**

Coordenação: Centro de Estudos de Sociologia e Estética Musical  
(C.E.S.E.M.) da Faculdade de Ciências Sociais  
e Humanas – Universidade Nova de Lisboa

Direcção: Prof. Doutor Manuel Pedro Ferreira

1. *Música, Estética e Sociedade nos Escritos de Jorge Peixinho*  
Cristina dos Anjos Raminhos Delgado Teixeira
2. *Constança Capdeville – Entre o Teatro e a Música*  
Maria João Serrão
3. *Interpretação Musical: Teoria e Praxis*  
Francisco Monteiro e Ângelo Martingo (coordenação)
4. *Investigação em Psicologia da Música – Estudos Críticos*  
Helena Rodrigues e Christopher Johnson (coordenação)

A publicar

- *Luigi Nono e a Nova Música: Inconformismo e Interpelação*  
Mário Vieira de Carvalho

Esta colecção é apoiada pela Fundação para a Ciência e a Tecnologia  
através do financiamento plurianual de unidades de investigação

# **INVESTIGAÇÃO EM PSICOLOGIA DA MÚSICA**

**– ESTUDOS CRÍTICOS –**

Coordenação:

Helena Rodrigues  
Christopher Johnson

Edições Colibri

Centro de Estudos de Sociologia e Estética Musical – UNL

*Biblioteca Nacional – Catalogação na Publicação*

Investigação em psicologia da música : estudos críticos / coord. Helena Rodrigues, Christopher Johnson. – (Ensaios musicológicos ; 4)

ISBN 978-972-772-764-3

I – RODRIGUES, Helena, 1962-

II – JOHNSON, Christopher, 1962-

CDU 78.01

159.922

Título: Investigação em Psicologia da Música  
– Estudos Críticos –

Editores: Helena Rodrigues e Christopher Johnson

Autores: Christopher Johnson, Debra Hedden, Helena Rodrigues,  
João Nogueira, Orlindo Gouveia Pereira, Wendy Sims

Edição: Edições Colibri / Centro de Estudos  
de Sociologia e Estética Musical

Tradução: Isabel Monteiro e Nuno Arrais

Revisão científica: Helena Rodrigues e Ângelo Martingo

Edição e revisão técnica: Tiago Paulos Veiga

Depósito legal: 266 998/07

Lisboa, 7 de Dezembro de 2007

# ÍNDICE

|                     |   |
|---------------------|---|
| AGRADECIMENTOS..... | 7 |
|---------------------|---|

## 1ª PARTE

|                                                                                                                                                     |   |
|-----------------------------------------------------------------------------------------------------------------------------------------------------|---|
| Reflexões para uma formação crítica e criativa no âmbito da Investigação em Psicologia da Música – Por Helena Rodrigues e Christopher Johnson ..... | 9 |
|-----------------------------------------------------------------------------------------------------------------------------------------------------|---|

## 2ª PARTE – ESTUDOS CRÍTICOS

|                                                                                                                                                                              |            |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| <b>2.1. The influence of training on the vocal ability of three year old children – De A. Jersild &amp; S. Bienstock .....</b>                                               | <b>13</b>  |
| 2.1.1. Resumo do artigo .....                                                                                                                                                | 13         |
| 2.1.2. Artigo original .....                                                                                                                                                 | 14         |
| 2.1.3. Estudo Crítico – Por Helena Rodrigues .....                                                                                                                           | 32         |
| <b>2.2. Music research in medical/ dental treatment: meta-analysis and clinical applications – De J. Standley .....</b>                                                      | <b>41</b>  |
| 2.2.1. Resumo do artigo / Abstract .....                                                                                                                                     | 41         |
| 2.2.2. Artigo original .....                                                                                                                                                 | 42         |
| 2.2.3. Estudo Crítico – Por Orlindo Gouveia Pereira .....                                                                                                                    | 67         |
| <b>2.3. Investigation of preschool children's comparative capability to sing songs with and without words – De L. Levinowitz .....</b>                                       | <b>71</b>  |
| 2.3.1. Resumo do artigo .....                                                                                                                                                | 71         |
| 2.3.2. Artigo original .....                                                                                                                                                 | 72         |
| 2.3.3. Estudo Crítico – Por Helena Rodrigues .....                                                                                                                           | 78         |
| <b>2.4. Effects of instruction and task format on preschool children's music concept discrimination – De W. L. Sims .....</b>                                                | <b>83</b>  |
| 2.4.1. Resumo do artigo / Abstract .....                                                                                                                                     | 83         |
| 2.4.2. Artigo original .....                                                                                                                                                 | 85         |
| 2.4.3. Estudo Crítico – Por Christopher Johnson .....                                                                                                                        | 97         |
| <b>2.5. The effect of music listening on physiological responses of premature infants in the NICU – De W. Cassidy &amp; J.M. Standley .....</b>                              | <b>103</b> |
| 2.5.1. Resumo do artigo / Abstract .....                                                                                                                                     | 103        |
| 2.5.2. Artigo original .....                                                                                                                                                 | 105        |
| 2.5.3. Estudo Crítico – Por Christopher Johnson .....                                                                                                                        | 119        |
| <b>2.6. The effectiveness of individual / small-group singing activities on kindergartners use of singing voice and developmental music aptitude – De J. Rutkowski .....</b> | <b>123</b> |

|                                                                                                                                                  |            |
|--------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| 2.6.1. Resumo do artigo / Abstract .....                                                                                                         | 123        |
| 2.6.2. Artigo original .....                                                                                                                     | 124        |
| 2.6.3. Estudo Crítico – Por Wendy Sims .....                                                                                                     | 136        |
| <b>2.7. The match between music excerpts and written descriptions by fifth and sixth graders – De P. J. Flowers .....</b>                        | <b>141</b> |
| 2.7.1. Resumo do artigo / Abstract .....                                                                                                         | 141        |
| 2.7.2. Artigo original .....                                                                                                                     | 142        |
| 2.7.3. Estudo Crítico – Por Debra Hedden .....                                                                                                   | 154        |
| <b>2.8. Faculty interjudge reliability of music performance evaluation – De M. J. Bergee .....</b>                                               | <b>161</b> |
| 2.8.1. Resumo do artigo / Abstract .....                                                                                                         | 161        |
| 2.8.2. Artigo original .....                                                                                                                     | 163        |
| 2.8.3. Estudo Crítico – Por Christopher Johnson .....                                                                                            | 173        |
| <b>2.9. The effect of vision and hearing loss on listeners perception of referential meaning in music – De A.-A. Darrow &amp; J. Novak .....</b> | <b>179</b> |
| 2.9.1. Resumo do artigo / Abstract .....                                                                                                         | 179        |
| 2.9.2. Artigo original .....                                                                                                                     | 181        |
| 2.9.3. Estudo Crítico – Por Christopher Johnson .....                                                                                            | 191        |
| <b>2.10. An investigation into the experience of musical performance anxiety in elementary school children – De C. Ryan .....</b>                | <b>195</b> |
| 2.10.1. Resumo do artigo / Abstract .....                                                                                                        | 195        |
| 2.10.2. Artigo original .....                                                                                                                    | 196        |
| 2.10.3. Estudo Crítico – Por João Nogueira .....                                                                                                 | 206        |

### 3ª PARTE

|                                                                                                                                           |            |
|-------------------------------------------------------------------------------------------------------------------------------------------|------------|
| <b>Fontes de pesquisa e informação bibliográfica para a realização de estudos de investigação no âmbito da Psicologia da Música .....</b> | <b>211</b> |
| 3.1. Fontes Fundamentais .....                                                                                                            | 212        |
| 3.2. Fontes Complementares .....                                                                                                          | 215        |
| 3.3. Bases de Dados .....                                                                                                                 | 218        |
| 3.4. Outros .....                                                                                                                         | 222        |
| <b>AUTORES DOS ESTUDOS CRÍTICOS .....</b>                                                                                                 | <b>225</b> |
| <b>COLABORADORES .....</b>                                                                                                                | <b>229</b> |
| <b>GLOSSÁRIO .....</b>                                                                                                                    | <b>231</b> |
| <b>BIBLIOGRAFIA .....</b>                                                                                                                 | <b>241</b> |



# AGRADECIMENTOS

A publicação deste livro só foi possível graças à colaboração de várias entidades e de vários colegas a quem queremos expressar os mais vivos agradecimentos.

Assim, agradece-se às revistas *Journal of Research in Music Education*, *Bulletin of the Council of Research in Music Education*, *Journal of Music Therapy*, *Child Development* e à organização do 16º *International Symposium on Research in Music Behavior* a autorização para a reprodução dos artigos que aqui se integram.

Aos Professores Leandro Almeida do Instituto de Psicologia e Ciências da Educação da Universidade do Minho, Margarida Pedroso Lima da Faculdade de Psicologia e Ciências da Educação da Universidade de Coimbra e a Manuel Lisboa da Faculdade de Ciências Sociais e Humanas agradece-se a disponibilidade com que sempre responderam a dúvidas que foram surgindo ao longo da preparação deste livro. São também devidos agradecimentos aos Professores Mário Vieira de Carvalho, Manuel Pedro Ferreira e Gabriela Cruz pela criação de boas condições de trabalho no âmbito do CESEM.

Aos colegas Debra Hedden, João Nogueira, Orlindo Gouveia Pereira e Wendy Sims agradece-se a generosidade que dedicaram aos estudos críticos efectuados. A Isabel Monteiro e Nuno Arrais o cuidado com que traduziram os estudos críticos efectuados pelos colegas americanos. A Ângelo Martingo o empenho com que colaborou na revisão científica. E a Tiago Paulos Veiga toda a paciência e sensibilidade com que ajudou este livro a nascer.

É devido ainda um agradecimento à Direcção da FCSH-UNL e ao Departamento de Ciências Musicais que permitiram a Helena Rodrigues o gozo de licença sabática no ano lectivo de 2005/06, período durante o qual se trabalhou nesta publicação.

Finalmente, são devidos agradecimentos muito especiais à Comissão Fulbright que possibilitou a colaboração de Christopher Johnson na leccionação da Área de Psicologia e Pedagogia Musical do Mestrado em Ciências Musicais, viabilizando um entendimento científico e humano muito produtivo entre professores e mestrandos, do qual este livro é apenas um dos resultados visíveis.



## 1ª PARTE

# REFLEXÕES PARA UMA FORMAÇÃO CRÍTICA E CRIATIVA NO ÂMBITO DA INVESTIGAÇÃO EM PSICOLOGIA DA MÚSICA

HELENA RODRIGUES E CHRISTOPHER JOHNSON

*“Penso que só há um caminho para a ciência ou para a filosofia: encontrar um problema, ver a sua beleza e apaixonar-se por ele; casar e viver feliz até que a morte vos separe – a não ser que encontrem um outro problema ainda mais fascinante, ou, evidentemente, a não ser que obtenham uma solução. Mas, mesmo que obtenham uma solução, poderão então descobrir, para vosso deleite, a existência de toda uma família de problemas-filhos, encantadores ainda que talvez difíceis, para cujo bem-estar poderão trabalhar, com um sentido, até ao fim dos vossos dias.”*

**Popper**  
[1989, p.3]

A Psicologia da Música é um campo de trabalho que tem vindo a ganhar terreno e suscitado o interesse crescente não só de músicos como também de todos aqueles que se interessam por aprofundar o comportamento humano em todas as suas dimensões.

O interesse pela "mente" musical é algo que podemos fazer remontar na história da humanidade tanto quanto nos leva o conhecimento sobre essa mesma história como também tanto quanto quisermos ser flexíveis na definição da disciplina. Há historiadores que recuam até Pitágoras e vão sucessivamente encontrando os primeiros enunciados de uma disciplina que no final do século XIX recebe um impulso decisivo com os trabalhos de Helmholtz dedicados ao estudo dos fenómenos perceptivos associados à audição. Os estudos efectuados por Seashore no âmbito da psicoacústica e da avaliação do talento musical e, entre outras, as suas obras *The Psychology of Musical Talent* e *The Psychology of Music*, publicadas no início do século XX, são também marcos históricos na emergência de uma nova disciplina.

Ou seja, os primeiros passos da Psicologia da Música dão-se em simultâneo com os primeiros passos da Psicologia. Tal não é de surpreender, pois o interesse geral pelo humano não poderia deixar de ser acompanhado pelo interesse por uma das componentes específicas da produção humana – a Música – que inclusivamente diferencia o ser humano de outras espécies.

Não é objectivo deste texto fazer uma História da Psicologia da Música, das suas variadas filiações e ramificações, nem tão pouco desenvolver outras reflexões de carácter filosófico e epistemológico sobre a aplicação do modelo científico aos fenómenos humanos e artísticos.

Pretendeu-se apenas referir, de modo pragmático, que, volvidos mais de cem anos, hoje é inquestionável o volume de trabalhos produzidos neste âmbito disciplinar. É cada vez maior o número de profissionais com diferentes tipos de formação – entre os quais teremos que destacar indivíduos ligados à Psicologia, à Educação, ao Ensino da Música, à Interpretação e Composição Musical e à Musicoterapia – que buscam na Psicologia da Música matéria para o seu enriquecimento teórico e para o aperfeiçoamento das suas práticas no seu quotidiano profissional.

Em Portugal, assiste-se também a este elevado interesse pela disciplina, a par, por exemplo, de um vertiginoso aumento de práticas educativas e terapêuticas no âmbito do desenvolvimento musical na primeira infância – a que, sem dúvida, deveria ser possível dar uma melhor atenção.



Na realidade, são inúmeros os dados objectivos que podem comprovar o elevado interesse de estudantes e profissionais de várias áreas profissionais quer pela Psicologia da Música em geral como pela intervenção em contextos de relação em que a Música, para além de prática artística em sentido estrito é também usada como *medium* relacional e comunicacional. Por exemplo, na intervenção com bebés e crianças em idade pré-escolar, com crianças com necessidades educativas especiais, com bebés prematuros, em contextos de apoio à terceira idade, em prisões, em ambientes socialmente desfavorecidos, em meio hospitalar, etc.

A par de outros contributos igualmente importantes no País, o carisma de Edwin Gordon e a divulgação dos seus trabalhos – recorde-se que a sua passagem pelo nosso País foi acompanhada pela publicação de duas das suas obras pelo Serviço de Educação da Fundação Calouste Gulbenkian – e, de uma forma especial, os Seminários dedicados a aspectos relativos à Investigação em Psicologia da Música que desde 1995 foi ministrando ora na Fundação Calouste Gulbenkian ora na Faculdade de Ciências Sociais e Humanas, têm ajudado a incrementar o entusiasmo por esta disciplina.

Mas, como em todos os processos de crescimento, é necessário diversificar e ampliar as referências informativas que podem alimentar o espírito crítico e criativo de todos quantos pretendem dedicar-se à investigação científica. Por outro lado, é necessário acompanhar as novas aprendizagens com meios prudentes capazes de canalizarem a energia e a motivação de forma adequada e amadurecida.

Numa actualidade caracterizada por uma enorme disseminação de informação nem sempre acompanhada pelos níveis de qualidade e rigor desejáveis, esta prudência deve traduzir-se no desenvolvimento de um elevado sentido crítico que tem de ser exercido continuamente de forma autónoma e inconformista, senão mesmo irreverente.

Não será descabido defender que actualmente os processos de economia de mercado afectam a circulação da produção científica de forma análoga à que é verificada noutras transacções. É, pois, necessário estar atento e saber seleccionar a informação tão facilmente disponível.

Esta é a primeira grande finalidade do livro que agora se publica: fomentar o exercício do sentido crítico junto daqueles que se interessam pela investigação nas Ciências Humanas e em particular em áreas afectas à Psicologia e à Música.

Evidentemente, só é possível exercitar o sentido crítico se previamente existir informação. No caso, referimo-nos concretamente a conhecimentos no âmbito da Metodologia da Investigação. Esta é uma obra que requer um conhecimento prévio dos métodos e técnicas de investigação usados nas Ciências do Comportamento. E neste campo há, sem dúvida, publicações de elevada qualidade editadas no nosso país.

No entanto, o conhecimento em abstracto dos métodos e técnicas de investigação não garante, por si só, a sua aplicação quer na leitura de artigos como na produção de trabalhos de investigação. De facto, a teorização da Metodologia da Investigação mostra-nos, por vezes, uma realidade idealizada, com conceitos e ideias que não são directa e facilmente transferíveis e aplicáveis quando se analisam trabalhos publicados. Muitas vezes, a leitura de artigos de investigação suscita no leitor dúvidas como, por exemplo, qual é a variável dependente ou independente na situação em análise, de que tipo de processo de amostragem se trata, a que tipo de procedimento se recorreu ou em que consiste um determinado processo de análise de dados.

O facto de grande parte da informação disponível ser de origem anglo-saxónica traz também problemas de linguagem. Mas, para lá de equívocos resultantes de questões de comunicação, há também questões decorrentes das dificuldades de compreensão de um vocabulário técnico, por vezes muito hermético. Algo que ultrapassa as simples e lineares questões de tradução.

No âmbito da Psicologia da Música há, sem dúvida, um trabalho a fazer no que toca à adopção e descodificação de vocabulário técnico utilizado.

Por vezes, assiste-se a uma prolixidade de vocabulário, inserida numa lógica de poder destinada a territorializar espaços conceptuais de trabalho, mas que no essencial serve apenas para replicar e duplicar a essência de conceitos pré-existentes com outras denominações supostamente destinadas a evidenciar novas formas de pensamento. Outras vezes, é certo, novas terminologias revelam opções epistemológicas de fundo. Certo é que, entre conceitos realmente inovadores que obrigam à criação de novas terminologias e entre algum *déjà vu* travestido de novo, confrontamo-nos, por vezes, nas Ciências Sociais e Humanas, com um palavroso labirinto



de vocabulário técnico que confunde, desorienta e oculta. Por vezes, o vocabulário é uma muralha que protege a fragilidade de alguns detentores de palavras, afastando outros que buscam algo de mais intrínseco e essencial no saber.

É, portanto, objectivo deste livro ajudar o leitor a aplicar conhecimentos da Metodologia da Investigação a situações concretas de investigação, guiando-o na compreensão de vocabulário técnico. E este objectivo leva-nos à explicação da organização de *Investigação em Psicologia da Música – Estudos críticos*.

A obra reúne dez artigos de investigação, reproduzidos com a gentil autorização de revistas prestigiadas como é o caso do *Journal of Research in Music Education*, *Journal of Music Therapy* e *Child Development* e de dois autores que apresentaram comunicações num Simpósio Internacional de Investigação. São artigos que datam de 1931 a 2005, que contemplam diversas abordagens metodológicas e cobrem temas diversificados. Acreditou-se ser importante adoptar este panorama abrangente e alargado quer pela necessidade de criar bases de conhecimento geral, necessário como base prévia para qualquer especialização posterior, como pela necessidade de abrir perspectivas no domínio. O facto de haver artigos que abordam a Música em diferentes tipos de população, desde a constituída por intérpretes à constituída por bebés prematuros, passando por crianças com problemas de visão e audição, é uma metáfora daquilo que acreditamos ser um caminho para o Futuro da formação de profissionais da Música que trabalhem em diferentes tipos de contextos: educação e terapia devem estar mais próximas na promoção saudável do desenvolvimento humano.

Cada um dos artigos é objecto de uma análise crítica efectuada por diferentes colaboradores e abordando parâmetros que visam simultaneamente a dissecação do artigo em termos objectivos e o exercício crítico do mesmo. Assim, embora deixando margem para alguma flexibilidade na exposição, a todos os colaboradores foi solicitada uma análise que contemplasse os seguintes parâmetros:

- a) Identificação da Finalidade e Problemas do estudo, relacionando-os respectivamente com as Conclusões e os Resultados;
- b) Categorização do estudo em termos de *design*;
- c) Análise crítica relativa à revisão bibliográfica, à metodologia (considerando sujeitos, instrumentos e procedimento), aos resultados e à discussão dos mesmos;
- d) Pertinência e contributos do artigo para o domínio de estudo, considerando possibilidades e limitações ;
- e) Comentário geral ou outros comentários.

Os artigos de investigação são publicados na língua original, sendo cada um precedido por um resumo traduzido do mesmo. A leitura de artigos de investigação em Psicologia da Música escritos em inglês é situação frequente com que os estudantes e profissionais se defrontam no seu quotidiano. Mas, justamente por causa das questões de tradução de vocabulário técnico a que atrás se fez referência, as análises críticas mereceram um cuidado especial tendo sido traduzidas e revistas por vários dos colaboradores que participaram na edição desta obra. Houve, efectivamente, um cuidado especial em tornar compreensível uma linguagem que por vezes se apresenta hermética.

Evidentemente, em caso algum se dispensa a leitura dos artigos originais.

As análises críticas que aqui se apresentam devem ser entendidas como guias de compreensão para cada artigo de investigação. No entanto, numa segunda fase devem suscitar outras análises críticas por parte do leitor. Convidamos, pois, o leitor a ler os artigos originais e a reflectir sobre eles usando os parâmetros acima descritos. O leitor poderá então confrontar a sua própria análise crítica com a que aqui é oferecida.

Desejável será que as próprias análises críticas que aqui se oferecem possam ser também alvo de crítica por parte do leitor exercendo-se assim a metacrítica.

Em suma, é nossa convicção que é necessário desenvolver um trabalho de profundidade, de atenção pelos detalhes, de selecção e reflexão crítica opondo-nos à atitude de superficialidade



voraz a que a massificação de publicação e a competitividade e pressão para a produção inevitavelmente conduzem.

O exercício do sentido crítico é a essência da criatividade. Vale a pena fazer novas propostas quando não nos revemos total e completamente nas que já existem. Mas ninguém cria a partir do nada. De facto, a criatividade é, muitas vezes, desarrumação, reorganização, focalização diferenciada noutras peças de um *puzzle* de informação previamente recebida.

Mas o exercício crítico, se esgotado em si próprio, pode tornar-se estéril. Encontra o seu sentido quando informa a síntese pessoal, a vontade própria de criar e inovar.

Esta é, pois, a outra grande finalidade do livro que agora se publica: inspirar aqueles que pretendem criar os seus próprios trabalhos de investigação, nomeadamente quando se propõem elaborar dissertações no âmbito da Psicologia e da Educação, do Ensino da Música, da Musicoterapia, de uma Didáctica das Artes ou da Interpretação Musical.

A par da dissecação pormenorizada de cada um dos artigos, do conhecimento das técnicas que lhes estão subjacentes (tarefa esta que é, evidentemente, inesgotável) e de uma visão abrangente sobre diferentes temas e possibilidades de abordagem, o leitor é convidado a libertar o seu pensamento num *brainstorming* criativo que lhe permita formular os seus próprios problemas de investigação, contrapor as suas próprias perspectivas, encontrar os seus interesses e as suas finalidades no âmbito da investigação que pretenda desenvolver. A sua voz interior.

Só este encontro apaixonado de que nos fala Popper permite que a elaboração de teses e dissertações seja mais do que o cumprimento de um dever para se adquirir um grau académico ou uma promoção profissional. A elaboração de uma dissertação deve ser motivo de júbilo e o início de uma paixão que pode, de diferentes maneiras, acompanhar a vida profissional de todos quantos lidam com a Música, a Educação e a Psicologia.

Esperamos, pois, com esta obra contribuir para o florescimento de uma paixão a cujo nascimento temos assistido em muitos dos nossos alunos.

A eles dedicamos este livro.

## 2ª PARTE

# ESTUDOS CRÍTICOS

2.1.

THE INFLUENCE OF TRAINING ON THE VOCAL  
ABILITY OF THREE YEAR OLD CHILDREN

EFEITOS DA INSTRUÇÃO MUSICAL NAS CAPACIDADES VOCAIS  
DE CRIANÇAS DE TRÊS ANOS DE IDADE

A. JERSILD & S. BIENSTOCK (1931)

CHILD DEVELOPMENT, VOL. II, Nº 4, DECEMBER, 272 - 291

### 2.1.1. Resumo do artigo

O estudo pretende avaliar o desempenho vocal de crianças de três anos de idade e a possibilidade de melhorar esse desempenho através da aprendizagem. Foram administrados um conjunto de testes avaliando capacidades de imitação de sons de altura diferente e de intervalos. Depois, ao longo de seis meses, um grupo de crianças foi sujeito a instrução musical. No final do estudo, os resultados destas crianças foram comparados com os de crianças não expostas a aprendizagem mas também elas sujeitas à administração daqueles testes. Paralelamente, foram efectuadas observações das vocalizações espontâneas das crianças. Verificou-se que as crianças que receberam instrução obtiveram classificações mais elevadas nos testes finais do que as do grupo de controlo e obtiveram-se dados relevantes para a caracterização do desempenho vocal de crianças desta faixa etária.

## 2.1.2. Artigo original

# The Influence of Training on the Vocal Ability of Three-Year-Old Children

ARTHUR T. JERSILD AND SYLVIA F. BIENSTOCK (1931)

The present investigation was designed to measure the singing ability of three year old children and to study the effects of training on the child's vocal performance.\*

Initial measurements were made of the child's ability to reproduce pitch, by the singing of separate tones, and his ability to reproduce intervals, by singing of two successive tones of different pitch. Training in singing was given to a group of children during a period of several months, while an equivalent control group was tested at the beginning and at the end of the investigation. A study was also made of the children's spontaneous vocalizations.

The study offers findings with regard to such factors as the notes and intervals which can most readily be sung by three year old children, individual differences in performance, and the influence of training on vocal performance. The data furthermore provide findings with regard to the relationship between children's vocal abilities and such factors as ability to learn the melodies and words of songs, general intelligence, parents' vocal range, and the range and content of spontaneous singing.

Forty-eight children, ranging in age from 31 months to 48 months, were used in the study. Eighteen of these were given training twice weekly during a period of six months. The 18 children completed a total of 40 practice periods required for each child.

The children who were not included in the training series were tested early in the study, and as many as could be found were retested at the end of the experiment. From their number 18 were selected to serve as a control group. Each child in this group was paired, as closely as possible, with respect to age and score on the initial tests, with a child in the group that completed the training series. The control group was used to find the degree to which the improvement exhibited by the trained group could be attributed to such factors as growth, increased familiarity with the procedure, more favorable rapport with the experimenter, and other incidental conditions rather than to actual increase in vocal ability through training.

References to published and unpublished experimental work in music can be found in the Twenty-eighth Year Book of the National Society for the Study of Education (6). References to the valuable studies conducted at the University of Iowa prior to and including the year 1928 are given by Seashore (4). The authors are not aware of published experimental results dealing with the main problems of the present study.

---

\* This study was conducted at the Child Development Institute, Teachers College, Columbia University. The authors are grateful to Prof. Lois H. Meek, director of the Institute, for aid in providing the facilities for the study, and to Prof. Esther Lloyd-Jones, Miss Ethel Gordon, and Mrs. Edna G. Krull for generous help in checking up the reliability of the experimental procedure. The authors are indebted to Miss E. Webster, director, and Mrs. A. Connors, teacher, of the Manhattanville Day Nursery, to Miss Katherine Burton, director of the Washington Heights Day Nursery, to Miss Christine Heinig, head-teacher of the nursery school of the Child Development Institute, and to Miss Helen Rogers and Miss Ruth Swift of the Greenwich House nursery school for help in procuring the children used in the study. They are grateful also to Dr. Ruth Arrington for providing supplementary records of the laughter and overt activity of children used in the study.



## PROCEDURE

### *Test of pitch and interval*

The first step in the study was to find the number of notes and intervals which each child could sing. Two tests, which hereafter be designated *pitch* and *interval*, were designated for this purpose.

The initial pitch test included the following 11 notes, beginning with Middle C:

C      D      E      F      G      A      B      C      D      E      F

The initial interval test included the following 12 intervals: Ascending and descending major and minor seconds and thirds, perfect fourths and perfect fifths.

When the pitch and interval tests were given, the children were taken in groups of two or three to a separate room and were seated at a table, facing the experimenter. They were encouraged to play on a small xylophone and permitted to inspect a picture book. The experimenter then proceeded with the test. A given note was sung by the experimenter and sounded on the xylophone or on a pitch-pipe. Two chromatic pitchpipes, one in the key of F, the other in the key of C, were used. The children were asked to listen and to sing the same tone. Various means of encouragement, including the use of the xylophone and the picture book, were employed. The children soon learned what was expected of them, and as time went on less and less extraneous motivation was needed.

Each item of the test was presented separately, and the responses of each child to each item were separately recorded. After a given tone had been presented, the experimenter kept sounding and singing the tone to prompt the child's efforts. If the child succeeded in reproducing the tone no prompt the child's efforts. If he tried to sing the tone and failed to reproduce it correctly or if he remained silent the experimenter continued to sing and sound the tone a maximum of eight times. Fewer than eight presentations were made if the child's attention could not be held. If the child reproduced the presented item he was credited with a 'hit'; if he tried but failed he was credited with an 'attempt'; if he continued to remain silent, his record on the particular item was entered as 'no attempt'. A record of the number of trials preceding success was also made. Each of the 11 notes of the pitch test was dealt with in this manner.

The pitch test was administered three times, on three different occasions, during the initial test series. When the record was scored, one credit was given to each note correctly reproduced. If the child sang a given note on only one test or on all three occasions when the test was administered he received one credit. The maximum score on the pitch test was, accordingly, 11.

The interval test was administered at two separate times after the completion of the pitch tests. The general procedure was the same as that described above, save that two notes constituting an interval were presented instead of separate notes as in the pitch tests. Each interval correctly reproduced during one or both administrations of the test was given one credit in the final score. The maximum score according to this procedure was 12.

### *Method of evaluating the child's responses*

Of first importance in a study of this kind is the need for a valid means of evaluating and scoring the responses made during the tests. Much time was spent in preliminary experimentation to find a procedure which could be used expeditiously in work with children at this early age and still yield reliable quantitative results. The use of a mechanical device, such as a tonoscope, for recording the children's responses was considered, but after preliminary work it was felt that the purpose of the study could be served without a device of this kind. The procedure of requiring the



experimenter to judge the correctness of the child's response was tried and appeared to give satisfactory results. Accordingly, during the course of the tests, the experimenter not only produced the notes or intervals which the child was required to sing, but also judged whether the child's responses correctly reproduced the tones which were being presented. In forming a judgment the experimenter made use of the musical instruments as a basis of comparison between the child's responses and the tone that was required. No response was recorded as correct unless in the opinion of the experimenter it was definitely in accord with the tone which was being sounded on the musical instrument. A description of methods designed to ascertain the validity of the experimenter's judgments follows.

### *Reliability of the judgments of the child's responses*

Various means were used in checking upon the reliability of the experimenter. First of all the experimenter was tested by another musician. The tester sang at random all the notes used on the initial tests given to the children, taking her note from a piano which was not within the hearing of the experimenter. She recorded each note thus sung. The experimenter in the meantime was required to identify (by aid of a pitch-pipe) and to record each note that was sung. After the test had been completed the tester's and the experimenter's records were compared, item by item. The experimenter made a perfect score on each of three repetitions of the test.

A more significant check on the reliability of the experimenter was obtained by introducing other musically trained examiners while tests were being given the children during the course of the study. The persons thus brought in took independent records of the children's performances while they were being tested by the experimenter.

On one occasion, 20 children were tested in this way, both for pitch and interval, by the experimenter and a second examiner. After the completion of the tests, the separate records were compared.

The pitch test on which simultaneous and independent records were made by the two examiners included the 11 notes of the initial pitch test. The total possible score for the 20 children was, accordingly, 220. An item by item comparison of the two independent records showed an agreement of over 96 per cent. (The judgments were in disagreement on only 7 notes in the total of 220.) The experimenter credited the children with 125 correct hits, one of which had been judged incorrect by the second examiner, while the second examiner credited them with 130 correct hits, six of which had been judged incorrect by the experimenter.

On the interval test, each child was tested by both examiners on the 12 intervals of the initial tests. The total possible score for the 20 children was 240. An item by item comparison of the records showed an agreement of 95 per cent. The two examiners agreed in judging 152 items as correctly reproduced and 75 items as incorrectly reproduced. The experimenter judged one response as correct which was judged incorrect by the second examiner, while the second examiner gave credit to twelve responses which were judged incorrect by the experimenter.

It can be observed that there was not only a high percentage of agreement between the two examiners, but also that the experimenter was more rigid in giving credit than was the second examiner.

The same method of checking upon the reliability of the records was used at a later stage of the experiment when still another musician was introduced to take simultaneous and independent records during the pitch tests. This musician, like the person previously introduced, possessed what is known as "absolute pitch." On this occasion, 10 children were tested on 11 notes, making a possible total score of 110. Once more an item by item comparison of the records was made at the end of the tests. The agreement in this instance was 96 per cent. Two items which had been credited as correct by the experimenter were recorded as failures by the second examiner, while on the other hand, this second examiner had given credit for 2 items which the experimenter had recorded as incorrect.

The above results indicate that the method of testing used in this study showed a high degree of reliability within the limits of the procedure employed.



The reliability of the tests, as measured by correlating the scores obtained by the experimenter on successive repetitions ranged from  $+.54 \pm .11$  to  $.95 \pm .02$ , with a median of  $.80 \pm .06$ .

It can be seen that the testing procedure corresponded quite closely to the type of test situation which an individual must face under practical circumstances, in school or even in performance of a professional character. The method, while showing high reliability when employed by different examiners did not, to be sure, give a measure of the exact vibration frequency of the tone produced by the child. The measurements were, however, sufficiently precise to measure not only individual differences in general range of musical performance, but also to test a given child's ability to reproduce the various steps of the musical scale.

#### *Extension of the tests*

As the study progressed, it appeared that the children were improving to such a degree that it was necessary to extend the range of notes and intervals used in the training series and in periodic retests. The tests were therefore augmented to include the original 11 notes used in the pitch test, plus the following notes: G, A, B, below Middle C, and G, A, B, and C, two octaves above Middle C, making a total of 18 notes. The interval test was augmented by the following intervals: ascending and descending major and minor sixths and sevenths, and octaves, making a total of 22 intervals.

In order to provide an instrument to cover these additions, a psaltery was added to the pitch-pipes and xylophone which had been employed from the beginning of the study. The psaltery was tuned daily.

#### *Methods of training*

Following the initial tests the experimental children were given further training in singing. Songs taken from music books for young children were used for this purpose. The children were brought to the music room, which adjoined the playground, two at a time. The experimenter would sing the songs and ask the children to join in the singing. As the study progressed special attention was given to difficult parts of the song.

The practice period was designed to last 10 minutes. Early in the study it was not found possible to maintain the interest of some of the children during the entire 10 minutes. Later in the study, when all the children were quite eager to partake in the music training, the period was sometimes extended to 15 minutes to compensate the earlier loss of training time.

At the beginning of the training series the children were permitted to play on the musical instruments, to improvise songs of their own or to try to sing the presented song more or less as they pleased. The experimenter again used various means to stimulate the child's interest in singing the songs that were used for training purposes; for example, when one of the songs, a lullaby, was presented, the child was encouraged to accompany the experimenter in going through the motions of putting a doll to sleep, while singing; another song, which ended with the words 'peek-a-boo,' likewise could readily be acted out to the accompaniment of singing. As time went on the children responded more readily without such incidental promptings.

Eight songs were used for training purposes. Three of these were used at the beginning of the training while the others were introduced one at a time as the study progressed. During training the policy was used to group a child who showed relatively much ability with one who was less competent. Furthermore, the children were rotated in such a way that each child would sing with a different child on different days. Each child received an average of two training periods a week. When tests were given the test took the place of the training period. If a child missed training periods by reason of absence from school, extra periods were allotted to him on his return. In addition to the initial and final tests, given in the months of December and May, respectively, the pitch and interval tests were introduced in January, February, and March. Since these tests are in themselves a form of training, the periods devoted to them are counted as a part of the child's



training series. Each child received forty training periods, including the periods devoted to the tests preceding the final test in May.

### *Spontaneous vocalizations*

A study of the children's spontaneous vocalizations constituted a separate project within the investigation.

The children were observed during the course of their free play at the nursery school during the morning hours. Each child was observed for a total period of 100 minutes, divided into at least ten intervals occurring at different times over a period of three to five months.

The investigator followed the child as closely and unobtrusively as possible. Whenever the child sang, hummed, chanted, or made outcries which could be heard clearly and which could be located on the musical scale a record of his response was made. Pitch-pipes were used to identify each response that was recorded. So far as possible, the verbal content of the child's responses, the activities accompanying his response, and the rhythmic pattern of his vocalization, if any could be observed, were also recorded.

This procedure was calculated to provide a general indication of the character of children's spontaneous singing, and of individual differences in the content of children's musical expressions during the course of their undirected activities. The results cannot be regarded as final measures of the absolute frequency of the children's singing, since notes sung in an undertone, or notes which could not be identified by reason of lack of clearness or distraction or other factors, were not entered upon the records.

## **RESULTS**

In the presentation of the results, the topics covered in the study will be treated in the following order: findings on the initial pitch and interval tests; the influence of practice on the child's performance; the extent and content of children's spontaneous vocalizations; correlation between various measurements.

The initial pitch tests were given to 48 children, and the initial interval tests to 47, including the 18 children subsequently used for training. The age range was 31 to 48 months, average 39 months. The average number of notes sung by 48 children on the three trials of the pitch test, containing the 11 notes from Middle C to F above C an octave higher was  $4.20 \pm 2.44$ . The average number of intervals sung by 47 children on the two trials of the interval test containing the 12 intervals (ascending and descending major and minor seconds and thirds, perfect fourths and fifths) was  $4.32 \pm 2.90$ . The scores for the experimental and control children used in the training study are given in a later table.

TABLE 1

*Number of children correctly reproducing the various notes and intervals on the initial tests, based on 3 repetitions of the pitch test, and 2 repetitions of the interval test, with allowance of only one credit for each note or interval sung during the test series*

| Pitch (48 cases) |    |    |    |    |    |    |    |    |   |   |
|------------------|----|----|----|----|----|----|----|----|---|---|
| MIDDLE C         | D  | E  | F  | G  | A  | B  | C  | D  | E | F |
| 19               | 28 | 30 | 27 | 29 | 22 | 13 | 12 | 11 | 6 | 5 |

| Interval (47 cases) |              |             |             |                |               |              |              |             |             |                |               |
|---------------------|--------------|-------------|-------------|----------------|---------------|--------------|--------------|-------------|-------------|----------------|---------------|
| ASCENDING           |              |             |             |                |               | DESCENDING   |              |             |             |                |               |
| Major second        | Minor second | Major third | Minor third | Perfect fourth | Perfect fifth | Major second | Minor second | Major third | Minor third | Perfect fourth | Perfect fifth |
| 24                  | 18           | 17          | 13          | 12             | 13            | 26           | 11           | 22          | 23          | 16             | 8             |

The results of the initial tests were examined to study the range of the children's voices, and to find where their voices were placed. Table 1 shows the number of children who were able to sing each of the notes and intervals covered in the tests.

It appears in table 1 that the notes from D to A inclusive were reproduced most readily by the children used in this study. It further appears that the closer intervals, such as the seconds and thirds, were reproduced more readily than the wider intervals. The descending intervals have a slight but not significant advantage over the corresponding ascending intervals.

The findings with regard to pitch differ somewhat from usual statements with regard to the voices of young children. The claim is usually made that songs for young children should fall within the compass extending from E first line to E fourth space. (See references 2, 3, 5.) The authors have been unable to find published experimental data to support this claim. Even less definite are many writings on methods of teaching singing to young children which state that songs should be within the compass of the child's voice, without giving definite indications of what this compass is.

The findings in the present study indicate that the voices of young children are placed lower than is commonly supposed. It appears that when actual tests were given the notes Middle C, and D above Middle C, (both of which are below the first line E), were sung considerably more readily than C and D an octave higher, and that very few children sang the E and F (an octave above Middle C) which have been regarded as well within the range of the child's voice.

It is true that the statements with regard to the range of children's voices usually pertain to children somewhat older than those used in this study. It is possible that the voices of children at the age of six favor a higher range than the voices of three year old children. But it may be said at this point, in anticipation of findings which will be presented later in greater detail, that it was found that following intensive training the children of this study scored higher on the notes below first line E than on the higher notes immediately beyond the fifth line F which has been regarded as well within the range of the voices of young children. On the final pitch tests given to the 18 children who received training the number of children who sang each of the 18 notes of the extended test was as follows:

| (Middle 1st line) |    |    |    |    |    |    |    |    |
|-------------------|----|----|----|----|----|----|----|----|
| G                 | A  | B  | C  | D  | E  | F  | G  | A  |
| 5                 | 15 | 17 | 18 | 18 | 18 | 18 | 18 | 18 |

| 5th line |    |    |    |    |    |    |    |   |
|----------|----|----|----|----|----|----|----|---|
| B        | C  | D  | E  | F  | G  | A  | B  | C |
| 18       | 17 | 18 | 17 | 15 | 15 | 14 | 13 | 9 |



In later tests of songs used for training purposes, the low notes (Middle C, D, Eb) again scored substantially higher than the high notes (D, E, F beyond C an octave above Middle C). It thus appears from the initial pitch tests, from final pitch tests given after training, and from the tests employing actual songs used in training, that three year old children, as represented in this investigation, can sing notes below the commonly accepted range quite as readily as some of the higher notes which have been regarded as well within the child's capacity.

This finding that children's voices are placed lower than is commonly supposed was unanticipated in the present study. At the beginning of the investigation, the experimenters planned their approach in accordance with the supposition that children's voices favored the higher notes. Early in the training period, however, it became evident that this procedure had to be revised. It was often found that by transposing the songs used for instruction a more ready response could be obtained. Children who had previously remained silent made efforts to sing after the songs had been transposed to a lower key. Only in the case of the spontaneous vocalizations, as will be indicated at a later point, did the children tend to favor a higher range of notes.

The results from the initial interval tests indicated that the closer intervals (i.e., the seconds and thirds) were reproduced more readily than the wider intervals (fourths and fifths). In the literature opinion is divided as to the relative ease with which children can sing narrow and wide intervals. The opinion is sometimes set forth, however, that the half-step or chromatic interval (minor second) should not be stressed in songs for the young child. Thorn (5) warns against the use of difficult half-steps; the New York Board of Education Music Syllabus (3) and Damrosch, Gartlan, and Gehrkins (1) do not formally introduce instruction in the use of chromatics in songs until the fifth year of school.

In table 1 it appears that the minor second was sung less frequently than the major second and the minor and major thirds, and slightly more frequently than the perfect fourths and fifths. The score for the minor second interval is relatively quite high, and does not tend to support the view that the chromatic interval should not be used in songs for young children. As will be shown in a later table dealing with the children's spontaneous vocalizations, the chromatic interval occurred more frequently than any other interval with the exception of the major second. On the final tests the minor second was sung by all of the 18 children who had received training.

### *The influence of training*

Table 2 shows a comparison between the experimental and the control groups of children used in the study of the effects of training. The ages for all children are calculated to January 1, 1931. The initial tests of the experimental children were completed in December 1930, while the controls were not completed until early in 1931. This discrepancy constitutes a defect in the experimental procedure which could not be overcome entirely. To compensate, in part, for this early discrepancy, the final tests of the controls were administered after final tests had been given to the experimental children. It can be seen further that there is a difference of over a month in the average ages of the two groups. Many of the children who had been given the initial tests and paired with children in the experimental groups were not obtainable for the final tests, due to contagious diseases. The use of other children who had been given the first tests and who were still available for the final tests increased the average age of the controls as indicated on the table. In the case of the interval tests, only sixteen controls were available for comparative purposes at the end of the study.

In spite of these shortcomings, a comparison of the two groups of children before and after training offers instructive findings.

The table gives results not only on the initial pitch and interval tests (Designated Test A, containing 11 and 12 notes and intervals respectively) but also on the extended tests introduced later in the study. The control group was measured on the extended scale only on the final tests. The initial pitch test was repeated three times, the initial interval test twice; the final pitch and interval tests were given three times to the controls, and twice to the experimental children.

In scoring the tests, only one credit was given for each item; if the child sang a given note correctly only once during the series or sang it on each repetition of the test, he received one credit for his performance in either case. Accordingly, the maximum score which a child could receive on



the initial and the final pitch test series was 11, and on the interval test series, 12. In the case of the extended tests, the maximum scores obtainable were 18 (pitch) and 22 (interval).

It appears in table 2 that the children who had received training scored considerably higher on the final tests than those who were tested only at the beginning and end of practice. The difference between the averages of the two groups on the final tests are reliable in all cases ( $\frac{\text{Difference}}{\text{S. D. difference}}$  ranges from 3.59 to 6.84).

S. D. difference

Tables 3 and 4 show the improvement in the average scores in pitch and interval during the course of training. Separate figures are given for the scores on the extended tests that were introduced two and a half months after the beginning of the study, when it became evident that most of the children were approaching a maximum score on the tests used at the beginning of the study.

The range of the scores on the initial pitch test represented in table 3 was 0 to 11 (maximum score, 11) on first presentation, and 8 to 11 on final presentation at the end of training. The range of the extended pitch test on first presentation was 2 to 16 (maximum 18), and on final presentation, 11 to 18. The scores on the interval test showed an initial range of 0 to 10 (maximum, 12), and on final presentation of the same test the range was 9 to 12. The extended interval test showed a range of 6 to 21, and 9 to 22 respectively on first and final presentations.

#### *Influence of training on variability.*

The results in tables 3 and 4 show a tendency toward a decline in relative variability as the training proceeds. Even in the case of the extended tests introduced late in the training period the relative variability is a good deal lower than on the initial more limited tests.

TABLE 2  
*Comparison between average scores of experimental and control children on pitch and interval tests given at beginning and end of study of effects of training*

|                                                                                                                | Experimental     | Control          |
|----------------------------------------------------------------------------------------------------------------|------------------|------------------|
| Pitch (18 pairs)                                                                                               |                  |                  |
| Average age in months . . . . .                                                                                | 38.18 $\pm$ 3.69 | 40.11 $\pm$ 3.93 |
| Average number notes correctly sung on first test (Test A, 11 notes) . . .                                     | 4.22 $\pm$ 2.44  | 4.22 $\pm$ 2.71  |
| Average number notes correctly sung on repetition of (Test A) at end of training . . . . .                     | 10.72 $\pm$ .11  | 6.44 $\pm$ 2.89  |
| Average notes correctly sung on extended test (Test B, 18 notes) at end of training . . . . .                  | 15.50 $\pm$ 1.78 | 8.00 $\pm$ 3.22  |
| Interval (16 pairs)                                                                                            |                  |                  |
| Average age in months . . . . .                                                                                | 38.75 $\pm$ 2.94 | 40.56 $\pm$ 3.69 |
| Average number of intervals correctly sung on initial test (Test A, 12 intervals) . . . . .                    | 4.25 $\pm$ 3.03  | 4.31 $\pm$ 2.89  |
| Average number of intervals correctly sung on repetition of Test A at end of training period . . . . .         | 11.50 $\pm$ .67  | 8.00 $\pm$ 2.50  |
| Average number of intervals correctly sung on extended test (Test B, 22 intervals) at end of training. . . . . | 17.0 $\pm$ 4.25  | 10.18 $\pm$ 4.26 |

The fact that there was a maximum limit to the score which a child could obtain on the tests, and the further fact that the same amount and character of training was given to all the children no doubt contributed largely to the apparent decline in variability as shown in tables 3 and 4. If the tests and the training had been graded according to individual differences exhibited early in the study it is possible that the improvement shown in final results would be even higher than the gains set forth in tables 3 and 4, and that individual differences would have remained more constant throughout the study. It is the marked improvement shown by the group as a whole in response to a training and test procedure constant for all children, rather than the findings with regard to changes in variability, that are of major interest in the results set forth in tables 3 and 4. (When the complete voice range of the 14 fathers and 15 mothers of the children who were given training were tested, the fathers sang an average of 20 notes, the mothers an average of 19. The children's average score of 15.5 notes on the final test, as indicated in table 3, does not differ widely from the average scores of the parents.)

TABLE 3

*Average scores on pitch test when repeated at various stages of training given to 18 3-year old children.  
Separate averages given for test as extended late in training period (Test B)*

| TEST A<br>11 ASCENDING NOTES FROM MIDDLE C TO F |                                           | TEST B<br>18 NOTES FROM G (BELOW MIDDLE C)<br>TO C (2 OCTAVES ABOVE MIDDLE C) |                                           |
|-------------------------------------------------|-------------------------------------------|-------------------------------------------------------------------------------|-------------------------------------------|
| Repetitions of test                             | Average number of<br>notes correctly sung | Repetitions of test                                                           | Average number of<br>notes correctly sung |
| December . . . . .                              | 4.22 $\pm$ 2.44                           |                                                                               |                                           |
| January . . . . .                               | 4.56 $\pm$ 2.90                           |                                                                               |                                           |
| February . . . . .                              | 6.50 $\pm$ 3.33                           |                                                                               |                                           |
| March . . . . .                                 | 8.78 $\pm$ 2.00                           | March . . . . .                                                               | 10.67 $\pm$ 2.78                          |
| May . . . . .                                   | 10.72 $\pm$ .11                           | May . . . . .                                                                 | 15.50 $\pm$ 1.78                          |

TABLE 4

*Average scores on interval test repeated at various stages of training given to 18 3year old children.  
Separate averages given for test as extended late in training period (Test B)*

| TEST A<br>ASCENDING AND DESCENDING MAJOR AND MINOR SECONDS,<br>MAJOR AND MINOR THIRDS; PERFECT FOURTHS AND FIFTHS<br>(TOTAL: 12 INTERVALS) |                                                  | TEST B<br>ASCENDING AND DESCENDING MAJOR AND MINOR SIXTHS, MAJOR<br>AND MINOR SEVENTHS, AND OCTAVE, PLUS INTERVALS OF TEST A<br>(TOTAL: 22) |                                                  |
|--------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------|
| Repetitions of test                                                                                                                        | Average number<br>of intervals<br>correctly sung | Repetitions of test                                                                                                                         | Average number of<br>intervals correctly<br>sung |
| December . . . . .                                                                                                                         | 3.94 $\pm$ 3.05                                  |                                                                                                                                             |                                                  |
| January . . . . .                                                                                                                          | 4.33 $\pm$ 2.89                                  |                                                                                                                                             |                                                  |
| February . . . . .                                                                                                                         | 8.05 $\pm$ 2.61                                  |                                                                                                                                             |                                                  |
| March . . . . .                                                                                                                            | 9.33 $\pm$ 2.33                                  | March . . . . .                                                                                                                             | 11.11 $\pm$ 3.44                                 |
| May . . . . .                                                                                                                              | 11.56 $\pm$ 1.00                                 | May . . . . .                                                                                                                               | 17.44 $\pm$ 4.00                                 |



### *Children's attitude toward prolonged training*

A further word should be said with regard to the response of young children to training in singing over a long period of time. The quantitative findings above have already indicated that the children reacted favourably to the procedure. During the course of the study evidences of the children's continued interest and willingness to cooperate appeared in many forms. At nearly all times the children willingly left their play to go to the room where the music was given. The children were never coerced. On rare occasions a child would temporize or refuse because of the attraction of the play in which he was engaged, but the following day he would often ask to be allowed to come to the music room. All of the children on many occasions asked for permission to come next, while many of them made daily requests for a turn at music. Sometimes a child would object to leaving the playground at the usual time prior to luncheon if he had not spent a period at music. On the whole, after the study was in progress, more persuasion was used in getting the child to leave the music room than in requesting him to come. The children tested for control purposes, although seen less frequently, showed an eager interest after brief acquaintanceship with the procedure.

Before the study was begun, the question as to whether the interest and cooperation of children could be maintained seemed almost as significant as the question regarding the possibility of improving children's performance in singing. The responsiveness shown both by the experimental and the control children used in this study, quite apart from any consideration of the objective findings, convince the investigators that much practical work in music can be done at this early age with encouraging results.

### *Tests covering songs used in training.*

Prior to the final pitch and interval tests, the children were tested on their ability to sing the songs which had been taught during training. The tests, covering the entire list of songs, were given twice on different days. Each child was tested individually. During the tests, the experimenter kept a record of the notes that were sung and of the child's reproduction of the rhythm of the songs. A stenographer kept a record of all the words spoken and sung by the experimenter and the children. At the conclusion of the tests, the records yielded a score for each child on reproduction of pitch, rhythm and words.

When the song tests were given, the child was first given opportunity to sing the songs to the accompaniment of an instrument but without verbal prompting. If he failed to proceed without help, or sang incorrectly, the experimenter prompted a second attempt. The stenographer made a record, in proper sequence, of all that was said. In scoring the child's performance two credits were given for each note reproduced without prompting, and one credit was given for a correct response following prompting. One credit was given for each word and rhythmic phrase correctly reproduced. In scoring the song which contained serial repetitions of a single word, credit was given to each group of identical words rather than to each separate word.

When the child's final score on these tests was computed, credit was given for each unit correctly reproduced, whether the item had been sung on both or only on one of the two presentations of the test. When the two repetitions of the song tests were scored separately, the correlations between the first and second presentations were as follows: Words,  $r = .83 \pm .05$ ; pitch,  $r = .75 \pm .07$ ; rhythm,  $r = .78 \pm .07$ .

The scores obtained on the final song tests were as follows: Pitch, range = 21 to 194; average =  $108 \pm 41$  (maximum possible score, 220). Words, range = 24-75; average =  $55. \pm 12$ . (maximum = 75). Rhythm, range = 5 to 14; average =  $10.33 \pm 2.62$  (maximum 14). Earlier in the training period tests were made of the children's ability to reproduce the notes of the songs first introduced. The maximum score on these earlier tests was 10. The average score obtained was  $4.67 \pm 2.96$ . The correlation between scores of pitch reproduction on these early tests and the final tests described above was  $.93 \pm .03$ .

At the end of the study all of the children could sing some of the songs and parts of others, while some of them could sing all the songs with little prompting. In a later section correlations between the scores on the above tests and other measures are given.

Examination of the results of the final song tests revealed that the low notes of the songs were reproduced more readily than the high notes. It appeared that the notes Middle C, D and Eb (which are supposed to fall below the child's favored range) scored respectively 68, 80 and 67 per



cent., while the higher notes D (above C an octave higher than Middle C), E and F (which have been regarded as well within the child's range) scored respectively only 50, 33, and 25 per cent.

### *Spontaneous vocalizations*

Records of the children's vocalizations while at play, obtained during a total of 100 minutes of observation of each child, with the procedure of recording only those notes which could definitely be verified by means of pitchpipes when the child sang, hummed, chanted, etc., provided a variety of findings.

The data were first examined to find the number of notes, intervals, and rhythmic measures recorded for each child. The notes sung by the child were scored by allowing one credit for each separate note recorded.

The records also provided a measure of the range of notes covered by each child in his spontaneous singing. This score represents the total number of different notes that were recorded for each child rather than the number of times each note was sung. A similar score was obtained for the number of different intervals which occurred in the child's vocalizations.

The data were further examined to find the number of times series of three notes or more which could definitely be placed within either the diatonic or chromatic scale occurred in the record of each child. Each record was also scored on the number of complete rhythmic measures which the observer had been able to identify clearly while the observations were being made.

Table 5 shows the averages and the average deviations of the various scores derived from the records of the spontaneous vocalizations of the eighteen children. Averages are given first for all vocalizations, including singing, calls, outcries, etc., while the second caption covers only the notes definitely sung, hummed or chanted. Averages are given for the total number of series of notes sung within a scale and separate averages are shown for the number of such series that could definitely be placed in the diatonic and chromatic scales. Series which might be regarded as either diatonic or chromatic were not included in the latter averages.

The results presented in Table 5 show wide individual differences. In a later section, the various measurements represented in the table are correlated with the results of tests and other measures.

TABLE 5

*Averages of the total number of notes and intervals, of the number of different notes and different intervals, rhythmic measures, and number of series of notes sung within a scale, occurring in records obtained during 100 minutes of observations of the spontaneous vocalizations of 18 children*

|               | Number of notes in all vocalizations | Number of notes sung, hummed, chanted | Number of different notes | Total number of intervals | Number of different intervals | Rhythmic measure | Scores of notes within a scale | Diatonic scale | Chromatic scale |
|---------------|--------------------------------------|---------------------------------------|---------------------------|---------------------------|-------------------------------|------------------|--------------------------------|----------------|-----------------|
| Range . . . . | 16 - 213                             | 7 - 207                               | 5 - 15                    | 2 - 45                    | 1 - 8                         | 0 - 11           | 0 - 10                         | 0 - 6          | 0 - 3           |
| Average . . . | 63                                   | 52                                    | 11.4                      | 12                        | 3.9                           | 3.9              | 3                              | 1.61           | 1.06            |
| A. D. . . . . | 35                                   | 37                                    | 2.2                       | 8.8                       | 2.3                           | 3.5              | 2.67                           | 1.48           | 1.11            |

TABLE 6

*Number of times each note occurred in the spontaneous singing, humming, chanting and outcries of 18 children during 100 minutes of observation devoted to each child*

|                                                     | MIDDLE C | C# | D  | D# | E  | F  | F# | G  | G# | A   | A#  | B   | C   | C# | D   | D# | E  | F  | ABOVE F |
|-----------------------------------------------------|----------|----|----|----|----|----|----|----|----|-----|-----|-----|-----|----|-----|----|----|----|---------|
| Singing . . . . .                                   | 0        | 3  | 5  | 4  | 10 | 63 | 30 | 53 | 61 | 172 | 81  | 62  | 84  | 28 | 72  | 3  | 12 | 0  | 3       |
| Chanting and humming . . . .                        | 0        | 0  | 2  | 0  | 8  | 6  | 5  | 6  | 17 | 16  | 15  | 20  | 40  | 8  | 25  | 10 | 6  | 9  | 0       |
| Outcries (including talking, calls, etc.) . . . . . | 0        | 0  | 4  | 4  | 0  | 2  | 5  | 10 | 12 | 38  | 21  | 29  | 24  | 6  | 22  | 8  | 8  | 6  | 2       |
| Total vocalizations . . .                           | 0        | 3  | 11 | 8  | 18 | 71 | 40 | 69 | 90 | 226 | 117 | 111 | 148 | 42 | 119 | 21 | 26 | 15 | 5       |

The figures showing the frequency of the diatonic and chromatic scales are small, but they are of interest in showing that children at this age do not limit their singing entirely to isolated notes.

Table 6 shows the frequency with which each of the various notes occurred in the spontaneous singing, humming, and chanting and in the children's calls, outcries, etc. Separate figures are given for frequencies in each category, as well as for all vocalizations combined.

The results shown in table 6 indicate that the spontaneous vocalizations tended to favor a relatively high range of notes. The notes from F above Middle C to D occurred most frequently. It is interesting to note that the children favored lower notes during the tests, as shown in previous sections, and somewhat higher notes when singing during their free play. It appears in table 6 that the higher notes occur relatively more frequently in the children's calls, outcries, etc., than in their actual singing.

TABLE 7  
*Frequency of occurrence of each interval recorded during observations  
of the children's spontaneous vocalizations*

| Ascending    |              |             |             |                |               | Descending   |              |             |             |                |               |
|--------------|--------------|-------------|-------------|----------------|---------------|--------------|--------------|-------------|-------------|----------------|---------------|
| Major second | Minor second | Major third | Minor third | Perfect fourth | Perfect fifth | Major second | Minor second | Major third | Minor third | Perfect fourth | Perfect fifth |
| 20           | 24           | 4           | 9           | 2              | 0             | 53           | 46           | 8           | 50          | 3              | 0             |

Table 7 shows the number of times each interval occurred in the records of spontaneous vocalizations.

The results presented in table 7 indicate that the descending intervals occurred more frequently than the ascending intervals in the children's spontaneous vocalizations. The closer intervals (seconds and thirds) scored considerably higher than the wider intervals (perfect fourths and fifths). This is in accord with the results obtained on the tests administered during the study.

It further appears in table 7 that the ascending and descending minor second or chromatic interval occurred more frequently than any other interval save the major second.

Reference to table 5 will show that series of three or more notes within the chromatic scale occurred almost as frequently as series within the diatonic scale. The results of the initial and final tests likewise showed a substantial score for the chromatic interval. It appears from all phases of the present study that chromatic intervals are not beyond the singing ability of the three- year old child.

#### *Intercorrelations in the musical behaviour of young children*

Table 8 shows the coefficients of correlation between various measures and scores obtained for the 18 children who were given training. All coefficients were obtained by the rank-difference method. The P.E.'s which are not entered, range from  $\pm .16$  with a coefficient of zero, to  $\pm .02$  with a coefficient of .95. To derive the measures represented by the caption "scores on pitch test (and interval test) after training," the children were ranked according to priority in obtaining a perfect score on the pitch and interval tests during the training series.

Most of the coefficients appearing in table 8 are positive. Intelligence quotient shows the lowest median coefficient, while the highest median coefficients are shown for total number of notes sung in spontaneous singing, and scores on rhythm and melody in the final song tests.

#### *Correlation between spontaneous singing, humming and chanting frequencies and frequency of laughter and overt activity.*

From other studies made with the same children, ratings were obtained on the frequency of the children's laughter and general overt activity. The correlation between children's laughter

frequencies and their scores on total number of notes sung, hummed or chanted during observations of spontaneous vocalizations was  $-.23$ ; between the latter measure and overt activity scores, the correlation was  $+.01$ . It would appear from these figures that there was no significant relationship between the children's tendency to sing, insofar as this tendency was measured in this study, and their tendency to laugh or to engage in physical activity.



TABLE 8

*Intercorrelations between various measurements obtained in the study of 18 trained children*

|                                                               | Pitch test before training | Interval test before training | Score on first test of songs | Pitch test after training | Interval test after training | Final song test: Melody | Final song test: Rhythm | Final song test: Words | Spontaneous singing: total notes sung | Spontaneous singing: number of different notes sung | Spontaneous singing: Total intervals | Spontaneous singing: Number of different intervals | Spontaneous singing: Total chromatics | Spontaneous singing: Total Rhythmic measures | Intelligence quotient |
|---------------------------------------------------------------|----------------------------|-------------------------------|------------------------------|---------------------------|------------------------------|-------------------------|-------------------------|------------------------|---------------------------------------|-----------------------------------------------------|--------------------------------------|----------------------------------------------------|---------------------------------------|----------------------------------------------|-----------------------|
| Score on pitch test before training . . . . .                 |                            |                               |                              |                           |                              |                         |                         |                        |                                       |                                                     |                                      |                                                    |                                       |                                              |                       |
| Score on interval test before training . . . . .              | .70                        |                               |                              |                           |                              |                         |                         |                        |                                       |                                                     |                                      |                                                    |                                       |                                              |                       |
| Score on first test of songs . . . . .                        | .52                        | .70                           |                              |                           |                              |                         |                         |                        |                                       |                                                     |                                      |                                                    |                                       |                                              |                       |
| Score on pitch test after training . . . . .                  | .56                        | .57                           | .79                          |                           |                              |                         |                         |                        |                                       |                                                     |                                      |                                                    |                                       |                                              |                       |
| Score on interval test after training . . . . .               | .47                        | .53                           | .68                          | .49                       |                              |                         |                         |                        |                                       |                                                     |                                      |                                                    |                                       |                                              |                       |
| Final song test: Melody . . . . .                             | .51                        | .58                           | .93                          | .76                       | .80                          |                         |                         |                        |                                       |                                                     |                                      |                                                    |                                       |                                              |                       |
| Final song test: Rhythm . . . . .                             | .39                        | .53                           | .92                          | .76                       | .57                          | .91                     |                         |                        |                                       |                                                     |                                      |                                                    |                                       |                                              |                       |
| Final song test: Words . . . . .                              | .49                        | .75                           | .81                          | .69                       | .70                          | .80                     | .73                     |                        |                                       |                                                     |                                      |                                                    |                                       |                                              |                       |
| Spontaneous singing: total notes sung . . . . .               | .09                        | .29                           | .52                          | .46                       | .48                          | .59                     | .63                     | .47                    |                                       |                                                     |                                      |                                                    |                                       |                                              |                       |
| Spontaneous singing: number of different notes sung . . . . . | .19                        | .00                           | .18                          | .38                       | .29                          | .29                     | .35                     | .22                    | .81                                   |                                                     |                                      |                                                    |                                       |                                              |                       |
| Spontaneous singing: total intervals . . . . .                | .05                        | .18                           | .39                          | .27                       | .41                          | .51                     | .55                     | .33                    | .93                                   | .72                                                 |                                      |                                                    |                                       |                                              |                       |
| Spontaneous singing: number of different intervals . . . . .  | .04                        | .18                           | .48                          | .26                       | .26                          | .50                     | .60                     | .37                    | .84                                   | .61                                                 | .93                                  |                                                    |                                       |                                              |                       |
| Spontaneous singing: total chromatics . . . . .               | .09                        | .12                           | .41                          | .34                       | .37                          | .45                     | .54                     | .32                    | .77                                   | .63                                                 | .85                                  | .85                                                |                                       |                                              |                       |
| Spontaneous singing: total rhythmic measures . . . . .        | .03                        | .39                           | .40                          | .23                       | .21                          | .29                     | .41                     | .32                    | .72                                   | .50                                                 | .67                                  | .62                                                | .53                                   |                                              |                       |
| Intelligence quotient . . . . .                               | .35                        | .42                           | .39                          | .46                       | .28                          | .25                     | .29                     | .31                    | .32                                   | .25                                                 | .32                                  | .35                                                | -.07                                  | .19                                          |                       |
| Median . . . . .                                              | .37                        | .475                          | .52                          | .475                      | .475                         | .545                    | .56                     | .48                    | .555                                  | .32                                                 | .46                                  | .49                                                | .43                                   | .395                                         | .315                  |

### *Correlations between parents and children*

As a further project, the voices of the children's parents were tested. To obtain a rating for the lower and upper ranges of the parents' voices, a numerical score was assigned to each note of the range covered by all the subjects. The highest and the lowest notes were given, respectively high and low numerical values. In rating the performance on the low notes, the highest rank was given to the lowest value. The correlations between the parents and children follow. Fifteen mothers, fourteen fathers, and thirteen cases of both parents combined, are represented in the correlations.

Child's score on first pitch test and voice  
range of mother = - .24; father + .21;  
both - .09

Child's score on final pitch test and voice  
range of mother = + .12; father + .20;  
both + .29

Child's ability to sing low notes on final test  
and lower range of voice of mother  
= + .17; father + .59; both + .62

Child's ability to sing high notes on final test  
and higher range of voice of mother  
= + .29; father + .17; both + .31

The coefficients above indicate that the child's final performance correlated higher with parents' range than did the child's first pitch score, and that children tend to resemble their parents more in the place of their voices (whether high or low) than in the range of their voices. The reliability of the coefficients is too low, however, to support any conclusions.

### *Sex differences*

The data were examined for evidences of sex differences. In the pitch tests, the boys scored consistently higher than the girls in number of notes sung and also in ability to sing both high and low notes. Since the number of cases was limited, this difference in favor of the boys may be due merely to chance.

## **MISCELLANEOUS OBSERVATIONS AND DISCUSSION OF RESULTS**

During the course of the study, certain general impressions were obtained, apart from the objective data. Some of these may be stated briefly.

It was frequently observed, in the case of children whose voices were placed low, that to concentrate the training on notes considerably higher than the child's favored range, and then to proceed downward toward this range, was more effective than to concentrate first upon notes immediately above the favored range.

It would appear that the general muscular adaptation involved in singing high notes differs somewhat from the motor pattern involved in the singing of low notes. The progression from low to high notes does not seem to constitute solely a continuous function but appears to entail also a shift in emphasis from one muscle group to another. In the present study it was often found that a child would be able during a period of time to sing some high notes and some low notes without being able to reproduce the notes in between. In a case such as this, it appeared that the child had 'found his voice' for high tones without, as yet, being able to make the transition between the muscle groups respectively involved in the singing of high and low tones.



This statement is offered only as a hypothesis. The authors do not have the empirical data necessary for a thorough treatment of the problem. It is possible that the phenomenon mentioned above may have some relationship with auditory tonal gaps.

Among factors responsible for poor performance in response to musical training, the authors would list the following: Inability to make the transition between high and low notes, as suggested above; presentation of songs pitched too high or too low for the particular child; the tendency of some children to sing too loudly; the failure of some children to give undivided attention to the particular items that are presented.

In the spontaneous singing, snatches of songs taught at home or at school occurred far less frequently than the children's own improvisations. The child's song sometimes related to his current activity or play, and sometimes contained only repetitions of syllables or words. These songs rarely exceeded three measures, but frequent repetitions of the same tune, sometimes with slight variations, were relatively common. The observer obtained the impression that the children sang somewhat more when alone than when in the company of several other children. The activities involving rhythmic motion, such as swinging and teetering, tended to be accompanied by singing more than other types of activity.

There was no evidence that the children strained their voices in singing the extended range of notes introduced into the training period. The practice of encouraging the children to sing softly no doubt helped to insure against strain.

The procedure of pairing a relatively good singer with a less capable child during the training periods seemed to be an effective device. Rivalry was not encouraged, nor did evidences of rivalry appear, but the presence of a child who would sing seemed to promote the efforts of other children.

## SUMMARY AND CONCLUSIONS

The present study was undertaken to investigate the ability of three year old children to reproduce pitch and interval, to study the effects of training on this ability, to investigate the content of children's spontaneous singing, and to find relationships between various phases of the child's singing behavior.

Initial tests were given to 48 children. Training was administered to 18 children over a period of 6 months. Each child received 40 ten-minute periods. Tests were given at intervals during the course of training. At the end of the study the trained children were tested and compared with children of a control group who had received no formal training since the administration of the initial tests. As a further project, each child who received practice was observed for a period of 100 minutes during his free play, and records were made of his spontaneous vocalizations.

The results show that on the initial tests the ascending notes from middle C to A were sung most readily by the children. This result indicates that children's voices are placed somewhat lower than has commonly been stated in previous writings. In subsequent song tests, and final pitch tests the same tendency to favor lower notes appeared.

The results further indicate that the narrow intervals (seconds and thirds) are sung more readily than the wider intervals (perfect fourths and fifths); that descending intervals tend to be sung more readily than ascending intervals; that the half-step or chromatic interval was sung quite readily in the test situation, and occurred frequently in the children's spontaneous vocalizations. This latter finding does not support the view that the chromatic interval should be excluded from songs for young children.

Training caused a marked improvement in the performance of the children. While the training was in progress it was found necessary to extend the range of notes and intervals covered in the study. On the three repetitions of the initial tests, 11 notes, extending from middle C to F above C an octave higher, were employed. The average number of notes initially sung by the 18 subjects used for training was  $4.22 \pm 2.44$ ; at the end of training the average score on this test was  $10.72 \pm .11$ .

The augmented test included the 11 notes of the initial test with four additional higher notes and three additional lower notes, extending from G below middle C to C two octaves above middle



C. The score on this extended test, when first introduced after two and a half months of training was  $10.67 \pm 2.76$ ; the final score on this test was  $15.50 \pm 1.78$ .

On the interval tests the results showed a similarly high improvement with training. In the two repetitions of the initial test, which included twelve intervals (ascending and descending major and minor seconds and thirds, perfect fourths and fifths) the children sang an average of  $3.94 \pm 3.05$  intervals, as compared with a final average score of  $11.56 \pm 1.00$  after training. An extended test, containing a total of 22 intervals, (including, in addition to the above, the ascending and descending major and minor sixths and sevenths, and octaves) was introduced after two and a half months of training. The average score on this test, when first introduced was  $11.11 \pm 3.44$ ; on the final test after further training, the average score was  $17.44 \pm 4.00$ .

The scores of the practiced children were reliably higher than the final scores of the control group.

The records of the children's spontaneous vocalizations showed wide individual differences. The chromatic interval occurred relatively quite frequently. The children tended to favor higher notes in their spontaneous vocalizations than in their performance in tests of songs and of pitch. Relative frequency of spontaneous singing, humming and chanting showed a negligible correlation with frequency of overt activity and of laughter, (records of which were made available through the courtesy of another investigator).

It appeared that children's voices tend to resemble the general trend of the parent's voice (whether high or low) more than their total pitch range. The reliability of all correlations between parents and children was low.

Apart from objective indications of the children's cooperation, as shown by progressive improvement in the test scores, the children gave evidence of interest in the musical exercises.

The results of the study as a whole indicate not only that children's performance in reproducing pitch and interval can be much improved through training, but also that children at this early age are more versatile in the use of their voices than has customarily been stated.

It is not claimed that the marked improvement shown by the children represents a change in capacity or an alteration of native ability. The present study does not answer the question as to whether the child who receives training in singing at an early age will have a permanent advantage over the child whose voice is not trained until later. The findings do suggest, however, that training at the age of three can be undertaken with promising results.

## REFERENCES\*

Damrosch, Walter; Gartlan, G. H.; and Gehrkins G. H.: Teacher's Book of the Universal School Music Series. New York, 1923.

Giddings, T. P.: Grade school music teaching. New York.

Course of study and syllabus in music for elementary schools. Board of Education, New York City, 1931.

Seashore, C. F.: The present status of research in the psychology of music. University of Iowa, Studies, 2, 1928, 1-29.

Thorn, A. G.: Music for young children. New York, 1929.

*Twenty-eighth year book of the national society for the study of education.* 1929, 720-722.

---

\* As referências bibliográficas deste artigo estão conforme o documento original.

## 2.1.3. Estudo Crítico – Por Helena Rodrigues

### I. Finalidade e problemas

#### a) Identificar finalidade e problemas do estudo analisando-os criticamente

Os autores afirmam que o estudo foi concebido para medir a capacidade de cantar das crianças de três anos de idade e para estudar os efeitos da aprendizagem no seu desempenho vocal. Estas parecem ser as finalidades da investigação relatada.

Entretanto, nas conclusões, os autores afirmam que o estudo foi efectuado para investigar a capacidade de crianças de três anos de idade para: reproduzir sons com uma dada altura; avaliar os efeitos da aprendizagem nesta capacidade; caracterizar o conteúdo do canto espontâneo das crianças e encontrar relações entre várias fases do comportamento vocal da criança (p.289). (Note-se que não é claro o que querem dizer com "encontrar relações entre várias fases do comportamento vocal da criança"). Estes parecem ser os problemas da investigação relatada.

Escreve-se "parecem" pois pretendendo ser rico em termos de recolha de informação no âmbito temático do desempenho vocal de crianças de três anos de idade, o estudo acaba por se tornar confuso em termos de identificação de finalidades e respectivos problemas.

Por vezes, converte-se numa "miscelânea" (a expressão é utilizada pelos próprios autores) de dados relacionados com o desenvolvimento vocal de crianças com três anos de idade, agregando questões e informações pouco focalizadas. Com o desenrolar do artigo vão-se incluindo outros "estudos" que se misturam e afastam do propósito inicial.

Não obstante, não deixa de ser um artigo estimulante, que nos faz pensar que, apesar de terem decorrido mais de setenta anos sobre a sua publicação, ainda sabemos objectivamente muito pouco sobre o desenvolvimento musical e vocal de crianças desta faixa etária.

#### b) Relacionar finalidades e problemas com conclusões e resultados do estudo

Os autores referem os seguintes resultados: os sons situados entre C3 e A3 eram facilmente entoados pelas crianças; a tessitura entre F3 a D4 era privilegiada nas vocalizações espontâneas; os intervalos de 2ª e 3ª eram mais facilmente entoados do que os de 4ª e 5ª; os intervalos descendentes eram mais facilmente entoados do que os ascendentes; os intervalos cromáticos eram facilmente cantados, ocorrendo frequentemente nas vocalizações espontâneas das crianças; a instrução causou melhorias no desempenho vocal das crianças; existem grandes diferenças individuais nas vocalizações espontâneas; as notas agudas ocorrem mais frequentemente nas vocalizações espontâneas que na situação de teste.

Estes dados relacionam-se com o desempenho vocal das crianças. Alguns deles estão directamente relacionados com os problemas acima enunciados. Descreve-se a tessitura e os intervalos mais facilmente produzidos por crianças desta faixa etária, evidenciando-se melhorias nestes aspectos após um período de instrução.

Se se aceitar que reproduzir sons e intervalos é sinónimo da capacidade de cantar das crianças então poderá dizer-se que finalidades e conclusões estão relacionadas.

No entanto, será que as capacidades de reproduzir sons e intervalos (basicamente estes são os critérios de medida para avaliar os efeitos da instrução) medem, efectivamente, a "capacidade de cantar"? Ou seja, se as finalidades não estão relacionadas com os problemas, como podem os resultados relacionar-se com as conclusões? Ou, em última análise, como podem as finalidades estar relacionadas com as conclusões?

Por outro lado, o estudo relata uma série de dados que se desviam dos problemas colocados inicialmente. De facto, procurou estabelecer-se relação entre as capacidades vocais das crianças e factores como a capacidade de aprendizagem de melodias e palavras das canções, a inteligência geral, a tessitura vocal dos Pais e a tessitura e conteúdo das vocalizações



espontâneas. Procurou-se, também, estabelecer relações entre a extensão vocal das crianças e a extensão vocal dos progenitores – embora o próprio refira isto como *"further project"* (cf. p. 288). E são relatados ainda outros dados que parecem não se enquadrar numa linha de pensamento coerente: *"the correlation between the children's laughter frequencies and their scores on total number of notes sung, hummed or chanted during observations of spontaneous vocalizations was -.23; (...) it would appear from these figures that there was no significant relationship between the children's tendency to sing, insofar as this tendency was measured in this study, and their tendency to laugh or to engage in physical activity"* (p. 288).

As finalidades do estudo são amplas e abarcam vários problemas. Mas são enunciados problemas cuja natureza se afigura diferente das finalidades. Acresce ainda que são relatados resultados que respondem a problemas não enunciados. Estes dois factos produzem confusão e informação desfocalizada.

Em suma, ressalta no artigo um conjunto diverso de dados coleccionados sem que se perceba o seu fio condutor, a sua finalidade. Desta indefinição básica decorrem várias fragilidades do estudo.

## **II. Reflexão crítica sobre aspectos metodológicos abordando aspectos relativos aos sujeitos, instrumentos e procedimentos do estudo.**

Relativamente à selecção dos sujeitos o estudo não fornece dados claros relativamente ao modo de selecção da amostra. Não é indicada a origem nem o sexo dos sujeitos. Embora o Grupo de Controle e o Grupo Experimental tenham sido constituídos com base no emparelhamento de idades e resultados no Pré-Teste, sendo os dois grupos constituídos por 18 sujeitos, no caso do teste de intervalos só foi possível recolher 16 resultados do Grupo de Controle para objectivos de comparação entre os grupos. O estudo refere que foram usados 48 sujeitos, tendo 18 crianças sido submetidas a instrução musical. Aparentemente, as 30 crianças restantes foram submetidas ao Pré-Teste a que se sujeitaram também as crianças do Grupo Experimental. Efectivamente, os autores revelam que em virtude de ter havido elevada morte experimental por parte dos sujeitos do Grupo de Controle inicialmente emparelhados com o Grupo Experimental, foram depois usados outros sujeitos a quem tinham sido também administrados os primeiros testes. O próprio estudo assinala ainda uma dificuldade relativa à recolha de dados dos sujeitos: o Pré-Teste do Grupo Experimental foi concluído mais cedo que o do Grupo de Controle. Para compensar isto, o Pós-Teste foi administrado ao Grupo de Controle também depois de administrado ao Grupo Experimental.

É certamente positivo que estes pormenores sejam revelados na descrição da investigação, sem falsos receios de deixar a nu algumas das suas fragilidades. Não é claro, contudo, na planificação do estudo porque é que também foram administrados os primeiros testes a outros sujeitos que à partida não estariam seleccionados para o Grupo de Controle. Isto é, quase se fica com a sensação que o "design" do estudo é posterior à recolha de dados.

Do mesmo modo, os instrumentos usados não são descritos exhaustivamente. Exemplos: seria importante conhecer exactamente que canções foram usadas na instrução musical; a referência à utilização de um xilofone nada esclarece quanto à tessitura do mesmo.

Em termos de procedimentos, registre-se que apesar de pretender ser um estudo experimental, devidamente controlado, o estudo recorre a algumas estratégias de carácter educativo que, embora podendo ter dificultado a padronização das situações, podem ter ajudado a criar um ambiente "natural", semelhante ao da sala de aula (por exemplo, as crianças foram colocadas em pares – uma com mais capacidade que outra – que iam sendo rodados).

Ainda em termos de procedimento, recorde-se, como exemplo, que para o teste de imitação de sons de altura diferente o experimentador cantava um som e depois tocava o mesmo num dos instrumentos musicais (o que em termos pedagógicos poderá ser discutível). Era pedido às crianças que escutassem e cantassem o som previamente ouvido. Cada item era apresentado por sua vez. O item era registado se a criança cantava correctamente, passando-se então ao item seguinte. Se falhasse ou não tentasse sequer, o experimentador insistia até um máximo de oito vezes. Se se não conseguisse manter a atenção da criança não se chegava às oito apresentações. O teste era classificado de acordo com três possibilidades: "consegue"; "tentou";



"não tentou". O teste de imitação de sons de altura diferentes foi administrado três vezes, em três ocasiões diferentes. A criança recebia sempre a pontuação 1 independentemente de ter acertado uma ou três vezes na tarefa. Perante o exposto fica-se com a sensação de que o procedimento era complexo e pouco prático.

Também o sistema de atribuição de pontuações parece susceptível de alguma ambiguidade (ou, pelo menos, não é devidamente justificado). De facto, sendo os testes repetidos, as crianças recebiam sempre a mesma pontuação – era atribuída a pontuação 1 – independentemente de a criança executar a tarefa correctamente uma, duas ou três vezes. De facto, se se pode argumentar que se a criança consegue executar a tarefa uma vez é porque de facto a consegue executar (podendo apenas estar inibida nas outras vezes em que se lhe pede a repetição da tarefa), também se pode argumentar que se a criança não foi capaz de repetir a execução da tarefa é porque se tratou de um sucesso aleatório. Assim, não se percebe por que razão foram repetidos os testes de altura e de intervalos. De igual modo, não se entende por que razão as crianças do Grupo de Controle foram testadas de modo diferente das do Grupo Experimental (*"the final pitch and interval tests were given three times to the controls, and twice to the experimental children"* (p. 280)). Analogamente, não se entende a seguinte afirmação: *"the control group was measured on the extended scale only on the final tests"* (p. 280).

A medida que o estudo se foi desenrolando o conteúdo das sessões de instrução musical e do teste foi alterado (alargando-se o âmbito da tessitura vocal e dos intervalos). Isto sucedeu porque à medida que o estudo se ia desenrolando as crianças iam melhorando o nível do seu desempenho, deixando o Pré-Teste de diferenciar o desempenho dos sujeitos, quer interindividualmente como intra-individualmente. Assim, foi necessário aumentar o número de sons e intervalos usados em ambos os testes, passando também a usar-se um xilofone para além do saltério e dos lamirés usados inicialmente.

O facto de se usar um Pós-Teste diferente do Pré-Teste (em condições diferentes no que toca também à utilização do xilofone que viria a substituir o saltério e o lamiré) leva-nos a questionar aspectos relativos ao procedimento experimental do estudo. Como estabelecer comparações entre resultados obtidos através de instrumentos de medição diferentes?

