

MARCOS SERGIO BELCZAK JUNIOR
GRR20163977

CAPACIDADE DE IMITAÇÃO DE BIBLIOTECAS DE CORDA EM QUARTETOS

Monografia apresentada à disciplina OA027 - Trabalho de Conclusão de Curso (I ou II) como requisito parcial à conclusão do Curso de Bacharelado/Licenciatura em Música - Departamento de Artes, Setor de Artes, Comunicação e Design da Universidade Federal do Paraná.

Orientador: Maurício Dottori

CURITIBA
2021

Dedico este trabalho a minha mãe, que me ensinou meus valores,
que me aceitou como sou,
e me inspirou a seguir meu caminho.

RESUMO

Esta pesquisa busca medir a capacidade discriminatória de músicos em relação a excertos musicais reais e virtuais, avaliando a possibilidade de recriação da expressividade musical humana por meios digitais. Para tal, performances de quatro quartetos de corda distintos foram selecionadas e replicadas digitalmente, com a utilização de instrumentos MIDI e bibliotecas de corda. Excertos dos áudios originais e das cópias digitais foram inseridos em um questionário, no qual 260 músicos, de expertises diversas, testaram suas habilidades de diferenciação entre as origens dos trechos tocados. A capacidade de imitação da expressividade humana pelos instrumentos virtuais foi medida utilizando o Jogo da Imitação, de Alan Turing, enquanto as tendências de resposta e os possíveis critérios utilizados foram mensurados e discutidos utilizando a Teoria da Detecção de Sinais.

Palavras-chave: instrumentos virtuais, expressividade musical, quartetos de corda, instrumentos MIDI, áudio digital, tecnologia e música.

ABSTRACT

This research aims to measure the overall discriminatory capability of musicians when it comes to real and virtual musical excerpts, questioning the possibility of replication of human musical expressivity by means of digital tools. Performances of four distinct string quartets were selected and recreated virtually with MIDI instruments and string libraries. Excerpts, extracted from both sources, were inserted in an online survey, in which 260 musicians of various expertises tested their abilities to discern between its origins. The capability of human musical expressiveness by virtual instruments was evaluated making use of Alan Turing's Imitation Game, while the participant's responsive tendencies and possible bias were measured and discussed using Signal Detection Theory.

Palavras-chave: virtual instruments, musical expressiveness, string quartet, MIDI instruments, digital audio, music and technology.

SUMÁRIO

| | |
|--|----|
| INTRODUÇÃO | 01 |
| 1 FUNDAMENTAÇÃO TEÓRICA | |
| 1.1 INSTRUMENTOS VIRTUAIS MODERNOS..... | 03 |
| 1.2 INSTRUMENTOS VIRTUAIS COMO FERRAMENTAS DE SIMULAÇÃO E IMITAÇÃO..... | 06 |
| 1.3 HUMANIZAÇÃO E EXPRESSIVIDADE..... | 08 |
| 2 MATERIAIS E MÉTODOS | |
| 2.1 PROCEDIMENTOS GERAIS..... | 10 |
| 2.2 MATERIAL SONORO..... | 14 |
| 2.3 FERRAMENTAS | 15 |
| 2.4 CRIAÇÃO DAS RÉPLICAS | 15 |
| 3 ANÁLISE DOS DADOS | |
| 3.1 ANÁLISE GERAL..... | 19 |
| 3.2 ANÁLISE POR GRUPOS..... | 20 |
| 4 DISCUSSÃO | |
| 4.1 RESULTADOS ENCONTRADOS X RESULTADOS ANTERIORES..... | 23 |
| 5 CONSIDERAÇÕES FINAIS | 28 |
| REFERÊNCIAS | 30 |

INTRODUÇÃO

Dos contos distópicos, nos quais máquinas coexistem com seres humanos, de Phillip K. Dick, aos insertos teóricos de Alan Turing sobre a natureza e a evolução da inteligência artificial, a capacidade do virtual e do eletrônico de se atualizarem em sempre crescentes níveis de complexidade e eficiência instiga os debates sobre a possibilidade de existir, em algum determinado momento da história, um objeto completamente sintético capaz de reproduzir, de forma completa e certa, a consciência humana em todas as suas formas.

Se o objeto em questão estaria, potencialmente, em consideração para uma completa mimetização do ser humano, talvez uma das provas mais derradeiras seja a capacidade de imitação dos setores mais subjetivos da mente humana, aqueles ligados às artes, à criatividade e à expressividade (Boden, M. 1998). A ideia de uma inteligência completamente sintética que possa imitar, por exemplo, o processo de composição de uma peça para orquestra ou a concepção de um artefato artístico está atrelada a uma miríade de debates, tanto teóricos quanto filosóficos, que vêm crescendo ao longo dos anos, conforme novas tecnologias baseadas na auto-análise parecem tornar essa realidade cada vez mais plausível.

Na música, um dos focos desse debate tem ocorrido na constante busca e demanda por instrumentos virtuais capazes de replicar o som de instrumentos acústicos com precisão e fidelidade. Enquanto algumas empresas investem em *samplers* capazes de reproduzir gravações de notas individuais captadas com extrema precisão e clareza, possibilitando que qualquer profissional com um teclado MIDI e um *software* de edição de áudio possam reproduzir um som de um instrumento de forma instantânea, outras procuram por soluções envolvendo sintetizadores extremamente complexos, cujos múltiplos osciladores e extensiva programação almejam recriar, de forma inteiramente artificial, todo o espectro sonoro de um instrumento.

A possibilidade de estar disponível, à extensão das mãos, o corpo sonoro e o total controle de qualquer instrumento existente, de poder recriar toda a complexidade de uma orquestra sem ao menos precisar sair de casa ou passar pelo extenuante processo de contratação de músicos, é uma perspectiva que seduz produtores musicais, arranjadores e compositores.

Pelo seu relativo baixo custo, facilidade de edição e praticidade, as bibliotecas de áudio focadas em instrumentos orquestrais têm dominado o mercado da trilha sonora. Uma pesquisa recente mostrou que 90% dos compositores profissionais atuantes no ramo das trilhas para filmes, jogos e comerciais da Nova Zelândia utilizam bibliotecas de áudio e instrumentos virtuais em seu dia a dia (Morgan, C. 2016). O livro *Acoustic and MIDI orchestration for the*

contemporary composer, escrito por Andrea Pejrolo e Richard Rosa já apontava, em 2007, para a crescente demanda do mercado audiovisual por instrumentos virtuais que pudessem, de certa forma, digitalizar completamente o processo da composição e concepção das peças orquestrais para filmes, sinalizando um interesse da indústria na possibilidade de substituir o extensivo processo de contratação, ensaios e gravação de uma orquestra por um único músico, capaz de manipular e recriar todos esses instrumentos de forma virtual.

Muitos compositores modernos de extremo sucesso, como Hans Zimmer e Danny Elfmann, têm endossado empresas que produzem bibliotecas de áudio prometendo qualidade e precisão impecáveis. Conforme programas de ensino musical, focados na composição para filmes, dedicam cada vez mais seus currículos à prática da orquestração virtual, é seguro dizer que boa parte da produção musical voltada para trilhas sonoras é, hoje, apresentada em sua forma final como um produto inteiramente virtual, composto, concebido e finalizado inteiramente no computador (Morgan, C. 2016).

A capacidade dos instrumentos virtuais de simularem com fidelidade a técnica, a expressividade e o som dos instrumentos reais é um tema em constante ascensão devido à popularidade dos mesmos. Uma pesquisa realizada em 2017 por Reinhard Kopiez, intitulada *replacing the orchestra - the discernibility of sample library and live orchestra sounds*, buscou quantificar e analisar a capacidade de discernimento de ouvintes no que tange a réplicas orquestrais virtuais e orquestras reais. Para tal, um questionário online, contendo excertos originais da peça *O Pássaro de Fogo*, de Stravinsky, e cópias completamente digitais, realizadas com bibliotecas de sample, foi desenvolvido e divulgado para mais de 600 pessoas com diferentes níveis de conhecimento musical.

A pesquisa encontrou resultados díspares entre dois grandes grupos de indivíduos. O primeiro grupo, composto por pessoas que foram capazes de deduzir a procedência dos excertos com 90% de precisão, continha em sua maioria maestros, produtores musicais e compositores experientes. O segundo grupo, contendo participantes que variavam entre estudantes universitários, músicos amadores, instrumentistas profissionais e professores de música, foi capaz de atingir uma margem de acerto de 68,25%.

