

Teste de Raciocínio Auditivo Musical (RAu): estudo inicial por meio da Teoria de Resposta ao Item

Fernando Pessotto – Universidade São Francisco, Itatiba, São Paulo, Brasil

Tatiana Freitas da Cunha – Universidade São Francisco, Itatiba, São Paulo, Brasil

Ricardo Primi – Universidade São Francisco, Itatiba, São Paulo, Brasil

Lucas de Francisco Carvalho – Universidade São Francisco, Itatiba, São Paulo, Brasil

Fabiano Koich Miguel – Universidade Estadual de Londrina, Londrina, Paraná, Brasil

Resumo

A presente pesquisa tem como objetivo buscar evidências de validade com base na estrutura interna e de critério para um instrumento de avaliação do processamento auditivo das habilidades musicais (Teste de Processamento Auditivo com Estímulos Musicais, RAu). Para tanto, foram avaliadas 162 pessoas de ambos os sexos, sendo 56,8% homens, com faixa etária entre 15 e 59 anos ($M=27,5$; $DP=9,01$). Os participantes foram divididos entre músicos ($N=24$), amadores ($N=62$) e leigos ($N=76$), de acordo com o nível de conhecimento em música. Por meio da análise *Full Information Factor Analysis*, verificou-se a dimensionalidade do instrumento, e também as propriedades dos itens, por meio da Teoria de Resposta ao Item (TRI). Além disso, buscou-se identificar a capacidade de discriminação entre os grupos de músicos e não-músicos. Os dados encontrados apontam evidências de que os itens medem uma dimensão principal ($\alpha=0,92$) com alta capacidade para diferenciar os grupos de músicos profissionais, amadores e leigos, obtendo-se um coeficiente de validade de critério de $r=0,68$. Os resultados indicam evidências positivas de precisão e validade para o RAu.

Palavras-chave: Processamento auditivo; Teoria de Resposta ao Item; Músicos; Avaliação da inteligência; Análise fatorial por informação completa.

Auditory Musical Reasoning Test: an initial study with Item Response Theory

Abstract

This study investigated internal structure and criterion validity of a test that aims at assessing auditory processing of musical ability (Auditory Musical Reasoning Test, RAu). 162 people of both sexes were evaluated, 56.8% men, aged between 15 and 59 years of age ($M=27.5$; $SD=9.01$). Participants were divided among musicians ($N=24$), amateurs ($N=62$) and lay people ($N=76$) according to the extension of their knowledge in music. Full Information Item Factor Analysis verified the dimensionality of the instrument and also the properties of the items via Item Response Theory (IRT). Furthermore, we sought to identify the ability to discriminate between professional musicians, amateurs and lay people. Data showed evidence that the items measure a major dimension ($\alpha=.92$) with high ability to differentiate groups of musicians, amateurs and lay people giving a criterion validity coefficient of $r=.68$. The results indicate positive evidence of reliability and validity for RAu test.

Keywords: Auditory processing; Item Response Theory; Musicians; Intelligence assessment; Full information factor analysis.

Test de Raciocinio Auditivo Musical (RAu): estudio inicial a través de la Teoría de Respuesta al Ítem

Resumen

La presente investigación tiene como objetivo buscar evidencias de validez basadas en la estructura interna y de criterio para un instrumento de evaluación del procesamiento auditivo de las habilidades musicales (Test de Procesamiento Auditivo con Estímulos Musicales, RAu). Para eso, fueron evaluadas 162 personas de ambos los sexos, siendo 56,8% hombres, con rango de edad entre 15 y 59 años ($M=27,5$; $DP=9,01$). Los participantes fueron divididos entre músicos ($N=24$), aficionados ($N=62$) y laicos ($N=76$) de acuerdo con el nivel de conocimiento en música. Por medio del análisis *Full Information Factor Analysis* se verificó la dimensionalidad del instrumento y también las propiedades de los ítems a través de la Teoría de Respuesta al Ítem (TRI). Además, se buscó identificar la capacidad de discriminación entre los grupos de músicos y no músicos. Los datos encontrados apuntan evidencias de que los ítems miden una dimensión principal ($\alpha=0,92$) con alta capacidad para diferenciar los grupos de músicos profesionales, aficionados y laicos obteniéndose un coeficiente de validez de criterio de $r=0,68$. Los resultados indican evidencias positivas de precisión y validez para el RAu.

Palabras clave: Procesamiento auditivo; Teoría de Respuesta al Ítem; Músicos; Evaluación de la inteligencia; Análisis factorial por información completa.

Aspectos compartilhados por habilidades atreladas à música e a diferentes tipos de raciocínio abrem um vasto campo de investigação dentro da área de estudo da avaliação da inteligência (Billhartz, Bruhn & Olson, 1999; Helmbold, Rammayer & Altenmüller, 2005; Schellenberg, 2004b), embora essa temática ainda seja pouco explorada na literatura (Cunha, 2007). Os estudos que buscam verificar relações entre algum tipo

de inteligência e música dividem-se, basicamente, em duas áreas (Schellenberg, 2003). A primeira tem como objetivo verificar os efeitos em curto prazo decorrentes de se ouvir músicas, como, por exemplo, o efeito Mozart (Hetland, 2000; Rauscher, Shaw & Ky, 1993); e a segunda tem como princípio investigar os efeitos em longo prazo relacionados ao estudo da música, ou seja, tratam da relação entre o aprendizado da música no

raciocínio e construtos relacionados (Franklin & cols., 2008; Schellenberg, 2004a).

O aprendizado da música, bem como os aspectos relacionados ao desenvolvimento de habilidades musicais, têm sido investigados por profissionais da área da educação musical, interessados em compreender os processos da inteligência que são subjacentes à aprendizagem da música. Algumas dessas especulações supõem que a inteligência geral é a base para o desenvolvimento de habilidades musicais (Radocy & Boyle, 1979; Shuter-Dyson & Gabriel, 1981). Diferentemente, alguns pesquisadores sugerem o uso do termo inteligência acadêmica, considerando a habilidade musical como um conhecimento especializado (Schoen, 1940; Wing, 1941) e, ainda, que habilidade musical seria componente da inteligência, mas independente de outras habilidades cognitivas (Gardner, 1983).