Por outro lado, também é certo que se se não o fizesse e se tivesse aplicado o Pré-Teste ao Grupo Experimental depois de este ter sido sujeito a instrução, não haveria lugar à diferenciação dos sujeitos e o teste seria inválido. O estudo alerta-nos para um problema de difícil resolução: um teste pode ser válido antes da implementação de uma intervenção educativa, passando a ser inválido depois da mesma (porque deixando de diferenciar os sujeitos deixa de ser fiável, e deixando de ser fiável deixa de poder ser válido). Por outro lado, a utilização de dois testes diferentes (um Pré-Teste antes da intervenção educativa, o Pós-Teste depois) cria problemas em termos de procedimento experimental, pois não faz sentido comparar resultados supostamente decorrentes de uma variável independente quando a variável dependente está a ser medida através de dois instrumentos diferentes.

Acresce ainda que ao contrário dos sujeitos do Grupo de Controle, os sujeitos do Grupo Experimental estiveram sempre em contacto com o experimentador. Esta familiaridade pode tê-los ajudado a desempenhar melhor os Pós-Testes, actuando assim como variável parasita. Do mesmo modo, o simples facto de o Grupo de Controle não estar sujeito a nenhum tratamento pode ser relevante. Para todos os efeitos, durante o período de instrução, as crianças do Grupo Experimental têm a noção de que estão a "aprender música" e esta noção pode contribuir para o seu sucesso nos Pós-Testes de desempenho vocal.

No que concerne à recolha de dados no âmbito do registo das vocalizações espontâneas das crianças, note-se que cada criança foi seguida e registada pelo observador durante as manhãs no recreio da sua escola num total de 100 minutos, distribuídos por, pelo menos, dez períodos de observação, em diferentes alturas, ao longo de três a cinco meses. É sem dúvida de registar o esforço dos autores, embora a limitação técnica de o observador ter de seguir a criança possa ter trazido alguns condicionalismos à espontaneidade das crianças. Muito provavelmente também, terá sido difícil assegurar condições de recolha idênticas para as produções vocais de todos os sujeitos. É curiosa a utilização da estenografia, o que revela a ênfase colocada na recolha de palavras. O conteúdo desta recolha parece ser uma "miscelânea" (no artigo diz-se que o observador registava padrões rítmicos, conteúdo verbal das respostas, actividades acompanhando as respostas, etc). Note-se, contudo, que os próprios autores assumem limitações nos registos, não devendo as mesmas ser consideradas como uma medida absoluta. Talvez os



autores se tenham interessado pelas vocalizações espontâneas das crianças pois de algum modo elas poderiam ajudar a criar uma linha de base a partir da qual se tirariam conclusões relativamente aos efeitos da aprendizagem. No entanto, o estudo deste problema mereceria, por si só, ser alvo de uma investigação autónoma.

Em suma, são várias as situações em que o estudo se revela confuso, carecendo de rigor. Por vezes, o "design" parece algo desgovernado, ao sabor das circunstâncias, o que contrasta com o carácter experimental e controlado que o estudo parece pretender.

#### **a) De que tipo ou tipos de estudo se trata**

O artigo reúne dois tipos de estudo.

Um é um estudo experimental em que se pretende verificar os efeitos da instrução musical (Variável Independente ou Variável Preditiva) sobre o desempenho vocal (Variável Dependente ou Variável Criterial), medido através de testes de reprodução de intervalos e de sons de altura diferente (critérios de medida). Este estudo visa comparar resultados do Grupo Experimental (que recebeu instrução) com resultados do Grupo de Controle (sem instrução).

O outro é um estudo de observação naturalista. Visa descrever o conteúdo das vocalizações espontâneas das crianças de três anos de idade na situação de recreio.

#### **b) Razão para testar o Grupo de Controle no princípio e no fim do estudo**

A utilização de um Grupo de Controle (que por definição é submetido aos mesmos Pré-Testes e Pós-Testes que o Grupo Experimental) destina-se a verificar se eventuais melhorias obtidas no desempenho do Grupo Experimental são uma consequência dos efeitos de instrução a que o grupo foi submetido ou se essas possíveis melhorias se devem a efeitos de crescimento/maturação e de familiaridade com os instrumentos de medida.

Ou seja, dado que ambos os grupos são sujeitos aos mesmos efeitos de crescimento, e dado que os procedimentos de medida se vão tornando igualmente familiares para ambos os grupos, se houver diferença entre os dois grupos ao nível do desempenho vocal (variável dependente) é legítimo assumir que essa diferença se deve à actuação da variável independente (instrução musical).

#### **c) Discussão acerca da validade e fiabilidade dos procedimentos usados na avaliação do desempenho vocal dos sujeitos**

No que respeita ao controlo da fiabilidade foram várias as estratégias usadas de forma a verificar previamente a fiabilidade do experimentador: utilizou-se um músico para testar as capacidades auditivas do experimentador (mas será que não se deveria testar também o músico?); utilizaram-se juízes independentes e ainda um outro músico com "ouvido absoluto" tendo-se verificado um nível de concordância elevado entre estes e o experimentador. Se por um lado estas medidas mostram um cuidado muito apreciável por parte dos autores em busca da objectividade necessária num estudo científico (cuidado este que por vezes escapa em artigos publicados na actualidade por prestigiadas revistas da especialidade), por outro lado a diversidade de estratégias utilizadas faz ressaltar o carácter pouco focalizado do artigo. A própria definição de "ouvido absoluto" não é pacífica.

A utilização de diferentes recursos instrumentais – voz, lamiré, xilofone, saltério – pode também ter contribuído para a não padronização das situações a testar.

No que toca ao registo das vocalizações espontâneas a tarefa parece especialmente falível por questões de praticabilidade (mesmo dispondo de meios sofisticados de gravação como os de que dispomos hoje, tal tarefa afigura-se complicada em termos práticos).



Ainda no que toca à fiabilidade, o estudo refere: *"the scores of the practiced children were reliably higher than the final scores of the control group"* (p. 290). Note-se, no entanto, que a fiabilidade neste caso não surge quantificada nem se explica como foi calculada.

Ao criar um teste com notas e intervalos para avaliar a capacidade de cantar dos sujeitos, desde logo se pode colocar em questão a validade de conteúdo da medida. Cantar é só entoar um som de altura determinada? Ser capaz de entoar dois sons distintos com maior ou menor diferença de alturas é cantar?

No que parece ser uma afirmação relativa à validade do teste, os autores afirmam *"testing procedure correspond quite closely to the type of test situation which an individual must face under practical circumstances, in school or even in performance of a professional character"* (p. 276). Haverá, porém, que colocar reservas relativamente à validade de conteúdo de um teste que para avaliar a capacidade de cantar das crianças se centra na avaliação de comportamentos parcelares como é o caso da avaliação da entoação de sons isolados e de intervalos. No mínimo, se se aceitar que efectivamente é verdade que o teste corresponde muito de perto às situações práticas com as quais os sujeitos se confrontam na prática, então dever-se-ia questionar a validade de conteúdo musical dessa escola ou dessa performance. Por outro lado, devemos ter presente que o conteúdo deste teste mais não é do que o reflexo do espírito da época – bastará termos presente a forma como, na época, a aptidão musical era testada por Seashore.

Finalmente, uma última reflexão que nos permite confrontar os conceitos de fiabilidade e validade: poderá ser relativamente fácil estabelecer a fiabilidade da medição do desempenho vocal no que toca à imitação de sons ou de intervalos através de testes como os que foram usados no presente estudo. Apesar de fiáveis, estes testes podem contudo ser inválidos enquanto medição da capacidade da criança cantar. Estando a validade em causa, a fiabilidade refere-se a uma medida que pode não fazer sentido.

Note-se, contudo, que o Quadro 8 do artigo (p.287) mostra a existência de correlações elevadas entre os dois tipos de desempenho (entoação de sons isolados e intervalos correlacionando-os com a entoação e canções) o que pode ser visto como uma medida de validade concorrente. Ou seja, em última análise, afinal pode fazer sentido avaliar a capacidade de cantar mediante a utilização de testes semelhantes aos usados neste estudo.

### III. Análise do enquadramento teórico do artigo

As referências teóricas do artigo são escassas, o que não é de admirar dado que se trata de um artigo datado de 1931, altura em que a Psicologia da Música dava ainda os primeiros passos. Citam-se alguns trabalhos que parecem ter um enfoque mais educativo que propriamente investigativo (caso de Thorn, Damrosch, Gartlan e Gehrkins, cf. p. 279). A propósito de trabalhos empíricos efectuados em Música mencionam-se o *Twenty-eighth Year Book of the National Society for the Study of Education* (1929) e um artigo de Seashore (1928) que refere estudos efectuados na Universidade de Iowa até 1928 – mas nenhum destes aborda o tipo de problemas abordados no presente estudo.

Apesar de o estudo não expressar premissas teóricas pré-existentes, nem assumidamente fundamentar as suas opções, a verdade é que o tipo de problemas que interessa aos autores denuncia, por si só, um determinado tipo de orientação teórica. O estudo enquadra-se e é contagiado pela filosofia atomista que caracterizava os primórdios da Psicologia da Música. Assim, a capacidade de cantar é decomposta em comportamentos parcelares como a entoação de intervalos e a produção de sons de altura diferente. Mas será que, não obstante poderem ser fiáveis, estes critérios são válidos enquanto medidas que reflectem e/ou predizem a "capacidade de cantar"?

Levantar esta questão é interrogarmo-nos sobre as premissas teóricas que sustentam esta investigação. Em última análise, estamos perante um problema filosófico de partida – e não nos podemos esquecer que estes determinam sempre quer a concepção como a apreciação de qualquer investigação.

Ou seja, afirmar que "capacidade de cantar" não é sinónimo de extensão vocal exprime também uma posição filosófica. No entanto, hoje dispomos de vários dados sujeitos à confirmação objectiva dos factos que nos levam a pensar noutro tipo de modelos. Por exemplo, à luz das teorias da *Gestalt* é-se levado a considerar que percebemos em termos de conjuntos de sons e de padrões – e, portanto, se tal acontece para a percepção é bem possível que a emissão vocal



se sustente no mesmo tipo de princípios. À luz da teoria de aprendizagem musical somos levados a pensar que a capacidade de cantar assenta numa combinação de instrução musical e capacidades de audição do indivíduo. Esta expressa-se por uma forma de pensamento musical contextualizado em que um conjunto de sons é diferente do significado particular de cada um isoladamente. Portanto, à luz deste conceito, a medição da capacidade de entoar intervalos ou de produzir sons isoladamente não é uma medida válida de avaliação da capacidade de cantar.

As teorias levam-nos à procura de factos. Os factos confirmam ou infirmam as teorias. Ao analisarmos os factos deste artigo que poderiam, aparentemente, apoiar a teoria de que a instrução melhorou a capacidade de cantar, somos confrontados com questões terminológicas de base, interrogando-nos sobre o significado de "capacidade de cantar".

Por vezes, somos levados a pensar que "uma palavra é uma teoria".

#### **IV. Reflexão sobre as possibilidades e limitações da investigação em Psicologia e Pedagogia Musical contribuir para a tomada de opções de carácter educativo, com base em excertos do artigo**

Uma grande parte das decisões no domínio educativo baseia-se em intuições, crenças e experiência adquirida. Poucas são as que se baseiam numa metodologia científica assente na observação e experimentação sistemática, da qual se retirem conclusões que possam aplicar-se à prática educativa. Este comentário aplica-se com toda a propriedade à prática da Pedagogia Musical, tantas vezes assente numa transmissão de saber dogmático em que ao longo do tempo os alunos vão imitando o Mestre, transformando-se em especialistas na arte de reproduzir intérpretes de interpretações recebidas.

O presente estudo tem o mérito de pôr à prova algumas das convicções que são inquestionavelmente assumidas quando se trata do desempenho vocal das crianças. Concretamente, os autores põem em causa a afirmação de uma das fontes bibliográficas citadas que defendia que o intervalo cromático não deveria ser usado nas canções até ao 5º ano de escolaridade. De facto, verificam que no final da instrução todos os sujeitos do grupo experimental eram capazes de cantar intervalos cromáticos e que o mesmo era dos mais frequentemente entoados nas vocalizações espontâneas das crianças. Citando os autores: *the score for the minor second interval is relatively quite high* (p. 279), *and does not support the view that the chromatic interval should be excluded from songs for young children* (p. 290).

Com base nos dados encontrados, os autores constatarem também que a colocação das vozes é mais grave do que o que habitualmente se supunha. Usando as suas próprias palavras: *"the findings with regard to pitch differ somewhat from usual statements with regard to the voices of young children. The claim is usually made that songs for young children should fall within the compass extending from E first line to E fourth space. The authors have been unable to find published experimental data to support this claim. Even less definite are many writings on methods of teaching singing to young children which to state that songs should be within the compass of the child's voice, without giving definite indications of what this compass is"* (p. 278).

Nos dias de hoje, são múltiplos os exemplos de situações musicais que usamos para instruir as crianças sem sabermos exactamente o que está ou não ao alcance delas e quais os progressos que vão realmente efectuando. Na actualidade, o "preconceito educativo" segundo o qual as crianças não conseguem entoar intervalos cromáticos vai ainda sendo passado em vários programas de formação de professores. É uma convicção sem sustentação empírica, baseada na transmissão de um saber dogmático.

Na prática tenho observado, pelo contrário, que mesmo antes dos três anos de idade há crianças que conseguem entoar canções com este intervalo e que se revelam particularmente interessadas pelo mesmo. Estudos como o aqui relatado, devidamente controlados, podem ajudar a aferir conhecimentos, questionando o saber dogmático e os dados da observação prática de cada professor.

Demonstrando que os sons compreendidos entre Ré3 a Lá3 eram melhor reproduzidos, que os intervalos de 2ª e 3ª eram mais facilmente reproduzidos e que os intervalos descendentes eram ligeiramente mais fáceis de reproduzir, o estudo elucida-nos de forma objectiva sobre características do desenvolvimento vocal das crianças desta faixa etária, podendo estes



conhecimentos ser integrados numa proposta educativa. (Curiosamente, Edwin Gordon na sua teoria de aprendizagem musical defende a utilização de padrões tonais com este tipo de características no estágio de aculturação. Dado que a sua teoria assenta em dados da observação empírica é bem possível que tenha efectuado experiências semelhantes chegando ao mesmo tipo de conclusões).

Mas o artigo vai ainda mais longe no que concerne a reflexões de carácter educativo. Por exemplo, refere que a análise dos resultados obtidos longitudinalmente revela uma diminuição na variabilidade à medida que a instrução se processa, argumentando que tal pode dever-se ao facto de ter sido oferecida a todas as crianças a mesma quantidade/ qualidade de treino e de haver um limite máximo no resultado possível de obter. Ou seja, se os testes e a instrução oferecida tivessem sido graduados de acordo com as diferenças individuais exibidas no início do estudo seria possível que os ganhos observados pudessem ainda ser melhores e que as diferenças individuais tivessem permanecido mais constantes ao longo do estudo. Esta é uma observação interessante a ter presente num debate sobre avaliação numa perspectiva criterial e normativa.

Os autores discutem quais os factores que poderão explicar o fraco desempenho de algumas crianças face à instrução musical e aconselham algumas estratégias de ensino (por exemplo, o procedimento de, durante o período de instrução, colocar uma criança que cante bem ao lado de uma com pior desempenho revelou-se uma medida eficaz; o procedimento de, nas crianças com colocação de voz grave, se concentrar o treino em notas bem acima desta tessitura e depois ir descendo em vez de procurar as notas vizinhas da sua tessitura, etc). À medida que a instrução se foi processando, os autores – contrariando expectativas fundadas no que era tradicionalmente dito relativo à extensão vocal das crianças – notaram que transpondo as canções se podia, em determinadas situações, obter uma resposta mais eficaz por parte das crianças. Recorde-se ainda que os testes foram sendo adaptados à instrução que foi sendo implementada e que foram contabilizados como períodos de instrução, dado que eles próprios foram considerados como uma forma de treino.

Finalmente, refira-se que os investigadores relatam ter observado grande interesse por parte das crianças a par de melhorias de desempenho evidenciadas, comentando que se pode fazer muito trabalho prático nesta idade. Embora este comentário possa parecer ingénuo, note-se que foi produzido em 1931. Devidamente contextualizado na época, é possível que tenha sido também revolucionário, desafiando ideias preconcebidas relativas à idade em que era suposto iniciar-se a educação musical de uma criança.

Em suma, a Pedagogia Musical pode e deve socorrer-se do pensamento científico e de dados da Psicologia da Música para fundamentar a sua prática. Seria, no entanto, impossível construir uma Pedagogia decompondo e sujeitando a provas científicas de verificação todo o seu saber. A intuição, a experiência do vivido e a observação directa são igualmente fontes de informação relevantes para a sua actuação. Este artigo mostra como ambas as abordagens podem contribuir para o aperfeiçoamento das práticas educativas.

## **V. Reflexões e problemas suscitados pelo estudo**

A recolha e análise das vocalizações espontâneas das crianças da faixa etária compreendida entre os dois e os quatro anos de idade é um tema insuficientemente explorado na literatura e com grande interesse. Efectivamente, trata-se de uma faixa etária muito rica em termos de produção vocal.

São frequentes as improvisações musicais, o cantarolar, a invenção de palavras, jogos de palavras, a repetição de palavras e sílabas, a invenção de letras para melodias conhecidas, a construção de novas melodias a partir da colagem de fragmentos de melodias, as repetições de frases e palavras com diferentes entoações, os diálogos cantados, as repetições da mesma canção, a associação destes jogos sonoros a actividades do seu dia a dia, etc. Enfim, uma série de comportamentos que podem fazer pensar que desenvolvimento musical e linguístico se apoiam mutuamente. Será que faz sentido observar o desenvolvimento de ambas as capacidades como faces de uma mesma moeda? Como se processa a autonomização de ambas as capacidades? Há diferenças individuais? Como? Quais? Que etapas? As produções vocais ditas "espontâneas"



são diferentes conforme tenham sido precedidas de actividades de carácter musical ou de carácter linguístico?

No texto, os autores referem: *"the observer obtained the impression that the children sang somewhat more when alone than when in the company of several other children"* (p.289) e *"the activities involving rhythmic motion, such as swinging and teetering, tended to be accompanied by singing more than other types of activity"* (p. 289). Efectivamente, verifica-se muitas vezes que é em situações de jogo solitário que, espontaneamente, se podem observar grandes produções vocais de tipo exploratório. Este tipo de comportamentos parecem ser essenciais em termos de desenvolvimento musical (e talvez mesmo em termos de desenvolvimento linguístico). Assegurar a "espontaneidade" dos comportamentos observados coloca dificuldades metodológicas em termos de recolha de dados. No entanto, a aprendizagem informal – e solitária – das palavras e da música (é a música que "ensina" a palavra? são as palavras que ajudam a memorizar/ explorar pequenos fragmentos musicais?) e a espontaneidade das produções vocais marcam fortemente o desenvolvimento na infância. Deveríamos conhecê-las melhor.

No estudo, a instrução musical foi realizada com canções com texto – estabelecendo-se critérios de avaliação que tinham em conta o facto da criança entoar com as palavras exactas ou idênticas. A questão da utilização de canções com texto *versus* canções sem texto sugere problemas que não estão ainda suficientemente explorados na literatura: memoriza-se melhor canções ensinadas com texto do que sem texto? "compreende-se" melhor musicalmente canções ensinadas com texto ou sem texto? imita-se melhor canções ensinadas com texto do que sem texto? há relação entre memorização, compreensão musical e imitação? a longo prazo há diferenças nestes desempenhos – ou seja, de que forma a inclusão de texto afecta o armazenamento da memória nestas três situações, conforme se trate de memória a curto prazo e a longo prazo? a afinação é afectada pelo facto das canções serem ensinadas com texto ou sem texto? e o desempenho rítmico? há relações específicas entre desempenho rítmico, desempenho tonal e utilização de palavras? há diferenças nestes tipos de desempenho devido a diferenças individuais de carácter musical e/ou linguístico?

Um outro problema para o qual o texto sensibiliza diz respeito à estabilidade dos ganhos de aprendizagem. Ou seja, a criança que recebeu instrução em canto numa idade precoce terá uma vantagem permanente sobre as crianças cujas vozes não foram treinadas senão mais tarde? Existe no desenvolvimento humano um "período crítico" para que a capacidade de cantar se inscreva no comportamento de forma definitiva, como algo de permanentemente adquirido? Que sucede se se procurar "ensinar a cantar" nesta idade sem que haja, antes, a prévia exposição da criança a exemplos de música cantada?

É certo que sobre estes assuntos temos a perspectiva e as convicções de vários pedagogos mas carecemos ainda de provas científicas que confirmem ou infirmem a veracidade das mesmas.

Finalmente, uma referência a exemplos deste estudo para ilustrar a diferença entre avaliação e medição. Recorde-se que cada criança era testada individualmente, três vezes, em dias diferentes, expondo-a ao item durante oito tentativas, no caso do teste de imitação de sons de altura diferente. A forma de atribuir a classificação era a seguinte: pontuação de 2 se a criança imitasse o som sem ajuda; 1 se fosse a seguir a ter ajuda. A pontuação era atribuída por tarefa, independentemente de a criança a ter cumprido uma, duas ou três vezes. A avaliação diz respeito às várias decisões subjectivas contidas neste procedimento, enquanto que a medição é algo de objectivo.

Concretamente, optar por estabelecer diferenças entre imitar um som sem ajuda ou com ajuda ou se a pontuação era atribuída por tarefa independentemente do sucesso do desempenho da criança se ter repetido, são decisões que se referem ao procedimento de avaliação (e que revelam sempre uma determinada idiosincrasia por parte do investigador). A decisão de pontuar com 2 ou com 1 um determinado comportamento é igualmente subjectiva (revelando também uma dada idiosincrasia mesmo que não expressa) mas a sua atribuição é objectiva em face do critério definido que determina quando é atribuída a classificação de 2 ou de 1. Ou seja, a medição visa a objectividade. A fiabilidade determina o grau de rigor desta objectividade. Mas não nos podemos esquecer que esta objectividade sucede ao estabelecimento de um critério que é de natureza subjectiva.



No caso concreto, bastaria, por exemplo, que o examinador determinasse que o desempenho das crianças não fosse avaliado por cada tarefa, mas sim pelo seu desempenho nos três dias distintos, para que os resultados fossem diferentes embora continuando a ser consequência de uma medição objectiva.

## 2.2.

### MUSIC RESEARCH IN MEDICAL / DENTAL TREATMENT: META-ANALYSIS AND CLINICAL APPLICATIONS

INVESTIGAÇÃO MUSICAL EM TRATAMENTOS MÉDICOS E ODONTOLÓGICOS:  
META-ANÁLISE E APLICAÇÕES CLÍNICAS

JAYME M. STANDLEY (1986)

JOURNAL OF MUSIC THERAPY, XXIII (2), 1986, 105 - 133

### 2.2.1. Resumo do artigo / Abstract

#### Resumo

Realiza-se uma meta-análise avaliando 208 variáveis, consideradas a partir de 98 estudos em que a Música foi utilizada como contingência com objectivos educativos e terapêuticos. Através de procedimentos estatísticos controlaram-se efeitos do tamanho das amostras, comparando-se então as variáveis de forma a determinar efeitos da utilização da música. Os benefícios atribuíveis à contingência musical cifram-se em mais três Desvios-Padrão ( $DP = 2.90$ ) do que o verificado nas situações de controle que permitiram delinear a linha-base. Os resultados demonstram que a Música é mais eficaz como contingência do que outros estímulos não musicais. Demonstram também que a utilização de Música é mais eficaz em situação contingencial do que em situação contínua. O emparelhamento de outros estímulos (como comida, aprovação ou estimulação visual) com a Música diminuiu a eficácia desta como reforço. Excepcionalmente, a Música funcionou como um reforço e simultaneamente como matéria em si mesma, sendo evidente o benefício para outros comportamentos académicos e sociais que foram medidos mas não reforçados contingencialmente com a Música.

#### Abstract

This meta-analysis evaluated 208 variables derived from 98 studies incorporating music as a contingency for education and therapeutic objectives. Variables were converted to effect sizes (ES) via statistical formulae and compared for determination of characteristics affecting reinforcement value of music. The overall benefits attributable to contingent music were almost three standard deviations greater than control/baseline conditions ( $ES = 2.90$ ). Results demonstrate that contingent music was more effective than contingent nonmusic stimuli used in these studies and more effective than continuous music. Pairing other stimuli such as food, approval, or visual stimulation with the music decreased its effectiveness as a reinforcer. Uniquely, music functioned as a reinforcer and simultaneously as subject matter, and there was evidence of generalized benefit to other academic and social behaviors that were measured but were not contingently reinforced with music.



### 2.2.2. Artigo Original

# A Meta-Analysis on the Effects of Music as Reinforcement for Education / Therapy Objectives

JAYNE M. STANDLEY (1986)

As early as 1955, Jeffrey reported success using contingent music to effect behavior change in children. He suggested more research into such use of music since, in his opinion, it had several advantages over other commonly used types of reinforcement. Music was considered preferable to candy, which satiated quickly and was deemed nutritionally unsound; it was also preferable to toys, which he felt were disruptive to the teaching process (Jeffrey, 1955). It was many years before a case study appeared in the literature that demonstrated the successful clinical application of a music contingency (Madsen & Madsen, 1968). In this report, an out-of-control juvenile delinquent male received guitar lessons for reducing hostility toward his mother and for completing his assigned household chores.

Music initiation, participation, or interruption as a contingency for behavior change is now an established education/therapy technique having a vast body of accumulated research. Meta-analysis is a statistical procedure that has become an accepted method for synthesizing such an extant body of literature, describing its characteristics and providing a basis for informed generalization about the results. Specifically, meta-analysis is the application of a variety of formulae to the results of a body of homogeneous research to compute effect sizes – quantitative summaries of the properties and findings of individual studies. Effect sizes can then be compared and contrasted across multiple variables (Glass, McGaw, & Smith, 1984) and these overall results, to some extent, generalized.

Meta-analysis has been criticized by some researchers and statisticians on the basis that it combines data from different studies conducted under varying conditions and that it promulgates an inherent bias through analysis of published articles reflecting screening via the referee process. Experts in the technique have argued that these criticisms are unfounded, since data are not directly pooled but converted via consistent formulae to effect sizes that are comparable (Mann, 1994) and since a meticulous meta-analysis uses all studies located on a specific topic, not just referred articles (Glass, McGaw, & Smith, 1984). It is the intent and purpose of the meta-analysis to synthesize results of individual studies and all effects, whether significant or not, to identify patterns or trends in a carefully defined area of research. Since an individual study can seldom definitively answer a research question, once a large body of research exists on a specified topic, then some method for synthesizing and interpreting these results seems needed. Despite detractors who prefer statistical analysis of individual studies, meta-analysis is currently being used extensively to effect policy in medicine and the social sciences, where it was endorsed by the National Research Council in 1992 (Mann, 1994).

Meta-analysis has become an accepted research technique in other fields, but there is little evidence that music researchers have used the technique. A recent electronic search of the computerized index PsycINFO (1984 to 1994) revealed 1,805 meta-analyses on a variety of topics in psychology, education, and sociology but included only one music meta-analysis, a 1986 work by Standley. Reasons for the dearth of this type of music research are unknown. Perhaps the field is so young that music researchers do not yet feel confident in the amount of research in specific areas as useful for this type of examination.

Although no meta-analyses on the use of contingent music in educational/therapy settings were located, there are meta-analyses on other contingencies for educational objectives. Wilkinson (1980) analyzed 14 process-product classroom studies and performed a meta-analysis on the



relationship between teacher praise and student achievement by converting correlation coefficients. Overall, she found a low correlation between praise and academic achievement ( $r = .076$ ). Correlations were higher for reading achievement ( $r = .107$ ) and for mathematics achievement ( $r = .178$ ) than for other academic areas. Negative correlations were found for academic behavior deemed to be of higher cognitive complexity (spelling, paragraph reading, and mathematics reasoning). Teacher praise was more highly correlated with academic achievement of people from lower socioeconomic backgrounds, whereas negative correlations were revealed for those from higher socioeconomic groups. Praise was also more effective in lower grades than in intermediate grades (Wilkinson, 1980).

Getsie, Langer, and Glass (1985) conducted a meta-analysis on the effects of feedback on discrimination learning in children by computing effect size (ES). Best performance resulted from reward plus punishment ( $ES = .24$ ), compared to reward alone ( $ES = -.01$ ) and punishment alone ( $ES = .20$ ). For bright, upper-class children in the fourth grade or higher, punishment alone was most effective. Overall, computed effects of feedback on learning were modest.

Another analysis synthesized results of 39 studies on all types of classroom reinforcement (Lysakowski & Walberg, 1980). It showed an overall effect size for social and tangible reinforcement to classroom learning equal to 1.17. This indicates that, in general, reinforced students had higher measures of classroom learning than did nonreinforced control students by a little more than one standard deviation. Additionally, there were no significant differences in effect by type of reinforcement categorized as follows: competence, being correct, social praise, contingent activities, tokens, check marks, tangibles, or edibles.

Tenenbaum and Goldring (1989) conducted a meta-analysis on 16 studies of motor skill learning that used what they termed enhanced instruction: cues, participation, reinforcement, and feedback/correctives. An overall mean effect size of .66 was found for studies in which investigators used enhanced instruction versus regular instruction methods. Mean effect sizes were somewhat better for females (.82) than for males (.58), and they were less for elementary students (.46) than for kindergarten (.68), junior high school (.59), and high school students (.81).

A meta-analysis on 36 studies concerning cognitive and behavior treatment of impulsivity in children diagnosed as hyperactive, learning disabled, or having attention deficit disorder, conduct disorder, or behavior disorder revealed that mean effect sizes varied greatly by outcome measure or dependent variable: response latency or measures of slowed behavior (1.73), direct observation or frequency measures (.83), response accuracy (.50), teacher ratings (.35), and parent ratings (-.10) (Baer & Nietzel, 1991). There was no discernible difference between effect sizes for studies in which subjects were trained in cognitive (.52) versus behavioral tasks (.50). Overall, treated groups improved approximately one-third to three-quarters of a standard deviation above control groups.

The only meta-analysis in the related areas of music education/therapy (Standley, 1986) included all uses of music in medical/dental treatment and was recently updated to include more contemporary references (Standley, 1996). The revised data pool from 92 studies and 233 dependent variables showed an overall effect size on primary variables due to music of 1.17. Live music ( $ES = 1.13$ ) was more effective than recorded music ( $ES = .86$ ). Studies using the subjects' preferred music showed an overall effect size of 1.40. Effects in medicine were differentiated by age only for infants, with their response ( $ES = .48$ ) being much lower than that of children and adolescents ( $ES = .95$ ) or that of adults ( $ES = .93$ ). Behavioral observation was the most conservative type of dependent measure ( $ES = .83$ ), with results close to those obtained through physiological measures ( $ES = .90$ ). Results were greatly differentiated by type of research design, with behavioral designs using subjects as their own control ( $ES = 1.28$ ) showing a greater effect than studies using an experimental/control group format ( $ES = .66$ ).

In summary, then, meta-analyses using ES as the statistical comparison conducted on the topic of reinforcement and learning have ranges in overall effect from .24 to 1.17. The only prior meta-analysis on music contained primarily noncontingent techniques and resulted in a similar level of effect ( $ES = 1.17$ ).

The purpose of this study was to quantify the diverse results of research on the effectiveness of contingent music in educational and therapy settings, to determine the magnitude of music's effect as a reinforcer, and to identify the circumstances that affect its reinforcement value.



## METHOD

### Procedure

Established practice in meta-analysis involves three primary steps (Getsie, Langer, & Glass, 1985). The first step is a complete literature search to locate all possible sources in the defined population of studies. It is important to include both published and unpublished sources in a meta-analysis to reduce possible biases in the analyzed data that might be created by referred screening of published results. Next, the relevant characteristics and results of the located articles are identified and categorized. The final step is to use a variety of statistical techniques to convert reported data to comparable effect size measures.

Studies qualified for inclusion in this meta-analysis if they (a) contained empirical data and were reported in English; (b) used music listening initiation, music interruption, or music performance as a contingency for behavior change; (c) used an educational or therapy objective as the dependent variable; and (d) reported results in a format amenable to replicated data analysis. The starting points for a comprehensive literature search were electronic searches of the following data bases: PsycINFO (1967 to 1994), PsycFIRST (1992 to January 1995), ERIC (1967 to 1994), Dissertation Abstracts (1861 to 1994), and MEDLINE (1985 to 1994). Additional articles were identified from the reference lists of articles obtained. The literature search located more than 100 studies on this topic, 98 of which met criteria for inclusion in the meta-analysis. These 98 studies yielded 208 measures of effects on dependent variables.

### Estimation of effect size

Each dependent variable measure reported in the selected studies was converted to an estimated effect size (ES) according to procedures outlined by Glass, McGaw, and Smith (1984). The ES represents that proportion of a standard deviation that quantifies the experimental effect on the two conditions; i.e., an ES of + 1.00 would indicate that the size of benefit due to the experimental condition was one standard deviation greater than that of the control condition. The effect sizes of experimental results in this analysis were estimated by contrasting the means of experimental/treatment conditions (Exp) and control/baseline conditions (Con) divided by the standard deviation (SD) of the control/baseline condition, as in the formula below

$$\frac{\bar{M}_{\text{Exp}} - \bar{M}_{\text{Con}}}{SD_{\text{Con}}} = (\text{Estimated Effect Size})$$

Data not reported in a format that included or allowed computation of means and SD were converted via a variety of published statistical formulae (Glass, McGaw, & Smith, 1984).

## RESULTS

Ninety-eight studies were included in this meta-analysis and are identified separately in the References section. Table 1 is a comprehensive outline and categorization of study variables grouped by type of education/therapeutic variable. Effect sizes were computed for 208 variables and ranged from  $ES = 35.44^1$  (contingent music effects compared to baseline measures of sitting posture of an adult with a mental disability) to  $ES = -7.05$  (effects of contingent music initiation compared to contingent music interruption on crying of a 10-year-old with a mental disability).

In Table 1, a negative ES value indicates that the second independent variable of the two comparisons was more effective than the first. Of the 208 effect sizes computed, there were 12

that resulted in negative values. Contingent music was less effective than candy (ES = -3.12), lights (ES = -1.60), the voice of infants' mothers (ES = -1.46), or the sound of a heartbeat (ES = -.25). Noncontingent music was less effective than baseline conditions for reducing disruptive school bus behavior (ES = -1.81) and for reducing crying of infants with colic (ES = -.23). Contingent music was less effective than baseline conditions for increasing treadmill duration of subjects with mental handicaps (ES = -.20) and for reducing self-stimulatory behaviors of subjects with profound mental disabilities (ES = -.04). Contingent music initiation was less effective than contingent music interruption for reducing crying and correcting poor posture with children with mental disabilities (ES = -7.05 and ES = -3.71, respectively). When contingent music was paired with tactile stimulation (ES = -1.22) and with music distortion (ES = -.27), it was less effective than baseline condition for subjects with mental disabilities.



**Table 1**  
**Variable Categorization and Mean Effect Size**

| KEY:                                |                                |
|-------------------------------------|--------------------------------|
| <b>Design:</b>                      | <b>Subject Characteristics</b> |
| ES = Effect Size                    | MD = Mental Disability         |
| Ss as Con = Subjects as Own Control | ED = Emotional Disability      |
| Exp = Experimental Group            | LD = Learning Disability       |
| Con = Control Group                 | BD = Behavior Disorder         |
|                                     | Hyper = Hyperactive            |
|                                     | CP = Cerbral Palsey            |
| <b>Music:</b>                       | <b>Other:</b>                  |
| CM = Contingent Music               | R = Reinforcer                 |
| M = Music                           | Elem. = Elementary             |
| C = Contingent                      |                                |

| EDUCATIONAL STUDIES                |           |                                    |                                                                       |                                                  |                               |
|------------------------------------|-----------|------------------------------------|-----------------------------------------------------------------------|--------------------------------------------------|-------------------------------|
| Author<br>(Date)                   | Design    | Subject<br>characteristic<br>/ age | Dependent<br>variables                                                | Independent<br>variables                         | ES                            |
| <b>ACADEMIC VARIABLES</b>          |           |                                    |                                                                       |                                                  |                               |
| Cotter &<br>Spradlin<br>(1971)     | Ss as Con | MH/ 8–16 yrs                       | Math                                                                  | CM                                               | 1.04                          |
| Dorow (1976)                       | Ss as Con | MH/ 20 yrs                         | Math<br>Music Listening                                               | CM via TV                                        | .84<br>7.37                   |
| Madsen<br>(1981)                   | Ss as Con | Normal/ 3rd gr                     | Math<br><br>Music information                                         | CM via TV<br>Books<br>Music vs Books             | 1.74<br>1.62<br>.55<br>2.47   |
| Madsen et al.<br>(1976)            | Exp/Con   | Normal/ 1st gr                     | Math                                                                  | CM via TV<br>Tutors<br>Music vs Tutors           | 5.60<br>3.88<br>11.98         |
| Madsen &<br>Forsythe<br>(1973)     | Exp/Con   | Normal/6th gr                      | Math                                                                  | CM via<br>Earphones                              | 1.16                          |
| Madsen et al.<br>(1975)            | Exp/Con   | Normal/ 5th gr                     | Math<br>Attentiveness<br>Generalized math<br>Generalized<br>attention | CM<br>CM                                         | 10.34<br>3.71<br>5.71<br>3.13 |
| McLaughlin &<br>Helm (1993)        | Ss as Con | Normal/ middle<br>school           | Math                                                                  | CM                                               | .25                           |
| Miller (1977)                      | Ss as Con | MH/ 9–14 yrs                       | Math                                                                  | CM-Preferred<br>CM-Non-<br>Preferred             | .51<br>.74                    |
| Miller et al.<br>(1974)            | Ss as Con | EH/ 10 – 13 yrs                    | Math                                                                  | CM/ Art<br>Nonpreferred<br>Activity              | 1.36<br>.54                   |
| Eisenstein<br>(1974)               | Ss as Con | Normal/ 7 – 8<br>yrs               | Reading in books<br>Reading from<br>flashcards                        | CM – Guitar<br>Lessons<br>CM – Guitar<br>Lessons | .76<br>1.14                   |
| Table continues on following pages |           |                                    |                                                                       |                                                  |                               |

| Author<br>(Date)                   | Design    | Subject<br>characteristic /<br>age | Dependent<br>variables                                                   | Independent<br>variables                                           | ES                                   |
|------------------------------------|-----------|------------------------------------|--------------------------------------------------------------------------|--------------------------------------------------------------------|--------------------------------------|
| <b>ACADEMIC VARIABLES</b>          |           |                                    |                                                                          |                                                                    |                                      |
| Gordon (1977)                      | Exp/ Con  | Normal/ 4th gr                     | Reading<br>Watkins Farnum<br>Test                                        | CM - Lessons                                                       | .76<br>.40                           |
| Greer et al.<br>(1975)             | Exp/ Con  | Normal/ 6th gr                     | Choice of<br>Reinforcer                                                  | Music vs. Bank<br>Music vs. Story<br>Music vs. Candy               | .03<br>.34<br>-3.12                  |
| Madsen &<br>Geringer<br>(1976)     | Exp/ Con  | Normal/ 5 yrs                      | Auditory<br>discrimination                                               | CM via TV<br>Playtime<br>Choice of R<br>Music vs. Playtime         | 1.20<br>1.22<br>1.14<br>0.2          |
| Jacobs (1976)                      | Ss as Con | Normal/ 3 yrs                      | Numbers                                                                  | CM                                                                 | 6.69                                 |
| Johnson &<br>Zinner                | Ss as Con | MH/ 12 yrs                         | Color<br>discrimination                                                  | CM<br>CM                                                           | .86<br>.86                           |
| Madsen et al.<br>(1988)            | Exp/ Con  | LD & BD/K & 5th<br>gr              | Academic Skills                                                          | CM for tutors                                                      | 1.55                                 |
| Steele (1971)                      | Ss as Con | Headstart/ 4-5 yrs                 | Reading<br>Participation                                                 | CM<br>CM                                                           | 7.30<br>2.19                         |
| Waldron<br>(1977)                  | Ss as Con | Developmental<br>Delay             | Telephone<br>numbers                                                     | CM / M cue<br>vs. M cue                                            | 1.03                                 |
| Greer et al.<br>(1971)             | Exp/Con   | Normal/ 6th gr                     | Vocal accuracy                                                           | Rock vs. Silence<br>ME series music<br>vs. Silence                 | .28<br>.05<br>.74                    |
| Madsen &<br>Madsen<br>(1972)       | Exp/Con   | Normal/ 6th gr                     | Atentiveness<br>Pitch acuity<br>Attending<br>Music choice - Soul         | CM<br>CM                                                           | .43<br>.60<br>.43                    |
| <b>SOCIAL VARIABLES</b>            |           |                                    |                                                                          |                                                                    |                                      |
| Bosco<br>(1974)                    | Ss as Con | Autistic/ 8 yrs                    | Social behavior                                                          | CM                                                                 | .86                                  |
| Gunter et al.<br>(1993)            | Ss as Con | Autistic/ 14 yrs                   | Stereotopy                                                               | CM<br>Continuous M<br>Cm vs.<br>Continuous M                       | 1.00<br>.77<br>3.30                  |
| Hanser (1972)                      | Ss as Con | EH/ Middle school                  | Social behavior                                                          | CM                                                                 | 4.40                                 |
| Hanser (1974)                      | Ss as Con | EH/ Elem.                          | Social behavior                                                          | CM                                                                 | 11.55                                |
| Hill et al.<br>(1989)              | Ss as Con | MH/ 17 yrs                         | Out-of-seat                                                              | CM<br>CM Rock<br>CM Jazz<br>CM Classical<br>CM Rap                 | 2.15<br>2.49<br>2.43<br>1.06<br>1.97 |
| Jellison et al.<br>(1984)          | Exp/ Con  | MH & Normal/<br>8-15 yrs           | Acceptance of<br>MH Ss<br>Positive Interactions<br>Positive Interactions | CM<br>CM vs. Lg group<br>(class)<br>CM vs. Lg group<br>(free time) | .25<br>2.96<br>.64                   |
| Madsen<br>(1982)                   | Ss as Con | Normal & BD/ 7th<br>& 8th gr       | On-task                                                                  | Approval<br>CM/ Approval                                           | 1.23<br>1.55                         |
| Madsen &<br>Wolfe                  | Exp/ Con  | College Ss                         | Body movement-<br>-reading<br>Body movement-<br>-listening               | CM<br>CM<br>Finger snapping<br>Negative Practice                   | .82<br>2.42<br>5.63<br>7.53          |
| Table continues on following pages |           |                                    |                                                                          |                                                                    |                                      |



| Author<br>(Date)                         | Design    | Subject<br>characteristic /<br>age | Dependent<br>variables                                                    | Independent<br>variables                        | ES                       |
|------------------------------------------|-----------|------------------------------------|---------------------------------------------------------------------------|-------------------------------------------------|--------------------------|
| <b>SOCIAL VARIABLES</b>                  |           |                                    |                                                                           |                                                 |                          |
| Wilson (1976)                            | SS as Con | MH/ 5-7 yrs                        | Disruptive behavior                                                       | CM                                              | 1.14                     |
| Wilson &<br>Hopkins<br>(1973)            | Ss as Con | Normal/ 7th & 8th<br>gr            | Home Ec Noise                                                             | CM                                              | 3.36                     |
| Wolfe (1982)                             | Exp/ Con  | Normal & Hyper/<br>3rd gr          | Body movement<br><br>Letters                                              | CM-Interruption<br>CM<br>CM-Interruption<br>CM  | .17<br>.04<br>.03<br>.10 |
| Yarbrough et<br>al. (1977)               | Exp/ Con  | Normal/ 5th gr                     | Attentiveness                                                             | CM                                              | .28                      |
| <b>MEDICAL STUDIES</b>                   |           |                                    |                                                                           |                                                 |                          |
| <b>PHYSICAL REHABILITATION VARIABLES</b> |           |                                    |                                                                           |                                                 |                          |
| Ball et al.<br>(1975)                    | Ss as Con | CP with MH/ 9-17<br>yrs            | Head Posture                                                              | CM                                              | 3.71                     |
| Corman<br>(1979)                         | Ss as Con | CP with MH/ 17-<br>24 yrs          | Head extension                                                            | CM                                              | 1.86                     |
| Grove et al.<br>(1975)                   | Ss as Con | CP with MH/ 10-<br>11 yrs          | Head extension                                                            | CM/ Approval                                    | 6.84                     |
| Kearney &<br>Fussy (1991)                | Ss as Con | Neurologic injury/<br>22 yrs       | Head Posture                                                              | CM                                              | 4.11                     |
| Macurik<br>(1979)                        | Ss as Con | CP with MH/<br>Elem.               | Head Posture                                                              | CM                                              | 2.02                     |
| Murphy et al.<br>(1979)                  | Ss as Con | CP with MH/ 14-<br>19 yrs          | Head Posture                                                              | CM                                              | 1.37                     |
| Wolfe (1980)                             | Ss as Con | CP/ 3-37 yrs                       | Head Posture                                                              | CM<br>CM/Tone<br>CM vs. CM/Tone                 | .34<br>.52<br>.10        |
| Gouvier et al.<br>(1985)                 | Ss as Con | Physical<br>Impairment/ 21<br>yrs  | Yelling during gait<br>training in PT                                     | CM/ TOK-BAK                                     | 4.19                     |
| Barrett (1962)                           | Ss as Con | Neurologic<br>impairment/ 38 yrs   | Uncontrollable tic                                                        | CM/ White noise                                 | 3.22                     |
| Holliday<br>(1987)                       | Ss as Con | CP/ 49 yrs                         | Increased physical<br>skill<br><br>Decreased<br>inappropriate<br>Behavior | CM<br><br>CM                                    | 6.51<br><br>.62          |
| <b>OTHER MEDICAL VARIABLES</b>           |           |                                    |                                                                           |                                                 |                          |
| Epstein et al.<br>(1974)                 | Ss as Con | Headaches/ 39<br>yrs               | EMG tension<br>Amt. of medication                                         | CM                                              | 2.48<br>2.04             |
| Boyle & Greer<br>(1983)                  | Ss as Con | Coma/ 31-56 yrs                    | Overt movement                                                            | CM                                              | 2.50                     |
| Buttersfield &<br>Siperstein<br>(1972)   | Ss as Con | Neonates/ 2 days                   | Sucking patterns                                                          | CM                                              | 2.00                     |
| DeCasper &<br>Carstens                   | Ss as Con | Neonates/ 3 days                   | Sucking patterns<br><br>Activity level                                    | CM<br>vs Continuous M<br>CM vs. Continuous<br>M | 1.84<br><br>1.70         |
| Table continues on following pages       |           |                                    |                                                                           |                                                 |                          |

| Author<br>(Date)                        | Design    | Subject<br>characteristic /<br>age | Dependent<br>variables           | Independent<br>variables                                                     | ES                                          |
|-----------------------------------------|-----------|------------------------------------|----------------------------------|------------------------------------------------------------------------------|---------------------------------------------|
| <b>OTHER MEDICAL VARIABLES</b>          |           |                                    |                                  |                                                                              |                                             |
| Etscheidt<br>(1989)                     | SS as Con | Infants with colic/<br>4-7 wks     | Cryings<br><br>Alertness         | CM<br>C Attention<br>CM/ C Attention<br>CM<br>C Attention<br>CM/ C attention | 1.72<br>2.54<br>3.36<br>.16<br>1.13<br>1.72 |
| Larson &<br>Ayllon (1990)               | Ss as Con | Infants with colic/<br>3-7 wks     | Crying                           | CM<br><br>Continuous M<br>CM vs. Continuous M                                | 4.95<br><br>-.23<br>4.48                    |
| <b>SPORT STUDIES</b>                    |           |                                    |                                  |                                                                              |                                             |
| Caouette &<br>Reid (1985)               | Exp/ Con  | MH/ 22-30 yrs                      | Stationary<br>bike duration      | CM<br>CM/ Lights<br>CM vs. Lights                                            | .72<br>1.32<br>-1.60                        |
| Caouette &<br>Reid                      | Ss as Con | MH/ 25 yrs                         | Stationary<br>bike duration      | CM-Preferred<br>CM vs. White Noise<br>CM vs. Pink Noise                      | .70<br>.13<br>.91                           |
| Ellis et al.<br>(1993)                  | Ss as Con | MH/ 18-19 yrs                      | Treadmill duration<br>Heart rate | CM<br>CM                                                                     | -.20<br>.22                                 |
| Hume &<br>Crossman<br>(1992)            | Ss as Con | Swimmers/ 12-16<br>yrs             | Practive behavior                | CM                                                                           | 2.26                                        |
| Silliman &<br>French                    | Exp/ Con  | MH/ 10-17 yrs                      | Soccer Kick                      | CM<br>C approval<br>CM vs. C approval                                        | 3.45<br>1.50<br>1.52                        |
| <b>DEVELOPMENTAL STUDIES</b>            |           |                                    |                                  |                                                                              |                                             |
| <b>ATTENTION/ INTERACTION VARIABLES</b> |           |                                    |                                  |                                                                              |                                             |
| Allen & Bryant<br>(1985)                | Ss as Con | MH/ 10-11 yrs                      | Crying                           | CM<br>Cm Interrupted<br>CM vs.<br>CM Interrupted                             | 4.81<br>7.78<br>-7.05                       |
|                                         |           |                                    | Sitting up                       | CM<br>CM Interrupted<br>CM vs. CM<br>Interrupted                             | .48<br>4.12<br>-3.71                        |
| Becker (1983)                           | Ss as Con | MH/ 34 yrs                         | Eye contact                      | CM<br>CM Distortion<br>CM vs.<br>CM/Distortion                               | 2.60<br>2.50<br>.12                         |
| Burch et al.<br>(1987)                  | Ss as Con | MH/ 31 yrs                         | Sitting posture                  | CM                                                                           | 35.44*                                      |
| Deutsch &<br>Parks (1978)               | Ss as Con | MH/ 14 yrs                         | Inappropriate<br>behavior        | CM                                                                           | 6.61                                        |
| Dewson &<br>Whiteley<br>(1987)          | Ss as Con | MH/ 11-29 yrs                      | Head turn                        | CM/ Slides/<br>Vibration                                                     | .60                                         |
| Dorow (1975)                            | Ss as Con | MH/ 9-15 yrs                       | Follow directions                | CM<br>CM/ Food<br>CM vs. CM/ Food                                            | 1.14<br>1.52<br>.42                         |
| Table continues on following pages      |           |                                    |                                  |                                                                              |                                             |



| DEVELOPMENTAL STUDIES            |           |                                    |                                                  |                                                                                                                                                     |                                                  |
|----------------------------------|-----------|------------------------------------|--------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------|
| ATTENTION/ INTERACTION VARIABLES |           |                                    |                                                  |                                                                                                                                                     |                                                  |
| Author<br>(Date)                 | Design    | Subject<br>characteristic /<br>age | Dependent<br>variables                           | Independent<br>variables                                                                                                                            | ES                                               |
| Dorow (1980)                     | Ss as Con | MH/ 15 yrs                         | Follow directions                                | CM/ Approval<br>CM/ Approval/ Food<br>C Approval<br>CM/ Approval-<br>Maintenance<br>C/ Approval/ Food-<br>Maintenance<br>C Approval-<br>Maintenance | 5.16<br>5.18<br>4.00<br><br>5.24<br>5.29<br>5.29 |
| Falb (1982)                      | Ss as Con | MH/ 6 mo.-3 yrs                    | Vasoconstriction                                 | CM<br>C heartbeat<br>CM vs. C Heartbeat                                                                                                             | 3.75<br>4.61<br>-.25                             |
| Finkelstein &<br>Ramey (1977)    | Ss as Con | Infants/ 8 mo.                     | Pull toy                                         | CM vs.<br>Continuous M                                                                                                                              | 2.69                                             |
| Holloway<br>(1980)               | Ss as Con | MH/ 30 yrs                         | Follow directions                                | CM<br>CM-Playing Instr<br>CM vs. CM-Playing                                                                                                         | 1.68<br>1.96<br>.28                              |
| Konarski<br>(1987)               | Ss as Con | MH/ 21-61 yrs                      | Pull lever                                       | CM vs. Slides<br>CM vs. Slides w/o<br>response<br>deprivation<br>CM/ Response dep.<br>vs. Slides                                                    | .96<br><br>1.34<br>1.58                          |
| Metzler (1974)                   | Exp/ Con  | MH/ 16-55 yrs                      | Imitative behavior                               | CM<br>Continuous M<br>Continuous M                                                                                                                  | 1.27<br>.09<br>.09                               |
| Ramos (1993)                     | Exp/ Con  | Abuse hot line<br>callers          | Sustain telephone<br>on-hold on for<br>counselor | Jazz vs. Relaxation M                                                                                                                               | .99                                              |
| Remington et<br>al. (1977)       | Ss as Con | MH/ 12.8 yrs                       | Press bar                                        | Blues M vs. Playing<br>drums                                                                                                                        | .62                                              |
| Standford &<br>Tustin (1974)     | Ss as Con | Abusive parent of<br>13-mo.-old    | Tolerate infant<br>crying                        | CM                                                                                                                                                  | 33.10*                                           |
| Saperston et<br>al. (1980)       | Exp/ Con  | MH/ 28 yrs                         | Reach for object                                 | CM vs. C juice                                                                                                                                      | .12                                              |
| Talkington &<br>Hall (1970)      | Exp/ Con  | MH/ Elem.                          | Echoic responses                                 | CM-Preferred<br>CM-Least preferred<br>CM-Most vs. Least<br>preferred                                                                                | 1.75<br>.77<br>1.26                              |
| Underhill &<br>Harris (1974)     | Ss as Con | EH/ 9-13 yrs                       | Imitation                                        | CM<br>Continuous M<br>CM vs. Continuous M                                                                                                           | 2.76<br>.01<br>2.28                              |
| Vollmer &<br>Iwata (1991)        | Ss as Con | MH/ 25-36 yrs                      | Follow directions                                | CM vs. Food<br>CM vs. Social Praise                                                                                                                 | 2.28<br>2.60                                     |
| SELF-HELP VARIABLES              |           |                                    |                                                  |                                                                                                                                                     |                                                  |
| Luiselli (1991)                  | Ss as Con | MH/ 10 yrs                         | Self-feeding                                     | CM/ Bright Light                                                                                                                                    | 2.00                                             |
| Johnson &<br>Babbitt (1993)      | Ss as Con | MH/ 4 yrs                          | Eating solid food                                | CM                                                                                                                                                  | 2.80                                             |
| Garwood<br>(1988)                | Ss as Con | MH/ 21 yrs                         | Enuresis                                         | CM/ Bell pad                                                                                                                                        | .10                                              |

Table continues on following pages

| DEVELOPMENTAL STUDIES      |           |                                    |                                 |                          |             |
|----------------------------|-----------|------------------------------------|---------------------------------|--------------------------|-------------|
| TRANSPORTATION VARIABLES   |           |                                    |                                 |                          |             |
| Author<br>(Date)           | Design    | Subject<br>characteristic /<br>age | Dependent<br>variables          | Independent<br>variables | ES          |
| Barmann et al. (1980)      | Ss as Con | MH/ 8 yrs                          | Disruptive school bus behavior  | CM                       | 17.60       |
|                            |           |                                    |                                 | CM vs. Continuous M      | 12.70       |
| McCarty et al. (1978)      | Ss as Con | EH/ 3-14 yrs                       | Disruptive school bus behavior  | CM                       | 1.57        |
|                            |           |                                    |                                 | Continuous M             | -1.81       |
| Reid et al. (1975)         | Ss as Con | Autistic/ 8-10                     | Disruptive car behavior         | CM<br>CM/ Ritalin        | .94<br>1.21 |
| Ritschl et al. (1972)      | Ss as Con | MH/ 7-15 yrs                       | Disruptive school bus behavior  | CM                       | 1.54        |
| STEREOTOPY VARIABLES       |           |                                    |                                 |                          |             |
| Borreson & Anderson (1982) | Ss as Con | MH/ 21 yrs                         | Rumination                      | CM                       | 2.16        |
|                            |           |                                    |                                 | CM/ Tactile stimulation  | -1.22       |
| Carrocio et al. (1976)     | Ss as Con | Schizophrenic/ 40 yrs              | Tics                            | C Guitar rental          | .81         |
| Davis et al. (1983)        | Ss as Con | MH/ 26 yrs                         | Rumination                      | CM-10"                   | 1.60        |
|                            |           |                                    |                                 | CM-30"                   | 1.89        |
|                            |           |                                    |                                 | CM/ Reprimand            | 2.33        |
|                            |           |                                    |                                 | Reprimand                | .99         |
|                            |           |                                    |                                 | CM vs. Reprimand         | 1.57        |
| CM-30" vs. CM-20"          | .60       |                                    |                                 |                          |             |
| Greenwald (1978)           | Ss as Con | MH/ 7-22 yrs                       | Self-stimulatory behavior       | Out-of-seat behavior     | 1.64        |
|                            |           |                                    |                                 | CM                       | -.04        |
| Hauck & Martin (1970)      | Ss as Con | Schizophrenic/ 5-9                 | Self-stimulatory behavior       | CM/DIstortion            | -.27        |
|                            |           |                                    |                                 | CM vs. CM/Distortion     | .20         |
|                            |           |                                    |                                 | CM vs. Continuous M      | .34         |
| Jorgenson (1971)           | Ss as Con | MH/ Elem.                          | Self-stimulatory behavior       | CM                       | 1.95        |
| Jorgenson (1974)           | Ss as Con | MH/ 9 yrs                          | Stereotopy<br>Follow directions | CM                       | 2.33        |
|                            |           |                                    |                                 | CM                       | 1.11        |
| McClure et al. (1986)      | Ss as Con | MH/ 9 yrs                          | Hand-to-mouth stereotopy        | CM                       | 7.07        |
| Ritchey (1977)             | Exp/ Con  | MH/ 17-41 yrs                      | Self-stimulatory behaviors      | CM                       | 1.04        |
| Ross (1974)                | Ss as Con | MH/ 13-22 yrs                      | Stereotopy                      | CM                       | 2.54        |
|                            |           |                                    |                                 | CM-Activity              | 3.55        |
|                            |           |                                    |                                 | C Play                   | 6.10        |
| Scruggs (1991)             | Ss as Con | Geriatric                          | Wandering                       | CM<br>M Activity         | .08<br>.29  |
| COMPLAINING VARIABLES      |           |                                    |                                 |                          |             |
| Cook & Freethy (1973)      | Ss as Con | Schizophrenic/ 43 yrs              | Complaining                     | CM                       | 1.14        |
|                            |           |                                    |                                 | CM/ Cue                  | 1.32        |
|                            |           |                                    |                                 | CM vs. CM/ Cue           | 2.90        |
| Williams & Dorow (1983)    | Ss as Con | Depression/ 79 yrs                 | Complaining                     | CM                       | 2.54        |
|                            |           |                                    |                                 | CM/ Reprimand            | 2.02        |
|                            |           |                                    |                                 | CM vs.<br>CM/ Reprimand  | .56         |
|                            |           | Table continues on following pages |                                 |                          |             |

Table continues on following pages



| DEVELOPMENTAL STUDIES          |           |                                                                                                                                |                             |                                           |                      |
|--------------------------------|-----------|--------------------------------------------------------------------------------------------------------------------------------|-----------------------------|-------------------------------------------|----------------------|
| WORK VARIABLES                 |           |                                                                                                                                |                             |                                           |                      |
| Author<br>(Date)               | Design    | Subject<br>characteristic /<br>age                                                                                             | Dependent<br>variables      | Independent<br>variables                  | ES                   |
| Bellamy &<br>Sontag            | Ss as Con | MH/ 11-12 yrs                                                                                                                  | Assembly line<br>production | CM – 20"                                  | 1.98                 |
|                                |           |                                                                                                                                |                             | CM – 30"                                  | .93                  |
|                                |           |                                                                                                                                |                             | CM – 20 vs.<br>CM – 30"                   | 1.64                 |
|                                |           |                                                                                                                                |                             | CM – 20 vs.<br>Continuous M               | 1.57                 |
|                                |           |                                                                                                                                |                             | CM – 30" vs.<br>Continuous M              | .43                  |
| Clegg (1982)                   | Ss as Con | MH/ 22-55 yrs                                                                                                                  | Work production             | CM<br>Continuous M<br>CM vs. Continuous M | 2.40<br>7.60<br>2.88 |
| Cotter (1971)                  | Exp/ Con  | MH/ 12-21 yrs                                                                                                                  | Work production             | CM vs. Continuous M                       | 1.50                 |
| Podvin (1967)                  | Ss as Con | MH/ Adults                                                                                                                     | Work production             | CM vs. Continuous M                       | 1.06                 |
| PREFERENCE VARIABLES           |           |                                                                                                                                |                             |                                           |                      |
| Greer et al.<br>(1974)         | Ss as Con | Normal/ K-6th gr                                                                                                               | Bar press                   | CM-Rock vs. CM-<br>Nonrock                | 4.50                 |
| Standley &<br>Madsen<br>(1990) | Ss as Con | Infants                                                                                                                        | Listening<br>preference     | CM vs. Mother's voice                     | -1.46                |
|                                |           |                                                                                                                                |                             | CM vs. Female Voice                       | 1.87                 |
|                                |           | <b>Note.</b> All effect size (ES) comparisons with baseline condition<br>unless otherwise stated.<br>* See Discussion section. |                             |                                           |                      |

Table 2  
Publication Analyses of Includes Studies (N=98)

*Publication Year*

|       |       |       |       |
|-------|-------|-------|-------|
| 1960s | 1970s | 1980s | 1990s |
| 2     | 53    | 29    | 14    |

*Publication Source*

|                |                     |                                        |                  |                         |                      |
|----------------|---------------------|----------------------------------------|------------------|-------------------------|----------------------|
| Music journals | Psychology journals | Mental retardation/ Education Journals | Medical Journals | Speech therapy Journals | Unpublished articles |
| 45             | 20                  | 16                                     | 2                | 1                       | 14                   |

*Reciprocal Citations of Included Studies By Journal Type*

|       |            |                              |         |                |
|-------|------------|------------------------------|---------|----------------|
| Music | Psychology | Mental retardation/Education | Medical | Speech Therapy |
| 249   | 18         | 52                           | 2       | 1              |

Table 2 shows the publication source and date of studies included in the meta-analysis. The bulk of the studies were published in the 1970s, with a substantial reduction in number published in the 1980s and 1990s, although the rate across the last 15 years appears to be stable. The source of publication data reveals that music researchers have had a major impact on the development of this topic by having contributed almost half the articles qualified for inclusion. Additionally, this table shows the results of an analysis of the references used by the included studies to determine whether they referenced others' prior work. The music journal authors frequently referenced the prior studies, with 249 citations. Authors of articles published in the mental retardation/education journals also frequently cited these studies ( $n = 48$ ). Authors publishing in psychology journals cited these studies only 18 times, although 20 of the studies in the meta-analysis were published in psychology journals. It seems apparent that educators and those publishing in the field of mental retardation research music as a contingency and quote a large number of prior studies. Psychology journal authors research this topic but do not often quote existing studies. From a historical perspective, it is interesting to note the five articles most referenced by the other meta-analysis studies: Barrett, 1962 ( $n = 22$ ); Madsen and Forsythe, 1973 ( $n = 20$ ); Steele, 1971 ( $n = 14$ ); Cotter, 1971 ( $n = 13$ ); and Greer, Randall, and Timberlake, 1971 ( $n = 12$ ).

Table 3 shows effect sizes for the variety of independent variables as compared to baseline conditions; also shown are effect sizes for contingent music variables as compared to other independent variables.



**Table 3**  
**Mean Effect Size for Independent Variable Comparisons (N = 208)**

*Experimental Conditions versus Baseline*

| Contingent music | Noncontingent Music | Contingent nonmusic stimuli | Contingent music paired with nonmusic stimuli | Generalized dependent variables | Music variations |
|------------------|---------------------|-----------------------------|-----------------------------------------------|---------------------------------|------------------|
| 2.90<br>n = 105  | 2.23<br>n = 9       | 2.37<br>n = 15              | 2.60<br>n = 22                                | 4.15<br>n = 2                   | 1.99<br>n = 4    |

*Contingent Music Conditions versus Other Conditions*

| Continuous music | Other music    | Nonmusic stimuli |
|------------------|----------------|------------------|
| 2.32<br>n = 13   | 0.24<br>n = 16 | 1.14<br>n = 22   |

The overall effect of contingent music conditions versus baseline conditions was 2.90. This effect is much larger than the 1.17 effect size previously established for social and tangible reinforcement to classroom learning (Lysakowski & Walberg, 1980). In this study, the benefits of noncontingent music over baseline conditions were found to be  $ES = 2.23$ . The prior medical meta-analysis on therapeutic uses of music that were primarily noncontingent revealed an overall effect size of 1.17 (Standley, 1996). Contingent music versus continuous music resulted in an  $ES = 2.32$ , and contingent music versus nonmusic stimuli as contingencies resulted in an  $ES = 1.14$ . In these comparisons, other contingent stimuli included progressive illumination of a string of lights, white noise, heartbeat, bank accounts, stories, candy, projected slides, books, tutor interactions, playtime, juice or food, praise or social approval, and the voices of infants' mothers.

Pairing other stimuli with contingent music ( $ES = 2.60$ ) slightly reduced its effectiveness. Contingent music paired with nonaversive stimuli such as food, approval, or visual stimulation ( $ES = 2.63$ ) was more effective than when cessation of contingent music was followed by aversive stimuli such as white noise, music distortion, or loud tones ( $ES = 1.84$ ).

There were two measures of generalized effect on variables not consequated (contingently reinforced) by the treatment condition ( $ES = 4.15$ ). In an educational setting, consequating an academic behavior positively affected social behavior and vice versa.

The 105 measures of the effects of contingent music as the sole reinforcement versus baseline conditions were further analyzed. Results are shown in Tables 4-10. Effects by intent of the music contingency are analyzed in Table 4. Procedures intended to increase behavior ( $ES = 2.97$ ) were slightly more effective than those intended to decrease behavior ( $ES = 2.77$ ).

**Table 4**  
**Mean Effect Size (ES) by Intent of Music Contingency (N = 105)**

|    | Contingency to increase behavior | Contingency to decrease behavior |
|----|----------------------------------|----------------------------------|
| ES | 2.97                             | 2.77                             |
| n  | 69                               | 36                               |

Table 5 shows differentiated effects by the techniques used to establish the music contingency: initiating music following the desired response ( $ES = 2.55$ ) versus interrupting ongoing music following an undesirable response ( $ES = 3.56$ ). It is apparent that music interruption is a more effective contingent music procedure by about one standard deviation.

**Table 5**  
*Mean Effect Size by Music Contingency Technique (N = 105)*

|          | Music<br>Initiation | Music<br>interruption |
|----------|---------------------|-----------------------|
| ES       | 2.55                | 3.56                  |
| <i>n</i> | 68                  | 37                    |

Timing of the contingency application is analyzed in Table 6. Immediate implementation of the contingency ( $ES = 3.38$ ) showed greater effect than delayed procedures: (a) those in which tokens, tickets, or points issued for desirable behavior could be exchanged later for music, and (b) those that required sustained correct responding for an extended period before participation in music activities ( $ES = 1.70$ ). As expected, delay of receipt of contingent music greatly reduced its effects on therapeutic/educational variables.

Table 7 shows analysis by type of educational or therapeutic dependent variable. Effects were greatest for physical rehabilitation ( $ES = 5.47$ ) and for developmental or daily living variables ( $3.40$ ). Music contingencies were somewhat less effective for social ( $ES = 2.04$ ) and academic variables ( $ES = 2.18$ ) measured in school settings, though these effect sizes are much greater than those found in prior meta-analyses of nonmusic reinforcement in the school setting. Music as a contingency was least effective in sports applications ( $ES = 1.39$ ). In this study, contingent music benefits in medicine ( $ES = 2.26$ ) were much greater than noncontingent music medical effects ( $ES = 1.17$ ) noted previously by Standley (1996).

**Table 6**  
*Mean Effect Size by Timing of Music Contingency (N = 105)*

|          | Immediate | Delayed |
|----------|-----------|---------|
| ES       | 3.38      | 1.70    |
| <i>n</i> | 75        | 30      |



Table 7  
Mean Effect Size by Educational / Therapeutic Objective (N = 105)

|    | EDUCATIONAL |          | DEVELOPMENTAL | MEDICAL | PHYSICAL       | SPORTS   |
|----|-------------|----------|---------------|---------|----------------|----------|
|    | Social      | Academic | Daily living  | Health  | Rehabilitation | Exercise |
| ES | 2.04        | 2.18     | 3.40          | 2.26    | 5.47           | 1.39     |
| n  | 20          | 24       | 38            | 7       | 11             | 5        |

Table 8  
Mean Effect Size by Age (N = 105)

|    | Infants - 4 yrs | 5 - 11 yrs | 12-14 yrs | 15-18 yrs | Adult | Elderly |
|----|-----------------|------------|-----------|-----------|-------|---------|
| ES | 3.51            | 2.53       | 1.96      | 1.08      | 4.51  | 1.31    |
| n  | 9               | 48         | 16        | 5         | 25    | 2       |

Table 9  
Mean Effect Size by Subject Characteristic (N = 105)

|    | Normal | Mentally<br>impaired | Emotionally<br>impaired | Medically/<br>physically<br>impaired |
|----|--------|----------------------|-------------------------|--------------------------------------|
| ES | 2.99   | 3.16                 | 2.38                    | 2.25                                 |
| n  | 32     | 48                   | 11                      | 14                                   |

Analyses by age (see Table 8) revealed decreasing effects from infancy (ES = 3.51) through each level of the school years to high school (ES = 1.08). Effect sizes then dramatically increased for uses with adults (ES = 4.51). Least effective were contingent uses of music with the elderly, specifically several studies incorporating Alzheimer's patients (ES = 1.31), although a benefit of greater than one standard deviation for music contingencies on problems of Alzheimer's patients would seem to be a clinically important result.

Table 9 shows effect size for subjects by ability to produce the desired response. The greatest effects were with subjects with mental disabilities (ES = 3.16) and subjects without disabling conditions (ES = 2.99), with little differentiation among other groups, i.e., those with emotional disabilities (ES = 2.38) or those with medical/physical impairments (ES = 2.25). It is interesting that contingent music effects were so large with people with mental disabilities, since the preponderance of these studies used subjects who were categorized as profoundly mentally disabled or multiply disabled.

Table 10 shows that effects varied greatly by type of design, with those studies using subjects as their own control (ES = 3.42) having much greater effect than those involving experimental/control group designs (1.62). A meta-analysis of applied single-subject research involving differential reinforcement (Prochnow-Lagrow, 1984) found a similarly high overall effect size (3.61), although this may be somewhat inflated since the authors modified the analysis procedure by omitting the first three data points for each baseline and treatment condition in order to avoid pollution of treatment results across conditions. In this meta-analysis, all baseline and treatment data points were analyzed. One reason for different effect sizes by design type may be that the studies using subjects as their own control were often single-subject or small-sample

studies with contingencies carefully tailored to individual preferences and reinforcement schedules. Such highly individualized techniques are a common aspect of behavioral procedures and are designed to enhance effectiveness.

Table 10  
*Effect Size by Design Type (N = 105)*

|          | Subjects as<br>own control | Experimental/<br>Control groups |
|----------|----------------------------|---------------------------------|
| ES       | 3.42                       | 1.62                            |
| <i>n</i> | 75                         | 30                              |

## DISCUSSION

The most conservative, descriptive metric of effect size was used in this study because of the variety of behavioral designs included in which independence between averaged data could not be assumed. Two results are indicated by asterisks in Table 1 and noted as exorbitantly high effect sizes. These reflect the floor-to-ceiling changes from almost no desired behavior under baseline conditions to immediate and high levels of the desired behavior. With only a single subject in each study, a control condition with extremely low variance results, and this produces a very large effect size when divided into the tremendous gain in behavior due to the experimental condition. As a statistical measure of standard deviation, these numbers are impossible large. As points on the scale of the common metric used for comparison in this study, they accurately reflect the unusual benefits of the music contingency established. The investigator determined that these data should be included, since the numbers carry no inherent value in themselves except within the context of the specific study from which they evolved.

One could argue that no single-subject designs should be included in an analysis of this type on the basis that very small *N*s create bias in the results. It seems to the investigator that that argument negates the validity and impact of single-subject research, which is central to behavioral treatment methodology. Although techniques are still evolving, the research literature does demonstrate precedence in including single-subject research in meta-analyses. To rule out the possibility of inflation due to size of *N*, however, a second analysis of all 208 variables was undertaken by the investigator using Hedges correction factor *K* (Glass, McGaw, & Smith, 1991, p. 113), a procedure for adjusting ES by the size of the smallest group sample in each study. In the great majority of data points, results were affected by only hundredths of a point, and this adjustment to the 208 variables did not substantially change the overall meta-analysis results.

In summary, these analyses demonstrate that contingent music is more effective than other contingent nonmusic stimuli and is more effective than continuous music, and that pairing other stimuli with contingent music decreases its effectiveness. Music is highly effective as a contingency for either increasing desirable behavior or reducing undesirable behavior, with slightly better results in increasing behavior. Music interruption is more effective than music initiation as the procedure for establishing the contingency, and immediate initiation of the music is more effective than delayed initiation in exchange for tokens or points. In this study, music contingencies were more effective in physical rehabilitation and with developmental behaviors than with medical or educational behaviors. Least effective were sports applications.

Adults and infants, compared with other age-groups, responded best to music contingencies, with effects declining from childhood through the primary and secondary-school years. Least effects were demonstrated for elderly patients, although effects at the levels demonstrated in these studies would be considered clinically significant. Individuals with mental disabilities and normal subjects demonstrated greater benefits than did those with emotional disabilities or medical/physical impairments. The majority of studies in this analysis used subjects



as their own control and demonstrated much greater effect than did studies involving experimental and control groups.

The effects due to contingent music revealed in this study certainly compare favorably to other types of reinforcement assessed through meta-analyses. Perhaps most interesting in this analysis was the widespread and creative use of contingent music across educational and therapeutic objectives. For example, in medicine, contingent music reinforced overt behaviors of comatose patients, decreased crying of infants with colic, improved neck strength and head posture of individuals with neurologic impairments, and reduced headache pain through biofeedback.

The most frequent uses of contingent music have occurred in the field of education. In regular education classrooms, music as subject matter or as pleasurable listening has reinforced other academic achievement, particularly math skills and reading. Contingent music has also reduced noise in Home Economics sewing classes, increased acceptance of children with disabilities in their mainstreamed classes, and improved aural discrimination, vocal pitch acuity, and music information and listening skills. In special education, contingent music has been used to increase self-feeding skill, decrease stereotypical behavior or disruptive behavior, decrease enuresis and rumination (habitual regurgitation), decrease out-of-seat behavior on the school bus, and increase classroom social behaviors such as following directions, imitating, and eye contact.

In physical education and sports, music was played over pool loudspeakers as a group contingency to increase on-task practice time for competitive swimmers, and music reinforcement via tape recorder was used on the soccer field to improve kick accuracy and in the workout room to increase duration of stationary bike riding.

In counseling sessions, one unique study from New Zealand helped an abusive father learn to tolerate the prolonged crying of his infant daughter, a behavior which had previously stimulated physical assault. Contingent music listening following a changing criterion tolerance for the taped crying was paired with videotapes of the child at play and during happy, contented moments (Sandford & Tustin, 1974).

In predicting the effects of music contingencies for future applications, it is interesting to note that several meta-analyses selected for comparison with this study showed decreased reinforcement effects across school grade levels (Getsie, Langer, & Glass, 1985; Wilkinson, 1980), as did this study. Perhaps this is due to preference changes as children develop (LeBlanc, 1981). However, from analysis of these contingent music studies, it seems apparent that such decreases may be more related to the educational values held by the teachers and therapists involved in the studies, namely, that more mature subjects should be expected to delay gratification and benefit from increasingly complex methods of reinforcement. In this study, delay of reinforcement greatly lessened effect, and the studies with older children seemed to incorporate more procedures that provided choices and delayed reinforcement.

A primary criticism of reinforcement for education and therapeutic endeavors has been the contention that reinforcement reduces intrinsic motivation, i.e., the self-engaging participation and pleasure in the activity itself. Cameron and Pierce (1994) conducted a meta-analysis on the relationship between reinforcement, reward, and intrinsic motivation. They concluded that the meta-analysis "results suggest that in the laboratory, overall reward does not negatively impact intrinsic motivation on any of the four measures analyzed" and that "it no longer seems appropriate to argue against the use of incentive systems in applied settings" (p. 394). In the current study, measures of music as subject matter increased simultaneously with music functioning as reinforcement for other academic subjects. Additionally, academic and social behaviors not contingently reinforced showed generalized benefits while other studies demonstrated students' frequent selection of music from a varied menu of possible reinforcers. This meta-analysis demonstrated that music is a powerful contingency, and results showed no evidence of negative effects on either students' academic performance or motivation.

This meta-analysis has shown that demonstrated effects of contingent music are profound, can be creatively designed, and are applicable across a wide variety of educational/therapeutic endeavors. Based on these studies, predicted effects in medicine and counseling seem particularly strong and, if the frequency of research is an indicator, much underutilized in clinical settings. Educational research on contingent music, however, is much more comprehensive in its documentation of uses and procedures. "Music for its own sake" certainly ought to be the core of music education programs; however, more use and creative application of such an effective

technique as music reinforcement would seem warranted, especially when it addresses contemporary political desire to improve the basic curriculum (music as reinforcement for other subject matter such as math and reading) while concomitantly improving music knowledge and skills. Clearly, effects are so beneficial across diverse fields of endeavor that knowledge of contingent music techniques would seem to be an important component of the professional skills of all music educators and music therapists.

Those seeking results of research often are frustrated by myriad separate and seemingly incomparable analyses and by the researchers' reluctance to generalize their results. Meta-analysis is a technique for global analysis and synthesis of a carefully defined topic and set of variables specifically for the purpose of generalization of results across multiple studies. Such analyses seem crucial to the professions of music education and therapy as we seek to increase our knowledge of effective techniques from general to specific applications and to affect educational and social policy. Music researchers have not yet incorporated meta-analysis procedures in large numbers, unlike investigators in a number of other fields. It is an additional research method that offers much to enhance our existing methods of inquiry, and pursuit of more research in this area would seem warranted.



## REFERENCES\*\*

- \* References marked with an asterisk indicate studies included in the meta-analysis.
- \* Allen, L. D., & Bryant, M. C. (1985). A multielement analysis of contingent versus contingent interrupted music. *Applied Research in Mental Retardation* 6 (1), 87-97.
- \* Ball, T. S., McCrady, R., & Hart, A. (1975). Automated reinforcement of head posture in two cerebral palsied retarded children. *Perceptual and Motor Skills*, 40, 619-622.
- Baer, R. A., & Nietzel, M. T. (1991). Cognitive and behavioral treatment of impulsivity in children: A meta-analytic review of the outcome literature. *Journal of Clinical Child Psychology*, 20 (4), 400-412.
- \* Barmann, B. C., Croyle-Barmann, C., & McLain, B. (1980). The use of contingent-interrupted music in the treatment of disruptive bus-riding behavior. *Journal of Applied Behavior Analysis*, 13, 693-698.
- \* Barrett, B. H. (1962). Reduction in rate of multiple tics by free operant conditioning methods. *The Journal of Nervous and Mental Disease*, 135 (3), 187-195.
- \* Becker, I. (1983). Control of acquisition of eye contact by distorted and undistorted music stimuli. *Journal of Music Therapy*, 20 (3), 132-142.
- \* Bellamy, T., & Sontag, E. (1973). Use of group contingent music to increase assembly line production rates of retarded students in a simulated sheltered workshop. *Journal of Music Therapy*, 10, 125-136.
- \* Borreson, P. M., & Anderson, J. L. (1982). The elimination of chronic rumination through a combination of procedures. *Mental Retardation*, 20 (1), 34-38.
- \* Bosco, B. K. (1974). *The use of radio control for the contingent application of music with a child exhibiting autistic behaviors: An exploratory study*. Unpublished master's thesis, Florida State University.
- \* Boyle, M. E., & Greer, R. D. (1983). Operant procedures and the comatose patient. *Journal of Applied Behavior Analysis*, 16, 3-12.
- \* Burch, M. R., Clegg, J. C., & Bailey, J. S. (1987). Automated contingent reinforcement of correct posture. *Research in Developmental Disabilities*, 8 (1), 15-20.
- \* Butterfield, E. C., & Siperstein, G. N. (1972). Influence of contingent auditory stimulation upon non-nutritional suckle. In J. F. Bosma (Ed.), *Third Symposium on Oral Sensation and Perception: The Mouth of the Infant* (pp. 313-333). Springfield, IL: Charles C Thomas.
- Cameron, J., & Pierce, W. D. (1994). Reinforcement, reward, and intrinsic motivation: A meta-analysis. *Review of Educational Research*, 64 (3), 363-423.
- \* Caouette, M., & Reid, G. (1985). Increasing the work output of severely retarded adults on a bicycle ergometer. *Education and Training of the Mentally Retarded*, 20 (4), 296-304.
- \* Caouette, M., & Reid, G. (1991). Influence of auditory stimulation on the physical work output of adults who are severely retarded. *Education and Training in Mental Retardation*, 26 (1), 43-52.

- \* Carroccio, D. F., Latham, S., & Carroccio, B. B. (1976). Rate-contingent guitar rental to decelerate stereotyped head/face-touching of an adult male psychiatric patient. *Behavior Therapy*, 7 (1), 104-109.
- \* Clegg, J. C. (1982). *The effect of non-contingent and contingent music on work production rate of mentally retarded adults in a work activity center*. Unpublished master's thesis, Florida State University.
- \* Cook, M., & Freethy, M. (1973). The use of music as a positive reinforcer to eliminate complaining behavior. *Journal of Music Therapy*, 10, 213-216.
- \* Corman, S. E. (1979). *Increasing head extension with the use of automated contingent music*. Unpublished doctoral dissertation, Western Michigan University.
- \* Cotter, V. W. (1971). Effects of music on performance of manual tasks with retarded adolescent females. *American Journal of Mental Deficiency*, 76 (2), 242-248.
- \* Cotter, V. W., & Spradlin, J. E. (1971). Effects of contingent music on retarded children's performance of addition computations. (Report No. 3). Parsons, KS: Parsons Research Center.
- \* Davis, W. B., Wieseler, N. A., & Hanzel, T. E. (1983). Reduction of rumination and out-of-seat behavior and generalization of treatment effects using a nonintrusive method. *Journal of Music Therapy*, 20 (3), 115-131.
- \* DeCasper, A. J., & Carstens, A. A. (1981). Contingencies of stimulation: Effects on learning and emotion in neonates. *Infant Behavior and Development*, 4, 19-35.
- \* Deutsch, M., & Parks, A. L. (1978). The use of contingent music to increase appropriate conversational speech. *Mental Retardation*, 16 (1), 33-36.
- \* Dewson, M. R., & Whiteley, J. H. (1987). Sensory reinforcement of head turning with nonambulatory, profoundly mentally retarded persons. *Research in Developmental Disabilities*, 8 (3), 413-426.
- \* Dorow, L. G. (1975). Conditioning music and approval as new reinforcers for imitative behavior with the severely retarded. *Journal of Music Therapy*, 12 (1), 30-39.
- \* Dorow, L. G. (1976). Televised music lessons as educational reinforcement for correct mathematical responses with the educable mentally retarded. *Journal of Music Therapy*, 13 (2), 77-86.
- \* Dorow, L. G. (1980). Generalization effects of newly conditioned reinforcers. *Journal of Music Therapy*, 15 (1), 8-14.
- \* Eisenstein, S. R. (1974). Effect of contingent guitar lessons on reading behavior. *Journal of Music Therapy*, 11, 138-146.
- \* Ellis, D. N., Cress, P. J., & Spellman, C. R. (1993). Training students with mental retardation to self-pace while exercising. *Adapted Physical Activity Quarterly*, 10 (20), 104-124.
- \* Epstein, L., Hersen, M., & Hemphill, D. (1974). Music feedback in the treatment of tension headache: An experimental case study. *Journal of Behavior Therapy and Experimental Psychiatry*, 5 (1), 59-63.
- \* Etscheidt, M. A. (1989). *Parent training to reduce the excessive crying associated with infant colic*. Unpublished doctoral dissertation, Georgia State University.