Os resultados foram então interpretados por Kopiez utilizando o *Teste de Turing* como base, colocando a orquestra virtual no papel da inteligência artificial, e os excertos originais no papel da inteligência humana. Concluiu-se que, em determinados grupos de indivíduos especializados em música, as orquestras digitais conseguiram, dentro da margem dada por Turing, imitar a orquestra real, mas músicos com dado maior grau de conhecimento musical conseguiram analisar nuances que possibilitam a distinção quase perfeita entre trechos reais e digitais.

Enquanto a análise geral dos dados pode ter soado promissora no que tange a capacidade de imitação dos instrumentos virtuais, é importante notar que as pesquisas realizadas, até o momento, levaram em consideração apenas grandes *ensembles*, como orquestras e big bands. Em grupos musicais com grande número de músicos, a expressividade individual tende a se diluir em prol de uma maior unidade sonora do conjunto. Grupos menores, como quartetos de corda, tendem a apresentar maior individualidade e expressividade (Glowski, D. 2013), pois o caráter intimista presente na formação de um quarteto permite maior liberdade de micro-variações rítmicas e timbrísticas, bem como maior fluidez nos ataques das notas.

Outro fator importante a se notar é o avanço tecnológico constante das últimas décadas. Lançamentos recentes das grandes empresas de instrumentos virtuais têm incorporado novas técnicas de gravação, tratamento, e manuseio dos *samples*, enquanto novas versões dos *samplers* permitem uma miríade de opções para a programação e controle dos instrumentos.

Levando em consideração os dois pontos prévios, essa pesquisa busca recriar o experimento levantado e idealizado por Kopiez, utilizando, porém, *ensembles* limitados a quartetos de cordas, com réplicas que dispõe das mais modernas técnicas e ferramentas de composição virtual. Ao escolher grupos menores, pode-se estudar a capacidade de replicação das características expressivas e grande variabilidade de técnicas e sonoridades possibilitadas pelas formações mais enxutas, bem como o quanto os avanços técnicos na produção desses instrumentos nos últimos anos influenciaram na qualidade geral e na capacidade de imitação dos instrumentos virtuais.

1 FUNDAMENTAÇÃO TEÓRICA

1.1 INSTRUMENTOS VIRTUAIS MODERNOS

A ideia de instrumentos capazes de manipular sons gravados previamente não é exatamente moderna. Modelos de teclados como o *mellotron*, nas décadas de 60 e 70, podiam disparar, com suas teclas, momentos específicos de gravações feitas em seu sistema analógico de fitas, tornando-se bastante popular na época, apesar de seu complexo manuseio e manutenção.

Enquanto em 1969 já podia se observar o primeiro modelo digital de um Sampler, consistindo no sistema EMS Musys, a tecnologia passou a se tornar disponível no mercado a partir de 1976 com o instrumento Computer Music Melodian, desenvolvido por Harry Mendell,

o qual substituiu o complexo sistema de fitas dos modelos como o *mellotron* por um sistema digital muito mais preciso, acessível e compacto. (Grogono, P. 1973)

Os sistemas digitais de *Samplers* rapidamente cresceram em popularidade, enquanto avanços na capacidade de processamento digital permitiram o desenvolvimento de instrumentos icônicos como LMD-649 e AKAi MPC60, as quais se tornaram ferramentas fundamentais no surgimento de estilos musicais que dominaram a mídia nas décadas de 80 e 90 (D’Errico, M. 2011).

Com a revolução da digitalização do áudio em meados da década de 90 e o surgimento do VST (tecnologia de estúdio virtual) e do VSTi (tecnologia do instrumento de estúdio virtual), as possibilidades trazidas pelo grande avanço tecnológico na capacidade de processamento dos computadores trouxeram uma nova dinâmica para os instrumentos *samplers*, agora completamente integrados em modernos sistemas digitais. Se até esse momento, tais instrumentos eram escolhidos de maneira estilística para fundamentar gêneros musicais como o Synth Wave e o Hip Hop (Exarchos, M. 2019), passa a existir a perspectiva de mapear completamente o som de um instrumento e possibilitar a reprodução dos mesmos por meio de controladores MIDI e computadores.

Em 2003, surgem as primeiras coleções de *samples* orquestrais, quando empresas como EastWest Quantum Leap buscaram capturar o som de todos os naipes de instrumentos e inseri-los em samplers digitais, programados para garantir a reprodução não só das notas, mas também das técnicas, variações naturais de timbre e intensidades. Esses produtos rapidamente tornaram-se populares na indústria musical, principalmente na área das trilhas sonoras, onde compositores foram apresentados à possibilidade de criar amostras de suas orquestrações, as quais poderiam ser enviadas de antemão aos diretores dos filmes, dando uma noção mais detalhada de como suas peças se relacionariam com as imagens (Morgan, C. 2016).

A eficácia e o dinamismo proporcionados pelos instrumentos orquestrais virtuais rapidamente estabeleceu a demanda na indústria musical cinematográfica por bibliotecas que pudessem não só serem usadas como demonstrações, mas também incluídas nas versões finais dos filmes. Cresce, a partir desse momento, o crivo em relação a qualidade e a fidelidade dos instrumentos virtuais. (Morgan, C. 2016)

Samplers modernos como o instrumento virtual Kontakt da empresa native instruments, operado por MIDI, permitem uma miríade de programações e roteamentos internos, possibilitando que diversos fatores atuem na reprodução de diferentes samples, como mudanças na intensidade e ataque das notas resultando no disparo de sons distintos. Para tirar vantagem da alta capacidade de armazenamento desses instrumentos, empresas desenvolvedoras de bibliotecas orquestrais operam um sistema bastante complexo e meticuloso, no qual todo o

espectro sonoro, variações de intensidade e técnicas estendidas de um instrumento são captados por gravações de milhares de *samples*. Transições entre as notas também são gravadas, resultando na possibilidade de programar movimentos como bends, legatos e slides, os quais aumentam a capacidade expressiva dos instrumentos virtuais (Klein, E. 2016).

Enquanto as variações de dinâmicas presentes nos ataques se mostraram influenciáveis na maneira como pessoas distinguem instrumentos virtuais de instrumentos reais (Hähnel, T. 2013), dados mais recentes mostram que a qualidade do som e a veracidade do timbre são os fatores pivotaes na percepção da origem de determinado áudio (Kopiez, R. 2017). Algumas das bibliotecas usadas neste trabalho são fruto de produções extensivas e meticulosas, passando pelo mais alto crivo, desde a seleção dos músicos e instrumentos a serem captados até a utilização dos melhores aparatos de gravação disponíveis no mercado.

Além de todos os aspectos abordados anteriormente, algumas bibliotecas dispõem de ferramentas alternativas que proporcionam ainda mais realismo para o som. Em seu artigo *Feigning Humanity*, Eve Klein discute, em detalhes, as funcionalidades dos instrumentos virtuais modernos:

Alguns instrumentos virtuais também incorporam uma gama de recursos para simular acústicas de ambiente (reverb de convolução), para modificar timbres instrumentais (equalizadores e filtros), para alterar o tempo (por exemplo, quantização), para alterar a altura (efeitos de transposição e controles de entonação humanizada), ou para efetuar o sinal de saída por meio da repetição, distorção ou transformação do sinal de áudio (efeitos de modulação, arpegiadores...). Instrumentos virtuais orquestrais, particularmente aqueles direcionados para as trilhas sonoras de filmes e videogames, podem incluir *Patches* de instrumentos que combinam síntese com sons sampleados para criar opções com um desenho de som mais abstrato/menos natural (Klein, E. 2016).

Por meio de parcerias com compositores famosos como Hans Zimmer e Danny Elfman, e também realizando colaborações com nomes influentes na música clássica, incluindo o violinista Joshua Bell e a cellista Tina Guo, empresas como Spitfire e Embertone entregam produtos de altíssima qualidade, os quais não apenas operam em proporcionar grande versatilidade e variedade de técnicas, mas também prometem a fidelidade sonora para com os instrumentos reais.

1.2 INSTRUMENTOS VIRTUAIS COMO FERRAMENTAS DE SIMULAÇÃO E IMITAÇÃO

Em seu livro *Simulacra e simulacrum*, o sociólogo Jean Baudrillard discute a natureza da simulação, bem como as relações existentes entre a simbologia, a iconoclastia e o pensamento social na criação de uma realidade compartilhada, a qual pode ou não, ser uma imitação abstrata fabricada pelo coletivo.