Sobre essa temática, Schellenberg (2004a) defende que o indivíduo pode exibir ganhos em relação à inteligência, que são provenientes das habilidades adquiridas em exposição às aulas de música, para os domínios cognitivos não-musicais, tais como a linguagem, a matemática e as habilidades visuoespaciais. Segundo o autor, as aulas de música parecem impulsionar muitas habilidades gerais, como, por exemplo, maior sensibilidade para lidar com sinais e informações do tipo gestálticas, habilidades motoras finas, que podem tender a transferirem-se a outros domínios cognitivos não-musicais, e, ainda, habilidades relevantes relacionadas ao sistema auditivo.

No que se refere ao sistema auditivo, Deutsch (1999) relata que a música fornece um espectro de mudanças acústicas rápidas e complexas, geralmente originadas da sobreposição de sons provenientes de diversas fontes. O sistema auditivo humano tem a tarefa de analisar o espectro e reconstruir a origem dos eventos sonoros, analogamente ao que acontece no nosso sistema visual, quando o mesmo interpreta um mosaico de luzes. Esse processo, chamado por Helmholtz (1859/1954) de inferência inconsciente, foi proposto no século XIX e muitos fenômenos relacionados às tarefas envolvendo o sistema auditivo podem ser compreendidos dessa maneira.

Exemplo disso é a relação de treinos musicais com domínios cognitivos não-musicais. O treino musical envolve longos períodos de atenção focada, prática diária, leitura de notação musical, memorização de trechos musicais extensos e a *expertise* em habilidades técnicas, tais como as motoras (Helmbold & cols., 2005). Supõe-se que essa combinação de processos cognitivos possa ter um impacto significativo no raciocínio, particularmente na infância, e que se estenda ao longo da vida (Schellenberg, 2004a).

Em estudo realizado por Bilhartz e colaboradores (1999), buscou-se investigar fatores relacionados às habilidades musicais e, mais especificamente, a relação dessas habilidades com o raciocínio espacial. Os autores sugeriram que o raciocínio espacial requer a habilidade de manter imagens mentais sem a ajuda de um modelo físico e combinar ou organizar essas imagens mentalmente. Assim, o processo mental de organizar elementos sequencial e espacialmente é muito utilizado em atividades mentais complexas e também por músicos, no desempenho em tarefas musicais.

Acerca dos mecanismos que determinam a organização mental de elementos, de acordo com Sloboda (2004), a principal característica da música é que, na informação musical, os sons estão sempre relacionados uns aos outros, nunca isolados. Para acontecer a percepção e o processamento cognitivo da música, as pessoas precisam perceber relações e identificar agrupamentos significativos dos sons que escutam. Em um nível elevado, tais agrupamentos podem ser complexos e extensos, resultando em agrupamentos compostos, por exemplo, por diversas notas musicais em relação. De outro modo, em um nível mais baixo, os agrupamentos podem ser relativamente simples, caracterizados por um padrão sonoro de pequenos conjuntos de notas musicais. Em ambos os casos, a percepção e o processamento dos estímulos sonoros estão relacionados com as habilidades auditivas.

Essas habilidades, de acordo com Stankov (1994), podem ser entendidas como habilidades cognitivas específicas, isto é, que dependem do som como *input*, do funcionamento do aparato auditivo humano e do córtex auditivo, responsável pelo processamento destas informações. As habilidades auditivas, então, refletem o nível em que as pessoas podem cognitivamente processar (memorizar, alterar, recuperar) e controlar a percepção dos inputs provenientes dos estímulos sonoros (Gustafsson & Undheim, 1996).

Essa definição assemelha-se ao conceito de processamento auditivo (*Ga*) encontrado no modelo das habilidades cognitivas de Carrol-Horn-Cattell (CHC). O CHC é um modelo hierárquico em três níveis ou estratos, caracterizado pelo progressivo aumento da especialização das capacidades cognitivas, partindo de um fator mais amplo, fator *g*, no estrato III, e seguindo em direção a fatores específicos, nos estratos I e II (McGrew, 2009). É utilizado como fundamentação para diferentes baterias de testes de inteligência em razão de sua concepção multidimensional que engloba diversos tipos de raciocínio, bem como as habilidades específicas ligadas a esses instrumentos (Flanagan & Ortiz, 2001).

O *Ga* é o processamento de *inputs* sonoros e está ligado à capacidade associada à percepção, análise e síntese de padrões sonoros, incluindo a linguagem oral, bem como a percepção de nuances em estruturas musicais complexas (Primi, 2003). Essa capacidade relaciona-se ao quão precisa e eficaz é esta manipulação de sons (Newton & McGrew, 2010). Nesse sentido, reflete o nível em que um indivíduo pode processar cognitivamente e controlar a percepção dos *inputs* provenientes de estímulos sonoros (Gustafsson & Undheim, 1996). De fato, vários autores (por exemplo, Cunha, 2007; Forgeard, Schlaug, Norton, Rosam & Iyengar, 2008; Gaab & Schlaug, 2003; Wallentin, Nielsen, Friis-Olivarius, Vuust & Vuust, 2010) que se interessam pelo assunto dividem suas amostras entre grupos de músicos e não-músicos, baseados na hipótese de que pessoas com formação em música apresentam maior capacidade no processamento auditivo em virtude de uma prática frequente e sistemática.

Com o objetivo de mensurar a habilidade musical, alguns autores desenvolveram instrumentos contendo estímulos musicais. Seashore (1942) buscou formular uma base psicológica da música definindo e mensurando habilidades específicas, como frequência, intensidade, duração e formato de onda, por meio do *Seashore's Measure of Musical Talents*, sendo considerado um ponto de partida para o estudo de estruturas mais complexas, como melodia, harmonia e ritmo. Edwards, Challis, Hankinson e Pirie (2000) viabilizaram um teste com a finalidade de mensurar a habilidade musical apresentando uma estrutura mais complexa, contendo ritmo, harmonia e tonalidade. Nele, o sujeito deve ouvir o estímulo duas vezes e em seguida escolher, dentre as possíveis respostas, a que mais se assemelha. O teste se demonstra longo, com aproximadamente 200 itens, e não foram encontrados registros sobre seu nome.