- \* Falb, M. E. (1982). *The use of operant procedures to condition vasoconstriction in profoundly mentally retarded (PMR) infants*. Unpublished master's thesis, Florida State University.
- \* Finkelstein, N. W., & Ramey, C. (1977). Learning to control the environment in infancy. *Child Development*, 48, 806-819.
- \* Garwood, E. C. (1988). The effect of contingent music in combination with a bell pad on enuresis of a mentally retarded adult. *Journal of Music Therapy*, 25 (2), 103-109.
- Getsie, R. L., Langer, P., & Glass, G. (1985). Meta-analysis of the effects of type and combination of feedback on children's discrimination learning. *Review of Educational Research*, 55 (1), 9-22.
- Glass, G., McGaw, B., & Smith, M. (1984). *Meta-analysis in social research*. Beverly Hills, CA: Sage Publications.
- \* Gordon, M. V. (1977). *The effect of contingent instrumental music instruction on the language reading behavior and musical performance ability of middle school students*. Unpublished doctoral dissertation, Columbia University Teachers College.
- \* Gouvier, W. D., Richards, J. S., Blanton, P. D., Janert, K., Rosen, L. A., & Drabman, R. S. (1985). Behavior modification in physical therapy. *Archives of Physical Medicine and Rehabilitation*, 66 (2), 113-116.
- \* Greenwald, M. A. (1978). The effectiveness of distorted music versus interrupted music to decrease self-stimulatory behaviors in profoundly retarded adolescents. *Journal of Music Therapy*, 15 (2), 58-66.
- \* Greer, R. D., Dorow, L., & Harrison, L. (1975). Aural discrimination instruction and the preferences of sixth-graders for music listening, story listening, and candy. In C. K. Madsen, R. D. Greer, & C. H. Madsen, Jr. (Eds.), *Research in music behavior* (pp. 97-108). New York: Teachers College, Columbia University Press.
- \* Greer, R. D., Dorow, L. & Randall, A. (1974). Music listening preferences of elementary school children. *Journal of Research in Music Education*, 22, 284-291.
- \* Greer, R. D., Randall, A., & Timberlake, C. (1971). The discriminate use of music listening as a contingency for improvement in vocal pitch acuity and attending behavior. *Bulletin of the Council for Research in Music Education*, no. 26, 10-18.
- \* Grove, D. N., Dalke, B. A., Fredericks, H. D., & Crowley, R. F. (1975). Establishing appropriate head positioning with mentally and physically handicapped children. *Behavioral Engineering*, 3 (2), 53-59.
- \* Gunter, P. L., Fox, J. J., McEvoy, M. A., Shores, R. E., & Denny, R. K. (1993). A case study of the reduction of aberrant, repetitive responses of an adolescent with autism. *Education and Treatment of Children*, 16 (2), 187-197.
- \* Hanser, S. B. (1972). *The effect of contingent music upon the behavior of a group of junior high school students*. Unpublished master's thesis, Florida State University.
- \* Hanser, S. B. (1974). Group-contingent music listening with emotionally disturbed boys. *Journal of Music Therapy*, 11 (4), 220-225.
- \* Hauck, L. P., & Martin, P. L. (1970). Music as a reinforcer in patient-controlled duration of time out. *Journal of Music Therapy*, 7, 43-53.

- \* Hill, J., Brantner, J., & Spreat, S. (1989). The effect of contingent music on the in-seat behavior of a blind young woman with profound mental retardation. *Education and Treatment of Children*, 12 (2), 165-173.
  - \* Holliday, A. M. (1987). *Music therapy and physical therapy to habilitate physical disabilities of young children*. Unpublished master's thesis, Florida State University.
  - \* Holloway, M. S. (1980). A comparison of passive and active music reinforcement to increase preacademic and motor skills in severely retarded children and adolescents. *Journal of Music Therapy*, 17 (2), 58-69.
  - \* Hume, K. M., & Crossman, J. (1992). Musical reinforcement of practice behaviors among competitive swimmers. *Journal of Applied Behavior Analysis*, 25, 665-670.
  - \* Jacobs, G. M. (1976). *The effect of music, social-tactile reinforcement, and verbal punishment on the visual discrimination learning of children of low income families*. Unpublished master's thesis, Florida State University.
- Jeffrey, W. E. (1955). New techniques for motivating and reinforcing children. *Science*, 121, 371.
- \* Jellison, J. A., Brooks, B., & Huck, A. (1984). Structuring small groups and music reinforcement to facilitate positive interactions and acceptance of severely handicapped students in the regular music classroom. *Journal of Research in Music Education*, 32, 243-264.
  - \* Johnson, C. R., & Babbitt, R. L. (1993). Antecedent manipulation in the treatment of primary solid food refusal. *Behavior Modification*, 17 (4), 510-521.
  - \* Johnson, J. M., & Zinner, C. C. (1974). Stimulus fading and schedule learning in generalizing and maintaining behaviors. *Journal of Music Therapy*, 11 (2), 84-96.
  - \* Jorgenson, H. (1971). Effects of contingent preferred music in reducing two stereotyped behaviors of a profoundly retarded child. *Journal of Music Therapy*, 8 (4), 139-145.
  - \* Jorgenson, H. (1974). The use of a contingent music activity to modify behaviors which interfere with learning. *Journal of Music Therapy*, 11, 41-46.
  - \* Kearney, S., & Fussey, I. (1991). The use of adapted leisure materials to reinforce correct head positioning in a brain-injured adult. *Brain Injury*, 5 (3), 295-302.
  - \* Konarski, E., Jr. (1987). Effects of response deprivation on the instrumental performance of mentally retarded persons. *American Journal of Mental Deficiency*, 91 (5), 537-542.
  - \* Larson, K., & Ayllon, T. (1990). The effects of contingent music and differential reinforcement on infantile colic. *Behaviour Research and Therapy*, 28 (2), 119-125.
- LeBlanc, A. (1981). Effects of style, tempo, and performing medium on children's music preference. *Journal of Research in Music Education*, 29, 143-156.
- \* Luiselli, J. K. (1991). Acquisition of self-feeding in a child with Lowe's syndrome. *Journal of Developmental and Physical Disabilities*, 3 (2), 181-189.
- Lysakowski, R. S., & Walberg, H. J. (1980). Classroom reinforcement. *Evaluation in Education: An International Review Series*, 4, 115-116.
- \* Macurik, K. M. (1979). An operant device to reinforce correct head position. *Journal of Behavior Therapy and Experimental Psychiatry*, 10, 237-239.



- \* Madsen, C. K. (1981). Music lessons and books as reinforcement alternatives for an academic task. *Journal of Research in Music Education*, 29, 103-110.
- \* Madsen, C. K. (1982). The effect of contingent teacher approval and withholding music performance on improving attentiveness [Special issue]. *Psychology of Music*, 76-81.
- \* Madsen, C. K., Dorow, L., Moore, R., & Womble, J. (1976). Effects of music via television as reinforcement for correct mathematics. *Journal of Research in Music Education*, 24, 51-59.
- \* Madsen, C. K., & Forsythe, J. L. (1973). Effect of contingent music listening on increases of mathematical responses. *Journal of Research in Music Education*, 21, 176-181.
- \* Madsen, C. K., & Geringer, J. (1976). Choice of televised music lessons versus free play in relationship to academic improvement. *Journal of Music Therapy*, 13 (4), 154-162.
- Madsen, C. K., & Madsen, C. H., Jr. (1968). Music as a behavior modification technique with a juvenile delinquent. *Journal of Music Therapy*, 5 (3), 72-76.
- \* Madsen, C. K., & Madsen, C. H., Jr. (1972). Selection of music listening or candy as a function of contingent versus noncontingent reinforcement and scale singing. *Journal of Music Therapy*, 9, 190-198.
- \* Madsen, C. K., Moore, R., Wagner, M., & Yarbrough, C. (1975). A comparison of music as reinforcement for correct mathematical responses versus music as reinforcement for attentiveness. *Journal of Music Therapy*, 12 (2), 84-95.
- \* Madsen, C. K., Smith, D. S., & Feeman, C. C., Jr. (1988). The use of music in cross-age tutoring within special education settings. *Journal of Music Therapy*, 25 (3), 135-144.
- \* Madsen, C. K., & Wolfe, D. E. (1979). The effect of interrupted music and incompatible responses on bodily movement and music attentiveness. *Journal of Music Therapy*, 16 (1), 17-30.
- Mann, C. C. (1994). Can meta-analysis make policy? *Science*, 266 (11), 960-962.
- \* McCarty, B., McElfresh, C., Rice, S., & Wilson, S. (1978). The effect of contingent background music on inappropriate bus behavior. *Journal of Music Therapy*, 15 (3), 150-156.
- \* McClure, J. T., Moss, R. A., McPeters, J. W., & Kirkpatrick, M. A. (1986). Reduction of hand mouthing by a boy with profound mental retardation. *Mental Retardation*, 24 (4), 219-222.
- \* McLaughlin, R. F., & Helm, J. L. (1993). Use of contingent music to increase academic performance of middle-school students. *Psychological Reports*, 72, 658.
- \* Metzler, R. K. (1974). The use of music as a reinforcer to increase imitative behavior in severely and profoundly retarded female residents. *Journal of Music Therapy*, 11 (2), 97-110.
- \* Miller, D. (1977). Effects of music listening contingencies on arithmetic performance and music preference of EMR children. *American Journal of Mental Deficiency*, 81 (4), 371-378.
- \* Miller, D. M., Dorow, L., & Greer, R. D. (1974). The contingent use of music and art for improving arithmetic scores. *Journal of Music Therapy*, 11, 57-64.
- \* Murphy, R., Doughty, N., & Nunes, D. (1979). Multielement designs: An alternative to reversal and multiple baseline evaluation strategies. *Mental Retardation*, 17 (1), 23-27.
- \* Podvin, M. G. (1967). The influence of music on the performance of a work task. *Journal of Music Therapy*, 4 (2), 52-56.

Prochnow-Lagrow, J. E. (1984). *Meta-analysis of applied single subject research utilizing differential reinforcement of behavior omission*. Unpublished doctoral dissertation, Northern Illinois University.

- \* Ramos, L. (1993). The effects of on-hold telephone music on the number of premature disconnections to a statewide protective services abuse hot line. *Journal of Music Therapy*, 30 (2), 119-129.
- \* Reid, D. H., Hill, B., Rawers, R., & Montegar, C. (1975). The use of contingent music in teaching social skills to a nonverbal, hyperactive boy. *Journal of Music Therapy*, 12 (1), 2-18.
- \* Remington, R. E., Foxen, T., & Hogg, J. (1977). Auditory reinforcement in profoundly retarded multiply handicapped children. *American Journal of Mental Deficiency*, 82 (3), 299-304.
- \* Ritchey, N. A., Jr. (1977). *A pleasant alternative to a punishment procedure using contingent loud noise to modify self-destructive and self-stimulative behaviors in the non-ambulatory, profoundly and severely retarded*. Unpublished master's thesis, Florida State University.
- \* Ritschl, C., Mongrella, J., & Presbie, R. J. (1972). Group time-out from rock and roll music and out-of-seat behavior of handicapped children while riding a school bus. *Psychological Reports*, 31, 967-973.
- \* Ross, L. E. (1974). *The effects of three treatments involving music on the stereotyped behaviors of profoundly retarded adults*. Unpublished master's thesis, Florida State University.
- \* Sandford, D. A., & Tustin, R. D. (1974). Behavioural treatment of parental assault on a child. *New Zealand Psychologist*, 2, 76-82.
- \* Saperston, B. M., Chan, R., Morphew, C. & Carsrud, K. (1980). Music listening versus juice as a reinforcement for learning in profoundly mentally retarded individuals. *Journal of Music Therapy*, 17 (4), 174-183.
- \* Scruggs, S. D. (1991). *The effects of structured music activities versus contingent music listening with verbal prompt on wandering behavior and cognition in geriatric patients with Alzheimer's disease*. Unpublished master's thesis, Florida State University.
- \* Silliman, L. M., & French, R. (1993). Use of selected reinforcers to improve the ball kicking of youths with profound mental retardation. *Adapted Physical Activity Quarterly*, 10 (1), 52-69.
- Standley, J. M. (1986). Music research in medical/dental treatment: Meta-analysis and clinical applications. *Journal of Music Therapy*, 23 (2), 56-122.
- Standley, J. M. (1996). Music research in medical/dental treatment: An update of a prior meta analysis. In C. E. Furman (Ed.), *Effectiveness of music therapy procedures: Documentation of research and clinical practice* (pp. 1-60). Silver Spring, MD: National Association for Music Therapy.
- \* Standley, J. M., & Madsen, C. K. (1990). Comparison of infant preferences and responses to auditory stimuli: Music, mother, and other female voice. *Journal of Music Therapy*, 27, 54-97.
- \* Steele, A. L. (1971). Contingent socio-music listening periods in a preschool setting. *Journal of Music Therapy*, 8, 131-139.
- \* Talkington, L. W., & Hall, S. M. (1970). A musical application of Premack's hypothesis to low verbal retardates. *Journal of Music Therapy*, 7 (3), 95-99.



- Tenenbaum, G., & Goldring, E. (1989). A meta-analysis of the effect of enhanced instruction: Cues, participation, reinforcement and feedback and correctives on motor skill learning. *Journal of Research and Development in Education*, 22 (3), 53-64.
- \* Underhill, K. K., & Harris, L. M. (1974). The effect of contingent music on establishing imitation in behaviorally disturbed retarded children. *Journal of Music Therapy*, 11, 156-166.
- \* Vollmer, T. R., & Iwata, B. (1991). Establishing operations and reinforcement effects. *Journal of Applied Behavior Analysis*, 24, 279-291.
- \* Waldron, G. M. (1977). *The effect of music as a structural prompt and as a reinforcer on learning a sequential task*. Unpublished master's thesis, University of Minnesota.
- Wilkinson, S. S. (1980). *The relationship of teacher praise and student achievement: A meta analysis of selected research*. Unpublished doctoral dissertation. University of Florida.
- \* Williams, G., & Dorow, L. G. (1983). Changes in complaints and non-complaints of a chronically depressed psychiatric patient as a function of an interrupted music/verbal feedback package. *Journal of Music Therapy*, 20 (3), 143-155.
- \* Wilson, C. V. (1976). The use of rock music as a reward in behavior therapy with children. *Journal of Music Therapy*, 13 (1), 39-48.
- \* Wilson, C. W., & Hopkins, B. (1973). The effects of contingent music on the intensity of noise in junior high home economics classes. *Journal of Applied Behavior Analysis*, 6, 269-275.
- \* Wolfe, D. E. (1980). The effect of automated interrupted music on head posturing of cerebral palsied individuals. *Journal of Music Therapy*, 17 (4), 184-206.
- \* Wolfe, D. E. (1982). The effect of interrupted and continuous music on bodily movement and task performance of third grade students. *Journal of Music Therapy*, 19 (2), 74-85.
- \* Yarbrough, C., Charboneau, M., & Wapnick, J. (1977). Music as reinforcement for correct math and attending in ability assigned math classes. *Journal of Music Therapy*, 14 (2), 77-88.

---

\*\* As referências bibliográficas deste artigo estão conforme o documento original.

## 2.2.3. Estudo Crítico – Por Orlindo Gouveia

### Finalidade e problemas

O artigo examina a aplicação da terapia pela música em Medicina, procedendo a uma revisão da literatura e a uma meta-análise dos estudos empíricos em que a música é utilizada em tratamentos médico-odontológicos. Depois, usa os resultados tendo em vista aplicações clínicas da música a contextos terapêuticos e programas hospitalares.

O que imediatamente se pode questionar é o que a própria autora (e os autores dos estudos examinados) entendem por *música*. Nem sequer há uma mera referência a género musical ou a outra categoria desse tipo. Os artigos usados na meta-análise também não esclarecem a questão, ficando-se por uma mera oposição entre, de um lado, música e não-música e, do outro, ruídos e vozes.

No final não há propriamente uma conclusão, mas antes um breve sumário das análises estatísticas. Além dele, apenas se faz notar que a investigação nesta área ocorreu principalmente nos últimos 5 anos (isto é, desde 1980), mediante a realização de estudos laboratoriais e clínicos, considerando-se categorias diagnósticas altamente diversificadas.

### Método

O interesse principal do estudo reside no uso da *meta-análise*. Como o próprio nome indica, a *meta-análise* é uma técnica estatística que se usa após terem sido feitas análises primárias e secundárias de dados directamente colhidos em estudos específicos, com a finalidade de determinar o que pode considerar-se comum nesse conjunto de estudos.

No presente caso, a autora seleccionou trinta artigos que mediram efeitos da *música* no decurso de tratamentos médicos e dentários reais.

A autora seguiu Getsic, Langer e Glass (1985), que indicam os três primeiros passos para se efectuar uma meta-análise. Primeiro, proceder a uma criteriosa revisão da literatura – incluindo trabalhos publicados ou não, por exemplo, teses – susceptíveis de serem incluídos, de acordo com a finalidade em vista. Segundo, determinar as características e resultados dos estudos seleccionados, identificando-os, descrevendo-os e categorizando-os. Terceiro, utilizar as técnicas estatísticas apropriadas.

No presente caso e seguindo-se Glass, McGaw e Smith (1984) foi usada a estimativa da dimensão do efeito (*estimative of effect size* (ES)).

ES representa a proporção do desvio padrão (standard deviation [SD]) que quantifica o efeito experimental em duas condições. Calcula-se usando a seguinte fórmula:

$$ES = \frac{\bar{X}_{\text{esp}} - \bar{X}_{\text{con}}}{SD_{\text{con}}}$$

em que

$\bar{X}_{\text{esp}}$  = média da condição ou grupo experimental

$\bar{X}_{\text{con}}$  = média da condição ou grupo de controlo

SD con = desvio padrão da condição ou grupo de controlo

(ES = + 1,00 indica que o grupo experimental pontuou 1 SD melhor que o grupo de controlo)



Uma vez que os estudos originais que se usam em meta-análise não são feitos a pensar nela, mas sim na sua lógica própria de tentar resolver um problema específico (usando, por isso, técnicas estatísticas variadas), pode acontecer que alguns deles não apresentem ou apresentem de forma incompleta os parâmetros média e desvio padrão.

Glass, McGaw e Smith (1984), entre outros, indicam soluções apropriadas para ultrapassar estas dificuldades. Deve, contudo, ter-se em conta que estas manipulações ulteriores aumentam a incerteza dos resultados desse modo obtidos. Aliás, deve ter-se sempre em conta que a regra fundamental da aplicação do cálculo estatístico se pode exprimir através do tradicional dito português: “não misturar alhos com bugalhos!”.

## **Revisão geral da literatura**

A revisão da literatura efectuada por J. Standley excede os trinta artigos que acabam por ser relacionados para uma meta-análise, o que tem o interesse de começar por equacionar o problema em termos gerais e históricos.

Segundo a autora, a mais antiga associação entre música e cuidados médicos surge no papiro Kahum. Podíamos acrescentar o de David tocando lira para o merencório Rei Saúl, ou as práticas mais sistemáticas descritas por Galeno, o médico do Imperador Marco Aurélio, e grande sistematizador da *Ars Medica*. Contudo, só para o final do século XIX, se começam a estudar os efeitos da música nos processos fisiológicos: Dogiel, em França, em 1810; Corning, nos E.U.A., em 1889; e Tarchanoff, na Rússia, em 1903. De acordo com as possibilidades da época, tratou-se de avaliar clinicamente a frequência respiratória do pulso e a tensão arterial. Estudos laboratoriais, ainda incipientes, foram conduzidos por Hyde (1924) e Vincent e Thompson (1929). Nesta vertente, a introdução de novas técnicas médicas de diagnóstico permitiu maior precisão nas medições. Wagner (1975) usou a electroencefalografia e Scartelli (1984) a electromiografia. Acrescenta-se que actualmente dispomos de novas tecnologias de imagem, com a tomografia axial computadorizada (TAC) e a ressonância electromagnética funcional (REMF), que nos permitem estudar a reacção dos diversos sectores do sistema nervoso no momento em que está a ser estimulado e que, certamente, irão revolucionar os estudos laboratoriais e, mesmo, os clínicos.

A introdução da música no armentário médico no hospital geral é creditada a Evan O’Neil Kane, em 1981. O uso da música no contexto de tratamentos odontológicos reais é anterior (Cherry e Pallin, 1948). Nos estudos que se seguiram, a música pode ser ambiente ou via auscultadores estereofónicos.

(A autora refere ainda várias outras aplicações da música em diversas situações médicas desde a sala de operações à sala de partos).

## **Revisão específica da literatura**

A revisão da literatura destinada à presente meta-análise, partiu de fontes diversas: pesquisa da autora, exame sistemático de publicações específicas (*Music Therapy* e *Journal of Music Therapy*), teses de doutoramento (*Dissertation Abstracts*), e bibliografias ou índices publicados por outros.

Foram identificados oitenta e um estudos passíveis de análise, mas só trinta acabaram por ser retidos. Nos outros cinquenta e um estudos faltava pelo menos um dos seguintes dados: inclusão de dados empíricos; utilização de situações de simulação (em vez de situações reais); uso de estímulos auditivos não musicais; por fim, surgiram ainda casos em que a referência aos resultados não era passível de aplicação a esta meta-análise.

## Categorização

Os trinta estudos seleccionados foram categorizados.

A categorização incluiu: identificação (quatro itens); tipo de problemas do doente; extensão da amostra; idade; género; número por grupo; desenho experimental; identificação das variáveis independentes e dependentes; presença ou ausência de dor.

As variáveis independentes incluíam: música, “não música”, ruídos ambiente, ruído branco, vozes humanas e bio-feedback.

As variáveis dependentes incluíam: auto-relatos (percepção da dor, de ansiedade, de eficácia, etc.); variáveis fisiológicas (pulso, tensão arterial, tensão muscular); variáveis comportamentais (gritar, gesticular, mover-se, etc.); considerando-se ainda o uso simultâneo de analgésicos e medicação.

Nas tabelas apresentadas, as cinquenta e cinco variáveis (!) dependentes estão muito menos categorizadas do que o que aqui se apresenta. Isto coloca alguns problemas interpretativos de equivalência ou não de situações designadas pela mesma categoria.

## Resultados

O resultado global revela a superioridade dos tratamentos em que se usa a música relativamente àqueles em que ela não é usada, dado que a média da estimativa da dimensão do efeito é  $ES = 0,98$ .

Em 54 das 55 variáveis dependentes, a música favoreceu o objectivo médico em  $ES = (0,97)$  para variáveis fisiológicas, em  $ES = (0,85)$  para variáveis psicológicas ou auto-relato, e em  $ES = (1,10)$  para variáveis comportamentais.

O maior efeito verificou-se nos pacientes odontológicos ( $ES = 2,26$ ), seguido dos cardíacos ( $ES = 1,14$ ) e cirúrgicos ( $ES = 1,13$ ). O efeito menor verificou-se em recém-nascidos ( $0,81$ ), pacientes oncológicos ( $0,55$ ) e em casos obstétricos ( $0,65$ ).

A combinação da música com a analgesia parece bastante positiva pois leva a um aumento de  $ES = 1,40$ .

## Discussão

O núcleo do presente artigo é a avaliação do efeito terapêutico adjuvante da *música* em diversas situações odontológicas e médico-cirúrgicas, através do uso da meta-análise dos resultados de trinta artigos seleccionados da literatura.

A meta-análise é um procedimento estatístico que permite discernir tendências gerais – sentido e magnitude – num campo de pesquisa alargado. Assim, os estudos empíricos singulares, que a originaram, podem ser encarados num contexto mais vasto. Além disso, permite discernir, com mais precisão, quais as direcções de investigação que vale a pena ou não prosseguir e, inclusivé, superar deficiências em futuros estudos singulares.

O problema principal do artigo em apreço, e que é comum a muitos outros, tem a ver com as variáveis independentes. A autora, como outros autores, parece não se aperceber da tremenda imprecisão inerente ao uso de uma variável designada *música*, *tout cour*, ou, de modo análogo, de uma condição designada *não-música*.

Sem entrar em discussões de carácter mais profundo fora do âmbito deste trabalho, a definição clássica de música é a de “uma sequência de sons que produz uma impressão harmoniosa”. Por oposição, fala-se de ruído quando os sons não produzem tal impressão e de ruído branco quando a sequência é totalmente aleatória.

As duas primeiras definições, de música e de ruído, são psicológicas, no sentido que é a impressão harmoniosa que decide. A terceira definição, de ruído branco, é física ou comunicativa (no ruído branco não se discernem sinais e, por isso, não há mensagem, embora possa haver meta-mensagem).



Parece evidente que quando se quer usar música como estímulo, quer no laboratório quer como adjuvante terapêutico, se tem de especificar de que música se trata. Não se sabe, porque não se estudou, mas creio que todos aceitarão que uma composição *heavy metal* e *Eine Kleine Nachtmusik*, em execução ou reprodução, devem provocar efeitos fisiopsicológicos muito diversos. Hoje, queremos saber mais. Queremos saber quais as diferenças na reacção neural. Nos extremos, sabemos alguma coisa. Sons musicais perto do limiar perceptivo inferior, mobilizam a atenção e criam um efeito de activação. Perto do limiar superior, para além dos 40.000 db provocam um efeito semelhante ao das drogas opiáceas. Daí, talvez, a população que acorre aos concertos *rock*!

Tudo isto implica que a definição mais precisa do *estímulo musical* tem de considerar, primeiro, os parâmetros físicos da sequência de sons, logo, os aspectos orgânicos prováveis, não podendo prescindir da dimensão psicológica (emoção e percepção da "impressão harmoniosa"), que leva à adesão ou repulsa pelo que é tocado.

Este último aspecto é particularmente sensível quando se pretende fazer terapia pela música. Podem estabelecer-se regras de tipo genérico mas se o paciente não gostar do que é tocado, será pouco provável a sua adesão à terapia. Isto é, tudo o que a investigação científica nos possa dizer, tem de ser caldeado pela relação terapêutica e a escolha final do estímulo musical personalizada. Quando se procede a tal personalização deverá, também, ser tido em conta a cultura musical (e outras) do sujeito.

As necessidades da investigação são umas, mas não necessariamente as mesmas que devem presidir ao estabelecimento de um plano terapêutico. Isto torna ainda mais complexa a tarefa da *investigação da terapêutica pela música*, pois será então necessário atender a duas lógicas. É impossível, à partida, esperar que duas sequências psico-fisicamente equivalentes produzam o mesmo resultado em sujeitos diferentes (ou, no mesmo, em diversas circunstâncias) e, inversamente, que o mesmo resultado não possa ser alcançado por sequências muito diversas. Só a experimentação rigorosa e equipada com os novos métodos de avaliar a resposta neurológica poderá dar resposta a estas interrogações.

A questão não está tanto em usar um grande número de variáveis dependentes. No estudo em apreço, a autora acabou por se defrontar com cinquenta e cinco, um número excessivo, mas que reflecte a falta de uma matriz comum aos investigadores nesta área. O agrupamento destas variáveis fisiológicas, psicológicas e comportamentais, mostrou diferenças insignificantes entre elas. Isto quer dizer que além do esforço classificativo é necessário no futuro proceder a uma hierarquização das variáveis.

## Contribuições

O presente estudo marca um momento na história da investigação em terapêutica pela música e parece encerrar um capítulo. Daqui para diante tudo tem de ser feito diferentemente. Cada autor não deverá continuar a sentir-se livre para usar as variáveis a bel-prazer ou que lhes são mais convenientes, na instituição em que se situa, mas antes proceder a uma estruturação teórica da investigação.

Vale a pena prosseguir estes estudos. No momento actual torna-se imperativo considerar as variáveis neuro-psicológicas.

## Comentário

Os estudos da Psicologia da Música, ou, em geral os da Psicologia da Arte, sempre foram pouco cultivados, dado a sua dificuldade intrínseca. De facto, o investigador tem de dominar dois campos do saber, a Psicologia e a arte em causa. Há poucas pessoas que orientam os estudos e a carreira nestes dois domínios, com igual proficiência. É um facto. Ele só pode ser superado através da constituição de equipas multidisciplinares. Como se referiu, a investigação em terapêutica pela música ainda é mais complexa e exigente. Não basta que alguém domine dois ramos do saber. É também requisito dominar uma arte, a mais antiga de todas: a arte de curar.

## 2.3.

### AN INVESTIGATION OF PRESCHOOL CHILDREN'S COMPARATIVE CAPABILITY TO SING SONGS WITH AND WITHOUT WORDS

ESTUDO COMPARATIVO SOBRE O DESEMPENHO VOCAL DE CANÇÕES COM E  
SEM PALAVRAS EM CRIANÇAS DO PRÉ-ESCOLAR

L. M. LEVINOWITZ (1989)

BULLETIN OF THE COUNCIL OF RESEARCH IN MUSIC EDUCATION, 100, 14 - 19

### 2.3.1. Resumo do artigo

#### Resumo

Este estudo visou verificar se uma canção é mais bem aprendida quando ensinada com palavras ou sem palavras. Visou ainda verificar se o nível de desenvolvimento linguístico interfere neste desempenho.

Os participantes foram 35 crianças com idades compreendidas entre os quatro e os cinco anos de idade. Após um período de instrução musical de cinco meses, integrando a aprendizagem de canções com palavras e sem palavras, as crianças foram submetidas à aprendizagem de duas canções – uma com palavras, outra sem palavras. Ao fim de um mês, procedeu-se à gravação das suas performances, posteriormente avaliadas por dois juízes. Foram também recolhidos dados relativos ao desempenho das crianças num teste de vocabulário.

Foram efectuados testes para análise das diferenças entre a canção aprendida com palavras e a canção aprendida sem palavras em termos de desempenho tonal e de desempenho rítmico. Os resultados mostraram um melhor nível de desempenho vocal em termos tonais na canção aprendida sem palavras. No entanto, no que toca ao desempenho rítmico não se verificaram diferenças.

A relação entre o desenvolvimento da linguagem da criança e a sua capacidade tonal e rítmica para cantar uma canção com palavras e uma canção sem palavras foi analisada através do cálculo de coeficientes de correlação, tendo-se obtido valores baixos.

A autora conclui defendendo a necessidade de utilização de canções sem palavras na instrução musical de crianças desta faixa etária.



# An investigation of preschool children's comparative capability to sing songs with and without words

L. M. LEVINOWITZ

## Introduction

The music capability of young children is not a new subject of investigation. Observational studies have been undertaken to discover and to classify young children's natural forms of music and to determine means of developing their tonal and rhythm capabilities (Moorhead and Pond, 1977; Holahan, 1984).

An important natural form of young children's music is spontaneous performance. Young children teach themselves music when they engage in spontaneous performance. Moreover, spontaneous performance may be the best indication of how young children think musically. Two types of spontaneous performance classified by Moorhead and Pond are group chant and spontaneous song (1977). The distinctions between chant and song are clear: 1) chant is sung most often by children in a group; spontaneous song by children in solitude, 2) the rhythmic structure of chant is rigid and associated with the children's body movements; the free, asymmetrical rhythmic structure of spontaneous song resembles plainchant, 3) the melody of chant is built around a recitation note which is inflected, for purposes of accentuation, downward a minor third; the melody of spontaneous song is composed of wide leaps as well as linear progressions.

The similarities between chant and song are perhaps less apparent but no less intriguing: 1) chant and spontaneous song are usually performed with a text. The words that comprise the text of the chant or song may be an expression of a thought, a statement of fact, or an imaginative arrangement of words that do not necessarily conform to the conventional manner of speech (Moorhead and Pond, 1977). That is, a child may use existing vocabulary, but the words have no orderly arrangement or syntax. 2) Although the words are apparent in young children's performances of chant and song, Moorhead and Pond thought that the music elements of those performances are of primary importance to the young child and the text is of secondary importance (Moorhead and Pond, 1977). Therefore, it is conceivable that teaching themselves rhythm is the primary concern of young children when they engage in spontaneous performances of chant, and teaching themselves melody is the primary concern of young children when they perform songs spontaneously.

Young children also learn music by performing songs by rote. Most, if not all, of the songs adults teach to children have a text. The text of these rote songs, however, is unlike the text of the child's spontaneous song. That is, the sentences of a rote song have syntax.

To the adult it is the text that is the appealing element of the song. Therefore, it is not surprising that the songs taught young children are chosen on the basis of the text rather than the music. Goetze concluded that the use of a neutral syllable such as "loo" rather than the text of a song improved the performance of young children (Goetze, 1985). The young child may be distracted by the orderly arrangement of the words of rote songs. The music elements, which should be the primary consideration become secondary. Perhaps rote songs that have no words should also be taught to the young child.

Furthermore, the young child's language development may interact with his ability to learn or perform a rote song. That is, a child whose language development is above average may not be

distracted by the words and, therefore be able to attend to the music elements of that song. The young child who is facile with the language might learn the words first and then associate the melody with those words. The child whose language development is average or below average, however, may be so distracted by the words that he is unable to attend to the musical elements of that song. An alternate explanation is that with less language facility a child first learns the melody and then associates the words with that particular melody.

The need is to know if there is a relationship between the young child's ability to sing rote songs and his language development. Does the child perform a rote song with words better than one without words? The answer to this question may give music educators a clue to how a young child learns music.

## Design and analysis

Thirty-five children from two classes of four and five-year-old children from St. Thomas's Nursery School in Fort Washington, Pennsylvania participated in the study. The children represented a population of white, upper-socioeconomic, nursery school students.

For five months, all children in both classes received music instruction from the investigator one day each week for thirty minutes. That music instruction included rhythm, movement, and rote-singing activities. During any one music class, half of the rote songs were sung with words; the other half were sung on a neutral syllable such as "bum" or an onomatopoetic sound such as "moo" or "drip." The repertoire of songs, with or without words included songs in major, harmonic minor, dorian, phrygian, lydian, mixolydian, and aeolian tonalities and in duple, triple, and unusual meters.

During the last month of instruction, two criterion songs were taught to the children in both classes (Figure 1); one song was sung with words, the other song was sung on the neutral syllable "bum." These songs had similar melodic and rhythmic contour and a similar harmonic structure. Each song was sung the same number of times by all children in both classes.

With words

The Jack in the box jumps up. The

Jack in the box goes flop. The Jack in the box goes

round and round the lid comes down with a plop.

Without words

Sung on "bum"



At the end of instruction the performance of each child was tape-recorded. Approximately one half of the children sang the criterion song with text first and the criterion song sung on a neutral syllable, second; the remaining children sang the criterion song sung on a neutral syllable first and the criterion song with text, second.

Two five-point rating scales were constructed to assess the children's tonal and rhythm singing achievement. Two judges used the rating scales to evaluate the children's performance. The combined rating and the combined rhythm rating of the two judges for the song with words and the song without words comprised the four singing achievement scores.

## **Rating scales**

### **Tonal**

The Child:

1. uses a singing voice.
2. sings small deviations around a pitch center.
3. sings, tonicizes one pitch, and sings intervals away from that pitch with correct melodic direction.
4. sings either tonic, dominant, subdominant or cadential patterns in tune.
5. sings all patterns in tune.

### **Rhythm**

The Child:

1. sings, but there is no evident rhythmicity; resembles plainchant.
2. sings one pattern consistently the same throughout the performance.
3. sings the song with a consistent tempo.
4. sings at least one phrase with a sense of meter.
5. sings the entire song with a sense of meter; the entire set of patterns are correct.

The Peabody Picture Vocabulary Test (PPVT) was used to assess the children's language development. The PPVT is a test designed to measure the receptive vocabulary of preschool and school age children. Each child was presented with four simple line drawings at a time and asked, "Put your finger on the word I have just said." The order of the words proceeded from common to rare. These standardized raw scores constituted the language development scores.

## **Analysis**

To interpret the young child's relative rhythm capability to sing a song with words and a song without words, the rhythm achievement scores were organized into a one-dimensional design. The same design was used for tonal achievement scores. Correlated sample t-tests were used to determine the extent of the differences.

The relationship between the young child's language development and his tonal and rhythm capability to sing a song with words and a song without words, was determined by first-order partial correlation coefficients with the children's ages in months held constant.

## Results and interpretation

### Reliability

The interjudge reliabilities of the tonal and the rhythm rating scales for the criterion song with words, the criterion song without words, and both songs combined are reported in Table 1. These reliabilities are high. A consistency between the ratings of the two judges indicates indirect validity for the tonal and the rhythm rating scales. For both the rhythm and the tonal rating scales, the interjudge reliabilities were higher for the song with words than for the song without words.

**Table 1**  
**Interjudge Reliabilities for the**  
**Tonal and Rhythm Rating Scales**

|          |     |                                                  |
|----------|-----|--------------------------------------------------|
| TWO      | .78 | TWO - tonal rating scale/<br>song without words  |
| TWW      | .93 |                                                  |
| COMBINED | .87 | TWW - tonal rating scale/<br>song with words     |
| RWO      | .84 | RWO - rhythm rating scale/<br>song without words |
| RWW      | .94 |                                                  |
| COMBINED | .90 | RWW - rhythm rating scale/<br>song with words    |

### Analysis of Rhythm Achievement Scores

Means, standard deviations, and correlated sample t-test data for the subjects are presented in Table 2. No difference was found between the children's rhythm performances of the criterion song with words and the criterion song without words.

**Table 2**  
**Means, Deviations, Correlated Sample t-Test Data**  
**Rhythm Achievement Criteria**

|       | <u>RWO</u>        | <u>RWW</u>         |
|-------|-------------------|--------------------|
| N     | 35                | 35                 |
| X     | 7.34              | 7.46               |
| S.D.  | 2.51              | 2.81               |
| ..... |                   |                    |
|       | <u>Difference</u> | <u>t-statistic</u> |
| d     | .029              | .087 <sup>ns</sup> |
| SSd   | 126.971           |                    |

The lack of a difference between the children's rhythm performances of the criterion song with words and the criterion song without words may have occurred for one of two reasons. Either there is no difference, or the rating scale was inadequate.



### ***Analysis of Tonal Achievement Scores***

A difference was found between the children's tonal performances of the criterion song with words and the criterion song without words ( $t = -3.43$ ,  $p < .01$ ). The song without words was performed better than the song with words. The result could be an error or children do learn the melody of a song better when they are not distracted by the words.

### ***Partial Correlation Coefficients***

The partial correlation coefficients are presented in Table 3. The correlation coefficients between children's language development and tonal and rhythm singing achievement are low. At best, only 4% of the variation in children's language development is accounted for by their tonal performance of a song with words. The PPVT is also correlated with Stanford-Binet I.Q. scores. Therefore, it is not surprising that the correlation between children's language development and their tonal or rhythm rote-singing achievement is similar to the correlation between music achievement and scholastic aptitude (Gordon, 1968; Hobbs, 1985).

Table 3  
First-Order Partial Correlation Coefficients

|      | TWW  | TWO  | RWW  | RWO  | PPVT |
|------|------|------|------|------|------|
| TWW  | 1.00 |      |      |      |      |
| TWO  | .80  | 1.00 |      |      |      |
| RWW  | .56  | .49  | 1.00 |      |      |
| RWO  | .60  | .64  | .64  | 1.00 |      |
| PPVT | .19  | .10  | .09  | .08  | 1.00 |

### **Conclusion**

Young children should receive rote-singing instruction that includes both songs with and without words. It cannot be concluded that there is a relationship between young children's language development and their tonal and rhythm performance of rote songs with and without words. Perhaps for young children to learn a song with words, two mental processes are necessary; one for audiation (inner-hearing) and the other for learning the words of the song.

## References\*

- Dale, Philip S. *Language development: structure and function*. New York: Holt, Rinehart, and Winston, 1976.
- Dunn, Lloyd M. *Peabody Picture Vocabulary Test*. Minnesota: American Guidance Service, Inc., 1959.
- Goetze, Mary. Factors affecting accuracy in children's singing. Unpublished doctoral dissertation, University of Colorado, 1985.
- Gordon, Edwin. A study of the efficacy of general intelligence and musical aptitude tests in predicting achievement in music. *Council for Research in Music Education*, 13 (1968), 40-45.
- Hobbs, Christine. A comparison of music aptitude, scholastic aptitude and academic achievement of young children. *Psychology of Music* 13 (1985), 93-98.
- Holahan, John M. The development of music syntax: Some observations of music babble in young children. Music in Early Childhood Conference. 28 June, 1984. Brigham Young University: Provo, Utah.
- Moorhead, Gladys E. and Donald Pond. *Music of young children*. Santa Barbara, California: Pillsbury Foundation for Advancement of Music Education, 1977.

---

\* As referências bibliográficas deste artigo estão conforme o documento original.



### 2.3.3. Estudo Crítico – Por Helena Rodrigues

#### Objectivos

O estudo pretende: 1) Verificar se uma canção aprendida com palavras é melhor ou pior desempenhada que uma canção aprendida sem palavras; 2) Verificar se o nível de desenvolvimento linguístico da criança tem influência no desempenho de uma canção, conforme se trate de uma canção com palavras ou sem palavras.

#### Design

Os sujeitos são submetidos à manipulação de uma Variável Independente (canção ensinada com ou sem palavras), ao longo de um período de instrução. O desempenho vocal das crianças é a Variável Dependente, actuando o seu nível de desenvolvimento linguístico como Variável Moderadora.

#### Revisão do artigo

##### Introdução e revisão

Na revisão da bibliografia, a autora refere essencialmente dois tipos de estudos que inspiram o presente artigo. No primeiro tipo de estudos, de carácter observacional, referem-se sobretudo os trabalhos de Moorhead e Pond (1977) que se encontram entre os primeiros trabalhos a caracterizar o desenvolvimento das capacidades musicais de crianças muito jovens. Estes trabalhos valorizam a observação em contexto natural dos comportamentos musicais ocorridos espontaneamente, estabelecendo distinções entre a canção e a entoação. Afirma-se que, apesar de se tratar de diferentes situações, em ambas as situações as crianças atribuem um valor secundário às palavras.

Depois, é referido outro estudo, levado a cabo por Goetze (1985), em que se concluiu que o desempenho vocal das crianças melhorava quando a canção era ensinada sem palavras.

A autora interroga-se se o nível de desenvolvimento linguístico poderá interferir com a sua capacidade de aprender ou desempenhar uma canção. Isto é, o facto da criança se distrair com as palavras presentes na canção ou, pelo contrário, o facto da criança se concentrar na melodia para depois aprender as palavras, poderão ser factores de influência no processo de aprendizagem de canções com palavras? O esclarecimento desta questão é o ponto de partida para o presente artigo.

A revisão da bibliografia não é extensa, mas de facto trata-se de um assunto pouco abordado na literatura, existindo poucos elementos relacionados com o tema. Este facto acrescenta interesse ao estudo efectuado pois trata-se de um contributo original. Sem dúvida que se trata de um tema fascinante que não recebeu ainda a atenção que merece na literatura especializada.

Seria desejável, todavia, que a idade das crianças a quem se referem os estudos efectuados tivesse sido referida. Presume-se que se trata também de crianças em idade pré-escolar. No entanto, a sua especificação seria importante. Efectivamente, a faixa etária compreendida entre os zero e os cinco anos de idade é muito rica em termos de aquisição musical e linguística e descrever o que sucede aos dezoito meses em termos de relação palavra / música

numa canção não é exactamente o mesmo do que sucede aos cinco anos de idade. Por outro lado, teria sido enriquecedor conhecer não só as conclusões de Goetze mas também minimamente algo da metodologia seguida pelo autor, tanto mais que este estudo parece ser uma sua replicação num dos seus objectivos.

Finalmente, consultando as referências bibliográficas da autora, percebe-se que esta preocupação relativa à existência ou não de palavras na aprendizagem de canções se enquadra dentro de ideias defendidas pela teoria de aprendizagem musical de Edwin Gordon. Esta fonte bibliográfica deveria, pois, ter sido incluída explicitamente na revisão teórica do artigo.

## Método

Participaram no estudo 35 crianças com idades compreendidas entre os quatro e os cinco anos de idade, a frequentar um jardim de infância de elevado nível sócio económico.

Ao longo de cinco meses as crianças foram submetidas a um período de instrução musical semanal com a duração de 30 minutos. Durante este período, metade das canções era aprendida com palavras enquanto que a outra metade era aprendida sem palavras. No último mês de instrução, foram ensinadas duas canções: uma das canções foi ensinada com palavras; a outra foi ensinada sem palavras.

No final do período de instrução procedeu-se à gravação de cada criança cantando individualmente cada uma destas canções: metade das crianças cantou a canção com palavras em primeiro lugar, seguindo-se a canção sem palavras; a outra metade procedeu de forma inversa.

A avaliação do desempenho vocal das crianças foi efectuada por dois juizes que usaram duas escalas de classificação – uma para a Avaliação Tonal, outra para a avaliação do Ritmo. Para a canção com palavras e para a canção sem palavras, foram usados os resultados tonais e rítmicos combinando as classificações de ambos os juizes, produzindo-se assim quatro resultados para a avaliação do desempenho vocal.

O desenvolvimento da linguagem das crianças foi avaliado através do teste *Peabody Picture Vocabulary Test* (PPVT). Trata-se de um teste que apresenta imagens, sendo a criança instruída para apontar a imagem correspondente à palavra ouvida.

A descrição do procedimento é correcta e deixa perceber o cuidado inerente a determinadas opções. Por exemplo, sendo o critério de medida deste estudo as duas canções usadas no último mês de instrução musical, é louvável que as mesmas tenham sido especificamente criadas para o efeito e que tenha havido a preocupação de criar duas melodias que supostamente apresentassem o mesmo tipo de dificuldades. De louvar também é a reprodução destas canções no artigo – este aspecto deveria ser obrigatório e óbvio, mas nem sempre se observa este cuidado.

De modo análogo, é de registar o cuidado havido no balanceamento da performance das crianças. Efectivamente, ao garantir que metade das crianças começou por cantar a canção com palavras seguindo-se a sem palavras e vice-versa, está-se a precaver efeitos causados pela sequência da apresentação o que poderia causar interferências no desempenho vocal das crianças.

É interessante que a autora tenha utilizado o mesmo grupo de crianças sujeitando-as a um período de instrução do qual constava a aprendizagem de canções com palavras e de canções sem palavras e que as tenha depois sujeito a uma avaliação idêntica. A autora poderia ter optado por escolher dois grupos de crianças: um a quem se ensinasse apenas canções com palavras e outro a quem se ensinasse apenas canções sem palavras. No entanto, a opção efectuada tem a vantagem de controlar melhor possíveis diferenças de partida entre os sujeitos.

Não é, no entanto, tão clara a razão por que a autora optou por realizar um período prévio de instrução musical de cinco meses em que ensinou canções com e sem palavras simultaneamente, em vez de passar directamente ao período de instrução musical em que foram ensinadas as canções do estudo propriamente dito. Terá sido para desta forma controlar melhor experiências prévias de aprendizagem dos sujeitos? Terá sido para familiarizar os sujeitos com um determinado tipo de prática educativa?



## Resultados

Antes de mais, refira-se que a autora reporta valores elevados para o cálculo da fiabilidade inter-juizes em todas as situações de avaliação (nas escalas de classificação tonal e rítmica para a canção aprendida com palavras, para a canção aprendida sem palavras e para ambas as canções respectivamente na situação tonal e na situação rítmica). Este cuidado é importante e mostra o rigor da autora relativamente às questões de medida – cuidado este nem sempre presente na literatura da especialidade.

Não obstante, questiona-se a afirmação «a fiabilidade entre as classificações dos dois juizes mostra validade indirecta das escalas de classificação tonal e rítmica» («a consistency between the ratings of the two judges indicates indirect validity for the tonal and the rhythm rating scales» (p. 18). Efectivamente, pode observar-se que os índices de fiabilidade observados são maiores para a situação de avaliação da canção com palavras do que para a de avaliação da canção sem palavras. Tal não é de estranhar, pois é mais fácil haver acordo entre os juizes na situação da canção com palavras, já que é relativamente fácil avaliar o desempenho das palavras. No entanto, este detalhe pode revelar uma certa "conspiração" no que está efectivamente a ser avaliado (recorde-se que as palavras não fazem explicitamente parte das escalas de classificação usadas) e daí as reticências colocadas relativamente à "validade indirecta"<sup>1</sup>.

Foram usados *Testes t* para analisar os níveis de desempenho vocal das crianças, em termos rítmico e tonal, comparando a situação de aprendizagem da canção com palavras com a de aprendizagem da canção sem palavras. Em termos tonais, observaram-se diferenças significativas no desempenho vocal da canção sem palavras face à canção com palavras. No entanto, não se verificaram diferenças na situação de avaliação rítmica.

Questiona-se a razão por que a autora apresenta os resultados de forma diferente: enquanto que na análise dos resultados de desempenho rítmico apresenta uma tabela com todos os dados, na análise dos resultados tonais limita-se a indicar « $t = -3.43$ ,  $p < .01$ ». (Isto é, há uma diferença de 3,43 entre as médias de ambos os grupos – esta diferença é negativa pois ter-se-á subtraído a média do resultado da canção sem palavras à da canção com palavras – sendo o nível de significância de .99. Ou seja, esta diferença poderá dever-se ao acaso em apenas 1%.)

Por outro lado, surpreendem os comentários que faz quando afirma, relativamente ao ritmo, «Ou efectivamente não há diferença, ou a escala de classificação é inadequada» («Either there is no difference or the rating scale was inadequate») (p.18) e, relativamente à avaliação tonal, «O resultado poderá ser um erro ou efectivamente as crianças aprendem melhor a melodia de uma canção quando não são distraídas com a utilização de palavras» («The result could be an error or children do learn the melody of a song better when they are not distracted by the words») (p.18). Surpreendem, na medida em que as contingências de qualquer investigação deveriam ser um pressuposto e, portanto, parecem desnecessários estes cuidados. De qualquer modo, é possível também olhar para estes comentários da autora como uma espécie de "excesso de zelo".

A relação entre o desenvolvimento da linguagem da criança e o seu desempenho vocal (tonal e rítmico) para cantar uma canção com palavras e uma canção sem palavras foi analisada através do cálculo de coeficientes de correlação entre os resultados obtidos pelas crianças no teste PPVT e nas *rating scales* de avaliação do desempenho vocal.

A tabela 3 apresenta os coeficientes de correlação existentes entre as várias situações testadas. (Naturalmente, o valor da correlação dos testes consigo próprios é 1.00 – por ex, veja-se coluna PPVT e intersecção com a fila PPVT, sucedendo o mesmo, evidentemente, com os outros testes).

O valor mais elevado resulta da correlação entre PPVT com TWW (situação de avaliação tonal da canção sem palavras), que é .19. No entanto, este é um valor muito baixo pois, tal como afirma a autora, significa que apenas 4% da variação do desenvolvimento linguístico se pode ficar a dever ao desempenho tonal de uma canção sem palavras. (Este valor de 4% é a raiz quadrada de .19)

A autora refere não ser surpreendente este resultado invocando um raciocínio silogístico: estando o PPVT correlacionado com os resultados do Stanford-Binet e não estando este relacionado com o desempenho musical, seria de esperar, logicamente, que o PPVT não estivesse associado ao desempenho musical.



## Conclusões

Em face dos resultados, a autora apresenta duas conclusões: 1) a instrução musical das crianças desta faixa etária deverá incluir canções com e sem palavras; 2) não se pode concluir que exista uma relação entre o desenvolvimento linguístico das crianças e o desempenho tonal e rítmico de canções com e sem palavras.

Em relação à segunda conclusão é interessante notar a sua formulação: a autora não escreve «não existe uma relação entre o desenvolvimento linguístico das crianças e o desempenho tonal e rítmico de canções com e sem palavras». Ou seja, poderá dizer-se que a autora partiu da hipótese que o nível de desenvolvimento linguístico da criança tem influência no desempenho de uma canção, conforme se trate de uma canção com texto ou sem texto. O estudo não confirma a hipótese mas também não a rejeita.

A autora termina chamando a atenção para a possibilidade de existência de dois processos mentais necessários para a aprendizagem de uma canção com palavras: um relacionado com a aprendizagem da audição, o outro com a aprendizagem das palavras de uma canção.

Relativamente à primeira conclusão há, sem dúvida, uma série de argumentos educativos de carácter filosófico que a poderão sustentar. No entanto, parece um pouco especulativo sustentar que a instrução musical das crianças deva incluir canções com e sem palavras apenas pelos resultados obtidos neste estudo. De facto, o design deste estudo contemplou um período de instrução musical prévio em que foram simultaneamente utilizadas canções com palavras e sem palavras. Não sabemos, portanto, o que sucederia se se tivesse usado apenas canções com palavras ou apenas canções sem palavras durante o período de instrução musical. Ter-se-ia chegado às mesmas conclusões? Não sabemos, também, o que sucederia se não tivesse existido esse período de instrução.

## Contributo para a área de estudo

O artigo apresenta a vantagem de questionar uma prática corrente (o ensino de canções com palavras), que é usada habitualmente pelos educadores. Portanto, antes de mais é de louvar a atitude de questionar algo que se dá por adquirido mas que nunca foi objecto de uma análise sistemática. Efectivamente, uma das razões que sustenta a necessidade de se efectuar investigação no âmbito da Psicologia e da Pedagogia Musical é, justamente, questionar práticas que a tradição implementou, e verificar empiricamente se há ou não sustentação científica para as mesmas.

Não obstante, talvez a conclusão devesse ser mais modesta, limitando-a aos resultados específicos encontrados. De facto, poderá a decisão entre ensinar canções com palavras e/ ou sem palavras basear-se apenas em dados que demonstrem que o desempenho vocal é melhor ou pior numa ou noutra situação?

Ao interrogar-se também sobre as relações entre desempenho vocal e desenvolvimento linguístico, o estudo abre questões de investigação interessantes. Efectivamente, muitas vezes vê-se argumentar a necessidade de implementação da educação musical no currículo baseando-a em ligações com outras áreas do conhecimento humano (como o desenvolvimento linguístico ou, muito frequentemente, o raciocínio matemático), sem que estas ligações tenham sido devidamente testadas.

Finalmente, refira-se que se trata de um artigo claro, conciso e sintético – e estas são qualidades que é sempre bom encontrar em artigos do âmbito da Investigação em Psicologia e Pedagogia Musical.



## Comentário

Este artigo faz reflectir sobre questões que podem ser inspiradoras para novas pistas de trabalho.

Por exemplo, detenhamo-nos na afirmação «é possível que as crianças estejam a ensinar-se a si próprias ritmo e melodia quando produzem espontaneamente cantos rítmicos e canções, respectivamente» («it is conceivable that teaching themselves rhythm is the primary concern of young children when they engage in spontaneous performances of chant, and teaching themselves melody is the primary concern of young children when they perform songs spontaneously») – (p. 15). Esta afirmação é interessante pois chama a atenção para algo a que se tem prestado pouca atenção: as produções vocais espontâneas das crianças na fase etária compreendida entre os dois e os cinco anos de idade (lamentavelmente, o período etário não é referido no artigo mas atendendo às investigações de Moorhead e Pond citadas, será lícito assumir que se trata desta faixa etária).

Sem dúvida que esta produção vocal espontânea, recolhida em contexto natural, pode constituir um bom domínio de pesquisa, com a vantagem de que, hoje, os meios de gravação e recolha são muito mais sofisticados do que os existentes na época de Moorhead e Pond. A afirmação distingue dois tipos de produção vocal espontânea: o canto rítmico e a canção. Esta distinção é interessante e deverá ser tida em consideração em trabalhos futuros que se proponham recolher e analisar a produção vocal espontânea das crianças na faixa etária compreendida entre os dois e os cinco anos de idade.

Por outro lado, ao pretender avaliar o desenvolvimento linguístico das crianças através do PPVT, o artigo acrescenta um novo ingrediente às reflexões que habitualmente se centram ou no desenvolvimento musical ou no desenvolvimento verbal.

Em suma, tomar em consideração aspectos referentes a estes três elementos – produção vocal espontânea (cantos rítmicos e canções), desenvolvimento musical e desenvolvimento linguístico – pode abrir pistas de conhecimento inovadoras quando se pretende descrever o desenvolvimento psicológico da criança.

O artigo pode inspirar ainda a que se produzam trabalhos semelhantes que contemplem a situação de compreensão e discriminação auditiva em vez da situação de produção vocal. Por exemplo, o que é melhor assimilado auditivamente (e tenha-se presente que a própria definição de "assimilação auditiva" pode revestir facetas várias): uma canção ensinada com palavras ou uma canção ensinada sem palavras<sup>2</sup>?

Finalmente, temas de estudo como este, em que a música se cruza com a palavra (sendo possível a existência de interações complexas entre aptidão musical tonal, aptidão musical rítmica e diferentes aspectos do desenvolvimento linguístico) merecem, decididamente, uma abordagem neurológica.

Este é um campo de estudo em expansão em que as equipas de investigação muito ganhariam se integrassem estudiosos capazes de reflectirem sobre problemáticas como as abordadas no presente artigo. A interpretação de uma canção com texto parece mobilizar simultaneamente capacidades de índole musical e de índole linguística. Saber como ambas se relacionam em termos de aprendizagem ao longo do desenvolvimento humano é algo que ainda está por desvendar, a que certamente a neurologia poderá oferecer um importante contributo. Cabe aos profissionais interessados pelas questões da Psicologia e da Pedagogia Musical ajudar a encontrar as perguntas certas. A este respeito, este artigo pode ser extremamente inspirador.

---

<sup>1</sup> Ou seja, apesar dos juízes avaliarem de forma idêntica (o que é um indício de fiabilidade), ambos podem estar equivocados da mesma forma. Podem, portanto, estar a cometer o mesmo tipo de erros de avaliação. Ou seja, ambos podem ser consistentes na sua avaliação incorrecta; pode haver fiabilidade e não existir validade. Não obstante, a situação esperada é que dois avaliadores independentes não cometam sistematicamente os mesmos erros estando, de facto, a avaliar o objecto. Por isso, a autora afirma que a fiabilidade entre os juízes é um indício de que ambos estão a avaliar o que se pretende (validade).

O que aqui se argumenta é que ao se verificar maior consistência na situação de existência de palavras na canção, ambos os juízes poderão, de facto, estar a avaliar algo que não é exactamente o conteúdo das *rating scales*. Portanto, neste caso específico, a fiabilidade poderá não ser necessariamente um "sintoma" indirecto de validade das *rating scales*.

<sup>2</sup> Para mais discussão sobre questões congéneres, consultar o primeiro capítulo da obra *Bebé Babá – da musicalidade dos afectos à música com bebés* de Rodrigues, H., Rodrigues, P. & Nunes, P. (2003).



## 2.4.

### EFFECTS OF INSTRUCTION AND TASK FORMAT ON PRESCHOOL CHILDREN'S MUSIC CONCEPT DISCRIMINATION

EFEITOS DA INSTRUÇÃO E DO TIPO DE TAREFAS NA DISCRIMINAÇÃO DE  
CONCEITOS MUSICAIS DE CRIANÇAS EM IDADE PRÉ-ESCOLAR

W. L. SIMS (1991)

JOURNAL OF RESEARCH IN MUSIC EDUCATION, 39, 298 - 310

#### 2.4.1. Resumo do artigo / Abstract

##### Resumo

Foram realizados dois estudos complementares para investigar a capacidade de crianças em idade pré-escolar identificarem um elemento ou uma combinação de elementos musicais, sendo, para tal, levadas a cabo actividades como a audição, movimento corporal e canto.

O primeiro estudo pretendeu verificar o efeito da instrução de curta duração em crianças em idade pré-escolar, analisando a sua capacidade de discriminação em tarefas de âmbito musical. Os sujeitos (N=30) foram divididos em dois grupos, sendo um o Grupo Experimental e o outro o Grupo de Controlo. Apenas este último foi submetido a um Pós-Teste, por forma a estabelecer termo de comparação com os resultados do Grupo Experimental. O Grupo Experimental foi submetido a um conjunto de quatro sessões (realizadas em pequenos grupos) por forma a ensinar a discriminar entre rápido/lento e calmo/agitado<sup>1</sup>, assim como a realizar discriminações em que fossem usadas as quatro combinações possíveis destas quatro características. Os resultados indicaram que a instrução dos sujeitos esteve significativamente relacionada com a capacidade de identificação e classificação das características musicais. O estudo mostrou também que os sujeitos foram notoriamente mais precisos nas tarefas de discriminação envolvendo uma característica apenas, por oposição às que envolviam a combinação de características.

O segundo estudo foi planeado como replicação do primeiro, acrescentando o desempenho vocal como uma resposta por parte dos alunos e a idade como variável. Todos os sujeitos (N=42) em idade pré-escolar receberam instrução idêntica à ministrada no primeiro estudo. Os resultados do teste auditivo indicaram novamente que os sujeitos foram significativamente mais eficazes na tarefa de discriminação em que estava em causa apenas uma característica. Também se comprovou que os alunos mais velhos obtiveram melhores resultados que os mais novos. Os sujeitos demonstraram mais facilidade em discriminar características musicais simultâneas a partir da sua própria interpretação vocal, do que a partir de exemplos gravados.

Com base nos dois estudos é possível concluir que: (1) alunos em idade pré-escolar conseguem aprender com facilidade a realizar tarefas de discriminação e categorização de características individuais; (2) a maioria das crianças desta idade poderá não estar preparada para levar a cabo tarefas que exijam atenção simultânea a mais que um elemento musical de cada vez; (3) nas crianças, a capacidade de discriminação baseada nas suas próprias performances poderá desenvolver-se mais cedo do que a competência para a discriminação em situação de audição; (4) poderão ser obtidas indicações iniciais acerca da capacidade de discriminação através da entoação vocal, em imitação simultânea com um dado modelo.



## Abstract

Two studies were completed to investigate young children's ability to identify single and combined musical elements in response to listening, movement, and singing activities.

Study 1 was an examination of the effects of short-term instruction on preschool children's ability to apply decentration to musical tasks. Subjects ( $N = 30$ ) were divided into (a) an experimental group, which participated in four small-group instruction sessions designed to teach the discriminations fast/slow and smooth/choppy as well as these elements four possible combinations and (b) a posttest-only control groups. Results indicated that instruction was significantly related to subjects ability to identify and label musical characteristics, and that subjects were significantly more accurate with single versus double discrimination responses. Study 2 was designed to replicate aspects of the previous study, with the addition of singing as a response mode and age as variable. All the preschool-age subjects ( $N = 42$ ) received the instruction component of Study 1. Results of the listenings test again indicated that subjects were significantly more successful with the single discrimination task and that older children scored significantly higher than younger children. The subjects were better able to label simultaneous musical characteristics in their own singing than in recorded listening examples. Conclusions drawn, based on both studies, include: (1) preschool-age children can easily learn to make and label single discriminations; (2) most young children may not be ready for music listening tasks requiring attention to more than one element at a time; (3) children's ability to make discriminations based on their own performance may develop earlier than their ability to make discrimination in listening situations; and (4) initial indications of discrimination may be demonstrated through singing in simultaneous imitation with a model.

## 2.4.2. Artigo Original

# Effects of instruction and task format on preschool children's music concept discrimination

W. L. Sims (1991)

Researchers have completed a variety of studies to identify and evaluate young children's musical skills, knowledge, and preferences, which have provided useful and enlightening insights into early childhood musical development. As Scott (1987) pointed out, however, the profession still has some critical needs, including a "solid body of literature determining either short- or long-term effects of music instruction with the young," as well as "theories of musical development which could serve to provide a framework for our thinking and a focus for our research" (p. 11).

One aspect of musical development that has been of particular interest to researchers is young children's ability to discriminate between contrasting musical elements such as loud/soft or fast/slow. According to Zimmerman (1984), "a developmental sequence pervades research findings in music concept formation, with concepts developing in the following order: volume, timbre, tempo, duration, pitch, and harmony" (p. 72). McDonald and Simons's (1989) review of the research literature resulted in the identification of a similar sequence. Most of the numerous studies on which these conclusions were based were examinations of young children's ability to discriminate an isolated musical element in the context of songs, musical excerpts, single pitches, or patterns. Since music is characterized by the simultaneous interaction of a number of elements, however, determining how young children respond to more than one element at a time seems important to developing a more complete understanding of children's responses to music.

Little is known yet about how young children respond to simultaneous musical elements, although music educators have suggested, and some research seems to indicate, that young children tend to "center" their attention on one characteristic of a musical stimulus to the exclusion of others (Hargreaves, 1986; Petzold, 1981; Shuter-Dyson & Gabriel, 1981; Sims, 1988; Young, 1982; Zimmerman, 1986). Although this conclusion is based primarily on examination and application of the child development theories of Jean Piaget, the extent to which his ideas pertain to musical development has yet to be determined empirically.

The most substantial work in this area is the body of research in which authors sought to relate Piaget's theories of conservation to the development of musical thought (see Serafine, 1980, and Hargreaves, 1986, for reviews of this literature). The results of these studies are inconclusive, however, and there is some question as to whether musical tasks can approximate or be analogous to Piagetian conservation tasks (Bartholomew, 1987; Hargreaves, 1986; Serafine, 1980). One question as to the appropriateness of the music tasks arises due to the mixed results of studies in which experimenters attempted to determine whether performance on musical conservation tasks can be improved with instruction (Hargreaves, 1986; Shuter-Dyson & Gabriel, 1981; Zimmerman, 1986). According to the theory, children progress through stages of development characterized by distinctive patterns of thinking and approaches to problem solving, the acquisition of which are developmental in nature and *not* amenable to direct training.

The Piagetian stage of development in to which most preschool-age children would fall is the preoperational stage. One characteristic of this period is the child's tendency to "center his attention exclusively on some single feature or limited portion of the stimulus array that is particularly salient and interesting to him, thereby neglecting other task-relevant features" (Flavell, 1977, p. 81). The ability to decenter, or to consider several characteristics of a stimulus simultaneously, is a basic perceptual skill. As it represents only one component of the conservation



response, it is a less complex and more easily isolated construct than conservation, and may be more appropriate for musical application. While the complete process of conservation has received much attention in the literature, there is little research available that addresses children's ability to make musical judgments and discriminations requiring decentration, or how and when this ability develops. As the ability to perceive and discriminate is considered critical to concept development (Gape, 1977; Greenberg, 1976), information about the development of decentration not only will provide knowledge relevant to developmental theory, but also might have an impact on techniques and materials selected to provide musical experiences and instruction for young children.

One study has directly addressed developmental aspects of the ability to apply decentration to a music listening task. Sims (1988) examined subjects' ability to demonstrate decentration during a nonverbal movement-to-music activity that required the untrained subjects to respond to two musical characteristics simultaneously. Results indicated that there were qualitative differences among the responses of four groups: preschoolers, first and second graders, third graders, and nonmusician college students. It was concluded that "there may be identifiable stages in the development of the ability to demonstrate decentration in response to a musical stimulus" (Sims, 1988, p. 125). Sims also noted that the relationship between these results and child development theory would be clarified by determining whether the discrimination required would be amenable to instruction. The first study reported in the present article examined the effects of shortterm instruction on the ability of preschool children to decenter, or make responses requiring two simultaneous musical discriminations. The study was designed to explore further the possible relationship between children's musical behavior and developmental theory, as well as to provide basic information about characteristics of young children's music concept discrimination skills.

## STUDY 1

### METHOD

The subjects for this study were children ranging in age from 4 years 0 months 1 to 5 years 1 month, enrolled in three preschools with similar programs. Music was included in the daily activities at all three schools, with greatest emphasis placed on rote singing of children's songs and playing singing games with fingerplay or motions. The children represented middle-class families with a variety of ethnic backgrounds. All children who were in attendance for each session served as subjects, resulting in a final  $N = 30$  (3 children were eliminated due to absences). Subjects were randomly divided in to an experimental group ( $n = 15$ ) and a posttest only control group ( $n = 15$ ). The average age of the children in each group was 4 years 6 months.

The experimental variable consisted of short-term instruction designed to teach the children to respond to two musical characteristics at the same time. Two pairs of contrasting characteristics related to tempo (fast/slow) and style of articulation (smooth/choppy) were selected since these discriminations were relevant to the musical task used by Sims (1988), which was being replicated. These characteristics also seem to work well because the two sets of terminology are different enough so as not to be inherently confusing to young children, as seems to be the case with language related to other concepts, such as high/low and loud/soft (McDonald & Simons, 1989). Although children's ability to make discriminations based on tempo has received much attention in the research literature, there has been little research addressing children's ability to discriminate between styles of articulation, and results have been inconclusive (Taebel, 1974, Young, 1982).

For this research, "fast" was operationally defined as examples where beats per minute equaled or exceeded 110, with "slow" examples defined as pieces with beats per minute of 80 or fewer. "Smooth" excerpts consisted of flowing, connected melodies or chords (e.g., the main theme from Smetana's "Moldau" or the cello theme from Camille Saint-Saens's *Carnival of the Animals* "Swan" section), whereas "choppy" was defined as short and detached notes or chords, including but not limited to staccato and pizzicato passages (e.g., phrases performed by the xylophone in the *Carnival of the Animals* "Fossils" section or the sections featuring piano from the "Javanaise" movement of Claude Bolling's *Suite for Flute and Jazz Piano*). Graduate students in



music education were used to establish that the examples clearly exemplified the desired characteristics.

Subjects in the experimental group were divided into small groups of three or four, and participated in four 20-minute small-group instructional sessions. Instruction was designed to provide a variety of experiences with each characteristic, and included many opportunities for the characteristics to be labeled verbally, as labeling and language are important in discrimination and concept learning (McMahon, 1986; Zimmerman, 1984). Instruction was sequenced as follows: Day 1- fast/slow; Day 2- review fast/slow, begin smooth/choppy; Day 3- review smooth/choppy, begin combinations fast and smooth, slow and smooth, fast and choppy, slow and choppy; Day 4-review combinations. All characteristics and their combinations were presented through speech, singing, movement, and listening activities, and experience with each element always preceded labeling.

A posttest was designed to elicit information related to several areas of investigation. One component was a listening test, consisting of short excerpts and procedures similar to those used during the instructional activities. The musical selections were different from those used during the lessons, to test for transfer of the concepts, rather than for the association of labels with specific pieces. For two items, subjects were asked to label the example as either "fast or slow," for two "smooth or choppy," and for two they were asked to listen for both pairs of characteristics, "smooth or choppy AND fast or slow." Subjects were told which set(s) of elements to listen for prior to listening, and were presented with the choices again after the excerpt was played. During the administration of this section of the posttest, no assistance with the terms or their usage was provided to children in either the experimental or control groups.

The second component of the posttest was the movement-to-music activity developed and tested by Sims (1988), based on the piece "Kangaroo" from *Carnival of the Animals*. Subjects were instructed that they would be hearing music that had been composed to describe a kangaroo, and to listen for the parts of the music that sounded like the kangaroo was hopping (short, detached tone clusters) and the parts that sounded like the kangaroo had stopped (sustained cadential chords). During the "hopping" sections, they were to hop like a kangaroo, but "stop to rest and look around" when the music sounded like the kangaroo had stopped hopping. Correct responses to these sections would demonstrate the smooth/choppy discrimination. The discrimination of primary interest, however, occurs during the last phrase of the short piece, during which the choppy "hopping" figures are played, but with a change in tempo. In the previous study, the young children stopped hopping at this point, apparently confused about the change and not seeming to recognize that the music was still characterized by the hopping figures, while many of the older children and all of the college students continued hopping. Of interest to the present study is whether the experimental subjects, after receiving instruction in which they experienced and verbalized that music could display two characteristics at once, would demonstrate this double discrimination by continuing to hop when the hopping sounds were slower. Prior to performance of the movement activity, practice was provided through the use of an experimenter-constructed recording, consisting of examples of "hopping" and "stopping" sounds similar to those in the stimulus piece, but with no change of tempo.

Upon completion of the movement activity, the children were asked to listen to the last phrase of the "Kangaroo" and to indicate whether it was "smooth or choppy." This final test component provided for a comparison of the movement response and verbal response to this phrase.

All posttests were administered in individual sessions lasting approximately 8 minutes and occurring three days after instruction for the experimental group. Verbal answers were recorded on answer sheets, and movement responses were recorded on videotape. A continuous observation procedure was used to transcribe the movement data for analysis (Sims, 1988). For each second of the "hopping" segments, the transcribers recorded on specially constructed observation forms whether or not the subject was hopping. Interobserver reliability, calculated for 57% of the responses, and was equal to .976.



## RESULTS

There were four questions on the posttest (excluding the final question about the "Kangaroo") that required the subjects to discriminate and label a single characteristic of a musical excerpt, and two double discrimination items that required two answers each, resulting in a total of four responses in this category, as well. A two-factor, repeated-measures analysis of variance (ANOVA) was used to compare the responses of the experimental and control groups on the single and double discrimination items. Results indicated significant main effects for treatment condition [ $F(1, 28) = 8.12, p < .05$ ] and for type of discrimination [ $F(1, 28) = 48.61, p < .05$ ], with no significant interaction (see Table 1). The experimental group scored significantly higher than the control group, and both groups scored significantly higher on single as compared to double discrimination items (see Table 2).

Table 1

*Summary Table: Repeated-Measures ANOVA Comparing Experimental and Control Groups on Single and Double Discrimination*

| Source of variance      | SS    | df | MS    | F      |
|-------------------------|-------|----|-------|--------|
| <i>Between subjects</i> | 29.73 | 29 |       |        |
| Groups                  | 6.66  | 1  | 6.66  | 8.12*  |
| Residual                | 23.07 | 28 | 0.82  |        |
| <i>Within subjects</i>  | 39.00 | 30 |       |        |
| Discrimination          | 24.06 | 1  | 24.06 | 48.61* |
| Group x Discrimination  | 1.08  | 1  | 1.08  | 2.18   |
| Residual                | 13.86 | 28 | 0.495 |        |
| <i>Total</i>            | 68.73 | 59 |       |        |

\* $p < .05$ .

All subjects were able to demonstrate the choppy/smooth (hopping/stopping) discrimination related to the movement activity by the end of the short practice example. Responses to the double discrimination section of the activity were considered to be correct if the subject hopped during or after the fifth second of the last phrase, the place at which the tempo change occurred. Seven experimental and five control subjects made correct responses, whereas the remaining 18 children, 60% of the subjects, stopped at the change in the music and did not resume hopping.

Table 2  
Results of Listening and Singing Discrimination Tasks

STUDY 1

|                    | <i>Listening discrimination</i> |                     |
|--------------------|---------------------------------|---------------------|
|                    | Single<br>(4 items)             | Double<br>(4 items) |
| Experimental group |                                 |                     |
| <i>M</i>           | 3.87                            | 2.33                |
| <i>SD</i>          | 0.35                            | 0.82                |
| % correct          | 96.67                           | 58.33               |
| Control group      |                                 |                     |
| <i>M</i>           | 2.93                            | 1.93                |
| <i>SD</i>          | 1.16                            | 0.70                |
| % correct          | 73.33                           | 48.33               |

STUDY 2

|                  | <i>Listening discrimination</i> |                     | <i>Double discrimination</i> |
|------------------|---------------------------------|---------------------|------------------------------|
|                  | Single<br>(4 items)             | Double<br>(4 items) | After singing<br>(8 items)   |
| Younger children |                                 |                     |                              |
| <i>M</i>         | 3.28                            | 1.81                | 4.90                         |
| <i>SD</i>        | 0.90                            | 0.68                | 1.14                         |
| % correct        | 82.14                           | 45.24               | 61.30                        |
| Older children   |                                 |                     |                              |
| <i>M</i>         | 3.76                            | 2.57                | 5.76                         |
| <i>SD</i>        | 0.62                            | 0.87                | 1.51                         |
| % correct        | 94.04                           | 64.28               | 72.02                        |

Thirteen children in the experimental group and 9 in the control group answered the final listening question, after the movement activity, with the correct response of "choppy." Of these 22 children, only 8 were correct in performance of their movement response, while 3 of the 8 children who answered "smooth," which was incorrect, did perform with correct movements.

## DISCUSSION

The results of this study clearly demonstrate that preschool children are capable of learning to identify and label musical characteristics with very few errors after a brief program of instruction. In examining the cell data (see Table 2), it is clear that the experimental group did very well on the single discrimination items (only 2 errors in 60 responses), making 23% more correct responses than the control group. Neither group was particularly successful with the double discrimination items, however, with both responding near the chance level, and only a 10% advantage for the experimental group. The dramatic difference in the experimental group's single versus double discrimination scores should particularly be noted. These data appear to be consistent with Piaget's theory – that young children tend to center on only one aspect of a stimulus at a time, and that the ability to decenter is acquired through development and experience rather than through instruction.



Only a few children in this study were successful in demonstrating double discrimination in the movement activity, and these responses did not seem to be dependent on participation in instruction or the ability to provide the correct verbal label. These results are similar to those reported by Sims (1988), and they reinforce the conclusion that most young children are not ready for, nor receptive to, music listening tasks requiring decentration responses.

The results of the control group are consistent with the body of research that has revealed that children do not express musical ideas well verbally (Flowers, 1984; McMahon, 1986; Petzold, 1981; Sims, 1986; Young, 1982). The success with which the experimental group *did* learn to do this, however, indicates that young children are receptive to learning and using language to describe music. The experimental group children were able to retain the terminology during the three days between instruction and testing. It would be informative for future research to determine the length of retention for this type of learning. Although the experimental group children had successfully performed songs and movements that used the combination of characteristics during instruction, they were not able to provide the correct labels during the posttest. While these results seem consistent with child development theory, the possibility that children could successfully demonstrate, rather than label, two musical characteristics simultaneously warrants further investigation, and will be examined in Study 2, which follows.

## STUDY 2

The purpose of Study 2 was to replicate and extend the first study, in part to investigate the speculation that children might be able to demonstrate combined musical elements through a nonverbal response mode, in this case, singing. Another purpose of Study 2 was to determine whether results similar to those of Study 1 would be obtained in a more usual large group instructional setting, as opposed to the small groups used in the first study. Of particular interest was the previous study's finding that the children learned to label concepts in a single discrimination task very easily. Although children's difficulty in providing verbal labels for musical concepts has been well documented in the research literature, most of the studies in this area did not incorporate instruction into the research design. Replication of the results in a more typical preschool classroom situation would strengthen the implications for early childhood music instruction.

One new aspect considered was whether the children could make use of two elements at a time while singing a familiar song in a purely imitative rote manner. Elicited imitation has been cited as a valuable research tool in the field of language development (Dale, 1972; Slobin & Welsh, 1971). According to Slobin and Welsh (1971), "in repeating a sentence, one must filter it through one's own productive system" (p. 176). An intriguing transfer to the music task may be that if children are able to imitate the desired characteristics, they are demonstrating through production an initial stage of discrimination or comprehension. A second, more sophisticated production behavior of interest was whether the children could sing a familiar song "on command" incorporating two characteristics specified by the teacher.

The singing response mode was of interest in the context of this replication for an additional reason. In the first study, the simultaneous discrimination task at which the children were not successful required them to answer questions in response to recordings of orchestral or piano music. One possibility, investigated in the current study, is that it may be easier for children to identify simultaneous musical characteristics in a version of a song they have just sung that exemplified those characteristics. It is possible that direct experience and interaction with the musical material might facilitate analysis and discrimination on a task that seems to be difficult for preschool children.



## METHOD

The subjects for this study were children enrolled in three classrooms at two preschools that are comparable to the preschools that provided subjects for Study 1. Although 56 children participated in one or more of the four instructional sessions, only the children in attendance for all of the sessions were considered subjects for the study, resulting in a total  $N$  of 42. The children ranged in age from 2 years 9 months to 5 years 11 months, with an average age of 4 years 3 months. Given the wide range of ages represented in this study, the effect of age on the activities tested here was of interest. In order to investigate possible age differences, two age-groups, younger ( $n = 21$ ,  $M$  age = 3 years 6 months), and older ( $n = 21$ ,  $M$  age = 5 years 0 months), were designated.

In the previous study, the effects of instruction were investigated using an experimental group/no-contact control group design, and results indicated that students receiving instruction were significantly and substantially better at performing the discrimination tasks than were control group students. Since instruction was not to be a variable in the present study, all subjects in this study participated in the instructional procedures so that they would be familiar with the terminology and types of activities.

Unlike the first project, in which instruction took place in small groups, subjects in the current study received instruction in their regular classes. During this study's four 20-minute instructional sessions, 14 to 22 children were present in each class, with class size varying due to absences. This was considered to represent a more typical setting for preschool music activities than was the small-group setting. The lesson plans used were the same as those taught in the earlier study.

All subjects were posttested individually by the experimenter within 3 to 5 days after the completion of instruction. As in the first study, the experimenter administered the test and was generally positive and encouraging but provided no specific feedback or reaction to the students' responses.

The first component of the test was the same listening task used in the previous study, consisting of short recorded musical excerpts, two of which the children were asked to label as "fast or slow," two as "smooth or choppy," and two as both "smooth or choppy AND fast or slow." The second and third components of the test were singing tasks and were designed specifically for the current project. All sung responses were tape recorded for subsequent analysis. Sung responses were of interest only as they related to the concepts being examined, and were not analyzed for pitch or rhythmic accuracy. The song "Rain, Rain, Go Away" was used for both singing tasks because it was short, familiar, and easy to sing. Throughout the lesson sequence, it had been sung by the children using the contrasting elements and their combinations. The song was reviewed with each subject prior to its use as a test activity.

One singing activity was designed to assess the children's ability to incorporate the elements studied into a song. Subjects were asked to sing the song four times, using each combination of fast/slow with smooth/choppy (e.g., "Please sing 'Rain Rain' for me fast and choppy"). Four orders of the paired elements were constructed and assigned to subjects at random to control for possible order effects. The method of analyzing these responses is described in the Results section.

For the second singing task, the children were asked to sing the song with the experimenter. They were instructed to "sing 'Rain Rain' with me, just the way that I sing it." The only cues available to the children were from the musical rendition itself, as the descriptive terminology was *not* used verbally to identify how the song would be performed. Since the children were expected to join in and match the model, their responses were imitative in nature. This process will be operationally defined as "simultaneous imitation," as opposed to echo-style imitation. A simultaneous format was selected so that the effects of memory inherent in an echo format would be minimized. Again as a control, four different orders of the combined characteristics were used, and assigned to subjects at random. After completing each version, the child was asked whether the song had been fast or slow, and smooth or choppy. There were two purposes of this activity: (1) to determine whether the children would be able to demonstrate the discrimination and vocal control necessary to match the combination of characteristics used by the model, and (2)



to determine whether the children could provide labels to describe characteristics of their own singing.