Eve Klein, no artigo “Feigning humanity”, aborda alguns dos princípios definidos por Baudrillard para contextualizar as funções e a natureza dos instrumentos virtuais:

Relembrando a citação de Baudrillard “a simulação instiga a diferença entre o verdadeiro e o falso, o real e o imaginário (1994:3), instrumentos virtuais orquestrais podem ser vistos de forma a desestabilizar a música orquestral acústica ao produzir qualidades suficientes da música orquestral ao ponto de instigar a distinção entre a criação musical acústica e virtual na mídia. Ao samplear, processar, efetuar, e de certa forma reproduzir os micro gestos de expressividade performática pela aplicação de instrumentos virtuais, os compositores desenvolveram sintaxes desses gestos que simbolizam a performance orquestral humana (Klein, E. 2016).

O artigo em questão estabelece o paralelo entre os conceitos propostos por Baudrillard e a intenção da indústria musical moderna de criar instrumentos virtuais orquestrais ditos realistas. Originalmente, tais instrumentos foram desenvolvidos como equipamentos capazes de criar demonstrações assertivas das versões finais, direcionados a compositores de trilha sonora que almejavam poder apresentar, de forma mais consistente e clara, suas ideias musicais para produtores e diretores de cinema (Morgan, C. 2016). Muitas questões, especialmente ligadas à praticidade e ao baixo custo da criação de composições utilizando apenas instrumentos virtuais, estabeleceram a demanda da indústria audiovisual pela substituição da orquestra real em benefício da virtual. Esta evolução da utilização das orquestras virtuais pode ser interpretada em linha com a ideia de Baudrillard de que as simulações, eventualmente, enfraquecem o sentido original do objeto em questão, podendo até substituí-lo por completo.

Outro exemplo da categorização de instrumentos virtuais em um contexto teórico pode ser observado na utilização do *Jogo da Imitação* (Turing, A. 1950) pelo pesquisador Reinhard Kopiez, em seu artigo “Replacing the Orchestra? — the discernibility of sample library and live orchestra sounds”. O Jogo da Imitação consiste em um modelo teórico, proposto por Alan Turing como uma forma de testar a capacidade de exibição de inteligência em máquinas que seja semelhante ou, até mesmo indistinguível, a de um ser humano. Este teste é relevante pois

estabelece uma maneira de quantizar o processo de pesquisa, uma vez que define um número fixo para que seus resultados sejam considerados atingidos: caso um terço dos participantes de determinado teste não consigam diferenciar a origem de determinado excerto como sendo humano ou virtual, pode-se estabelecer a existência de uma Inteligência Artificial, capaz de dialogar com a Inteligência Humana em termos semelhantes (Turing, A. 1950).

Kopiez adapta o *Jogo da Imitação*, colocando as bibliotecas de samples como o elemento virtual e as orquestras reais como o elemento humano, propondo que, em um determinado teste, caso os participantes do mesmo não sejam capazes de diferenciar, com ao menos 70% de precisão, as origens de dados excertos sonoros, pode-se concluir a capacidade dos instrumentos virtuais de imitar os trejeitos humanos e a inteligência musical que caracterizam a performance orquestral original.

Os resultados do trabalho de Kopiez mostraram uma divisão distinta entre os valores de dois grupos. O primeiro, formado por pessoas de considerada menor expertise musical, atingiu 68,75% de precisão na distinção entre os exemplos virtuais e reais. De acordo com o autor, esse número sinalizaria a existência da imitação dos trejeitos e características humanas na performance digital. Porém, um segundo grupo, formado por indivíduos de grande expertise musical, obteve resultados cuja precisão atingiu 80%, desconsiderando a hipótese da imitação nos instrumentos virtuais, uma vez que existem, claramente, nuances e diferenças que podem ser captadas por profissionais com mais refinamento na escuta musical.

Não se pode desconsiderar, no entanto, a constante evolução, não só dos instrumentos virtuais, com técnicas de captação, desenvolvimento e execução cada vez mais complexas e refinadas, mas também das técnicas de composição e programação virtuais modernas. Os instrumentos e bibliotecas utilizados neste trabalho não existiam na época do desenvolvimento da pesquisa de Kopiez, e são considerados, tanto pela indústria quanto pelos profissionais atuantes no mercado, como tecnologias e produtos muito superiores aos usados anteriormente.

Simulações se caracterizam pela replicação de um ou mais fatores reais que possam delimitar uma ideia ou um objeto (Baudrillard, J. 1994). Ao estabelecer que instrumentos virtuais orquestrais são, por definição, concebidos como potenciais simulações de instrumentos acústicos, e estes últimos se atrelam ao ato humano da performance, cabe a discussão sobre quais fatores, sejam eles timbrísticos, expressivos ou acústicos, contribuem para a percepção da natureza de um determinado trecho como sendo humana e natural, em contrapartida a soar robótica ou virtual.

1.3 HUMANIZAÇÃO E EXPRESSIVIDADE

Existem muitas divergências conceituais e opinativas no que tange ao tópico dos elementos de expressividade na música, mas é consensual a ideia de que, ao realizar uma performance, um músico busca expressar uma ideia ou um sentimento por meio de seu instrumento (Woody, R.H. 2000). Estudos mostram que é possível a captação, com a utilização de microfones e equipamentos de gravação, não só do som mas também dos elementos conjuntos que podem caracterizar uma performance como expressiva (Clarke, E.F. 1993). Se uma gravação pode captar a expressividade, esta é, então, algo tangível e concreto, que pode eventualmente ser replicado de forma artificial (Klein, E. 2016).

Se um instrumento virtual é idealizado como uma possível simulação de um objeto real, este deve, por definição, ser capaz de possibilitar a reprodução da expressividade dos instrumentos orquestrais acústicos. Mas quais seriam os elementos que comunicam e enganam os ouvintes quanto a noção da expressividade humana em uma performance digital?

Palmer define a expressividade como o microcosmos de variações em altura, intensidade, tempo e timbre que caracterizam uma performance como humana (1989). Variações de tempo incluem: mudanças métricas, micro e macro alterações rítmicas e exageros de notações musicais (pausas, codas, suspensões e síncopas). Variações de altura podem ser caracterizadas como flutuações naturais de afinação (principalmente em instrumentos não-temperados), técnicas de passagem de notas (legatos, portamentos, *slides*) e técnicas que desestabilizam a afinação (*pitch bend*). Variações de dinâmica, decorrentes de ideias musicais (*crescendo*, *sforzando*, *diminuendo*) ou micro variações naturais presentes no ato da performance (pequenas diferenças de volumes entre notas sequências). Variações timbrísticas, representadas como técnicas específicas que alteram o som, ou mesmo características acústicas do instrumento que podem alterar o timbre em diferentes regiões do seu espectro sonoro. (Klein, E. 2013)

A influência das micro-variações na percepção da humanização de uma performance pode ser vista em determinadas pesquisas direcionadas. Um estudo de 2007, publicado por Michael Oehler e Christoph Reuter, mostrou que a adição de um vibrato modelado sinteticamente em simulações de instrumentos de sopro aumentou significativamente a percepção dos ouvintes em relação ao realismo e a humanidade dos áudios. Esse estudo não lidava, no entanto, com linhas melódicas ou mesmo performances, consistindo apenas em notas sintéticas isoladas comparadas lado a lado com notas geradas por fontes acústicas.

Simulações virtuais de performances barrocas apresentaram resultados distintos, dependendo de suas técnicas de produção. Modelos gerados de forma puramente digital, utilizando um algoritmo que apenas transcreveu fragmentos das peças originais para a linguagem MIDI, mostraram-se de fácil distinção para com suas imagens reais. Modelos que foram gerados artificialmente mas que tiveram, em sequência, parâmetros como intensidade, tempo e entonação manipulados de forma manual provaram-se como uma tarefa muito mais difícil de distinção, pelo menos entre indivíduos com menor grau de expertise musical. Finalmente, uma etapa final do estudo, a qual consistia de áudios originais e réplicas virtuais, ambos reproduzidos propositalmente em baixíssima resolução, revelaram-se distinguíveis apenas para grupos de considerada maior expertise musical, como produtores musicais e maestros (Hähnel, T. 2013).

Os resultados encontrados previamente estão em linha com pesquisas mais recentes, como o artigo de Kopiez, o qual também observou que, enquanto a simulação de micro variações é essencial para a percepção da humanização quando avaliada por grupos de menores expertises, indivíduos com grande proficiência musical e habilidades de escuta bastante desenvolvidas são capazes de distinguir, com precisão praticamente perfeita, a procedência de excertos sonoros como sendo virtuais ou reais (Kopiez, R. 2017).