Peretz, Champod e Hyde (2003) desenvolveram a Bateria Montreal de Avaliação da Amusia (MBEA), incorporando seis subtestes, sendo a tarefa a ser executada semelhante ao descrito anteriormente, ou seja, o sujeito deve julgar se dois estímulos apresentados são similares ou não. O *Gordon's "Advanced Measure of Musical Audiation"* – AMMA (Gordon, 2007), considerado um teste para avaliar o talento musical de forma específica, compreende em sua estrutura mudanças melódicas e rítmicas que o sujeito deve discriminar, como, por exemplo, identificar a dissonância presente em um acorde. Com a finalidade de criar um teste mais curto e que pudesse avaliar a habilidade musical e ser aplicável a músicos e não-músicos, Wallentin e cols. (2010) desenvolveram o *Musical Ear Test* (MET), composto de 104 itens

divididos, igualmente, em dois subtestes, sendo um de frases melódicas e outro de frases rítmicas. A tarefa a ser desempenhada é discriminar, entre dois estímulos apresentados, se são ou não semelhantes.

Os testes mencionados foram desenvolvidos com propósitos específicos no contexto da educação musical. Pode-se fazer uma analogia com os tipos de métodos de avaliação da inteligência ora focalizados na avaliação de desempenho acadêmico, entendido como conhecimento estruturado a partir de experiências de aprendizagem em um domínio específico – ora focalizado na medida de processos cognitivos –, como os testes de inteligência que buscam avaliar raciocínio, memória, processamento visual e auditivo, dentre outros (Ackerman, 1996). Os testes mencionados vêm de uma tradição pedagógica, derivados de uma matriz em que a habilidade musical é entendida como conhecimentos específicos organizados. O presente estudo apresenta o desenvolvimento do Teste de Raciocínio Auditivo Musical – RAu (Cunha & cols., 2006, descrito na seção instrumentos), que segue uma tradição psicológica – psicométrica e cognitiva –, buscando-se medir os processos de raciocínio geral aplicados a estímulos de conteúdo musical, mas cujas tarefas se assemelham a testes de inteligência como a Bateria de Provas de Raciocínio (BPR-5, Primi & Almeida, 2000, Primi, Silva, Santana, Muniz & Almeida, no prelo). O teste RAu pretende mensurar o processamento auditivo segundo o modelo CHC, buscando-se preencher uma lacuna existente quanto à disponibilidade de testes de avaliação cognitiva desse fator amplo descrito no modelo. Especificamente, este estudo objetiva fazer uma investigação inicial das evidências de validade com base na estrutura interna dos itens criados para o RAu e sua relação com critério externo: o nível de treinamento musical recebido.

Método

Participantes

Participaram deste estudo 162 pessoas de ambos os sexos, sendo 56,8% homens. A faixa etária variou de 15 a 59 anos ($M=27,5$; $DP=9,01$). Os participantes possuíam diferentes níveis de conhecimento sobre música, sendo classificados como músicos (G1; $N=24$), amadores (G2; $N=62$) e leigos (G3; $N=76$). O critério para inclusão do G1 foi ser estudante ou formado em curso universitário de música. Para G2, os indivíduos deveriam ter algum tipo de experiência amadora com música como atividade de lazer ou *hobby*. Por fim, para G3, os participantes deveriam não ter qualquer experiência relacionada ao estudo musical.

Instrumentos

Teste de Processamento Auditivo com Estímulos Musicais – Rau (Cunha & cols., 2006)

Inicialmente, para a construção do RAu, foram utilizados instrumentos musicais e dois *softwares* de gravação e edição de áudio. Os instrumentos foram conectados a uma placa de som ligada ao computador, para a gravação dos itens. Ao final da gravação e consequente criação dos itens, os mesmos foram transformados no formato *wave sound (.wav)*, para então ser colocados no *software* final. Assim, foram desenvolvidos 54 itens, 20 com tarefas do tipo seriação e 34 com tarefas do tipo analogia, divididos em 3 subtestes. Os 20 itens do tipo seriação compuseram o subteste Séries. Por sua vez, os 34 itens do tipo analogia foram distribuídos em dois subtestes separados: Analogias 1 (20 itens) e Analogias 2 (14 itens). Vale ressaltar que, além desses itens, foram criados três para cada subteste, para serem utilizados como exemplos nas instruções. Os itens foram agrupados no *software* desenvolvido especificamente para esta pesquisa.

Para todos os subtestes, os itens foram construídos com base nas concepções teóricas do raciocínio indutivo, mas foram delineados por meio do sistema Embretson de desenho cognitivo das tarefas (Embretson, 1994). Esse sistema determina os passos para a criação de itens de testes com base na teoria da psicologia cognitiva sobre o construto. A partir desses pressupostos, o delineamento dos itens foi realizado de acordo com a relação entre a quantidade de elementos (melodias ou acordes) apresentados e a quantidade de transformações sofridas (isto é, mudanças sofridas pelos elementos). Exemplo de um item composto por elemento sofrendo transformação seria uma melodia sendo apresentada e posteriormente sendo tocada numa tonalidade mais aguda.

Para os subtestes Séries e Analogias 1, os 10 itens inicialmente desenvolvidos para cada um possuíam uma estrutura, ou seja, um tipo de pergunta e três possíveis respostas. Nos itens desses subtestes havia apenas um elemento e uma transformação, variando-se basicamente o tipo de transformação, como, por exemplo, finalizações de melodias e transposição de intervalos. Esses itens foram manipulados em relação aos timbres, para modificar os seus níveis de dificuldade. Ou seja, enquanto os 10 itens criados inicialmente possuíam, cada um deles, um mesmo timbre na pergunta e nas respostas (por exemplo, a pergunta sendo tocada com o som de um piano e as possíveis respostas também), os mesmos itens foram manipulados para que fossem inseridos timbres diferentes (uma flauta na pergunta e um piano, um som sintetizado e um trompete como os tipos de timbres

das opções de resposta), ou seja, um timbre para a pergunta e um timbre diferente para cada uma das três alternativas de resposta. Dessa forma, em cada teste, 10 itens foram transformados em 20.