## RESULTS

A two-factor, repeated measures analysis of variance (ANOVA) was used to compare the mean scores on the single and double discrimination listening task for the older and younger age-groups. Results indicated that older children scored significantly higher than the younger children [ $F(1, 40) = 13.43, p < .05$ ], and other scores were significantly higher for single versus double discrimination items [ $F(1, 40) = 61.21, p < .05$ ], with no significant interaction (see Table 3). These results, particularly those of the older children, are quite comparable to those obtained in Study 1 with respect to means and percentages correct (see Table 2).

There were four items on the posttest that required the children to sing with the experimenter and then identify both the fast/slow and smooth/choppy characteristics of the version sung. This resulted in eight "double discrimination after singing" items. No significant difference was found between age groups for these items,  $t = 1.16, p > .05$  (see Table 2). All children approximated the tempo and style of articulation of the model when singing with her.

Table 3

*Summary Table: Repeated-Measures ANOVA Comparing Younger and Older Children on Single and Double Discrimination.*

| Source of variance      | SS    | df | MS    | F      |
|-------------------------|-------|----|-------|--------|
| <i>Between subjects</i> | 32.29 | 41 |       |        |
| Age                     | 8.06  | 1  | 8.06  | 13.43  |
| Residual                | 24.23 | 40 | 0.60  |        |
| <i>Within subjects</i>  | 62.00 | 42 |       |        |
| Discrimination          | 37.34 | 1  | 37.34 | 61.21* |
| Age x Discrimination    | 0.41  | 1  | 0.41  | 0.67   |
| Residual                | 24.25 | 40 | 0.61  |        |
| <i>Total</i>            | 94.29 | 83 |       |        |

\* $p < .05$ .

The fourth component of the posttest was the "sing on demand" section. Subjects were asked to sing "Rain Rain" four times, each with a different combination of the two sets of musical elements being studied. Some children would not sing the song alone or stated that they could not sing it in the manner requested and refused to try. Therefore, data will be reported for only the 27 children who completed all four versions. These subjects included 63% of the younger children,  $n=12$ , and 65% of the older children,  $n=15$ .

Upon listening to the tape-recorded sung responses, it immediately becomes apparent that the children were not successful at incorporating the characteristics requested into their singing. The most informative way to analyze these responses was to examine the four responses made by each child, to determine what, if any, relationship existed among each set of responses. No child's performance reflected the correct relationships for all four versions. About half (14) of the subjects sang the song almost exactly the same way all four times. Of the remaining 13 subjects, only 4 correctly incorporated the contrast between smooth and choppy, and another 2 approximated a correct relationship between tempos. The responses of the remaining 7 children reflected changes among versions, but these were incorrect changes, all related to dynamic level. Three children sang louder in response to the "fast" instruction, two sang louder when the request to sing



"choppy" was included in the instructions, and two remaining children sang louder for one or two examples, but without a pattern evident.

To explore the tempo responses further, each response was timed to the nearest second. The number of beats per minute was calculated based on the 8 beats constituting the song (8 beats times 60 seconds divided by number of seconds of singing). These results should be considered to be estimates only, however, because many performances did not maintain a steady beat. The average beats per minute for each of the four versions were approximately: fast/smooth, 127; fast/choppy, 104; slow/smooth, 105; slow/choppy, 102. Inspection of these data reveal little systematic variation among tempos, other than for the fast/smooth version, which was performed about 25% faster than the other versions.

## DISCUSSION

One purpose of Study 2 was to replicate the teaching and testing procedures used with the experimental group in Study 1. As presented in Table 2, the single discrimination listening responses of the "older" group in Study 2 ( $M$  age = 5 years 0 months) were quite comparable to the responses of the slightly younger experimental group in the previous study ( $M$  age = 4 years 6 months). Study 2's "younger" group ( $M$  age = 3 years 6 months), while scoring a bit lower than the experimental group in Study 1, scored higher than the Study 1 no-instruction control group, even though the average age of the control group was one year older ( $M$  age = 4 years 6 months). These data, while further reflecting age differences in discrimination ability, also seem to suggest that single discrimination instruction can be very effective for 4- and 5-year-olds, and moderately effective even with 3-year-old-children. It is also conceivable that the large-group setting proved less effective for the youngest children.

The subjects were almost equally unsuccessful in responding to the double discrimination listening task in both studies, although the group of children with the highest average age (Study 2 "older" group) performed a bit better than the other three groups. It may be concluded that Study 2 generally did replicate the results of Study 1. The children in both studies performed single discriminations quite successfully after instruction, but most responded at near the change level on double discrimination items.

This replication strengthens the prior conclusion that most young children are not ready for music listening tasks requiring attention to more than one element at a time, and supports the speculation that the ability to do this may be acquired through development and experience rather than through direct instruction. The age differences described in this study and the substantially higher scores of the older children versus younger on the double discrimination listening task provide additional support for this conclusion. Teachers should be aware that preschoolage children tend to focus on one element during music listening tasks, perhaps to the exclusion of others. Activities during listening experiences, such as movements directly related to a specific musical characteristic (Sims, 1986), might be planned to ensure that the students are attending to the *desired* element or contrasts. Based on the replication of the success with which the children learned to apply terminology to single discriminations, and the fact that this occurred in the context of both large- and small-group classroom settings, it may be concluded that the early years, which are a time of rapid language acquisition, would seem to be an ideal time to begin building children's music vocabulary and discrimination skills.

The effectiveness of singing to facilitate children's ability to demonstrate decentration may be examined by comparing the scores on the listening and singing double-discrimination labeling tasks. On the listening task, the younger children made correct responses 45% of the time, and the older, 64%. After singing a version of the song that incorporated two of the elements, however, younger children were able to label what they had just sung with 61% accuracy, and older children were correct 72% of the time. While the "after singing" scores might be considered only moderately successful, they are substantially more successful than the listening-only format, particularly for the younger children. One reasonable conclusion based on this comparison may be that children's ability to make discriminations based on their own performance is developmental in nature, and is evidenced earlier than their ability to make discriminations in more abstract listening situations. It also may be true that even if they hadn't been singing themselves, the a cappella melody would



have elicited a higher number of correct responses than the stimuli consisting of more complex orchestral and piano pieces. Familiarity with the musical stimulus also may have been a factor affecting this task. Further research on the effects of stimulus complexity and familiarity on young children's discrimination responses might provide teachers with valuable methodological and sequencing information.

The children were all able to approximate the characteristics combined by the model in the singing of a familiar song even though they were not as successful in labeling these characteristics. Their ability to do this reflects their ability to make the appropriate discriminations, albeit at an initial level. This finding reinforces the validity of child centered teaching methodologies that advocate an "experience first" approach to learning, and corroborates a number of studies that have suggested that children may perceive more than they can label (McDonald & Simons, 1989). It seems that although it might not be an effective use of a teacher's time to try to teach double discriminations directly to preschool children, a valuable foundation for future discrimination learning might be laid by incorporating double discrimination experiences into singing activities. Research to explore children's ability to combine characteristics in movement would help determine if movement could also provide a productive methodology for early interaction with combined characteristics.

The preschool-age children in this study were not successful with the task that required them to sing a familiar song using two specified characteristics. There was some indication that "fast" in combination with "smooth" was the easiest characteristic for the children to understand and/or produce, possibly because that is the most "normal" way of singing from among the four combinations presented. Since the children could incorporate two characteristics into the simultaneous imitation activity, it seems that their difficulty with the "sing on demand task" was not the result of a vocal control problem, but rather was another reflection of the difficulty young children have in responding to verbal descriptors of music concepts. Because these data were obtained from the relatively small number of students who attempted all four versions ( $n = 27$ ), these results should be viewed with some caution, although the number of the remaining children who indicated that they "did not know how" to sing in the manner requested actually seems to lend support to these conclusions. The possibility that preschool-age children might be successful in demonstrating one specified characteristic at a time in singing was not examined in this study, but would be a valuable area for further research.

The subjects in this study demonstrated discrimination through simultaneous imitation in singing with more success than through their independent production of singing responses. As mentioned previously, researchers in language acquisition have found that elicited imitation can be a valuable research tool. This may be true for research in music education, as well, as a child's imitation ability may provide an initial nonverbal indication of the child's perception or production ability. Further research to explore the relationship between imitation and production abilities, as well as the usefulness of both simultaneous and echo-style imitation as assessment tools in music, seems warranted.

Based on the results of these two studies, as well as research and theory cited above, a tentative sequence of development for the acquisition of music concept discrimination in listening and singing might be identified. It is speculated that the ability to produce sung responses imitating a specific musical characteristic is demonstrated first, followed by the ability to label the characteristic in the child's own performance and then in listening examples, with the ability to apply the specified concept to the production of a song developing last. In each case, the ability to make specific single discrimination responses would most likely precede the ability to make discriminations requiring the combination of these concepts or musical elements. This hypothesized sequence has the potential to contribute to the formulation of a developmental framework for the acquisition of music discriminations or concepts, and warrants further investigation. The exploration of the validity of these ideas certainly provides a wealth of possibilities for future research.

Given the attention currently being focused on musical experiences during the early childhood years as important to future musical development, there is still relatively little known about how children actually develop music skills and concepts. Determining sequences of development, as well as the types of activities and instruction that might be the most appropriate or effective for promoting musical growth and understanding in young children, should be an important goal for researchers in music education.



## REFERENCES\*

- Bartholomew, D. (1987). Problems with Piagetian conservation and musical objects. *Bulletin of the Council for Research in Music Education*, 93, 2740.
- Dale, P. S. (1972). *Language development*. Hinsdale, IL: Dryden Press.
- Flavell, J. H. (1977). *Cognitive development*. Englewood Cliffs, NJ: Prentice-Hall.
- Flowers, P. J. (1984). Attention to elements of music and effect of instruction in vocabulary on written descriptions of music by children and undergraduates. *Psychology of Music*, 12, 17-24.
- Gagne, R. M. (1977). *The conditions of learning*. New York: Holt, Rinehart, & Winston.
- Greenberg, M. (1976). Research in music in early childhood education: A survey with recommendations. *Bulletin of the Council for Research in Music Education*, no. 45, 1-20.
- Hargreaves, D. J. (1986). *The developmental psychology of music*. Cambridge, England: Cambridge University Press.
- McDonald, D. T., & Simons, G. M. (1989). *Musical growth and development: Birth through six*. New York: Schirmer.
- McMahon, O. (1986). Implications of recent research into aspects of music in early childhood. *International Society for Music Education Yearbook*, 13, 161-164.
- Petzold, R. G. (1981). Child development. In *Documentary report of the Ann Arbor Symposium* (pp. 42-48). Reston, VA: Music Educators National Conference.
- Scott, C. R. (1987). Reaching the young child through music. *International Music Education*, 14, 11-20.
- Serafine, M. L. (1980). Piagetian research in music. *Bulletin of the Council for Research in Music Education*, 62, 1-21.
- Shuter-Dyson, R., & Gabriel, C. (1981). *The psychology of musical ability*. London: Methuen.
- Sims, W. L. (1986). The effect of high versus low teacher affect and passive versus active student activity during music listening on preschool children's attention, piece preference, time spent listening, and piece recognition. *Journal of Research in Music Education*, 34, 173-191.
- Sims, W. L. (1988). Movement responses of pre-school children, primary grade children, and pre-service teachers to characteristics of musical phrases. *Psychology of Music*, 16, 110-127.
- Slobin, D. I., & Welsh, C. A. (1971). Elicited imitation as a research tool in developmental psycholinguistics. In C. S. Lavatelli (Ed.), *Language training in early childhood education* (pp. 170-185). Urbana: University of Illinois Press.
- Taebel, D. K. (1974). The effect of various instructional modes on children's performance of musical concept tasks. *Journal of Research in Music Education*, 22, 170-83.
- Young, L. P. (1982). An investigation of young children's music concept development using nonverbal and manipulative techniques. *Dissertation Abstracts International*, 43, 1345A. (University Microfilms No. DA8222205).



Zimmerman, M. P. (1984). The relevance of Piagetian theory for music education. *International Journal of Music Education*, 3, 31-34.

Zimmerman, M. P. (1986). Music development in middle childhood: a summary of selected research studies. *Bulletin of the Council for Research in Music Education*, 86, 18-35.

---

\* As referências bibliográficas deste artigo estão conforme o documento original.

### 2.4.3. Estudo Crítico – Por Christopher Johnson

#### Objectivos

A finalidade desta pesquisa foi examinar as relações possíveis entre o comportamento musical das crianças e a teoria do desenvolvimento. O primeiro estudo testou especificamente os efeitos da instrução na capacidade das crianças indicarem a percepção e discriminação de conceitos musicais. O segundo estudo procurou determinar se as respostas não verbais poderiam produzir diferenças em termos de discriminação. Este estudo estendeu o projecto a um contexto mais natural (i.e. a instrução foi proporcionada num grupo mais alargado).

#### Design

O primeiro estudo foi de natureza experimental, tendo sido constituídos dois Grupos para o efeito: um Experimental e outro de Controle. Ao Grupo Experimental foi ministrada instrução em pequenos grupos; enquanto o Grupo de Controle apenas foi submetido à realização de um Pós-Teste. O segundo estudo foi também de cariz experimental, tendo os sujeitos sido divididos em dois grupos: um constituído por crianças mais novas e outro por crianças mais velhas. Neste caso, a instrução foi ministrada em grandes grupos.

#### Revisão do artigo

#### Introdução e Revisão

Sims inicia a sua exposição defendendo a necessidade de compreender as competências, conhecimentos e preferências musicais das crianças. Depois, justifica a necessidade crítica de um esclarecimento dos efeitos de curto e longo termo da instrução. Esta necessidade é mais especificamente direccionada para um interesse na discriminação de características musicais dicotómicas específicas. A autora faz então uma revisão de dados e teorias presentes na bibliografia, relativa à ordem pela qual os estudantes adquirem e desenvolvem conceitos musicais, fazendo notar que a música contém em simultâneo vários elementos, mesmo quando estamos concentrados em aspectos individuais do estímulo.

De seguida, Sims discute dados da investigação que referem a tendência das crianças (em idade pré-escolar) para centrarem a sua atenção num elemento particular, em vez de se "descentrarem", contrapondo diversos elementos em simultâneo. A autora relaciona este fenómeno com as teorias de conservação de Piaget<sup>2</sup> e o estágio de desenvolvimento dos alunos, que neste caso corresponderá ao período pré-operatório. Uma das características deste estágio é, precisamente, a de as crianças centrarem a sua atenção num elemento único, sem se darem conta de outros aspectos que decorrem ao mesmo tempo.

A autora conclui, argumentando ser importante saber mais sobre este tema, que poderá ser relevante para o aperfeiçoamento das teorias do desenvolvimento e ter repercussões sobre as técnicas e materiais utilizados no ensino das crianças.



## Estudo 1

### Método

Os participantes nesta pesquisa foram 30 crianças, com idades compreendidas entre os 48 e os 61 meses, provenientes de três infantários com características semelhantes. Estas foram divididas de forma aleatória em dois grupos. O Grupo Experimental recebeu instrução musical, enquanto o Grupo de Controlo foi apenas submetido a um Pós-Teste.

A Variável Independente deste estudo foi constituída por um breve período de instrução musical, planeada para ensinar as crianças a reagir a duas características musicais em simultâneo. Os dois emparelhamentos escolhidos para esta circunstância foram: *rápido/lento* e *calmo/agitado*. Enquanto o primeiro par envolve *tempo* (uma área largamente estudada), o segundo compreende o estilo (área esta menos estudada que a primeira). A investigadora definiu operacionalmente cada termo da lição de maneira a evitar qualquer ambiguidade que confundisse os alunos. Todos os exemplos estudados eram diametralmente opostos nas suas características. Os estímulos usados foram previamente mostrados a estudantes universitários, tendo a identificação das características destes estímulos merecido o seu acordo absoluto.

As crianças do Grupo Experimental foram divididas em grupos de três e quatro sujeitos e submetidas a aulas com a duração de 20 minutos. As metodologias utilizadas nas aulas incluíram actividades como canto, movimento corporal, audição e explicação teórica. Foram utilizadas três Variáveis Dependentes, administradas individualmente às crianças, com a duração de 8 minutos.

Na medição da primeira Variável Dependente pediu-se aos alunos para ouvirem e classificarem excertos musicais. Os primeiros dois excertos foram classificados ou como rápidos ou como lentos. Os dois seguintes foram testados quanto ao estilo: calmo ou agitado. Nos excertos finais pediu-se aos alunos que classificassem os excertos usando combinações destas quatro características.

A segunda variável incluiu movimento corporal em combinação com a audição musical, tendo as respostas sido aferidas a partir da observação do grau de concordância dos seus movimentos em relação ao excerto musical apresentado.

A terceira variável consistiu na apresentação do mesmo excerto utilizado no segundo teste, tendo-se pedido às crianças que se pronunciassem quanto ao estilo do excerto ouvido.

### Resultados do Estudo 1

Os resultados da primeira avaliação foram obtidos a partir da utilização da análise da variância de uma medida repetida (ANOVA). Os dados obtidos indicaram uma diferença significativa em duas áreas: (1) o Grupo Experimental obteve melhores resultados do que o Grupo de Controlo e (2) as crianças foram mais eficazes quando lhes foi solicitada a discriminação de uma característica individualmente do que quando tiveram que identificar duas características combinadas do mesmo excerto. Não se verificou qualquer interacção entre estas duas variáveis.

Na apreciação a partir do movimento corporal, foram em número semelhante os alunos de cada grupo que fizeram o seu movimento corporal corresponder aos excertos ouvidos na situação de discriminação de duas características combinadas. No entanto, 60% dos sujeitos de cada grupo fizeram-no de forma incorrecta, indiciando que as aulas não tiveram influência significativa na execução da tarefa.

No terceiro teste, em que os alunos tiveram que identificar características de estilo na mesma secção da peça utilizada no teste anterior, foi possível verificar um grau mais elevado de eficácia, tendo sido obtidas vinte e duas respostas correctas. No entanto, quando estas respostas verbais foram comparadas com as do teste de movimento corporal, constatou-se que apenas oito coincidiram nas duas ocasiões. Concomitantemente, três das oito crianças que responderam verbalmente de forma errada, aplicaram correctamente os movimentos aos excertos ouvidos. Presume-se assim que estes testes não terão sido muito fiáveis e que algumas das respostas terão sido fruto de escolhas feitas ao acaso.



## Discussão do Estudo 1

O primeiro aspecto realçado por Sims é que a Variável Independente – instrução específica de curta duração – demonstrou claramente ter tido repercussões na capacidade de caracterização de aspectos dicotómicos da música, quando estes são apresentados de forma independente. O Grupo Experimental apenas cometeu dois erros em sessenta respostas dadas. O Grupo de Controle foi 23% menos preciso.

Embora o Grupo Experimental tenha cumprido melhor a tarefa de discriminação de características combinadas, nenhum dos grupos obteve resultados que nos garantam não se tratarem de respostas ao acaso. Sims indica que a falta de capacidade para discriminar características combinadas poderá apoiar a teoria Piagetiana, segundo a qual, as crianças que atravessam o estágio pré-operacional, quando estão a realizar uma dada actividade, se centram em apenas um aspecto específico do estímulo.

O restante debate proposto por Sims centra-se no movimento e subsequente categorização da parte do excerto que exigia a discriminação de características combinadas. A autora reflecte então sobre certos aspectos das respostas que levaram a resultados inconclusivos e que determinaram a realização do segundo estudo.

## Estudo 2

### Método

No segundo estudo, Sims modificou alguns aspectos presentes no estudo anterior. Primeiramente modificou a estrutura do Grupo Experimental. No Estudo 1 tínhamos presente um Grupo Experimental e outro de Controle. Neste segundo estudo, as crianças foram divididas em dois grupos com idades diferentes. Visto estar já claramente demonstrado que a discriminação de duas características musicais opostas pode ser ensinada, neste estudo este factor foi mantido, tendo todos os participantes sido sujeitos ao mesmo tipo de instrução.

A autora também modificou ligeiramente a Variável Dependente. Embora o conteúdo se tenha mantido, a forma de aplicação foi modificada. Enquanto no primeiro estudo os alunos receberam instrução em pequenos grupos de 3 ou 4 elementos, no presente estudo os alunos foram enquadrados de forma mais comum, tendo recebido formação em conjunto, numa turma. As avaliações das Variáveis Dependentes também foram de alvo de modificações. O primeiro teste foi mantido de forma idêntica ao utilizado no Estudo 1, mas os outros dois, dada a sua inconsistência, foram totalmente readaptados.

Assim, neste estudo a segunda tarefa para os estudantes foi constituída pela entoação de uma canção familiar. A argumentação da autora baseou-se no facto de considerar que, se as crianças conseguem copiar aspectos dos elementos musicais alvo, predefinidos e presentes no exemplo utilizado, está-se, conseqüentemente, perante a demonstração da fase inicial de discriminação.

A terceira tarefa também foi alterada no sentido de levar as crianças a verbalizarem, categorizando uma melodia com os dois pares de características dicotómicas referidas anteriormente. Porém, desta feita, as crianças caracterizavam a canção que tinham acabado de cantar, categorizando-a a partir de aspectos que iam sendo alterados no processo de memorização/ imitação do estímulo musical reapresentado.

Acrescentou-se ainda um quarto aspecto ao Pós-Teste. Neste caso fez-se com que as crianças cantassem uma canção que lhes era familiar, quatro vezes consecutivas. De cada uma das vezes, foi solicitado aos sujeitos que incorporassem na sua interpretação diferentes elementos dos pares dicotómicos. Por exemplo, dando-lhes directrizes como: «Desta vez canta lento e agitado».



## Resultados do Estudo 2

Os resultados deste estudo apresentam semelhanças e diferenças relativamente aos do primeiro estudo.

O grupo constituído pelos alunos mais velhos foi melhor sucedido no Pré-Teste, quer na discriminação de características consideradas individualmente como na discriminação de características combinadas. Todos os sujeitos obtiveram melhores resultados na discriminação de características consideradas individualmente. Já no que toca à tarefa de discriminação de duas características combinadas, as respostas parecem não ter sido mais do que ocasionais.

As diferenças ocorreram nos dois aspectos seguintes do Pós-Teste: (1) todos os sujeitos foram capazes de cantar imitando as características (tempo e estilo) apresentadas pelo investigador – isto veio demonstrar que a primeira fase da discriminação de características combinadas estava concretizada; (2) depois, quando lhes foi pedido para verbalizarem as características da canção que acabavam de cantar, foram melhor sucedidos que na tarefa auditiva do Estudo 1. Nesta situação, não foram encontradas diferenças de relevância nos dois grupos etários.

O último aspecto relativo ao Pós-Teste dizia respeito à incorporação dos pares de características musicais anteriormente referidos na interpretação vocal das suas canções. Os resultados mostraram que muitos alunos se recusaram a participar nesta actividade. Dos que tentaram, nenhum foi capaz de interpretar a canção incorporando correctamente as quatro variantes de combinações das características.

## Discussão do Estudo 2

No primeiro parágrafo da discussão de resultados, Sims ilustra o bom desempenho dos dois grupos do segundo estudo, por oposição ao Grupo de Controlo do primeiro estudo. Embora o grupo de alunos de faixa etária mais baixa não tenha sido tão eficiente como o Grupo Experimental do primeiro estudo, este era constituído por crianças significativamente mais novas, tendo superado o Grupo de Controlo do primeiro estudo, constituído por crianças 12 meses mais velhas. Ficou assim evidenciado que, se na tarefa de discriminação de características individuais existe alguma questão relacionada com aspectos do desenvolvimento, ela foi superada através da instrução. Ambos os grupos deste segundo estudo foram menos eficazes na discriminação de elementos combinados na situação de audição, embora o grupo constituído pelos alunos mais velhos tenha sido o grupo mais bem sucedido dos quatro grupos que participaram em ambos os estudos.

Sims afirma que este estudo reforça as conclusões anteriores, segundo as quais as crianças não estão habilitadas a fazer discriminações de duas características em simultâneo. Apesar de não ousar esclarecer se o factor em falta é do âmbito do desenvolvimento ou da experiência, a autora conclui que a instrução direccionada para a realização desta tarefa não é suficiente para melhorar aquela capacidade.

Um aspecto que se destaca é o de verificar que a eficiência das crianças foi significativamente melhor quando a tarefa foi proposta depois de cantarem em vez de depois de ouvirem. Esta observação reforça muitos outros estudos que indicam que as pessoas aprendem com a aplicação prática e não pela simples exposição à informação. Embora muitos outros factores possam estar implicados na melhoria da discriminação – como, por exemplo, a simplicidade da melodia – o aspecto activo da participação dos sujeitos parece ser um elemento chave.



## Contribuição para a Área de Estudo

Esta investigação proporciona uma série de contribuições para a educação musical. Em primeiro lugar, mostra que os alunos podem ser ensinados a identificar elementos separados numa performance musical. Porém, crianças de faixa etária mais baixa parecem não ter experiência, desenvolvimento ou outras especificidades, que lhes permitam identificar duas características musicais em simultâneo. Pelo menos, não apenas através de instrução especificamente direccionada para a realização desta tarefa.

A segunda contribuição importante deste estudo envolve os diversos testes realizados e o modo como estes se relacionam com um possível modelo de desenvolvimento de discriminações mais complexas. Sims especula afirmando que «fica demonstrado que em primeiro lugar surge a capacidade para imitar características musicais específicas; depois, segue-se a capacidade de identificar um determinado elemento na sua própria performance e só depois em situações de audição; finalmente, surge a capacidade de aplicar estes conceitos específicos a uma canção que é produzida independentemente».

Esta é a base de uma teoria do desenvolvimento da aquisição da discriminação musical. Embora este estudo envolva crianças de tenra idade, afigura-se como pertinente a aplicação do mesmo modelo (ou de outro similar) a alunos mais velhos. A autora apresenta assim um trabalho que merece ser continuado em futuras investigações.

## Comentário

Este artigo inicia-se da melhor forma e desenvolve-se com qualidade. A introdução e revisão foram exemplarmente elaboradas. Sims redigiu este ensaio de forma clara e lógica, destacando o objecto de investigação e as teorias da área de estudo que com ele se relacionam. Não há dúvida de que a autora conhece a bibliografia envolvida e que identifica claramente a pertinência dos presentes estudos.

A Metodologia do primeiro estudo identifica com clareza o assunto a desenvolver e procedimentos do projecto. Embora a planificação das lições não tenha sido completamente descrita, torna-se evidente que se quisermos replicar este estudo, esta informação será prontamente disponibilizada. Foi muito importante a definição dos termos utilizados no estudo, escrita com grande proficiência. Finalmente, a explicação relativa à medição das Variáveis Dependentes permite ao leitor saber exactamente o que foi testado e como se executou o processo.

Os resultados são claros e concisos. As duas tabelas apresentadas são informativas, e o texto segue de perto a estrutura estabelecida na metodologia. Os testes estatísticos são integralmente explicados, sendo as tabelas facilmente inteligíveis e úteis para a compreensão dos resultados.

O único aspecto da primeira secção de discussão de resultados que parece sair do contexto geral, é o facto de Sims se precipitar na atribuição da falta de sucesso dos alunos na tarefa de discriminação de características combinadas justificando que esta se deve a questões de desenvolvimento. É necessário ter em consideração que, neste projecto, é solicitado às crianças que realizem complexas tarefas de análise musical. Assim sendo, pode haver aqui uma curva de aprendizagem que ainda não se sobrepôs à prática e à experiência. A autora não reconhece totalmente que a falta de experiência pode estabelecer correspondência directa com a falta de desenvolvimento. O Grupo Experimental teve melhores resultados neste aspecto do que o Grupo de Controlo, o que pode indiciar que, com mais treino, estes sujeitos poderiam ter resultados significativamente melhores do que os obtidos por acaso. Nenhum professor de formação musical esperaria que os seus alunos, com tão pouca preparação e experiência, executassem com sucesso um teste de formação auditiva. Assim sendo, será que se poderia esperar resultados diferentes dos obtidos, quando nos encontramos em presença de tão complexa tarefa?

O segundo estudo é plenamente justificado na secção que sucede o primeiro. Nesta fica claro para o leitor o que a autora se propõe modificar e a argumentação para tais alterações.



Contudo, apenas se torna evidente no final do artigo o fundamento para que as opções tomadas sustentem uma possível teoria para o desenvolvimento da discriminação musical. Um aspecto adicional, que não ficou completamente esclarecido, é o de ter sido acrescentada a quarta secção dos testes da Variável Dependente, não se percebendo em que parte do estudo foi aplicada. É o segundo aspecto proposto no método, mas aparece como quarto nos resultados. Já que este é um aspecto de grande importância no final da investigação, seria bom que tivesse sido convenientemente explicitado.

A discussão do Estudo 2 é excelente. Sims utiliza os dados para relacionar os dois estudos entre si e confronta-os também com trabalhos de outros investigadores. O seu aconselhamento aos professores é notável, propondo a aplicação directa e imediata, dos conhecimentos obtidos ao contexto de sala de aula. A autora revê todos os aspectos da investigação e resultados, guiando-nos pela sua argumentação especulativa.

Este é um artigo de referência. A autora teve extremo cuidado com todas as suas considerações, sendo os presentes projectos perfeitamente lógicos nas suas pretensões, procedimentos e conclusões. As ideias e opções são meticulosamente apresentadas. A investigadora demonstrou ter confortável mestria na abordagem da área de estudo e do processo científico. O leitor fica com a sensação de que este projecto foi realizado com absoluto rigor. Um dos aspectos mais positivos deste artigo é a teoria apresentada. Fica claro que Sims conhece aprofundadamente a literatura existente sobre a educação musical na infância e o desenvolvimento infantil em geral. Começa o seu relatório com fundamentação teórica rigorosa e termina-o com a *sua* própria teoria (se bem que não a designe como tal) sobre como as crianças aprendem a discriminar.

O mais importante é que esta teoria se fundamenta nos dados recolhidos. Muitas das teorias descritas em revistas da especialidade são pouco mais do que idealizações concebidas pelos autores. Embora bem construídas, estão muitas vezes fundadas no vazio e desfasadas da realidade. Por seu lado, este artigo mostra uma progressão clara de pensamento entre uma teoria e outra mais aprofundada, baseada em dados científicos concretos. Estamos perante uma apresentação verdadeiramente consistente.

---

N.T.

1- Note-se que no original se encontram as designações *smooth* e *choppy* que seriam mais apropriadamente traduzidas por "regular" e "irregular". Em sua vez, optou-se por traduzir aqueles termos por "calmo" e "agitado", uma vez que estes termos estarão mais próximos do vocabulário das crianças que constituem os sujeitos desta investigação e nos parecem adequados no presente contexto.

2 – Jean Piaget estabeleceu uma teoria do desenvolvimento cognitivo, em que todas as crianças se desenvolvem de acordo com um conjunto de estádios que se sucedem segundo uma ordem predeterminada. Estes estão divididos por patamares etários, com características específicas diferenciadas. Nesta investigação, os sujeitos encontram-se claramente no período pré-operatório que, segundo Piaget, vai dos 2 aos 7 anos de idade e se caracteriza, entre outras coisas, pelo desenvolvimento da linguagem, o pensamento simbólico e egocentrismo.

# 2.5.

## THE EFFECT OF MUSIC LISTENING ON PHYSIOLOGICAL RESPONSES OF PREMATURE INFANTS IN THE NICU

EFEITOS DA AUDIÇÃO MUSICAL NAS RESPOSTAS FISIOLÓGICAS DE BEBÉS PREMATUROS NUMA UNIDADE DE CUIDADOS INTENSIVOS

J.W. CASSIDY & J.M. STANDLEY (1995)

JOURNAL OF MUSIC THERAPY, XXXII (4), 208 - 227

### 2.5.1. Resumo do artigo / Abstract

#### Resumo

Participam neste estudo vinte bebés prematuros na sua primeira semana de vida. Estes bebés tinham entre vinte e quatro e trinta semanas de tempo de gestação, apresentavam peso inferior à média estavam a ser oxigenados numa UNCI (Unidade Neonatal de Cuidados Intensivos). Todos os bebés foram sujeitos a um ABR (*Auditory Brainstem Response*), teste auditivo para assegurar que as respostas audiológicas eram consistentes com uma audição normal. Ao longo da experiência, dez dos bebés ouviram canções de embalar através de auscultadores Bio-Logic ALGO Ear Couplers TM [auscultadores especiais utilizados em UNCI, n.t.] enquanto que os outros dez serviram de Grupo de Controle. O procedimento experimental ocorreu durante três dias e seguiu um plano do tipo ABABABABA: cinco segmentos de silêncio (A), alternando com quatro segmentos de música (B). Em ambas as condições (silêncio ou música), cada segmento tinha a duração de quatro minutos. Os dados foram recolhidos durante este procedimento, tendo sido registados em cada minuto os níveis de saturação de oxigénio, a frequência cardíaca, a frequência respiratória e o número de episódios de apneia/bradicardia. Obteve-se assim uma linha de base de dados correspondente aos 36 minutos da aplicação das condições do tratamento. Os resultados indicaram que a música não era contra-indicada na primeira semana de vida destes bebés prematuros, para os quais a estimulação sensorial é habitualmente restrita. De facto, a música teve efeitos positivos nos níveis de saturação do oxigénio, na frequência cardíaca e na frequência respiratória. Não foi observado qualquer aumento de episódios de apneia/bradicardia na sequência do tratamento musical.

#### Abstract

In this study 20 low birthweight infants of 24-30 weeks gestation age, who were being oxygenated in a Neonatal Intensive Care Unit (NICU), served as subjects during their first week of life. Ten infants listened to lullabies through Bio-logic insert earphones with ALGO Ear Couplers and 10 infants served as control subjects. All subjects passed an auditory brainstem response (ABR) procedure to insure that audiological responses were consistent with normal hearing. Experimental treatment occurred across 3 days and was conducted in an ABABABABA design, with data collected during five segments of four minutes each of silence alternating with four segments of four minutes each of music. Oxygen saturation levels, heart rate, respiratory rate, and number of apnea / bradycardia episodes were recorded once per minute for the duration of baseline and treatment conditions (36 minutes). Results indicated that music was not contraindicated in the first week of life for these very low birthweight infants for whom sensory



stimulation is usually restricted. In fact, music had noticeably positive effects on oxygen saturation levels, heart rate, and respiratory rate. No increase in apnea I bradycardia episodes following music treatment were observed.

---

The authors wish to express appreciation to the staff at Woman's Hospital in Baton Rouge, Louisiana for their support in this project. In particular, we would like to thank Steven B. Spedale, MD for his guidance throughout the project and Karen M. Ditty, CCC-A for her invaluable help with equipment and testing. This study was supported in part by a Summer Faculty Research Stipend awarded by the Council of Research at Louisiana State University.

## 2.5.2. Artigo Original

# The Effect of Music Listening on Physiological Responses of Premature Infants in the NICU

Jane W. Cassidy & Jayne M. Standley (1995)

Singing lullabies is an established custom within our culture to soothe, comfort, and pacify newborns. Research studies have documented these soothing effects and have shown that the typical neonate is an active recipient of these and other audiological stimuli at all stages of development, even prenatally (Standley & Madsen, 1990). All prolonged auditory stimuli with a regular rhythm (lullabies, metronome, heartbeat) seem to pacify normal infants equally well. Initially, it was believed that heart beat was a superior calming stimulus for infants due to claims by Salk (1960), but this theory has been rejected by more contemporary research (Detterman, 1978).

It has been learned that one to two day old newborns attend to and discriminate auditory stimuli (Blass, 1987), learn to suck differentially for the preferred stimulus of their mother's voice (DeCasper & Fifer, 1980), quiet to sustained auditory sounds (Brackbill, Adams, Crowell, & Gray, 1966), and recognize stories that were read to them in the womb (Kolata, 1984). Polverini-Rey (1992) demonstrated that fetuses not only respond to external auditory stimuli, but learn from and remember them long after birth. In this study lullabies played daily during the nine weeks prior to birth differentially decreased the crying of the newborns at four weeks of age.

Recently, neonatology personnel have become interested in the use of music to reduce the distress of fragile, premature infants (White-Traut & Tubeszewski, 1986). Such infants often demonstrate a high risk for complications and a failure to thrive (Leonard, 1993; Oehler, 1993; Standley, 1991). Their behavioral state affects their physiological well being. Infant distress is indicated by frequent sleep disruptions, easily elicited startle responses, frequent agitation and crying, and, subsequently, decreased blood oxygenation levels (Gordin, 1990). Some premature infants are so fragile that even the approach of medical rounds to their isolette causes increased bradycardia and apnea episodes (Gorski, Davison, & Brazelton, 1979).

Despite their neurological immaturity and overresponsiveness to all stimuli, premature infants are necessarily treated in isolettes with average, sustained decibel levels of 70-80 dB. Intermittent increases of 15-20 dB due to equipment alarms, personnel voices, and knocks on the isolette to arouse apneic infants are common (Gottfried & Hodgman, 1984). It has been assumed that sustained environmental auditory stimulation at this decibel level is aversive and distressing to the premature infant. Some research has investigated whether music played loudly enough to mask these aversive auditory stimuli might function to reduce distress and help infants thrive. Caine (1992) demonstrated that lullabies played between 70-80 dB in the NBICU (Newborn Intensive Care Unit) isolettes three times per day significantly reduced the total length of hospital stay by an average of 5 days, Lorch, Lorch, Diefendorf, and Earl (1994) demonstrated that infants 33 to 35 weeks post conception showed differential physiological responses to music heard at 78 dB.

Aversive auditory stimuli are especially detrimental to oxygenated, low birthweight infants since respiratory regularity and subsequent oxygen saturation levels are directly affected by the infant's behavioral state and degree of pain (Gordin, 1990). Keeping oxygen saturation levels high with the lowest amount of oxygen possible is an important medical objective since extended reliance on high levels of oxygen can cause retrolental fibroplasia, permanent damage to the infant's eyes (Mullins, 1986). Additionally, weaning from use of a respirator to unaided breathing, usually accomplished as quickly as possible, is particularly stressful for the infant, and is usually accompanied by extensive distress symptoms (Field, 1987).



Research has now begun investigating the effects of music listening on physiological responses of premature infants. Collins and Kuck (1991) observed 17 premature infants (gestational age 24-37 weeks) who were within the first two weeks of their life. The infants were stimulated to an agitated state and observed for 10 minutes during silence followed by music combined with intrauterine sounds for 10 minutes. Results showed that the babies' oxygen saturation levels were significantly increased by this one 10-minute trial of auditory stimulation, though arterial pressure and heart rate were not significantly affected. This study utilized no control group, however, so it cannot be determined whether music or the passage of time alone accounted for the infants becoming calm and having increased oxygen saturation levels.

Standley and Moore (1993) conducted a similar study to determine the effectiveness of music versus newborns' preferred auditory stimulus, their mother's voice (DeCasper & Fifer, 1980).

Effects were also measured across several days. Two groups of NBICU low birthweight infants heard the selected auditory stimuli for 20 minutes for each of three days via earphones intended to focus the auditory stimuli. Effects on oxygen saturation levels and frequency of Oximeter alarms were observed. Results showed differential responses to the two auditory stimuli as time progressed. On the first day, music infants had somewhat higher oxygen saturation levels, but these effects disappeared by Days 2 and 3 of the study. Infants hearing music had fewer occurrences of Oximeter alarms than did those hearing their mothers' voice. It was also noted that the very premature infants (gestational age prior to 30 weeks) seemed to have increased episodes of apnea and bradycardia following music stimulation. It was theorized that perhaps the duration of the music or the decibel level via earphones was overstimulating these most fragile infants. Results of this study support the suggestion that determining age protocols, decibel levels, and presentation schedules for music stimulation of the premature infant is an important clinical consideration (Schaefer, Hatcher, & Barglow, 1980).

Prior research in this area raises the following questions. Is there a minimum gestational age criterion for therapeutic benefit from auditory stimulation? It is known that the human cochlea forms by the 20th week of gestation. It has been observed that the fetus actively responds to external auditory stimuli from the 26th week (Marlowe, 1982) with consistent responses present after 28 weeks (Birnholtz & Benacerraf, 1983). Auditory brainstem responses are obtained as early as 26 weeks gestation (Cox, Martin, Carlo, & Hack, 1993). How does the newborn of only 26-30 weeks gestational age respond to music stimulation intended to mask aversive environmental auditory stimuli?

Would an intermittent schedule of music presentation reduce overresponsiveness, increase awareness, and prolong any calming effects during the first week of life for very low birthweight infants? Though evidence suggests that environmental factors influence dendrite growth in the brain, research has not yet isolated how early in the premature infant's development any auditory stimuli might be utilized therapeutically (Campbell, 1985).

The purpose of the current study was to ascertain the effect of an intermittent schedule of music stimulation (4 minutes on/4 minutes off) via Bio-logic insert earphones and ear couplers on physiological responses of 24 to 30 week gestational age, oxygenated infants in their first week of life.

### Procedures

Twenty premature infants served as subjects in this study. Infants were patients in a 24 bed Level III Neonatal Intensive Care Unit (NICU) designed to provide a consistently supportive and nurturing environment that promotes the optimal developmental outcome for each infant. Minimum stimulation is a standing requirement for all infants in this unit. All infants were observed during the first week (but not during the first 24 hours) of life. All infants were receiving prophylactic, ototoxic drugs. Once infants were off assisted ventilation, an auditory brainstem response (ABR) procedure was completed by a certified audiologist (CCC-A) to insure that auditory responses were consistent with normal hearing. Prior to data collection, parents or guardians of each child were given a written and oral description of the purpose and procedures for the study. Equipment to be used during the listening treatment was demonstrated and parents received the opportunity to listen to



the music. At this time, permission for each child to be a subject in the research project was solicited and obtained.

Criteria for selection of subjects included being of 24 to 30 weeks gestational age, weighing less than three pounds, and receiving oxygen to assist breathing. Seventy-four babies meeting these criteria were identified over the eleven month data collection period. Infants who were removed from ventilation during the course of data collection or were sedated during the 4 hours prior to data collection on any given day ( $n = 34$ ) were dropped from the study. Infants were not included if they were sedated due to the impact of such medications on their physiological responses and behavioral state (Paap & Nahata, 1990). Other infants were not included because parents did not sign the consent form ( $n = 8$ ), they were deemed medically unstable by hospital personnel ( $n = 6$ ), a researcher was unavailable for three consecutive days during the first week of life ( $n = 5$ ), or the ABR test was failed ( $n = 1$ ). Of the remaining infants, 10 listened to recorded music and 10 served as control subjects. Groups were balanced with regard to gestational age, chronological age, weight, and sex.

Subjects were individually observed for three consecutive days while on radiant warmers. In order to keep conditions constant and to accommodate the regular routines of hospital personnel, observations were made exclusively during the hours of 3 p.m. and 6 p.m. daily. No other interventions were performed during the observation periods.

The music stimulus tape used in this project was 30 minutes in duration and consisted of six commercially recorded original lullabies sung by a female vocalist with orchestral accompaniment. The selections included *Drift Away*, *Rhapsody in Baby Blue*, *Nightlights*, *I'll Love You Forever*, *Dreamship*, and *Loving My Baby Goodnight*.

Auditory stimuli for experimental subjects were presented binaurally via a battery-operated Sony Walkman (WM-D3) stereo cassette recorder placed outside of the radiant warmer and connected to Bio-logic insert earphones with an insert earphone cable which was placed alongside the radiant warmer. The tips of the insert phones were then attached to ALGO Ear Couplers. These sterile, disposable couplers, which serve the same purpose as adult headphones, consist of a hard flat shell which is outlined on the inner side with foam rubber. The outside of this foam rubber is coated with a sticky substance which, when placed against the head of an infant, seals around the pinna. Secure placement of these couplers reduces background noise by 14 dB SPL at 2000 Hz (Kileny, 1988). The music was played at an average 80 dB SPL to mask aversive auditory stimuli and to facilitate attentiveness (Caine, 1992; Lorch et al., 1994). At each initiation and cessation of music, sounds were faded in or out over two seconds to avoid causing sudden extreme changes in auditory stimuli.

Experimental treatment was conducted in an ABABABABA design, with data collected during five segments of four minutes each of silence alternating with four segments of four minutes each of music. Data were collected on control subjects for 36 minutes as well.

Data collection occurred once per minute for the duration of baseline and treatment. Data included oxygen saturation levels (estimation of the amount of oxygen carried by the blood) read from an Ohmeda Biox 3700 Pulse Oximeter and heart rate and respiratory rate read from a Dyna Scope DS 3300 monitor. Data were also recorded on the number of apnea (a pause in breathing) and bradycardia (extreme decrease in heart rate) episodes lasting 10 seconds or more as charted by nurses in the 24 hour period prior to and following observations.

## Results

The purpose of this study was to observe physiological responses of severely premature infants under music listening conditions as compared to normal environmental sounds. Every minute for 36 minutes during three consecutive days data were collected on oxygen saturation levels, heart rate, and respiratory rate. Also noted from the medical charts were instances of apnea and bradycardia for the 24 hours preceding and following data collection.

Important to the use of music with premature infants whose neurological development is incomplete is whether or not auditory stimuli are too stimulating and therefore contraindicated. Contraindications based on physiological responses could include continuous large differences between mean rates with music subjects being less healthy or less stable than control subjects and



greater variability among subjects in the music group during data collection. Therefore, both means and standards deviations were compared between groups and across time.

A two-way ANOVA with repeated measures (treatment group by minute) was calculated on each day for each dependent measure. Of these nine statistical tests there was only one result indicating statistically significant differences ( $p < .05$ ) due to treatment, either as a main effect or interaction. On Day 1 there was a significant difference due to treatment for the dependent measure of oxygen saturation [ $F(1, 18) = 7.7, p = .0125$ ]. As is clearly seen in Figure 1, infants in the music condition consistently had higher (more preferable) oxygen saturation levels ( $M = 95.8\%$ ) than infants in the control group ( $M = 93.8\%$ ). Means for both groups were well within an acceptable clinical range.

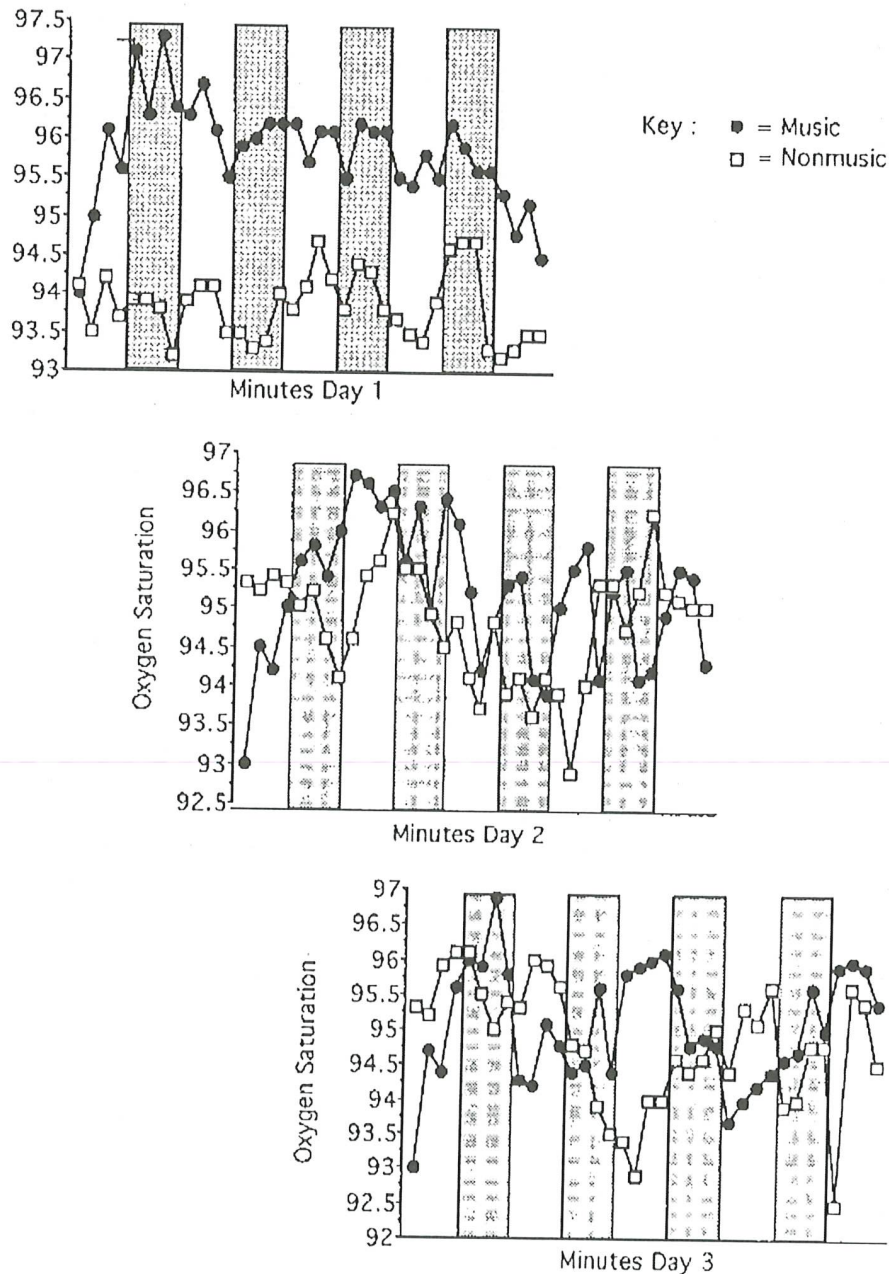


FIGURE 1.  
Mean oxygen saturation levels across minutes and days.

Close inspection of means and standard deviations across dependent measures, days, and minutes reveals further information. Mean oxygen saturation levels for the 20 subjects ( $n = 10$  per group) are displayed in Figure 1. Each of the three days is represented by a graph which defines the 36 minutes of data collection. Shaded areas of the graphs indicate 4-minute intervals during which the experimental subjects listened to music.

It is clear on Day 1 that from the second minute of data collection, experimental subjects experienced a dramatic increase in oxygen saturation levels. This increase was exaggerated during the first episode of music listening, then remained high and stable across the remaining minutes. On Days 2 and 3 the levels were similarly variable for both groups, however, the same trend of dramatic increase for music subjects at the start of data collection each day was apparent. During the majority of time intervals throughout the study, music subjects had higher oxygen saturation levels than did nonmusic subjects.

Comparisons of standard deviations between groups both within each minute and across the 36 minutes on each day for oxygen saturation indicate no differential effect due to treatment. Inspection of graphs in Figure 2 indicate that one group is not consistently more homogeneous at each minute nor more stable across the duration of data collection.

Mean heart rates by group across 36 minutes for each day are displayed in Figure 3. Again, shaded areas of the graphs indicate music listening intervals for the experimental subjects. On Day 1, all subjects were similar during the first baseline interval. Once the music conditions began, the subjects listening to music experienced a dramatic decrease in heart rate that remained lower than that of control subjects for the majority of subsequent time intervals on Day 1. The opposite was apparent on Days 2 and 3 with music subjects experiencing predominantly higher heart rates than nonmusic subjects.

Standard deviations across the 36 minutes for heart rate are presented in Figure 4. Deviations for the music group are clearly consistently lower and more stable than for the control group after the initial baseline on Day 1 and Day 2 of data collection. The differences between groups are less obvious on Day 3.

Graphs of respiratory rate means are displayed in Figure 5. On Day 1, from minutes 4 to 15, experimental infants experienced lower respiratory rates than did control infants. This became less clear during the last 21 minutes. On Day 2, the mean rates were primarily lower for the experimental infants and their rates were much more stable across the 36 minutes than were the rates of control infants. On Day 3 both groups experienced similar variability in respiratory rates (although the range is smaller than that of the previous days).



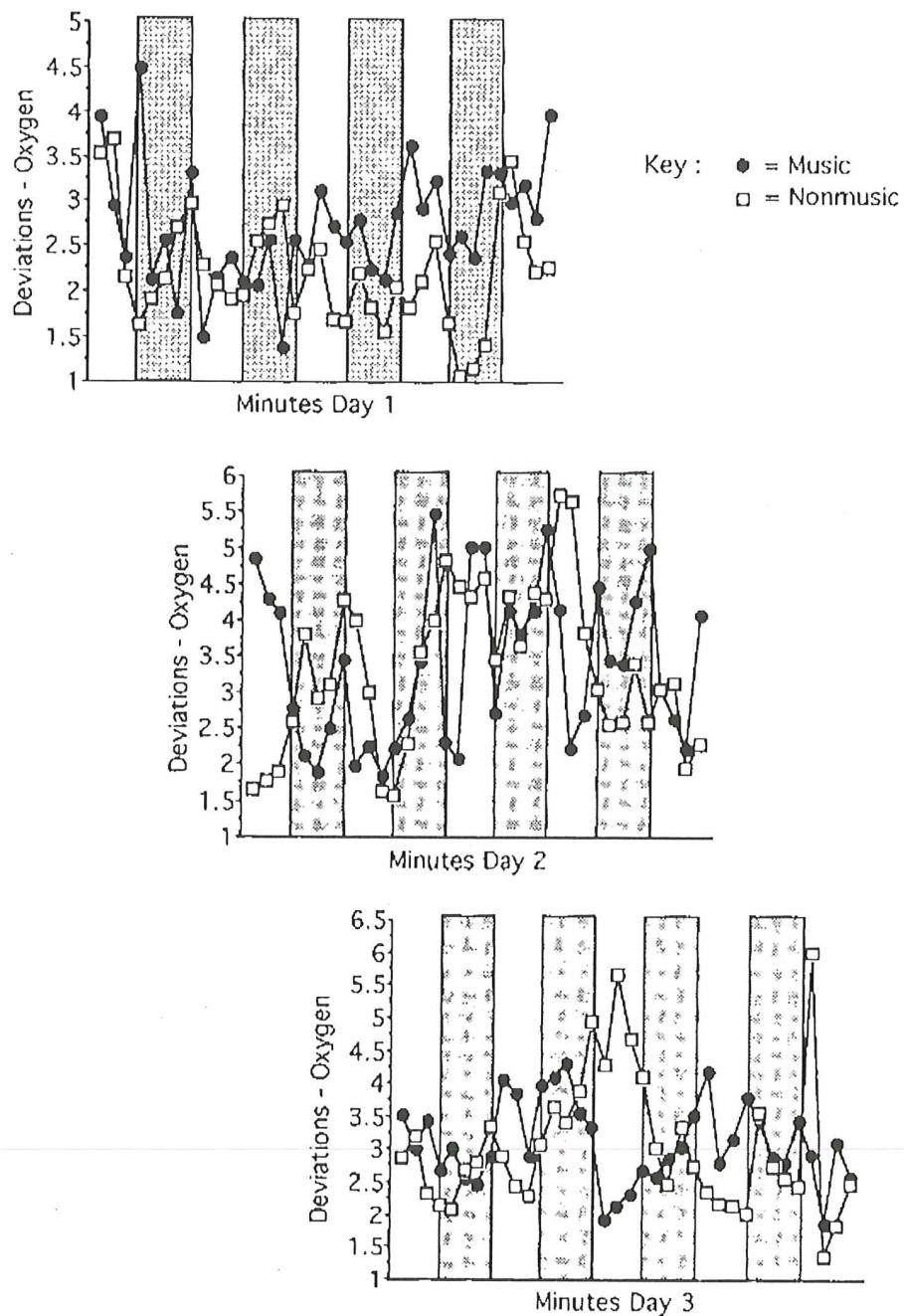


FIGURE 2.  
Standard deviations of oxygen saturation levels across minutes and days.

Graphs of standard deviations for respiratory rate are displayed in Figure 6. On Day 2 music subjects had consistently lower and more stable deviations than control subjects from the first minute of baseline. On the other days there was no clear difference.

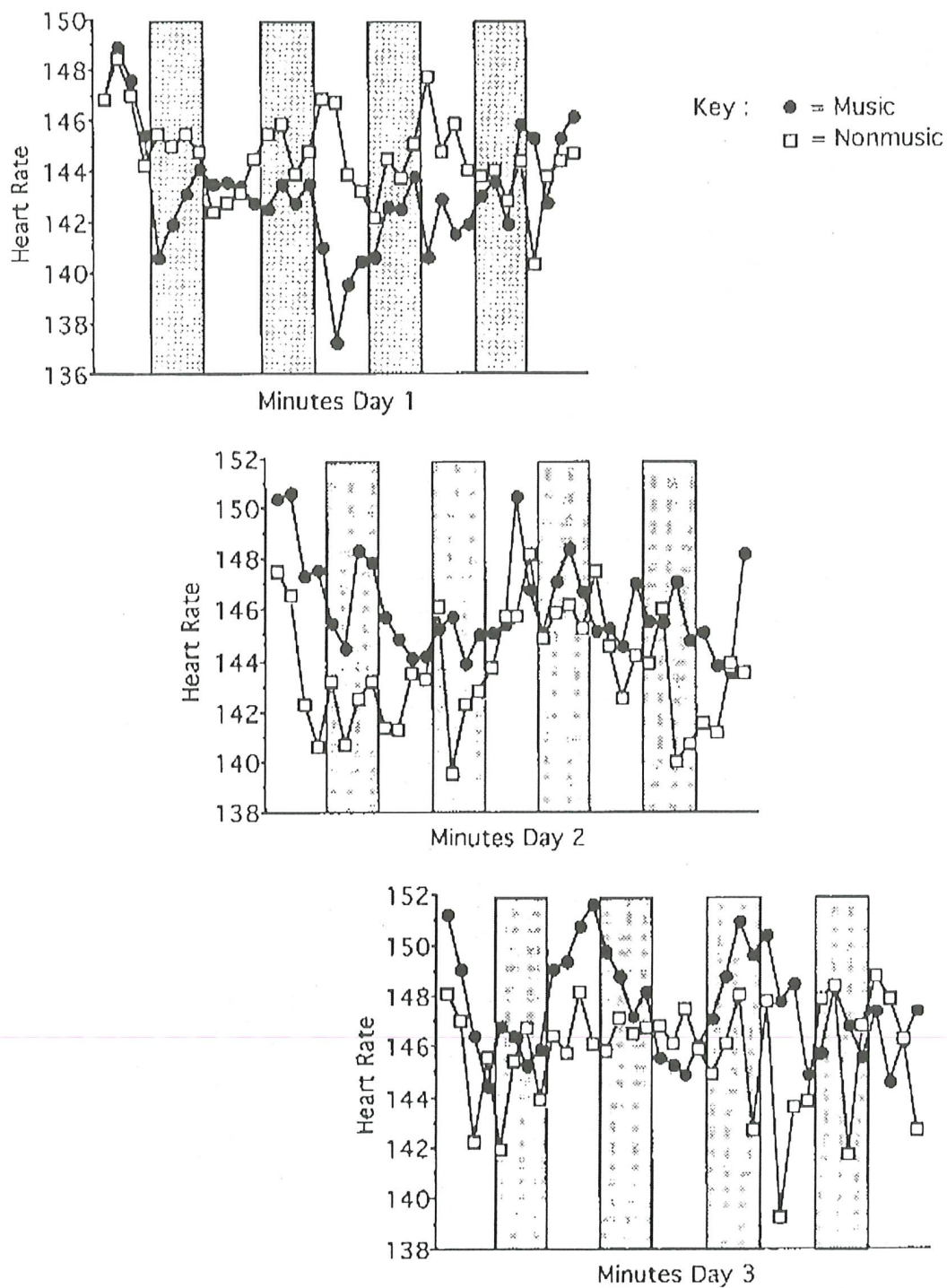


FIGURE 3.  
Mean heart rates across minutes and days.

Information from the infants' medical charts is presented in Table 1. Included in this table are the sex of each infant, gestational age (40 weeks is considered full term), weight in grams on first day of data collection (approximately 28 grams is equal to one ounce), day of infant's life on the first day of data collection, and the number of instances of apnea (A) and bradycardia (B) for the 24 hours preceding each day of data collection and following the last day. The groups are



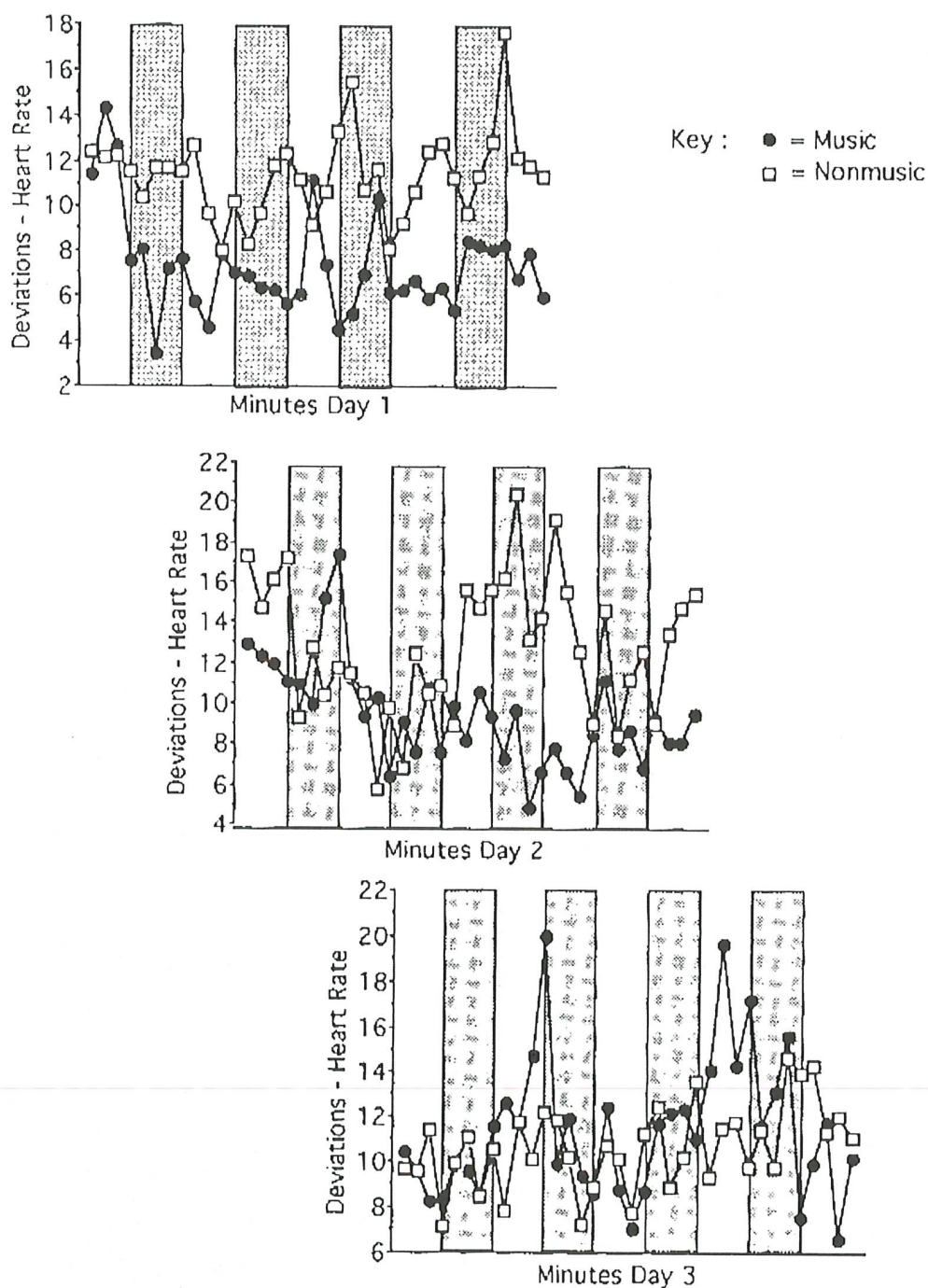


FIGURE 4.  
Standard deviations of heart rates across minutes and days.

similar in the mean gestational age, birth weight (slightly more than one ounce difference between means), and day of life on first day of data collection. Mean instances of apnea/bradycardia are similar across days and between groups. There was little difference between groups in apnea/bradycardia levels with three music versus two nonmusic subjects who experienced more than two episodes of bradycardia in a given 24 hour time span.

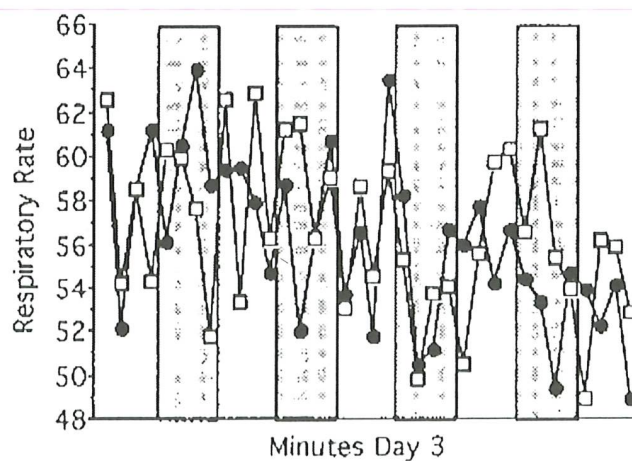
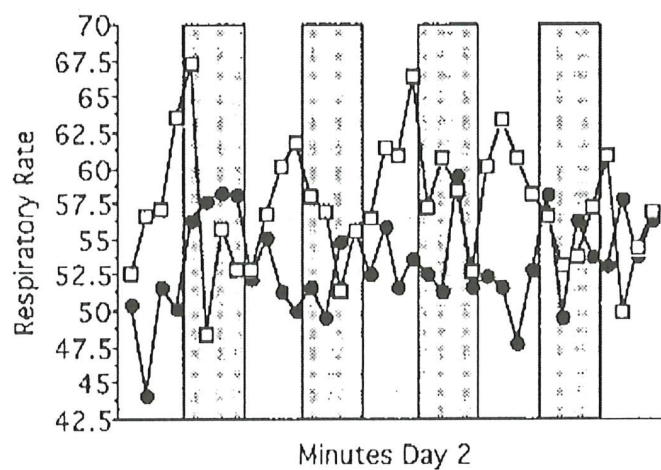
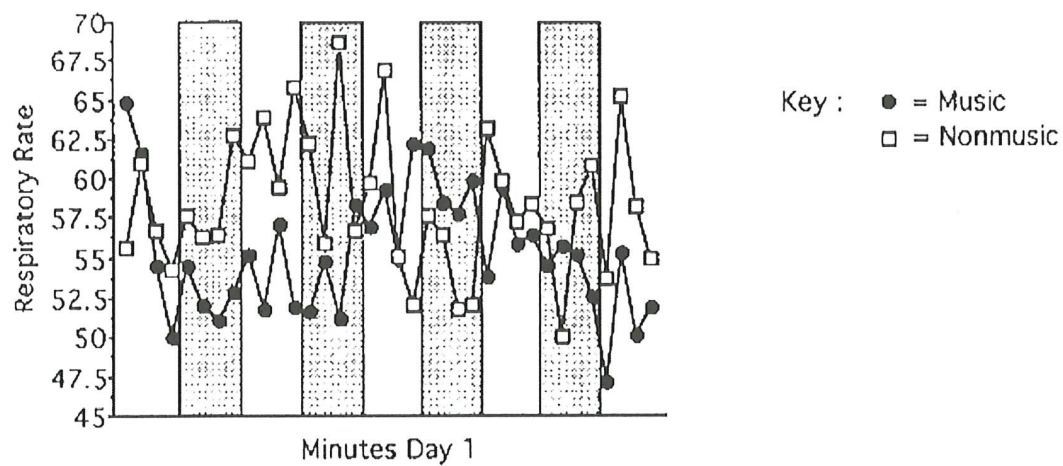


FIGURE 5.  
Mean respiratory rates across minutes and days.



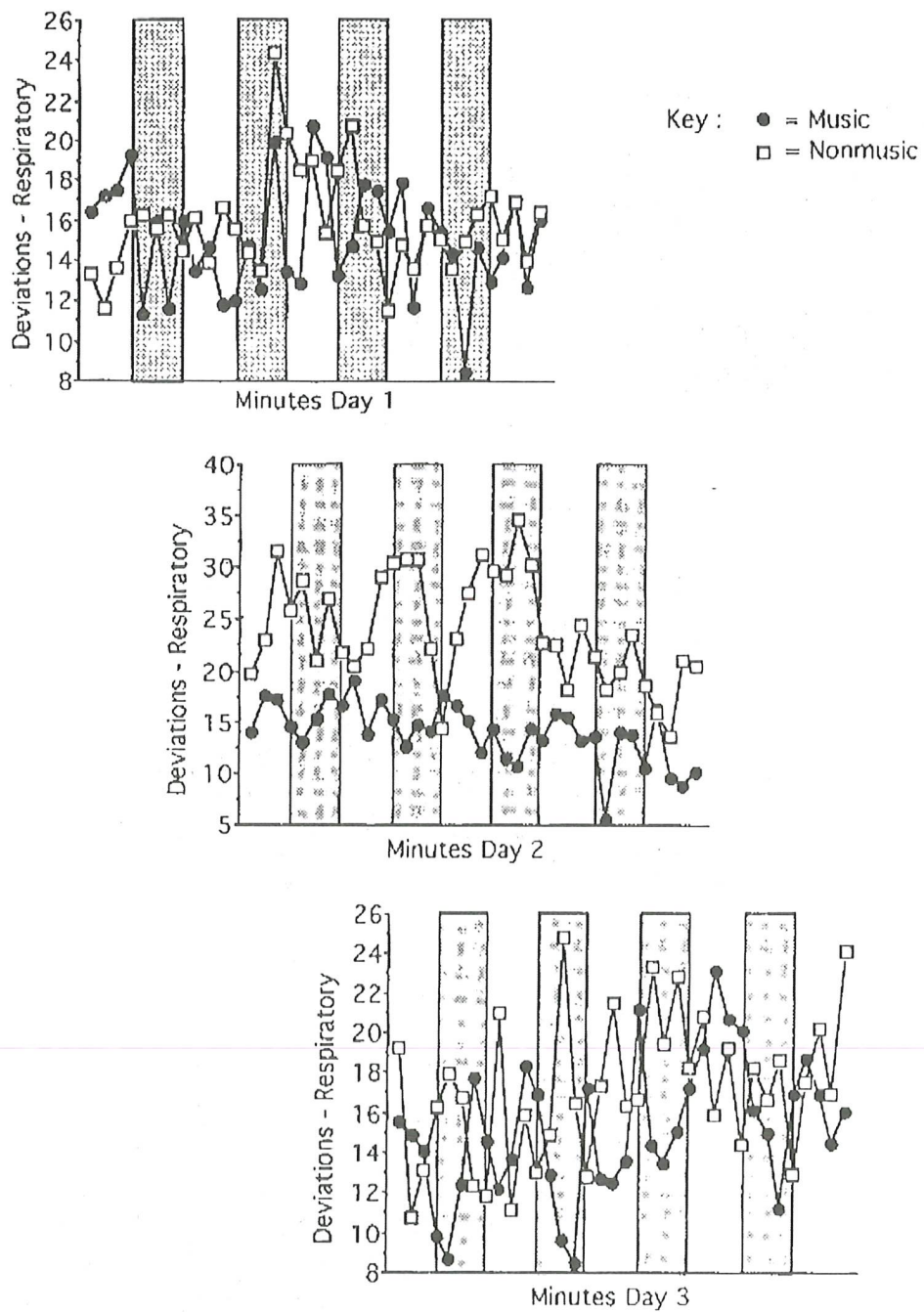


FIGURE 6.  
Standard deviations of respiratory rates across minutes and days.

Instances of apnea were almost nonexistent since the majority of subjects were receiving machine assisted ventilation. Subjects who experienced apnea were those who were being weaned from machine assisted ventilation.

**Table 1**  
***Instances of Apnea / Bradycardia and Demographic Data***

| <b>Music Subjects</b> |                   |                |                                    |               |               |               |                |
|-----------------------|-------------------|----------------|------------------------------------|---------------|---------------|---------------|----------------|
| Sex                   | Gest. Age (weeks) | Weight (grams) | Age on 1 <sup>st</sup> Day of Data | A/B Pre Day 1 | A/B Pre Day 2 | A/B Pre Day 3 | A/B Post Day 3 |
| F                     | 29                | 1,100          | 3                                  | 0             | 0             | 0             | 0              |
| M                     | 27                | 720            | 3                                  | 0             | 0A/1B         | 0             | 0A/1B          |
| M                     | 27                | 875            | 3                                  | 1A/1B         | 0             | 2A/3B         | 0              |
| F                     | 28                | 855            | 5                                  | 0             | 0             | 0             | 0              |
| M                     | 28                | 1,242          | 2                                  | 0             | 0             | 0             | 0A/2B          |
| M                     | 28                | 1,175          | 2                                  | 0             | 0             | 0             | 0              |
| M                     | 27                | 950            | 2                                  | 0             | 0             | 0             | 3A/3B          |
| M                     | 27                | 880            | 3                                  | 0             | 0             | 0             | 0              |
| F                     | 24                | 545            | 4                                  | 0             | 0A/1B         | 0A/4B         | 0A/2B          |
| F                     | 27                | 1,015          | 2                                  | 0             | 0             | 0             | 0              |
| Mean                  | 27.2              | 938.4          | 2.9                                |               |               |               |                |

| <b>Nonmusic Subjects</b> |                   |                |                                    |               |               |               |                |
|--------------------------|-------------------|----------------|------------------------------------|---------------|---------------|---------------|----------------|
| Sex                      | Gest. Age (weeks) | Weight (grams) | Age on 1 <sup>st</sup> Day of Data | A/B Pre Day 1 | A/B Pre Day 2 | A/B Pre Day 3 | A/B Post Day 3 |
| F                        | 24                | 555            | 4                                  | 0             | 0             | 0             | 0              |
| F                        | 28                | 925            | 5                                  | 0             | 0A/1B         | 6A/9B         | 5A/7B          |
| M                        | 29                | 1,325          | 3                                  | 0             | 0A/2B         | 4A/7B         | 0A/1B          |
| M                        | 27                | 998            | 2                                  | 0             | 0             | 0             | 0              |
| M                        | 26                | 900            | 3                                  | 0             | 0             | 0             | 0              |
| M                        | 28                | 810            | 3                                  | 0             | 0             | 0             | 0              |
| F                        | 30                | 1,200          | 3                                  | 0             | 0             | 0             | 0              |
| F                        | 26                | 645            | 3                                  | 0             | 0             | 0A/1B         | 0              |
| F                        | 26                | 910            | 2                                  | 0             | 0             | 0             | 0              |
| M                        | 26                | 925            | 2                                  | 0             | 0             | 0             | 0              |
| Mean                     | 27                | 919.3          | 3.0                                |               |               |               |                |

## Discussion

Music immediately and positively affected oxygen saturation levels, heart rate and respiration rate on Day 1. Acclimation to the music stimuli resulted in more minimal effects on these physiological responses on Days 2 and 3. These findings are consistent with prior research in this area. This study demonstrated no negative effects on apnea/bradycardia. It also clarifies several issues for clinical application of music in the NICU.

The most important information for music therapists from these data may be that music listening is not contraindicated in the first week of life for even these very fragile premature babies. In fact, when there were consistent differences between groups, music subjects had more medically acceptable responses and were more stable. Inspection of standard deviations gives added clinical importance to these results. From a homeostasis perspective (referring to the ability of the physiological system to maintain internal stability) it is clinically important that there appears to be less variability among music subjects on the three dependent measures and the effect of music does not deteriorate across days as quickly or as much as is apparent from the means. Under peaceful sleeping conditions, neonates exhibit higher and more stable oxygen saturation levels, slower and more stable respiratory rates, and more stable heart rates (although the actual heart rate is dependent on numerous variables including gender and medication). To the extent that these trends are noticeable in these data, music appears to relax and comfort these usually irritable infants.

As in the Standley and Moore study (1993) it seems that music has its greatest effect on the first day of presentation. After the novelty has worn off, effects appear less noticeable as perhaps these infants are acclimating to their normally noisy environment by "turning off" sound



perception. Since it appears that music is not contraindicated, it may be one tool to nurture bonding under conditions where new parents are unable to hold, feed, bathe, and in many cases touch their critically ill infants. In providing tapes of favorite music and a small tape recorder (with carefully controlled duration of stimuli, tempos, and dynamic levels), parents can experience some control over their child's environment. These infants also may respond positively long term to the calming influence of familiar music as did the infants in the Polverini-Rey study (1992) who, as fetuses with similar gestational ages to the infants in this study, listened to lullaby music.

Unlike the findings of Standley and Moore (1993), comparable infants in this study did not experience an increase of apnea and bradycardia following the cessation of music. This may be due to a larger sample of severely premature infants in the current study. Some infants, due to myriad medical problems, experience higher rates of apnea/bradycardia. However, there were two other major differences between the studies: the presentation schedule for music stimulation and the mode of presentation.

Perhaps in the previous study the 20-minute duration of the music listening experience was overstimulating for some of these sensitive infants. As a rule, all critically premature infants are placed on minimal stimulation during the early weeks of life, however, environmental auditory stimulation is much more difficult to control in an intensive care unit than other sensory stimuli. Therefore, one might expect that presentation of music would, at the very least, be no more stimulating than a normal environment and at best would be soothing, be predictable, and would mask other environmental sounds. In this study, reducing the total music listening time to 16 minutes and presenting it in 4-minute segments alternating with 4 minutes of silence did not appear to negatively affect instances of apnea and bradycardia. Perhaps a tape constructed such that each piece was faded in and out and rest time was provided between each selection would be the most appropriate stimulus. This issue of controlling music duration as well as tempo and dynamic changes to avoid overstimulating very low birth weight infants is certainly one that music therapists should consider carefully in clinical NICU applications.

In the present study, ear cups attached to audiological insert phones were used to present the music stimulus to infants. With this equipment we were insured that each infant was being presented the stimulus at the same carefully controlled decibel level and that the decibel level was comfortable for the infant. In previous studies music has been presented via phonopad earphones (Standley & Moore, 1983) and free field (Caine, 1992; Collins & Kuck, 1991; Lorch et al., 1994). While phonopad earphones and ear cups serve the same function, different considerations must be given to the stimuli presented via each piece of equipment.

Phonopad earphones are positioned over each ear in the fold of the infant's knitted cap. Using this mode of presentation the ear canal is blocked off thereby creating a very small resonating chamber. Results from research in audiology indicate that an average of 7-15 dB correction is necessary for more accurate approximations of real-ear measurements of newborns and infants due to differences in ear canal volume (Stelmachowicz, 1991). This means that an adult ear with a relatively large ear canal (i.e., resonating chamber) would hear an 80 dB sound as 80 dB, whereas a newborn would hear that same tone as anywhere from 87 to 95 dB, or 7 to 15 dB louder. Premature infants ear canals are even smaller and therefore the difference is probably greater. The use of ear cups in the present study eliminated this problem by expanding the resonating chamber to include the ear canal and the ear cup itself. Using the ear cups, decibel levels could be set without regard to the size of ear canals.

Obviously, presentation mode needs to be carefully considered with this population and all modes have some inherent drawbacks. Audiological insert phones are exorbitantly expensive as are the ear cups which are sterile and need to be discarded after each infant has used them. Phonopad earphones are difficult to place in exactly the same position for all infants, and ear canal sizes vary by infant making calculations of appropriate dB levels very difficult. A tape recorder placed free field in a radiant warmer may not mask environmental noise as effectively as ear couplers or earphones.

Music in the NICU appears to be immediately beneficial on the physiological responses of very low birthweight infants, though acclimation to the auditory stimulus is evident. It is becoming clearer, through empirical data, that carefully controlled musical stimuli have no short term detrimental effects as measured by physiological responses. Further research is warranted to refine clinical techniques and procedures and to determine long term effects.

## REFERENCES\*

- Birnholz, J., & Benacerraf, B. (1983). The development of human fetal hearing. *Science*, 222, 516-518.
- Blass, E. (1987). What babies know, and noises parents make. *Science*, 237, 726.
- Brackbill, Y., Adams, G., Crowell, D. H., & Gray, M. L. (1966). Arousal level in neonates and preschool children under continuous auditory stimulation. *Journal of Experimental Child Psychology*, 4, 178-188.
- Caine, J. (1992). The effects of music on the selected stress behaviors, weight, caloric and formula intake, and length of hospital stay of premature and low birth weight neonates in a newborn intensive care unit. *Journal of Music Therapy*, 28 (4), 180-192.
- Campbell, L. (1985). The very low birth weight infant: Sensory experience and development. *Topics in Clinical Nursing*, January, 19-33.
- Collins, S. K., & Kuck, K. (1991). Music therapy in the Neonatal Intensive Care Unit. *Neonatal Network*, 9 (6), 23-26.
- Cox, L. C., Martin, R. J., Carlo, W. A., & Hack, M. (1993). Early ABRs in infants undergoing assisted ventilation. *Journal of the American Academy of Audiology*, 4, 13-17.
- DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science*, 208, 1174-1176.
- Detterman, D. (1978). The effect of heartbeat sound on neonatal crying. *Infant Behavior and Development*, 1, 36-48.
- Field, T. (1987). Alleviating stress in intensive-care unit neonates. *Journal of American Obstetrical Association*, 9, 129-135.
- Gordin, P. C. (1990). Assessing and managing agitation in a critically ill infant. *MCN*, 15, 26-32.
- Gorski, P., Davison, M., & Brazelton, T. (1979). Stages of behavioral organization in the high-risk neonate: Theoretical and clinical considerations. *Seminars in Perinatology*, 3 (1), 61-72.
- Gottfried, A. W., & Hodgman, J. E. (1984). How intensive is newborn intensive care? *Pediatrics*, 74, 292-294.
- Kileny, P. R. (1988). New insight on infant ABR hearing screening. *Scandinavian Audiology Supplement*, 30, 81-88.
- Kolata, G. (1984). Studying learning in the womb. *Science*, 225, 302-303.
- Leonard, J. E. (1993). Music therapy: Fertile ground for application of research in practice. *Neonatal Network*, 12 (2), 47-48.
- Lorch, C. A., Lorch, V., Diefendorf, A. O., & Earl, P. W. (1994). Effect of stimulative and sedative music on systolic blood pressure, heart rate, and respiratory rate in premature infants. *Journal of Music Therapy*, 31 (2), 105-118.
- Marlowe, J. A. (1982). Hearing in infancy: The development of auditory skills and the audiological evaluation. *Seminars in Speech, Language, & Hearing*, 3 (1), 28-44.



- Mullins, N. L. (Ed.). (1986). *Mosby's Medical & Nursing Dictionary*. St. Louis: The C. V. Mosby Co.
- Oehler, J. M. (1993). Developmental care of low birth weight infants. *The Nursing Clinics of North America*, 28 (2), 289-301.
- Paap, C. M., & Nahata, M. C. (1990). Clinical Pharmacokinetics of Antibacterial Drugs in Neonates. *Clinical Pharmacokinetics*, 19 (4), 280-318.
- Polverini-Rey, R. A. (1992). *Intrauterine musical learning: The soothing effect on newborns of a lullaby learned prenatally*. Unpublished doctoral dissertation, California School of Professional Psychology, Los Angeles.
- Salk, L. (1960). The effects of the normal heartbeat sound on the behavior of the newborn infant: Implications for mental health. *World Mental Health*, 12 (4), 168-175.
- Schaefer, M., Hatcher, P., & Barglow, P. (1980). Prematurity and infant stimulation: Review of research. *Child Psychiatry and Human Development*, 10 (4), 199-212.
- Standley, J. (1991). The role of music in pacification/stimulation of premature infants with low birthweights. *Music Therapy Perspectives*, 9, 19-25.
- Standley, J., & Madsen, C. (1990). Comparison of infant preferences and responses to auditory stimuli: Music, mother, and other female voice. *Journal of Music Therapy*, 27 (2), 54-97.
- Standley, J., & Moore, R. (1993, April). *The effect of music vs. mother's voice on NBICU infants' oxygen saturation levels and frequency of bradycardia/apnea episodes*. Paper presented at the Tenth National Research in Music Behavior Symposium, Tuscaloosa, AL.
- Stelmachowicz, P. G. (1991). Current issues in pediatric amplification. *Pediatric Amplification*, 1-17.
- White-Traut, R., & Tubeszewski, K. (1986). Multimodal stimulation of the premature infant. *Journal of Pediatric Nursing*, 1 (2), 90-95.