Enquanto é fácil de extrair, dos trabalhos anteriores, a conclusão de que a expertise musical e, por conseguinte, a presença de uma audição crítica e bastante perceptiva, influenciam na distinção entre instrumentos acústicos e suas réplicas virtuais, é importante notar que tais exclamações sinalizam, e isto é também observado pelos autores de ambos os artigos, que além das diferenças de expressividade, a qualidade sonora dos instrumentos virtuais também atua na distinção entre performances de diferentes procedências, com potenciais nuances timbrísticas sendo perceptíveis aos ouvidos mais treinados.

Elementos como ambientação e espacialização também podem contribuir para a caracterização de um excerto como natural. Uma pesquisa realizada em 2017, intitulada *The plausibility of a string quartet performance in virtual reality*, publicada por Ilias Bergström, verificou que, ao disponibilizar para “espectadores virtuais”, os quais usavam aparelhos de realidade virtual para assistir a uma performance completamente digital de um quarteto de cordas, controles que podiam realçar ou reduzir certos aspectos da apresentação, o parâmetro que mais influenciou positivamente na percepção do grupo como convincente e humanizado foi a espacialização, ou seja, o realce da noção da distribuição dos instrumentos em um determinado espaço e a relação sonora dos mesmos com o ambiente em que estão inseridos.

Mesmo existindo catalogadas uma extensa lista de fatores que contribuem para a simulação da expressividade, é importante notar que tais fatores nunca se aplicam isolados em uma performance (Clarke, E.F. 1993). Um músico, na intenção de comunicar suas ideias,

utilizará de complexas combinações dos elementos citados previamente, fato este que se torna um grande desafio para o desenvolvimento de instrumentos virtuais que possibilitem a total reprodução dos trejeitos humanos no ato da performance musical. Enquanto a captação e programação de elementos expressivos de forma isolada é relativamente fácil, a inserção dos mesmos em um sistema completamente funcional, o qual permite a transição fluente e a combinação de diversas técnicas, em diferentes níveis de intensidade, dinâmica e acentuação, requer um nível de processamento digital muito amplo, além de demandar um grande número de samples captados, bem como uma extensa programação dos mesmos em um instrumento que apresente as funcionalidades requeridas. (Klein, E. 2013)

Ainda assim, pesquisas como a de Kopiez revelam resultados promissores quando avaliam as capacidades de distinção entre indivíduos de menor expertise musical, e o autor ainda discute o fato de que, indivíduos como quase nenhum aprofundamento musical apresentariam resultados ainda menos precisos quando encumbidos da tarefa de discernir performances reais de virtuais. Considerando que existem estimativas de que 70% de toda a produção musical voltada para trilhas cinematográficas é feita exclusivamente com instrumentos virtuais, os quais integram, inclusive, as versões finais dos filmes, é possível dizer que os instrumentos virtuais já atingem um nível de simulação considerável (Morgan, C. 2016).

É importante notar que as pesquisas realizadas até o momento utilizaram performances realizadas por grupos orquestrais. De acordo com Andrea Pejrolo, em seu livro *Acoustic and Midi Orchestration for the Contemporary composer*, a simulação de características expressivas em *ensembles* grandes como orquestras é muito mais fácil que a imitação de trejeitos específicos de instrumentos individuais, uma vez que o corpo sonoro dos naipes de instrumentos ameniza a percepção das micro variações previamente abordadas neste trabalho. (Pejrolo, A. 2007)

A utilização de quartetos de corda como base para a criação das réplicas virtuais é uma maneira de evidenciar as características expressivas dos instrumentos, e permite uma análise com maior foco nos elementos que denotam a humanidade em um excerto sonoro.

2 MATERIAIS E MÉTODOS

2.1 PROCEDIMENTOS GERAIS

Esta pesquisa usa como base o modelo técnico apresentado por Kopiez em seu artigo *Replacing the Orchestra? — the discernibility of sample library and live orchestra sounds*, no qual um relatório *online*, contendo excertos musicais originados de performances orquestrais

humanas e replicações virtuais da peça *O Pássaro de Fogo*, são apresentados aos participantes, os quais devem classificá-los em relação a sua origem utilizando apenas a escuta musical.

Adaptações ao modelo de Kopiez consistem no foco em um grupo reduzido de instrumentos como o quarteto de cordas, bem como a utilização de diferentes gravações e peças, em vez de uma única fonte. Esta última mudança busca reduzir a influência de fatores alheios a performance, como pequenas divergências na simulação da reverberação, ruídos como tosses/passos e artefatos sonoros repetitivos decorrentes da gravação, que possam servir como “pistas” para a identificação da procedência dos áudios e desvincular o foco do participante em utilizar sua escuta para a escolha das respostas.

O questionário foi desenvolvido utilizando a ferramenta Google Forms. Este meio foi escolhido pois apresenta todas as funcionalidades desejadas, de forma simples e gratuita. A pesquisa consiste em duas etapas. Em uma primeira instância, o participante é apresentado ao tema e recebe algumas recomendações para a realização do teste. Posteriormente, o participante responde a 16 perguntas, as quais consistem de excertos musicais retirados de performances reais e réplicas virtuais, que devem ser diferenciadas presente as opções “real” e “virtual”. A segunda etapa engloba questões extras como profissão, elementos auditivos que influenciaram na escuta musical e um espaço aberto, no qual os participantes podem descrever seu processo de decisão e escolha das respostas. Essas questões permitem a análise crítica e contextualizada dos dados obtidos, facilitando a divisão em diferentes grupos de distintas expertises musicais e seus respectivos resultados.

Os resultados dos questionários serão analisados de duas formas. A primeira atribui às perguntas respostas corretas, e contabiliza as mesmas para atingir uma média coletiva. Esta média é, então, analisada utilizando os conceitos do *Jogo da Imitação*. O teorema, proposto por Alan Turing em 1958, determina que a existência de inteligência artificial seria atingida se, ao responderem um teste composto por perguntas executadas por ambos seres humanos e computadores, participantes não conseguissem diferenciar as procedências dos estímulos com ao menos 70% de precisão.

O teste é adaptado aos moldes e terminologias deste trabalho, colocando as orquestras reais no papel do ser humano e os instrumentos virtuais no papel dos computadores. A capacidade da replicação da expressividade humana seria a equivalente à inteligência artificial, colocada em questão no teorema original.

Como método para a análise mais detalhada dos dados obtidos, a teoria da detecção de sinais foi escolhida. Esta teoria permite a quantização de fatores que seriam comumente tidos como subjetivos, e é muito utilizada em pesquisas que lidam com tomadas de decisões guiadas por opiniões, experiências pessoais e critérios individuais (Kellen,D. 2008).

A teoria define que um determinado teste discriminatório, composto de um único estímulo por pergunta e duas possíveis respostas, resultará em quatro alternativas: acertos, omissões, falsos alarmes e rejeições corretas.

Traduzindo para os termos aplicados neste trabalho, um acerto seria a classificação correta de um excerto musical original como sendo real. Uma omissão seria a classificação de um trecho real como virtual. Um falso alarme seria a resposta que qualifica como real um instrumento virtual, enquanto uma rejeição correta seria a discriminação assertiva de um instrumento virtual. Os quatro parâmetros podem ser melhor visualizados na tabela abaixo:

| | | |
|--------------|--------------|------------------|
| | Respostas | Respostas |
| Estímulos | Real | Virtual |
| Real (s1) | Acerto | Omissão |
| Virtual (s2) | Falso alarme | Rejeição correta |

A teoria define que a tomada de decisões em um teste discriminatório é feita por meio de duas variáveis: sensibilidade e especificidade. A primeira consiste na capacidade do indivíduo em identificar e detectar atributos que o levem a diferenciar duas opções, bem como a existência de fatores como configurações de atributos que possam encaminhar o indivíduo a discriminar, de forma errônea ou enganosa, a origem dos estímulos. O valor da sensibilidade pode ser teorizado utilizando os valores proporcionais de Acertos e Falsos Alarmes em relação à totalidade dos respectivos estímulos, presentes na pesquisa. Para obter o grau de variância, necessário para a análise completa e a disposição gráfica dos resultados, os valores são convertidos para a tabela de Distribuição Normal de transformada de Z. A equação para o cálculo da sensibilidade, também chamada de d' ou “d-prime”, é a seguinte:

$$d' = z(\text{Acertos}/s1) - z(\text{alarmes falsos}/s2)$$

Os valores obtidos dessa equação são interpretados como a capacidade de um indivíduo ou grupo em identificar elementos presentes nos estímulos, bem como utilizá-los de forma eficaz para tomar uma decisão de resposta. Quanto maior o número obtido, maior a capacidade

de discernibilidade do indivíduo em relação aos estímulos apresentados.