Em relação a Analogias 2, foram inicialmente desenvolvidos 7 itens, posteriormente manipulados em relação aos seus timbres para totalizar 14. Nesse caso, havia a presença de mais de um elemento em alguns itens, portanto, a manipulação de timbres foi diferente e a estrutura do subteste se apresentou de outra forma, a saber, o sujeito deveria escutar duas melodias distintas, havendo um botão para cada uma, e encontrar, em quatro alternativas possíveis, as que apresentavam o mesmo padrão (analogia) identificado no item. Para aqueles com apenas um elemento, seguiu-se o mesmo critério usado em Séries e Analogias 1. Já para os itens com mais de um elemento, utilizaram-se timbres distintos entre os elementos. Por exemplo, num item composto por duas melodias sofrendo uma alteração, uma poderia ser tocada por um piano e a outra por uma flauta. As opções de respostas teriam os mesmos timbres da pergunta.

Procedimentos e análise dos dados

Após a aprovação do projeto pelo Comitê de Ética da Universidade São Francisco, os sujeitos assinaram o Termo de Consentimento Livre e Esclarecido (TCLE) e, para aqueles menores de idade, solicitou-se a autorização dos responsáveis. As coletas se deram em instituições de ensino superior, escolas de música ou residências dos participantes. Ocorreram de modo individual e coletivo e foram realizadas por meio de computadores, tanto de uso pessoal dos pesquisadores quanto dos participantes, observando sempre que as máquinas estivessem equipadas com caixas de som funcionando. Não houve tempo mínimo ou máximo para responder ao instrumento e ao questionário.

Para atingir o objetivo deste estudo, foram realizadas análises utilizando procedimentos da Teoria de Resposta ao Item (TRI). Em um primeiro momento, na busca por evidências de validade com base na estrutura interna e pela consistência interna, realizou-se uma análise fatorial dos itens do instrumento por informação completa com o emprego do programa TESTFACT (Wilson, Wood & Gibbons, 1991). Esse programa emprega algoritmos da Teoria de Resposta ao Item para análise de vetores de resposta para extrair os parâmetros da análise fatorial a partir do parâmetro de discriminação da TRI. Ele também emprega análises iniciais de matrizes de correlação tetracórica para evitar problemas conhecidos que ocorrem ao se analisarem matrizes de correlação Pearson com variáveis binárias (acerto e erro) que

tenham uma heterogeneidade de distribuição (Primi & Almeida, 1998). Essa análise buscou verificar a estrutura dimensional subjacente aos itens. O TESTFACT emprega um modelo de TRI de três parâmetros, mas não estima os valores do parâmetro c , que precisam ser informados. Assim, uma análise prévia foi feita no programa XCALIBRE 4.1 (Guyer & Thompson, 2011) para se estimarem os parâmetros segundo o modelo logístico de 3 parâmetros, dos quais os valores de c foram usados na calibração feita pelo TESTFACT. Para manter a comparabilidade com o modelo de Rasch, que foi estimado em seguida, a constante D foi mantida em 1 para se ter um modelo logístico, e, além disso, centralizaram-se as escalas com média de b (índices de dificuldade) igual a zero. Na sequência, empregou-se o modelo de Rasch para analisar e calibrar os parâmetros de dificuldade e de habilidade dos sujeitos. Também foi investigado o ajuste dos itens com o emprego desse modelo. Finalmente, efetuou-se uma ANOVA buscando-se comparar o desempenho dos grupos testando-se a validade de critério.

Resultados e discussão

Na busca por evidências de validade com base na estrutura interna, por meio da análise fatorial por informação completa (*Full Information Factor Analysis*), foram verificadas as soluções com base em 1, 2 e 3 fatores, limite estabelecido de acordo com o número de subtestes do instrumento. Considerando as magnitudes moderadas e altas de correlação entre os fatores, optou-se pela extração de somente 1 fator, solução essa também indicada pela análise do gráfico *scree plot*, cujos valores dos primeiros cinco *eigenvalues* foram: 17,3; 3,2; 2,4; 2,3; 2,12, sendo que o primeiro *eigenvalue* é 5,3 vezes maior que o segundo, sugerindo a presença de um fator principal. Esse fator explicou 45,23% da variância total e apresentou alto índice de consistência

interna que, extraído pelo *software* TESTFACT, foi igual a 0,92.

Considerando o delineamento proposto neste estudo, não é possível aprofundar-se acerca da interpretação do fator encontrado. Sendo assim, outras pesquisas devem buscar esclarecer esse aspecto a partir de outros delineamentos de pesquisa. Estudos que correlacionem o RAu com medidas de inteligência fluida (Gf), bem como com marcadores de inteligência cristalizada (Gc), poderão fornecer maiores esclarecimentos acerca do significado do fator que se destacou na análise fatorial aqui apresentada.

Na sequência, foram realizadas análises buscando calibrar os parâmetros dos itens do instrumento por meio do modelo de Rasch (ou modelo de 1 parâmetro), que tem como base a Teoria de Resposta ao Item (TRI). Para o conjunto de itens como um todo, o índice de fidedignidade encontrado foi de 0,89 (e o alfa de Cronbach calculado foi de 0,91), indicando alta consistência entre os itens e um nível aceitável de erro.

Sobre a análise dos itens, apresentam-se na Tabela 1 os parâmetros psicométricos obtidos nas três análises efetuadas. A primeira coluna apresenta o nome dos itens (GAS: séries; GL1: Analogias 1; GL2: Analogias 2) e, em seguida, os parâmetros psicométricos clássicos, índice de dificuldade (ID) e correlação item-total (r). Nas três próximas colunas, são apresentados os parâmetros de dificuldade (b) da análise de Rasch calibrados pelo WINSTEPS e os índices de ajuste *infit* e *outfit*. Em seguida, são apresentados os parâmetros a , b e c do modelo de 3 parâmetros da TRI calibrados no XCALIBRE e, nas últimas duas colunas, a comunalidade e a carga fatorial final obtida no TESTFACT na dimensão extraída. Nota-se, nas análises feitas no XCALIBRE e TESTFACT, a eliminação de alguns itens que não atingiram os critérios definidos.