---

\* As referências bibliográficas deste artigo estão conforme o documento original.

### 2.5.3. Estudo Crítico – Por Christopher Johnson

#### Objectivo

O objectivo desta investigação era averiguar o efeito de um plano intermitente de estimulação musical (quatro minutos de música / quatro minutos de silêncio) através de auscultadores Bio-Logic e Ear Couplers, nas respostas fisiológicas de bebés prematuros, com vinte e quatro a trinta semanas de tempo de gestação, durante a sua primeira semana de vida, altura esta em que se encontravam sujeitos a oxigenação.

#### Design

O plano desta investigação consistiu numa experiência com dois grupos semelhantes.

Os dois grupos foram formados e equilibrados com base na idade de gestação, idade cronológica, peso e sexo. O grupo A foi sujeito a estimulação musical (Grupo Experimental); o grupo B não (Grupo de Controle).

#### Revisão do artigo

Este artigo começa com uma revisão bibliográfica sobre os benefícios calmantes da música nas crianças normais. Estes benefícios são também reconhecidos ao nível do período pré-natal. Revelaram-se mais calmantes do que outros estímulos até agora mais considerados (como, por exemplo, batimentos cardíacos).

Esta exposição é seguida por uma revisão de todas as capacidades dos bebés, incluindo, entre outras, a atenção, a discriminação de sons e o reconhecimento de experiências repetidas.

Segue-se uma discussão sobre estudos mais recentes em que a música aparece não só como calmante mas também como possível factor de alívio do sofrimento. A partir daqui estabelecem-se pontes para considerações acerca do efeito da música em estados comportamentais.

A atenção é focada, em seguida, no meio ambiente do bebé prematuro – a incubadora e o isolamento. Segue-se uma revisão da pesquisa existente sobre o efeito da música na melhoria do ambiente dentro da incubadora que, apesar de protectora, pode causar sofrimento. Cassidy e Standley revêem depois a literatura relativa à importância de uma oxigenação apropriada da criança de maneira a estruturarem o procedimento para o seu próprio estudo.

Os dois parágrafos seguintes revêem cuidadosamente a bibliografia que aborda a investigação sobre os efeitos da audição musical nas respostas fisiológicas dos bebés prematuros. Isto prepara os outros dois últimos parágrafos nos quais se encontram as questões que conduziram à presente investigação. Estes quatro parágrafos, apesar de descritos abreviadamente, formam o argumento lógico para a existência da presente investigação, e são a essência desta revisão.

#### Método

As características dos participantes estão descritas no início da Metodologia. Há muitos detalhes que fornecem informação sobre procedimentos médicos que são utilizados no tratamento destes bebés. No parágrafo seguinte, Cassidy e Standley discutem os critérios específicos usados para seleccionarem os participantes. O rigor utilizado é visível no facto de apenas terem escolhido



vinte participantes em setenta e quatro potenciais voluntários. Além disso, os critérios de agrupamento são apresentados. Foram feitos dois grupos de dez bebês cada, grupos esses correspondentes e baseados em idade de gestação, idade cronológica, peso e sexo.

O terceiro parágrafo descreve a programação do projecto. Todas as observações foram agendadas de forma a serem completadas em blocos de três horas, em três dias consecutivos. O parágrafo seguinte descreve o estímulo musical usado, constituído por uma gravação de trinta minutos, contendo faixas alternadas de quatro minutos de música e quatro de silêncio. Aqui também são descritas com algum detalhe as canções utilizadas no projecto.

O quinto parágrafo descreve todo o equipamento usado no projecto. As autoras descrevem a forma como a música foi apresentada às crianças e a intensidade usada. Devido à sensibilidade auditiva dos participantes, a intensidade foi monitorizada com extremo cuidado.

O sexto parágrafo amplia a descrição do esquema de estímulo através da ilustração dos segmentos de música e silêncio. É explicado como e quando são recolhidos os dados: uma vez por minuto, durante trinta e seis minutos, tanto para os participantes do Grupo Experimental como para os do Grupo de Controle. O sétimo parágrafo descreve as variáveis dependentes, que são os níveis de saturação de oxigénio, a frequência cardíaca e a frequência respiratória. Também foram recolhidos dados sobre a apneia/bradicardia durante toda a experiência.

## Resultados

Nesta secção, as autoras descrevem os dados recolhidos dos participantes, nas três variáveis dependentes, em cada minuto, para os três dias. Descrevem também as diferenças em termos de desvio-padrão. Os resultados são então relacionados com dois factores indiciadores de desconforto e sofrimento – a apneia e a bradicardia.

Os primeiros testes relatados dizem respeito à oxigenação. Para o tratamento de dados, procedeu-se à análise da variância (ANOVA) com medições repetidas para cada dia e para cada uma das variáveis dependentes. A análise dos resultados do primeiro dia revelou uma diferença significativa nas crianças do Grupo Experimental, que mostraram um maior nível de saturação de oxigénio (o que é desejável). Este resultado significativo não foi visível nos dois dias seguintes, embora na maior parte do tempo os níveis de saturação de oxigénio tenham sido mais elevados para o Grupo Experimental. A comparações de desvios-padrão mostrou que nenhum dos dois grupos era mais homogéneo ou consistente que o outro.

Os testes de frequência cardíaca não indicaram diferenças significativas entre os dois grupos. No entanto, no primeiro dia houve uma tendência para as frequências cardíacas do Grupo Experimental serem mais lentas que as do Grupo de Controle. Esta tendência foi invertida nos dois dias seguintes do estudo. Os resultados dos desvios-padrão mostraram que o Grupo Experimental era mais estável do que o Grupo de Controle ao longo do estudo, embora em menor grau no terceiro dia.

Os dados sobre a frequência respiratória não foram significativamente diferentes entre os dois grupos. No entanto, no início do primeiro dia, o Grupo Experimental apresentou uma frequência respiratória consistentemente mais baixa. Este efeito só permaneceu durante cerca de metade do tratamento. No segundo dia os dados indicaram que as crianças do Grupo Experimental estavam ligeiramente mais estáveis, mas não num grau significativo. No terceiro dia não existiram diferenças, embora a variação para todos tenha sido menor.

Os últimos testes apresentados foram os da apneia/bradicardia. Não houve diferenças importantes ou significativas entre os dois grupos em nenhuma destas duas variáveis dependentes.



## Discussão

Cassidy e Standley começam a discussão dos resultados salientando o único dado estatístico significativo que obtiveram (níveis de saturação de oxigénio, no primeiro dia). Fazem notar que este resultado é consistente com pesquisas prévias, assim como a habituação das crianças à música nos dias dois e três. O outro aspecto do estudo que elas realçam é o facto de não se terem apresentado efeitos adversos relacionados com a presença da música (apneia/bradicardia). No parágrafo seguinte, as autoras mostram como este aspecto é crucial, uma vez que isto significa que a audição de música, mesmo em prematuros frágeis, em fases muito precoces do seu desenvolvimento, não é contra-indicada.

Enfatizam, ainda, o modo como os participantes do Grupo Experimental tiveram melhores respostas a nível médico e demonstraram maior estabilidade que o Grupo de Controle.

Os resultados obtidos são também examinados à luz de estudos anteriores. As autoras pensam que apesar de se instalar uma certa habituação à música por parte dos bebés, num quadro onde os pais não lhes podem pegar, alimentar, dar banho ou tocar, talvez a partilha da música seja uma forma através da qual possam criar laços com o seu novo bebé.

Uma parte deste estudo difere dos anteriores. Neste, os bebés não sofreram de episódios de apneia ou bradicardia depois do estímulo musical ter cessado. Isto foi visto como muito positivo. As autoras levantaram a hipótese de que esta diferença nos resultados possa ter a sua razão de ser no esquema de horário montado e/ou no modo de apresentação. Assim, colocam a hipótese de que o formato de apresentação do estímulo usado num estudo anterior – constituído por vinte minutos seguidos de música – possa ter sido sobre-estimulante. Neste caso, o formato utilizado – constituído por quatro minutos de música alternando com quatro de silêncio/etc. – poderá ser mais apropriado a esta população.

As autoras examinam os benefícios do equipamento experimental utilizado, comparando-o com outros. Concluem que os estudos nesta área têm introduzido melhorias nas formas de fazer ouvir música a crianças deste tamanho e idade.

Concluem que a música nas UNCI<sup>1</sup> parece ter um efeito benéfico imediato, até para os bebés prematuros mais novos. Se for feito com métodos controlados, os dados empíricos parecem suportar esta prática clínica através das medições das respostas fisiológicas encontradas.

## Contribuições para a área de estudo

O aspecto mais saliente deste estudo foi o de mostrar que os benefícios da música em UNCI, já estabelecidos anteriormente, podiam também ser aplicados e praticados com os mais novos dos bebés prematuros. Embora só uma diferença significativa tenha sido registada entre os dois grupos, o aumento da estabilidade dos bebés e a não existência de contra-indicações, faz desta prática, pela primeira vez, uma alternativa viável para esta população.

## Comentário

A revisão bibliográfica relativamente a trabalhos de investigação realizados com bebés prematuros é excelente. As autoras fazem um trabalho excepcional na descrição do campo da terapia musical aplicado aos bebés prematuros. O conhecimento prévio é amplo e a revisão de estudos anteriores mostra uma linha nítida de saber que estabelece a necessidade da presente investigação.

A verdadeira vantagem da secção de metodologia é que tudo está perfeitamente articulado, não deixando margem para dúvidas quanto ao que foi feito e à forma como foi levado a cabo. As autoras são obviamente pessoas que compreendem inteiramente a área na qual trabalham. Percebem o que é necessário considerar, controlar e ajustar. Os procedimentos são claros, não só para outros investigadores que queiram replicar a experiência, como para os



terapeutas que queiram reproduzir as actividades destas investigadoras em UNCI, criando apropriadamente quer os estímulos musicais como também as suas formas de apresentação. Todos os detalhes dos estímulos, o equipamento, a regularidade da apresentação, os tipos de dados a serem registados e a informação adicional de relevância, são apresentados de forma clara e acessível.

A secção dos resultados está bem feita e as investigadoras não fazem extrapolações indevidas. No entanto, como não havia contra-indicações quanto aos comportamentos relacionados com stress, não se percebe porque é que não consideraram modificar o alfa (isto é, o nível de significância). Mesmo que a probabilidade de um erro estatístico tipo I fosse aumentada (isto é, mesmo que a hipótese nula tivesse sido rejeitada quando deveria ter sido aceite), neste caso havia salvaguardas. Se elas quisessem ter identificado uma diferença que não existiu, sabendo, no mínimo, que os comportamentos indesejados não aumentariam, reclamar um falso positivo não teria sido nocivo para o estudo. Considerando o pequeno número de participantes, compensar esta falta de potência com um alfa maior (portanto, exigindo um nível de significância menor) poderia ter sido uma boa escolha.

A discussão dos resultados é um excelente trabalho ao colocar todas as conclusões em perspectiva. As autoras explicam o que é que os dados significam no contexto real. Não confundem significância estatística com importância ou significância prática. Além de discutirem estes aspectos específicos, colocam este estudo no contexto de todo um corpo de pesquisas nesta área. Neste sentido, este estudo cresce em importância, ao interpretar técnicas anteriores através de escolhas feitas aqui.

Assim, toda esta área de investigação é colocada em conjunto, para começar a consolidar a prática da terapia musical nas UNCI como um corpo de investigação, até ao momento da publicação.

É de assinalar este estudo no panorama da investigação, quer pela profundidade do resumo de toda uma área de estudos, quer pelas questões levantadas.

---

<sup>1</sup> Unidade Neonatal de Cuidados Intensivos (UNCI) -Tradução de Neonatal Intensive Care Unit (NICU)

## 2.6.

### THE EFFECTIVENESS OF INDIVIDUAL / SMALL-GROUP SINGING ACTIVITIES ON KINDERGARTNERS USE OF SINGING VOICE AND DEVELOPMENTAL MUSIC APTITUDE

A EFICÁCIA DE ACTIVIDADES CANTADAS INDIVIDUALMENTE OU EM PEQUENOS  
GRUPOS NO DESENVOLVIMENTO DA APTIDÃO MUSICAL E NA VOZ CANTADA DE  
CRIANÇAS NO JARDIM DE INFÂNCIA

J. RUTKOWSKI (1996)

JOURNAL OF RESEARCH IN MUSIC EDUCATION, 44, 353 - 368

### 2.6.1. Resumo do artigo / Abstract

#### Resumo

Neste estudo foi investigado: o efeito de actividades de prática vocal realizadas individualmente ou em pequenos grupos no desenvolvimento musical e na capacidade de utilizar a voz cantada em contexto de jardim de infância; a relação entre a aptidão musical (tonal) e o desenvolvimento vocal; o período de tempo necessário à instrução para fornecer resultados significativos.

Foram utilizadas as mesmas canções, actividades e jogos com um Grupo Experimental e com um Grupo de Controle. No entanto, o ensino ministrado ao Grupo Experimental implicou participações individuais e em pequenos grupos.

Os resultados mostraram que as actividades individuais e em pequenos grupos tiveram efeito no desenvolvimento musical das crianças no uso das suas vozes cantadas. Mostrou também que era necessário um ano inteiro de ensino para que o resultado fosse efectivo. O estudo mostrou ainda que existe uma pequena relação entre a aptidão musical (tonal) e a capacidade de cantar.

#### Abstract

The following were investigated in this study: the effect of individual/small-group singing activities used within the traditional large-group setting on kindergartners' use of singing voice and developmental music aptitude; the strength of relationship between tonal aptitude and use of singing voice; the length of time over which instruction is needed to provide significant results. Songs, activities, and games were the same for both the control and treatment groups; however, instruction for the treatment group involved small-group and individual participation in the singing activities. Results showed that the treatment did have an effect on the children's development of use of singing voice, that an entire year of instruction was needed for the treatment to be effective, and that a very small relationship seems to exist between use of singing voice and developmental tonal aptitude.



## 2.6.2. Artigo Original

# The effectiveness of individual / small-group singing activities on kindergartners use of singing voice and developmental music aptitude

J. Rutkowski

The ability to sing in a church, at sporting events, with the radio, or even in the shower is taken for granted by many. However, approximately 18% of children in Grade 6 and below are unable to sing simple songs and are considered problem singers (Franklin & Franklin, 1983; Gould, 1968; Roberts & Davies, 1975; Romaine, 1961). This percentage has increased over the years (Goetze, Cooper, & Brown, 1990). Furthermore, these children are often labeled "monotones" by parents, teachers, and society. As adults, these individuals cannot successfully participate in the most basic singing activities expected of the general population. Due to the social nature of many of these activities, those who are unable to sing may suffer embarrassment or humiliation. They are denied the most personal form of music expression available to human beings.

The assumption that the ability to sing is an innate talent rather than a trainable skill is suspect. If one can speak, one can sing (Goetze, Cooper, & Brown, 1990; Gordon, 1971; Rutkowski, 1989). Furthermore, the relationship between a child's ability to use a singing voice and his or her music aptitude seems to be small (Rutkowski, 1986). Those with above-average music aptitude may in fact be labeled "nonsingers." However, since these individuals do not know how to use a singing voice and are, therefore, not able to reproduce a melodic line, they are often assumed to have low music aptitude, even by their music teachers.

Various remediation strategies to teach singing have been identified. Gould (1968), Roberts and Davies (1975, 1976), and Romaine (1961) have investigated the effectiveness of a "speech-to-song" method in which the child first gains control of his or her speech pitch, and then is given a personal note (identified by the teacher) from which the singing range can be expanded. The conclusions of these researchers indicated that this method was successful. Buckton (1977) and Wolner and Pyle (1933) concluded that vocal training was superior to pitch-discrimination and tonal-memory methods. Richner (1976) also found training with an emphasis on singing to be most effective. Cobes (1970) concluded that reinforcement, even for performances approximating a criterion pitch, was successful. Phillips's results have indicated breath control techniques as effective (1985, 1992). Gordon (1971, 1979) suggested that the nonsinger yell or call out in order to use the muscles needed for singing and that the out-of-tune singer engage in echo responses that involve chordal patterns. Roberts and Davies (1975) and Pedersen and Pedersen (1970) also found chord-tone patterns, as opposed to diatonic patterns, easier for the developing singer. McNitt (1972) found this technique appropriate for adult nonsingers as well.

Jarjisian (1981) concluded that practice with both pentatonic and diatonic pitch patterns, rather than only one or the other, enhanced rote singing achievement. Singing songs written in a variety of modes in addition to major and minor has also been found to improve children's singing performance (DeYarman, 1972; Miller, 1975). Hale (1977) discovered melodic reinforcement to be more effective initially with the uncertain singer than was harmonic reinforcement. In a similar investigation, Sterling (1985) found children's singing performances to be more accurate when

---

The author wishes to thank Pennsylvania State University, which supported the pilot study described in this article with a Penn State Research Initiation Grant and the main study with a Penn State College of Arts and Architecture Faculty Research Grant. Joanne Rutkowski is an associate professor of music education in the School of Music, Pennsylvania State University, University Park, PA 16802-1901. Copyright © 1996 by Music Educators National Conference.



melodic replication or traditional tonal harmonic accompaniment was used, as opposed to chromatic or dissonant harmonic accompaniment. However, she also concluded that the nature of the melody itself had a significant effect regardless of the type of accompaniment used. A tape recorder (Klemish, 1974), a vertical keyboard (Jones, 1979), and kazoos (Runfola, 1981) have all been successfully used to assist the nonsinger. Whitman (1971), however, found that a keyboard instrument did not help the singing achievement of fourth-, fifth-, and sixth-grade students.

Individual or small-group arrangements for singing instruction have also been investigated and found to be more productive than large-group work (Gaiser, 1961; Goetze, 1985; Jersild & Bienstock, 1931, 1934; Richner, 1976; Smale, 1987; Updegraff, Heliger, & Learned, 1938). However, Smith (1963) indicated the impracticality of individual work and found large-group training effective for improving singing achievement of kindergartners in his study. Gordon (1985) concluded from a survey of the literature that "most strategies have only a few studies to support their effectiveness" (p. 14) and "further study is needed concerning the effectiveness of all identified strategies" (p. 16). Based on the percentage of problem singers in elementary schools, it also seems that these strategies have not been widely adopted.

Effective, practical strategies for assisting young problem singers need to be identified. Individual remediation within a large-group/classroom setting may provide the benefits of individual/small-group instruction, but would be more practical for teachers who see large numbers of students, often only once a week. However, this strategy has not been empirically investigated. In addition, "because the development of music aptitude takes place in the formative years from birth (or possibly before birth) through approximately age nine" (Gordon, 1986, p. 3) and because short-term music instruction has been shown to significantly influence 5-year-olds' developmental music aptitude levels (Flohr, 1981), it seems appropriate to monitor developmental aptitude levels when conducting research with children younger than age 9. Therefore, the purpose of this study was to (1) investigate the comparative effectiveness of two teaching strategies, large-group and individual/small-group singing activities, on kindergartners' use of singing voice and developmental music aptitude, (2) monitor children's singing voice achievement during the instructional period to determine the length of time over which instruction is needed to produce improvement of the singing voice, and (3) monitor children's developmental music aptitude levels during the instructional period in order to determine what effect, if any, the type of instruction has on the children's developmental aptitudes.

## PILOT STUDY

A pilot study was conducted (Rutkowski, 1990b) to preliminarily test the hypotheses, provide "ideas, approaches, and clues not foreseen prior to the pilot study" (Issac & Michael, 1982, p. 34), and provide "enough data ... to make a sound decision on the advisability of going ahead with the main study" (Issac & Michael, 1982, p. 35). The following questions were posed upon which decisions were made for the main study:

- 1 – Will there be significant differences between the control and treatment groups developmental music aptitude scores (as measured by the Primary Measures of Music Audiation) after 4 months and/or 9 months of instruction?
- 2 – Will there be significant differences between the control and treatment groups use of singing voice (as measured by the Singing Voice Development Measure) after 4 months and/or 9 months of instruction?
- 3 – Are the procedures feasible for a study with large numbers of subjects?
- 4 – Are the raters reliable when scoring children's use of singing voice?



## METHOD

Kindergartners ( $n = 14$ ) from a parochial elementary school located in a small Pennsylvania town participated in the pilot study. All subjects were 5 years old when the study began.

A randomized control-group pretest-posttest design was used (Isaac & Michael, 1982). The lessons were designed by the researcher and taught by a graduate research assistant who had 2 years experience teaching general and choral music grades K-12. The researcher and assistant met frequently throughout the pilot study to discuss previous lessons, design new lessons, and ensure that treatment and control groups were receiving appropriate instruction. Lessons comprised large-group singing activities and rhythmic/movement activities. Songs, activities, and games were the same for both the control and treatment groups. However, instruction for the treatment group also included activities that encouraged small-group and individual singing. Small-group singing was introduced prior to individual singing. For example, the children participated in a rondo to line up at the door at the end of music class. The teacher would sing a child's name on a descending major or minor arpeggio pattern; the children would all respond with "go to the door" on the same tonal pattern while that child lined up at the door. This large-group activity was used with the control group. After the children in the treatment group were familiar with the activity, they progressed to small-group singing. If a girl's name was sung, only the girls would respond with "go to the door." Other small groups were also used; for example, only the students sitting in the same row as the child whose name was sung would sing back. When the children were comfortable with singing in small groups, individual singing was used. For example, if the teacher sang a child's name, only that child responded with "go to the door." Opportunities for these types of small-group/individual singing responses were presented at least twice in every treatment-group lesson.

Subjects were randomly assigned to either the control or treatment group. The children were scheduled for two 30-minute music classes per week. For purposes of the pilot study, during this music time, the children received instruction for 15 minutes with the music teacher, who was trained in music but not music education, and for 15 minutes with the research assistant. Although it would have been more appropriate for all instruction to be administered by the research assistant, this was not possible given the school climate. The control group received instruction from the research assistant first while the treatment group was receiving instruction from the music teacher. After 15 minutes, the groups switched. All children received the same instruction from the music teacher, which focused on music activities other than singing.

A test battery was administered by the research assistant three times: prior to treatment, midway through the treatment period (in December), and at the end of the treatment period (in May). The battery consisted of the Primary Measures of Music Audiation (PMMA), a standardized instrument designed to measure young children's developmental music aptitude (Gordon, 1979), and the Singing Voice Development Measure (SVDM), an instrument designed to measure children's use of singing voice. The SVDM was developed based on the hypothesis that "use of singing voice is a separate but requisite behavior to the ability to sing with accurate intonation" (Rutkowski, 1990a, p. 93). Singing-voice behaviors were identified and a rating scale developed. Children received one overall rating between 1 and 5 that reflected their performance of all the patterns. The SVDM has been shown to be a valid measure of young children's use of singing voice (Rutkowski, 1984, 1986, 1990a). The SVDM rating scale consisted of the following:

- 1 – "Presinger" does not sing but chants the song text.
- 2 – "Speaking-range singer" sustains tones and exhibits some sensitivity to pitch but remains in the speaking voice range (usually A2, the A below middle C, to C3, middle C).
- 3 – "Uncertain singer" wavers between speaking and singing voice and uses a limited range when in singing voice (usually up to F3).
- 4 – "Initial range singer" exhibits use of initial singing range (usually D3 to A3).
- 5 – "Singer" exhibits use of extended range (sings beyond the register lift: B3-flat and above).



The Tonal (PMMAT) and Rhythm (PMMAR) sections of PMMA were administered on separate days under controlled conditions as specified in the test manual. The children took the test together; they were not divided into control and treatment groups.

After the children completed each section of PMMA, the patterns for SVDM were practiced. See Figure 1 for the patterns used. The research assistant used the following procedure for individual testing:

- 1 – Play the first pattern on the tone bells or piano ("see the bird").
- 2 – Sing the first pattern for the children with the text. Do not use any accompaniment.
- 3 – Have the children echo. Again, do not use any accompaniment.
- 4 – Repeat steps 1-3 with each pattern.
- 5 – Do not pause in between any of the above steps.

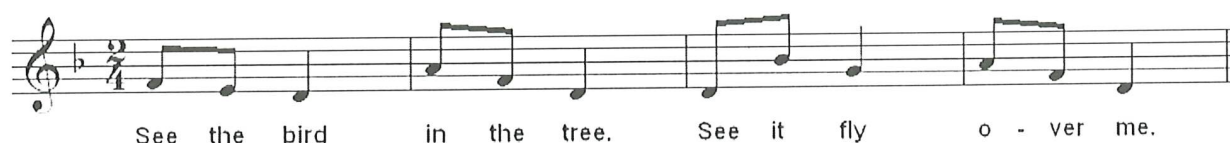


Figure 1. Singing patterns used for the Singing Voice Development Measure (SVDM).

After these two practice sessions, the children's voices were individually tested. Each child reported to a familiar, private room where his or her voice was tape-recorded as he or she echoed the research assistant singing the SVDM patterns. For the pretest, the procedure was repeated several days later. It was believed that two evaluations would provide a more accurate assessment of the children's use of singing voice. However, this procedure was very time-consuming and would be particularly difficult with a large sample. It was decided that it would not be necessary to record the children's voices twice for the midtest and posttest if a significant difference did not exist between these two performances on the pretest. A tape recording was prepared that presented the first pretest performances followed by the second pretest performances. Two raters, both experienced in rating children's singing voice, were employed to score this tape recording. The raters were unaware of which test they were listening to. They were asked to rate the two sets of performances on the tape according to the SVDM rating scale, assigning one rating for each child's performance of all four patterns. The raters completed their evaluations independently. A *t*-test was then performed. Since there was not a significant difference between these scores, the children's voices were recorded only once each for the midtest and posttest.

PMMAT and PMMAR scores were recorded for the pretests, midtests, and posttests. The two raters who scored the SVDM pretests were employed to score the tape recordings of the midtest and posttest performances. The tape presented all the children's performances randomly in the following order: midtest, posttest, and midtest (repeated to investigate internal consistency of each rater). The rating procedures were the same as for the pretest.

## RESULTS

All reliability coefficients for SVDM were satisfactory. Intrarater reliability coefficients were .82 and .77. Interrater reliability coefficients were .84 on the pretest, .95 on the midtest, and .93 on the posttest. Analyses of covariance, with the pretest as the covariate, were performed to make decisions regarding the necessary frequency of testing for the main study. Significant differences existed between groups for PMMAT on the midtest and posttest, for PMMAR on the midtest, and for SVDM on the posttest. The treatment group means were higher on each of these tests.

It was concluded that the main study should be conducted, that the procedures were feasible for use with larger numbers of children, and that the raters were reliable in their use of SVDM.



## MAIN STUDY

Based on the results of the pilot study, it was decided to conduct the main study with a larger number of students but with one teacher providing "formal" music instruction. Specific questions were:

- 1 – Will there be significant differences between the control and treatment groups developmental music (tonal) aptitude scores, as measured by PMMA (Gordon, 1979), after 4 months and/or 9 months of instruction? If a difference exists, is there a significant interaction for midtest and/or posttest scores between groups and levels of pretest aptitude?
- 2 – Will there be significant differences between the control and treatment groups use of singing voice, as measured by the SVDM (Rutkowski, 1984, 1986, 1990a), after 9 months of instruction? If a difference exists, is there a significant interaction for posttest scores between groups and levels of pretest singing voice achievement?
- 3 – What is the strength of the relationship between tonal aptitude and use of singing voice?

## METHOD

Kindergartners (n=99) from a public elementary school in a mid-sized central Pennsylvania town participated in the main study, which began in September 1989 and utilized a nonrandomized control-group pretest-posttest design (Isaac & Michael, 1982). Intact classes were randomly assigned to either the control or treatment group. As in the pilot study, all music lessons consisted of large-group singing activities and rhythmic/movement activities. Songs, activities, and games were the same for both the control and treatment groups. However, instruction for the treatment group involved small-group and individual participation in the singing activities. Small-group singing was again introduced prior to individual singing. Lesson plans used for the main study were the same lesson plans used for the pilot study. However, in the main study, the children received music instruction from their regular music teacher for 30 minutes once a week.

Prior to treatment, a test battery, comprised of PMMA tonal subtest (PMMAT) (Gordon, 1979) and SVDM, was administered by the music teacher. The PMMAT was also administered halfway through the treatment period (in January) and at the end of the treatment period (in June). Since results of the pilot study did not reveal any significant differences on the rhythm portion of PMMA after 9 months of instruction and since administration of that subtest would have taken three of the instructional class periods (pretest, midtest, posttest), it was decided to only administer the tonal subtest for the main study. It was also decided to administer PMMAT three times, because significant tonal aptitude differences existed halfway through the treatment period as well as at the end of the treatment period in the pilot study. The SVDM was administered prior to treatment and at the end of the treatment period. It was decided to not administer this test halfway through the treatment period since results of the pilot study did not reveal significant singing-voice differences until the end of the treatment period. The time-consuming nature of the SVDM administration also contributed to this decision. Tests were administered by the regular music teacher following the same procedures used in the pilot study.

The regular music teacher scored PMMAT and recorded each child's score. Two raters, who used SVDM to rate the children's singing voices in the pilot study, were employed to score SVDM for the main study. A tape recording that presented the posttest performances first, followed by the pretest performances, was prepared for each rater. The raters were unaware of the test order. Instructions given to each rater were the same as those for the pilot study.

**Table 1**  
Means and Standard Deviations for Control and Treatment Groups on Dependent Measures and Covariates

| Test     | Control group (n = 49) |       |       | Treatment group (n = 50) |       |      |
|----------|------------------------|-------|-------|--------------------------|-------|------|
|          | X                      | Adj X | SD    | X                        | Adj X | SD   |
| PMMAT    |                        |       |       |                          |       |      |
| Pretest  | 23.18                  | 23.60 | 10.27 | 24.52                    | 25.34 | 9.07 |
| Midtest  | 28.47                  | 28.69 | 5.42  | 30.24                    | 30.59 | 5.69 |
| Posttest | 30.41                  | 30.58 | 4.99  | 30.94                    | 31.73 | 6.55 |
| SVDM     |                        |       |       |                          |       |      |
| Pretest  | 5.88                   | 5.87  | 1.99  | 6.20                     | 6.16  | 2.10 |
| Posttest | 5.45                   | 5.40  | 2.51  | 6.46                     | 6.48  | 2.10 |

*Note.* PMMAT = Primary Measures of Music Audiation Tonal subtest; SVDM = Singing Voice Development Measure.

## RESULTS

Correlation coefficients, derived from a two-way mixed model analysis of variance (ANOVA) (Bintig, 1980), were computed to determine the agreement between raters when scoring for SVDM. The high coefficients of .90 for the pretest and .99 for the posttest indicate that the raters used the SVDM similarly when rating the children's singing-voice performance.

Means and standard deviations for control and treatment groups on PMMAT pretests, midtests, and posttests and SVDM pretests and posttests are presented in Table 1. Analyses of

**Table 2**  
ANCOVA Results for Group Differences on Dependent Measures

| Test/Source   | df | Adjusted sum of squares | Mean square | F     | p <   |
|---------------|----|-------------------------|-------------|-------|-------|
| PMMA midtest  |    |                         |             |       |       |
| Covariance    | 1  | 699.42                  | 699.42      | 29.22 | 0.000 |
| Group         | 1  | 48.28                   | 48.28       | 2.02  | 0.159 |
| Error         | 96 | 2297.91                 | 23.94       |       |       |
| Total         | 98 | 3074.91                 |             |       |       |
| PMMA Posttest |    |                         |             |       |       |
| Covariance    | 1  | 890.85                  | 890.85      | 35.49 | 0.000 |
| Group         | 1  | 0.32                    | 0.32        | 0.01  | 0.911 |
| Error         | 96 | 2409.81                 | 25.10       |       |       |
| Total         | 98 | 3307.66                 |             |       |       |
| SVDM Posttest |    |                         |             |       |       |
| Covariance    | 1  | 216.33                  | 216.33      | 68.72 | 0.000 |
| Group         | 1  | 14.79                   | 14.79       | 4.70  | 0.033 |
| Error         | 96 | 302.21                  | 3.15        |       |       |
| Total         | 98 | 543.84                  |             |       |       |

*Note.* PMMA = Primary Measures of Music Audiation; SVDM = Singing Voice Development Measure.

covariance (ANCOVAs), with pretest PMMAT and SVDM as covariates, were performed to determine the answers to Research Questions 1 and 2 (Table 2). No significant tonal aptitude differences existed between groups, although both groups had gains in tonal aptitude mean scores during the treatment period. Since no significant differences existed, it was not necessary to perform any further analysis. Significant differences existed between groups for SVDM on the posttest: The treatment group mean was higher than the control group mean. A two-way ANOVA was then performed to determine if the treatment was more effective for those students who had low, middle, and/or high singing achievement at the time of the pretest (Table 3). Low singing achievement was defined as a combined rater score from 2 to 4; middle singing achievement as a score between 5 and 7; high singing achievement as a score between 8 and 10. No significant interaction existed indicating that the treatment effect was not influenced by students pretest singing achievement.

Pearson product-moment correlations were computed to determine the answer to Research Question 3 (Table 4). All correlation coefficients were low and suggest a very small relationship between use of singing voice and developmental tonal aptitude.



## CONCLUSIONS

It was concluded that individual/small-group instruction within the traditional large-group/classroom setting did have an effect on the children's use of singing voice, regardless of their pretest singing achievement. The posttest SVDM mean score for the treatment group was 6.46. When this score is divided by 2 (because the mean reflects the sum of two raters scores), a singing voice level of 3.23 is derived. Similarly, the average singing-voice level for the control group was 2.73. Therefore, the average child in the treatment group was an "uncertain/initial range singer," whereas the average child in the control group was a "speaking range/uncertain singer" at the conclusion of the instructional period. Furthermore, the control-group mean decreased from the pretest to the posttest. Significant differences between the groups may have been a result of this decrease in singing voice achievement for the control group rather than the effectiveness of individual/small-group singing. This decrease, however, raises a concern that ineffective instruction may actually be detrimental to children's singing-voice development.

**Table 3**  
*ANOVA Results for SVDM Posttest by Group and SVDM Pretest Levels*

| Source        | df | Adjusted sum of squares | Adjusted mean square | F     | p <   |
|---------------|----|-------------------------|----------------------|-------|-------|
| Group         | 1  | 16.880                  | 16.880               | 4.85  | 0.030 |
| Level         | 2  | 190.314                 | 95.157               | 27.35 | 0.000 |
| Level x Level | 2  | 6.502                   | 3.251                | 0.93  | 0.396 |
| Error         | 93 | 323.556                 | 3.497                |       |       |
| Total         | 98 | 543.838                 |                      |       |       |

Note: SVDM = Singing Voice Development Measure.

The treatment did not have an effect on the children's developmental tonal aptitude as it seemed to have in the pilot study, although percentile means increased for both groups over the course of the main study. Teacher effect is always a consideration. One teacher taught all the children for the pilot study and a different teacher taught all the children for the main study. However, frequency and length of instruction is also of concern. In the pilot study, the children received music instruction twice a week for a total of 60 minutes: two 15-minute sessions with a research assistant engaging the children in singing and rhythmic activities and two 15-minute sessions with another teacher who engaged the children in listening activities. In the main study, the children received music instruction with a specialist once a week for 30 minutes. It is not known what music experiences the children received in their regular kindergarten classroom. Perhaps shorter, more frequent sessions are more effective.

**Table 4**  
*SVDM & PMMAT Correlation Coefficients Matrix*

|               | PMMAT (N = 99) |         |          |
|---------------|----------------|---------|----------|
|               | Pretest        | Midtest | Posttest |
| SVDM pretest  | .175           | .135    | .182     |
| SVDM posttest | .205           | .166    | .207     |

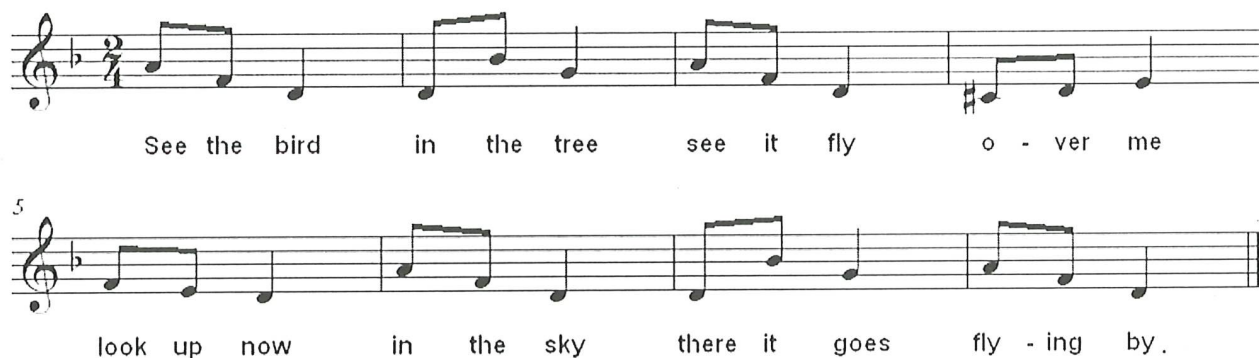
Note. PMMAT = Primary Measures of Music Audiation tonal subtest; SVDM = Singing Voice Development Measure.

The results of this study concur with previous results (Rutkowski, 1986) and illustrate that the relationship between use of singing voice and developmental tonal aptitude seems very small. However, the limited variability of SVDM, due to only 5 scoring levels and 4 tonal patterns, may be a factor in these low correlations. The recommendations provided for revising SVDM allow for further investigation of this apparent minimal relationship. Nevertheless, the notion that singing ability reflects a child's music ability seems to be in question. Teachers, as well as researchers, need to be cognizant of this and not assume that a child's singing performance is an indication of his or her tonal potential for learning music.

## RECOMMENDATIONS

Based on the results of this study, it is recommended that teachers include individual/small-group singing activities in their music lessons. Individual singing has consistently been shown to result in more accurate singing and to improve singing achievement (Gaiser, 1961; Goetze, 1985; Jersild & Bienstock, 1931, 1934; Richner, 1976; Smale, 1987; Updegraff, Heliger, & Learned, 1938). It seems that these techniques are also effective when applied in the traditional large-group classroom setting.

Since differences in singing voice as measured by SVDM did not emerge at the fourth-month testing in the pilot study but were significant after 8 months of instruction, it seems that research with young children conducted for only a short period of time may result in premature conclusions. Had the pilot study been conducted for only half the school year, the conclusions would have been quite different and the main study would probably not have been conducted.



Additional research is encouraged, especially studies investigating the effect of frequency and amount of instruction on use of singing voice and developmental music aptitude. Perhaps shorter, more frequent music sessions are more desirable. Techniques for tape-recording individual children's singing during large-group singing activities should also be investigated. Singing voice scores for SVDM were based on a child's ability to sing alone. It may be that scores for children instructed through individual/small-group techniques were significantly higher because these children were comfortable singing alone. Individual testing procedures are extremely time-consuming and tedious. Perhaps the small sample sizes often used in singing research and teachers reluctance to monitor singing-voice development are consequences of these testing procedures.

It is also recommended that any research investigating children's music abilities and perceptions through singing be conducted with children who have use of the singing voice. If a child does not have use of the singing voice, he or she cannot use that instrument to express music ability or perception. Therefore, results of studies involving singing as a measurement technique for musical abilities may be in question.

The raters commented that many children seem to be inconsistent in their use of singing voice. Often children fluctuate between two stages on the SVDM rather than exhibiting one behavior consistently. It has been recommended that the SVDM scale be expanded to allow for this fluctuation. It has also been recommended that the number of tonal patterns be increased to portray a more accurate picture of each child's use of singing voice and to allow for more variability in the measure. Expansion of the scale in these ways would also result in a more sensitive instrument providing an even more accurate measure of children's use of singing voice. The following revised SVDM scale and tonal patterns (see Figure 2) are proposed and should be the subject of future research:



- 1 – "Presinger" does not sing, but chants the song text.
- 1.5 – "Inconsistent Speaking-Range Singer" sometimes chants, sometimes sustains tones and exhibits some sensitivity to pitch but remains in the speaking voice range (usually A2 to C3).
- 2 – "Speaking-Range Singer" sustains tones and exhibits some sensitivity to pitch but remains in the speaking-voice range (usually A2 to C3).
- 2.5 – "Inconsistent Limited-Range Singer" wavers between speaking and singing voice and uses a limited range when in singing voice (usually up to F3).
- 3 – "Limited-Range Singer" exhibits use of limited singing range (usually D3 to F3).
- 3.5 – "Inconsistent Initial Range Singer" sometimes only exhibits use of limited singing range, but other times exhibits use of initial singing range (usually D3 to A3).
- 4 – "Initial Range Singer" exhibits use of initial singing range (usually D3 to A3).
- 4.5 – "Inconsistent Singer" sometimes only exhibits use of initial singing range, but other times exhibits use of extended singing range (sings beyond the register lift: B3-flat and above).
- 5 – "Singer" exhibits use of extended singing range (sings beyond the register lift: B3-flat and above).

Although this revised scale actually has nine levels, it would seem cumbersome for a rater to learn nine behaviors. Therefore, the initial five levels remain: Children who fluctuate between any two levels receive a score of the lower of the two plus ".5." For example, a child who often uses a speaking range voice (2) but occasionally jumps into a limited singing range (3) would be scored as "2.5."

## SUMMARY

Research over the last 50 years has shown that all children can be taught to sing and that individual or small-group arrangements are more effective than large-group settings. However, with the percentage of problem singers increasing, these strategies have apparently not found application in the classroom. Teachers seem reluctant to ask children to sing alone or in small groups. Identification of effective games and activities that encourage individual and small-group singing in a nonthreatening classroom setting is necessary. These techniques would perhaps be more readily adopted by teachers. This study confirmed that use of these techniques has a significant effect on development of children's singing voices.

## REFERENCES\*

- Bintig, A. (1980). The efficiency of various estimations of reliability of rating scales. *Educational and Psychological Measurement*, 40, 619-643.
- Buckton, R. M. (1977). A comparison of the effects of vocal and instrumental instruction on the development of melodic and vocal abilities in young children. *Psychology of Music*, 5 (1), 36-47.
- Cobes, C. J. (1970). The conditioning of a pitch response using uncertain singers. *Dissertation Abstracts International*, 31, 1829A-1830A. (University Microfilms No. 70-19,398).
- DeYarman, R. M. (1972). An experimental analysis of the development of rhythmic and tonal capabilities of kindergarten and first grade children. In E. E. Gordon (Ed.), *Experimental research in the psychology of music: 8* (pp. 1-44). Iowa City: University of Iowa Press.
- Flohr, J. M. (1981). Short-term music instruction and young children's developmental music aptitude. *Journal of Research in Music Education*, 29, 219-223.
- Franklin, E., & Franklin, A. D. (1983). The uncertain singer. *Update*, 1 (3), 3-6.
- Gaiser, P. E. (1961). *A study of tone-matching techniques as remedial instruction for non singers*. Unpublished doctoral dissertation, University of Oregon, Eugene, OR.
- Goetze, M. (1985). Factors affecting accuracy in children's singing. *Dissertation Abstracts International*, 46, 2955A. (University Microfilms No. DA8528488).
- Goetze, M., Cooper, N., & Brown, C. (1990). Recent research on singing in the general music classroom. *Bulletin of the Council for Research in Music Education*, no. 104, 16-37.
- Gordon, D. S. (1985). A survey of literature and practice in assisting the pitch-defective singer in the elementary school. *Pennsylvania Music Educators Association: Bulletin of Research in Music Education*, 16, 11-18.
- Gordon, E. E. (1971). *The psychology of music teaching*. Englewood Cliffs, NJ: Prentice-Hall.
- Gordon, E. E. (1979). *Primary measures of music audiation*. Chicago: G.I.A. Publications.
- Gordon, E. E. (1986). *Primary measures of music audiation and the intermediate measures of music audiation*. Chicago: G.I.A. Publications.
- Gould, A. O. (1968). *Finding and learning to use the singing voice: A manual for teachers* (Report No. OEC-6-10-016). Macomb, IL: Western Illinois University. (ERIC Document Reproduction Service No. ED 025 531).
- Hale, M. R. (1977). An experimental study of the comparative effectiveness of harmonic and melodic accompaniment in singing as it relates to the development of a sense of tonality. *Bulletin of the Council for Research in Music Education*, no. 53, 23-30.
- Isaac, S., & Michael, W. B. (1982). *Handbook in research and evaluation*. San Diego, CA: Edits Publishers.
- Jarjisian, C. S. (1981). The effects of pentatonic and/or diatonic pitch pattern instruction on the rote-singing achievement of young children. *Dissertation Abstracts International*, 42, 2015A. (University Microfilms No. 8124581).



- Jersild, A. T., & Bienstock, S. F. (1931). The influence of training on the vocal ability of three-year-old children. *Child Development*, 2 (40), 272-291.
- Jersild, A. T., & Bienstock, S. F. (1934). A study of the development of children's ability to sing. *Journal of Educational Psychology*, 25, 481-503.
- Jones, M. (1979). Using a vertical-keyboard instrument with the uncertain singer. *Journal of Research in Music Education*, 27, 173-184.
- Klemish, J. (1974). Treating the uncertain singer through the use of the tape recorder. *Bulletin of the Council for Research in Music Education*, no. 37, 36-45.
- McNitt, H. H. F. (1972). A self-instructional program for adult uncertain singers. *Dissertation Abstracts International*, 32, 6479A. (University Microfilms No. 72- 13, 896).
- Miller, P. H. (1975). An experimental analysis of the development of tonal capabilities of first grade children. In E. E. Gordon (Ed.), *Experimental research in the psychology of music: 10* (pp. 77-97). Iowa City: University of Iowa Press.
- Pedersen, D. M., & Pedersen, N. O. (1970). The relationship between pitch recognition and vocal pitch production in sixth-grade students. *Journal of Research in Music Education*, 18, 265-272.
- Phillips, K. H. (1985). The effects of group breath-control training on the singing ability of elementary students. *Journal of Research in Music Education*, 33, 179-191.
- Phillips, K. H. (1992). *Teaching kids to sing*. New York, NY: Schirmer Books.
- Richner, S. S. (1976). The effect of classroom and remedial methods of music instruction on the ability of inaccurate singers, in the third, fourth, and fifth grades, to reproduce pitches. *Dissertation Abstracts International*, 37, 1447A, (University Microfilms No. 76-19,898).
- Roberts, E., & Davies, A. D. M. (1975). The response of "monotones" to a programme of remedial training. *Journal of Research in Music Education*, 23, 227-239.
- Roberts, E., & Davies, A. D. M. (1976). A method of extending the vocal range of "monotone" schoolchildren. *Psychology of Music*, 4, 29-43.
- Romaine, W. B. (1961). *Developing singers from non-singers*. Unpublished doctoral dissertation, Teachers College, Columbia University, New York, NY.
- Runfola, M. (1981). *An investigation of a technique for identifying early childhood uncertain singers*. Paper presented at the national meeting of the Music Educators National Conference, Minneapolis, MN.
- Rutkowski, J. (1984). *Development of a rating scale to assess individual children's use of the vocal instrument*. Unpublished manuscript, State University of New York at Buffalo, Buffalo, NY.
- Rutkowski, J. (1986). The effect of restricted song range on kindergarten children's use of singing voice and developmental music aptitude. *Dissertation Abstracts International*, 47, 4706A. (University Microfilms No. DA8619357).
- Rutkowski, J. (1989). The problem singer: What does research suggest? *General Music Today*, 2(3), 24-29.

- Rutkowski, J. (1990a). The measurement and evaluation of children's singing voice development. *The Quarterly: Center for Research in Music Learning and Teaching*, 1 (1-2), 81-95.
- Rutkowski, J. (1990b). *The comparative effectiveness of individual and group singing activities on kindergarten children's use of singing voice and developmental music aptitude*. Paper presented at the national meeting of the Music Educators National Conference, Washington, DC.
- Smale, M. J. (1987). An investigation of pitch accuracy of four-and five-year-old singers. *Dissertation Abstracts International*, 48, 2013A. (University Microfilms No. DA8723851).
- Smith, R. B. (1963). The effect of group vocal training on the singing ability of nursery school children. *Journal of Research in Music Education*, 11, 137-141.
- Sterling, P. A. (1985). A developmental study of the effects of accompanying harmonic context on children's vocal pitch accuracy of familiar melodies. *Dissertation Abstracts International*, 45, 2436A. (University Microfilms No.DA8425896).
- Updegraff, R., Heliger, D., & Learned, J. (1938). The effect of training upon singing ability and musical interest of three, four, and five-year-old children. *University of Iowa Studies in Child Welfare*, 14, 83-131.
- Whitman, C. F. (1971). The effects of piano instruction on vocal pitch accuracy in fourth, fifth, and sixth grade students. *Dissertation Abstracts International*, 31, 5453A-5454A. (University Microfilms No. 71-17,438).
- Wolner, M., & Pyle, W. H. (1933). An experiment in individual training of pitch deficient children. *Journal of Educational Psychology*, 24, 602-608.

---

\* As referências bibliográficas deste artigo estão conforme o documento original.



### 2.6.3. Estudo Crítico – Por Wendy Syms

#### Objectivo e âmbito do estudo

A premissa subjacente a este estudo é que todas as crianças deveriam estar em condições de desenvolver a capacidade de cantar, dado que esta é uma capacidade comum a todos os seres humanos. Os educadores na área da música continuam à procura de metodologias eficazes para ajudar as crianças a desenvolver as suas vozes cantadas.

Um objectivo particular deste estudo foi examinar a eficácia de actividades cantadas no seio de grupos grandes versus grupos pequenos, ou individualmente, de modo a determinar as estratégias que melhor podem ajudar a desenvolver a voz cantada das crianças. A autora estava também interessada em saber quanto tempo seria necessário para que esta instrução tivesse efeito. Um outro objectivo era determinar se a prática vocal estava relacionada com o desenvolvimento da aptidão musical. A autora explica que, apesar de alguns autores e professores de música acreditarem que as crianças incapazes de reproduzir melodias têm baixa aptidão musical, existem poucas evidências empíricas que provem ou refutem esta ideia. Este aspecto do estudo foi baseado no trabalho de Edwin Gordon, autor do teste *Primary Measures of Music Audiation – PMMA*. Gordon defende que a aptidão musical de uma criança pode ser determinada através deste teste e que a aptidão musical continua a desenvolver-se até cerca dos 9 anos de idade, período após o qual estabiliza. Gordon acredita que a aptidão musical se pode desenvolver se se melhorar o canto das crianças antes desse período crítico terminar.

Este estudo mostra os resultados de duas experiências: um estudo piloto e um estudo principal. A variável independente para ambos os estudos foi o tipo de prática vocal proporcionada às crianças nas suas aulas de música.

A variável independente, na condição experimental, apresentava dois tipos de instrução: prática vocal ministrada no seio de um grupo grande e prática vocal ministrada num grupo pequeno e em actividades individuais nas aulas. Na condição de controle apresentava a instrução vocal apenas na situação de grupo grande.

Usaram-se três variáveis dependentes: *Singing Voice Development Measure (SVDM)* – *Medida do Desenvolvimento da Voz Cantada* – e os dois sub-testes do *PMMA* (Tonal e Rítmico). Tanto no estudo piloto como no estudo principal, foram administrados testes antes (Pré-Testes), a meio (testes intercalares) e no final do período de instrução (Pós-Testes). No estudo piloto, os testes anteriormente descritos foram aplicados nestas três situações. Depois, no estudo principal, baseado nos resultados do estudo piloto, o *SVDM* não foi usado a meio do período de instrução e o subteste Ritmo do *PMMA* nunca foi usado.

#### Design:

##### Pré-Teste / Testes intercalares / Pós-Testes

D1, D2, D3\* E D1\*, D2, D3\* E D1, D2, D3\*  
D1, D2, D3\* C D1\*, D2, D3\* C D1, D2, D3\*

Variáveis Dependentes (usadas para os Pré-Testes, testes intercalares e Pós-Testes):

D1 – SVDM (Medições do Desenvolvimento da Voz Cantada)  
D2 – PMMA (Tonal)  
D3 – PMMA (Ritmo)

Variáveis Independentes:

E – Grupo Experimental – Aulas incluindo prática vocal em grupos grandes, grupos pequenos e individualmente.  
C – Grupo de Controlo – Aulas incluindo prática vocal apenas em grupos grandes.

\* Só aplicado no estudo piloto.

## Revisão

A revisão bibliográfica salienta a importância da obtenção de mais informação acerca das formas como os professores podem ajudar as crianças a desenvolver as vozes cantadas. A revisão abrange desde aspectos genéricos como a necessidade de um indivíduo cantar, como parte integrante da sociedade, até assuntos mais específicos sobre o papel da instrução e da capacidade em aprender a cantar. Os métodos para correcção da desafinação e o desenvolvimento de cantores são também abordados.

A autora torna-se, então, mais específica e focaliza-se neste estudo, abordando o tipo de actividades através das quais a prática vocal pode ser proporcionada. A revisão bibliográfica conclui com a argumentação da necessidade do estudo e seus objectivos.

## Método

### Sujeitos

#### *Estudo piloto:*

Os participantes são crianças de cinco anos de idade, a frequentar um jardim de infância (primeiro ano de educação formal nos EUA),  $n=19$ . As crianças foram distribuídas aleatoriamente, sendo colocadas ou no Grupo Experimental ou no Grupo de Controle.

#### *Estudo principal:*

Os participantes são crianças de cinco anos de idade, a frequentar um jardim de infância,  $n=99$ . Turmas intactas foram submetidas às condições experimentais ou às condições de controle.

## Instrumentos

### a) *SVDM (Singing Voice Development Measure) – Medida do Desenvolvimento da Voz Cantada*

Este instrumento foi desenvolvido por Rutkowski para avaliar a voz cantada das crianças, em alternativa a outros instrumentos de avaliação mais centrados na afinação ou na capacidade de emitir sons com uma dada altura.

Para a administração deste instrumento, as crianças devem imitar um conjunto de padrões melódicos, procedendo-se então à gravação das suas performances vocais. Estas respostas são então avaliadas por um júri que utiliza uma escala de classificação de 1 a 5, respectivamente: 1- pré-cantor; 2- cantor com tessitura de fala; 3- cantor incerto; 4- cantor com tessitura inicial; 5- cantor. A autora baseia-se em investigações anteriores, que mostraram que esta seria uma medida válida para avaliar crianças de tenra idade.

No estudo piloto, os coeficientes de fiabilidade interjuizes (coeficientes de correlação indicando a concordância entre as classificações atribuídas por diferentes juizes) revelaram-se satisfatórios. No estudo principal, estes índices são elevados (Pré-Teste=.90, Pós-Teste=.99).

### b) *PMMA*

É um teste estandardizado para avaliar o desenvolvimento da aptidão musical de crianças. No estudo piloto foram administrados e analisados separadamente os dois sub-testes: o Tonal (*PMMAT*) e o Rítmico (*PMMAR*). No estudo principal, no entanto, só foi utilizado o Teste Tonal. Os itens destes testes são basicamente tarefas de discriminação auditiva. No caso dos itens do Sub-Teste Tonal, as crianças ouvem dois curtos padrões tonais, devendo responder se são iguais ou



diferentes. De forma similar, no Sub-Teste Rítmico, as crianças comparam um par de padrões rítmicos devendo responder se são iguais ou diferentes.

## **Procedimento**

### *Estudo piloto:*

Os participantes do estudo piloto foram sujeitos a um Pré-Teste. Foram colocados aleatoriamente num Grupo Experimental ou num Grupo de Controle. Tiveram aulas de prática vocal duas vezes por semana durante cerca de quatro meses, ao fim dos quais foram testados novamente. Em seguida receberam mais quatro meses de aulas e no termo da instrução foram testados com um Pós-Teste.

Ambos os grupos receberam uma instrução musical similar, mas a turma do grupo experimental oferecia às crianças a possibilidade de participar em jogos e actividades vocais, em pequenos grupos ou individualmente, pelo menos duas vezes por aula, no contexto de uma aula normal de música. Para permitir estas actividades, as crianças foram colocadas em pequenos grupos de forma aleatória. A aula, de 30 minutos, foi dividida em duas partes de 15 minutos, para que um assistente de investigação ensinasse canto a metade da turma, enquanto o professor de música continuava com o resto da turma, fazendo actividades de música não cantadas. Ao fim de 15 minutos trocavam de grupo.

### *Estudo principal:*

No estudo principal foi usado o mesmo plano de aulas que no estudo piloto. No entanto, ao visar-se um maior número de alunos, foi necessária a utilização de turmas completas. Assim, as turmas foram colocadas aleatoriamente no Grupo de Controle ou no Grupo Experimental e receberam, do professor habitual, aulas de música de 30 minutos uma vez por semana durante todo o ano escolar (cerca de 9 meses).

## **Resultados**

### *Estudo piloto:*

Foram efectuadas análises de covariância para os testes intercalares e para os Pós-Testes, usando o Pré-Teste como covariante.

Os resultados indicaram que o Grupo Experimental teve um desempenho significativamente melhor do que o Grupo de Controle nas seguintes situações: no *PMMAT* do teste intercalar e do Pós-Teste; no *PMMAR* do teste intercalar; no Pós-Teste *SVDM*. Isto forneceu informação que a autora usou para decidir com que frequência administrar os testes no estudo principal.

### *Estudo principal:*

Os resultados indicaram que ambos os grupos tiveram ganhos nos resultados do *PMMAT* ao longo do estudo, não tendo sido encontradas diferenças significativas entre os dois grupos. Porém, no Pós-Teste *SVDM*, o Grupo Experimental teve um desempenho significativamente superior ao Grupo de Controle. Análises posteriores indicaram que este resultado não foi influenciado pelo *SVDM* dos estudantes no Pré-Teste (ou seja, estes estudantes não tinham um desempenho mais elevado no *SVDM* logo à partida, no início do estudo). As correlações de Pearson calculadas entre o *PMMAT* e o *SVDM* foram baixas, o que indica uma pequena relação entre o uso da voz cantada e o desenvolvimento da aptidão musical.

## Conclusões

A autora explica os resultados relacionados com os testes de aptidão tonal e o canto. Concluiu que o ensino em grupos pequenos / individualmente que decorreu durante as aulas normais de música teve um efeito positivo no desenvolvimento da voz cantada das crianças. Com efeito, no Grupo de Controle houve um decréscimo dos resultados *SVDM* entre o Pré-Teste e o Pós-Teste, enquanto no Grupo Experimental essas pontuações aumentaram.

O estudo piloto indicia uma relação positiva entre os resultados no teste *PMMAT* e a instrução musical recebida, tal não se verificando no estudo principal. A autora especula sobre as diferenças encontradas e atribui-as ao facto de os professores serem diferentes nos dois casos, ou às diferenças de planificação tendo em conta que no estudo piloto as aulas eram mais curtas embora mais frequentes.

A autora explica que a fraca relação encontrada entre o uso da voz cantada e o desenvolvimento da aptidão musical tonal poderá dever-se à limitada escala (entre 1 e 5) da pontuação do *SVDM*. Mas, mesmo com esta advertência, ela conclui que os professores não devem presumir que o desempenho vocal das crianças é uma forma de avaliação da aptidão musical.

A partir dos resultados do estudo, são fornecidas recomendações para investigadores e professores. Apoiada nas conclusões do estudo piloto, que indicou que meio ano lectivo não é suficiente para que um efeito seja visível, essas recomendações defendem a necessidade de uma duração adequada de aplicação dos procedimentos nos estudos experimentais com crianças pequenas.

É recomendada mais investigação sobre a duração e frequência óptima das aulas de música.

A autora sugere que investigações deste tipo poderiam ser mais válidas se fossem limitadas a crianças que já sabem usar as suas vozes cantadas em vez de incluir crianças que ainda estão em estádios diferentes de desenvolvimento vocal, como era o caso da maioria das crianças deste estudo. A autora recomenda também a revisão do *SVDM*, aumentando a sua sensibilidade discriminativa como instrumento de medição, de modo a melhor caracterizar crianças cujas performances flutuavam entre níveis da escala.

No seu sumário, a autora reitera a conclusão de que o ensino vocal individual ou em grupos pequenos tem efeitos positivos no seio de uma turma normal de música. Também acentua a necessidade de materiais e actividades que encorajem os professores a usarem a instrução de canto individualmente e em grupos pequenos, nas aulas de música.

Este estudo contribui de diversas formas para os campos da pedagogia e da investigação. Desde a sua publicação, a metodologia de ensino empregue e recomendada tem sido efectivamente mais aceite, dado que o estudo fornece uma boa validação para essas práticas.

O *SVDM – Singing Voice Development Measure*, revisto, é um excelente instrumento para investigadores e professores que desejem um meio sistemático, conveniente e acessível, de medir o desenvolvimento vocal infantil. Esta escala poderá ser usada facilmente pelo professor enquanto observa o comportamento vocal das crianças no contexto da aula de música.

Metodologicamente, a investigação aqui descrita faz uma importante advertência: alguns estudos falham na identificação de métodos ou de materiais de ensino válidos porque a duração da sua implementação não é feita num período de tempo suficientemente longo para gerar efeitos.

## Contribuição para a área de estudo

Este estudo fornece um bom modelo de investigação.

A revisão bibliográfica é minuciosa e resume bem um grande corpo de investigação. Foi bem estruturada, indo de um conteúdo mais generalista até especificidades mais directamente relacionadas com este estudo. Esta secção faz uma boa fundamentação, baseada em aplicações lógicas da bibliografia relacionada com a investigação.

Um aspecto potencialmente problemático da metodologia foi a utilização de turmas completas no estudo principal, uma vez que as turmas desenvolvem frequentemente



“personalidades” distintas que podem afectar as respostas decorrentes da instrução que lhes é ministrada.

No entanto, esta necessidade de utilização de turmas inteiras é frequentemente uma necessidade prática em investigação baseada em meio escolar, uma vez que as escolas se mostram compreensivelmente relutantes em quebrar os seus horários normais e atribuir aleatoriamente as crianças a turmas diferentes.

Por outro lado, o uso de grupos intactos fornece uma visão realista, já que é dessa forma que o ensino é ministrado nas escolas, e talvez a única opção para os investigadores poderem trabalhar com crianças em grupos de aulas de música. A investigadora fez o melhor possível para contrabalançar efeitos negativos decorrentes do uso de turmas inteiras. Nomeadamente: 1) usou diversas turmas (note-se que o número de turmas usadas deveria estar especificado no artigo; dado o número de participantes por grupo, parecem terem sido quatro turmas); 2) atribuiu as turmas ao Grupo de Controle e ao Grupo Experimental de forma aleatória; 3) utilizou sempre o mesmo professor para todo o ensino ministrado em cada um dos estudos.

A utilização de vários instrumentos de avaliação, o exame dos efeitos da instrução ao longo do tempo em cada resultado dos testes e a análise de possíveis relações entre os vários resultados dos testes, proporcionou ampla informação que pode ser retida deste estudo. Como já foi referido, o *SVDM* é um instrumento útil e valioso.

Uma palavra de prudência é necessária em relação ao uso do *PMMA* como medida de “aptidão”. A definição de aptidão musical baseada num conjunto de itens que testam apenas a discriminação melódica ou rítmica pode ser problemática. O potencial musical das crianças é certamente baseado em mais do que as capacidades requeridas neste estudo – escutar, lembrar e comparar padrões ouvidos fora de um contexto musical.

Parece razoável considerar os sub-testes *PMMA* como medidas de discriminação auditiva musical, mas tecer outras conclusões baseando-as nestes testes poderá ser inapropriado. As conclusões deste estudo teriam sido tão válidas e significativas – e, na minha opinião, até mesmo mais precisas – se o termo “desenvolvimento da aptidão musical” fosse substituído por “discriminação tonal” e/ou “discriminação rítmica”.

A quantidade de tempo aplicada neste estudo é certamente uma das características fortes deste projecto. Não é muito frequente a realização de estudos que implicam instrução durante um ano, e ainda menos durante dois. Além disso, a utilização de um estudo piloto tão extenso não é comum. Devido à análise cuidadosa e às ideias retiradas dos procedimentos e dados do estudo piloto, a investigadora foi capaz de tomar decisões que melhoraram a metodologia e a interpretação do estudo principal.

Pode ser difícil relatar, num só artigo, dois estudos relacionados, mas a autora foi bem sucedida nesta tarefa. O resultado é um artigo que está claramente organizado e bem escrito. As análises de dados estão explicadas com clareza e foram seleccionadas para responder às questões da investigação. É de notar que as conclusões e recomendações estão equilibradas, reflectidas e apropriadamente focadas nos resultados obtidos. Há muito a aprender com este artigo, não só pelo seu conteúdo interessante e útil, mas também pela excelente estrutura do relatório escrito, em si.

# 2.7.

## THE MATCH BETWEEN MUSIC EXCERPTS AND WRITTEN DESCRIPTIONS BY FIFTH AND SIXTH GRADERS

CORRESPONDÊNCIA ENTRE EXCERTOS MUSICAIS E DESCRIÇÕES  
ESCRITAS REALIZADAS POR ALUNOS DOS 5º E 6º ANOS DE ESCOLARIDADE

P.J. FLOWERS (2000)

JOURNAL OF RESEARCH IN MUSIC EDUCATION, 48 (3), 262 - 277

### 2.7.1. Resumo do artigo / Abstract

#### Resumo

O propósito deste estudo foi examinar o efeito da instrução musical na descrição escrita de excertos musicais. Alunos dos 5º e 6º anos receberam quatro dias de instrução sobre descrição musical, centrada em elementos analíticos, metáforas/emoções e vocabulário relacionado com o *tempo*<sup>1</sup>. Em seguida foi-lhes proposta a audição de seis excertos musicais, organizados e divididos por instrumentação e estilo, sobre os quais redigiram breves descrições. Estas descrições foram então comparadas com outras, realizadas por estudantes dos mesmos anos mas sem instrução prévia. Não foram encontradas diferenças significativas nas descrições de alunos que se encontravam em anos iguais. Contudo, os alunos do quinto ano que receberam instrução utilizaram mais vocábulos e elementos musicais analíticos nas suas descrições do que os que não foram alvo de qualquer tratamento. No escalonamento estabelecido por um grupo de juízes que fizeram corresponder as descrições realizadas pelos sujeitos com os excertos originais, não se verificaram diferenças importantes entre alunos do mesmo ano. No entanto, os alunos do 6º ano foram mais eficazes do que os do 5º ano. Na maioria dos casos, quando os excertos não corresponderam correctamente, os sujeitos mostraram-se confundidos por outros excertos com instrumentação igual, corroborando a prioridade de nomear instrumentos na execução de tarefas de descrição musical.

#### Abstract

The purpose of this study was to examine the effect of instruction in writing about music on descriptions of music excerpts. Fifth and sixth graders received 4 days of instruction in music description focusing on metaphor and emotion, analytic music elements, and temporal language. Then, they wrote brief descriptions of six excerpts that were counterbalanced for instrumentation and style. Their descriptions were compared to those of students who did not receive instruction. There were no significant differences between children receiving and not receiving instruction within each grade level, except that the taught fifth graders mentioned significantly more musical elements than did fifth graders who did not receive instruction. There were no significant differences within grade levels in the rate at which expert readers matched descriptions to the original music excerpt, but sixth graders significantly outperformed the fifth graders. When excerpts were not correctly matched, they were usually misassigned to the excerpt with the same instrumentation, corroborating the priority of naming instruments in music description tasks.



## 2.7.2. Artigo original

# The Match between Music Excerpts and Written Descriptions by Fifth and Sixth Graders

P.J. Flowers (2000)

The purpose of the present study was to examine the integration of music and language arts in the elementary school curriculum, specifically, student writing about music using three types of description – (1) metaphor and emotion, (2) analytic elements of music, and (3) temporal change and sequence – to communicate information that would enable a reader with musical expertise to identify the music about which the description was written. This project was designed to extend previous research in vocabulary development, analytic description of musical elements, and responses to written description tasks. In addition, studies using nonverbal measures of focus of attention on musical elements were relevant in interpreting the outcomes of the study.

## COMING TO TERMS

### Vocabulary and Language in Describing Music

Music vocabulary and descriptive abilities do not develop in isolation from other life experiences, and much of what children know comes from incidental learning outside of school. There is a high degree of consistency in words used by first graders to describe music experiences and other experiences (Flowers, 1998), suggesting that, at least in the early years, most words are not reserved for a particular discipline. Words describing musical elements are among this vocabulary, including the well-known terms "loud," "soft," "fast," and "slow" (Cassidy, 1993; Hair, 1981), as well as such polysemous words as "high" and "long." Children and adults know the names of many individual instruments and their families, and describing instrumentation often constitutes a large proportion of their analytic responses to music (Bundra, 1993; Cassidy 1994; Ellis, 1999; Flowers, 1984). Beyond using basic terms for instrumentation, dynamics, and tempo, children and even adults struggle for words to describe melody, pitch, texture, and mode (which they sometimes try to explain), and they have virtually no vocabulary with which to describe harmony (which is rarely mentioned at all) (Cassidy & Speer 1990; Costa-Giomi, 1994a, 1994b, 1996; Flowers, 1983a, 1983b, 1984; Hair, 1981). With or without specific music vocabulary, adult listeners often attempt to describe analytic elements of music, particularly following music instruction (Cassidy & Speer, 1990; Ellis, 1999; Flowers, 1983a) and when no visual stimuli are present (Geringer, Cassidy & Byo, 1996).

Beyond technical music vocabulary, there exists a rich lexicon that is used in daily language and may be applied to music. Listeners use figurative language and describe emotions to communicate their personal experiences and associations with music (Bundra, 1993; Flowers, 1983a, 1984; Smith, 1996). In examining cultural differences in music description, Morrison and Yeh (1999) found that university students from the United States focused more on analytic elements than did Asian undergraduates, who made more metaphorical statements. Instructional focus on technical vocabulary and consistently applied terms have led some researchers to conclude that the "good stuff" gets thrown away; that is, the imaginative and unique descriptions of sound and emotion are sometimes ignored (Flowers, 1996; Furman, Hersh, & Shiraishi, 1995).

In addition to looking at music vocabulary and description, several researchers have focused on the process and product of writing about music in different grade levels and settings. The use of extramusical or figurative descriptions associated with musical understanding and



terminology has been cited as one of the most promising outcomes of these efforts (Bergee & Crawford, 1995; Larson & Merrion, 1987; O'Brien, 1992). Furthermore, writing about music might help enhance memory of orchestral excerpts when unique figurative language is used (Flowers, 1985).

## Focus of Attention

Researchers have used both verbal and nonverbal means of determining which musical elements are perceived as most salient in music listening tasks. A substantial body of literature exists suggesting that verbal report is not a good measure of what children hear (Hair, 1985) and favoring ongoing behavioral tasks over verbal description as a measure of music perception or focus of attention (Madsen & Geringer, 1990). Although some attempts have been made to encourage description beyond typical static responses by repeated hearing and describing (Aiello, 1996) and by ongoing nonverbal measures followed by verbal analysis (Flowers, 1983a), there are definite limitations to the generalizations that can be made from written description to perception. In fact, there are numerous discrepancies between what students say about music and how they perform on nonverbal tasks. For example, children and nonmusician adults often refer to timbre, tempo, dynamics, and sometimes melody in written tasks, yet research using a continuous response digital interface (CRDI) has shown, for instance, that timbre is not usually considered the most salient feature of music, and when it does receive focus of attention, there is a commensurately lower preference for that excerpt (Madsen & Geringer, 1990; Geringer & Madsen, 1995/1996). Comparisons of studies using different measurement tools suggest restraint in using verbalization as a dependent variable for listening. Still, oral and written expression are valued behaviors that represent a viable avenue of research about music and experience.

The present study was designed to replicate and extend findings of a previous investigation in which children in Grades 2-5, undergraduate education majors, and undergraduate music majors listened to music and attempted to write descriptions that could be matched back to the original excerpt after being randomly ordered (Flowers, 1996). Results of that study showed that subjects at all age levels wrote descriptions that were matched to excerpts beyond rates of chance, with an increase at every age level except Grade 2 and Grade 3. Descriptions by fifth graders were matched correctly 41.5% of the time, and those of music majors were correctly matched in 92% of the cases. Qualitative observations were made to distinguish descriptions of the college undergraduates (who were more successful) from those of the children (who were less successful). One noticeable difference was that the undergraduates described musical change over time, using such temporal language as "begins," "returns," "climax," "later," and "ends." They also described balance of voices and texture, again incorporating relational language. The children's descriptions of musical elements were more static; rarely did they talk about temporal change or sequence. Another noticeable difference was the breadth of categories encompassed by the undergraduates' descriptions. The undergraduates used technical vocabulary, nontechnical vocabulary, and included references to imagery, emotion, and metaphor. Typically, the children named instruments, passed judgment, or described an emotion. Rarely did they do all three. In this study, similar to previous research (Flowers, 1988), the 3rd-5th graders were redundant in their application of terms to excerpts – relatively few of the frequently used words were applied uniquely to any one excerpt. The most familiar terms were not used to distinguish one excerpt from another and their liberal application across excerpts may have undermined the match-to-excerpt task.

The present study has built on knowledge of how children write about music with additional attention to the use of temporal language in their written descriptions. Research questions focused on (1) the effect of instruction in writing about music on subsequent student descriptions and expert evaluators' match of descriptions to excerpts after random ordering; (2) characteristics of the excerpts (instrumentation or style) that were commonly mismatched; (3) word count; and (4) vocabulary of successful students.



## METHOD

The study was conducted in a suburban elementary school in central Ohio. A total of 84 children from two fifth-grade classes (19 students each) and two sixth-grade classes (23 students each) participated in the project. Forty-nine percent of the participants were girls, and 51% were boys. Most of the children had received general music instruction since kindergarten, and some were enrolled in private piano lessons, band, or string instruction. The study was completed while the school was pilot-testing an arts schedule in which children attended music class daily for a week followed by 3 weeks in other arts before the next music rotation. Instruction was provided to one fifth-grade class (5-1) and one sixth-grade class (6-1) daily, Tuesday through Friday, during 1 week in January. Children in the other two classes (5-2 and 6-2) wrote music descriptions in a single class period later that month.

### Daily Instruction

Instruction to children in 5-1 and 6-1 focused on writing about music emphasizing metaphors, musical elements, and temporal language. Each lesson lasted about 45 minutes and followed the same general plan: (1) presentation and discussion of the lesson objective, including sample terms and phrases that could be used in music descriptions, (2) listening to a piece of music and jotting down ideas for the description, (3) a second listening to the same piece of music followed by a quiet period of writing, and (4) voluntary reading aloud of written descriptions. Four lessons were presented on four successive days. Daily objectives and listening selections were:

- \* Day 1: To create metaphors and describe emotions evoked by "Hungarian Dance no. 5" by Brahms (2 minutes, 43 seconds);
- \* Day 2: To analyze and describe musical elements in "Dance of the Tumblers" from *The Snow Maiden* by Rimsky-Korsakoff (3 minutes, 50 seconds);
- \* Day 3: To write a story using temporal language to describe a sequence of events based on "Waltz" from *Sleeping Beauty* by Tchaikovsky (4 minutes, 36 seconds);
- \* Day 4: To analyze and describe musical elements, and to incorporate temporal language focusing on musical change and sequence in "Comedians Gallop" by Kabalevsky (1 minute, 51 seconds).<sup>1</sup>

Objectives focusing on metaphor, emotion, and musical elements were included because of the preponderance of these types of description in previous studies. The lessons objectives focusing on change or sequence were designed to help the children refine their metaphoric and music analytic skills and to facilitate a closer match between music excerpts and their descriptions. On Days 3 and 4, students were to include temporal language in their writing, not limiting their writing to a static description of sound, image, or emotion.

Data were collected on the frequency with which the children used (1) metaphor or emotion, (2) analytic musical elements, and (3) temporal language in their daily writing assignments. Discrete instances of metaphor or description of emotion were easily counted. Descriptions of musical elements were also obvious. In some cases, children mentioned terms that they had learned in class, for example "the music has texture," but said little to relate them to specific characteristics. These examples were counted as description of a musical element because they showed awareness or attention to that element, even though not much information was provided. There was no attempt to categorize responses as correct or incorrect. Descriptions of temporal change or sequence were the most subtle. Frequency was based only on explicit use of temporal language. Such words as "then," "next," "finally," "still," "all of a sudden," and "again" denoted sequence and change, as did the musical terms "crescendo," "gets higher," and so on. Some types of description were omitted from analysis, including nonspecific terms such as "nice" and



"different," and judgments such as "good" and "enjoyable." Only the three categories of metaphor/emotion, analytic musical elements, and temporal change/sequence were considered in analysis of the daily writing assignments. Descriptions were categorized and counted two times, 1 week apart, and any necessary adjustments were made to insure consistent application of categorical definitions.

### Six Excerpts (Posttest)

Following the four days of instruction in classes 5-1 and 6-1, children in all four classes (5-1, 5-2, 6-1, 6-2) listened to an audiotape of six excerpts and wrote short descriptions of each. As in a previous study (Flowers, 1996), they were instructed to describe the music so that someone who had not heard the music would know what it sounded like. They were told that their descriptions would be mixed up and someone would try to guess which excerpt they were describing. Each excerpt was played, and then the tape was paused until all children had stopped writing before going on to the next excerpt. The excerpts were about one minute in length. Three were from the Classical period and three were from the 20th century; two were performed on solo piano, two were by wind ensemble, and two were by string orchestra. The six excerpts were: (1) Beethoven, Sonata no. 6 in F for Piano, op. 10, no. 2; (2) Mozart, Serenade for Wind Octet in C minor, K. 388, "Nacht Musique"; (3) Debussy, "Cloches à travers les feuilles" from *Images*, Book 2; (4) Stamitz, Clarinet Concerto no. 10 in B-flat (string introduction); (5) Stravinsky, *Symphonies of Wind Instruments*; and (6) Bartok, Divertimento (string orchestra).

The six descriptions written by children in the four classes were studied in a variety of ways:

- (1) A primary question of the study was to determine the rate at which the written descriptions by fifth and sixth graders could be matched to the intended excerpt, and whether instruction in writing would increase the rate of correct identifications. To answer this question, the six descriptions written by each child were printed in random order and ten evaluators (in-service music teachers) attempted to correctly match them to the original excerpt. Evaluators were given an audiotape of the six excerpts and were allowed to play it as often as necessary to reorder the 6 descriptions of all 84 children. Number of correct matches was determined, and then group and grade level comparisons were made.
- (2) Assignment of excerpt to description by the 10 evaluators provided information about mismatches as well as correct matches; for example, were excerpts with similar instrumentation or similar style most often mismatched?
- (3) Total number of words was counted and compared across the four classes.
- (4) Descriptions of metaphor/emotion, analytic musical elements, and temporal change/sequence were categorized and counted using the same procedure as in analysis of the daily assignments. Comparisons were made between children receiving daily instruction and those receiving no instruction within each grade level.
- (5) Finally, written descriptions of the nine most successful students were studied for vocabulary or other evidence that contributed to their success.

## RESULTS

*Daily Instruction.* Children in classes 5-1 and 6-1 wrote daily descriptions of music as instructed: Day 1 – emotions and metaphors; Day 2 – analytic musical elements; Day 3 – a story with change or sequence of events; and Day 4 – temporal change or sequence of musical elements. Both classes succeeded at these assignments as shown by descriptive data presented in Table 1. For the most part, there was a low rate of description except in the category representing that day's writing objective. This is consistent with previous research in which children have focused on a single element or type of description to the exclusion of others. Day 4 is an



interesting case because two categories – musical elements and temporal language – were required. Clearly, these are not mutually exclusive, but were coded separately in the present study.

On Day 4, both fifth and sixth graders decreased in frequency of musical elements compared with Day 2, and in frequency of temporal change/sequence compared with Day 3. Yet, collectively, there is evidence that they described change or sequence of musical elements as was requested.

Table 1

*Means and Standard Deviations for Use of Metaphor and Emotion, Analytic Musical Elements, and Temporal Language by Day and Class*

|                                  | Day 1 | Day 2 | Day 3 | Day 4 | 6 Descriptions |
|----------------------------------|-------|-------|-------|-------|----------------|
| <b>Grade 5-1 (n = 19)</b>        |       |       |       |       |                |
| <i>Metaphor &amp; Emotion</i>    |       |       |       |       |                |
| Mean                             | 5.42  | .68   | .50   | .42   | 5.74           |
| SD                               | 2.93  | 1.29  | .86   | .77   | 5.13           |
| <i>Analytic Musical Elements</i> |       |       |       |       |                |
| Mean                             | .47   | 8.47  | .56   | 6.84  | 11.68*         |
| SD                               | 1.07  | 3.31  | 1.20  | 3.10  | 8.58           |
| <i>Temporal Language</i>         |       |       |       |       |                |
| Mean                             | .05   | 1.58  | 2.67  | 1.00  | 2.37           |
| SD                               | .23   | 2.06  | 2.14  | 1.70  | 3.80           |
| <b>Grade 5-2 (n = 19)</b>        |       |       |       |       |                |
| <i>Metaphor &amp; Emotion</i>    |       |       |       |       |                |
| Mean                             |       |       |       |       | 8.63           |
| SD                               |       |       |       |       | 6.48           |
| <i>Analytic Musical Elements</i> |       |       |       |       |                |
| Mean                             |       |       |       |       | 5.95 *         |
| SD                               |       |       |       |       | 4.01           |
| <i>Temporal Language</i>         |       |       |       |       |                |
| Mean                             |       |       |       |       | 0.90           |
| SD                               |       |       |       |       | 1.37           |
| <b>Grade 6-1 (n = 23)</b>        |       |       |       |       |                |
| <i>Metaphor &amp; Emotion</i>    |       |       |       |       |                |
| Mean                             | 6.05  | .91   | .61   | .35   | 6.39           |
| SD                               | 3.26  | 1.78  | .99   | .78   | 8.38           |
| <i>Analytic Musical Elements</i> |       |       |       |       |                |
| Mean                             | 1.05  | 10.83 | .70   | 8.65  | 17.39          |
| SD                               | 1.66  | 4.84  | 1.89  | 4.42  | 9.74           |
| <i>Temporal Language</i>         |       |       |       |       |                |
| Mean                             | .29   | .83   | 5.52  | 2.26  | 3.74           |
| SD                               | .90   | 1.47  | 3.59  | 2.80  | 5.28           |
| <b>Grade 6-2 (n = 23)</b>        |       |       |       |       |                |
| <i>Metaphor &amp; Emotion</i>    |       |       |       |       |                |
| Mean                             |       |       |       |       | 5.83           |
| SD                               |       |       |       |       | 5.03           |
| <i>Analytic Musical Elements</i> |       |       |       |       |                |
| Mean                             |       |       |       |       | 13.78          |
| SD                               |       |       |       |       | 6.16           |
| <i>Temporal Language</i>         |       |       |       |       |                |
| Mean                             |       |       |       |       | 4.61           |
| SD                               |       |       |       |       | 4.14           |

\*  $t(25) = 2.64, p < .02$ . Data in boldface represent the type of instruction provided on a given day.