Especificidade, também chamado de tendência, bias ou enviesamento, representa a propensão de um indivíduo ou grupo em escolher certas respostas em prejuízo de outras. Este parâmetro existe pois diferentes configurações de Acertos e Alarmes falsos podem resultar em valores de Sensibilidade (d') idênticos, e este fator pode prejudicar a análise de dados. Faz-se uso então, de uma equação auxiliar que busca medir os critérios recorrentes que causam o enviesamento individual e a predominância de certas opções. Esta equação é uma relação de adição entre as proporções de acertos e falsos alarmes, e pode ser classificada de duas maneiras: conservadora, quando há um número reduzido de falsos alarmes, resultando em um número elevado de omissões; liberal, quando existe a predominância de falsos alarmes nas respostas.

A equação para o cálculo da especificidade, denominada C é a seguinte:

$$C = -0.5 \cdot [z(\text{acertos}/s1) + z(\text{alarmes falsos}/s2)]$$

Para esta equação, os resultados interpretam-se da seguinte forma: se $C > 0$, o critério é conservador; se $C < 0$, o critério é liberal; se $C = 0$, o critério é denominado neutro, e será então analisado como a inexistência de critério, pois não estabelece ou representa uma tendência de escolha de resposta.

Outras equações, propostas pela TDS, não foram utilizadas nesse trabalho, pois não influenciam na análise geral dos dados para esse projeto em específico. Como forma de calcular a capacidade de discernimento entre sons de origem virtual e sons de origem real, um teste naturalmente incerto, é importante a utilização de parâmetros que materializam as tendências individuais de escolha das respostas, uma vez que as mesmas estão atreladas a diferentes fatores pessoais como vivências, escutas musicais, especificidade auditiva e capacidades distintivas (Kellen, 2003).

Em contrapartida, a quantificação bruta dos dados, atrelada ao *Jogo da imitação*, mantém o caráter estatístico da pesquisa e permite a análise direta sobre a capacidade de imitação da expressividade humana. Aplicar ambas as teorias em análises separadas e, posteriormente, conjuntas, permite o aprofundamento teórico embasado por fatores matemáticos e psicossociais, os quais englobam a miríade de fatores relacionados à tomada de decisões em relação a esse tema.

2.2 MATERIAL SONORO

Para o material sonoro utilizado, levou-se em consideração fatores como proeminência de diferentes técnicas no decorrer da composição, alta execução da performance e existência de gravações de boa qualidade sonora, disponíveis em plataformas digitais. Foram escolhidas as seguintes peças, com seus respectivos movimentos, motivos de escolha e apresentações selecionadas:

Do *Quarteto de cordas* em fá maior, de Maurice Ravel. Foram selecionados para replicação com instrumentos virtuais um trecho do primeiro movimento, o qual apresenta notas curtas intercaladas com melodias longas, bem como o trecho inicial do segundo movimento, cujo tema é composto inteiramente da técnica pizzicato. A performance escolhida para a reprodução dos áudios reais no questionário foi realizada pelo Quarteto Hagen, em apresentação no auditório Mozarteum em 2000 (disponível em: <https://www.youtube.com/watch?v=ehQMe57TPPM>).

Quarteto de cordas em mi menor, op 121, Gabriel Fauré. Desta peça, foi selecionado um trecho do terceiro movimento, que consiste em passagens legato rápidas, as quais transitam por todo o naipe das cordas, acompanhadas de ritmo pulsante tocado em pizzicato pelas outras vozes. A performance escolhida foi executada em 27 de agosto de 2018, pelo quarteto Ebène, no festival Wissembourg (disponível em: <https://tinyurl.com/sud6dpyf>)

Quarteto de cordas nº 3 em Lá maior, op 41, de Robert Schumann. Os trechos utilizados são do primeiro e do terceiro movimento, consistindo em um curto trecho com notas longas, suaves e lentas, e outro trecho com passagens em staccato intercaladas com ataques rápidos em *legato* na melodia principal. A performance escolhida foi executada em 27 de agosto de 2013, pelo quarteto Ebène, no festival Wissembourg, disponível no canal do youtube do festival.

Quarteto de cordas nº 2, “*cartas íntimas*”, de Léo Janáček. Foram retirados um trecho do primeiro movimento, o qual utiliza diferentes estilos de arco e técnicas para atingir um som único, um trecho do segundo movimento, com passagens rápidas e presença de surdina, e, finalmente, o trecho inicial do quarto movimento, que usa técnicas como *staccato* e *marcato*. A performance escolhida foi retirada do quinto festival ChamberFest Cleveland, executada em 24 de junho de 2016 no auditório Mixon, parte do Instituto de música de Cleveland, disponível no youtube, no canal ChamberFest Cleveland.

Os trechos originais, presentes no questionário, foram retirados dessas mesmas performances, levando em conta fatores como qualidade técnica, qualidade sonora, existência e inexistência de ruídos externos como tosses, movimentações e falas que pudessem influenciar a decisão do participante.

2.3 FERRAMENTAS

As réplicas dos trechos escolhidos foram criadas de forma inteiramente virtual. A plataforma de áudio digital da escolha foi o programa Ableton Live, versão 11.0. Este programa permite o controle de múltiplos canais MIDI simultaneamente, facilitando a edição do quarteto como grupo, além de apresentar uma interface dinâmica que proporciona maior fluidez no trabalho .

As bibliotecas de áudio escolhidas foram: Joshua Bell solo violin (embertone); Fischer Viola (embertone); Chris Hein solo Cello (best service); Chris hein solo Viola (best service) e Tina Guo Acoustic Cello (Cinesamples). Nenhuma adição ou alteração foi executada na programação das bibliotecas. Essas escolhas foram baseadas em recomendações e listas que apontam tais bibliotecas como líderes no mercado em termos de qualidade sonora, captação dos *samples* e fidelidade técnica.

O sampleador utilizado para reproduzir as bibliotecas foi o Kontakt Player (native instruments), versão seis, pois as bibliotecas selecionadas foram criadas e otimizadas para utilização com esse plug-in. Ferramentas digitais adicionais foram: Filtro multi bandas Pro-Q 3 (fabfilter), *Reverb* de convolução Altiverb (audio ease), *Stereo Imager* (waves), Izotope Ozone 7, Retro-color 20 (XLN audio).

Todas as ferramentas virtuais foram manuseadas em um notebook Dell Inspiron I7, auxiliado por uma placa de áudio Behringer UMC 204. O áudio foi monitorado usando duas caixas Yamaha HS7 e um fone de ouvido semiaberto, modelo AKG K240 mkII.

Para a realização de algumas automações, foi utilizada uma controladora MIDI Arturia Minilab mkII.

2.4 CRIAÇÃO DAS RÉPLICAS

Para iniciar as imitações, os áudios originais das performances escolhidas foram gravados e abertos no programa de edição Ableton Live, no qual os trechos selecionados, todos variando de 15 a 30 segundos de música, foram isolados e usados como base para estabelecer o andamento médio das versões digitais.

Após os primeiros procedimentos, os áudios originais foram silenciados no programa, e criou-se uma primeira versão digital dos trechos, consistindo em uma transcrição literal das partituras dos trechos para a linguagem MIDI, a qual não levou em conta fatores como variações aleatórias de andamento, pausas não-intencionais, grandes alterações de dinâmica ou mesmo

técnicas diferenciadas notadas na partitura. Cada instrumento virtual apresenta uma programação individual, que o faz lidar de formas diferentes com relação ao disparo, transição e posicionamento das notas. A transcrição inicial garante o mapeamento desses parâmetros, e impede que inconsistências indesejadas, decorrentes das particularidades de cada biblioteca, atrapalhem a programação das peças nos próximos passos.

Na terceira etapa, foram adicionadas as variações de técnicas ao mapeamento MIDI. As bibliotecas utilizadas permitem o acesso a diferentes técnicas com a utilização de *Key Switches*, notas situadas entre dó -2 e dó 0, as quais não estão presentes no campo harmônico dos instrumentos originais, mas, ao serem tocadas, ativam um diferente conjunto de samples respectivos a técnicas variadas como pizzicato, staccato e tremolo.