Tabela 1. Parâmetros psicométricos dos itens do RAu

(Continua)

Item	ID	r	b	$Infit$	$Outfit$	a	b	c	h^2	Carga
GAS01	0,83	0,31	-1,12	1,01	1,13	1,08	-1,13	0,17	0,22	0,47
GAS02	0,76	0,49	-0,59	0,87	0,68	1,54	-0,44	0,17	0,63	0,79
GAS03	0,90	0,35	-1,79	0,92	0,61	1,46	-1,47	0,17	0,56	0,75
GAS04	0,86	0,39	-1,34	0,90	0,69	1,43	-1,12	0,17	0,53	0,73
GAS05	0,85	0,42	-1,28	0,86	0,71	1,45	-1,08	0,16	0,48	0,70
GAS06	0,56	0,24	0,50	1,25	1,27	0,82	0,78	0,18	0,10	0,31
GAS07	0,49	0,34	0,87	1,14	1,15	1,31	1,30	0,19	0,93	0,96
GAS08	0,59	0,15	0,34	1,31	1,61	0,69	0,52	0,17	0,02	0,16
GAS09	0,60	0,31	0,31	1,16	1,16	0,88	0,45	0,17	0,18	0,42
GAS10	0,47	0,31	0,97	1,16	1,25					
GAS11	0,60	0,36	0,31	1,10	1,05	1,03	0,49	0,18	0,28	0,53
GAS12	0,69	0,50	-0,19	0,89	0,77	1,50	-0,09	0,16	0,63	0,79

Tabela 1. Parâmetros psicométricos dos itens do RAu (Continuação)

Item	ID	r	b	Infit	Outfit	a	b	c	h ²	Carga
GAS13	0,96	0,21	-2,73	0,98	0,69	1,31	-2,31	0,17	0,23	0,48
GAS14	0,84	0,33	-1,17	0,98	0,92	1,18	-1,11	0,17	0,36	0,60
GAS15	0,81	0,38	-0,93	0,95	0,86	1,17	-0,89	0,17	0,33	0,57
GAS16	0,62	0,34	0,21	1,11	1,12	0,92	0,33	0,17	0,18	0,43
GAS17	0,44	0,54	1,09	0,88	0,91	1,88	1,15	0,17	0,92	0,96
GAS18	0,39	0,24	1,38	1,27	1,31					
GAS19	0,51	0,31	0,75	1,17	1,21	0,94	1,09	0,18	0,16	0,40
GAS20	0,40	0,44	1,32	1,03	1,00					
GL101	0,79	0,45	-0,81	0,90	0,68	1,47	-0,65	0,17	0,59	0,77
GL102	0,27	0,57	2,08	0,81	0,90	1,90	1,80	0,14	0,83	0,91
GL103	0,69	0,42	-0,15	0,97	0,95	1,12	-0,12	0,17	0,33	0,58
GL104	0,92	0,30	-2,00	0,94	0,60	1,37	-1,70	0,17	0,43	0,66
GL105	0,49	-0,12	0,85	1,66	2,01					
GL106	0,64	0,51	0,08	0,89	0,84	1,42	0,14	0,16	0,56	0,75
GL107	0,77	0,40	-0,68	0,92	1,04	1,15	-0,68	0,16	0,30	0,55
GL108	0,66	0,29	0,02	1,13	1,27	0,86	0,04	0,17	0,14	0,37
GL109	0,87	0,31	-1,42	1,01	0,74	1,18	-1,32	0,17	0,33	0,58
GL110	0,69	0,49	-0,15	0,90	0,79	1,38	-0,07	0,17	0,55	0,74
GL111	0,57	0,56	0,44	0,86	0,77	1,50	0,50	0,16	0,67	0,82
GL112	0,32	0,38	1,78	1,01	1,33	1,95	1,89	0,16	0,89	0,94
GL113	0,70	0,41	-0,22	0,96	1,05	1,05	-0,23	0,17	0,29	0,54
GL114	0,94	0,34	-2,30	0,88	0,48	1,63	-1,76	0,17	0,57	0,76
GL115	0,57	-0,13	0,47	1,62	2,13					
GL116	0,68	0,29	-0,12	1,14	1,11	0,93	-0,05	0,17	0,18	0,42
GL117	0,80	0,45	-0,85	0,85	0,77	1,39	-0,75	0,16	0,44	0,67
GL118	0,79	0,20	-0,81	1,10	1,61	0,84	-1,02	0,17	0,10	0,31
GL119	0,76	0,45	-0,60	0,93	0,71	1,47	-0,44	0,17	0,59	0,77
GL120	0,75	0,36	-0,52	1,01	1,00	1,04	-0,53	0,17	0,24	0,49
GL201	0,63	0,55	0,14	0,84	0,78	1,58	0,18	0,16	0,61	0,78
GL202	0,60	0,59	0,30	0,80	0,73	1,62	0,28	0,16	0,62	0,79
GL203	0,47	0,59	0,97	0,83	0,79	1,66	0,94	0,16	0,70	0,84
GL204	0,65	0,49	0,04	0,89	0,92	1,26	0,07	0,16	0,44	0,67
GL205	0,39	0,38	1,36	1,10	1,15	1,16	1,63	0,16	0,35	0,59
GL206	0,64	0,56	0,10	0,83	0,73	1,59	0,14	0,16	0,64	0,80
GL207	0,64	0,42	0,10	1,01	0,93	1,22	0,22	0,17	0,39	0,62
GL208	0,67	0,49	-0,10	0,91	0,82	1,39	0,00	0,17	0,53	0,73
GL209	0,58	0,64	0,37	0,75	0,67	1,94	0,34	0,15	0,76	0,87
GL210	0,26	0,49	2,13	0,91	0,98	1,64	2,10	0,15	0,32	0,56
GL211	0,53	0,58	0,62	0,84	0,78	1,82	0,64	0,16	0,50	0,71
GL212	0,39	0,61	1,36	0,81	0,76	1,81	1,26	0,15	0,50	0,70
GL213	0,72	0,53	-0,39	0,83	0,68	1,64	-0,27	0,16	0,56	0,75
GL214	0,46	0,62	1,00	0,79	0,74	1,74	0,92	0,15	0,52	0,72

Adotando-se como critério de desajuste os valores acima de 1,3 (Linacre & Wright, 1994; Smith, 1996; Wright & Linacre, 1994), os resultados apontam que a

maioria dos itens apresentaram níveis de *infit* e *outfit* adequados, isto é, inferiores a 1,3 e correlação item-total variando entre -0,13 e 0,62, em ambos os casos

indicando adequação e consistência interna entre os itens. Entretanto, cinco itens apresentaram magnitude inferior a 0,30 para correlação item-total, dois dos quais (5 e 15 do subteste Analogias 1) apresentaram magnitude negativa, indicando uma idiosincrasia nesses itens, já que não se alinham com os restantes do teste. Vale lembrar que o conteúdo desses dois itens é o mesmo, variando apenas os timbres utilizados, conforme descrito na seção instrumentos. Os itens que apresentaram valores de *infit* e *outfit* elevados e magnitudes de correlação negativa serão descartados da versão final do RAU.