## Six Excerpts (Posttest) Results

The primary research question pertaining to the six descriptions (the final writing task) that were randomly ordered and subsequently matched to excerpt by 10 evaluators was whether the daily instruction would increase the rate of matching compared with students at the same grade level who received no instruction, or whether there would be a grade level difference. Ancillary questions concerned whether focus on style or instrumentation was evident in mismatches; how many total words were used across the six descriptions; what was the effect of daily instruction on types of description; and what were the characteristics of the work of successful students.

The chi-square statistic was used to compare the two classes within each grade level on number of correct matches between excerpt and written descriptions. There was no significant difference between the number of correct matches of descriptions written by children in 5-1 ( $M = 41\%$  correct) and children in 5-2 ( $M = 47\%$ ) ( $\chi^2 = 4.32$ ,  $df = 1$ , n.s.); similarly, there was no significant difference between the number of correct matches of descriptions written by children in 6-1 ( $M = 51\%$  correct) and children in 6-2 ( $M = 55\%$ ) ( $\chi^2 = 2.30$ ,  $df = 1$ , n.s.). Apparently the instruction in writing about music did not improve the ability of students to write descriptions that could be matched to excerpt beyond those of students who received no instruction. There was, however, a significant grade level difference with an average of 45% correct for the fifth graders, and 53% for the sixth graders ( $\chi^2 = 78.38$ ,  $df = 1$ ,  $p < .001$ ).

The six music excerpts that the children described were selected with attention to instrumentation and style. Excerpts 1 (Beethoven) and 3 (Debussy) were performed on solo piano; excerpts 2 (Mozart) and 5 (Stravinsky) were performed by wind ensemble; and excerpts 4 (Stamitz) and 6 (Bartok) were performed by string orchestra. Excerpts 1, 2, and 4 represented the classical period and excerpts 3, 5, and 6 represented the 20th century. Table 2 shows that in all cases, the most consistent matches were to the targeted excerpt, ranging from 32% correct (fifth graders, Mozart) to 67% correct (sixth graders, Beethoven). Descriptions of the piano excerpts (Beethoven and Debussy) were the most easily matched at both grade levels. In all but one case, mismatched descriptions were usually assigned to the excerpt characterized by the same instrumentation. There was little evidence of mismatching according to style.

In a previous study (Flowers, 1996), children used more words in describing music at successive grade levels, and they achieved more success in having their descriptions correctly matched as word frequency increased. This was not so clearly the case in the present study. While there was a word frequency difference in favor of the students receiving instruction (5-1 and 6-1) [ $F(1, 80) = 2.87$ ,  $p = .053$ ], this was eclipsed by the significant interaction between group (instruction, no instruction) and grade level (fifth, sixth) [ $F(1, 80) = 7.07$ ,  $p = .009$ ]. Fifth graders who received daily instruction used the most words across their 6 descriptions ( $M = 48.84$ ), and fifth graders who received no instruction used the fewest ( $M = 29.11$ ). The sixth-grade groups were similar to one another in total word count (6-1  $M = 44.09$ , and 6-2  $M = 47.04$ ).

The content of the six written descriptions by all groups was coded and categorized exactly as were the daily descriptions written by the children who received instruction, that is, as metaphor/emotion, analytic musical element, or temporal change/sequence. Then, comparisons were made between the children receiving instruction and those not receiving instruction within each grade level (refer to Table 1). Again, the greatest differences between instruction and no instruction are apparent in the fifth grade (5-1 and 5-2, respectively). While there were no significant differences between the fifth-grade classes in use of metaphor/emotion and temporal language, there was a large and significant difference in description of musical elements (5-1  $M = 11.68$  and 5-2  $M = 5.95$ ,  $t(25) = 2.62$ ,  $p < .02$ ). Fifth graders who received instruction wrote mostly about analytic musical elements in their final six descriptions, whereas students who did not receive instruction wrote more about metaphor and emotion. In general, the sixth graders had higher frequencies in all categories than did the fifth graders, although they did not use significantly more words. There were no significant differences in number of references to metaphor/emotion, analytic musical elements, and temporal change/sequence between sixth grade students receiving and not receiving instruction.



**Table 2**  
**Assignment of Descriptions to Excerpt**

1 = Beethoven, 2 = Mozart, 3 = Debussy, 4 = Stamitz, 5 = Stravinsky, 6 = Bartok

Solo Piano ..... 1 and 3                      Classical ..... 1, 2, and 4  
Wind Ensemble ..... 2 and 5                  20th Century ..... 3, 5, and 6  
String Orchestra ..... 4 and 6

|                   | Beethoven        | Mozart           | Debussy          | Stamitz          | Stravinsky       | Bartok           |
|-------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| <b>Beethoven</b>  |                  |                  |                  |                  |                  |                  |
| Grade 5           | <b>219 (58%)</b> | 22 (6%)          | 44 (12%)         | 67 (18%)         | 9 (2%)           | 18 (5%)          |
| Grade 6           | <b>307 (67%)</b> | 11 (2%)          | 95 (21%)         | 14 (3%)          | 18 (4%)          | 13 (3%)          |
| <b>Mozart</b>     |                  |                  |                  |                  |                  |                  |
| Grade 5           | 24 (6%)          | <b>121 (32%)</b> | 35 (9%)          | 29 (8%)          | 104 (27%)        | 67 (18%)         |
| Grade 6           | 12 (3%)          | <b>185 (40%)</b> | 27 (6%)          | 26 (6%)          | 122 (27%)        | 86 (19%)         |
| <b>Debussy</b>    |                  |                  |                  |                  |                  |                  |
| Grade 5           | 48 (13%)         | 35 (9%)          | <b>225 (59%)</b> | 14 (4%)          | 41 (11%)         | 17 (4%)          |
| Grade 6           | 90 (20%)         | 18 (4%)          | <b>303 (66%)</b> | 9 (2%)           | 23 (5%)          | 17 (4%)          |
| <b>Stamitz</b>    |                  |                  |                  |                  |                  |                  |
| Grade 5           | 51 (13%)         | 38 (10%)         | 19 (5%)          | <b>157 (41%)</b> | 25 (7%)          | 90 (24%)         |
| Grade 6           | 30 (7%)          | 49 (11%)         | 8 (2%)           | <b>257 (56%)</b> | 33 (7%)          | 82 (18%)         |
| <b>Stravinsky</b> |                  |                  |                  |                  |                  |                  |
| Grade 5           | 12 (3%)          | 91 (24%)         | 40 (11%)         | 32 (8%)          | <b>154 (41%)</b> | 51 (13%)         |
| Grade 6           | 11 (2%)          | 127 (28%)        | 27 (6%)          | 25 (5%)          | <b>205 (45%)</b> | 63 (14%)         |
| <b>Bartok</b>     |                  |                  |                  |                  |                  |                  |
| Grade 5           | 25 (7%)          | 70 (18%)         | 18 (5%)          | 81 (21%)         | 47 (12%)         | <b>139 (37%)</b> |
| Grade 6           | 13 (3%)          | 65 (14%)         | 3(1%)            | 120(26%)         | 61 (13%)         | <b>198 (43%)</b> |

Finally, the work of the most successful students was examined for evidence that might be helpful in understanding characteristics of identifiable music descriptions. Nine children whose descriptions were perfectly matched by at least four of the ten judges were included in this analysis. Seven of these children were girls and two were boys. None came from group 5-1 (instruction); four were from 5-2 (no instruction); three were from 6-1 (instruction); and 2 were from 6-2 (no instruction). Five of the children's descriptions were matched perfectly by four of the ten judges; four of the children's descriptions were matched perfectly to excerpt by five of judges; and no child's descriptions were matched by six or more judges. Given the lack of significant differences between instructional groups and the small number of successful students under consideration (n=9), group generalizations are not recommended on the basis of this sample. Rather, the purpose was to look at characteristics of relatively more identifiable written descriptions as a basis for speculation and future research.

Overall, 74% of the descriptions written by the nine most successful students were correctly matched to excerpt. They used an average of 47 words across the six descriptions, ranging from 21 to 85. They made an average of 7.11 references to metaphor/emotion, ranging from 0-20; 13.89 references to analytic musical elements, ranging from 1-29; and 3.44 references to temporal change/sequence, ranging from 0-14. In word count and type, their descriptions were not much different from those of the other children. Finally, a list of every word they wrote for each excerpt was compiled and ordered by frequency of use. Words used two or more times per excerpt are presented in Table 3, and those that were applied only to that excerpt are in boldface.

Table 3

*Word Used Two or More Times by the Most Successful Students (n = 9)*

**#1 Beethoven**

piano – 9  
 then – 5  
 gets – 4  
 loud – 3  
 louder – 3  
**soft – 3**  
 fast – 2  
 happy – 2  
**melodies – 2**  
 slow – 2  
 relaxing – 2

**#2 Mozart**

woodwind(s) – 4  
**instrument – 3**  
 loud – 3  
**sad – 3**  
**eerie – 2**  
 fast – 2  
 gets – 2  
**staccato – 2**  
 then – 2

**#3 Debussy**

dream(y) – 5  
 piano – 5  
 slow – 3  
 fast – 2  
 gets – 2  
 played – 2  
 quiet – 2

**#4 Stamitz**

fast – 4  
 loud – 4  
 happy – 3  
 violin(s) – 3  
**dancing – 2**  
 forte – 2  
 high – 2  
 played – 2  
 string(s) – 2  
 woodwind(s) – 2

**#5 Stravinsky**

high – 5  
 low – 3  
**horn(s) – 2**  
 sounds – 2  
**squeaky – 2**  
 then – 2  
 weird – 2

**#6 Bartok**

loud – 3  
 low – 3  
**same – 3**  
 sounds – 3  
**background – 2**  
**family – 2**  
 fast – 2  
**jumpy – 2**  
**mezzo – 2**  
 string(s) – 2  
**time – 2**  
 violin(s) – 2

*Note: Words in boldface type were applied only to that excerpt.*

## DISCUSSION

Previous research about music description has focused on categorizing types of verbal responses and on determining focus of attention or perception from participants comments. It is doubtful whether a writing-about-music task is a sufficient means of assessing what children perceive or conceptualize. In fact, writing and listening are probably mutually distracting, even with pauses to focus on each behavior separately. Thus, it was not the purpose of the present study to fully classify all aspects of the descriptions, nor was it to infer what children hear through evidence of the written word. The purpose of daily instruction was to encourage children to listen carefully and communicate through written language features of the music that would be understandable by another person. The goal was to tell something that was identifiable in relation to the specific music excerpt. Whether these features were described metaphorically, emotionally, or analytically was left to the child. This objective required listening, thinking, and writing.



The primary research purpose of the study was to determine whether instruction in writing about music would produce descriptions that were more clearly representative of the music than those written by typical children who had received no instruction. The four daily writing assignments to classes 5-1 and 6-1 did not present new information to these children, although writing unique descriptions of music may have been a novel activity. Several children commented that they had learned about metaphors in second grade. They were prolific in their class contributions of emotional descriptors; there was evidence throughout the music room of instruction in musical elements; and writing a story as a sequence of events was familiar. The requirement of incorporating temporal language into their descriptions of musical elements on day 4 was somewhat of a new idea, but it seemed to be easily understood. Their ability to fulfill these writing assignments was shown by a rise in the frequency of the relevant category on the day the assignment was introduced. Nevertheless, practice in describing metaphor and emotion, analytic musical elements, and temporal events in daily writing did not distinguish subsequent descriptions from those of children who received no instruction, with the exception of the fifth graders dramatic increase in attention to analytic elements following instruction. Given the similarity of descriptions by children receiving and not receiving instruction, it follows that they were matched to excerpt at about the same rates.

Despite the lack of difference between instructional groups in the rate at which their descriptions were matched to excerpt, there was a significant grade level difference, with sixth graders descriptions being correctly matched an average of 8% more frequently than those of fifth graders. Previous research (Flowers, 1996) showed an increase in matching rate at successive grade levels, 3rd through fifth grade. It is especially interesting to observe the similarity between the fifth grades in this and the previous study (45% and 41.5% correctly matched, respectively). However, unlike students in the present setting, those in the 1996 study were from an urban school in which there was no systematic music instruction, and they wrote about 6 entirely different excerpts. This evidence points to a developmental factor in writing descriptions that clearly represent specific music excerpts, but how children's verbal prowess increases through the grades in relation to this task is not apparent from results of the present study.

Additional factors were examined to shed light on how unique and specific written descriptions might be developed. In the 1996 study, not only did students increase their rate of "matchability" at each grade level, but they also used more words. Results of the present study do not support the idea that more words are better, however, with students in 5-1 using the most words in their six descriptions but being matched to excerpt at the lowest rate. Furthermore, the nine most successful students were similar to the larger population in word count and type of description. The consistent application of vocabulary words to specific excerpts and not to others may help to identify the match between a description and its intended excerpt. Lack of uniquely applied descriptors has been observed in previous student work (Flowers, 1988, 1996). There is a tendency for children to overuse the music vocabulary words that they do know, applying them liberally to each excerpt without attempting to distinguish one piece from another through the precise use of language. The present study attempted to encourage unique descriptions by (1) giving practice in writing, and (2) making the final writing task a matching game, yet many children attempted to list everything they knew or heard, regardless of the most salient features of the excerpt.

While focus of attention was not the primary question in the study, analysis of mis-matches highlighted the importance of naming instruments in written descriptions. Previous research (Bundra, 1993; Cassidy, 1994; Cassidy & Speer, 1990; Ellis, 1999; Flowers, 1983a, 1983b, 1984) has shown that instrumentation often receives the most verbal attention among musical elements; it is usually the most frequently named or labeled characteristic. Instrument names are also among the most common vocabulary found in first grade basal series music books (Flowers, 1998). So, it was no surprise that in the present study, when a description was mismatched, it was most often assigned to the excerpt with the same instrumentation. For example, when descriptions of the Mozart wind ensemble excerpt were not correctly matched, they were most often assigned to the Stravinsky excerpt rather than either of the other two examples from the same style period (Beethoven or Stamitz). Despite the well-documented verbal attention that children and adults give to instrumentation, timbre as a predominantly salient characteristic seems to have a small or even inverse relationship to preference (Geringer & Madsen, 1995/1996) and aesthetic experience (Madsen, 1997). Perhaps this seeming inconsistency – high verbal attention to instrumentation but

little connection between timbre and preferential or aesthetic response – points to the difficulty that people have in explaining what is important about music and describing its emotional impact. Or, perhaps, it reflects the typical general music curriculum in which students study vocabulary pertaining to music elements, omitting discussion of emotional or temporal events.

Several teaching implications are supported by this study and its related literature. First, writing about music may provide opportunities for focused listening, development of music and other vocabulary, practice in writing mechanics, and communication skill. Writing about music is a good vehicle for integrating music with other academic objectives, and it seems to enhance rather than detract from music learning objectives. Furthermore, music provides an opportunity to concentrate on temporal events, an area that seems to be neglected given the proclivities of children to focus on static descriptions rather than patterns of change. Children in the present study were creative in their use of metaphor, open in their description of emotion, and abundant in their attention to analytic musical elements. Further refinement of instruction and research might focus on shaping these outpourings into more meaningful prose that communicates about the music and about personal experience. Obviously more research is needed to understand the developmental, social, perceptual, and educational factors that affect such communication.

---

#### NOTE

1. All listening selections were taken from the London compact disc 444 107-2, "Spectacular Dances," performed by the Royal Philharmonic Orchestra or the London Symphony Orchestra with Stanley Black.



## REFERENCES\*

- Aiello, R. (1996, April). *Rehearing music: A perceptual paradox*. Paper presented at the National Biennial In-Service Conference of Music Educators National Conference, Kansas City, MO.
- Bergee, M. J., & Crawford, J. L. (1995). Developing writing-across-the-curriculum projects in eighth grade band: An observational case study. *Missouri Journal of Research in Music Education*, 32, 1-23.
- Bundra, J. I. (1993). *A study of music listening processes through the verbal reports of school-aged children*. Doctoral dissertation, Northwestern University. (University Microfilms No. 9415701)
- Cassidy, J. W. (1994, April). *Children's written descriptions of music*. Paper presented at the National Biennial In-Service Conference of Music Educators National Conference, Cincinnati, OH.
- Cassidy, J. W., & Colwell, C. M. (1993, April). *Identifying characteristics of music: Children's ability to make multiple discriminations*. Paper presented at the 10th National Symposium on Research in Music Behavior, Tuscaloosa, AL.
- Cassidy, J. W., & Speer, D. R. (1990). Music terminology: A transfer from knowledge to practical use. *Bulletin of the Council for Research in Music Education*, no. 106, 11-21.
- Costa-Giomi, E. (1994a). Recognition of chord changes by 4- and 5-year-old American and Argentine children. *Journal of Research in Music Education*, 42, 68-85.
- Costa-Giomi, E. (1994b). Effect of timbre and register modifications of musical stimuli on young children's identification of chord changes. *Bulletin of the Council for Research in Music Education*, no. 121, 1-15.
- Costa-Giomi, E. (1996). Mode discrimination abilities among of young children. *Psychology of Music*, 24, 184-198.
- Ellis, M. C. (1999). Spoken responses about music excerpts before and after a course in music appreciation. *Contributions to Music Education*, 26, 7-32.
- Flowers, P. J. (1983a). The effect of instruction in vocabulary and listening on nonmusicians descriptions of changes in music. *Journal of Research in Music Education*, 31, 179-189.
- Flowers, P. J. (1983b). Aspects of concert attendance behavior of undergraduate nonmusic majors. *Contributions to Music Education*, 10, 19-26.
- Flowers, P. J. (1984). Attention to elements of music and effect of instruction in vocabulary on written descriptions of music by children and undergraduates. *Psychology of Music*, 12, 17-24.
- Flowers, P. J. (1985). The effect of written descriptions on memory of fifth graders and undergraduates for orchestral excerpts. In C. K. Madsen & C. A. Prickett (Eds.), *Applications of research in music behavior* (pp. 32-42). Tuscaloosa: The University of Alabama Press.
- Flowers, P. J. (1988). The effects of teaching and learning experiences, tempo, and mode on undergraduates' and children's symphonic music preferences. *Journal of Research in Music Education*, 36, 19-34.

- Flowers, P. J. (1996, August). *Uniqueness and redundancy in written descriptions of music excerpts by children and undergraduates*. Research presented at the 4th International Conference on Music Perception and Cognition, McGill University, Montreal, Canada.
- Flowers, P. J. (1998). Music Vocabulary of First Grade Children: Words Listed for Instruction and Their Actual Use. *Journal of Research in Music Education*, 46, 5-15.
- Furman, C. E., Hersh, S., & Shiraishi, I. M. (1995, April). *Verbal preferences and affect descriptors of junior high listeners*. Paper presented at the 11th National Symposium on Research in Music Behavior, Granville, OH.
- Geringer, J. M., Cassidy, J. W., & Byo, J. L. (1996). Effects of music with video on responses of nomusic majors: An exploratory study. *Journal of Research in Music Education*, 44, 240-251.
- Geringer, J. M., & Madsen, C. K. (1995/1996). Focus of attention to elements: Listening patterns of musicians and nonmusicians. *Bulletin of the Council for Research in Music Education*, no. 127, 80-87.
- Hair, H. I. (1981). Verbal identification of music concepts. *Journal of Research in Music Education*, 29, 11-21.
- Hair, H. I. (1985). Children's responses to music stimuli: Verbal/nonverbal, aural/visual modes. In C. K. Madsen & C. A. Prickett (Eds.), *Applications of research in music behavior* (pp. 59 – 70). Tuscaloosa: University of Alabama Press.
- Larson, C. M., & Merrion, M. (1987). Documenting the aesthetic experience: The music journal. In T. Fulwiler (Ed.), *The journal book* (pp. 254-260). Portsmouth, NH: Boynton/Cook Publishers.
- Madsen, C. K. (1997). Focus of attention and aesthetic response. *Journal of Research in Music Education*, 45, 80-89.
- Madsen, C. K., & Geringer, J. M. (1990). Differential patterns of music listening: Focus of attention of musicians versus nonmusicians. *Bulletin of the Council for Research in Music Education*, no. 105, 45-57.
- Morrison, S. J., & Yeh, C. S. (1999). Preference responses and use of written descriptors among music and nonmusic majors in the United States, Hong Kong, and the People's Republic of China. *Journal of Research in Music Education*, 47, 5-17.
- O'Brien, W. (1992). The effects of figurative language in music listening instruction. *Contributions to Music Education*, 19, 20-31.
- Smith, W. H. (1996, April). *The diversity of the music listening experience as reflected in thought-listed responses of first-year undergraduate music majors*. Paper presented at the National Biennial In-Service Conference of Music Educators National Conference, Kansas City, MO.

---

\* As referências bibliográficas deste artigo estão conforme o documento original.



### 2.7.3. Estudo Crítico – Por Debra Hedden

#### Objectivo

Este projecto foi elaborado no sentido de replicar e alargar o âmbito de um estudo anteriormente realizado pela autora. O seu objectivo primordial foi o de determinar se alunos de música dos 5º e 6º anos conseguiriam melhorar a sua descrição escrita sobre excertos musicais, quando ministrado ensino específico para o efeito.

#### Design

Um grupo, constituído por uma turma do 5º ano e uma turma do 6º ano, que não recebe tratamento, é utilizado como termo de comparação (grupo de controle) para o grupo-alvo, que é submetido a instrução musical e avaliado no termo desta.

#### Revisão do artigo

#### Introdução e revisão

Flowers coloca o problema da integração de ambas as artes, música e escrita, através da redacção sobre música. Esta é sustentada por três componentes descritivas: (a) metáforas e emoções, (b) elementos analíticos musicais e (c) *tempo* – alterações e sequência.

As descrições são utilizadas para "comunicar informação que permita ao leitor com conhecimentos musicais, identificar a música sobre a qual se está a escrever". Um dos objectivos deste estudo foi entrelaçar música e linguagem, a partir de descrições escritas da música que os estudantes ouviram.

A revisão bibliográfica apresentada proporciona informação valiosa sobre o tipo de linguagem normalmente utilizada em descrições musicais. A autora sustenta que o vocabulário e o aperfeiçoamento descritivo não são reciprocamente exclusivos, antes surgem através de "experiências de vida". Para mais, sugere que há vocábulos descritivos comuns que as crianças aplicam, independentemente da especificidade do assunto a retratar, por exemplo "forte", "suave", "rápido", "lento", etc. Tanto crianças como adultos demonstram destreza na identificação de instrumentos, evidenciando mais dificuldades na detecção de descritores relativos a aspectos melódicos, harmónicos ou de textura.

Complementando esta informação, é-nos dito que o ouvinte utiliza as suas experiências pessoais, envolvendo emoções e incluindo representações figurativas, quando escreve sobre música. A cultura pode desempenhar um papel importante no género de elementos descritivos aplicados: um estudo de 1999 sugere que alunos americanos e asiáticos se concentram em elementos distintos: os americanos focalizam-se mais na componente analítica, enquanto os asiáticos se centram em elementos metafóricos. Alguma investigação adicional indica que alguns ouvintes têm preferência pela linguagem técnica, podendo com esta substituir elementos descritivos de carácter criativo.

Outros estudos fizeram a abordagem do processo de redacção e o resultado final da escrita, a diversos níveis. Os resultados indicaram que a descrição figurativa pode ser uma ajuda efectiva para a melhoria das capacidades utilizadas no desenvolvimento destes dois aspectos, contribuindo também para a memorização de excertos orquestrais. Nestes estudos, as tarefas verbais e não verbais oferecem informação em sintonia com os resultados da audição. Em contraste com as descrições verbais, as apreciações comportamentais pareceram reflectir melhor a atenção dos ouvintes. Flowers adverte o leitor para o facto de poderem existir divergências entre as descrições verbais e não verbais da percepção de uma mesma música por parte dos

estudantes. Não obstante a existência de tais discrepâncias quando as crianças respondem a tarefas auditivas, a autora defende a perspectiva da realização de pesquisa sobre descrições musicais verbais e escritas, como valor maior para investigação em educação musical.

O presente estudo foi planeado no sentido de replicar parte de um estudo anterior (Flowers, 1996), no qual foi pedido a estudantes universitários de música, estudantes universitários de educação e crianças frequentando entre o 2º ano e o 5º ano, que fizessem descrições musicais escritas após audição de excertos, propostos de forma aleatória.

## Método

Este estudo esteve dividido em duas fases distintas. Numa primeira etapa participaram duas turmas do 5º ano e duas turmas do 6º ano ( $N = 84$ ). Uma turma de cada ano recebeu preparação diária de escrita descritiva sobre música, por um período consecutivo de quatro dias. As aulas distribuídas por quatro sessões, de 45 minutos cada, constavam de: (a) introdução ao vocabulário, (b) audição musical para treino da descrição escrita, (c) uma segunda oportunidade para ouvir a obra, seguida de tempo limitado para a escrita, e (d) partilha verbal voluntária das descrições escritas. Durante as sessões de instrução, as duas turmas que receberam preparação redigiram descrições com objectivos diários específicos.

Na primeira sessão, a *Dança Húngara nº5* de Brahms foi o foco de atenção para os alunos. Depois de ouvirem cerca de três minutos desta peça, foi-lhes solicitado que “criassem metáforas e descrevessem emoções”. No segundo dia a tarefa incluiu a análise e descrição de elementos musicais da ópera *A Donzela da Neve* de Rimsky-Korsakov (aproximadamente quatro minutos). Na sessão do terceiro dia foram ouvidos excertos do bailado *A Bela Adormecida* de Tchaikovsky (aproximadamente quatro minutos e meio), sobre a qual os alunos escreveram uma história, empregando linguagem relativa a aspectos temporais e sequenciais. Finalmente, na quarta sessão, a obra *Comedian's Gallop* de Kabalevsky (aproximadamente dois minutos) constituiu matéria para os alunos analisarem e descreverem elementos musicais e aplicarem linguagem temporal, de acordo com as alterações e sequências musicais.

As outras duas turmas não receberam qualquer instrução. No entanto, parte dos alunos tinha tido aulas particulares e a maioria tinha conhecimentos musicais genéricos, adquiridos desde o jardim de infância. Depois de as duas turmas terem terminado a preparação, todas as classes envolvidas no estudo ouviram seis excertos musicais. Os excertos, constituídos por obras de Beethoven, Mozart, Debussy, Stamitz, Stravinsky e Bartok, possuíam diferentes tipos de instrumentação e tinham, cada um, a duração média de um minuto. Os estudantes foram então instruídos a elaborar uma descrição que, quando proposta a um leitor externo, pudesse ser identificada com os excertos apresentados.

A segunda fase foi constituída pela análise de nove das melhores descrições escritas pelos estudantes. Em cada fase do estudo dez juízes fizeram uma contagem de palavras por descrição, que foi catalogada de acordo com a presença de elementos metafóricos/emocionais, componentes analítico musicais e alterações de *tempo*/sequência. Para além disso, as nove melhores descrições foram avaliadas quanto ao vocabulário utilizado e outros indicadores de eficácia.

## Resultados

As duas turmas que receberam instrução musical demonstraram fraco desempenho ao nível da descrição musical, apesar das sessões terem parecido úteis. No quarto dia, quando as descrições dos alunos das várias turmas foram misturadas e analisadas, tornou-se difícil reconhecer que alunos haviam sido sujeitos à instrução musical prévia.

A análise de dados concentrou-se: (a) no grau de correspondência entre descrições e excertos; (b) na percentagem de identificações correctas; (c) na facilidade com que os juízes ( $N = 10$ ) estabeleceram correspondências com os excertos; (d) na comparação dos estudantes por turmas e por anos; (e) na interpretação dos equívocos (normalmente devido à semelhança da



instrumentação), (f) na contagem do total de palavras empregues, (g) na contagem e categorização de metáforas/emoções, elementos analíticos musicais, e linguagem temporal/sequencial; e, (h) na análise das nove melhores descrições em termos de vocabulário.

As descrições dos alunos foram comparadas por Grupo (Controle vs. Experimental) e por Anos de Idade. Os dados daí resultantes incluíram a frequência da utilização de metáforas/emoções, elementos analíticos e linguagem de natureza temporal. Algumas vezes as crianças optaram pela utilização de termos musicais apenas relacionada com a simples menção dos vocábulos, sugerindo uma aplicação superficial e falta de compreensão dos mesmos. Mais que exactidão, o propósito do estudo foi objectivar descodificadores. Palavras com valor descritivo pouco concreto, como, por exemplo, “lindo” e “bom”, não foram incluídas na análise.

Os descritores foram agrupados em três conjuntos principais. Foram contabilizados e confirmados duas vezes para se assegurar a concordância da análise. A análise  $\chi^2$  foi conduzida no sentido de comparar quatro categorias de combinações entre descrições e excertos. Não foram encontradas diferenças relevantes entre as turmas dos mesmos anos, quer instruídas, quer não. Este resultado em particular, sugere que a preparação não melhorou a capacidade dos alunos para redigir descrições de acordo com os excertos. No entanto, ocorreu uma diferença significativa no que toca à comparação dos resultados por anos, uma vez que os alunos do 6º ano que tiveram instrução prévia obtiveram resultados 8% mais satisfatórios que os colegas do 5º ano.

A descrição dos excertos de piano foi a mais correcta de todos os grupos de instrumentos. Apenas uma das descrições dos excertos de piano estava em desacordo com as demais, devido a um equívoco na descrição do estilo desse excerto. As restantes discordâncias ocorreram devido a equívocos de correspondência na instrumentação.

A variedade na utilização de palavras deveu-se mais ao ano lectivo ao qual os alunos pertencem, e menos ao facto dos alunos estarem inseridos no grupo com instrução prévia ou no grupo de controle. No caso específico dos alunos do 5º ano, os alunos sujeitos a instrução prévia utilizaram mais palavras do que os colegas do mesmo ano pertencentes ao grupo de controle. Nenhuma das turmas do 5º ano diferiu muito em relação à utilização de metáforas/emoções ou descrição de *tempo*, mas demonstraram variações significativas no que diz respeito à identificação de elementos musicais. A turma de 5º ano que foi sujeita ao processo de preparação fez muito mais uso de elementos analíticos. Em contrapartida, os alunos da outra turma do mesmo ano, sem qualquer instrução, caracterizaram-se pela aplicação de metáforas.

Embora a contagem tenha demonstrado que os alunos do 6º ano recorreram a um maior número de palavras, não foram encontradas diferenças significativas no que diz respeito à frequência das categorias utilizadas.

Na segunda fase do estudo, foram analisadas as descrições de nove dos melhores alunos. As descrições dos alunos ( $N = 9$ ) levaram à identificação do excerto musical por, pelo menos, 4 juízes, representando alunos de todas as classes, excepto da turma de 5º ano que teve instrução. A correspondência entre descrições e excertos foi de 74%, com uma utilização de quarenta e sete palavras idênticas em seis dos textos. Foi observada uma forte utilização de elementos analíticos, com pouca recorrência a elementos do tipo metáforas/emoções e mudanças de tempo/sequência. Acima de tudo, as descrições eram bastante semelhantes às de outros estudantes dos diversos grupos.

## Discussão

Flowers argumenta convincentemente que o seu estudo foi delineado com o objectivo de facilitar a “audição, o pensamento e a escrita” das crianças, em vez de se centrar em comportamentos independentes observados na escrita e na audição. No entanto, cada um destes objectos poderia ter sido estudado isoladamente. A autora preocupou-se em que o estudo tratasse de uma descrição musical escrita facilmente inteligível para o leitor. Por esta razão foi dada particular atenção à instrução no âmbito da escrita, de forma a obter descrições que demonstrem maior representatividade musical.

As crianças tiveram oportunidade de decidir sobre os elementos que queriam descrever, tendo podido optar entre o metafórico, o emocional e o analítico. Os ensinamentos de escrita foram estruturados no sentido de fornecer informação sobre o efeito dos mesmos. A maioria dos



alunos teve aulas genéricas de música desde que entraram para a escola, tendo de antemão alguns conhecimentos sobre linguagem artística. Dado que os alunos tiveram a possibilidade de optar pelo tipo de descrição, é provável que só tenham aplicado a categoria da sua preferência. Para que, num futuro próximo, seja possível determinar as capacidades de utilização de outros descodificadores, será prejudicial a focalização apenas em elementos específicos escolhidos pelos sujeitos.

Constata-se neste estudo a tentativa de cruzamento de conhecimentos provenientes de conhecimentos genéricos, adquiridos informalmente, e de novos conhecimentos adquiridos através de instrução. Porque não houve diferenças significativas entre os dois grupos, no que diz respeito ao rigor descritivo, parece que a instrução prévia não melhorou a capacidade de correspondência com os excertos. De facto, as classes sem instrução específica obtiveram melhores resultados no que toca à correspondência com os excertos. Contudo, os alunos com instrução preparatória revelaram maior frequência na aplicação diária dos descodificadores com um objectivo preciso.

Também foi possível observar um impressionante aumento da presença de elementos analíticos, utilizados nos textos dos alunos de 5º ano. Esta parece ser uma parte relevante da informação, pelo facto de a atenção alargada e consistente, poder fornecer mais competências e comodidade nas tarefas de descrição musical. Apesar de os grupos instruídos e não instruídos não terem diferido quanto à correspondência entre descrições e excertos, a discrepância expressiva entre anos demonstrou que os grupos de alunos mais velhos foram mais capazes de fazer corresponder descrições com excertos (maior frequência na ordem dos 8%). Os alunos do 5º ano com preparação preliminar manifestaram maior frequência de palavras nas suas descrições, mas foram os menos precisos na correspondência com os excertos.

A maior precisão das discriminações dos alunos de 6º ano sugere que o desenvolvimento ocorre com a idade, com a maturidade e, possivelmente, com a experiência de audição e de escrita. Assim sendo, esta é uma informação essencial para os educadores de música, no sentido de planificarem as actividades de audição de modo a introduzirem vocabulário, de forma gradual e de acordo com o nível de desenvolvimento em que os alunos se encontrem. Estas actividades poderão fornecer melhor compreensão, à medida que os alunos experimentem mais actividades de audição e de escrita, através de um treino progressivo e consistente.

No estudo anterior de Flowers (1996), tinha sido demonstrado que o número de palavras, bem como a correspondência, aumenta a cada ano em que se encontravam os alunos. Comparativamente, os resultados do presente estudo demonstraram que os alunos instruídos da turma do 5º ano aplicaram o maior número de palavras, mas obtiveram o pior resultado no que diz respeito à correspondência. Transparece assim que estes estudantes se concentraram mais nas palavras e menos no rigor da execução da tarefa. Talvez o professor de música devesse planear as actividades auditivas abordando a atenção a cada elemento separadamente, até que os educandos atinjam maturidade suficiente para possibilitar o cruzamento de actividades e a justaposição de aprendizagens.

Quando foram examinadas as descrições dos 9 alunos melhor sucedidos, os resultados indicaram que o seu trabalho era similar ao da restante amostra. É pois possível inferir que se o professor fizer uma aplicação consistente de vocabulário específico para cada excerto, poderá concorrer para a melhoria das competências de correspondência dos alunos, que era um dos objectivos deste estudo. Esta poderá ter sido uma das razões que levou a que os alunos reduzissem a utilização de palavras comuns, incrementando a aplicação de linguagem específica.

Embora o estudo tenha sido desenhado de forma a oferecer treino de escrita e jogo de correspondência, os sujeitos foram pouco diligentes ao utilizarem linguagem corrente e vocabulário comum a todas as descrições em detrimento de uma caracterização minuciosa e singular de cada excerto. As faltas de correspondência demonstraram que os sujeitos identificaram incorrectamente os excertos. Por outro lado, estas faltas de correspondência também mostram que os sujeitos confundiram determinados excertos musicais com outros que apresentavam instrumentação semelhante.

Pese embora a instrumentação ser um elemento comum aos mais elementares livros sobre música e uma especificidade que tanto adultos como crianças reconhecem imediatamente, Flowers faz notar que se verificou neste estudo uma generalizada falta de atenção a aspectos tímbricos e de natureza estética. A autora refere que, dentro do contexto geral, talvez este seja o resultado da dificuldade que as pessoas têm em descrever música através da componente



emocional e/ou haver falta de atenção para com estes aspectos em particular nos *currícula* musicais.

São muitas as implicações pedagógicas que, segundo Flowers, aqui se adivinham. A escrita oferece às crianças a possibilidade de se concentrarem na audição musical, aumentar o vocabulário com esta relacionado, desenvolvendo técnicas de escrita e de comunicação expressiva. Por outro lado, a música facilita o enlace entre a escrita sobre música e outras áreas, promovendo a aprendizagem musical. Finalmente, a atenção deverá ser reforçada em torno dos aspectos *tempo* e sequência, elementos que normalmente são subvalorizados.

### **Contribuição para a Área de Estudo**

Este estudo fornece informação importante sobre instrução de curta duração, resumindo o seu efeito na escrita sobre música. Os resultados sugerem que a brevidade da instrução foi insuficiente para que se obtivessem diferenças significativas quanto à correspondência entre a escrita descritiva e os excertos musicais propostos. Por outro lado, Flowers alerta-nos para o cuidado a ter na generalização da percepção auditiva a partir de exemplos escritos, pois há competências dissimuladas em cada comportamento.

Ao constatar que ocorreu uma discrepância nos resultados entre os alunos de 5º ano e 6º ano, os professores poderão ter interesse em reflectir se cada nível de desenvolvimento dos alunos deverá requerer especificidades em termos de profundidade, amplitude e concentração na audição e na escrita.

Tendo em conta que as crianças utilizaram léxico semelhante ao longo das descrições, os educadores deveriam considerar a utilização de vocabulário particular, relacionado com o estilo e/ou género, no sentido de os alunos aprenderem termos que os ajudem a reconhecer aspectos estruturais e estilísticos de diferentes tipos de música.

Como hipótese, poder-se-iam destinar as lições de um dia à abordagem da história e instrumentação e as sessões de um dia subsequente à focalização em aspectos estruturais e estilísticos. Por outras palavras, as experiências de audição múltipla de uma peça musical seriam vantajosas para o cultivo da clareza de entendimento, aumento de vocabulário e melhoria das capacidades de audição e escrita.

### **Comentário**

Este artigo inclui uma excelente introdução, que delinea o problema com transparência. Flowers integra, com sucesso, a pesquisa por si realizada anteriormente, enquadrando-a com outros estudos pertinentes à revisão bibliográfica. Estamos, no entanto, em presença de um estudo com situações demográficas deliberadamente distintas do precedente. A autora explana também a dificuldade que a descrição musical representa para adultos, estudantes e crianças.

Os juízes forneceram informação relevante quanto aos equívocos que ocorreram, pois quase todas as descrições que não combinavam tinham a mesma instrumentação. A investigadora delimitou os excertos a três meios performativos e dois períodos históricos. A contagem de palavras e a categorização dos descodificadores foi aplicada às quatro turmas. Em seguida, as categorizações foram comparadas via análise não-paramétrica. Nessa ocasião, a frequência foi um meio valioso para a tradução dos dados. Para análise estatística dos descritores foram utilizados testes *t*, tendo os resultados sido minuciosamente apresentados e devidamente descritos.

Foi feita a apresentação das nove descrições melhor sucedidas, embora a sua finalidade tenha ficado pouco esclarecida. Flowers defende que foi útil a determinação dos melhores trabalhos. No entanto, estes só foram considerados perfeitamente coincidentes com os excertos por menos de metade dos juízes. Teria sido útil perceber melhor o significado de “melhor sucedidos”, por que apenas foram escolhidos este nove, e ainda como tudo isto contribuiu para os resultados genéricos do estudo.

Os professores de Música poderiam considerar um esforço concertado no sentido de se variar a instrumentação ao longo da actividade pedagógica para que, sem esforço, as crianças se pudessem concentrar noutros elementos. Fornecidas aos alunos as opções e os tipos de descrição disponíveis, é possível que a concretização da tarefa se tenha tornado mais dispersa que o necessário.

Os sujeitos foram instruídos a escrever descrições que facilmente fossem entendidas por qualquer pessoa. Por isso, é possível que as crianças se tenham concentrado na caligrafia, ortografia, gramática, ou na combinação destes factores, ao invés do teor descritivo.

Se tivermos em conta esta permuta de conteúdos, não é surpreendente que tenha havido falta de resultados significativos respeitantes à correspondência. Talvez os elementos apresentados tenham oferecido demasiadas opções de descrição. A brevidade da instrução, enquanto fonte de informação valiosa, poderia produzir resultados diferentes se fosse estendida por algumas semanas, focalizando a atenção em apenas um descodificador por sessão. Esta é uma ideia potencial para um futuro estudo.

Quando a informação respeitante ao desenvolvimento das crianças é adicionada à equação, pareceria que uma maior duração da preparação seria benéfica para determinar se a instrução tem efeito na descrição musical escrita.

Acima de tudo, o estudo parece ter sido bem conduzido, tendo havido mais interesse na informação genérica do que em procedimentos diferenciados. A autora respondeu às questões da sua pesquisa e forneceu uma grande quantidade de informação, complementada pelos seus trabalhos anteriores. No futuro parece importante estender este estudo a todos os graus de aprendizagem, focando aspectos mais particulares, para melhorar a compreensão da audição, da escrita e da descrição musical.

Em jeito de conclusão, este estudo providencia informação importante no que diz respeito a níveis intermédios de aprendizagem, estimulando futuras investigações. Flowers sugere que mais estudos seriam benéficos para uma maior compreensão dos factores que afectam a aprendizagem e o desenvolvimento musical, nomeadamente de natureza social, perceptiva e pedagógica.

---

**N.T.**

1- A palavra *tempo* refere-se, neste contexto, ao significado musical do vocábulo de origem italiana, com o intuito de definir a velocidade a que uma determinada peça musical deve ser tocada, mas com a indicação de aspectos relacionados com o carácter. Ex: *Allegro* – significa rápido, mas também alegre.





### 2.8.1. Resumo do artigo

#### Resumo

O estudo da performance musical em contexto real mantém-se uma área de investigação pouco desenvolvida. Este estudo é uma abordagem desse tipo, averiguando a fiabilidade interjuízes na avaliação de performance musical numa grande universidade (Universidade do Missouri, Columbia – E.U.A). Os participantes, elementos dos júris<sup>1</sup> nas provas de final de semestre, eram professores das diversas famílias de instrumentos: metais ( $n=4$ ), percussão ( $n=2$ ), madeiras ( $n=5$ ), canto ( $n=5$ ), piano ( $n=3$ ) e cordas ( $n=5$ ). Cada jurado completou uma escala de classificação com critérios específicos e atribuiu uma nota global a cada performance, por ordem alfabética<sup>2</sup>. Estas classificações não foram partilhadas com os outros participantes ou com os alunos. A fiabilidade interjuízes foi calculada a partir dos resultados totais e dos resultados parciais obtidos nas escalas de classificação e a partir das classificações globais atribuídas pelos juízes. Todas as permutas possíveis de dois, três e quatro juízes foram analisadas estatisticamente, tendo sido calculadas correlações médias, desvios-padrão e distribuições das amostragens. A fiabilidade interjuízes apresentou valores positivos, independentemente do número de juízes que constituía cada painel. Os coeficientes de fiabilidade obtidos a partir dos resultados totais nas escalas de classificação, bem como os obtidos a partir das classificações alfabéticas, foram estatisticamente significativos. Em todos os grupos (ressalva para a percussão que, com  $n=2$  participantes, impôs um critério de significância muito apertado) foram observados coeficientes de fiabilidade significativos nas subescalas<sup>3</sup>, com excepção de uma das subescalas da voz. No caso dos júris constituídos por um maior número de juízes (quatro ou cinco), os resultados parecem ser consistentes, não sendo, porém, significativamente mais fiáveis do que os da avaliação efectuada através da classificação alfabética. Não se verificou um decréscimo na média da fiabilidade, em proporção com a diminuição do número de juízes. No entanto, as permutas de dois e três juízes tenderam, em média, a evidenciar maior variabilidade e distribuição, e menor uniformidade do que os grupos constituídos por quatro e cinco juízes. Não se verificaram diferenças em termos de fiabilidade em função dos níveis de experiência profissional dos avaliadores. Neste contexto é recomendada a utilização de um júri constituído por um mínimo de cinco juízes para a avaliação da performance musical.



## Abstract

Assessment of music performance in authentic contexts remains an underinvestigated area of research. This study is an examination of one such context, the interjudge reliability of faculty evaluation of end-of-semester applied music performances. Brass ( $n = 4$ ), percussion ( $n = 2$ ), woodwind ( $n = 5$ ), voice ( $n = 5$ ), piano ( $n = 3$ ), and string ( $n = 5$ ) instructors evaluating a recent semester's applied music juries at a large university participated in the study. Each evaluator completed a criterion-specific rating scale for each performer and assigned each performance a global letter grade not shared with other evaluators or with the performer. Interjudge reliability was determined for each group's rating scale total scores, subscale scores, and the letter-grade assessment. All possible permutations of two, three, and four were examined for interjudge reliability, and averaged correlations, standard deviations, and ranges were determined. Full-panel interjudge reliability was consistently good regardless of panel size. All total score reliability coefficients were statistically significant, as were all coefficients for the global letter-grade assessment. All subscale reliabilities for all groups except Percussion (which, with an  $n$  of 2, had a stringent significance criterion) were statistically significant, with the exception of the Suitability subscale in Voice. For larger panels ( $ns$  of 4 and 5), rating scale total score reliability was consistently but not greatly higher than reliability for the letter-grade assessment. There was no decrease of average reliability as group size incrementally decreased. Permutations of two and three evaluators, however tended on average to exhibit more variability, greater range, and less uniformity than did groups of four and five. No differences in reliability were noted among levels of experience or between teaching assistants and faculty members. Use of a minimum of five adjudicators for performance evaluation in this context was recommended.

## 2.8.2. Artigo original

# Faculty interjudge reliability of music performance evaluation

M. J. Bergee (2003)

Performing is considered central to what one must know and be able to do if one is to learn music. As evidence, MENC: The National Association for Music Education (1996) has advocated "performing of music, alone and with others" as one of its foundational content standards, part of the National Standards for Music Education. Not surprisingly, then, courses of study leading to degrees in music and music education strongly emphasize acquiring performance knowledge and skills. According to the National Association of Schools of Music (1999), competence on at least one major performing medium should be expected for students pursuing any baccalaureate degree in music, with competence specified as performing a cross-section of repertory, developing technical skills, and acquiring the ability to sight-read. In its new licensing standards for teaching specialists in the arts, the Interstate New Teacher Assessment and Support Consortium (2002, p. 17) has recommended that music teaching specialists be "able to read, write, and perform music fluently. They [should be able to] sing or play an instrument with a high degree of expressive skill."

As with all core aspects of a standards-based curriculum, issues of reliability and validity in assessment arise. Standards and assessment "inescapably go hand in hand" (Music Educators National Conference, 1996, p.1). Standards have no meaning without appraisal of student progress toward meeting objectives (Stewart, 2002). Ideally, instruction, learning, and assessment constitute fully interactive systems, with each informing the others (Saunders & Holohan, 1997; Zdzinski, 1991).

With performance assessment, one must resolve its inherent potential for unreliability. Performance essentially is ephemeral; applying common values to a phenomenon so individual and shifting is difficult (Dixon, 2000). Only human observers are capable of evaluating the musical worth of a sound or sequence of sounds (Radocy, 1986). Evaluators, however, often have little clue as to how they arrive at given performance ratings (Fiske, 1983).

A growing number of researchers (e.g., Abeles, 1973; Bergee, 1988a; Horowitz, 1994; Jones, 1986; Nichols, 1991; Saunders & Holohan, 1997; Zdzinski, 2002) have advocated that performance assessment under the correct circumstances exhibit good reliability and validity. These investigators have succeeded in developing reliable and valid rating scales or rubrics for music performance. Groups of adjudicators have used these assessment devices to evaluate performances in given domains, often securing strong interjudge reliability and criterion-related validity.

Such studies have contributed substantively to the research base in performance assessment. Their chief purpose, however, was to clarify performance in a given domain and to construct valid assessment devices. To a substantial degree, assessment as found in these studies was *ex situ*. Performance assessment in more authentic, ecologically relevant contexts remains underinvestigated. Among a host of others, evaluation of applied music (private instruction) in higher education settings is one such underinvestigated context. Beyond their role in determining whether students have met learning objectives in performance, such evaluations have served as a standard of comparison for peer and self-evaluation in previous studies (Bergee, 1993, 1997).

In one of these studies (Bergee, 1997), however, I uncovered a pronounced lack of reliability among some faculty evaluators, which raises concern about the viability of such assessment and strongly suggests further examination. In the present study, therefore, I examined the interjudge reliability of faculty evaluation of end-of-semester jury performances in applied



music. If such examination is to be of use, it must consider three circumstances frequently encountered in actual situations: variability in size of adjudication groups, mode of evaluation, and adjudicator experience. With concern for variability in size of adjudication groups, such groups often are smaller than the seven or more Fiske (1975, 1983) has recommended. Panels of two or three are common. Thus, it seemed prudent to examine interjudge reliability across varied sizes of judging groups.

A second circumstance is the mode of evaluation employed. In jury evaluations, faculty members commonly write anecdotal comments and assign performances a letter grade. Owing to its absence of structure, such evaluation potentially exemplifies the "irregular and uncontrolled observations" (Abeles, 1973, p. 246) that frequently compromise performance evaluation. I therefore examined interjudge reliability in two modes of evaluation, criteria-specific rating scales and a global approach involving the assigning of letter grades to performances. Based on his interjudge reliability findings, Fiske (1975, 1977) has recommended global assessments. Rating scales, on the other hand, yield particularized information and encourage attention to all relevant performance aspects. As examiners in a recent study of conservatory performance assessment were divided in their preferences for criteria-based versus holistic assessment (Stanley, Brooker, & Gilbert, 2002), I chose to use and study both evaluation modes.

Pertaining to the third circumstance, adjudicators are commonly viewed as increasing in consistency as they gain experience. Faculty evaluation panels (as the ones in this study typify) often vary considerably in quantity of experience, however. I therefore studied whether lack of prior experience evaluating juries seemed to influence interjudge reliability.

## METHOD

### Assessment Instruments

Researchers often have used factor analysis to develop rating scales with good construct and content validity. First, an inclusive set of items describing aspects of performance in a given area (e.g., clarinet, voice) is analyzed for an underlying factor structure. After such a structure is determined and stabilized, items with high and relatively unitary loadings on a given factor are chosen to define factor subscales. These items then are paired with Likert-type, categorical response scales (a common technique for assessing performance in research contexts; e.g., Britten, 2002; Wapnick, Flowers, Alegant, & Jasinskis, 1993), and judges use the rating scale to assess performances and establish interjudge reliability. Such *facet-factorial* (after Butt & Fiske, 1968) rating scales have been developed in several solo performance domains, including brass (Bergee, 1988a), snare drum (Nichols, 1991), clarinet (Abeles, 1973), voice (Jones, 1986), and stringed instruments (Zdzinsky & Barnes, 2002).

*Brass.* For a study described in Bergee 1988a, I developed a rating scale originally for euphonium and tuba performance. I grouped 27 items under a four-factor structure: (a) Interpretation/Musical Effect, (b) Tone Quality/Intonation, (c) Technique, and (d) Rhythm/Tempo. Total score interjudge reliabilities were .94, .99, and .98 for three panels of judges; subscale (i.e., items under each factor) reliability estimates ranged from .89 to .99. In this same study (Bergee, 1998a), I determined criterion-related validity by comparing scale outcomes with a magnitude estimation evaluation and also with scores obtained from the MENC adjudication form for solo wind instruments. In a subsequent study (Bergee, 1988b), I found no loss of reliability when the scale was used to evaluate jury performances of all brass instruments.

*Percussion.* Nichols (1991) grouped 18 items (6 apiece) under three factors: (a) Technique/Rhythm, (b) Interpretation, and (c) Tone Quality. He reported a split-halves reliability of .91, interjudge reliabilities of subscales ranging from .47 for Tone Quality to .81 for Interpretation, and a total score interjudge reliability of .69. Concurrent validity, determined by a paired comparisons procedure, was .79.

Although Nichols's original scale was intended for snare drum, for this study's purposes it seemed feasible to modify the scale to assess mallet performance as well. (Juries required performances on both.) Items remained as they appeared on Nichols's scale with one exception



under Tone Quality, from "drum tone sounded muffled" to a more general "tone was characteristic." I decided to retain Nichols's Tone Quality subscale in spite of its relatively low reliability rather than develop a new assessment instrument. This study's resulting higher coefficient (.69) partially validated this decision, although the reliability of the Tone Quality caption remained lower than the other two (both .90).

Woodwinds. Abeles (1973) used the facet-factorial approach to develop a rating scale for clarinet performance. He grouped 30 items under six factors: (a) Interpretation/Musical Effect, (b) Intonation, (c) Rhythm/Continuity, (d) Tempo, (e) Articulation, and (f) Tone. Subscale reliabilities ranged from .58 to .98; total score reliabilities ranged from .93 to .98. Criterion-related validity was established via a paired-comparisons procedure. A subsequent study (Bergee, 1989) revealed good validity and interjudge reliability when the scale also was used for evaluating saxophone performance. As no rating scale has yet been developed to assess woodwind performance in general, I modified Abeles's scale for this purpose. No modification outside of the Articulation subscale was required; under Articulation, I used the three items dealing with generalized aspects (e.g., "free from tonguing noise") and omitted the two dealing specifically with clarinet articulation ("squeaked"; "tonguing produced thunkie sound").

Voice. Jones (1986), in developing a rating scale for solo vocal performance, established five factors: (a) Interpretation/Musical Effect, (b) Tone/Musicianship, (c) Technique, (d) Suitability/Ensemble, and (e) Diction. The scale produced high interjudge reliability total score estimates (>.89). Subscale reliabilities generally were good, but some adjudication groups reliability outcomes for the Suitability/ Ensemble factor were low. Jones determined criterion-related validity by comparing scale scores with evaluations using the MENC-NIMAC Vocal Solo Adjudication Form.

Piano. The facet-factorial approach has not been used to develop a rating scale for solo piano performance. For this study, I worked directly with piano faculty to develop such a scale. I first examined literature in piano performance adjudication (e.g., Wapnick et al., 1993) and compiled a set of items describing piano performance. Redundant items were eliminated and others modified for clarity. I then took the items to two members of the piano faculty, who examined them for appropriateness and completeness. Together, we grouped items under three broad categories that seemed best to explain the items: (a) Interpretation/Musical Effect, (b) Rhythm/Tempo, and (c) Technique. With the piano faculty's assistance, I chose the three most representative items under each category to define subscales. Use of the scale in this study resulted in Cronbach's alphas ranging from .85 to .96 for total scores, and from .72 to .91 for subscales.

Strings. Recently, Zdzinsky and Barnes (2002) developed a scale for stringed instrument performance. They grouped 28 items under five factors: (a) Interpretation/Musical Effect, (b) Articulation/Tone, (c) Intonation, (d) Rhythm/Tempo, and (e) Vibrato. Reliability for subscale and total scores ranged from .87 to .94. They established criterion-related validity by comparing scale scores with global ratings obtained via the MENC rating form for stringed instruments. They also compared scale outcomes with those obtained by a magnitude estimation procedure.

Because time was of the essence in these jury performances (scheduled at 10-minute intervals), and to address concerns expressed by some of the area coordinators about the amount of time required for evaluations, I limited all subscales to three items apiece. I had intended to choose the three subscale items exhibiting the highest loadings on a given factor; frequently, however, I encountered substantial redundancy. I ultimately selected subscale items that addressed clearly distinct performance aspects. Also, to economize evaluators' time, I converted all items to positive wording. (Items in the original scales had been worded either positively or negatively, with wording choice randomly determined.)

I paired each item with five-option, Likert-type scales whose response categories included SD (Strongly Disagree), D (Disagree), N (Neutral), A (Agree), and SA (Strongly Agree). In addition to completing a rating scale for each performer, evaluators assigned a grade to each performance ranging from A+ (an *excellent performance in all respects*) to F (an *exceedingly poor performance in all respects*). These grades were for the study's purposes only. On a separate form, each faculty evaluator awarded an additional grade and wrote comments shared with the performer as a summative assessment of the jury performance. I specified to evaluators that the students would not see the study's grades and that they would not be compared with the students' actual grades.



On the study's form, evaluators indicated whether they were faculty members or graduate teaching assistants and number of years' prior experience evaluating juries.

For quantification, I converted categorical responses to a 1-5 scale (Strongly Disagree = 1 to Strongly Agree = 5). I converted the grade assigned to a 13-point scale (F = 1, D- = 2, etc., to A+ = 13).

## Participants and Procedures

At the conclusion of a recent academic semester, I asked those brass ( $n = 4$ ), percussion ( $n = 2$ ), woodwind ( $n = 5$ ), voice ( $n = 5$ ), piano ( $n = 3$ ), and string ( $n = 5$ ) faculty members and teaching assistants who were to evaluate that semester's juries (of graduate and undergraduate music majors and minors only) to participate in the study. All agreed. (For percussion juries, the panel usually consists of the percussion faculty member and one alternating member of the brass faculty. Thus, I tallied one participant twice – as a member of the brass group and again as a member of the percussion group. It was this individual's first time evaluating percussion juries.) Of the evaluators, two, both in the voice area, were graduate teaching assistants. One of the faculty members in the string area was part time but had occupied her position for over 20 years. The remaining evaluators all were full-time faculty members. Of the 24 evaluators, nine indicated 0 to 5 years of previous experience evaluating juries, six 6 to 10 years, and another six 11 years or more. Three did not respond to this question.

Before the jury performances, I met with each area's faculty, distributed the evaluation instruments described, and briefly discussed them. I asked the evaluators not to put their names or the names of the students on the forms. I requested instructors to read each item carefully and to respond to each scale, and I asked them not to discuss the jury performances with one another while completing the forms.

Each evaluation form took about 1 minute to complete. At the conclusion of the evaluations, I collected the completed forms from each area coordinator. I had requested that each area complete 16 evaluations, but I was able to use only 12 to 15 depending on the area (see Table 1). On occasion, one or more of the evaluators had left the juries before completing all 16 forms or had left and returned.

## RESULTS

Table 1 shows all interjudge reliability results (Coefficients of Concordance, *Ws*). To determine interjudge reliability for panels of fewer than the full complement of judges, I calculated coefficients for all possible permutations of two to four as appropriate, with averaged coefficients, standard deviations, ranges, and lowest and highest coefficients reported in Table 1.

Full-panel interjudge reliability was consistently good regardless of panel size, which ranged from an  $n$  of 2 for Percussion to an  $n$  of 5 for Woodwind, Voice, and Strings. All subscale interjudge reliabilities for all groups except Percussion were statistically significant, with the exception of the Suitability subscale in Voice. (Owing to its  $n$  of 2, Percussion had an exceptionally stringent statistical significance criterion.) All rating scale total score interjudge reliability coefficients were statistically significant, as were all for the global letter grade assessment. For full-group panels with  $ns$  of 4 and 5, rating scale total-score reliability was consistently but not greatly higher than reliability for the letter grade assessment. Average standard deviations for permutations of four were low and relatively uniform, whereas averaged standard deviations for permutations of two and three tended to be higher and less uniform.

As Table 1 shows, there was no loss of average reliability as group size incrementally decreased. With this incremental decrease, however, average variability and range from lowest to highest coefficients showed an incremental increase.

I examined differences among levels of experience by looking for a consistent pattern of lower (or higher) coefficients among permutations containing the less-experienced evaluators. I found no such pattern. Because only two of the evaluators were graduate teaching assistants and



both were members of the voice group, I compared teaching assistant evaluation to faculty evaluation within the voice area and found no discernible pattern there as well.

## DISCUSSION

In contrast with Payne's (1997) assertion that "establishing a strong reliability correlation [in performance adjudication] may not be possible," interjudge reliability in this study was stable and consistent, even among most pairwise comparisons of evaluators. Interjudge reliability was good for rating scale total scores, subscales, and the global letter grade assessment alike. Somewhat in contrast with Fiske (1975, 1977), subscale score reliabilities compared favorably with the global grade letter assessment, especially among the larger panels. Amount of prior experience had no apparent effect on reliability.

When asked during postevaluation discussions what may have led to such consistency, especially given the wide range of experience and lack of training among evaluators, virtually all area coordinators mentioned that their colleagues treat one another with collegiality and respect, and that they seem to share a commitment to student success. Veteran faculty informally help younger or less-experienced faculty members learn expectations and routines for evaluation. Such helpfulness and collegiality also was noted among the most reliable faculty evaluators in an earlier study (Bergee, 1997). Therefore, more experienced faculty might consider periodic efforts to share values and expectations for student performance with newer faculty. Perhaps veteran faculty could sponsor informal sessions acquainting newer faculty with these values and expectations (and reaffirming them among themselves). Among other salutary effects, such sessions may help to build a sense of collegiality among newer faculty.

Evaluators in this study found the criteria-specific rating scales comfortable to use. There were no negative postparticipation comments about the process, even among faculty who had expressed some initial reluctance. Evaluators readily familiarized themselves with the forms, and evaluations proceeded smoothly. Appreciation was expressed for the format. Given the compressed schedule, circling response categories seemed to effectively balance effort with information.

Some evaluators commented that the forms "got them thinking" about performance aspects they might not have considered solely through the writing of anecdotal comments. An assessment process involving only the writing of brief, unstructured comments does not allow evaluators an opportunity to deliberate on all performance aspects. Because of time constraints, evaluators may tend to dwell on weaknesses. Criteria-specific rating scales, on the other hand, are more comprehensive, encouraging attention to all aspects of the performance and providing balanced feedback to performers.

Not surprising, total score interjudge reliabilities in this study were somewhat lower on average than the exceedingly high total score reliabilities reported in earlier studies (Abeles, 1971; Bergee, 1988a; Jones, 1986; Nichols, 1991; Zdzinski & Barnes, 2002). With the exception of percussion, all full-group coefficients in this study were between .70 and .80. A "real-world" performance assessment along with the researcher's modifications of the scales and choice of analysis may have influenced this. Use of the nonparametric *W* may have driven down coefficients; the original studies frequently used Hoyt's analysis of variance technique. Limiting subscale items also may have had an effect. Doubling the number of items via Spearman-Brown brought reliability coefficients roughly to levels reported in the original studies. Therefore, in future research, more subscale items should be used, perhaps by using the scales in their original forms. Also, the original random assignment of positive and negative wording should be restored. A response set – in this case, a tendency to use the same circle to mark a response category for several adjacent items – was clearly evident among some evaluators. Many faculty, however, would have expressed strong reservations at completing a lengthy scale for each performer under such strict scheduling constraints. Evaluators might consider using criterion-specific rating scales in lieu of the more common regimen of writing comments, because rating scale total score reliabilities in this study tended to be higher than letter-grade reliabilities, especially among the larger panels.



Table 1

Interjudge Reliability (Coefficient of Concordance, *W*) among Brass, Percussion, Woodwind, Voice, Piano, and String Faculty Groups

| Brass ( <sup>n</sup> fac = 4, <sup>n</sup> perf = 15)      |                                           |              |                       |           |             |         |         |         |
|------------------------------------------------------------|-------------------------------------------|--------------|-----------------------|-----------|-------------|---------|---------|---------|
|                                                            | Interp/ME                                 | Tone/Inton   | Technique             | Rhy/Tempo | Total Score | Grade   |         |         |
|                                                            | .64**                                     | .58**        | <i>n</i> = 4<br>.67** | .44**     | .72**       | .68**   |         |         |
|                                                            | ns = 3 (averaged across all permutations) |              |                       |           |             |         |         |         |
| <i>M</i>                                                   | .68                                       | .63          | .71                   | .50       | .75         | .72     |         |         |
| ( <i>SD</i> )                                              | (.05)                                     | (.04)        | (.05)                 | (.07)     | (.03)       | (.07)   |         |         |
| Rg                                                         | .11                                       | .09          | .10                   | .17       | .05         | .12     |         |         |
| LoHi                                                       | .64-.75                                   | .59-.68      | .68-.78               | .42-.59   | .74-.79     | .70-.82 |         |         |
|                                                            | ns = 2 (averaged)                         |              |                       |           |             |         |         |         |
| <i>M</i>                                                   | .76                                       | .72          | .78                   | .62       | .82         | .79     |         |         |
| ( <i>SD</i> )                                              | (.07)                                     | (.07)        | (.06)                 | (.11)     | (.04)       | (.09)   |         |         |
| Rg                                                         | .21                                       | .20          | .13                   | .27       | .12         | .22     |         |         |
| LoHi                                                       | .66-.87                                   | .60-.80      | .71-.84               | .50-.77   | .76-.88     | .66-.89 |         |         |
| Percussion ( <sup>n</sup> fac = 2, <sup>n</sup> perf = 12) |                                           |              |                       |           |             |         |         |         |
|                                                            | Interpretation                            | Tone Quality | Tech/Rhythm           |           | Total       | Grade   |         |         |
|                                                            | .90                                       | .69          | .90                   |           | .93*        | .96*    |         |         |
| Woodwind ( <sup>n</sup> fac = 5, <sup>n</sup> perf = 12)   |                                           |              |                       |           |             |         |         |         |
|                                                            | Interp                                    | Tone         | Rh/Cn                 | Inton     | Tempo       | Artic   | Total   | Grade   |
|                                                            | .67**                                     | .61**        | .67**                 | .58**     | .69**       | .51**   | .79**   | .65**   |
|                                                            | <sup>n</sup> s = 4 (averaged)             |              |                       |           |             |         |         |         |
| <i>M</i>                                                   | .69                                       | .64          | .69                   | .61       | .71         | .54     | .80     | .67     |
| ( <i>SD</i> )                                              | (.02)                                     | (.05)        | (.05)                 | (.05)     | (.05)       | (.06)   | (.02)   | (.06)   |
| Rg                                                         | .04                                       | .12          | .11                   | .13       | .11         | .16     | .05     | .15     |
| LoHi                                                       | .67-.71                                   | .59-.71      | .63-.74               | .55-.68   | .67-.78     | .47-.63 | .78-.63 | .63-.78 |
|                                                            | <sup>n</sup> s = 3 (averaged)             |              |                       |           |             |         |         |         |
| <i>M</i>                                                   | .73                                       | .68          | .73                   | .65       | .74         | .60     | .83     | .71     |
| ( <i>SD</i> )                                              | (.03)                                     | (.07)        | (.07)                 | (.08)     | (.07)       | (.09)   | (.04)   | (.09)   |
| Rg                                                         | .10                                       | .20          | .20                   | .23       | .18         | .29     | .12     | .23     |
| LoHi                                                       | .68-.78                                   | .57-.77      | .61-.81               | .53-.76   | .66-.84     | .46-.75 | .76-.88 | .59-.82 |

(Table 1 continues on next page)

Table 1 (continued)

| Woodwind ( <sup>n</sup> fac = 5, <sup>n</sup> perf = 12) |           |           |           |         |             |         |         |         |
|----------------------------------------------------------|-----------|-----------|-----------|---------|-------------|---------|---------|---------|
|                                                          | Interp    | Tone      | Rh/Cn     | Inton   | Tempo       | Artic   | Total   | Grade   |
| ns = 2 (averaged)                                        |           |           |           |         |             |         |         |         |
| M                                                        | .79       | .76       | .80       | .74     | .81         | .70     | .87     | .78     |
| (SD)                                                     | (.05)     | (.10)     | (.09)     | (.11)   | (.09)       | (.11)   | (.05)   | (.11)   |
| Rg                                                       | .17       | .31       | .24       | .38     | .28         | .39     | .14     | .31     |
| LoHi                                                     | .69-.86   | .56-.87   | .65-.89   | .50-.78 | .65-.93     | .51-.90 | .78-.92 | .57-.88 |
| Voice ( <sup>n</sup> fac = 5, <sup>n</sup> perf = 12)    |           |           |           |         |             |         |         |         |
|                                                          | Intp/ME   | Tn/ME     | Tech      | Suit    | Dict        | Total   | Grade   |         |
| n = 5                                                    |           |           |           |         |             |         |         |         |
|                                                          | .78**     | .59**     | .62**     | .38     | .60**       | .70**   | .58**   |         |
| ns = 4 (averaged)                                        |           |           |           |         |             |         |         |         |
| M                                                        | .79       | .61       | .64       | .42     | .62         | .72     | .60     |         |
| (SD)                                                     | (.05)     | (.05)     | (.06)     | (.09)   | (.06)       | (.06)   | (0.09)  |         |
| Rg                                                       | .13       | .12       | .13       | .20     | .16         | .16     | .20     |         |
| LoHi                                                     | .74-.87   | .58-.68   | .59-.72   | .34-.54 | .54-.70     | .66-.82 | .53-.73 |         |
| ns = 3 (averaged)                                        |           |           |           |         |             |         |         |         |
| M                                                        | .81       | .66       | .68       | .48     | .66         | .74     | .65     |         |
| (SD)                                                     | (.07)     | (.07)     | (.09)     | (.12)   | (.09)       | (.09)   | (0.12)  |         |
| Rg                                                       | .21       | .23       | .30       | .38     | .31         | .23     | .42     |         |
| LoHi                                                     | .71-.92   | .57-.80   | .56-.86   | .33-.71 | .56-.87     | .63-.86 | .48-.90 |         |
| ns = 2 (averaged)                                        |           |           |           |         |             |         |         |         |
| M                                                        | .86       | .74       | .76       | .61     | .75         | .81     | .73     |         |
| (SD)                                                     | (.08)     | (.11)     | (.11)     | (.15)   | (.13)       | (.10)   | (0.14)  |         |
| Rg                                                       | .23       | .35       | .31       | .47     | .37         | .29     | .38     |         |
| LoHi                                                     | .73-.96   | .60-.95   | .60-.91   | .42-.89 | .57-.94     | .64-.93 | .56-.94 |         |
| Piano ( <sup>n</sup> fac = 3, <sup>n</sup> perf = 14)    |           |           |           |         |             |         |         |         |
|                                                          | Interp/ME | Rhy/Tempo | Technique |         | Total Score |         | Grade   |         |
| n = 3                                                    |           |           |           |         |             |         |         |         |
|                                                          | .76**     | .67**     | .87**     |         | .76**       |         | .90**   |         |
| ns = 2 (averaged)                                        |           |           |           |         |             |         |         |         |
| M                                                        | .82       | .75       | .90       |         | .82         |         | .92     |         |
| (SD)                                                     | (.06)     | (.10)     | (0.05)    |         | (0.08)      |         | (0.03)  |         |
| Rg                                                       | .11       | .18       | .08       |         | .15         |         | .05     |         |
| LoHi                                                     | .77-.88   | .68-.86   | .87-.95   |         | .74-.89     |         | .90-.95 |         |



Table 1 (concluded)

| Strings ( $n_{\text{fac}} = 5$ , $n_{\text{perf}} = 16$ ) |         |          |         |                      |         |         |
|-----------------------------------------------------------|---------|----------|---------|----------------------|---------|---------|
|                                                           | Interp  | Art/Tone | Inton   | Rhy/Temp             | Vibrato | Total   |
|                                                           |         |          |         |                      |         | Grade   |
|                                                           | .58**   | .57**    | .70**   | $n = 5$<br>.50**     | .67**   | .71**   |
|                                                           |         |          |         |                      |         | .69**   |
|                                                           |         |          |         | $n_s = 4$ (averaged) |         |         |
| <i>M</i>                                                  | .61     | .59      | .72     | .53                  | .69     | .73     |
| ( <i>SD</i> )                                             | (.07)   | (.02)    | (.04)   | (.04)                | (.04)   | (.04)   |
| <i>Rg</i>                                                 | .18     | .05      | .11     | .10                  | .09     | .11     |
| <i>LoHi</i>                                               | .52-.70 | .58-.63  | .68-.79 | .48-.58              | .66-.75 | .68-.79 |
|                                                           |         |          |         | $n_s = 3$ (averaged) |         |         |
| <i>M</i>                                                  | .65     | .64      | .75     | .60                  | .73     | .76     |
| ( <i>SD</i> )                                             | (.09)   | (.03)    | (.06)   | (.06)                | (.05)   | (.06)   |
| <i>Rg</i>                                                 | .26     | .11      | .19     | .16                  | .19     | .16     |
| <i>LoHi</i>                                               | .50-.76 | .59-.70  | .66-.85 | .48-.64              | .64-.83 | .67-.83 |
|                                                           |         |          |         | $n_s = 2$ (averaged) |         |         |
| <i>M</i>                                                  | .74     | .73      | .81     | .69                  | .80     | .82     |
| ( <i>SD</i> )                                             | (.11)   | (.05)    | (.07)   | (.08)                | (.07)   | (.08)   |
| <i>Rg</i>                                                 | .31     | .19      | .22     | .23                  | .22     | .18     |
| <i>LoHi</i>                                               | .54-.85 | .63-.82  | .68-.90 | .56-.79              | .66-.88 | .70-.88 |

Note. Refer to text for full names of subscale. *M* = mean, *SD* = standard deviation, *Rg* = range, *LoHi* = lowest and highest coefficients among the permutations. (\*)  $p < .05$ . (\*\*)  $p < .01$ .

In agreement with Fiske (1975, 1977), this study's findings point to an improvement in stability as adjudication groups increase in size. Variability and range consistently decreased as panel size increased. This outcome suggests that the frequent use in juries of small panels of two or three evaluators should be reconsidered: subscale and total score reliabilities exhibited by some permutations were below minimally acceptable levels. At some point, stability inevitably levels off, perhaps at Fiske's (1975, 1983) recommended minimum of seven. If evaluation groups indeed should be seven or greater (a determination further study should make), additional faculty may need to be recruited. According to Fiske (1975), raters need not be specialists in the area. In this study, for example, one of the percussion evaluators was a brass faculty member. Those agreeing to evaluate, however, should commit to remain for the duration of the juries. Occasional absences leave some students with smaller evaluation panels and their evaluation scores with a correspondingly higher risk of unreliability.

Taking the conditions surrounding jury evaluation into consideration, interjudge reliability in this study generally was good. Faculty, students, and researchers may be allowed some confidence in the consistency of this form of evaluation, although such an assertion must be considered provisional until corroborating data are obtained. In support of Mills's (1987) earlier observation about the scarcity of research in this area, further effort should be directed toward refining performance structures, developing valid assessments in all domains, and testing the efficacy of these assessment tools in performance contexts.

## REFERENCES

- Abeles, H. F. (1973). Development and validation of a clarinet performance adjudication rating scale. *Journal of Research in Music Education*, 21, 246-255.
- Bergee, M.J. (1988a). An application of the facet-factorial approach to scale construction in the development of a rating scale for euphonium and tuba music performance (Doctoral dissertation, University of Kansas, 1987). *Dissertation Abstracts International*, 49, 1086A.
- Bergee, M.J. (1988b). Use of an objectively constructed rating scale for the evaluation of brass juries: A criterion-related study. *Missouri Journal of Research in Music Education*, 5 (5), 6-15.
- Bergee, M.J. (1989). An investigation into the efficacy of using an objectively constructed rating scale for the evaluation of university-level single-reed juries. *Missouri Journal of Research in Music Education*, 26, 74-91.
- Bergee, M.J. (1993). A comparison of faculty, peer, and self-evaluation of applied brass jury performances. *Journal of Research in Music Education*, 41, 19-27.
- Bergee, M.J. (1997). Relationships among faculty, peer, and self-evaluation of applied jury performances. *Journal of Research in Music Education*, 45, 601-612.
- Britten, R. L. (2002). Instrumentalists' assessment of solo performance with compact disc, piano, or no accompaniment. *Journal of Research in Music Education*, 50, 63-74.
- Butt, D. S., & Fiske, D. W. (1968). Comparison of strategies in developing scales for dominance. *Psychological Bulletin*, 70, 505-519.
- Dixon, S. (2000, November). *Assessing the performer*. Paper presented at the Palatine Assessment Workshop, Lancaster, England.
- Fiske, H. E. (1975). Judge-group differences in the rating of secondary school trumpet performances. *Journal of Research in Music Education*, 23, 186-196.
- Fiske, H. E. (1977). Relationship of selected factors in trumpet performance adjudication reliability. *Journal of Research in Music Education*, 25, 256-263.
- Fiske, H. E. (1983). Judging musical performances: Method or madness? *Update: Applications of Research in Music Education*, 1 (3), 7-10.
- Horowitz, R. A. (1994). The development of a rating scale for jazz guitar improvisation performance (Doctoral dissertation, Columbia University Teachers College). *Dissertation Abstracts International*, 55, 3443A.
- Interstate New Teacher Assessment and Support Consortium. (2002). *Model standards for licensing classroom teachers and specialists in the arts: A resource for state dialogue*. Washington, DC: Council of Chief State School Officers.
- Jones, H. (1986). An application of the facet-factorial approach to scale construction in the development of a rating scale for high school solo vocal performance (Doctoral dissertation, University of Oklahoma). *Dissertation Abstracts International*, 47, 1230A.
- Mills, J. (1987). Assessing solo musical performance – A preliminary study. *Bulletin of the Council for Research in Music Education*, no. 91, 119-125.



- Music Educators National Conference. (1996). *Performance standards for music*. Reston, VA: Author.
- National Association of Schools of Music. (1999). *1999-2000 handbook*. Reston, VA: Author.
- Nichols, J. P. (1991). A factor analysis approach to the development of a rating scale for snare drum performance. *Dialogue in Instrumental Music Education*, 15, 11-31.
- Payne, B. (1997). A review of research on band competition. *Journal of Band Research*, 33, 1-19.
- Radocy, R. E. (1986). On quantifying the uncountable in musical behavior. *Bulletin of the Council for Research in Music Education*, no. 88, 22-31.
- Saunders, T. C., & Holohan, J. M. (1997). Criteria-specific rating scales in the evaluation of high school instrumental performance. *Journal of Research in Music Education*, 45, 259-272.
- Stanley, M., Brooker, R., & Gilbert, R. (2002). Examiner perceptions of using criteria in music performance assessment. *Research Studies in Music Education*, 18, 43-52.
- Stewart, J. L. (2002, April). *Performance evaluation in music research*. Paper presented at the National Biennial In-Service Conference of MENC: The National Association for Music Education, Nashville, TN.
- Wapnick, J., Flowers, P., Alegant, M., & Jasinskis, L. (1993). Consistency in piano performance evaluation. *Journal of Research in Music Education*, 41, 282-292.
- Zdzinski, S. F. (1991). Measurement of solo instrumental music performance: A review of literature. *Bulletin of the Council for Research in Music Education*, no. 109, 47-58.
- Zdzinski, S. F., & Barnes, G. V. (2002). Development and validation of a string performance rating scale. *Journal of Research in Music Education*, 50, 245-255.

---

\* As referências bibliográficas deste artigo estão conforme o documento original.

## 2.8.3. Estudo Crítico – Por Christopher Johnson

### Objectivo

O objectivo desta investigação foi determinar, *in situ*<sup>4</sup>, a fiabilidade interjuízes numa situação de avaliação, constituída por diferentes grupos de examinadores de provas de performance musical, aquando das provas semestrais numa grande Universidade. Este foi um esforço para aprofundar investigações anteriores sobre o rigor na avaliação, considerando o número de juízes que devem constituir um júri, as suas diferenças de estilo e o seu nível de experiência.

### Design

Relativamente ao design encontramos-nos perante um estudo descritivo. O investigador adoptou cinco tabelas de apreciação previamente utilizadas e experimentadas e criou uma nova especificamente para esta investigação. Cada grupo de examinadores aplicou estas escalas de classificação de forma independente. Os júris não trocaram informação entre si, nem com os alunos que se apresentaram para avaliação. Os dados usados na análise estatística são os resultados e comentários obtidos através das grelhas de avaliação utilizadas.

### Revisão do Artigo

### Introdução e revisão

Bergee começa por nos esclarecer, com argumentação consistente, a importância da performance musical na educação musical. Esta opinião é sustentada pela ideia de que a performance musical é parte essencial dos padrões de educação musical estabelecidos na *Music Educators National Conference Standards* (1996)<sup>5</sup>.

O autor estabelece uma concordância lógica entre a existência de um determinado padrão de performance e a necessidade de avaliar com fiabilidade de modo a determinar se esse padrão foi efectivamente atingido. Em seguida remete-nos para a ideia de que, apesar de a capacidade de avaliar individualmente uma performance musical fazer parte da condição humana, poucos são aqueles que conseguem traduzir racionalmente o que os levou a determinada apreciação. Assim sendo, o autor considera que este tipo de avaliação não deverá ser considerada válida ou fiável, estabelecendo assim a necessidade da realização do presente estudo.

Foram também passadas em revista as investigações feitas até à data, dando conta da lacuna bibliográfica existente, nomeadamente no que diz respeito a investigações com metodologia descritiva e *in situ*, a qual alicerça este estudo. Finalmente, Bergee expõe e analisa as variáveis externas que acredita deverem ser analisadas em investigações relativas ao estudo da fiabilidade interjuízes em situações de avaliação da performance musical no seu contexto natural. Estas variáveis são: 1) a variabilidade de tamanho dos grupos de examinadores; 2) o instrumento de avaliação padronizado a utilizar ao longo do processo de avaliação; 3) a experiência do avaliador.

### Método

Bergee separa a metodologia em duas partes distintas. A primeira descreve a investigação prévia sobre os instrumentos de avaliação e subsequente adaptação às necessidades desta investigação. Este processo foi aplicado aos instrumentos de avaliação utilizados para os Metais,



Madeiras, Percussão, Voz e Cordas. Não havendo instrumento de avaliação para Piano, foi criado um, utilizando os outros como modelo. Esta primeira secção sobre metodologia termina com a explicação de algumas alterações e ajustamentos feitos nas escalas de classificação, justificando também a opção pela escala de classificação baseada numa ordem alfabética.

Na segunda parte, Bergee retrata os participantes e os procedimentos a que foram sujeitos, descrevendo com algum detalhe o facto de todos eles terem concordado em colaborar. No entanto, mais à frente, neste mesmo capítulo, o autor afirma que alguns dos participantes demonstraram alguma relutância em aceitar as condições impostas e em cooperarem com o estudo, uma vez que o que aqui é avaliado faz parte das suas tarefas profissionais.

Em seguida, o autor expõe todas as instruções relativas aos procedimentos. Estranhamente optou por não fazer qualquer referência aos alunos que foram avaliados. Tal poderá ficar a dever-se ao facto de estes serem apenas o estímulo e não os objectos do estudo. Porém, uma breve exposição deste assunto talvez tivesse sido útil à análise posterior dos resultados.

## Resultados

Os resultados deste artigo são bastante concisos. A apresentação geral é estabelecida numa única tabela bastante extensa, que constitui um bom resumo do estudo. Bergee analisa os dados através dos Coeficientes de Concordância<sup>6</sup> – teste estatístico adequado ao tratamento de dados quando organizados de forma ordinal. Embora também fosse ajustado fazer o cálculo da fiabilidade baseado numa medida intervalar (como é o caso do *Alpha de Crombach*, usado habitualmente), sem dúvida que a opção de Bergee foi a mais adequada.

Nos parágrafos seguintes são-nos apresentados os resultados mais revelantes. Estes resultados indicam que, na generalidade dos casos, a fiabilidade interjuízes foi bastante elevada, como se pode verificar através dos coeficientes de fiabilidade das escalas de classificação. O autor não faz uma descrição muito detalhada dos seus próprios resultados, já expostos na supracitada tabela, mas aponta para aspectos únicos em torno dos mesmos, que escapariam a um leitor menos avisado. Também salienta aspectos relacionados com a supracitada terceira variável externa: a experiência do avaliador. Embora não apresente quaisquer valores em concreto, Bergee indica que não foram encontradas diferenças que apontem para a influência deste factor.