A quarta etapa consiste no pareamento do andamento e dos intervalos de tempo das réplicas com os trechos originais. Conforme abordado previamente, micro variações no andamento e ataque das notas são características da performance humana (Palmer, C. 2013), e a replicação das mesmas no ambiente virtual é importante para comunicar ao ouvinte a noção da humanidade na peça (Clarke, E.F, 2000). Para isso, foram realizadas automações no metrônomo do programa, de forma que os instrumentos MIDI pudessem se alinhar com as variações de andamento naturais decorrentes da performance humana. Após esse processo, o *Grid* (sistema paramétrico que estabelece a posição das notas MIDI em relação ao tempo dentro de uma DAW) do programa foi desligado e as notas foram ajustadas de forma a serem sincronizadas com os ataques dos trechos originais.

A automação do tempo no programa pode ser interpretada como uma forma de traduzir, para o ambiente digital, certos aspectos naturais e intencionais de uma performance humana. Conforme explorado por Eric Clarke em seu trabalho *imitating and evaluating real and transformed musical performances*, ocorrências como micro variações constantes, que se repetem ao longo de determinados motivos, são fatores naturais e artísticos que desviam da notação musical, muitas vezes não se encontrando presentes nas partituras originais, mas essenciais para a comunicação da expressividade humana em uma performance. Como exemplo, pode-se observar um dos trechos utilizados neste trabalho, retirado do quarto movimento do *Quarteto de cordas Cartas Íntimas*, escrito por Léos Janacek.

IV.

Allegro $\text{♩} = 112$

FONTE: Masters music, 1955.

Enquanto na notação original não existe nenhuma sinalização de variação de tempo, a performance, usada como base para a criação da réplica virtual, apresentava uma figura temporal recorrente, na qual o andamento da música se tornava mais rápido conforme o motivo caminhava da primeira nota até o último compasso desse tema. Esta micro variação se mostrou constante para cada repetição ou reapresentação desse mesmo motivo e foi, portanto, tida como um elemento de micro variação intencional, o qual fazia parte da expressividade artística do grupo em questão. Este elemento foi replicado virtualmente pela automação do metrônomo na DAW, como uma forma de abordar a existência de micro variações musicais que se executam de forma grupal.

Outros tipos de variação, atreladas à disposição dos ataques das notas em relação ao andamento relativo de determinado trecho, são interpretadas como micro variações aleatórias: imperfeições naturais, decorrentes das limitações humanas quanto à execução de um motivo musical de forma completamente precisa (Palmer, C. 2013). Estas variações são, então, traduzidas para as performances digitais pelo desligamento do *grid* e disposição livre das notas no sistema MIDI.

Com as técnicas já mapeadas e os fragmentos sincronizados, faz-se a programação e pareamento das dinâmicas, consistindo na quinta etapa, dividida em três partes. Em primeira instância, foram mapeadas as variações de intensidade e as marcações originais de dinâmicas notadas na partitura. Este processo é feito de forma individual em cada canal MIDI, utilizando da automação de parâmetros dentro do *plugin* Kontakt que controlam propriedades como intensidade, volume e expressão. Em sequência, automações adicionais foram realizadas usando como base os áudios originais, levando em conta fatores como variações expressivas e

divergências sonoras intencionais decorrentes da linguagem musical do *performer*. O foco dessa etapa está em garantir que as réplicas obtenham um grau de humanização e dinamismo relativos aos das performances humanas. Finalmente, um terceiro controle de volume ocorre, dessa vez, levando em conta a sonoridade dos quatro instrumentos como um todo. Esse último procedimento visa controlar possíveis disparidades de volume resultantes das constantes automações e variações de técnicas, balanceando os quatro canais MIDI para que soem com maior unidade.

A sexta etapa consiste na criação do espaço sonoro e da ambientação. Este é um processo importante pois as bibliotecas de cordas, cada qual gravadas em diferentes ambientes com diferentes formas de captação, quando tocadas em conjunto tendem a soar desconexas, mesmo com o mapeamento preciso das técnicas, dinâmicas e expressões das performances originais (Hähnel, T. 2013). O espaço em que se executa uma performance, captado nos microfones, ajuda a firmar o conceito da unidade, do quarteto como um grupo musical, e tem um papel fundamental no estabelecimento da unidade entre os instrumentos (Bergström, I. 2017). O primeiro passo consistiu na simulação virtual de diferentes espaços, utilizando plugins de reverberação como Altiverb (audio ease) e True Verb (eventide). Esses plugins foram adicionados em um grupo contendo a saída dos quatro canais MIDI e suas relações *DRY/WET* foram manuseadas de forma a simular uma captação próxima natural de um quarteto, a qual consiste de ambos os sons diretos gerados pelos instrumentos e suas respectivas reflexões no ambiente.

Ainda na sexta etapa, plugins como S1 *Stereo Imager* (Waves) e PanPot (GoodHertz) são adicionados na cadeia para trabalhar com a espacialização do som, auxiliando na percepção da localização dos instrumentos no espaço criado. Cada réplica de quarteto recebeu um espaço diferente, garantindo que a reverberação não fosse usada como uma forma alternativa de identificar os excertos no questionário.

A última etapa consiste no tratamento do áudio, envolvendo adição de efeitos por grupo, mixagem e masterização. Aqui, são levados em conta fatores como diferenças naturais de volume entre as bibliotecas, pequenos problemas de fases em stereo e conceitos gerais de mixagem, como disputas entre frequências e inconsistências harmônicas, possivelmente geradas pelas adições dos *plugins*. Estes ajustes são realizados ao final do processo para garantir que a expressividade original, programada nas bibliotecas, não seja alterada ou inibida por fatores externos, como controles de volume e espectro sonoro. Foi utilizado o equalizador ProQ 3 (Fabfilter) para ajustar frequências conflitantes, e o plugin Retro Color 20 (XLN) foi empregado para simular um filtro de microfone e também adicionar um ruído ambiente, fatores que ajudam a reforçar a ideia da gravação de uma performance. O supressor harmônico Soothe, da empresa

oeksound, foi adicionado para controlar possíveis desbalanceamentos no espectro sonoro, gerados pela adição de múltiplos efeitos em cadeia, garantindo uma sonoridade controlada para a mixagem.

As réplicas e os trechos originais, selecionados para o trabalho, foram masterizados seguindo as recomendações de volume para plataformas de *streaming* digital, definidas pela União Europeia de Radiodifusão (relatório geral, 2020). Todas as réplicas foram exportadas do programa em formato WAV, 44.1k hz e 16 bits, garantindo a maior qualidade e fidelidade sonora.

3 ANÁLISE DOS DADOS

3.1 ANÁLISE GERAL

Ao total, foram contabilizadas 281 respostas, coletadas ao longo de duas semanas. Testes incompletos foram desconsiderados, resultando em 260 respostas válidas. Os dados foram retirados dos formulários e realocados para uma planilha do Excel, onde as respostas foram analisadas de forma total e, posteriormente, de forma específica, distinguindo entre as expertises musicais selecionadas ao fim do teste.

É importante reiterar que os termos “respostas corretas” e “acertos” são pertencentes a diferentes métodos analíticos. O primeiro termo faz parte da análise bruta dos dados, levando em conta a pontuação recebida pelos participantes de acordo com o número de vezes em que o mesmo corretamente discerniu a procedência dos excertos. O termo “acertos” faz parte da metodologia da detecção de sinais, e consiste em uma das quatro possíveis classificações resultantes da escolha de uma resposta.

A análise geral bruta das respostas coletadas resultou em uma taxa média de respostas corretas de 62.75%, valor abaixo do limite de 70% estabelecido por Turing no *Jogo da Imitação*. A taxa de acertos (*Hits*) correspondeu a 61%, enquanto a taxa de alarmes falsos atingiu um valor de 32%. Estes valores podem ser utilizados, na teoria da detecção de sinais, para o cálculo das variáveis Sensibilidade (d') e Critério (C) conforme as seguintes equações:

$$d' = z(0.61) - z(0.32) \rightarrow d' = 0.74$$

$$C = -0.5 \times [z(0.61) + z(0.32)] \rightarrow C = 0.094$$

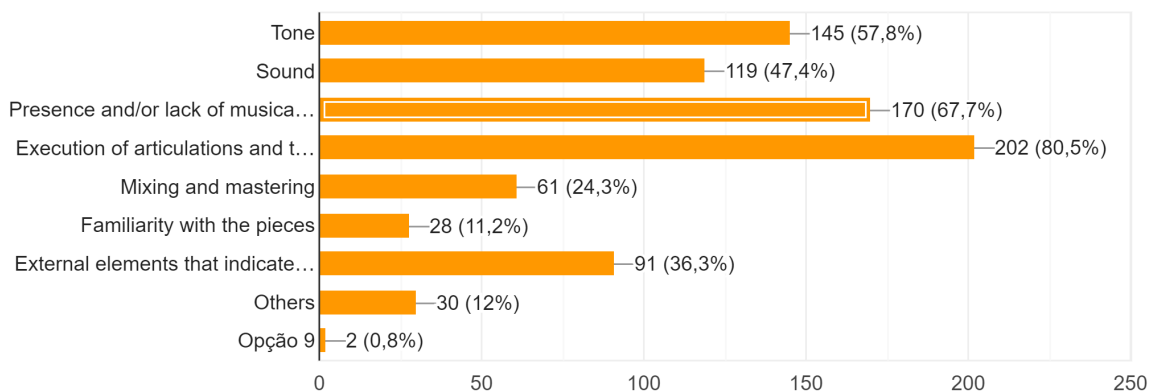
Este valor da especificidade é qualificado como conservador, caracterizado pela predominância de omissões em contrapartida a um baixo número de falsos alarmes.