Na Figura 1 apresenta-se o mapa de itens ou mapa do construto obtido no programa WINSTEPS, representado pela escala do construto na posição vertical, com a distribuição dos sujeitos ao lado esquerdo e dos itens ao lado direito, posicionados em função de seu índice de dificuldade. Pode-se verificar o quanto os itens representam os diversos níveis do construto e a posição da amostra em relação aos itens. Os valores M, S e T indicam, respectivamente, a média, 1 e 2 desvios padrão do teta (esquerda) e dos índices de dificuldade (direita). Nota-se uma boa distribuição dos itens ao longo da escala, porém há uma quantidade grande de itens fáceis cobrindo a região de -1 DP e uma lacuna na região acima de +2 DP. É interessante notar que, embora o RAU seja um teste que consiga diferenciar pessoas com educação formal em música (como se verá adiante), ao mesmo tempo ele consegue mensurar mais precisamente pessoas leigas – com baixo nível de habilidade – sua habilidade de processamento auditivo. Isso sugere que o teste foi bem sucedido em operacionalizar um componente de raciocínio e processamento auditivo que não dependa de conhecimento organizado adquirido em educação formal em música.

No que concerne aos índices de dificuldade, o delineamento proposto para os itens do RAU pressupunha que as alterações de timbre poderiam interferir no nível de dificuldade, sendo ele considerado um elemento a mais para ser codificado pelo sujeito. Sendo assim, a partir de uma estrutura de item semelhante, a mudança de timbre poderia provocar mudança no grau de dificuldade da tarefa. De acordo com os resultados, foi possível observar que, para os itens do subteste Séries, a mudança do timbre aumentou o nível de dificuldade de 8 itens; em Analogias 1, a mudança no nível de dificuldade foi observada em apenas 2 desses itens e, em Analogias 2, esse efeito foi observado em 5 dos itens. Apesar desse efeito em relação ao timbre não ter sido observado em todos os itens, é possível verificar que essas alterações pareceram correlacionar-se ao nível de dificuldade dos itens desenvolvidos para o RAU.

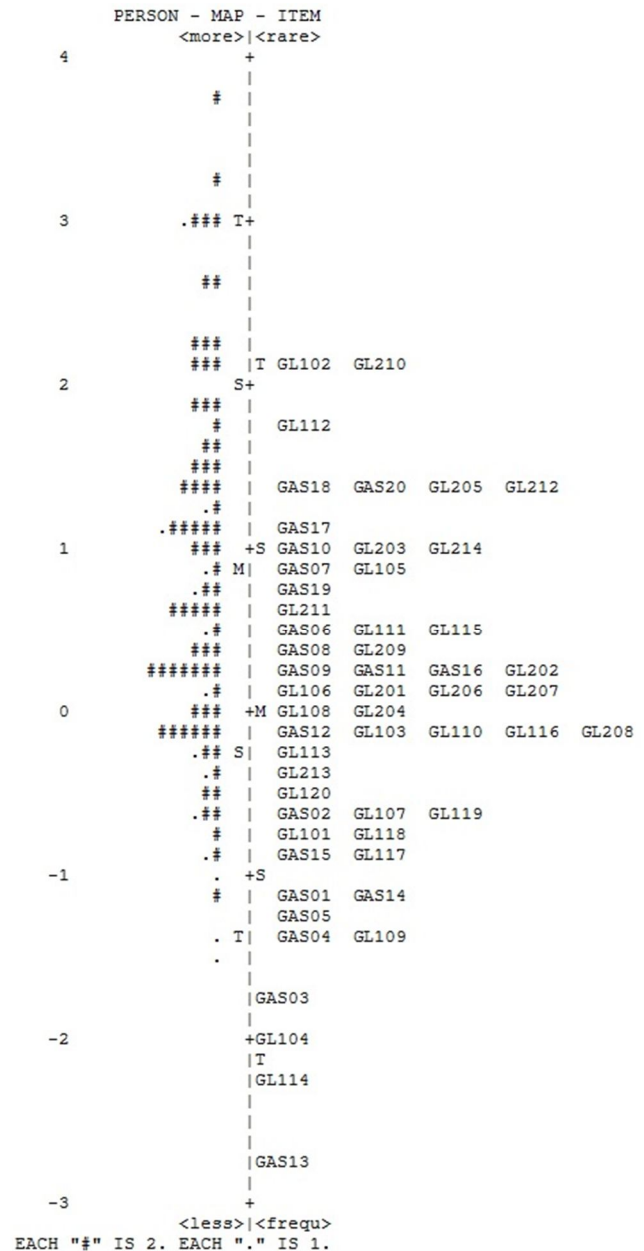


Figura 1. Mapa do construto do teste RAU.

Ainda, na continuidade, hipotetizou-se que músicos teriam notas mais elevadas no RAU em razão de seu conhecimento específico. Como os componentes principais das tarefas do RAU são o raciocínio – descoberta das relações entre os estímulos – e o conhecimento específico – associado ao conteúdo musical dos estímulos –, esperava-se que os músicos tivessem maior nota por se saírem melhor no componente específico. A análise de variância (ANOVA), tendo como variável dependente o teta dos sujeitos obtidos pelo modelo de Rasch e a variável grupo, testando-se as diferenças de média entre músicos, amadores e leigos, confirmou a diferença

significativa entre o desempenho dos três grupos no RAu [$F=63,975$; $p<0,01$; η^2 parcial= $0,44$], indicando que 44% da variância das notas é explicada pela

variável grupo do participantes, o que equivale a um coeficiente de validade de critério de 0,67 considerado muito alto (Hemphill, 2003; Meyer & cols., 2001).