## Discussão

Nesta secção, Bergee compara os seus resultados com os encontrados em investigações precedentes, com o mesmo objecto de estudo. São expostas algumas razões para as diferenças encontradas. A impressão geral sobre o alto grau de fiabilidade encontrada na avaliação interjuízes, segundo o autor, advém do facto de os participantes obedecerem a uma conduta ética colegial de respeito mútuo. Este aspecto condicionou o processo de avaliação das performances e por certo exerceu um peso considerável no nível de uniformidade alcançada.

A subsequente revisão dos procedimentos da experiência resultou na conclusão de que tudo tinha corrido de acordo com o planeado, obtendo-se, consequentemente, um bom projecto. Este estudo estimulou também o interesse dos sujeitos pelas grelhas de avaliação utilizadas, por permitirem uma avaliação válida e rigorosa das performances dos alunos.

Genericamente, os coeficientes obtidos por Bergee foram numericamente mais baixos que os encontrados em trabalhos anteriores. No entanto, este é um item enganador, já que não tem em conta dois aspectos críticos deste estudo: primeiro porque este se realizou *in situ*; segundo, porque os estudos anteriores que apresentaram coeficientes de fiabilidade mais elevados, utilizaram diferentes procedimentos de análise estatística. Nesses casos a utilização de testes paramétricos deu lugar a um incremento artificial dos resultados. Por isso há que relativizar o facto de os coeficientes de fiabilidade encontrados no presente estudo serem mais baixos, visto que



estes testes são mais rigorosos e resultam da aplicação de procedimentos analíticos decorrentes da aplicação de testes não paramétricos.

As recomendações finais feitas pelo autor vão ao encontro das proferidas por Fiske (1975, 1977), no sentido de se utilizar o maior número possível de juízes, para garantir uma apreciação consistente e correcta da performance musical. Surpreendentemente, é defendido que a experiência dos juízes não é tão importante quanto o seu número. Bergee termina o artigo com um lugar comum, sugerindo que é necessário que se faça mais pesquisa nesta área.

### **Contribuição para a área de estudo**

A principal contribuição deste estudo é a demonstração clara de que a performance musical pode ser avaliada de forma fiável em contextos reais desde que seja fornecido um bom instrumento de avaliação e um grupo de avaliadores devidamente preparado. Esta observação aplica-se às avaliações globais e às subescalas. Este é, precisamente, um dos aspectos relevantes deste estudo. De facto, se de um modo geral a investigação tem demonstrado que o único resultado fiável é o resultado global, neste caso específico as subescalas do instrumento de avaliação também se revelaram fiáveis.

O segundo contributo importante desta investigação é o comentário positivo dos avaliadores quanto à utilização de um método de apreciação mais formal. Em muitas situações em que a performance musical é avaliada, os examinadores simplesmente contribuem com breves comentários pouco fundamentados e objectivos, procedendo a uma classificação ordinal ou alfabética, sem relação com qualquer modelo relativo aos aspectos específicos a considerar. As grelhas de avaliação usadas neste estudo são de fácil utilização, de tal forma que os elementos do júri as acolheram de bom agrado e, em vez de ajuizarem apenas a primeira coisa que ouviram, cuidaram de considerar todas as vertentes da performance que lhes eram propostas. Assim sendo, a utilização destas grelhas seria benéfica para uma avaliação mais correcta e válida dos alunos submetidos aos Standards Nacionais para a Educação Musical estabelecidos pela organização americana *Music Educators National Conference* (MENC).

### **Comentário**

Há muitos aspectos neste artigo que fazem dele uma peça de investigação importante. A Introdução e a Revisão Literária, embora curtas, são completas. Aqui estão referenciados vinte e sete artigos, o que demonstra que Bergee é um profundo conhecedor da literatura existente, usando cuidadosamente esse capital na construção do seu projecto. Não se perde em análise de ideias supérfluas para convencer pelo volume de informação, construindo a estrutura deste trabalho de forma lógica, sintética e sustentada.

A metodologia utilizada pressupõe opções que deverão ser escrutinadas. Em primeiro lugar, Bergee utilizou escalas de avaliação apropriadas às várias secções de instrumentos, usando, portanto, diferentes medidas durante o processo. Por um lado, isto traduz-se numa sensação de solidez científica, uma vez que cada instrumento é submetido a um tratamento de medição específico de acordo com a família a que pertence, tal e qual como aconteceu durante a experiência. Porém, não teria sido possível adaptar uma das escalas de avaliação de forma a ser possível utilizar a mesma medida entre as várias famílias de instrumentos? Em termos práticos faria todo o sentido uniformizar o tipo de grelha usada para avaliar as várias famílias de instrumentos.

De facto, poder-se-á argumentar que quando não existia um instrumento de avaliação específico (i.e., o da área de Piano), Bergee e o grupo de Professores colaboradores criaram um, baseado nos preexistentes. Além disso, nenhum dos instrumentos de avaliação preexistentes foi utilizado na totalidade porque, para se tornar exequível, foi necessário cortar diversos itens. Por outro lado, o objectivo de Bergee era perceber se a fiabilidade da avaliação interjuízes poderia ser alcançada através da utilização de utensílios cientificamente comprovados, não estando



propriamente empenhado em criar um conjunto de instrumentos de avaliação. Portanto, a sua posição é sustentável, ainda que outras ideias possam igualmente ser consideradas.

Por razões de ordem prática, compreende-se que Bergee só tenha utilizado professores da sua universidade. Conseguir o contributo de outros seria muito mais difícil, se não impossível. Contudo, questiona-se aqui se professores de outras instituições seriam tão fiáveis quanto os que participaram no estudo. Ao analisarmos a argumentação do autor para o elevado grau de fiabilidade, defendendo que esta se deve ao facto de existir um respeito ético e colegial entre os participantes, podemos fazer-nos a seguinte questão: Será que o grau de fiabilidade das avaliações se manteria elevado caso o respeito entre participantes fosse menor? Será que não há outra explicação para os resultados?

Por outro lado, para o leitor teria sido útil saber mais acerca do processo de escolha aleatória dos estudantes utilizados como estímulo. Foram escolhidos aleatoriamente de entre todos os alunos da instituição? Eram alunos que já tinham sido avaliados em exames anteriores? Eram todos estudantes do primeiro ano ou eram alunos avançados em final de licenciatura? Terão sido aleatoriamente escolhidos de entre anos diferentes? Embora estes dados não tivessem tido, com certeza, influência sobre os elevados coeficientes de fiabilidade encontrados, o acesso a esta informação teria sido oportuno. Em todo o caso, presume-se que foi escolhida uma amostra aleatória dos alunos que se encontravam a ter aulas, aquando da realização do estudo.

A análise dos resultados é magnificamente realizada. Bergee é um investigador que domina a estatística com sabedoria, analisando os seus dados de forma irrepreensível. As informações recolhidas a partir da Escala de Likert, não são intervalares, mas sim ordinais. Por consequência, é necessária a utilização de testes não paramétricos. Esta é uma regra frequentemente violada. Esta opção em relação à análise estatística eleva o estudo de Bergee. A sua compilação de diferentes permutas e combinações de juízes é também inovadora e elucidativa. Mais uma vez, estamos perante um excelente tratamento estatístico de um complexo agregado de dados.

Bergee confronta os seus resultados com os de outras investigações anteriores e dá-nos um entendimento claro da forma como estas estão relacionadas com o seu estudo. Não se apoia nos seus resultados para se evidenciar de outras investigações. Faz sugestões sobre o caminho a seguir e oferece um panorama do que sabemos até ao momento sobre o estudo da fiabilidade interjuízes e a avaliação da performance musical em geral.

Já que o autor definiu as grelhas de avaliação e os colegas demonstraram um grande entusiasmo em relação à aplicação destas novas ferramentas, seria interessante saber se, naquela instituição, as continuam a utilizar de forma regular. Isto poderia lançar luz sobre o nível de apoio que o investigador realmente teve no que diz respeito à utilização destes instrumentos de avaliação. Pelo facto de existir uma forma melhor e mais eficaz de fazer as coisas, não significa que as pessoas alterem a sua forma habitual de trabalhar. Por certo este ainda será um longo caminho a percorrer.

Estamos perante um excelente artigo de investigação científica. O autor escolheu criteriosamente as palavras a utilizar na exposição de cada ideia. À medida que os pensamentos são expressos, o leitor fica com a ideia de que o estudo está a ser realizado com idêntico rigor. Isto poderá explicar a razão pela qual este projecto teve tanto sucesso. Finalmente, o aspecto mais importante a sublinhar neste artigo é, possivelmente, o seu contributo para a prática docente.

---

N.T.

- 1 – O termo Júri designa um grupo constituído por juízes ou jurados. Ou seja, os termos *juiz* ou *jurado* aplicam-se a cada membro constituinte de um Júri.
- 2 – De acordo com o artigo (p.142) esta escala foi estabelecida de A+ (excelente performance em todos os aspectos) até F (performance extremamente pobre em todos os aspectos).
- 3 – Subescalas – Referência a aspectos específicos da escala de classificação. Ou seja, a avaliação da performance do instrumento é efectuada de acordo com uma escala de classificação constituída pela avaliação de aspectos específicos (subescalas). À escala correspondem os resultados totais e à subescala os resultados parciais.
- 4 – *In situ* – do latim – no lugar, em contexto real.
- 5 – Para informação mais detalhada sobre os standards para educação musical estabelecidos nesta conferência, consultar o site oficial da *The National Association for Music Education* – MENC, em [www.menc.org](http://www.menc.org).
- 6 – A concordância de coeficientes (*Ws*) é a medida estatística para o cálculo de fiabilidade para dados ordinais, tal como o *Alpha de Crombach* é usado para o cálculo de fiabilidade de dados intervalares. O teste da concordância de coeficientes é do tipo não paramétrico. Este tipo de teste estatístico não necessita de distribuição normal e variância igual entre os grupos, por oposição aos testes paramétricos, que assentam em medidas intervalares com distribuição normal, variância homogênea e, intervalos contínuos e iguais. Os testes não-paramétricos aplicam-se normalmente a amostras baseadas em medidas ordinais e nominais, sendo muito úteis para cálculos rápidos, com amostras ou muito grandes ou muito pequenas.

## Bibliografia

Fiske, H. E.(1975). Judge-group differences in the rating of secondary school trompe performance, *Journal of Research in Music Education*, 23, 186-196.

Fiske, H. E.(1977). Relationship of selected factors in trumpet performance adjudication reliability. *Journal of Research in Music Education*, 25, 256-263.

Music Educators National Conference (1996) *Performance standards for music*. Reston, VA: Author.





# 2.9.

## THE EFFECT OF VISION AND HEARING LOSS ON LISTENERS PERCEPTION OF REFERENTIAL MEANING IN MUSIC

O EFEITO DA PERDA DE VISÃO E AUDIÇÃO DOS OUVINTES NA  
PERCEPÇÃO DE SIGNIFICADOS REFERENCIAIS EM MÚSICA

A.-A. DARROW & J. NOVAK (2005)

COMUNICAÇÃO APRESENTADA NO 16<sup>o</sup> INTERNATIONAL SYMPOSIUM ON RESEARCH IN MUSIC BEHAVIOR, ST.  
LOUIS, MO

### 2.9.1. Resumo do artigo / Abstract

#### Resumo

O objectivo deste estudo foi examinar o efeito da perda de visão ou audição dos ouvintes na percepção de significados referenciais em música.

Os participantes eram estudantes de uma escola estatal para surdos e cegos e estudantes com visão e audição normais que frequentavam escolas vizinhas ( $N = 96$ ). Os estímulos musicais eram seis excertos extraídos da obra *O Carnaval dos Animais* de Saint Saens e foram ordenados aleatoriamente. Cada estímulo tinha a duração de 37 segundos. Para a escolha dos excertos foi tido em consideração o seu uso em estudos similares e também a inequívoca indicação do compositor quanto aos significados extra-musicais da obra.

Depois de se proceder às adaptações necessárias para efectuar o estudo com participantes com perda de visão ou audição, foi pedido a todos os participantes que seleccionassem uma imagem suscitada por cada um dos fragmentos musicais.

A pergunta da investigação era a seguinte: "Os estudantes com perda de visão ou audição associam as mesmas imagens visuais à música que os estudantes sem perdas sensoriais?" Os dados, analisados segundo o procedimento ANOVA, procuraram examinar os efeitos da condição sensorial, bem como os factores idade e sexo.

Verificou-se que a condição sensorial tinha um efeito significativo nos resultados. A análise dos resultados mostrou que as imagens seleccionadas por participantes sem perdas de audição ou de visão coincidiam significativamente mais com as imagens sugeridas pelo compositor do que as seleccionadas pelos outros participantes. Não foram encontrados efeitos significativos nem da idade nem do sexo. Não foram encontradas interações significativas, também.

Dados sumários indicaram que as imagens seleccionadas eram mais ou menos fáceis de identificar, consoante as condições sensoriais dos participantes. Os dados revelaram também graus de dificuldade e padrões de confusão similares nas diversas idades e condições sensoriais, indicando que as respostas dos participantes não eram aleatórias e que alguns significados referenciais em música são convencionais.

#### Abstract

The purpose of the present study was to examine the effect of vision and hearing loss on listeners perception of referential meaning in music.



Participants were students at a state school for the deaf and blind, and students with typical hearing and vision who attended neighboring public schools (N = 96). The music stimuli consisted of six 37-second randomly ordered excerpts from Saint Saëns, *Carnival of the Animals*. The excerpts were chosen because of their use in similar studies and the composer's clearly intended meaning.

After allowing for appropriate procedural accommodations for participants with hearing or vision loss, all participants were asked to select the image portrayed by the music.

A Univariate ANOVA was computed to address the research question, "Do students with vision or hearing loss assign the same visual images to music as students without such sensory losses?" Data were analyzed to examine the effects of sensory condition as well as age and gender.

A significant main effect was found for sensory condition, with follow up tests indicating that participants with typical hearing and vision agreed with the composer's intended meaning significantly more often than did participants with vision or hearing loss. No significant main effects were found for gender or age, and no significant interactions were found.

Summary data indicated that selected images were more easily identified, or were more difficult to identify across conditions. The data also revealed an order of difficulty and patterns of confusion that were similar across sensory conditions and ages, indicating participant responses were not random, and that some referential meaning in music is conventional.

## 2.9.2. Artigo Original

# The Effect of Vision and Hearing Loss on Listeners Perception of Referential Meaning in Music

A.-A. Darrow & J. Novak (2005)

There are two traditional, but opposing aesthetic positions related to music – absolutist and referentialist perspectives (Radocy & Boyle, 1997). The absolutist position is that music refers to nothing outside of itself; that is, there is no meaning in music beyond that inherent in the musical sounds themselves (Meyer, 1956). The opposing or referentialist position is that music can refer to nonmusical events or concepts, and therefore, can arouse mental images (Langer, 1959). Though it is generally agreed that music has the power to evoke images in the listener's mind, the perception of such images is dependent upon listeners ability to aurally discriminate musical detail, and to mentally organize what has been discriminated. The role of these abilities in determining imagery in music has important implications for listeners with vision or hearing loss. Furthermore, Deschenes (1995) suggests that images do not occur logically because of the music, but rather as symbols of the listener's experiences – a suggestion that has implications for listeners with vision or hearing loss as well.

Though all music has the potential to evoke images in the listener, one genre of music, program music, is composed specifically for the purposes of prompting images or ideas in the mind of the listener. Program music is instrumental music that endeavors to arouse mental pictures in the mind of the listener (Lagassé, 2001). The "program" of a piece is the story or subject, often indicated by the piece's title. In one form or another, program music has existed at least since the 1600s (Hoffman, 1997). Early examples of program music include well-known pieces such as Antonio Vivaldi's *Four Seasons*. Although listeners familiar with such well-known pieces may readily agree with the composers intended images, musical sounds cannot convey concrete meaning in the way that words can. Music does have the capacity to evoke shared associations; however, it will inevitably evoke different associations for its listeners, and associations that perhaps vary from the composer's intended meaning (Wagner & Darrow, 1983; Trainor & Trehub, 1992).

Interested in the ability of listeners with typical hearing and vision to decode referential meaning in music, Trainor and Trehub (1992) and Wagner and Darrow (1983) examined listeners visual interpretation of selected music. Using Prokofiev's *Peter and the Wolf*, Trainor and Trehub found that 3-, 4- and 6-year-old children matched pictures to the appropriate musical excerpts significantly better than chance, but were more adept at decoding the intended meaning in some excerpts than others. Using Saint Saëns, *Carnival of the Animals*, they found again that children's performances were better than chance at identifying the composer's intended imagery. Wagner and Darrow (1983) were interested in developmental factors related to listeners visual interpretation of music. They examined children's, adolescents, and adults ability to identify a composer's intended images. Also using Saint Saëns, *Carnival of the Animals*, they found that age, and thus life experiences, were significant factors that aided the listener in decoding the composer's musical message. This study suggests that music is subject to a range of interpretations that are perhaps more specific to the unique experiences of the listener than to the characteristics of the music (Deschenes, 1995; Wagner & Darrow, 1983).

Students with a sensory loss experience music, and indeed the world around them, differently than do individuals with typical hearing and vision. Structuring music experiences for students who have vision loss or hearing loss requires attention to typical instructional issues –



such as the students intellectual, musical, and social development, but in addition, special attention to issues such as proper presentation of the musical stimuli, alternative means for student response, and appropriate adaptations to the environment (Adamek & Darrow, in press). Also to be considered is how the mind, which is receiving limited sensory information, organizes perceptual material from musical sounds. Musical concept formation relies on a child's ability to make judgments about what he or she hears. Visual symbols are used most often to represent musical concepts. Students with vision loss or hearing loss are both at a disadvantage, either receptively or expressively, to interpret music via visual representations.

Children typically learn much about the world around them by visual observation. Most children with limited vision have learned to use other senses to acquire much of the information that is transmitted in their environment; however, there is still information that is strictly visual and cannot be accessed in any other way. Children who are blind or who have severe visual impairments frequently use imagery words for which they have no sensory experience – such as when describing a blue sky. These imagery words are a type of echolalia that Cursforth (1951) coined “verbalisms.” There is no evidence to suggest, however, that such expressions are used without some meaning to children with severe visual limitations (Adamek & Darrow, in press). There are other words in these children's vocabulary – such as animals, which they may also use, but without hands-on experiences, their understanding is reliant upon whatever verbal descriptions they have been given.

Walker (1981) devised a test using pictures that were theorized to “look like” musical sounds. Using Braille answer sheets that were embossed to “feel like” the pictorial representations of musical sounds, Walker (1985) tested 14 subjects who were congenitally blind. He found that these subjects selected “felt” images that corresponded to the visual images chosen by sighted subjects, suggesting that musical sounds produce inherent “images” in the listener. However, in a subsequent study by Madsen and Darrow (1989), they found that subjects with vision loss evidenced different “images” than did the subjects in Walker's study, indicating that aural-related mental images may be relevant only for persons with certain backgrounds and not at all relevant for some populations.

Music therapists and music educators have long advocated the use of visual symbols – such as listening maps or stick notation, as a means to guide children's interpretation of musical sounds. In auditory testing situations, children with hearing loss often use graphic representations to illustrate what they hear. Music educators and therapists of students with hearing loss have also suggested the use of visual symbols to interpret musical sounds (Darrow & Gfeller, 1991). However, even students with severe hearing loss have some aural access to music through their residual hearing. Using their residual hearing, they can detect certain frequencies, timbres, and intensities of music (Darrow, 1991, 1992). Many individuals with hearing loss have learned to make discriminations about the music they hear and to interpret and attach meaning to it (Darrow, 1993, Darrow & Cohen, 1991; Gfeller, 2001). How the music is interpreted and the meaning attached to it, however, is often different from the typical listener. Using face drawings to depict emotions, Darrow (in press) found that children with hearing loss assigned significantly different emotions to music than did children with typical hearing.

The purpose of the present study was to investigate whether students with vision or hearing loss assign the same images to music as students without such sensory losses. For individuals who have hearing loss, how much of the musical complexities can be missed and the visual image still conveyed? We know that individuals who have hearing loss often perceive music differently from typical hearing listeners, yet they do have, in some areas, similar musical aptitude, preference, and performance abilities (Darrow, 1987, 1991, Gfeller & Lansing, 1991, 1992; Gfeller, et al., 1998). Can such individuals, even with missing musical information, decode the intended visual image in ways similar to hearing individuals? Is some music, because of its compositional features, more reliable to convey images? For individuals who have vision loss, what form does musical imagery take? Can it be analogous to or representative of objects they may have never seen? Can students with vision loss learn sufficiently about the characteristics of visual subject matter through verbal descriptions or hands-on experiences such that they can determine their musical representation?



## Method

### Participants

Students at a state school for the deaf and blind and students at neighboring public schools served as participants ( $N = 96$ ). Of the 64 participants at the state school for the deaf and blind, 30 were students with hearing loss who had aided or unaided hearing thresholds within the speech range, and 34 were students with low vision ( $N = 17$ ) or students who were blind ( $N = 17$ ). Participants with typical hearing and vision ( $N = 32$ ) were matched on age and gender, as much as possible, to the participants with vision and hearing loss. Ages ranged from 8 years to 18 years with a mean age of 13 years. Fifty-six males and 40 female students participated in the study.

### Music Stimuli and Response Format

The music stimuli consisted of six different 37-second randomly ordered excerpts from Saint Saëns, *Carnival of the Animals* (Ondrej Lenárd, Slovak Radio Symphony). The excerpts were chosen because of their use in similar studies, and because of the composer's clearly intended meaning, which is evident in their titles: The Aquarium (fish), The Elephant, The Mule, The Aviary (bird), March of the Royal Lion, and the Kangaroo. A list of the excerpts is found in Table 1. The response format was pictorial representations of the excerpts for participants with typical hearing and vision, and for participants with hearing loss. The pictures consisted of black line drawings on white paper. Response sheets were enlarged for participants with low vision. Participants who were blind indicated their response to each excerpt (selected from a six-alternative forced choice response set) using a Perkins Braille. No spoken narration was used on the recording except a male voice announcing the number of the excerpt that followed.

### Procedures

Participants with hearing loss, wearing their personal hearing aids, were tested in a typical classroom listening environment in small groups of three to four students. Participants were seated in front of two stereo speakers set up on tables. Participants were encouraged to feel the speakers or tables if they wished. The volume of each excerpt was set at the highest level not resulting in distortion of the music. Participants were encouraged to sit at whatever distance from the speakers resulted in their most comfortable listening level (MCL). After listening to each excerpt, participants were asked to circle the picture that best described the image portrayed by the music. The response set for all subjects was six-alternative forced choice.

Participants with vision loss were tested in a typical classroom listening environment, also in groups of three to four students. The volume was set at the typical classroom listening level for music activities. After listening to each excerpt, participants were asked to either – circle the (enlarged) picture that best described the image portrayed by the music, or to Braille their response given the six-alternative forced choice response set.

Participants with typical hearing and vision were tested in their music classrooms on the classroom stereo. The volume was set at the typical classroom listening level for music activities. After listening to each excerpt, participants were asked to circle the picture that best described the image portrayed by the music.



## Results

A preliminary t-test revealed no significance differences between the responses of participants with low vision and participants who were blind; subsequently, these participants responses were grouped together for further analyses. A Univariate ANOVA was computed to address the research question, "Do students with vision or hearing loss assign the same visual images to music as students without such sensory losses?" Data were analyzed to examine the effects of sensory condition (hearing loss, vision loss, typical hearing and vision) as well as age (younger /8-13 years, older/14-18 years) and gender (male, female).

A significant main effect was found for the sensory condition ( $F(2, 96) = 16.05, p < .05$ ). See Table 2. Follow-up tests indicated that participants with typical hearing and vision agreed with the composer's intended meaning significantly more often than did participants with vision or hearing loss. Those data are shown in Table 3. No significant main effects were found for gender or age, and no significant interactions were found.

The data were also examined for correct responses to individual items by sensory condition groups. A summary of those figures appears in Table 4. These data indicate that selected images were more easily identified (Aviary), or were more difficult to identify (Mule) across conditions. The data also revealed an order of difficulty and patterns of confusion that were similar across sensory conditions and ages. A series of Chi-Square tests indicated there was no significant difference between groups for individual excerpts, with the exception of excerpt #1, the Aquarium ( $\chi^2 = 8.64, df = 2, \alpha = .01$ ). On this excerpt, participants with hearing loss agreed with composer's intended image significantly less often than did participants with vision loss or typical hearing and vision. These data would indicate that sensory condition group differences were due to overall performance and not to individual test items.

## Discussion

The purpose of the present study was to investigate whether students with vision or hearing loss assign the same images to music as students without such sensory losses. The findings would indicate that they do not assign the same images to music; however, participants with vision and hearing loss did respond similarly to participants with typical hearing and vision in that selected images were more easily identified (Aviary), or were more difficult to identify (Mule), regardless of their sensory conditions. The data also revealed an order of difficulty and patterns of confusion that were similar across sensory conditions and ages, indicating participant responses were not random, and that some referential meaning in music is conventional.

For all sensory condition groups, there was no significant difference between younger and older participants, or between males and females. Although some research has indicated differences in subjects ability to identify intended images based on age (Wagner & Darrow, 1983), the findings of the present investigation would seem to corroborate other research (Trainor & Trehub, 1992) where no age effects were found, though performances became more accurate with increasing age.

The summary data in Table 4 suggests there are compositional dimensions of music – such as tempo, rhythm, texture, timbre, or pitch that influence the participants perception of imagery in music. The flute timbre, loudness, and rapid tempo of the Aviary excerpt successfully conveyed the composer's intent to many participants. Likewise, some of the same compositional features – though softer timbre and volume, also made the Aquarium excerpt salient for many participants, but most likely audibly inaccessible for participants with hearing loss (see Table 4). Several adaptive strategies can be used to assist children with hearing loss in accessing the compositional components of music that are intended to portray images. Such strategies include providing tactile stimulation and visual displays to supplement the music (Darrow, 1992; Darrow & Goll, 1989; Oval Window Audio, 2004). Music used to portray images may also need to be transposed. Even children with little hearing are able to hear tones that occur within their range of residual hearing.



The hearing range of children with hearing loss is highly variable and individualized; however, their range of residual hearing is generally in the lower frequencies (Johnson, Benson, & Seaton, 1997).

Overall, the Aviary and Aquarium excerpts received the most correct responses, perhaps because participants were more familiar with these images, perhaps even having fish or birds in their homes. However, responses to these two excerpts also showed a pattern of confusion. Many children selected the bird for the Aquarium excerpt. Because this error was common, it is unlikely that these responses were random. Errors may not have been because children failed to appreciate the composer's imitative techniques, but rather confused what aspect of the bird the composer was trying to imitate: Was it the "fluttering" of bird wings (made by a flute), or the gentle "gliding" movement of a bird (made by Celeste)? Another common pattern of confusion was between the Lion and the Elephant excerpts. It may be that confusion was not due to ambiguity on the part of the composer, but rather participants had similar impressions of these animals – such as large, bold, majestic, lumbering.

All participants performed poorly on the Mule excerpt, where the glissando of a violin bow was meant to express the vocal sounds of a mule. Clearly, such a composition technique would have little meaning for participants with a hearing loss, or even hearing participants who have never heard a mule. The same rationalization could be made for responses to the Lion excerpt, where piano arpeggios are meant to express the roar of the lion, though it is perhaps more likely that participants have heard a lion roar than heard a mule, and indeed more participants identified the Lion excerpt correctly.

The hearing loss group agreed with the composer's intended meaning significantly less often than did participants with vision or hearing loss (Table 4), indicating that perhaps hearing loss is a greater deterrent to perceiving imagery in music than is vision loss. There are a number of reasons that may explain why the hearing loss group responded differently than the other two groups. The first and most obvious reason is, that even with considerable amplification, they may not have heard, or more likely, listened to the excerpts as well as hearing participants. While the audiograms of these participants indicate that they could *hear* the music, an audiogram is no indication of a child's ability to *listen*. The complex task of listening is a much more involved task than the physical act of hearing. The child who has hearing loss must learn through varied experiences to use his or her residual hearing to interpret sounds, attach meaning to them, and develop listening rules, cues and strategies.

Some people believe that when a child's vision is absent, his or her hearing is naturally enhanced; however, the listening skills of children who are blind are developed just as all listening skills are – through practice. By virtue of their disability, children who are blind must practice and utilize their listening skills more earnestly than most children. There are some data, however, that suggest the earlier the onset of blindness, the more perceptive the individual's listening skills (Gougoux, Lepore, Lassonde, Voss, Zatorre, & Beline, 2004). It is likely that the difficulty for students with vision loss was not in hearing or listening to the music, but in determining the corresponding image.

All interpretations of music are guided by the listener's experiences. Beyond the constraints of their disability, the unique life experiences of participants with vision and hearing loss may have also played a role in their performance. The environment of a state school for the deaf and blind is purposefully structured so that students do not need to rely on their residual vision or hearing, and therefore are less practiced at interpreting what they see or hear. For students who are deaf, there is usually little or no attention given to sounds, and music is typically not a part of daily life. It is likely that most of the participants with hearing loss had never been asked before to determine imagery in music, or even to attend to music with a specific sense of purpose. Likewise, in a school for the blind, even students with partial sight are rarely asked to envision images.

The process of matching a picture to music requires that listeners first have an image in their mind. Since experience precedes imagery, a participant would need to have had various experiences with the selected animals to form an image – whatever form that image might take. It is perhaps then lack of appropriate experiences that influenced the responses of some participants, rather than their vision loss or an inability to form mental images. For example, the Kangaroo excerpt relies on the rhythm of the music to express the movement of a kangaroo. Students who were blind did poorly on this excerpt, perhaps because it is unlikely that any students would have had hands-on experience with such an animal, and certainly not to the extent that they would be familiar with its movements. They may have *heard* that a kangaroo hops, but how clearly



that information is integrated into their thinking would be difficult to conjecture. Lack of appropriate experiences would also influence the responses of participants with hearing loss since the intended musical metaphors, even when participants could hear them, would be lost without such experiences.

## Clinical Implications

There are many concepts, such as object identification, that children typically learn through vision; therefore, these concepts are never specifically taught. Even for children who are not blind, but who have low vision, objects are seldom seen clearly. Vicarious learning then is at risk for all children with vision loss. The world meets sighted children half way. Curious to explore what they see, children move toward objects and people, ever increasing their knowledge of the world around them. Children without sight are no less curious; however, they require an environment that is aurally and tactually inviting. When structuring music activities for students with vision loss, children who are blind or who have low vision must be taught primarily through direct instruction and hands-on experiences. These experiences should be frequent and varied enough that children learn to make generalizations. For example, when children sing songs about animals, large pictures or stuffed animals, or perhaps even real animals would enhance the meaning of the lyrics, not just for those with vision loss.

Although many music students who are blind can easily learn music by listening to others play or sing, it is to their benefit to provide them with enlarged scores or tactile translations of the printed music – the shapes of notes and symbols, and the linear format of music. By doing so, the student is provided practice in interpreting the sighted world, in addition to the fact that much of this tactile translation information will later transfer to reading Braille music. Children should be encouraged to use their residual vision, and to visually analyze what they see. Like listening, perceiving is a skill that must be practiced.

Because hearing loss is a disability that is not readily apparent, its impact on a child's development is often underestimated. Successful communication – the exchange of meaningful information, is central to children's understanding of the world around them. Without communication, the resulting isolation can affect all aspects of a child's psychological, social, and intellectual development. When structuring music activities for children with hearing loss, visual support materials are essential. Visual representations of music allow children to define or to clarify what they hear. For example, when children sing songs about animals that include imitations of their sounds, such as *Old MacDonald*, – again, pictures or stuffed animals, would bring meaning to these sounds, and the song. Other suggestions for prompting musical imagery with children who have sensory loss include providing them with frequent and varied music listening experiences, posing questions and prompting discussion by asking children to: describe what they hear, listen for musical analogies, relate the music to everyday experiences, or even to draw images of what they hear.

In summary, this study indicates that some children can assign concrete meaning to music by means of imitative or metaphorical compositional techniques, and thus, do have the capacity to understand referential meaning in music. As music therapists and educators, it is important to provide frequent opportunities for all children to listen to music and to utilize their sense of imagination. "Listening to music must mean more than opening the ears – it should mean opening the mind to imagery" (Zerull, 1992, p. 28). We must, however, be mindful of our students unique characteristics and experiences, and plan accordingly.

## References

- Adamek, M., & Darrow, A. A. (in press). *Music in special education*. St. Louis: MMB Publishers.
- Cutsforth, T. D. (1951). *The blind in school and society*. New York: American Foundation for the Blind.
- Darrow, A. A. (in press). The role of music in deaf culture: Deaf students' perception of emotion in music. *Journal of Music Therapy*.
- Darrow, A.A. (1993). The role of music in Deaf culture: Implications for music educators. *Journal of Research in Music Education*, 41 (2), 93-110.
- Darrow, A. A. (1992). The effect of vibrotactile stimuli on the recognition of pitch change by hearing impaired children. *Journal of Music Therapy*, 29, 103-112.
- Darrow, A.A. (1991). An assessment and comparison of hearing impaired children's preference for timbre and musical instruments. *Journal of Music Therapy*, 28, 48-59.
- Darrow, A. A. (1987). An investigative study: The effect of hearing impairment on music aptitude. *Journal of Music Therapy*, 24, 88-96.
- Darrow, A. A., & Cohen, N. (1991). The effect of programmed pitch practice and private instruction on the vocal reproduction accuracy of hearing impaired children: Two case studies. In K. Bruscia (Ed.), *Case Studies in Music Therapy* (pp. 191-204). Barcelona Publishers: Philadelphia, PA.
- Darrow, A. A., & Gfeller, K. (1991). A study of public school music programs mainstreaming hearing impaired students. *Journal of Music Therapy*, 28, 23-39.
- Darrow, A. A., & Goll, H. (1989). The effect of vibrotactile stimuli via the SOMATRON™ on the identification of rhythmic concepts by hearing impaired children. *Journal of Music Therapy*, 26, 115-124.
- Deschenes, B. (1995). Music and symbols. *Music Therapy Perspectives*, 13, 40-45.
- Gfeller, K. (2001). Aural Rehabilitation of Music Listening for Adult Cochlear Implant Recipients: Addressing Learner Characteristics. *Music Therapy Perspectives*, 19, 88-95.
- Gfeller, K.E. & Lansing, C.R. (1991). Melodic, rhythmic, and timbral perception of adult cochlear implant users. *Journal of Speech and Hearing Research*, 34, 916-920.
- Gfeller, K. E. & Lansing, C. R. (1992). Musical perception of cochlear implant users as measured by the primary measures of music audiation: An item analysis. *Journal of Music Therapy*, 29, 18-39.
- Gfeller, K., Knutson, J.F., Woodworth, G., Witt, S., & DeBus, B. (1998). Timbral recognition and appraisal by adult cochlear implant users and normally hearing adults. *Journal of the American Academy of Audiology*, 9, 1-19.
- Gougoux, F., Lepore, F., Lassonde, M., Voss, P., Zatorre, R.J., & Belin, R. (2004). Pitch discrimination in the early blind: People blinded in infancy have sharper listening skills than those who lost their sight later. *Nature*, 430, 309.



- Hoffman, M. (1997). *The NPR® Classical Music Companion: Terms and Concepts from A to Z*. Boston, MA: Houghton Mifflin Company.
- Johnson, C., Benson, P., Seaton, J. (1997). *Educational Audiology Handbook, 1st Edition*. San Diego: Singular Publishing Group, Inc.
- Lagassé, P. (Ed.) (2004). *The Columbia Encyclopedia*, 6<sup>th</sup> ed. New York: Columbia University Press.
- Langer, S. (1959). *Philosophy in a new key*. Cambridge, MA: Harvard University Press.
- Madsen, C. K., & Darrow, A. A. (1989). The relationship between music aptitude and sound conceptualization of the visually impaired. *Journal of Music Therapy*, 26, 71-78.
- Meyer, L. B. (1956). *Emotion and meaning in music*. Chicago, IL: University of Chicago Press.
- Oval Window Audio (2004). *Multisensory Sound Lab*, Retrieved March 3, 2005 from <http://www.ovalwindowaudio.com/>
- Radocy, R. E., & Boyle, J. D. (1997). *Psychological Foundations of Musical Behaviors*. Springfield, IL: Charles C. Thomas Publisher.
- Trainor, L. J., & Trehub, S. E. (1992). The development of referential meaning in music. *Music Perception*, 9, 455-470.
- Wagner, M. J., & Darrow, A. A. (1983). A comparison of the suggested musical images of a composer and the images chosen by child, adolescent, and adult listeners. *Contributions to Music Education*, 9, 29-35.
- Walker, A. R. (1981). The presence of internalized images of musical sounds. *Bulletin of the Council for Research in Music Education*, 66-67, 107-112.
- Walker, A. R. (1985). Mental imagery and musical concepts: Some evidence from the congenitally blind. *Bulletin of the Council for Research in Music Education*, 85, 227-238.
- Zerull, D. (1992). Just imagine. . . improving the band experience. *Music Educators Journal*, 78b, 25-28.

**Table 1 - Order of Musical Excerpts**

| Order | Recording                    | Test Stimuli        | Response Sheet |
|-------|------------------------------|---------------------|----------------|
| 1     | March of the Royal Lion (#1) | The Aquarium (Fish) | Kangaroo       |
| 2     | The Elephant (#5)            | Elephant            | Mule           |
| 3     | The Kangaroo (#6)            | Mules               | Lion           |
| 4     | The Aquarium (#7)            | The Aviary (Bird)   | Elephant       |
| 5     | The Mules (#8)               | Lion                | Fish           |
| 6     | The Aviary (#10)             | Kangaroo            | Bird           |

**Table 2 – Univariate ANOVA of Correct Responses by Sensory Condition, Age, and Gender**

| Source                 | <i>F</i> | <i>df</i> | <i>p</i> |
|------------------------|----------|-----------|----------|
| Sensory Condition      | 16.06    | 2         | .000*    |
| Gender                 | .55      | 1         | .462     |
| Age                    | .61      | 1         | .437     |
| Condition/ Gender      | .12      | 2         | .883     |
| Condition/ Age         | 2.54     | 2         | .085     |
| Gender/ Age            | 1.29     | 1         | .260     |
| Condition/ Gender/ Age | .92      | 2         | .403     |

Note: \*indicates a significant difference,  $p < .05$

**Table 3 - Post Hoc Follow up Test to ANOVA of Sensory Group Performances**

| Sensory Groups | Sensory Condition Means |             |                  |
|----------------|-------------------------|-------------|------------------|
| Group          | Hearing Loss            | Vision Loss | Hearing & Vision |
|                | 1.90                    | 2.56        | 3.75             |

Note: Underline indicates no significant difference,  $p < .05$



**Table 4 - Summary Table of Correct Responses by Sensory Conditions and Excerpts**

| Excerpts | Sensory Conditions |              |                  | Total Correct for Excerpt  |
|----------|--------------------|--------------|------------------|----------------------------|
|          | Vision Loss        | Hearing Loss | Hearing & Vision |                            |
| Aquarium | 25                 | 7            | 27               | = 59                       |
| Elephant | 17                 | 11           | 22               | = 50                       |
| Mule     | 6                  | 6            | 8                | = 20                       |
| Aviary   | 24                 | 12           | 25               | = 61                       |
| Lion     | 9                  | 12           | 20               | = 41                       |
| Kangaroo | 7                  | 9            | 15               | = 31                       |
|          | 88                 | 57           | 117              | = Total Correct for Groups |

### 2.9.3. Estudo Crítico – Por Christopher Johnson

#### Objectivo

O objectivo do estudo era investigar se estudantes com perda de visão ou de audição atribuíam as mesmas imagens à música, que estudantes sem essas perdas.

Para os indivíduos com perda de audição, as questões que conduziram esta investigação foram: (1) Qual a quantidade mínima de complexidade musical necessária para para que se mantenha a evocação da imagem visual?; (2) Podem, estes indivíduos, mesmo com informação musical perdida, descodificar a imagem visual pretendida da mesma forma que os indivíduos sem perda auditiva? (3) Podem determinadas características da composição tornar mais viável a associação de imagens?

Para os indivíduos com perda de visão, as questões que conduziram a investigação foram: (1) Que forma toma a imagética musical?; (2) Pode a imagética musical evocar objectos nunca percebidos pela visão?; (3) Podem os estudantes com perda de visão aprender o suficiente acerca de assuntos visuais, através de descrições verbais ou de experiências tácteis, de forma a determinar a sua representação visual?

#### Design

O plano desta investigação é quasi-experimental. A variável independente era a perda de visão ou de audição (ou ausência das mesmas) dos participantes. A variável dependente era a atribuição de imagens ao estímulo musical. Outras variáveis independentes examinadas foram o sexo e a idade dos participantes.

#### Revisão do Artigo

#### Introdução e revisão bibliográfica

A introdução a este artigo revê as Escolas do pensamento em estética da música. É um comentário fascinante, uma vez que parece claro que pessoas que desenvolveram teorias estéticas, não consideraram as populações com as incapacidades abordadas neste estudo. Embora Darrow e Novak não vão ao ponto de dizer que este hiato existente nestas filosofias põe em causa toda a teoria, muitos leitores poderiam colocar questões que até agora não foram postas.

No primeiro parágrafo da revisão, Darrow e Novak começam por construir o seu argumento à volta da ideia da intenção programática do compositor e das descodificações que o ouvinte faz acerca dos programas referenciais em música. Estas conexões são exploradas na bibliografia existente em relação a certas obras programáticas e às associações que estas obras suscitam. No quarto e quinto parágrafos, os autores abordam a problemática dos estudantes com perda sensorial. São levantadas questões como a experiência e a compreensão musical. São também consideradas questões relacionadas com a interpretação musical ao nível da recepção e da expressão.

No sexto parágrafo, Darrow e Novak revêem estudos onde estudantes com alguma deficiência visual tinham de utilizar diferentes tipos de imagens. Um desses estudos descobriu que os estudantes tinham mais tendência para escolher imagens “sensoriais” (“feel like” e “look like”) em relação à música. Um outro estudo encontrou resultados contrários, com os estudantes a escolherem imagens diferentes daquelas encontradas na investigação anterior. A última área de revisão destinou-se à pesquisa sobre a imagética relacionada com a música, em estudantes com



perdas parciais da audição. Existindo já bastante investigação nesta área, salienta-se o facto de Darrow e Novak conseguirem resumir, num único parágrafo, um grande corpo desses dados existentes.

## Método

Dois grupos de participantes eram oriundos de uma escola para crianças com perdas de visão ou de audição. O terceiro grupo era formado por alunos sem perdas de audição ou de visão, mas geograficamente próximos da escola de ensino especial. O número total de participantes foi de noventa e seis (96), dos quais, cinquenta e seis (56) eram do sexo masculino e quarenta (40) eram do sexo feminino. As idades dos participantes estavam compreendidas entre os oito (8) e os dezoito (18) anos, sendo a média de idades de treze (13) anos.

A investigadora escolheu seis excertos de trinta e sete (37) segundos cada, retirados da obra *O Carnaval dos Animais*, de Saint Saens. A escolha desta obra deveu-se a dois factores essenciais: (1) o facto de Saint Saens ter tido uma clara intenção de evocar imagens nessa obra; (2) e o facto dessa obra já ter sido utilizada no passado por muitos outros investigadores da mesma área.

Os seis excertos escolhidos foram retirados dos andamentos com os seguintes títulos: *O Aquário*, *O Elefante*, *O Burro*, *O Aviário (galinhas)*, *Marcha do Leão Real* e *O Canguru*. Foram feitas adaptações para as formas de resposta, utilizando-se traçados simples para os participantes invisuais ou com visão muito limitada. Para aqueles com baixa visão, as imagens foram ampliadas. Para os estudantes completamente invisuais, recorreu-se a uma folha, em *Braille*, com seis opções de resposta.

Os procedimentos habituais para a escuta e sinalização de respostas (através de círculos) foram os normais. Os indivíduos com perda de audição foram testados individualmente de forma a adaptar a sua incapacidade à circunstância. Os estudantes com alguma perda de visão foram testados numa sala de aula típica, assim como os participantes com audição e visão normais. Todos os participantes ouviram um excerto e depois assinalaram a imagem que julgaram melhor representar o fragmento musical apresentado. O mesmo processo foi repetido para os cinco excertos seguintes.

## Resultados

Os resultados procuram responder a duas questões: a primeira é relativa à diferenças entre os grupos; a segunda é relativa à observação dos dados para cada excerto.

Os dados foram analisados através do procedimento "Univariate ANOVA" para examinar o número de respostas correctas em cada grupo, bem como o sexo e idade de cada participante. Foram encontradas diferenças significativas para a variável *grupo*, o mesmo não acontecendo para as variáveis *sexo* e a *idade*. Para além disso não se verificaram interacções entre as variáveis. Os testes seguintes indicaram que os participantes sem perda de visão e audição tiveram mais respostas concordantes com o significado pretendido pelo compositor, do que os outros participantes. Neste aspecto, a diferença entre participantes com e sem perda de visão e audição foi significativa.

Os dados foram também examinados no sentido de se detectarem, em cada excerto, diferenças entre os grupos. Na maior parte dos casos, os níveis de respostas correctas foram encontrados em todos os grupos, com excepção do excerto "O Aquário". Neste caso, o grupo com perda de audição teve resultados significativamente piores que os outros grupos.

As questões mais abrangentes da investigação não foram levantadas na secção dos resultados, nem sujeitas a tratamento estatístico, tendo, no entanto, sido tratadas na secção seguinte, como se passa a descrever.



## Discussão

No primeiro parágrafo, Darrow e Novak reiteraram o propósito do projecto e explicaram como é que os estudantes com perdas de visão e audição responderam. Embora este grupo não tenha respondido de forma tão precisa, responderam de forma similar ao grupo sem perdas de audição e visão. Os excertos que eram mais frequentemente identificados de forma correcta por um grupo também o eram pelo outro. A investigadora discute depois o motivo da inexistência de efeitos do sexo e da idade e compara os resultados com investigações anteriores.

Nos quatro parágrafos seguintes, Darrow e Novak procuram perceber o motivo pelo qual os estudantes são capazes de discernir melhor as intenções do compositor em certos excertos do que noutros. Examina as características composicionais, bem como a forma como os animais podem ter sido percebidos pelos participantes (por exemplo, os leões e os elefantes podem ter sido percebidos como grandes, majestosos, cheios, pesados, etc.).

As autoras iniciam então a discussão sobre as diferenças de resultados obtidas nas respostas dos estudantes com perdas sensoriais. No caso dos indivíduos com perda de audição, pode acontecer que eles comecem a ouvir o estímulo pelo simples facto de se aumentar o volume do som, o que não significa que tenham realmente aprendido a escutar. De facto, o aumento do volume de som pode não ser um factor determinante na melhoria da percepção sonora de um estímulo, uma vez que um estudante com perdas sensoriais pode ter aprendido a compensar a incapacidade de receber estímulos auditivos, não necessariamente sabendo analisá-los quando os passa a ouvir. Também é possível que os estudantes com perda de visão possam compensar essa perda concentrando-se na audição, promovendo essa compensação com o treino e/ou com a atenção redobrada aos estímulos auditivos. Além disso, é ainda apontada a forte possibilidade de as imagens visuais dos animais ou os seus nomes (que eram dados a escolher e a relacionar com a música) poderem não ter o mesmo significado para estudantes com perdas sensoriais (visuais ou auditivas), relativamente aos estudantes sem este tipo de perdas sensoriais.

Darrow e Novak incluem ainda uma curta secção sobre as implicações clínicas relacionadas com estes resultados. Nessa secção, discute-se a forma como, mediante o fornecimento de novos materiais e experiências, se poderão ajudar pessoas com deficiências visuais (e também auditivas) a conseguirem uma melhor interpretação e compreensão da música com a qual contactam.

## Contribuições para a área de estudo

Este artigo contribui em vários aspectos para os domínios da educação musical e da musicoterapia. Houve bastantes estudos que utilizaram excertos de Saint Saens. Claramente, este estudo alarga os dados sobre audição e imagética a uma nova população. Darrow tem um extenso registo de pesquisas sobre crianças com deficiências, e este estudo é mais um passo na compreensão destas crianças de forma mais completa.

Os participantes puderam responder correctamente a muitos dos excertos (que fornecem muita informação auditiva), excepto no excerto *O Aquário*, que não foi prontamente identificado. Depois da discussão destes resultados, as autoras adicionaram uma secção intitulada "Implicações Clínicas". Embora essas implicações clínicas sejam dirigidas especificamente a musicoterapeutas, qualquer professor de música que tenha estudantes com deficiências similares poderá facilmente incorporar essas sugestões nas suas aulas. Essas sugestões não são apenas para populações com perdas auditivas, mas também servem para enquadrar deficientes visuais. Existem muitos livros sobre o ensino especial de crianças. No entanto, raramente se encontra um estudo baseado em dados experimentais que, simultaneamente, seja acompanhado de aplicações específicas baseadas nesses conhecimentos novos.



## Comentário

Um dos contributos importantes deste estudo é a interessante revisão bibliográfica efectuada pelas autoras. Darrow e Novak indicam os nomes mais sonantes da filosofia da música e expõem de forma aliciante e sucinta as ideias essenciais desses filósofos, e as práticas associadas a essas teorias. No seguimento, Darrow e Novak criticam esses filósofos por serem responsáveis pela elaboração de filosofias que deixam de lado uma parte significativa da humanidade.

Nos dois últimos parágrafos, as autoras determinam quais as áreas do conhecimento relacionadas com as deficiências em estudo e, em seguida, definem com clareza os objectivos e a pertinência do estudo que se propuseram a realizar.

Um ponto a criticar neste estudo é o facto de, na sua declaração de objectivos, serem apresentadas questões pertinentes que, embora estando presentes na discussão, acabaram por não ser todas tratadas como verdadeiras questões da investigação.

Não obstante a existência de algumas opções subjectivas, a metodologia adoptada é, na sua globalidade, clara e linear. Embora não seja clara a forma como a lista de escolha obrigatória foi apresentada aos estudantes invisuais, devemos presumir – tendo em conta que Darrow é especialista em estudos sobre crianças com necessidades especiais, para além de também estar bem documentada na área de metodologia de investigação – que os testes terão sido aplicados com rigor. Todos os outros aspectos estão suficientemente bem descritos, de forma a possibilitar eventuais replicações.

A secção dos resultados é também bastante transparente. Os testes usados são claramente explicados e são apresentadas razões para a sua utilização.

Destacam-se dois resultados de particular interesse. Em primeiro lugar, realça-se o facto de os estudantes “normais” terem tido melhor desempenho que os estudantes com deficiências visuais ou auditivas, sem diferenças significativas entre os estudantes com deficiências visuais e auditivas. O segundo resultado importante diz-nos que os diferentes excertos musicais podem conter características composicionais que os tornam mais ou menos difíceis de identificar. Efectivamente, os três grupos acertaram com sucesso nas tarefas pedidas aproximadamente pela mesma ordem. Isto dá peso ao último argumento da autora, respeitante à educação dos estudantes com deficiências.

O grande mérito desta investigação encontra-se na secção «Discussão», onde são apresentados os esclarecimentos mais relevantes, e onde são feitas importantes interpretações de significado.

Outra contribuição muito importante deste estudo é a reflexão acerca das possíveis formas de se ensinarem estudantes com perdas de visão ou audição a fazerem as associações com maior precisão. Darrow e Novak conseguem informar-nos sobre capacidades importantes que as pessoas com deficiência auditiva ou visual podem ter, informando-nos ainda sobre formas concretas de actuação para melhorar o ensino dessas pessoas.

Em suma, estamos perante uma investigação muito pertinente, que resulta num excelente artigo, escrito com clareza e cuidado, contribuindo valiosamente para a sua área de estudo.

---

(1) Pode-se entender como significados referenciais em música, os significados semânticos extra-musicais que uma determinada música pode invocar. Habitualmente este tipo de música, denomina-se de *música programática*. As invocações são propostas pelo compositor através de títulos ou textos. Neste caso, são títulos com nomes de animais, que pretendem provocar na imaginação das crianças as imagens desses animais. (N.T.)

# 2.10.

## AN INVESTIGATION INTO THE EXPERIENCE OF MUSICAL PERFORMANCE ANXIETY IN ELEMENTARY SCHOOL CHILDREN

ESTUDO DA ANSIEDADE NO DESEMPENHO MUSICAL  
EM CRIANÇAS DO ENSINO BÁSICO

C. RYAN (2005)

COMUNICAÇÃO APRESENTADA NO 16<sup>o</sup> INTERNATIONAL SYMPOSIUM ON RESEARCH IN MUSIC BEHAVIOR, ST.  
LOUIS, MO

### 2.10.1. Resumo do artigo / Abstract

#### Resumo

A ansiedade face ao desempenho musical é um problema comum nos músicos. A investigação empírica com músicos adultos tem vindo a crescer, ao passo que nas crianças esta área tem sido descurada. Este estudo examina a experiência de ansiedade face ao desempenho musical em crianças. Cento e setenta e três crianças, do 3<sup>o</sup> ano ao 7<sup>o</sup> ano de escolaridade, responderam à forma “estado” do *State-Trait Anxiety Inventory for Children* (STAIC) (Inventário de Ansiedade Estado-Traço para Crianças) em dois momentos na escola: durante um dia normal e no dia de um concerto importante. A ansiedade-traço foi também medida na primeira ocasião. Os resultados mostraram que a ansiedade-estado foi significativamente maior no dia do concerto e que estava relacionada com o nível de ansiedade-traço das crianças. Não foram encontradas diferenças para o tipo de aula de música. Diferenças relativas ao género foram aparentes nos anos mais avançados.

#### Abstract

Performance anxiety is a common problem for musicians. While empirical research on adult musicians is growing, child musicians have largely been overlooked. This study examined the experience of performance anxiety in children. 173 children in grades three through seven completed the state form of the *State-Trait Anxiety Inventory for Children* at two points in time: during a regular school day and on the day of a major school concert. Trait anxiety was also measured on the first occasion. It was found that state anxiety was significantly higher on the day of the school concert and was related to children's level of trait anxiety. No differences were found for type of music class. Gender differences were apparent in the upper elementary grades.  
Key words: performance, anxiety, children, music, school.



# **An Investigation into the Experience of Musical Performance Anxiety in Elementary School Children**

**C. Ryan (2005)**

Musical performance anxiety has been identified as one of the most prevalent problems among adult musicians (Cooper, 1989; Dews & Williams, 1989; Fishbein, Middlestadt, Ottati, Straus & Ellis, 1988; Kemenade, Son, & Heesch, 1985; Kendrick, Craig, Lawson & Davidson, 1982; Lockwood, 1989; Wesner, Noyes & Davis, 1990). Described by Salmon (1990) as "the experience of distressful apprehension about and/or actual impairment of performance skills, in a public context, to a degree unwarranted given the individual's musical aptitude, training, and level of preparation," performance anxiety can negatively affect musicians opportunities for success and has been blamed for the premature termination of otherwise promising musical careers.

A growing body of literature has suggested various personal (e.g. perfectionism, external locus of control, low self-esteem, low stress hardiness, and auditory style) and environmental (e.g. audience size and stature, absence of the musical score, presence of peers) factors that seem to be related to a heightened experience of musical performance anxiety (Brodsky, Sloboda, and Waterman, 1994; Hutterer, 1980; LeBlanc, Jin, Obert, and Siivola, 1997; Leglar, 1979; Mor, Day, Flett, and Hewitt, 1985; Nagel, 1988; Salmon, Shook, Lombart, and Berenson, 1995; Tobacyk and Downs, 1986). However, most of the research literature has focused on treatments for reducing debilitating musical performance anxiety through either psychological or pharmacological approaches (c.f. Appel, 1976; Brantigan, Brantigan, and Joseph, 1982; Gates, Saegert, Wilson, Johnson, Shepherd, and Hearne, 1985; James, Burgoyne, and Savage, 1983; Kendrick, Craig, Lawson, and Davidson, 1982; Sweeney and Horan, 1982). Given the large number of adult musicians who suffer from performance anxiety and the clear research interest in effective anxiety reduction treatments, it would seem beneficial to examine this issue in younger musicians, for whom early intervention strategies might reduce the potential for developing debilitating anxiety as they mature.

While performance is a key component of music education, minimal attention has been paid to the experiences of children with regard to performance anxiety. In contrast, the related performance-evaluative domains of test anxiety and sports performance anxiety have long identified children as sufferers, and have even considered treatment approaches specifically for children (c.f. Fleege, Charlesworth, Burts & Hart, 1992; Hembree, 1988; Kass and Gish, 1991; Passer, 1983; Scanlan & Lewthwaite, 1984; Simon & Martens, 1979; Smith, Smoll & Barnett, 1995; Terry, Coakley & Karageorghis, 1995; Zatz & Chassin, 1985). In fact, one of the earliest studies pointing to the issue of anxiety in children with respect to musical performance comes from the sports psychology literature. Simon and Martens (1979) examined the anxiety responses of 749 9-14 year old boys participating in one of three types of activities: required school activities (academic test, physical education class competition), nonrequired school musical activities (solo instrumental performance and band competition), and nonrequired nonschool sports activities (individual and team). They found that, of 11 activities, by far the greatest anxiety was reported by



the boys performing a solo on a musical instrument; performing with a band was responsible for the highest anxiety among group activities.

With respect to solo musical performances, Ryan has conducted a series of studies on sixth-grade children performing in a piano recital. She found that performance anxiety was negatively related to self-esteem, that children's heart rates rose significantly through each stage of the recital (sitting stage-side, walking onstage, performing), and that over half of the children exceeded the 50<sup>th</sup> percentile of anxiety norms (Ryan, 1998). Additionally, she found that gender seems to play a key role in the anxiety experiences of these 12-year old musicians (Ryan, 2000, 2004). While girls heart rates followed a steady rise from baseline through each stage of the recital, boys heart rates remained relatively low – even when next-in-line to play – until they were actually on stage, at which time their heart rates spiked rapidly and exceeded the girls. The behavioral, physiological, performance quality, and self-report measures of girls tended to move synchronously in a predictable way according to anxiety level; however the boys measures displayed a lack of synchrony. Boys with high heart rates immediately prior to performing displayed few anxious behaviors and ultimately performed poorly. Boys who maintained lower heart rates prior to performing, on the contrary, displayed many anxious behaviors, but performed better. This lack of synchrony between the three systems of anxiety response has been noted throughout the literature; however previous research has generally not examined musicians by gender. This study seems to suggest that the desynchrony may, in fact, exist only for male musicians.

While the previous studies on children have examined performance anxiety in solo contexts, a few studies have also looked at children's and adolescents responses to ensemble experiences. Rothlisberger (1992) examined eighth grade students performance anxiety as they auditioned for honours band. He found that among high trait-anxious students, those whose school music programs were highly regarded tended to experience higher levels of performance anxiety than those whose music programs were less well regarded. This seems to suggest the possibility that expectation to perform well may play a role in students experience of performance anxiety.

Audience size and stature may also be related to increased anxiety in young musicians. LeBlanc et. al (1997) examined performance anxiety in high school instrumentalists performing under three audience conditions: a) alone in a practice room; 2) in a rehearsal room with one researcher; and 3) in a rehearsal room with several researchers and peers. They found that self-reported anxiety rose significantly for each higher gradation of audience size, but that heart rates rose significantly only under the condition with the largest audience presence. Gender differences were also noted – females had significantly higher heart rates than males in two out of three conditions and reported higher anxiety than males under the largest audience condition.

Elementary and junior high school musicians have also been shown to experience performance anxiety. Ryan (1999) examined state anxiety in 313 elementary and junior high band students. She found that the greatest number of high anxiety students (50%) were in sixth-grade. Between grades eight and nine gender differences were observed whereby eighth-grade girls reported higher anxiety than eighth-grade boys, but the reverse was true in ninth-grade. A relationship between gender, anxiety, adolescence, and social awareness was postulated.

Though few in number, these studies seem to indicate that performance anxiety is experienced by musicians as young as sixth-grade. While this is a common age for children to begin instruction on band instruments, individual music instruction often begins well before sixth grade. In fact, Nagel (1987) found that 90% of adult musicians began instruction before the age of 12; 46% before the age of seven. Additionally, children are generally expected to perform in a choral capacity in school musical events from the outset of elementary school, often during the preschool years. These performances are typically viewed as positive experiences for children – opportunities to show parents and teachers what they have learned and to express themselves publicly through music. However, the question remains as to whether these performances might cause undue anxiety for some children. The purpose of this study was to 1) examine the school concert experience of elementary school children with respect to performance anxiety, and 2) to determine whether children younger than sixth-grade experience musical performance anxiety.



## METHOD

### *Participants*

349 children ages 5-13 participated in this study. Students attended a private, co-educational Pre-kindergarten – Grade 12 school in a large urban center. The school was physically divided into a junior school (Pre-K-Grade 7) and a senior school (Grades 8-12). Music was mandatory for all students in the junior school, with general music classes from Pre-K to Grade 4 and the option of general music or band in grades 5-7. All students in grades 1-7 participated in two full-school concerts each year – one in December and one in May. Children also took part in grade-specific and ensemble-specific concerts at least once per year.

Homeroom teachers were asked to administer the state form of the State-Trait Anxiety Inventory for Children at two times – once during a regular school class in December, and again on the day of the December concert. The trait form of the inventory was also completed on the first occasion. Due to scheduling conflicts, some teachers had students complete the original forms, but did not return the concert-day forms. Only data from children who completed forms on both days was used in this study – 190 children fulfilled this requirement. Since the anxiety measure used in this study was not designed for children in the early primary grades, the data from grade one and two students was not included in the analysis. No grade two students completed all forms, but 12 grade one children did – their data was examined for exploratory purposes only. The resulting number of children in grades three through seven who were included in the analysis was 173, with the following number of boys and girls in each grade: grade 3 = 20 (9m/11f); grade 4 = 41 (27m/14f); grade 5 = 64 (45m/19f); grade 6 = 29 (16m/13f); grade 7 = 19 (10m/9f). There were 66 females and 107 males.

### *Measure*

The State-Trait Anxiety Inventory for Children (STAIC) measures two forms of anxiety – state and trait – in elementary school children, and has been used frequently in anxiety research. The A-state form measures situational anxiety and asks children to indicate how they feel in a specific situation. It contains 20 statements to which students must respond with one of three degree levels of a given anxious or nonanxious feeling (for example, “I feel: very calm, calm, not calm”). The A-trait form measures anxiety as a personality construct – a relatively stable measure of a child’s tendency to be anxious. It contains 20 complete statements addressing anxious or nonanxious feelings and requires students to note the frequency with which they experience that feeling (for example, “I feel like crying: hardly-ever, sometimes, often”). Its parent version, the State-Trait Anxiety Inventory, has been used extensively throughout the literature. The STAIC is a modified version of the STAI, with simplified language for use with children. Its internal consistency, test-retest reliability, and test validity are noted in the STAIC manual (Spielberger, 1973) and are considered satisfactory.

For this study, instructions on the first A-state form (the regular school day) were to respond to the questions according to how you feel ‘right now, at this very moment’. For the concert day, children were instructed to respond according to ‘how you feel when walking onstage at a school concert’.

### *Procedure*

Both A-state and A-trait forms were administered to students by their homeroom teacher during a regular school day in December – one week before the school concert. Teachers were given written instructions as well as verbal instructions during a staff meeting. Prior to completing the forms, students were asked to read along silently while the teacher read the instructions out



loud. On the day of the concert, teachers administered the second A-state form, with instructions specific to the performance situation. The concert was a formal event with all students, teachers, and their families in the audience. It took place in the school gym following normal school procedures. The program consisted of selections for band, choir, recorder ensemble, Orff ensemble, and massed choir.

Trait anxiety scores were transformed into anxiety levels (low, moderate, and high) according to the norms given for each age in the STAIC manual. The STAIC has been normed for children in grades four through six; however the instruction manual indicates that it can be used with children just above and below these grades. For this reason, it was decided to analyze only the data from children in grades three through seven, who could reasonably fit within these parameters. The norms for grade six were also applied also to grade seven students; the norms for grade four were applied to grade three students.

## RESULTS

A repeated measures analysis of variance with three independent variables (grade, gender, anxiety level) was calculated on the A-state scores for students in grades three through seven. Significant results were found for state anxiety ( $F=4.07$ ;  $df=1, 144$ ;  $p < 0.05$ ), indicating that children across all grades perceived greater anxiety on the day of the concert. As seen in table 1 the variability was also higher on the concert day, indicating a wider spread in children's responses. Significant results were also noted for anxiety level ( $F=9.2$ ;  $df=2, 144$ ;  $p < .001$ ). Table 2 shows that mean state anxiety scores across both testing times were lowest for children in the low trait anxiety group, moderate for the moderate trait anxiety group, and highest for the high trait anxiety group. No differences were noted for grade or gender.

Since students had the option to study a band instrument beginning in grade five, a repeated measures analysis of variance was used to compare the A-state scores of children in these grades, to determine whether participation in band, as compared with general music class, affected their anxiety response to the concert situation<sup>1</sup>. Gender, grade, and music class (band or general music) were the independent variables. No main effects were noted. An interaction between grade and gender, however, approached significance ( $p < 0.055$ ). As seen in table 3, grade five boys had higher anxiety than grade five girls at both testing times, and both genders increased in anxiety from school day to concert day. In grade six, boys reported a small decrease in anxiety on the concert day, but girls reported a moderate increase. In grade seven, boys reported a substantial increase in anxiety, while girls reported a small decrease.

## DISCUSSION

Research on musical performance anxiety has, until recently, focused almost exclusively on adult performers. While previous studies on the experience of musical performance anxiety in school-age musicians is limited, findings have consistently indicated that children as young as sixth-grade do experience it, and that their experiences differ according to age and gender. The purpose of the present study was to test for the existence of musical performance anxiety in children younger than sixth grade, and to do so in the context of a musical experience common to many children – that of the school concert.

The finding that state anxiety rose significantly from the regular school day to the day of the school concert, and did not interact with grade or gender, seems to indicate that children as young as third grade do experience an increased level of anxiety prior to a school concert. This is consistent with findings in the test anxiety literature which indicate that test anxiety is firmly established in children by the end of second grade. Unfortunately, the lack of an appropriate measure for examining performance anxiety in young children made it impossible to examine the responses of first and second grade students with any degree of reliability. The grade one students trait anxiety scores were generally quite high and no discernable pattern could be gleaned from their state responses to the two situations. Some children with very high trait scores indicated high



anxiety on the regular school day, but low anxiety on the concert day; the opposite was also true: students with low trait anxiety reported low state anxiety on day one, but high state anxiety on day two.

Given the results of this study and the literature on test anxiety, it is likely that a similar pattern of response might be found in second grade students as was found in the older children. However, it is also conceivable that performance anxiety may manifest younger than test anxiety, as children often begin participating in musical performances younger than they begin taking tests. The development of an appropriate measure for examining this issue in young musicians would be beneficial to the overall development of research in the area of performance anxiety. With a better understanding of the origins and early stages of performance anxiety, we may be able to develop teaching strategies that better equip young musicians with the tools to offset the potentially negative effects of performance anxiety.

Not surprisingly, performance anxiety across all grades was related to students anxiety-proneness, as measured by the A-trait scale of the STAIC. Children with higher trait anxiety tended to experience greater state anxiety on the day of the concert. This finding may provide valuable information for music teachers. Trait anxiety might potentially serve as an indicator of how individual students are likely to be affected by performance experiences. Such information could potentially allow teachers to tailor their teaching approaches and performance preparation to meet the needs of individual students. This finding also suggests that training in anxiety management to reduce overall stress and anxiety might be effective in reducing children's experience of situational anxiety – in particular performance anxiety. With more and more attention being paid to the highly structured and stressful lives of children, and research indicating that test anxiety, athletic performance anxiety, and now musical performance anxiety are common among elementary students, such training might be a well-timed intervention.

The finding that the type of music class attended did not seem to affect late elementary students experience of performance anxiety is not surprising, given that all students in this school were active participants in the music program and in school concerts. However, it should be noted that the proportion of children in this school who participated in band, particularly in grade five, far outweighed the proportion of children who did not. An examination of the means indicates that the children in grades five and seven who did not participate in band reported greater performance anxiety than those who did not. While this finding was not statistically significant, it is possible that some children elected not to participate in band (where performance is a more frequent occurrence than the bi-yearly concert) because they found performing to be a stressful event. Since unequal group sizes may have contributed to an inconclusive result on this issue, further research seems to be warranted.

As in previous studies, gender differences were apparent in the late elementary grades. These differences were not apparent in grades three and four. In grade five, children of both genders rose in anxiety from school day to concert day. It is somewhat surprising to find that in grade five boys perceived higher anxiety than girls both on the regular school day and on the concert day. Previous research on children and adults has indicated that females tend to report higher levels of anxiety than males (LeBlanc et. al, 1997; Nagel, 1988; Widmer et. al, 1997; Abel & Larkin, 1990). Given that the youngest musicians examined prior to this study were sixth-graders, this result invites further investigation into gender responses in students under the age of twelve.

Perhaps the most interesting findings with respect to gender pertain to students in grades six and seven. Grade six boys reported a small decrease in anxiety on the concert day, but grade seven boys reported a large increase. Grade six girls reported a moderate increase, but grade seven girls reported a small decrease. With the exception of grade seven girls, all other groups increased in variability from the school day to the concert day. The decreases in anxiety for grade six boys and grade seven girls are difficult to interpret – previous research has indicated that grade six and seven students do experience performance anxiety, in both solo and ensemble contexts (Ryan, 1998, 1999, 2000, 2004). It may be that a more sensitive measure is needed when examining performance anxiety in children and adolescents.

The rise in anxiety in grade six girls and grade seven boys is similar to findings in previous research indicating an earlier heightening of anxiety response by adolescent girls than by adolescent boys (Ryan, 1999). Gender differences with respect to musical performance anxiety have been found in three studies on child musicians (Ryan, 1999, 2000, 2004) and highlight the need to examine musicians according to gender, rather than as one homogeneous group, as has

been typically done in work with adult musicians. It seems that from at least sixth grade musicians may have different anxiety responses depending upon whether they are male or female. We may need to develop gender-specific strategies for training young musicians for the performance experience.

Results of this study need to be taken with caution. As a pilot study, children from only one school were sampled. It is possible that a more random sampling of children from many schools would produce different findings. As well, the unequal number of boys and girls and band versus general music students makes it difficult to draw conclusions about the findings pertaining to these groups. As a preliminary investigation, only one measure (psychological) was used. A more comprehensive study using behavioural and physiological measures, in addition to the psychological measure, will allow us to get a better picture of the issue. Future research on this topic will provide more information as to the role of gender and age with respect to children's experience of musical performance anxiety. Moreover, the effect of such anxiety on the quality of children's performances and their enjoyment of the performance experience should be investigated.



## REFERENCES

- Abel, J. L. and Larkin, K. T. (1990). Anticipation of performance among musicians: Physiological arousal, confidence, and state-anxiety. *Psychology of Music*, 18, 171-182.
- Appel, S. S. (1976). Anticipation of performance among musicians: Physiological arousal, confidence, and state-anxiety. *Psychology of Music* 171-182.
- Brantigan, C. O., Brantigan, T. A., and Joseph, N. (1982). Effect of beta blockade and beta stimulation on stage fright. *The American Journal of Medicine*, 72 (1), 88-94.
- Brodsky, W., Sloboda, J. A., and Waterman, M. G. (1994). An exploratory investigation into auditory style as a correlate and predictor of music performance anxiety. *Medical Problems of Performing Artists*, December, 101-112.
- Cooper, C. L. (1989). Popular musicians under pressure. *Psychology of Music*, 17, 22-36.
- Dews, C. L.B. and Williams, M. S. (1989). Student musicians' personality styles, stresses and coping patterns. *Psychology of Music*, 17, 37-47.
- Fishbein, M., Middlestadt, S. E., Ottati, V., Straus, S., and Ellis, A. (1988). Medical problems among ISCOM musicians: Overview of a national survey. *Medical Problems of Performing Artists*, 3, 1-8.
- Fleege, P. O., Charlesworth, R., Burts, D. C., and Hart, C. H. (1992). Stress begins in kindergarten: A look at behavior during standardized testing. *Journal of Research in Childhood Education*, 7 (1), 20-26.
- Gates, G. A., Saegert, J., Wilson, N., Johnson, L., Shepherd, A., and Hearne, E., M. (1985). Effect of beta blockade on singing performance. *Annals of Otolaryngology, Rhinology, and Laryngology*, 94 (1), 570-574.
- Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety. *Review of Educational Research*, 58 (1), 47-77.
- Hutterer, J. (1980). A structural analysis of the performance anxiety syndrome as experienced among solo musicians. *Dissertation Abstracts International*, 41, 354B.
- James, I. M., Burgoyne, W., and Savage, I. T. (1983). Effect of pindolol on stress-related disturbances of musical performance: Preliminary communication. *Journal of the Royal Society of Medicine*, 76 (2), 194-196.
- Kass, R. G. and Gish, J. M. (1991). Positive reframing and the test performance of test anxious children. *Psychology in the Schools*, 28, 43-52.
- Kemenade, J. F. L. M. van, and Heesch, N. C. A. van (1995). Performance anxiety among professional musicians in symphonic orchestras: A self-report study. *Psychological Reports*, 77, 555-562.
- Kendrick, M. J., Craig, K. D., Lawson, D. M., and Davidson, P. P. (1982). Cognitive and behavioral therapy for musical performance anxiety. *Journal of Consulting and Clinical Psychology*, 50 (3), 353-362.

- LeBlanc, A., Jin, Y. C., Obert, M., and Siivola, C. (1997). Effect of audience on music performance anxiety. *Journal of Research in Music Education*, 45, 480-496.
- Leglar, M. (1979). Measurement of indicators of anxiety levels under varying conditions of musician performance. *Dissertation Abstracts International*, 39A, 5201-5202.
- Lockwood, A. H. (1989). Medical problems of musicians. *New England Journal of Medicine*, 320, 221-227.
- Mor, S., Day, H. I., Flett, G. L., and Hewitt, P. L. (1995). Perfectionism, control, and components of performance anxiety in professional artists. *Cognitive Therapy and Research*, 19 (2), 207-225.
- Nagel, J. J. (1988). In pursuit of perfection: Career choice and performance anxiety in musicians. *Medical Problems of Performing Artists*, December, 141-145.
- Nagel, J. J. (1987). An examination of commitment to career in music: Implications for alienation from vocational choice. *Dissertation Abstracts International*, 42, 5-A, 1154-1155.
- Passer, M. W. (1983). Fear of failure, fear of evaluation, perceived competence, and self-esteem in competitive-trait-anxious children. *Journal of Sport Psychology*, 5, 172-188.
- Rothlisberger, D. J. (1992). Effects of video modeling preparation on student instrumental audition performance achievement and performance anxiety. *Dissertation Abstracts International*, 53, 2287A-2288A.
- Ryan, C., (2004). Gender differences in children's experience of musical performance anxiety. *Psychology of Music*, 32 (1), 89-103.
- Ryan, C. (2000). *A Study of the Differential Responses of Male and Female Children to Musical Performance Anxiety*. Doctoral Dissertation, McGill University.
- Ryan, C. (1999). Musical performance anxiety: Are junior high band students affected? *Opus* 41 (1), 8-9.
- Ryan, C. (1998). Exploring musical performance anxiety in children. *Medical Problems of Performing Artists*, 13, 83-88.
- Salmon, P., Shook, C. P., Lombart, K. G., and Berenson, G. (1995). Performance impairments, injuries, and stress hardiness in a sample of keyboard and other instrumentalists. *Medical Problems of Performing Artists*, December, 140-146.
- Salmon, P. (1990). A psychological perspective on musical performance anxiety: A review of the literature. *Medical Problems of Performing Artists*, 5, 2-11.
- Scanlan, T. K. and Lewthwaite, R. (1984). Social psychological aspects of competition for male youth sport participants: I. Predictors of competitive stress. *Journal of Sport Psychology*, 6, 208-226.
- Simon, J. A. and Martens, R. (1979). Children's anxiety in sport and nonsport evaluative activities. *Journal of Sport Psychology*, 1, 160-169.
- Smith, R. E., Smoll, F.L., and Barnett, N.P. (1995). Reduction of children's sport anxiety through social support and stress-reduction training for coaches. *Journal of Applied Developmental Psychology*, 16 (1), 125-142.
- Spielberger, C. D., Edwards, D. C., Lushene, R. E., Montuori, J., and Platzek, D. (1973). *State-Trait Anxiety Inventory for Children*. Palo Alto, CA: Consulting Psychologists Press, inc.



- Sweeney, G. A. and Horan, J. J. (1982). Separate and combined effects of cue-controlled relaxation and cognitive restructuring in the treatment of musical performance anxiety. *Journal of Counseling Psychology*, 29 (5), 486-497.
- Terry, P., Coakley, L., and Karageorghis, C. (1995). Effects of intervention upon precompetition state anxiety in elite junior tennis players: The relevance of the matching hypothesis. *Perceptual and Motor Skills*, 81, 287-296.
- Tobacyk, J. J., and Downs, A. (1986). Personal construct threat and irrational beliefs as cognitive predictors of increases in musical performance anxiety. *Journal of Personality and Social Psychology*, 51 (4), 779-782.
- Wesner, R. B., Noyes, R., and Davis, T. L. (1990). The occurrence of performance anxiety among musicians. *Journal of Affective Disorders*, 18, 177-185.
- Widmer, S., Conway, A., Cohen, S., and Davies, P. (1997). Hyperventilation: A correlate and predictor of debilitating performance anxiety in musicians. *Medical Problems of Performing Artists*, December, 97-106.
- Zatz, S. and Chassin, L. (1985). Cognitions of test-anxious children under naturalistic test-taking conditions. *Journal of Consulting and Clinical Psychology*, 53 (3), 393-401.

Table 1

Mean state anxiety scores and standard deviations for children in each grade during a regular school day and on the day of the school concert

|                | 3             | 4             | 5             | 6             | 7             |
|----------------|---------------|---------------|---------------|---------------|---------------|
| <b>School</b>  | 29.9<br>(4.1) | 29.3<br>(6.4) | 29.6<br>(6)   | 28.6<br>(5.2) | 30<br>(7.8)   |
| <b>Concert</b> | 32.6<br>(8.9) | 30.6<br>(6.6) | 31.7<br>(8.5) | 29<br>(6.2)   | 32.5<br>(7.7) |

Table 2

Mean state anxiety scores for low, moderate, and high trait anxious students during a regular school day and on the day of the school concert

|                         | <b>School</b> | <b>Concert</b> |
|-------------------------|---------------|----------------|
| <b>Low anxious</b>      | 27.4<br>(3.8) | 28.4<br>(5.3)  |
| <b>Moderate anxious</b> | 30.8<br>(6.6) | 33.2<br>(8.7)  |
| <b>High anxious</b>     | 31.7<br>(7.2) | 33.8<br>(8.7)  |

Table 3

Mean STAIC (state) scores for students in grades five through seven.

| Grade | Male         |               | Female       |               |
|-------|--------------|---------------|--------------|---------------|
|       | STAIC school | STAIC concert | STAIC school | STAIC concert |
| 5     | 30 (6.1)     | 32.3 (8.8)    | 28.9 (6)     | 30.4 (7.9)    |
| 6     | 27.6 (3.3)   | 27.1 (4)      | 29.9 (6.9)   | 31.2 (7.6)    |
| 7     | 28 (6)       | 33 (10)       | 32.3 (9.3)   | 31.7 (3.9)    |



## 2.10.3. Estudo Crítico – Por João Nogueira

### Introdução

Na introdução são referidos vários estudos relacionados com o tema em estudo: a ansiedade relativa ao desempenho musical. Define-se o conceito e os factores pessoais e ambientais que com ele se relacionam. Refere-se ainda que a maioria dos estudos se centra nos tratamentos psicológicos ou farmacológicos para a redução dessa ansiedade.

Dada a prevalência do problema nos adultos e o interesse na sua redução, afirma-se que poderá ser benéfica uma intervenção precoce para evitar o desenvolvimento da ansiedade debilitadora em adulto. Contrariamente ao domínio da ansiedade em testes e no desporto, as crianças não têm sido estudadas relativamente à ansiedade em provas musicais.

O artigo dá um exemplo concreto de uma investigação que coloca a execução musical a solo e a execução musical com banda como as actividades mais ansiogéneas de entre uma lista de várias actividades consideradas ansiogénicas (referência de 1979). Depois revê estudos mais recentes (de 1998 a 2004) sobre a ansiedade musical relacionando-a negativamente com a auto-estima, e refere como os aspectos fisiológicos (ritmo cardíaco, p. ex.) são moderados pelo género. Neste caso, a sincronia entre os três sistemas da ansiedade, comportamental, fisiológico e verbal, não existe nos rapazes.

De seguida, refere estudos que sugerem um papel das expectativas para o desempenho, para o tamanho e o tipo de audiência e a adolescência.

Habitualmente, no 1º ciclo de escolaridade cantar é uma experiência positiva, mas para algumas crianças poderá causar bastante ansiedade. O propósito deste estudo é saber se as crianças sentem ansiedade no desempenho musical e se esta ansiedade existe antes do 6º ano de escolaridade.

Nesta revisão bem elaborada da literatura sobre ansiedade no desempenho musical, a autora assume que essa ansiedade é prejudicial. Aqui parece haver alguma confusão relativamente ao papel motivacional da ansiedade. A ansiedade é essencial para o empenhamento e reflecte a importância dos acontecimentos. Logo, se os recitais são importantes, *devem* ser precedidos de ansiedade antecipatória. É esta que motiva para uma maior concentração, compenetração e aplicação, que resultará num melhor desempenho.

Os padrões de ansiedade (e de outras emoções) ao longo da preparação e antes, durante e depois das provas desportivas têm uma enorme variabilidade individual (Hanin, 2003). Uma despreocupação, talvez por excesso de confiança, antes das provas não favorece a preparação (para quê estudar ou treinar se já sei ou se a prova não é importante?) e não antecipa (preparando o corpo e a mente) o que sucederá nas provas. É a ansiedade *durante* a prova a que pode interferir com o desempenho. Ou seja, de acordo com a chamada lei de Dobson-Yerkes, a relação entre a ansiedade (stress) e o desempenho é curvilínea (em U invertido). Isto é, são os níveis moderados de ansiedade que estão associados aos melhores níveis de desempenho. Demasiado pouca ansiedade não chega para motivar um bom desempenho. O aumento da ansiedade traduz-se num melhor desempenho até um nível óptimo, a partir do qual mais ansiedade começa a perturbar o desempenho. É a esta ansiedade excessiva que deveria referir-se este estudo. Claro que demasiada ansiedade antecipatória poderá também interferir com a preparação.

### Método

### Participantes

Trata-se de uma amostra de conveniência. Dos trezentos e quarenta e nove (349) alunos inicialmente inquiridos, apenas cento e noventa (190) completaram a 2ª avaliação e destes,

dezassete (17) eram de anos não abrangidos pelas normas do instrumento utilizado. Nestas condições, a distribuição dos alunos por ano de escolaridade e de género não foi igual para todos os grupos. O estudo decorreu numa escola apenas.