Analisando as perguntas de forma individual, a questão com a maior taxa de respostas corretas atingiu 83%, e consistia de um fragmento do quarteto de cordas de Janacek, replicado por instrumentos virtuais. A pergunta com maior taxa de erros foi, também, um trecho de Janacek, desta vez tocado por instrumentos reais e erroneamente classificado como virtual pelos participantes, com uma taxa de 57%.

No que tange às opções apresentadas ao fim do questionário, relativas aos elementos que influenciaram na escolha das respostas, as mais escolhidas foram: execução das articulações e técnicas, com 80%; presença e/ou falta de expressividade, 67%; timbre, com 58% e sonoridade, com 47%.

As opções tidas como de menor relevância para a escolha das respostas consistem em: familiaridade com as peças tocadas, com um valor de 11%; mixagem e masterização, com 24% de escolha e finalmente, presença de barulhos/ruídos alheios as peças, tido por 36% das pessoas como um fator decisivo para a distinção entre instrumentos reais e virtuais.

Quais fatores influenciaram na escolha das respostas?



Traduzido de cima para baixo: Timbre, Sonoridade, Presença e/ou falta de expressividade musical, execução das articulações, mixagem/masterização, familiaridade com as peças, elementos externos como ruídos e outros.

3.1 ANÁLISE POR GRUPOS

A distribuição das profissões/expertises musicais, coletadas ao fim do trabalho, é a seguinte: Músicos amadores/hobby, 23,8%; Produtor musical, 13,8%; Compositor/Arranjador, 11,9%; Bacharel/Mestre/PhD em música, 11,5%; Instrumentista, 11,2%; Estudante de música, 10,8%; Compositor de trilha sonora, 7,3%.

Outras categorias, consistindo de Maestro/conductor, pesquisador em música, artista de mixagem/masterização, *sound designer* e técnico de áudio, somaram 9,7% das respostas, e não foram analisadas individualmente de forma detalhada, devido à sua baixa taxa amostral. Após uma coleta preliminar dos dados, as categorias “Artista de mixagem/masterização” e “Professor de música”, foram aglomeradas, respectivamente, às categorias “Produtor Musical” e “Estudante de música”, usando como parâmetro a similaridade da expertise musical, bem como tendências semelhantes para a escolha das respostas.

O grupo Bacharel/Mestre/PhD em música teve a maior taxa de respostas corretas, figurando em 71%. A taxa de acertos e falsos alarmes ficou, respectivamente, em 65% e 22%, resultando em um cálculo de sensibilidade de 1,114. O cálculo da especificidade resultou em um critério de 0.185, definido pela teoria como conservador, delineando uma tendência do grupo em não caracterizar instrumentos virtuais como sendo reais. Dentre as opções apresentadas ao final do trabalho, “Presença ou falta de expressividade musical” e “timbre” foram as respostas mais recorrentes para esse grupo, presentes em 100% dos questionários enviados por participantes dessa classe.

Apenas um outro grupo foi capaz de pontuar acima do limite de 70%, estabelecido por Turing. Os participantes que marcaram suas expertises musicais como “Compositor de trilha sonora” tiveram uma taxa de respostas corretas de exatamente 70%. As taxas de acerto e falsos alarmes ficaram, respectivamente, 67% e 26%, resultando em um cálculo de sensibilidade de 1,083. A especificidade, com um valor de 0.102, qualifica o critério desse grupo como conservador.

O grupo “Compositor/Arranjador” apresentou uma taxa de 69% de respostas corretas, podendo ser considerada na margem de erro de 1 ponto, proposta por Turing. Com uma taxa de acertos de 68% e uma taxa de falsos alarmes de 30%, este grupo atingiu uma sensibilidade de 0,9 e uma especificidade de 0,028. Enquanto ainda tido como conservador, valores muito próximos ao zero podem ser interpretados como um critério relativamente neutro.

Por apresentarem resultados e padrões de escolha semelhantes, os grupos “Estudante de música” e “professor de música” foram analisados conjuntamente. Juntos, assumiram uma taxa de respostas corretas de 64%. As taxas de acertos e falsos alarmes ficaram, respectivamente, 62% e 32%, resultando em uma sensibilidade de 0,77 e uma especificidade conservadora de 0,081.

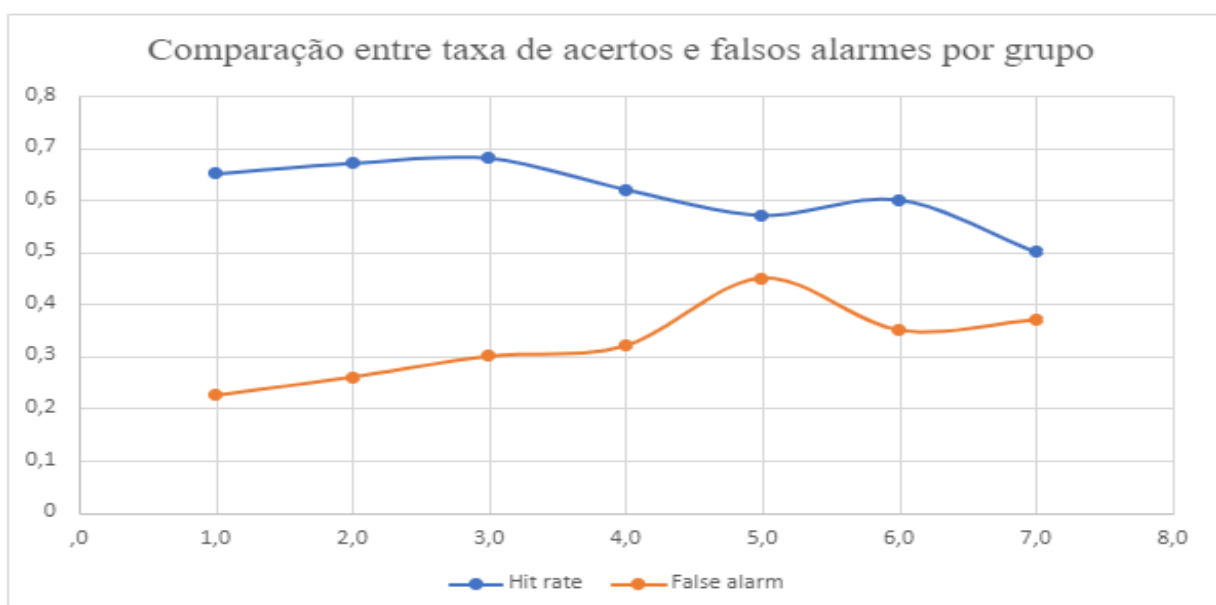
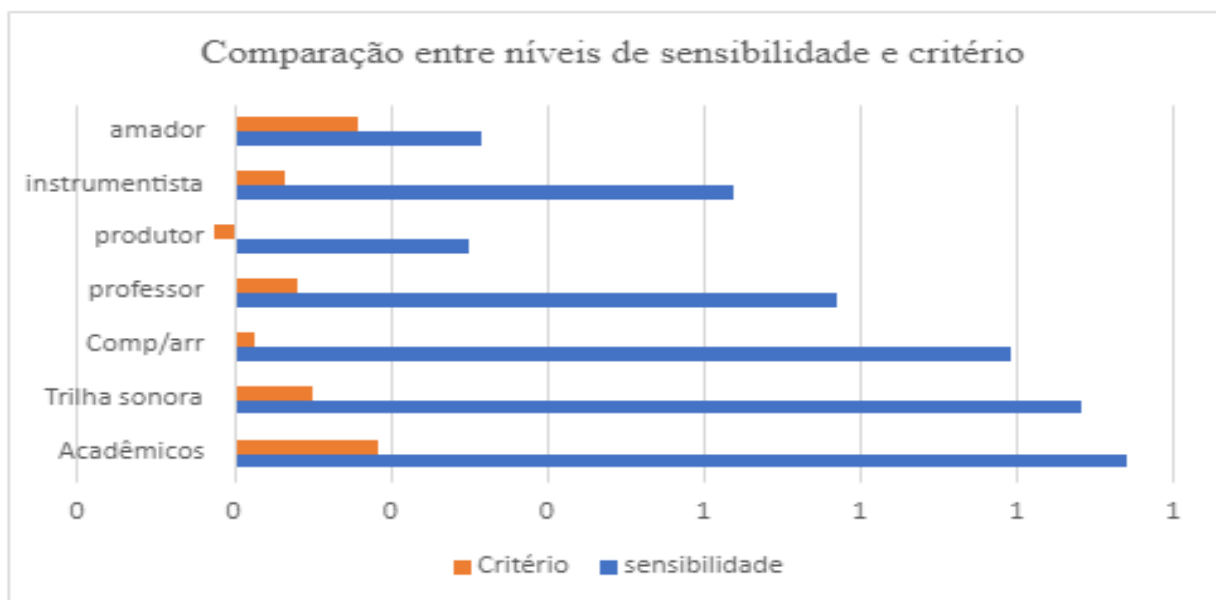
Outras categorias que apresentaram resultados semelhantes e foram, portanto, agrupadas, consistem de “Artista de Mixagem/Masterização” e “Produtor Musical”. Com 56% de respostas corretas, este grupo foi o único a apresentar um critério considerado liberal, com um valor de -0,025, mas que também pode ser tido como relativamente neutro. Os acertos e falsos