Tabela 2. Dados descritivos dos grupos

Grupos	N	Média	DP	Mínimo	Máximo
G1	24	3,00	1,58	-0,54	5,86
G2	62	1,64	1,08	-1,14	3,87
G3	76	0,40	0,74	-1,37	2,58
Total	162	1,26	1,38	-1,37	5,86

Complementando os dados encontrados na ANOVA, foi verificado o teste *post hoc* (*Student-Newman-Keuls*) para investigação das diferenças entre cada um dos três grupos componentes da amostra. Foi possível verificar que os três grupos se distinguem significativamente no RAu ($p=0,05$). Com base nesses dados, é possível identificar a existência de um padrão linear de diferença entre leigos, amadores e músicos, formando-se três grupos distintos de acordo com o nível de estudo em música, isto é, os indivíduos do grupo de músicos apresentaram média superior no instrumento em relação aos demais grupos; o grupo de amadores, média intermediária; e o grupo de leigos, a menor média, disposição entre grupos que era esperada para um instrumento que avalia o processamento auditivo. Embora o teste tenha sido relativamente mais fácil para a amostra, ele mostrou possuir uma boa capacidade de diferenciar níveis gradativos de contato com música, indicando que, além do raciocínio geral e processamento auditivo, o teste operacionalizou um componente específico associado ao conhecimento adquirido em experiência formal com a música. Nesse sentido, os resultados indicam evidências positivas de validade de critério, na medida em que o RAu foi capaz de separar, pelo desempenho, os grupos com diferentes graus de treinamento/instrução em música. Mesmo não sendo possível chegar a uma conclusão mais detalhada, pôde-se demonstrar que há diferenças no desempenho nas tarefas cognitivas propostas pelo instrumento quando são comparadas pessoas que foram expostas à instrução musical com pessoas que não foram expostas.

Considerações finais

A habilidade musical é um construto ainda pouco estudado na Psicologia. Nesse sentido, o processamento auditivo (*Ga*) apresentado na estrutura do modelo CHC pode ser considerado um elemento-chave na compreensão desses processos, por se tratar essencialmente de processamento cognitivo a partir de estímulos sonoros, constituindo uma característica fundamental para um bom desempenho musical. Por

isso, novos instrumentos voltados à avaliação desse fator têm uma relevância muito grande.

Os resultados deste estudo preliminar do RAu são promissores ao indicar que as tarefas criadas avaliam uma dimensão principal (estrutura interna) e apresentaram consistência e homogeneidade adequadas. Além disso, a escala formada pelos itens foi capaz de discriminar os grupos de participantes com diferentes graus de aprendizagem formal em música. Assim, conclui-se que o RAu apresentou boas qualidades psicométricas, podendo-se constituir em uma importante ferramenta para a pesquisa e também para o trabalho profissional, tanto como um teste cognitivo do fator *Ga* como também como uma ferramenta para estudo da habilidade musical. Evidentemente, novos estudos são necessários, especialmente aqueles que apliquem o RAu com outros testes representando vários fatores amplos da inteligência segundo o modelo CHC, para que se possa aprender mais sobre sua validade de construto.

Referências

- Ackerman, P. L. (1996). A theory of intellectual development: process, personality, interests and knowledge. *Intelligence*, 22, 227-257.
- Billhartz, T. D., Bruhn, R. A., & Olson, J. E. (1999). The effect of early music training on child cognitive development. *Journal of Applied Developmental Psychology*, 20(4), 615-636.
- Cunha, T. F. (2007). *Desenvolvimento de um Teste de Processamento Auditivo com estímulos musicais* (Dissertação de mestrado). Universidade São Francisco, Itatiba, São Paulo, Brasil.
- Cunha, T. F., Primi, R., Berberian, A., Ambiel, R. A. M., Pessotto, F., & Miguel, F. K. (2006). *Teste de Processamento Auditivo com Estímulos Musicais*. Itatiba: Laboratório de Avaliação Psicológica e Educacional – LabAPE, Universidade São Francisco.