## **Instrumentos**

A medida utilizada – STAIC – é uma medida válida e garantida, para os anos analisados. Isto é, que mede o que é suposto medir, a ansiedade (validade), e que o faz de uma forma consistente e fiável (garantia). Apesar de o manual da prova dever referir estas características psicométricas, em cada estudo é necessário analisar como se comporta o instrumento, para podermos afirmar que os resultados são válidos e garantidos. Esta escala não dá, no entanto, qualquer indicação sobre a qualidade da ansiedade (perturbadora ou facilitadora).

Eventualmente, o facto de se apresentarem instruções diferentes para o mesmo instrumento no dia de escola e no dia do concerto poderia ter sido incluído na rubrica do procedimento.

## **Procedimento**

Os questionários foram aplicados pelos professores, após um encontro com os investigadores. A transformação dos resultados em níveis de ansiedade deveria ser referida na secção seguinte, já que diz respeito à análise de dados. Nesta secção deve ser descrito o que foi feito durante a recolha de dados. Todas as operações posteriores que não interfiram com essa recolha (por exemplo, dividir os participantes em níveis para lhes atribuir uma tarefa influencia a recolha de dados) devem ser colocados na secção seguinte.

Temos um *design* longitudinal com medidas repetidas em que se compara a ansiedade nas duas ocasiões, no dia normal de escola e no dia do concerto, em função do ano de escolaridade, do género e do nível de ansiedade-traço. No entanto, o estudo é exploratório, e, por isso, não há preocupação em garantir a igualdade de sujeitos nos vários grupos. Num procedimento experimental teríamos de ter grupos iguais, por exemplo, sessenta (60) de cada ano, sendo estes metade de cada género e nestes trinta (30), dez (10) de cada nível de ansiedade.

## **Resultados**

Ao comparar os mesmos sujeitos em duas ocasiões diferentes, importa considerar esse facto utilizando um procedimento com medidas repetidas. A análise de variância (ANOVA) utiliza-se para comparar as médias de três ou mais grupos. Este procedimento é uma extensão do teste *t* de Student (pseudónimo do seu inventor, que trabalhava para a fábrica Guinness, com a qual tinha um contrato de exclusividade) que compara as médias entre dois grupos. Estes podem ser independentes (cada observação é independente das outras, isto é, sujeitos diferentes nos dois grupos) ou relacionados (cada sujeito é observado duas vezes e comparam-se as duas ocasiões). O teste estatístico é significativo, ou seja, há diferenças significativas entre a ansiedade nos dois momentos, embora sem efeito do ano nem do género. Assim, há mais ansiedade nos dias de concerto do que nos dias normais de escola, independentemente do ano de escolaridade e do género.

Mas, se examinarmos o quadro 1, no 6º ano a diferença parece mínima (0,4). Ou seja, talvez o resultado global não se verifique em todos os anos. Também quando se encontra o efeito do nível de ansiedade-traço e se afirma que a diferença é baixa nos menos ansiosos, moderada nos moderadamente ansiosos e elevada nos muito ansiosos, não parece ser tido em conta que a diferença entre os momentos é maior nos moderadamente ansiosos. Além disso, os níveis atingidos não parecem muito diferentes entre os níveis mais elevados (33,2 e 33,8 no quadro 2).



Neste sentido, seria útil comparar estes 2 resultados com um teste t para amostras relacionadas ou emparelhadas (comparar a média dos mesmos sujeitos nas duas ocasiões). Estranha-se a utilização da ANOVA num estudo correlacional (apenas são observadas as variáveis, não havendo qualquer controlo ou manipulação) e de não haver nenhum coeficiente de correlação. Seria de esperar uma correlação positiva entre a ansiedade-traço e a diferença entre a ansiedade-estado nas duas ocasiões para afirmar a relação.

Considerando o tipo de classe de música em vez da ansiedade-traço, não se encontram efeitos principais. No entanto, o artigo fala de que uma interacção quase atinge a significância ( $p < 0,055$ ).  $P$  é o nível de risco que se admite para errar ao aceitar uma relação ou uma diferença inexistente. Ou seja, que o resultado obtido não é uma flutuação aleatória dentro do conjunto de todos os resultados possíveis, mas reflecte mesmo uma relação ou diferença. Quando se usa  $p = 0,05$  quer dizer que, em cem (100) estudos idênticos apenas cinco (5) obteriam esse resultado como uma variação normal. Isto quer dizer que dificilmente se poderá atribuir à sorte esse resultado, logo, o efeito é estatisticamente real. Um risco máximo de 0,05 é o que se convencionou para as ciências sociais. Noutras áreas esse risco terá de ser muito mais reduzido. Não podemos voar com uma probabilidade de 5% de queda do aparelho...

A significância estatística depende grandemente do número de participantes. Daí a necessidade de um número razoavelmente amplo para que se diluam os possíveis enviesamentos da amostra. Este número varia em função dos procedimentos, mas oscila entre o mínimo de cem (100) para a estatística paramétrica (que exige a normalidade da distribuição) e de trinta (30) para a estatística não-paramétrica. Quando existem várias dimensões para comparar, no mínimo, deverão ter cinco sujeitos em cada grupo ou por item de um questionário. Para aumentar o risco e não deixar escapar uma relação apenas por motivos estatísticos (não atinge a significância porque há poucos sujeitos, mas o efeito é real) será necessário haver salvaguardas ao nível do desenho experimental. Se o *design* fosse experimental talvez se justificasse explorar essa proximidade de 0,05, mas neste caso as condições de utilização do teste já foram em muito levadas ao limite (exigência de células com efectivos iguais).

## Discussão

Na discussão, a autora começa por afirmar que a ansiedade-estado aumenta nos dias de concerto sem interagir com o género ou com o ano de escolaridade. Isto para crianças a partir do 3º ano, o que é consistente com a literatura que diz que a partir do fim do 2ºano já existe esse aumento de ansiedade.

Lamentando a inexistência de instrumentos adequados a crianças mais novas, refere os resultados inconsistentes do STAIC no 1º e no 2º ano (estes dados não são descritos nos resultados, por isso não podemos perceber a sua inconsistência). Especula, em seguida, acerca da ansiedade no desempenho musical ser anterior à ansiedade aos testes, por estes começarem mais tarde na vida da criança, e acerca dos benefícios do desenvolvimento de uma medida apropriada para os mais novos. Conhecendo melhor as origens da ansiedade poderão ser desenvolvidas técnicas de ensino que equipem melhor as crianças para lidar com os potenciais efeitos negativos dessa ansiedade.

Chama ainda a atenção para o facto de que os mais ansiosos (em termos de traço) tenderão a ser mais afectados, mas talvez a informação seja a de que os menos ansiosos são menos afectados, sendo os outros igualmente afectados. Isto justifica uma atenção talvez a todos, não com programas de intervenção, mas com treino de gestão da ansiedade, de modo a prevenir os problemas nas vidas altamente estruturadas e exigentes de muitas crianças.

Relativamente ao tipo de aula (participação ou não na banda), os dados não são fornecidos na rubrica dos resultados e o erro na frase: -“os que não participam são mais ansiosos do que os que não participam”- não pode ser corrigido. Além disso, a interpretação apresentada é abusiva porque é dito que os resultados não são estatisticamente significativos. O que deveria ser concluído é que não há diferenças entre os tipos de aula.

Os dados relativos ao género interpretam o que parece e ignoram a falta de significância estatística. Talvez para essas comparações fosse importante utilizar testes t para dar algum suporte estatístico às aparências. A argumentação aqui segue mais os estudos anteriores do que

os resultados desta investigação. O que deveria ser concluído é que não há diferenças entre os meninos e as meninas.

A nota final de chamada de atenção para que estes resultados, ou melhor, estas interpretações, sejam lidos com cuidado mostra a consciência da autora acerca de todos os problemas com esta investigação. Na última frase, a autora assinala a falta do que poderia tornar importante este estudo: a determinação do impacto da ansiedade no desempenho musical.

## Referências

Esta secção é bastante completa e é de assinalar apenas a necessidade de corrigir a repetição do título das duas primeiras referências.

## Apreciação geral

Em suma, esta é uma investigação pouco preparada em termos teóricos, no sentido em que procura apenas “ver” se existe mais ansiedade nos dias de concerto. A questão de saber se esse aumento é importante, isto é, se interfere com o desempenho, motivando ou perturbando, é descurada. O controlo das condições de aplicação foi deixado aos professores das crianças e não houve uma selecção que permitisse uma igualdade nos grupos a comparar. Para que serve isto? Não é clara a finalidade deste estudo, para além de cumprir algum requisito curricular para a carreira ou para um concurso.

O principal problema na análise de dados reside na interpretação de resultados estatisticamente não significativos, especulando com os dados da literatura. Para um estudo exploratório sem o controlo experimental exigido pelo método dos inquéritos para estabelecimento de normas, e não havendo igualdade numérica entre os grupos comparados, seria preferível a recolha de mais informação acerca de cada criança. Nomeadamente, percebendo o papel motivador ou perturbador da ansiedade no concerto.

O artigo tem o mérito de reconhecer na conclusão as suas fragilidades.

## Referência bibliográfica

Hanin, Y. L. (2003). Performance Related Emotional States in Sport: A Qualitative Analysis. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* [On-line Journal], 4(1). Disponível em <http://www.qualitative-research.net/fqs-texte/1-03/1-03hanin-e.htm> [Acedido a 3 de Outubro de 2005].





### 3ª PARTE

## Fontes de pesquisa e informação bibliográfica para a realização de estudos de investigação no âmbito da Psicologia da Música

Por Helena Rodrigues e Tiago Paulos Veiga

Neste capítulo apresentam-se algumas sugestões de fontes de pesquisa e informação bibliográfica relevantes para a realização de estudos de investigação nas áreas da Psicologia da Música e da Educação Musical. A informação aqui disponível não pretende ser exaustiva.

Efectivamente, seria desejável o acesso a informação representativa de outras perspectivas, para além da anglófona. Por exemplo, os países da ex-União Soviética têm concedido, nos seus planos de formação de professores, grande atenção à Pedagogia e à Psicologia da Música mas o seu trabalho ainda não se repercutiu ou foi divulgado nas publicações correntemente disponibilizadas no Ocidente.

De qualquer modo, ainda que limitado, este capítulo poderá ajudar o jovem pesquisador a abrir perspectivas e a organizar o seu trabalho. Os recursos aqui mencionados deverão ser encarados apenas como um ponto de partida para a construção de uma base de recursos informativos de acordo com temas de trabalho específicos.

Assim, este capítulo está organizado segundo quatro tipos de sugestões de consulta: Fontes Fundamentais, Fontes Complementares, Bases de Dados e Outros.

Em Fontes Fundamentais estão indicadas publicações específicas das áreas da Psicologia da Música e da Educação Musical ou onde podem ser encontrados artigos especializados dessas áreas. Note-se que, de uma forma geral, os principais periódicos mencionados nesta secção estão associados a organizações que o jovem investigador deverá igualmente conhecer.

Em Fontes Complementares são mencionadas algumas revistas especializadas de elevado prestígio cujo âmbito de publicação é muito abrangente e onde foram encontrados vários artigos das áreas em questão. Se se tiver em conta que, à data de publicação deste livro, o *International Index to Music Periodicals* e o *Social Sciences Citation Index* referem a existência de, respectivamente, centenas de periódicos na área da Música e de milhares na área das Ciências Sociais, é evidente que a selecção aqui apresentada é muitíssimo limitada. De qualquer modo, pareceu-nos útil a sua inclusão pois pode orientar o jovem investigador a explorar fontes e pistas de trabalho em direcções concretas.

Por outro lado, pesquisas por tema e autor e por periódicos das áreas da Psicologia e da Educação, próximas do objecto de trabalho desta publicação, podem abrir o leque das publicações cuja consulta pode ser relevante. As bases de dados referidas no ponto 3 deste capítulo auxiliarão neste tipo de pesquisas.

Ou seja, as "fontes complementares" apresentadas são uma mera pista de trabalho, devendo ser muito mais alargada em função da área de pesquisa do investigador. Por exemplo, se o investigador decidir desenvolver um trabalho relativo à influência da utilização de drogas na performance musical, deverá – naturalmente – alargar a sua pesquisa bibliográfica a revistas do âmbito da Medicina, da Farmacologia, da Adição, da Sociologia, etc. Se decidir desenvolver um trabalho relativo ao estudo do desenvolvimento musical na infância, então a pesquisa por revistas da área da Psicologia, da Pediatria e da Educação será igualmente obrigatória. Etc.

A seguir, as Bases de Dados apresentadas podem ser um recurso precioso para quem quiser usufruir em pleno das potencialidades de pesquisa que a informática oferece.

Finalmente, em Outros, mencionam-se algumas informações que podem ser de grande ajuda para o jovem investigador. A selecção de sites aí incluída é quase arbitrária, mas optou-se



por a incluir devido ao interesse dos respectivos sites e na certeza de que podem constituir um bom ponto de partida para a navegação na net. Relativamente a isto, recorda-se que por vezes os sites indicados podem não estar a permitir o acesso directo. Uma forma de contornar o problema pode ser inserir directamente o nome da organização ou periódico procurado no Google.

### 3.1. Fontes Fundamentais

#### ***British Journal of Music Education***

([www.journals.cambridge.org/action/displayJournal?jid=BME](http://www.journals.cambridge.org/action/displayJournal?jid=BME))

O *British Journal of Music Education* pertence à Cambridge University Press, responsável pela publicação de várias dezenas de outros jornais, nas mais variadas áreas de estudo. No site acima referenciado é disponibilizado o índice dos volumes publicados pelo *BJME*. Os resumos estão disponíveis gratuitamente, e o preço de compra do texto integral dos artigos está assinalado.

#### ***British Journal of Music Therapy***

([www.bsmt.org/journal.htm](http://www.bsmt.org/journal.htm))

O *British Journal of Music Therapy* – *BJMT* pertence à British Society of Music Therapy – BSMT. No endereço [www.bsmt.org/vol%201%20nos%201-2.htm](http://www.bsmt.org/vol%201%20nos%201-2.htm) temos o índice dos volumes publicados pelo *BJMT*, e os resumos de todos os artigos são disponibilizados gratuitamente.

#### ***Bulletin of the Council for Research in Music Education***

([www.crme.uiuc.edu/index.php](http://www.crme.uiuc.edu/index.php))

O *Bulletin of the Council for Research in Music Education* fornece um fórum onde a pesquisa contemporânea está acessível a todos os que tenham interesse sobre o universo da Educação Musical. O *Bulletin* contém pesquisas recentes e revisões de anteriores pesquisas. Em [www.crme.uiuc.edu/abstract.php](http://www.crme.uiuc.edu/abstract.php) temos a possibilidade de procurar artigos através de palavras-chave, tendo acesso gratuito aos respectivos abstracts.

#### ***Contributions To Music Education***

([www.cme.webhop.org/](http://www.cme.webhop.org/))

O *Contributions To Music Education* pertence à Ohio Music Education Association. A partir da página inicial da revista temos acesso ao índice de todos os volumes já publicados, sendo possível aceder aos títulos de todos os artigos publicados entre 1973 e 1988. A partir de 1989 (Volume 16) são também disponibilizados os abstracts de todos os artigos.

#### ***International Journal of Music Education***

(<http://publications.isme.org/virtuemart/34.html>)

O *International Journal of Music Education* é o jornal oficial da International Society of Music Education – ISME. No endereço <http://ijm.sagepub.com/archive/> temos o índice dos volumes publicados pelo *International Journal of Music Education*. Alguns resumos são disponibilizados gratuitamente, sendo possível o acesso pago a artigos específicos mediante o pagamento de uma verba devidamente assinalada.

### ***Journal of Music Therapy***

([www.musictherapy.org/products/pubs.html](http://www.musictherapy.org/products/pubs.html))

O *Journal of Music Therapy* pertence à American Music Therapy Association. No endereço <http://apt.allenpress.com/perlserv/?request=get-archive&issn=0022-2917&ct=1> temos o índice dos volumes publicados, com abstracts gratuitos de todos os artigos.

### ***Music Education Research***

([www.tandf.co.uk/journals/titles/14613808.asp](http://www.tandf.co.uk/journals/titles/14613808.asp))

O *Music Education Research* é um jornal internacional que beneficia da contribuição de uma larga comunidade de pesquisadores em todas as áreas da educação musical.

No endereço [www.ingentaconnect.com/content/routledg/cmue](http://www.ingentaconnect.com/content/routledg/cmue) temos o índice dos artigos publicados, e é disponibilizado o acesso gratuito aos resumos dos artigos desta publicação.

### ***Music Perception***

([www.ucpressjournals.com/journal.asp?iissn=0730-7829](http://www.ucpressjournals.com/journal.asp?iissn=0730-7829))

A revista *Music Perception* pertence à Society of Music Perception and Cognition, dedicada ao fomento dos conhecimentos na área da percepção e cognição musical. A partir do site da revista conseguimos ter acesso ao índice dos volumes publicados pelo *Music Perception*. Os resumos estão disponíveis gratuitamente.

### ***Music Scientiae***

(<http://musicweb.hmt-hannover.de/escom/english/MusicScE/MSstart.htm>)

O *Music Scientiae* pertence à European Society for the Cognitive Sciences of Music – ESCOM, dedicada a estudos nas áreas da psicologia, pedagogia, e filosofia da música. A partir do site da revista conseguimos ter acesso ao índice dos volumes publicados pelo *Music Scientiae*. Os resumos estão disponíveis gratuitamente.

### ***Nordic Journal of Music Therapy***

([www.njmt.no/](http://www.njmt.no/))

O *Nordic Journal of Music Therapy* publica artigos sobre a musicoterapia praticada nos países nórdicos (Dinamarca, Finlândia, Islândia, Noruega, Suécia, Países Bálticos, Estônia, Lituânia).

A língua inglesa é a principal língua deste jornal. No endereço [www.njmt.no/contentslist.html](http://www.njmt.no/contentslist.html) temos o índice dos volumes publicados pelo *Nordic Journal of Music Therapy*. Alguns resumos e textos integrais são disponibilizados gratuitamente.

### ***Performance Online***

([www.performanceonline.org](http://www.performanceonline.org))

Revista de interpretação musical. Em português e inglês.



## **Periódicos publicados pela National Association for Music Education – MENC**

([www.menc.org/publication/articles/journals.html](http://www.menc.org/publication/articles/journals.html))

A organização americana intitulada National Association for Music Education publica seis periódicos relacionados com a Educação Musical.

Estes periódicos estão indicados abaixo bem como os respectivos motores de busca. O último (JRME) é o mais especializado em termos de investigação.

a) *General Music Today – GMT* ([www.menc.org/publication/articles/jguindsrch.htm](http://www.menc.org/publication/articles/jguindsrch.htm))

b) *Journal of Music Teacher Education – JMTE* (idem)

c) *Update: Applications of research in Music Education* (idem)

d) *Teaching Music – TM* ([www.menc.org/publication/articles/tmindex.html](http://www.menc.org/publication/articles/tmindex.html))

e) *Music Educators Journal* ([www.menc.org/publication/articles/mejindex.html](http://www.menc.org/publication/articles/mejindex.html))

f) *Journal of Research in Music Education* ([www.menc.org/publication/articles/jrmeind.html](http://www.menc.org/publication/articles/jrmeind.html))

O acesso aos resumos dos artigos está disponível mediante a subscrição dos respectivos jornais.

## ***Psychology of Music***

(<http://www.sagepub.co.uk/journalsProdManSub.nav?prodId=Journal201640>)

O *Psychology of Music* é publicado pela Society for Education, Music and Psychology Research – SEMPRES. O *Psychology of Music* é oferecido gratuitamente aos membros da SEMPRES.

No endereço <http://www.sagepub.co.uk/journalsProdDesc.nav?prodId=Journal201640> é possível a subscrição de apenas um número do jornal.

No endereço <http://pom.sagepub.com/archive/> temos um arquivo onde os resumos dos artigos publicados estão disponíveis gratuitamente. O preço da subscrição dos artigos está devidamente assinalado.

## ***Philosophy of Music Education Review***

([http://muse.jhu.edu/journals/philosophy\\_of\\_music\\_education\\_review/](http://muse.jhu.edu/journals/philosophy_of_music_education_review/))

O *Philosophy of Music Education Review* procura disseminar o trabalho de pesquisa no âmbito da filosofia da música e da educação musical. No endereço [http://muse.jhu.edu/demo/philosophy\\_of\\_music\\_education\\_review/v011/11.1index.pdf](http://muse.jhu.edu/demo/philosophy_of_music_education_review/v011/11.1index.pdf) temos o index dos artigos publicados. Na página do jornal é disponibilizado o acesso gratuito a alguns resumos.

## ***Research and Issues in Music Education***

([www.stthomas.edu/rimeonline/vol1/purpose.htm](http://www.stthomas.edu/rimeonline/vol1/purpose.htm))

Na página desta publicação temos acesso gratuito aos abstracts dos artigos já publicados.

## ***Revista da Associação Brasileira de Educação Musical***

A Associação Brasileira de Educação Musical ([www.abem.clic3.net/](http://www.abem.clic3.net/)), fundada em 1991, disponibiliza o índice de todos os volumes da sua revista. A partir do quinto volume, é facultado o acesso aos resumos dos artigos.

## ***Revista da Associação Portuguesa de Educação Musical***

(<http://www.apem.jazznet.pt/page14/page14.html>)

A *Revista da Associação Portuguesa de Educação Musical – APEM* (designada no passado por Boletim da APEM) é uma publicação desta organização, que é o organismo representante da International Society of Music Education (ISME) em Portugal. É

distribuída gratuitamente aos respectivos sócios; os números antigos podem ser adquiridos na Sede da APEM.

### **Revista Música, Psicologia e Educação**

([www.cipem.wordpress.com/revista/](http://www.cipem.wordpress.com/revista/))

A Revista Música, Psicologia e Educação é uma publicação do CIPEM – Centro de Investigação em Psicologia da Música e Educação Musical, pertencente ao Departamento de Artes e Motricidade Humana da Escola Superior de Educação do Instituto Politécnico do Porto. Na página acima referida, são disponibilizadas as edições dos anos entre 1999 e 2004, estando disponível o texto integral dos artigos, ou, pelo menos, uma versão PDF dos respectivos resumos.

## **3.2. Fontes Complementares**

### ***Annals of the New York Academy of Sciences***

([www.nyas.org/annals/browse.asp](http://www.nyas.org/annals/browse.asp))

O *Annals of the New York Academy of Sciences* é a colecção de publicações da New York Academy of Sciences (NYAS).

Na página acima referida é disponibilizado o catálogo de todas as revistas publicadas pela NYAS. A partir deste site é possível fazer-se a pesquisa por palavra-chave das revistas que poderão interessar a uma determinada pesquisa.

### ***Australian Association for Research in Music Education***

([www.deakin.edu.au/education/music-ed/aarme/](http://www.deakin.edu.au/education/music-ed/aarme/))

No endereço [www.deakin.edu.au/education/music-ed/aarme/doc/Indexed\\_Abstracts.pdf](http://www.deakin.edu.au/education/music-ed/aarme/doc/Indexed_Abstracts.pdf), a Australian Association for Research in Music Education oferece-nos acesso a um index e aos abstracts dos artigos publicados nas conferências organizadas por esta associação.

Na página inicial são ainda disponibilizadas ligações para outros sites australianos de educação musical.

### ***Brain Research***

([www.elsevier.com/wps/find/journaldescription.cws\\_home/506048/description#description](http://www.elsevier.com/wps/find/journaldescription.cws_home/506048/description#description))

A *Brain Research* é uma revista internacional multidisciplinar dedicada à investigação no domínio das ciências do cérebro, nos quais se incluem estudos no âmbito da música e psicologia.

Em [www.sciencedirect.com/science/journal/00068993](http://www.sciencedirect.com/science/journal/00068993) é-nos apresentado o index a partir do qual temos acesso gratuito aos abstracts de todos os artigos. É possível fazer-se a pesquisa rápida de artigos pelo título, palavras-chave ou autor.

### ***Child: Care, Health and Development***

([www.blackwellpublishing.com/journal.asp?ref=0305-1862&site=1](http://www.blackwellpublishing.com/journal.asp?ref=0305-1862&site=1))

A *Child: Care, Health and Development* publica artigos sobre todos os aspectos da saúde e desenvolvimento de crianças: desenvolvimento emocional, social, físico e psicomotor; factores sociais e ambientais perturbadores do desenvolvimento e saúde da criança. São também feitas avaliações de intervenções específicas em crianças, e analisadas medidas políticas e organizacionais do âmbito educativo. Em [www.blackwell-synergy.com/loi/cch](http://www.blackwell-synergy.com/loi/cch) temos o índice de todos os volumes já publicados. Os abstracts estão disponíveis gratuitamente. Os textos integrais requerem subscrição.



### ***Finnish Music Quarterly***

([www.fmq.fi/index.html](http://www.fmq.fi/index.html))

O *Finnish Music Quarterly* divulga a cultura musical finlandesa desde 1985, abordando ainda tópicos do domínio da educação musical. Em [www.fmq.fi/articles.html](http://www.fmq.fi/articles.html) temos o índice de alguns artigos publicados.

### ***International Journal of Community Music***

([www.intljcm.com/](http://www.intljcm.com/))

O *International Journal of Community Music* é publicação académica gratuita que publica artigos de pesquisa, debates, resenhas e apontamentos concernentes à Música na Comunidade.

Em [www.intljcm.com/current.html](http://www.intljcm.com/current.html) temos acesso aos textos integrais do último volume desta publicação, e em [http://www.intljcm.com/past\\_issue.html](http://www.intljcm.com/past_issue.html) podemos consultar os textos integrais dos volumes anteriores.

### ***International Journal of Research in Choral Singing***

O *International Journal of Research in Choral Singing* é uma publicação do Consortium for Choral Research, um grupo internacional de pesquisa dedicado a investigadores que sejam também músicos coralistas. No site da organização ([www.choralresearch.org/](http://www.choralresearch.org/)) temos acesso gratuito aos textos integrais dos volumes já publicados.

### ***Journal of Band Research***

([www.americanbandmasters.org/jbr/](http://www.americanbandmasters.org/jbr/))

O *Journal of Band Research* pertence à The American Bandmasters Association.

Na página do jornal temos acesso gratuito (em formato PDF ou Word) à edição de 2007 desta publicação.

### ***Journal of Cognitive Neuroscience***

(<http://jocn.mitpress.org/>)

Este jornal pertence à The MIT Press ([www.mitpress.mit.edu/main/home/default.asp](http://www.mitpress.mit.edu/main/home/default.asp)).

Em <http://jocn.mitpress.org/contents-by-date.0.shtml> podemos consultar o índice dos volumes já publicados, tendo acesso gratuito aos abstracts e acesso ao texto integral mediante subscrição.

Em <http://jocn.mitpress.org/search.dtl> é possível fazer-se uma pesquisa rápida (por palavras-chave) ou pesquisa avançada (autor, data, número da revista, etc).

### ***Journal of Historical Research in Music Education***

([www.utc.edu/Faculty/William-Lee/Journal.html](http://www.utc.edu/Faculty/William-Lee/Journal.html))

O *Journal of Historical Research in Music Education* – JHRME, suportado pelo Music Department of the University of Tennessee at Chattanooga ([www.utc.edu/Faculty/William-Lee/index.html](http://www.utc.edu/Faculty/William-Lee/index.html)), publica artigos e revisões sobre todos os assuntos concernentes à História da Educação Musical.

Apesar de na página do JHRME não termos acesso ao índice ou a quaisquer abstracts dos artigos já publicados, a página do Music Department of the University of Tennessee at Chattanooga revela-se muito útil por disponibilizar informações importantes acerca desta área de estudo: links para outros sites acerca da História da Educação Musical; referências a outros jornais; bibliografia variada sobre História da Educação Musical; referências a arquivos e museus importantes.

### ***Journal of Technology in Music Learning***

([www.auburn.edu/outreach/dl/ctmu/jtml/page2.html](http://www.auburn.edu/outreach/dl/ctmu/jtml/page2.html))

O *Journal of Technology in Music Learning*, fundado em 1999, dedica-se à publicação de todo o tipo de artigos no domínio da investigação acerca da aplicação de tecnologia na Educação Musical.

Em [www.auburn.edu/outreach/dl/ctmu/jtml/Content\\_of\\_Past.html](http://www.auburn.edu/outreach/dl/ctmu/jtml/Content_of_Past.html) temos acesso aos índices dos artigos já publicados. Não são disponibilizados abstracts.

### ***Nature***

([www.nature.com/nature/index.html](http://www.nature.com/nature/index.html))

A Revista *Nature* pertence ao Nature Publishing Group ([www.nature.com/npg/](http://www.nature.com/npg/)).

No site da Revista é possível fazer-se uma pesquisa rápida (por palavras-chave) ou pesquisa avançada (autor, data, número da revista, etc).

### ***Psicologia em Estudo***

([www.scielo.br/scielo.php?script=sci\\_serial&pid=1413-7372&lng=es](http://www.scielo.br/scielo.php?script=sci_serial&pid=1413-7372&lng=es))

Em [www.scielo.br/scielo.php?script=sci\\_issues&pid=1413-7372&lng=es&nrm=iso](http://www.scielo.br/scielo.php?script=sci_issues&pid=1413-7372&lng=es&nrm=iso), a revista brasileira *Psicologia em Estudo* disponibiliza o índice e o texto integral dos volumes publicados entre 2001 e 2007.

No endereço [www.scielo.br/cgi-bin/wxis.exe/iah/?IsisScript=iah/iah.xis&base=article^dpe&format=iso.pft&lang=e&limit=1413-7372](http://www.scielo.br/cgi-bin/wxis.exe/iah/?IsisScript=iah/iah.xis&base=article^dpe&format=iso.pft&lang=e&limit=1413-7372) é possível pesquisar-se um artigo por palavras-chave, autor, assunto, ano de publicação, etc.

### ***Revista Internacional Latino Americana de Musicoterapia***

([www.geocities.com/Paris/Metro/8395/ilajmt-e.html](http://www.geocities.com/Paris/Metro/8395/ilajmt-e.html))

A *Revista Internacional Latino Americana de Musicoterapia* pertence à ADIMU – Âmbito de Asistencia, Docencia e Investigación en Musicoterapia, que procura promover o desenvolvimento da musicoterapia enquanto disciplina científica e saída profissional.

No site da revista são apresentados resumos dos artigos já publicados.

### ***Science***

([www.sciencemag.org/](http://www.sciencemag.org/))

A revista *Science* pertence à Advancing Science Serving Society ([www.aaas.org/](http://www.aaas.org/)). No site da Revista é possível fazer-se uma pesquisa rápida (por palavras-chave) ou pesquisa avançada (autor, data, número da revista, etc).

### ***The Music Quarterly***

([www.mq.oxfordjournals.org/](http://www.mq.oxfordjournals.org/))

O *The Music Quarterly*, fundado em 1915, foi a primeira publicação erudita no domínio musical nos Estados Unidos. Em [www.mq.oxfordjournals.org/archive](http://www.mq.oxfordjournals.org/archive) temos o índice dos volumes entre 1915 e 2006. Em [www.mq.oxfordjournals.org/search.dtl](http://www.mq.oxfordjournals.org/search.dtl) é-nos permitido fazer pesquisa pelo autor, título ou palavra-chave.

### ***The Neuroscientist***

(<http://sagepub.com/journalsProdDesc.nav?prodId=Journal200902>)

O *The Neuroscientist* pertence à Sage Publications ([www.online.sagepub.com/](http://www.online.sagepub.com/)).



Em <http://nro.sagepub.com/> podemos consultar o índice dos volumes já publicados, tendo acesso gratuito aos abstracts e acesso ao texto integral mediante subscrição.

### ***Visions of Research***

(<http://www-usr.rider.edu/~vrme/>)

O *Visions of Research* é uma publicação oferecida como serviço público pela New Jersey Music Educators Association, organização afiliada do MENC: The National Association for Music Education, disponibilizando-se, a partir do seu site, acesso totalmente gratuito aos textos integrais de todos os artigos já publicados.

### ***Zero to Three***

([www.zerotothree.org](http://www.zerotothree.org))

*Zero to Three* é uma organização sem fins lucrativos que pretende ter como missão informar, educar e auxiliar os adultos no exemplo que são para os bebés. No site desta organização estão disponibilizadas informações variadas acerca do desenvolvimento das crianças desde o nascimento até aos três anos de idade (por exemplo: desenvolvimento do cérebro, saúde mental do bebé, saúde e nutrição, sono, etc.).

## **3.3. Bases de dados**

Nesta secção apresentam-se catálogos informáticos que ajudam a identificar e localizar publicações científicas. Nem todos são gratuitos pelo que, sempre que possível e a título meramente indicativo, mencionam-se instituições que, à data de publicação deste livro, tinham subscrito os respectivos serviços.

### **Blackwell-Synergy**

([www.blackwell-synergy.com/](http://www.blackwell-synergy.com/))

O projecto Blackwell-synergy disponibiliza mais de 1 milhão de artigos de mais de 850 jornais, das mais variadas áreas científicas (Economia, Finanças, Engenharia e Construção, Ciências da Saúde, Humanidades, Direito, Matemática, Medicina, Ciências Sociais, Artes, etc.).

Em [www.blackwell-synergy.com/action/showJournals](http://www.blackwell-synergy.com/action/showJournals) temos acesso à lista das publicações disponíveis.

Em [www.blackwell-synergy.com/search/simple](http://www.blackwell-synergy.com/search/simple) é possível fazer-se uma pesquisa por palavra-chave, autor, título, revista, etc.

O acesso ao motor de pesquisa é gratuito, bem como a leitura dos abstracts dos artigos pesquisados. A leitura dos textos integrais requer subscrição.

### **B-ON**

([www.b-on.pt](http://www.b-on.pt))

A B-On reúne as principais editoras de revistas científicas internacionais, oferecendo um vasto conjunto de artigos nos mais variados domínios do conhecimento científico, com o intuito de estimular e incrementar a produção científica nacional. A pesquisa pode ser feita em vários tipos de fontes (bases referenciais, periódicos online, catálogos online, bases de texto integral, etc.).

### **Catálogo Colectivo de Publicações Periódicas existentes em Bibliotecas Portuguesas** ([www.fct.mctes.pt/infodoc/sidbd/bdcat.htm](http://www.fct.mctes.pt/infodoc/sidbd/bdcat.htm))

A Fundação para a Ciência e a Tecnologia – FCT disponibiliza o acesso a uma base de dados que possibilita a localização de publicações periódicas e nacionais e estrangeiras existentes em bibliotecas do país.

### **Directory of Open Access Journals** (<http://www.doaj.org/>)

O Directory of Open Access Journals disponibiliza o acesso aos textos integrais de numerosas publicações científicas, nas mais variadas áreas e línguas. São disponibilizados mais de 138000 artigos.

### **Domínio Público** ([www.dominiopublico.gov.br](http://www.dominiopublico.gov.br))

O Portal Domínio Público foi lançado em 2004 pelo Ministério da Educação brasileiro, com o intuito de disponibilizar uma biblioteca virtual cuja informação possa ser referência para professores, alunos, pesquisadores, e para a população em geral. Neste site podemos aceder a obras literárias, artísticas e científicas (na forma de texto, som, imagem e vídeo) do património cultural brasileiro e universal. Entre os documentos disponibilizados incluem-se artigos, teses e dissertações sobre música e educação musical.

### **ERIC – Education Resources Information Center** ([www.eric.ed.gov/](http://www.eric.ed.gov/))

O projecto ERIC é uma Base de Dados Online que fornece referências sobre mais de um milhão de documentos na área da educação, incluindo sempre que possível links directos para os textos integrais dos artigos referenciados (o ERIC contém um colecção crescente de artigos integrais gratuitos em formato Adobe PDF format). As referências indexadas no ERIC incluem: artigos de jornais, livros, escritos de conferências, análises técnicas, e outros materiais relacionados com a área da educação.

Os jornais indexados no ERIC podem ser consultados no endereço: [www.eric.ed.gov/ERICWebPortal/Home.portal?nfpb=true&pageLabel=JournalPage&logoUrlLink=false](http://www.eric.ed.gov/ERICWebPortal/Home.portal?nfpb=true&pageLabel=JournalPage&logoUrlLink=false). Neste endereço são disponibilizados os dados essenciais dos jornais indexados (título, ISSN, editor, frequência de publicação, descrição sumária do jornal, títulos de artigos disponíveis online, total de entradas no ERIC).

A pesquisa de artigos específicos (por nome de artigo ou palavras-chave) pode ser feita a partir da página inicial.

De entre as instituições portuguesas que disponibilizam a consulta do ERIC incluem-se, por exemplo: a Faculdade de Ciências Sociais e Humanas da Universidade Nova de Lisboa, a Universidade Católica do Porto, a Universidade de Aveiro, a Universidade do Minho, a Faculdade de Psicologia e Ciências da Educação da Universidade de Lisboa, a Escola Superior de Educação de Setúbal, a Faculdade de Psicologia e Ciências da Educação da Universidade de Coimbra.

### **Free Full Text** ([www.freefulltext.com](http://www.freefulltext.com))

O FreeFullText fornece links directos para mais de 7000 periódicos especializados (incluindo alguns da área musical) com acesso online totalmente gratuito através da internet (embora alguns destes acessos impliquem um registo gratuito online).



### **Ingenta Connect**

([www.ingentaconnect.com](http://www.ingentaconnect.com))

O projecto Ingenta Connect disponibiliza mais de 22 milhões de documentos de mais de 31000 publicações, das mais variadas áreas científicas (agricultura, artes e humanidades, biologia e ciências da vida, química, economia, matemática, medicina, filosofia, sociologia, etc).

Em [www.ingentaconnect.com/content](http://www.ingentaconnect.com/content) temos acesso à lista das publicações disponíveis.

O acesso ao motor de pesquisa é gratuito, bem como a leitura dos abstracts dos artigos pesquisados. A leitura dos textos integrais requer subscrição.

### **International Index to Music Periodicals**

(<http://iimp.chadwyck.com/marketing/index.jsp>)

O International Index to Music Periodicals é uma base de dados que fornece os índices e abstracts para um largo número de jornais sobre música, quer de âmbito académico, quer de âmbito popular. Tem mais de meio milhão de artigos indexados, com os mais variados temas: educação musical, estudos sobre a performance musical, composição, teoria, etnomusicologia, jazz, e todas as formas de música popular. A versão de textos integrais, IIMP Full Text, contém aproximadamente 120 publicações em texto integral, com novo conteúdo adicionado mensalmente.

Este serviço requer subscrição (todas as informações estão disponíveis no site).

### **JSTOR – The Scholarly Journal Archive**

([www.jstor.org](http://www.jstor.org))

O JSTOR é um arquivo electrónico que se dedica essencialmente à digitalização, catalogação e preservação digital de documentos de variadas áreas científicas, pertencentes a editoras e instituições de diversos países. O arquivo do JSTOR é disponibilizado mediante subscrição.

Em [www.jstor.org/about/participants\\_intl.html](http://www.jstor.org/about/participants_intl.html) temos a lista das instituições que participam neste projecto (incluindo instituições portuguesas).

Em [www.jstor.org/about/all.list.html](http://www.jstor.org/about/all.list.html) temos a lista completa dos jornais e revistas associadas ao JSTOR.

### **Music Education Resource Base**

([www.fmpweb.hsd.uvic.ca/merb/](http://www.fmpweb.hsd.uvic.ca/merb/))

O Music Education Resource Base – MERB, que inclui o Canadian Music Index – CMI é suportado pela University of Victoria – Faculty of Education ([www.educ.uvic.ca/](http://www.educ.uvic.ca/)).

O MERB/CMI é uma base de dados bibliográfica de mais de 31.000 fontes sobre música e educação musical de 35 publicações canadianas e internacionais e de outras fontes, cobrindo um período de publicações entre 1956 e a presente data. A lista dos jornais abrangidos por esta base de dados pode ser consultada em [www.fmpweb.hsd.uvic.ca/merb/journals.htm](http://www.fmpweb.hsd.uvic.ca/merb/journals.htm). Sendo somente uma base de dados bibliográfica, o MERB/CMI não permite a consulta dos artigos ou dos seus abstracts, não deixando, contudo, de ser uma ferramenta muito útil para se procurarem artigos de um determinado autor ou sobre um tema específico.

Na página [www.fmpweb.hsd.uvic.ca/merb/search.htm](http://www.fmpweb.hsd.uvic.ca/merb/search.htm) é possível pesquisar-se gratuitamente artigos por título do artigo, autor, designação da revista ou jornal, fonte ou assunto, ou, palavra-chave.

### **PORBASE – Base Nacional de Dados Bibliográficos**

(<http://opac.porbase.org/#focus>)

O PORBASE é a Base Nacional de Dados Bibliográficos, um catálogo colectivo online das bibliotecas portuguesas. De forma muito simples e rápida, este site permite-nos pesquisar por palavra-chave artigos ou livros disponíveis numa ou em várias das bibliotecas nacionais, apresentando-nos de imediato as bibliotecas onde esses artigos ou livros poderão ser consultados.

### **Project Muse – Scholarly Journals Online**

(<http://muse.jhu.edu/>)

O Project Muse disponibiliza online textos integrais de mais de 300 publicações de mais de 60 editoras académicas, em vários domínios como as humanidades, as artes e as ciências sociais.

Em <http://muse.jhu.edu/journals/> temos acesso à lista das publicações disponíveis.

Em <http://muse.jhu.edu/search/search.cgi> é possível fazer-se uma pesquisa por palavra-chave, autor, título, revista, etc. O acesso ao motor de pesquisa é gratuito, bem como a leitura dos abstracts dos artigos pesquisados. A leitura dos textos integrais requer subscrição.

### **Pro Quest Digital Dissertation**

([www.il.proquest.com/](http://www.il.proquest.com/))

A Pro Quest Digital Dissertation é uma base de dados e motor de busca que fornece acesso a mais de 125 bilhões de páginas digitais do mundo académico, em múltiplas áreas, desde as artes à literatura, ciências sociais, tecnologia e medicina.

### **Psycinfo**

(<http://psycinfo.apa.org/psycinfo/>)

A PsycInfo pertence à American Psychological Association ([www.apa.org](http://www.apa.org)).

A PsycINFO é a maior base de dados mundial na área da psicologia, fornecendo resumos e referências sobre mais de 1800 revistas de psicologia e de disciplinas associadas (como a medicina, a psiquiatria, a educação, as ciências sociais e humanas, a neurociência, a linguística, e a farmacologia, etc.).

O acesso a esta Base de Dados não é gratuito. De entre as instituições portuguesas que disponibilizam a consulta da PsycInfo incluem-se: Biblioteca do ISCTE, Biblioteca da Faculdade de Medicina da Universidade do Porto, Faculdade de Psicologia e Ciências da Educação da Universidade de Coimbra, Universidade do Algarve, Instituto Superior de Psicologia Aplicada – ISPA.

### **PUBMED**

([www.ncbi.nlm.nih.gov/sites/entrez](http://www.ncbi.nlm.nih.gov/sites/entrez))

A PUBMED é uma base de dados que fornece acesso a referências bibliográficas na área da literatura biomédica. Através do motor de busca disponibilizado podemos aceder a vários artigos sobre Música e Psicologia.

### **RILM – International Repertory of Music Literature**

(<http://www.rilm.org/>)

O RILM (procura ser um guia abrangente sobre publicações de música de todo o mundo, nas mais variadas áreas de interesse do domínio musical. A base de dados do RILM contém mais de 400.000 entradas (resumos concisos de artigos, livros, bibliografias, catálogos, dissertações, filmes, vídeos, iconografias, comentários críticos de obras



completas, gravações etnográficas, registos de conferências, fontes electrónicas, e resenhas).  
Serviço disponível mediante subscrição.

### **Science Direct**

[www.sciencedirect.com](http://www.sciencedirect.com))

A Science Direct é uma biblioteca digital que nos permite pesquisar artigos dos mais variados domínios científicos (Ciências Físicas e Engenharia, Ciências da Vida, Ciências da Saúde, Ciências Sociais e Humanas), incluindo a música.

A partir do site inicial é possível fazer-se uma pesquisa por palavra-chave, autor, título, revista, etc.

Este é um dos recursos da B-On.

## **3.4. Outros**

### **ADIMU – Âmbito de Asistencia, Docencia e Investigación en Musicoterapia**

[www.geocities.com/Paris/Metro/8395/spanish.html](http://www.geocities.com/Paris/Metro/8395/spanish.html)

### **American Orff-Schulwerk Association**

[www.aosa.org](http://www.aosa.org)

### **American Music Therapy Association**

[www.musictherapy.org/](http://www.musictherapy.org/)

### **American Psychological Association – APA**

[www.apa.org/](http://www.apa.org/)

### **Associação Portuguesa de Ciências Musicais**

[www.apcm.fcsh.unl.pt/](http://www.apcm.fcsh.unl.pt/)

### **Associação Portuguesa de Educação Musical – APEM**

[www.apem.jazznet.pt](http://www.apem.jazznet.pt)

### **Association for Technology in Music Instruction**

[www.atmionline.org/](http://www.atmionline.org/)

### **Australian Association for Research in Music Education**

[www.deakin.edu.au/education/music-ed/aarme/index.php](http://www.deakin.edu.au/education/music-ed/aarme/index.php)

### **BABYLABS**

<http://br.geocities.com/costagomes/babylabs.html>

Site concebido por Carlos Gomes, reunindo informação sobre babylabs existentes na Europa e América do Norte.

### **British Society of Music Therapy – BSMT**

[www.bsmt.org](http://www.bsmt.org)

### **Centro de Estudos de Sociologia e Estética Musical – CESEM**

[www.fcsh.unl.pt/cesem/](http://www.fcsh.unl.pt/cesem/)

O CESEM foi criado para promover a investigação nas áreas da Sociologia da Música e da Estética Musical, em articulação com outras disciplinas das Ciências Musicais e com as Ciências Sociais e Humanas em geral.

Para além de um centro de documentação, o CESEM dispõe de uma máquina que permite a digitalização de fichas e microfiches, o que pode constituir um importante recurso no acesso a bibliografia.

**Centro de Investigação em Psicologia da Música e Educação Musical – CIPEM**  
[www.cipem.wordpress.com/about/](http://www.cipem.wordpress.com/about/)

**Centro Ward de Lisboa – Júlia d'Almendra**  
<http://centroward.no.sapo.pt>

**Companhia de Música Teatral**  
[www.musicateatral.com](http://www.musicateatral.com)

**European Society for the Cognitive Sciences of Music – ESCOM**  
<http://musicweb.hmt-hannover.de/escom/english/index.htm>

**INET – Instituto de Etnomusicologia – Centro de Estudos de Música e Dança**  
[www.fcsh.unl.pt/inet/](http://www.fcsh.unl.pt/inet/)

**International Association of Jazz Educators**  
[www.iaje.org](http://www.iaje.org)

**International Society of Music Education – ISME**  
[www.isme.org/](http://www.isme.org/)

**Meloteca – Sítio de Música e Artes**  
[www.meloteca.com/](http://www.meloteca.com/)

**Centro de Informação da Música Portuguesa**  
[www.mic.pt/port/apresentacao.html](http://www.mic.pt/port/apresentacao.html)

**Music Teachers National Association**  
[www.mtna.org](http://www.mtna.org)

**Music Therapy World**  
[www.musictherapyworld.net/](http://www.musictherapyworld.net/)

**National Association for Music Education – MENC**  
[www.menc.org](http://www.menc.org)

**New York Academy of Sciences**  
[www.nyas.org/index.asp](http://www.nyas.org/index.asp)

**Ohio Music Education Association**  
[www.omea-ohio.org/](http://www.omea-ohio.org/)

**Organization of American Kodály Educators**  
[www.oake.org](http://www.oake.org)

**Società Italiana per l'Educazione Musicale – SIEM**  
[www.siem-online.it/](http://www.siem-online.it/)

**Society of Music Perception and Cognition**  
[www.musicperception.org/](http://www.musicperception.org/)

**Society for Education, Music and Psychology Research – SEMPRES**  
[www.sempre.org.uk/](http://www.sempre.org.uk/)



**Technological Directions in Music Learning ejournal**  
<http://music.utsa.edu:16080/tdml/>

**The American Bandmasters Association**  
[www.americanbandmasters.org/](http://www.americanbandmasters.org/)

**The Gordon Institute for Music Learning**  
[www.giml.org/](http://www.giml.org/)

**TI-ME: Technology Institute for Music Educators**  
[www.ti-me.org](http://www.ti-me.org)

# AUTORES DOS ESTUDOS CRÍTICOS

## CHRISTOPHER M. JOHNSON

Obteve os graus académicos de licenciatura, mestrado e doutoramento na Universidade do Estado da Florida. Desde 1992, lecciona na Universidade do Kansas no âmbito da Educação Musical e da Musicoterapia. Nesta universidade exerce as funções de Associate Dean for Academic Affairs and Research in the School of Fine Arts.

É editor do *International Journal of Music Education: Research* e integrou a direcção editorial da revista *Journal of Research in Music Education*. Publicou artigos em muitas publicações, das quais se destacam: *Journal of Research in Music Education*, *International Journal of Music Education*, *Journal of Music Therapy*, *Psychomusicology*, *Journal of Band Research*, *Contributions to Music Education*, *Southeastern Journal of Music Education*, e no *Bulletin of the Council for Research in Music Education*. Integra o comité executivo do Music Educators National Conference (MENC) e tem sido um membro activo da *International Society for Music Education* (ISME).

Possuindo uma vasta experiência de ensino, Christopher Johnson recebeu os prémios Ned N. Fleming Award for Excellence in Teaching e J. William Fulbright Scholar, respectivamente em reconhecimento da qualidade do seu ensino universitário e pelos seus trabalhos de investigação.

## DEBRA HEDDEN

É Professora Associada de Educação Musical na Universidade do Kansas. Integra a direcção editorial do *Journal of Music Teacher Education*, tendo exercido cargos executivos no Music Educators National Conference (MENC).

Efectuou os seus estudos na Universidade de Iowa e na Universidade de Northern Iowa. A sua experiência ao nível do ensino superior inclui cursos de licenciatura em educação musical, principalmente em metodologia e pedagogia, tendo-se especializado no ensino à distância.

Tem participado em diversos simpósios e conferências, tendo publicado em revistas como *Applications of Research in Music Education*, *Contributions to Music Education*, o *Music Educators Journal*, o *Orff Echo*, entre outros.

Organizou diversos festivais corais e dirigiu o Northern Iowa Children's Choir e o KU Youth Chorus, tendo a qualidade do seu ensino obtido o reconhecimento do *Music Educators National Conference*.



## **HELENA RODRIGUES**

Diplomada com o Curso Superior de Piano do Conservatório de Música do Porto, efectuou também formação nas áreas de Canto, Música de Câmara, Expressão Dramática, Musicoterapia e Pedagogia Musical.

Licenciada em Psicologia pela Faculdade de Psicologia e Ciências da Educação da Universidade do Porto, Mestre em Ciências da Educação pela Faculdade de Psicologia e Ciências da Educação da Universidade de Coimbra. Doutorada em Psicologia por esta universidade.

Professora-Auxiliar da Faculdade de Ciências Sociais e Humanas da Universidade Nova de Lisboa e membro do CESEM (Centro de Estudos de Sociologia e Estética Musical). Desde 1994 que trabalha em estreita colaboração com Edwin Gordon, sendo responsável pela introdução das suas ideias no nosso País. Tem publicados diversos trabalhos de natureza científica, tendo vindo a privilegiar a infância e o estudo do desenvolvimento musical e psicológico como campo de trabalho privilegiado.

Directora-artística da Companhia de Música Teatral ([www.musicateatral.com](http://www.musicateatral.com)), participa regularmente na concepção e interpretação dos trabalhos artísticos deste grupo.

## **JOÃO NOGUEIRA**

Psicólogo, doutor (Especialidade de Psicologia da Educação pela Universidade Nova de Lisboa) e mestre (área de Psicologia da Educação pela Universidade de Lisboa) em Ciências da Educação.

Professor-auxiliar da Faculdade de Ciências Sociais e Humanas da UNL.

Lecciona Psicologia Educacional na formação inicial de professores e Métodos de Investigação nos mestrados em Ciências da Educação e de Psicologia e Pedagogia Musical, além de Comunicação Interpessoal naquele mestrado. Orienta vários trabalhos que envolvem a autoconfiança dos professores, alunos e dos músicos.

Músico autodidacta, é contrabaixista dos Lundum Ensemble e da Lisbon Swingers Jazz Big Band.

## **ORLINDO GOUVEIA PEREIRA**

Médico psiquiatra, doutor em Psicologia pela Universidade de Brandeis (U.S.A).

Professor Catedrático de Psicologia da Arte na Universidade Nova de Lisboa, Universidade de Évora e Lusíada do Porto.

Conferencista convidado na Área de Psicologia e Pedagogia Musical do Mestrado em Ciências Musicais da Faculdade de Ciências Sociais da Universidade Nova de Lisboa. Autor de livros e artigos científicos e artísticos.

## **WENDY L. SIMS**

Efectuou os seus estudos na Universidade do Estado de Kent, tendo realizado o Doutoramento em Educação Musical na Universidade do Estado da Florida.

Professora e Directora de Educação Musical na Universidade do Missouri-Columbia. Possuindo uma vasta experiência de ensino, quer a nível universitário como ao nível do pré-escolar, é especialista em investigação da educação musical na primeira infância.

Apresenta *workshops* e conferências, publicando regularmente em jornais e revistas nacionais e internacionais. Editou o livro "Strategies for Teaching Prekindergarten Music" e co-editou o livro "Music in Prekindergarten", ambos publicados pela MENC-The National Association for Music Education (USA).

Desempenhou cargos directivos no Music Educators National Conference e na International Society for Music Education.





## COLABORADORES

### ÂNGELO MARTINGO

*Doctor of Philosophy in Music* – PhD (University of Sheffield); *Master of Music in performance* (Reading); *Diplome Supérieur D'Enseignement* (École Normale de Musique); Bacharel em Instrumento – piano (ESMAE).

Investigador e Membro do Conselho Científico do Centro do Centro de Estudos em Sociologia e Estética Musical (CESEM). Docente e Coordenador Científico do Curso de Música do Instituto Superior de Estudos Interculturais e Transdisciplinares (ISEIT-Mirandela).

Co-editor de *Interpretação Musical* (Lisboa: Colibri, 2006); Co-editor e co-fundador da revista electrónica *Performance On-line – Revista de Interpretação Musical* ([www.performanceonline.org](http://www.performanceonline.org)).

### ISABEL MONTEIRO

Concluiu os seus estudos superiores de composição no Conservatório de Música do Porto e na ESMAE – Porto. Fez um mestrado sobre Varèse no IRCAM / EHESS em Paris. Actualmente dedica-se ao ensino de crianças e jovens nas disciplinas de Educação Musical e Acústica.

Tendo como principal área de interesses de investigação a Psicoacústica e em particular o estudo do timbre, frequentou a parte curricular da Área de Psicologia e Pedagogia Musical do Mestrado em Ciências Musicais da Faculdade de Ciências Sociais e Humanas da Universidade Nova de Lisboa.

### NUNO ARRAIS

É professor de Contrabaixo na Escola Profissional Artística do Vale do Ave e na Escola Profissional de Música de Viana do Castelo.

Estudou Piano no Centro de Cultura Musical e Contrabaixo na Escola Profissional Artística do Vale do Ave. É bacharel pela Academia Nacional Superior de Orquestra e licenciado pela Escola Superior de Artes do Espectáculo do Instituto Politécnico do Porto. Em Portugal e no estrangeiro estudou com professores como: Alexandre Samardjiev, Florian Petzborn, António Aguiar, Ercole de Conca, Ubaldo Fioravanti, Giuseppe Etorre e Paul Ellison.

Como instrumentista é regular a sua apresentação pública em recitais com piano e com a Companhia de Música Teatral.

Interessando-se pelo estudo das emoções na performance musical, concluiu recentemente a dissertação para a obtenção do grau de Mestre em Ciências Musicais na especialidade de



Psicologia e Pedagogia Musical na Faculdade de Ciências Sociais e Humanas da Universidade Nova de Lisboa.

## **TIAGO PAULOS VEIGA**

Estudou guitarra (eléctrica e clássica) e improvisação musical na Escola Interartes de Cascais, com o Professor Jorge Lee.

É Licenciado em Ciências Musicais pela Faculdade de Ciências Sociais e Humanas da Universidade Nova de Lisboa (FCSH-UNL), tendo aí concluído também a Pós-Graduação em Educação.

Trabalhando regularmente com bebés e crianças em idade pré-escolar e escolar, é actualmente bolseiro da FCT no âmbito do projecto “Desenvolvimento Musical na Infância e Primeira Infância” a decorrer no Centro de Estudos de Sociologia e Estética Musical da FCSH-UNL.

# GLOSSÁRIO

## A

### **Amostra de conveniência (Convenience sample)**

Amostra facilmente disponível.

### **Análise de dados (Data analysis)**

Conjunto dos processos de organização e tratamento estatístico que permitem visualizar, classificar, descrever e interpretar os dados colhidos junto dos sujeitos de modo a torná-los compreensíveis.

### **Análise de covariância (Analysis of covariance ou ANCOVA)**

Técnica estatística usada para igualizar os grupos em uma ou mais variáveis quando se testa a significância estatística. Ou seja, os dados relativos a uma dada variável em estudo são ajustados em função de diferenças pré-existentes à experiência. Esta análise permite reduzir o enviesamento decorrente dessa parte de variabilidade pré-existente.

A análise multivariada da covariância (MANCOVA) é uma extensão da análise de covariância.

### **Análise de Variância (Analysis of Variance ou ANOVA ou F Test)**

Técnica da estatística paramétrica usada para determinar a significância estatística das diferenças entre médias. Pode ser usada com dois ou mais grupos, não necessariamente com o mesmo número de sujeitos.

A análise de variância "one way" é usada para comparar grupos entre si.

A análise de variância "two-way" é usada para comparar três ou mais grupos em duas variáveis diferentes simultaneamente. Ou seja, os mesmos indivíduos ou grupos são comparados nas duas condições.

A análise multivariada da variância (MANOVA) é uma extensão da análise de variância.

### **Análise paramétrica**

Ver estatística paramétrica.

### **Análise não paramétrica**

Ver estatística não paramétrica.

## B

### **Balanceamento (Counterbalanced design)**

Desenho experimental em que todos os grupos recebem o tratamento, mas por diferente ordem de forma a minorar os próprios efeitos da ordem de administração do tratamento.

### **Biofeedback (Reaferição biológica)**

Circuito de retroacção electrónica de fenómenos corporais, geralmente não consciencializados. O sujeito, recebendo informação por esta via artificial, é convidado a responder a ela, num determinado sentido, de acordo com as instruções do experimentador. Por exemplo, o ritmo cardíaco é tornado audível e o sujeito deve actuar sobre ele, tentando aumentá-lo ou diminuí-lo.

## C

### **Canto rítmico (Chant)**

Interpretação vocal da parte rítmica de uma melodia, fazendo-o com inflexões expressivas mas sem diferenças de altura do som. Note-se que esta é uma designação adoptada para o substantivo inglês, uma vez que não existe um termo equivalente em português. Ver também entoar.



**Coefficiente Alfa (Coefficient alpha)**

Índice de fiabilidade que pode ser pensado como a média de todos os índices de fiabilidade calculados segundo o método das duas metades corrigidos através da fórmula de Spearman-Brown.

**Coefficiente de correlação (Correlation coefficient (r))**

Valor que indica o grau de relação entre duas variáveis quantitativas medidas num mesmo grupo de sujeitos. Pode apresentar números decimais entre -1 (correlação negativa perfeita) e +1 (correlação positiva perfeita); um coeficiente de 0 expressa uma ausência de relação entre as duas variáveis.

O coeficiente de correlação de Pearson (Pearson r) é adequado apenas quando as variáveis em questão são contínuas, intervalares ou expressas em números racionais. Portanto, não serve para variáveis ordinais (por ex., não pode ser utilizado se uma das variáveis é o nível de instrução) ou nominais.

**Coefficiente de fiabilidade (Reliability coefficient)**

Índice da consistência de resultados obtidos através do mesmo instrumento e que é o *ratio* entre o resultado verdadeiro e a variância total dos resultados do teste. Dependendo do tipo de consistência que se pretende estudar e das características do instrumento, podem ser utilizados diversos métodos de cálculo de um coeficiente de fiabilidade.

Por exemplo, o **coeficiente de equivalência ou teste das formas paralelas** determina-se utilizando duas formas paralelas do instrumento de medida, com o mesmo grupo de sujeitos e calculando um coeficiente de correlação entre os resultados que os sujeitos obtiveram com cada uma das duas formas. O **coeficiente de estabilidade ou teste-reteste** é calculado a partir dos resultados obtidos através da administração repetida do mesmo teste com o mesmo grupo de sujeitos. O coeficiente de fiabilidade pode ainda determinar-se segundo o **método das duas metades**, situação em que as duas metades de um teste (por exemplo, uma constituída pelos itens ímpares e outra pelos itens pares) são correlacionadas uma com a outra, sendo esta correlação ajustada através da fórmula de Spearman-Brown.

**Coefficiente de validade (Validity coefficient)**

Correlação entre o teste e o critério (que pode ser concorrente ou preditivo).

**Conclusão (Conclusion)**

Consequência lógica deduzida da análise dos dados. Parte de um trabalho de investigação que resume os principais resultados da investigação, as suas implicações e limites, podendo apresentar sugestões para outras investigações.

**Contextos reais (Authentic contexts)**

*In situ*, em contexto real (com metodologia descritiva).

**Contingência (Contingency)**

Relação entre o comportamento de um organismo e acontecimentos externos (usados normalmente como reforços) que se sucedem ao comportamento e que aumentam ou diminuem a probabilidade de um comportamento semelhante ocorrer no futuro.

**Construto (Construct)**

Construção teórica relativa a um traço ou característica psicológica supostamente existente.

**Correlação (Correlation)**

Relação entre duas variáveis que se influenciam mutuamente. A correlação não expressa causalidade. A correlação é positiva quando os resultados seguem o mesmo tipo de progressão entre ambas as variáveis. É negativa quando baixos resultados numa variável estão associados a resultados elevados noutra variável. É nula quando não existe relação entre os resultados das duas variáveis.

**Curva de aprendizagem (Learning Curve)**

A curva de aprendizagem refere uma relação entre a duração e o progresso da aprendizagem. O conceito remonta aos trabalhos de Hermann Ebbinghaus que no século XIX estabeleceu as chamadas "curva de memorização" e "curva de esquecimento" relacionando memorização e esquecimento com outros factores. Tem sido aplicado a outros contextos. Por exemplo, no campo da performance musical é interessante notar que a partir de uma certa altura o excesso de repetição no âmbito da prática musical pode produzir efeitos inversos no que concerne aos efeitos sobre a aprendizagem.

## D

### Dados (Data)

Informações acerca de uma amostra ou de uma população.

### Definição do problema (Problem statement)

Conjunto de proposições que mostram a finalidade (**purpose**) específica da investigação, as variáveis em estudo e especificamente uma relação entre estas variáveis; inclui a descrição do background e a justificação do estudo.

### Definição operacional

Definição de uma variável em termos observáveis e mensuráveis, em que se especificam as actividades necessárias para medir e classificar objectivamente uma variável.

### Desenho experimental ou plano experimental (Experimental design)

Plano de investigação que visa obter respostas válidas para as hipóteses de investigação formuladas e que se caracteriza por fornecer informações sobre os sujeitos, a amostragem, os instrumentos de medida, os métodos de colheita de dados e as técnicas de análise de dados que o investigador pretende utilizar.

### Desenho de investigação (Design)

Ver plano de investigação.

### Discussão (Discussion)

Reflexão sobre os resultados do estudo, integrando-os numa perspectiva abrangente e incluindo limitações do mesmo; parte de um relatório de investigação que oferece uma interpretação dos resultados da investigação.

### Distribuição t de student

Teste estatístico paramétrico, utilizado para estudar a diferença entre duas médias.

## E

### Efeito de Hawthorne (Hawthorne Effect)

Efeito decorrente da modificação do comportamento dos participantes de uma investigação pelo facto de estarem a participar na mesma, implicando repercussões ao nível das observações relativas à variável dependente.

### Entoar (To chant)

De acordo com o Dicionário de Língua Portuguesa, a palavra entoar significa dar tom, fazer soar, solfejar. Neste livro – bem como noutras traduções de livros ingleses – a palavra entoar tem sido adoptada como tradução do verbo *To chant*, para o qual não existe, de facto, um termo equivalente em português. Em termos práticos, o vocábulo inglês designa a possibilidade de interpretar vocalmente a parte rítmica (apenas, isto é, sem contar com diferenças na altura do som) de uma melodia, fazendo-o com inflexões expressivas.

### Escala de Lickert (Lickert scale)

Escala de medida que fornece uma indicação da opinião ou atitude do sujeito expressando o seu grau (geralmente entre 1 e 5 ou entre 1 e 7, portanto, é uma escala contínua) de acordo ou de desacordo relativamente a cada um dos enunciados propostos.

### Estatística descritiva (Descriptive statistics)

Medidas estatísticas que descrevem e resumem os resultados de um estudo. Incluem medidas de tendência central (por ex. a média), variabilidade (por ex. o desvio padrão) e correlação.

### Estatística inferencial (Inferential statistics)

Parte da Estatística cujas técnicas de análise de dados, baseadas na teoria das probabilidades, permitem deduzir se os resultados encontrados em amostras podem ser generalizados para a população. Existem estatísticas paramétricas e não paramétricas.



**Estatística não-paramétrica (Non-parametric statistics)**

Tipo de testagem estatística que não necessita de distribuição normal e variância igual entre os grupos, por oposição aos testes paramétricos, que assentam em medidas intervalares com distribuição normal, variância homogénea e intervalos contínuos e iguais. A testagem não-paramétrica aplica-se normalmente a grandes amostras, baseando-se em medidas ordinais e nominais.

**Estatística paramétrica (Parametric statistics)**

Método de estatística inferencial cuja utilização se baseia no pressuposto de que a amostra é representativa de uma população normal.

**Estatisticamente significativo (Statistically significant)**

Conclusão de que muito provavelmente os resultados obtidos não se devem ao acaso ou a erros de amostragem; ou seja, é bem provável que uma determinada correlação ou diferença observada exista, de facto, na população.

**Estudo longitudinal (Longitudinal study)**

Estudo realizado ao longo de um certo período de tempo visando a colheita de dados relativos aos mesmos participantes de uma investigação em diversas ocasiões.

**Estudo piloto (Pilot study)**

Estudo prévio realizado em escala reduzida e que visa determinar a viabilidade de uma investigação futura, identificar possíveis problemas e assegurar que os métodos e os objectivos são adequados.

**Erro de medida (Measurement error)**

Variação nos resultados devido a falta de fidelidade do instrumento de medida e que faz com que o resultado obtido seja diferente do resultado verdadeiro.

**Erro Tipo I ou Erro Alfa (Type Error I or Alpha Error)**

Erro que consiste em concluir sobre a existência de uma diferença real entre duas populações quando esta diferença não existe na realidade. Ou seja, a hipótese nula é rejeitada quando deveria ser aceite.

**Erro Tipo II ou Erro Beta (Type Error II or Beta Error)**

Erro que consiste em concluir que há ausência de uma diferença real entre duas populações quando na realidade existe uma diferença. Ou seja, a hipótese nula é aceite, quando deveria ser rejeitada.

**Experiência (Experiment)**

Ver investigação experimental.

## F

**Fiabilidade ou fidelidade (Reliability)**

Característica de um equipamento ou de um instrumento (por exemplo, um teste ou uma grelha de observação) que expressa a consistência da medida. Isto é, tais instrumentos asseguram que se colhe os mesmos dados cada vez que os mesmos fenómenos se produzem.

**Fidelidade ou fiabilidade interjuízes (Interjudge reliability ou inter-scorer reliability)**

Consistência entre as observações ou classificações atribuídas por dois ou mais juízes independentes.

**Fidelidade ou fiabilidade intrajuízes (Intrajudge reliability)**

Consistência entre as observações ou classificações atribuídas pelo mesmo juiz colhidas em momentos de avaliação diferentes.

**Fonte primária (Primary source)**

Documento que provém directamente do autor e cujo conteúdo é original.

**Fonte secundária (Secondary source)**

Documento que consiste numa revisão, descrição, organização e interpretação de trabalhos de fonte primária já publicados.

**Fonte terciária (Tertiary source)**

Categoria de documentos que inclui elementos de fonte primária e secundária. Por ex.: dicionários.

### **Fórmula de Spearman-Brown (Spearman-Brown prophecy formula)**

Fórmula usada para ajustar a correlação entre as duas metades de um teste (e que dá o coeficiente de fiabilidade calculado segundo o método das duas metades) de forma a compensar o tamanho do teste.

## **G**

### **Grelhas de avaliação (Forms)**

Material usado na avaliação de uma determinada de acção e que define os comportamentos de observação e os respectivos critérios de anotação e/ou classificação.

### **Grupo de Controle (Control group)**

Grupo de indivíduos semelhante aos indivíduos do Grupo Experimental mas não submetido à intervenção cujos efeitos se pretende avaliar.

### **Grupo Experimental (Experimental group)**

Grupo submetido ao tratamento que está a ser estudado.

## **H**

### **Hipótese (Assumption or hypothesis)**

Previsão dos resultados do estudo; afirmação carecida de confirmação relativa à ocorrência de determinado tipo de relação esperada entre duas ou mais variáveis; as hipótese (que podem também ser expressas sob a forma de problemas ou questões) devem ser descritas no início de um trabalho científico.

### **Hipótese de investigação (Research hypothesis)**

Afirmação que prediz a(s) relação(ões) esperada(s) entre duas ou mais variáveis. É uma resposta plausível para a(s) questão (ões) da investigação.

### **Hipótese nula (Null hypothesis)**

Enunciado que prediz a ausência de relação ou de diferença estatisticamente significativa entre grupos de sujeito numa dada variável dependente. (Ou seja, afirma-se que uma dada diferença obtida num determinado parâmetro é devida a um erro de amostragem ou é devida ao acaso). Habitualmente, a hipótese nula é formulada para ser rejeitada, devendo depois testar-se a hipótese alternativa.

## **I**

### **Instrument (Instrument)**

Forma de recolher dados de modo sistemático como, por exemplo, um teste, um questionário ou uma grelha de entrevista.

### **Interacção (Interaction)**

Influência de duas variáveis sobre uma terceira. Efeito criado por combinação entre duas ou mais variáveis independentes.

### **Interpretação (Interpretation)**

Etapas de uma investigação que visa compreender os dados colhidos.

### **Intervalo de confiança (Confidence interval)**

Intervalo numérico que estabelece as margens em que provavelmente se situa o resultado verdadeiro. Ou seja, o intervalo de confiança fornece o resultado mais elevado e o resultado mais baixo. O resultado verdadeiro situar-se-á, com um determinado grau ou nível de confiança, neste âmbito intervalar.

### **Investigação bibliográfica (Bibliographic research)**

Investigação baseada em fontes bibliográficas.

### **Investigação empírica (Empirical research)**

Investigação baseada em evidências observáveis.



**Investigação aplicada (Applied Research)**

Investigação cujo principal objectivo é encontrar uma aplicação prática de novos conhecimentos.

**Investigação experimental (Experiment)**

Estudo de investigação em que o procedimento é rigorosamente controlado e replicável e em que uma ou mais variáveis independentes são manipuladas sistematicamente pelo investigador de modo a determinar os seus efeitos sobre a(s) variável(eis) dependentes.

**Investigação fundamental (Fundamental research)**

Investigação que visa aumentar o domínio do saber sem se preocupar com as aplicações práticas imediatas.

**Investigação quase-experimental (Quasi-experimental research)**

Estudo da relação entre variáveis após manipulação da variável independente, em que os sujeitos não foram seleccionados ao acaso e/ ou em que as variáveis parasitas escapam ao controle do investigador.

## K

**Kuder-Richardson formula 20 (Kuder-Richardson formula 20)**

Fórmula adequada para calcular o índice de fiabilidade em situações em que cada ítem do teste é classificado como certo (1) ou errado (0).

## L

**Linha de base (Baseline)**

Relatório gráfico de medidas registadas antes de uma intervenção.

## M

**Medir (To measure)**

Atribuir valores numéricos a variáveis (sejam elas objectos ou traços psicológicos) segundo regras estabelecidas, com o objectivo de representar as características dessas variáveis.

**Metodologia (Methodology)**

Conjunto de métodos e técnicas que guiam o processo de investigação. Parte de um relatório de investigação que descreve os sujeitos, instrumentos e procedimento utilizados no âmbito dessa investigação.

**Método científico (Scientific method)**

Sistema de princípios e questionamento que inclui a identificação do problema, a formulação de hipóteses, a recolha e análise de dados visando testar essas hipóteses e a divulgação desses resultados, de modo a que outros cientistas possam verificar, aprender, analisar, replicar e criar.

## N

**Nível de confiança (Level of confidence)**

Nível de probabilidade pré-definido segundo o qual um determinado intervalo de confiança inclui o parâmetro em estudo. Este nível de probabilidade estabelece o número de vezes que, em 100 ou 1000, um resultado ocorre por acaso.

Na investigação em Psicologia da Música é frequente a utilização de níveis de confiança de 95 e 99%. Isto é, o resultado obtido poderá ter ocorrido por acaso 1 ou 5 vezes em 100. Ou, expresso de outro modo, quando se usa  $p=0.05$  (ou  $p > .05$ ) quer dizer que em 100 estudos idênticos apenas 5 obteriam esses resultados como uma variação normal. Ou seja, dificilmente se poderá atribuir à sorte esse resultado; há uma grande probabilidade de o resultado obtido não ter sido fruto do acaso e, portanto, de o efeito ser estatisticamente real.

Nas ciências sociais convencionou-se arriscar esta margem de erro. No entanto, noutras áreas esse risco terá de ser muito mais reduzido (por ex. não podemos voar com uma probabilidade de 5% de queda do avião).

### **Nível de significância (Level of significance)**

Probabilidade de rejeitar a hipótese nula quando esta é verdadeira.

### **Nível de significância Alfa $\alpha$ (Alfa Level of significance)**

Valor numérico associado ao risco de se enganar, rejeitando a hipótese nula quando na realidade esta hipótese é verdadeira (ou seja, não há relação ou diferenças entre variáveis e diz-se que há). Por exemplo, se o nível de significação alfa é fixado a 0.05 (pode também escrever-se apenas .05), o investigador aceita a probabilidade de se enganar 5 vezes em cada 100 ao rejeitar a hipótese nula. Se o nível de significação alfa é fixado a 0.01, (pode, também, escrever-se apenas .01) o investigador aceita a probabilidade de se enganar 1 vez em cada 100 ao rejeitar a hipótese nula. Portanto, quanto mais exigente for o nível de significação alfa (ou seja, quanto menor for o seu valor), menor é a probabilidade de se enganar.

## **P**

### **Plano experimental (Experimental Plan)**

Ver desenho experimental.

### **Plano de investigação ou desenho da investigação (Research design)**

Plano e estratégia geral para recolha de dados visando uma resposta válida para as questões de investigação ou para as hipóteses formuladas.

### **Pós-Teste (Post-test)**

Administração de um teste ou de outro tipo de medição depois de aplicar a intervenção experimental.

### **Pré-Teste (Pre-test)**

Administração de um teste ou de outro tipo de medição antes de aplicar a intervenção experimental. Ensaio de um teste, instrumento de medida ou de um equipamento antes da sua administração em maior escala.

### **Potência ou robustez de um teste (Robust Testing)**

Probabilidade de rejeitar a hipótese nula quando ela é falsa. Isto é, probabilidade de não cometer um erro tipo II (beta).

### **Procedimento (Procedures)**

Descrição detalhada de como foi (ou será) realizado o estudo.

## **R**

### **Relatório de Investigação (Research report)**

Relato escrito de uma investigação. É normalmente constituído pelas seguintes partes: introdução; revisão teórica (eventualmente subdividida numa secção em que se faz a revisão da literatura relativa a questões mais gerais relativas ao problema em estudo e outra em que se faz uma revisão mais específica e focalizada); metodologia (incluindo os procedimentos de análise); apresentação dos resultados; discussão e interpretação dos resultados; conclusões; referências bibliográficas. (Uma tese ou dissertação inclui estes elementos).

### **Revisão da literatura (Literature review)**

Identificação e exame profundo, sistemático e crítico de fontes bibliográficas contendo informação relativa ao problema de investigação que é objecto de estudo.

### **Resultados (Results)**

Explicação da análise dos dados recolhidos que visam responder às questões ou hipóteses da investigação. Parte de um relatório de investigação que resume (pode incluir tabelas e gráficos), de forma objectiva e sem interpretação, a análise dos dados.



## **Resumo (Summary)**

Descrição sumária do relatório de investigação, devendo incluir uma frase para definir o objectivo do estudo; outra para caracterizar os sujeitos, instrumentos e o desenho da investigação; outra para referir os resultados mais importantes do estudo, mencionando os procedimentos de análise; e, por último, a discussão dos resultados, apresentando uma síntese das principais conclusões e suas implicações.

## **S**

### **Seleção aleatória (Random assignment)**

Processo de seleccionar indivíduos ou grupos aleatoriamente para serem submetidos a diferentes condições de tratamento.

## **T**

### **Teste de aptidão (Aptitude test)**

Teste usado para prever a performance do indivíduo numa situação futura num domínio específico e claramente definido do comportamento.

### **Teste de desempenho ou teste de realização (Achievement test)**

Teste usado para medir a competência ou performance do sujeito em determinada área de conhecimento ou domínio do saber num dado momento e que é resultado quer do seu potencial inato como da sua aprendizagem e treino específico. Teste que mede o grau de sucesso de uma determinada aprendizagem.

### **Teste de Qui Quadrado (*Chi-square test*)**

Teste não paramétrico de significância estatística usado para examinar diferenças com variáveis nominais (portanto, os dados são registos de frequências). É utilizado em dois tipos de circunstâncias: a) Para comparar frequências observadas com frequências esperadas visando verificar se são significativamente diferentes; b) Para verificar a existência de uma relação entre duas variáveis (ou, talvez melhor, estimar se duas variáveis são independentes).

### **Teste de significância estatística (Statistical significance tests)**

Testes estatísticos utilizados para determinar se é provável que os resultados obtidos com uma amostra sejam representativos da população.

Operação que consiste em determinar se a diferença observada entre dois grupos é suficientemente grande para que se possa atribuir ao tratamento experimental e não ao acaso.

### **Teste t (T test)**

Meio de avaliação estatística cujo objectivo é a comparação entre duas médias, para determinar se a probabilidade da diferença entre as mesmas é efectiva ou ocasional.

### **Tratamento experimental ou intervenção experimental (Experimental treatment)**

Manipulação da Variável Independente num estudo experimental que, em princípio, deverá produzir mudanças na Variável Dependente.

## **V**

### **Variância (Variance)**

Índice estatístico que reflecte o grau de dispersão num grupo de resultados.

### **Validade (Validity)**

Capacidade para um instrumento (teste, grelha de observação, por exemplo) medir o que se propõe medir. Um teste é válido se permite tirar conclusões adequadas, significativas e úteis para o fim a que se destina.

### **Validade concorrente (Concurrent validity)**

Tipo de validade em que a medida de critério é obtida ao mesmo tempo que os resultados do teste.

### **Validade de conteúdo (Content validity)**

Tipo de validade que define em que medida as questões e/ou itens do teste ou as tarefas avaliadas são representativas do universo de comportamento que o teste pretende medir.

**Validade de critério ou validade criterial ou validade referenciada a um critério (Criterion-related validity)**

Tipo de validade que demonstra que o teste é eficaz para estimar o desempenho do sujeito numa determinada situação.

**Validade externa (External validity)**

Qualidade de uma investigação que apresenta um grau de exactidão suficiente para que os resultados obtidos possam ser generalizados.

**Validade interna (Internal validity)**

Qualidade de uma investigação científica que permite afirmar que os resultados obtidos podem ser atribuídos à manipulação da variável independente.

**Validade preditiva (Predictive validity)**

Tipo de validade de critério em que as medidas de critério são obtidas posteriormente (meses ou anos) à colheita dos resultados do teste. (Por exemplo, se se comprovar que um teste de aptidão musical detecta bem os sujeitos que virão a ser bons músicos é porque tem validade preditiva).

**Variabilidade (Variability)**

Quantidade de dispersão dos resultados em torno de um valor central.

**Variável (Variable)**

Situação, comportamento ou característica que varia, assumindo pelo menos dois valores.

**Variável dependente (Dependent variable)**

Variável que se supõe ser afectada pela Variável Independente. Resultado dependente do efeito da manipulação da Variável Independente. Pode, eventualmente, ser chamada "Variável critério".

**Variável independente (Independent variable)**

Variável que afecta – ou se supõe afectar – a Variável Dependente em estudo e cujo efeito se pretende determinar através do desenho de investigação. Condição controlada e manipulada pelo investigador, com a finalidade de estudar os seus efeitos na variável dependente. Pode ser também chamada "variável experimental" ou "variável tratamento".

**Variável parasita ou variável estranha ou variável confundida (Extraneous variable)**

Variável não controlada e que pode explicar os resultados. Isto é, variável presente na investigação por ausência de controle por parte do investigador e que pode inesperadamente produzir efeitos sobre a variável dependente, falseando os resultados.

**Variável moderadora ou variável atributo (Moderator variable)**

Variável que pode ou não pode ser controlada mas que tem um efeito no estudo de investigação e que pode servir para descrever uma amostra.





## BIBLIOGRAFIA

- Almeida, L. & Freire, T. (2000) – *Metodologia da Investigação em Psicologia e Educação*. Braga: Psiquilíbrios.
- APA (1994) – *Publication Manual of the American Psychological Association*. Washington, DC: APA.
- Carmo, H. & Ferreira, M. (1998) – *Metodologia da Investigação. Guia para Auto-Aprendizagem*. Lisboa: Universidade Aberta.
- Cozby, P. (2003) – *Métodos de Pesquisa em Ciências do Comportamento*. São Paulo: Ed. Atlas.
- D' Oliveira (2005) – *Teses e Dissertações*. Lisboa: Editora RH.
- Gordon, E. (2005) – *Music Education Research – Taking a Panoptic Measure of Reality*. Chicago: GIA Publications.
- Gross, R. (2003) – *Key Studies in Psychology*. London: Hodder & Stoughton.
- Kerlinger (2003) – *Metodologia da Pesquisa em Ciências Sociais*. São Paulo: E.P.U.
- Madsen, C. & Madsen, C. (1978) – *Experimental Research in Music*. Raleigh: CPC, Inc.
- Popper, K. (1989). *Em Busca de Um Mundo Melhor*. Lisboa: Ed. Fragmentos.
- Rossi, J. et al (1999) – *Os Métodos de Investigação em Psicologia*. Lisboa: Instituto Piaget.
- Tuckman, B. (1994) – *Manual de Investigação em Educação*. Lisboa: Fundação Calouste Gulbenkian.





Execução Gráfica  
Colibri – Artes Gráficas  
Faculdade de Letras  
Alameda da Universidade  
1600-214 Lisboa  
Telef. / Fax 21 796 40 38  
[www.edi-colibri.pt](http://www.edi-colibri.pt)  
[colibri@edi-colibri.pt](mailto:colibri@edi-colibri.pt)