- Deutsch, D. (1999). Grouping mechanisms in music. Em D. Deutsch (Org.), *The Psychology of Music* (pp. 299-348). Nova Iorque: Academic Press.
- Edwards, A. D. N., Challis, B. P., Hankinson, J. C. K., & Pirie, F. L. (2000). Development of a standard test of musical ability for participants in auditory interface testing. Em P. R. Cook (Ed.), *Proceedings of the 2000 International Conference on Auditory Display (ICAD)*. Georgia Institute of Technology: International Community for Auditory Display.
- Embretson, S. (1994). Applications of cognitive design systems to test development. Em C. R. Reynolds (Ed.), *Cognitive assessment: a multidisciplinary perspective* (107-135). Nova Iorque: Plenum Press.
- Flanagan, D. P., & Ortiz, S. O. (2001). *Essentials of cross-battery assessment*. Nova Iorque: John Willey & Sons, Inc.
- Forgeard, M., Schlaug, G., Norton, A., Rosam, C., & Iyengar, U. (2008). The relation between music and phonological processing in normal-reading children and children with dyslexia. *Music and Phonological Processing*, 25(4), 383-390.
- Franklin, M. S., Rattray, K., Moore, K.S., Moher, J., Yip, C. Y., & Jonides, J. (2008). The effects of musical training on verbal memory. *Psychology of Music*, 36(3), 353-365.
- Gaab, N., & Schlaug, G. (2003). Musicians differ from nonmusicians in brain activation despite performance matching. *Annals of the New York Academy of Sciences*, 999, 385-388.
- Gardner, H. (1983). *Frames of mind*. Nova Iorque: Basic Books.
- Gordon, E. E. (2007). *Learning sequences in music: a contemporary music learning theory*. Chicago: GIA Publications.
- Guyer, R., & Thompson, N. A., (2011). *User's manual for XCalibre 4.1*. St. Paul MN: Assessment Systems Corporation.
- Gustafsson, J. E., & Undheim, J. O. (1996). Individual differences in cognitive functions. Em D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 186-242). Nova Iorque: Macmillan.
- Helmbold, N., Rammayer, T., & Altenmüller, E. (2005). Differences in primary mental abilities between musicians and nonmusicians. *Journal of Individual Differences*, 26(2), 74-85.
- Helmholtz, H. von (1859/1954). *On the sensations of tones as a physiological basis for the theory of music*. (2ª ed.). Nova Iorque: Dover.
- Hemphill, J. F. (2003). Interpreting the magnitudes of correlation coefficients. *American Psychologist*, 58, 78-79.
- Hetland, L. (2000). Listening to music enhances spatial-temporal reasoning: evidence for the "Mozart effect." *Journal of Aesthetic Education*, 34(3-4), 105-148.
- Linacre, J. M., & Wright, B. D. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8(2), 370.
- McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, 37(1), 1-10.
- Meyer, G. J., Finn, S. E., Eyde, L., Kay, G. G., Moreland, K. L., Dies, R. R., Eisman, E. J., Kubiszyn, T. W., & Reed, G. M. (2001). Psychological testing and psychological assessment: a review of evidence and issues. *American Psychologist*, 56, 128-165.
- Newton, J. H., & McGrew, K. S. (2010). Introduction to the special issue: current research in Cattell-Horn-Carroll-based assessment. *Psychology in the Schools*, 47(7), 621-634.
- Peretz, I., Champod, A. S., & Hyde, K. (2003). Varieties of musical disorders. The Montreal Battery of Evaluation of Amusia. *Annals of the New York Academy of Sciences*, 999, 58-75.
- Primi, R. (2003). Inteligência: avanços nos modelos teóricos e nos instrumentos de medida. *Avaliação Psicológica*, 1, 67-77.
- Primi, R., & Almeida, L. (1998). Considerações sobre a análise fatorial de itens com resposta dicotômica. *Psicologia Teoria Investigação e Prática*, 3, 225-234.
- Primi, R., & Almeida, L. S. (2000). *Baterias de Provas de Raciocínio (BPR-5): manual técnico*. São Paulo: Casa do Psicólogo.
- Primi, R., Silva, M. C. R., Santana, P. R., Muniz, M., & Almeida, L. S. (no prelo). The use of the bi-factor model to test the uni-dimensionality of a battery of reasoning tests. *Psicothema* (Oviedo).
- Radocy, R. E., & Boyle, J. D. (1979). *Psychological foundations of musical behavior*. Springfield: CC Thomas.

- Rauscher, F. H., Shaw, G. L., & Ky, K. N. (1993). Music and spatial task performance. *Nature*, 365, 611.
- Schellenberg, E. G. (2003). Does exposure to music have beneficial side effects? Em I. Peretz & R. J. Zatorre (Eds.), *The cognitive neuroscience of music* (pp. 430-448). Oxford: Oxford University Press.
- Schellenberg, E. G. (2004a). Does exposure to music have beneficial side effects? Em R. J. Zatorre & I. Peretz (Eds.), *The cognitive neuroscience of music* (pp. 430-448). Nova Iorque: Oxford.
- Schellenberg, E. G. (2004b). Music lessons enhance IQ. *Psychological Science*, 15, 511-514.
- Schoen, M. (1940). *The psychology of music: a survey for teacher and musician*. Oxford: Ronald Press.
- Seashore, C. E. (1942). Science in music. *Science*, 95(2469), 417-422.
- Shuter-Dyson, R., & Gabriel, C. (1981). *The psychology of musical ability*. Londres: Methuen.
- Sloboda, J. A. (2004). *The musical mind: the cognitive psychology of music*. Nova Iorque: Oxford University Press.
- Smith, R. M. (1996) Polytomous mean-square fit statistics. *Rasch Measurement Transactions*, 10(3), 516-517.
- Stankov, L. (1994). Auditory abilities. Em R. J. Sternberg (Ed.), *Encyclopedia of human intelligence* (pp. 157-162). Nova Iorque: Macmillan.
- Wallentin, M., Nielsen, A. H., Friis-Olivarius, M., Vuust, C., & Vuust, P. (2010). The Musical Ear Test, a new reliable test for measuring musical competence. *Learning and Individual Differences*, 20, 188-196.
- Wilson, D. T., Wood, R., & Gibbons, R. (1991). *TESTFACT: test scoring, item statistics, and item factor analysis*. Chicago, IL: Scientific Software International.
- Wing, H. D. (1941). A factorial study of music tests. *British Journal of Psychology*, 31, 341-355.
- Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8(3), 370.

Recebido em 05/08/2012
Reformulado em 08/09/2012
Aprovado em 13/10/2012

Pesquisa realizada com apoio financeiro da CAPES, CNPq e FAPESP.

Sobre os autores:

Fernando Pessotto é psicólogo, mestre (bolsista FAPESP) e doutorando pela Universidade São Francisco (USF). Atua como docente do curso de graduação em psicologia no Centro Universitário Salesiano de Americana (Unisal).

Tatiana Freitas da Cunha Lima é psicóloga pela Faculdade Ruy Barbosa e mestre pela Universidade São Francisco (bolsista CAPES), atuando como recrutadora em empresa de tecnologia.

Ricardo Primi é psicólogo pela PUCCampinas, doutor em Psicologia Escolar e do Desenvolvimento Humano pela Universidade de São Paulo com parte desenvolvida na Yale University (EUA) e docente do Programa de Pós-Graduação *Stricto Sensu* em Psicologia da Universidade São Francisco. Recebe financiamento do CNPq (produtividade em pesquisa), FAPESP e CAPES.

Lucas de Francisco Carvalho é psicólogo, mestre e doutor pela Universidade São Francisco (USF) e docente do Programa de Pós-Graduação *Stricto Sensu* da Universidade São Francisco (USF).

Fabiano Koich Miguel é psicólogo pela Universidade Presbiteriana Mackenzie e mestre e doutor pela Universidade São Francisco. Atualmente é professor adjunto da Universidade Estadual de Londrina (UEL). Recebe financiamento CNPq e Fundação Araucária.

Contato com os autores:

Programa de Pós Graduação *Stricto Sensu* em Psicologia da Universidade São Francisco
Rua Alexandre Rodrigues, 45 – CEP 13251-900 – Itatiba-SP.
E-mail: rprimi@mac.com