alarmes, respectivamente, 57% e 45%, resultam de uma sensibilidade de 0,3, representando a taxa mais baixa dentre as categorias presentes.

O grupo “Músico Amador” atingiu valores semelhantes à categoria anterior: com 57% de respostas corretas, 50% de acertos e 37% de falsos alarmes, os valores de sensibilidade e especificidade somaram, respectivamente, 0,31 e 0,16.

Finalmente, o grupo “Instrumentistas” apresentou uma taxa de respostas corretas de 61%. Os valores de Acertos, 60%, e falsos alarmes, 33,5%, resultam em uma sensibilidade de 0,63 e um critério de 0,66.

Uma comparação entre os diferentes valores de sensibilidade e critério obtidos por grupo, bem como a relação de acertos e falsos alarmes, podem ser observados nos gráficos a seguir:



Uma categoria com resultados interessantes, mas que não será considerada aqui por sua baixa representação (apenas 3 respostas), é o grupo de Maestros. Este grupo apresentou 94% de respostas corretas, não manifestando nenhum falso alarme ao longo dos testes enviados. Uma vez que as outras categorias também apresentaram casos isolados em que todas as respostas foram corretamente assinaladas, seria inconclusivo realizar qualquer análise comparativa entre este grupo e os outros, dada a diferença amostral.

É interessante observar que existiu uma predominância da escolha das opções “Timbre”, “Presença e/ou falta de expressividade musical” e “articulações” dentre as três categorias que apresentaram os melhores resultados no teste. Fatores como “elementos externos como ruídos” e “sonoridade” se mostraram mais recorrentes entre grupos com pontuações menores no questionário.

4 DISCUSSÃO

4.1 RESULTADOS ENCONTRADOS X RESULTADOS ANTERIORES

Esta pesquisa encontrou resultados que destoam dos dados previamente encontrados por outros trabalhos. As taxas de respostas corretas foram significativamente mais baixas do que os exemplos prévios, sinalizando, com uma porcentagem de 62,75%, uma maior dificuldade na diferenciação entre exemplos reais e virtuais. Este resultado foi abaixo do esperado, uma vez que as pesquisas encontradas na literatura de apoio sinalizaram, muitas vezes, uma alta capacidade de distinção entre instrumentos reais e virtuais.

Nota-se ainda que esta pesquisa lidou com instrumentos solo, em contrapartida às pesquisas prévias, que trabalharam com ensembles maiores, cujas bibliotecas de áudio eram compostas de naipes de instrumentos. Replicar a sonoridade de instrumentos individuais é comumente associada a uma tarefa muito mais desafiadora do que trabalhar com naipes, uma vez que a expressividade dos instrumentos solo é tida como mais acentuada e intrincada de se replicar virtualmente (Pejrolo, A. 2007). Logo, era de se esperar que esta pesquisa obtivesse níveis de diferenciação maiores do que os previamente apresentados, mas o mesmo não se verificou. É possível que esta mesma pesquisa, se aplicada agora com a utilização de bibliotecas orquestrais, obtivesse resultados ainda mais significativos para os instrumentos virtuais.

Em seu trabalho *Musical Plausibility: estimation and synchronization*, publicado em 2013, Tilo Hahnel encontrou que a qualidade sonora do excerto em questão pode ser um dos

fatores que induzem os ouvintes a classificarem uma performance como sendo real ou virtual, sendo que amostras de baixíssima qualidade sonora se revelaram indiscerníveis quanto a sua procedência.

Enquanto não existiram, nos questionários desenvolvidos para esta pesquisa, alguma medida ou parâmetro para analisar, de forma direta, a qualidade dos áudios e sua influência na percepção sonora, uma análise sobre os valores de Falsos Alarmes (32%) e Omissões (39%) revela uma tendência dos participantes em classificarem, com maior frequência, performances reais como virtuais, em contrapartida a classificarem performances virtuais como reais. Esta tendência pode sinalizar que um dos parâmetros auditivos utilizados para a escolha das respostas foi a qualidade sonora, com performances de concebida menor qualidade sonora sendo classificadas como virtuais. É importante notar, porém, que as opções “mixagem/masterização” e “sonoridade” figuraram entre as menos escolhidas dentre os fatores que influenciaram na distinção, sendo mais comuns em grupos que obtiveram menores níveis de sensibilidade.

Um fator trazido por pesquisas anteriores e detalhado no trabalho *Replacing the Orchestra? – The Discernibility of Sample Library and Live Orchestra Sounds*, publicado em 2017 por Reinhard Kopiez, é a diferença de discernimento entre expertises musicais diversas. O trabalho em questão encontrou resultados que se dispuseram em dois grupos: o primeiro, composto por produtores musicais, arranjadores, compositores e maestros, obteve uma taxa de respostas corretas de 80% e um valor de sensibilidade igual a 2,09, enquanto um segundo grupo, formado por estudantes de música, músicos amadores, instrumentistas e professores de música, atingiu uma taxa de 68%, com um valor de sensibilidade de 1,37. Em ambos os grupos, os valores de especificidade se igualaram a zero, significando a inexistência de uma tendência de enviesamento entre performances reais e virtuais.

Os resultados desta pesquisa diferem dos encontrados por Kopiez, uma vez que nenhum grupo, mesmo os classificados como de alta expertise musical, atingiram valores acima de 71%, com a média geral se encontrando quase 10 pontos abaixo dos resultados totais de Kopiez. Um dos fatores que influenciam nesta disparidade entre as pesquisas está na presença da especificidade como uma variável a ser considerada, enquanto nas pesquisas anteriores, os valores para essa medida foram quase sempre neutros. A análise dos valores de sensibilidade e especificidade desta pesquisa revelam que, ao ouvir os exemplos e responder os questionários, participantes demonstraram uma tendência a classificar erroneamente trechos reais como virtuais. Este fator pode estar relacionado com a presença de quatro diferentes performances reais nesta pesquisa, uma vez que a pesquisa de Kopiez utilizou apenas uma gravação real em seu questionário, fragmentada em diferentes trechos.

Retirar os excertos reais de quatro fontes diferentes pode ter criado uma situação em que os participantes, quando confrontados com a incerteza e a qualidade geral das réplicas digitais, passaram a classificar trechos reais, com menor qualidade sonora e performática, como sendo digitais. Este argumento pode encontrar suporte ao se analisar algumas das respostas deixadas no espaço aberto após o questionário, no qual os participantes tinham a opção de deixar um texto relatando sua experiência e metodologia ao responder o teste. O fragmento abaixo é uma tradução de um comentário deixado por um arranjador/compositor, o qual terminou o teste com uma taxa de acertos de 78%:

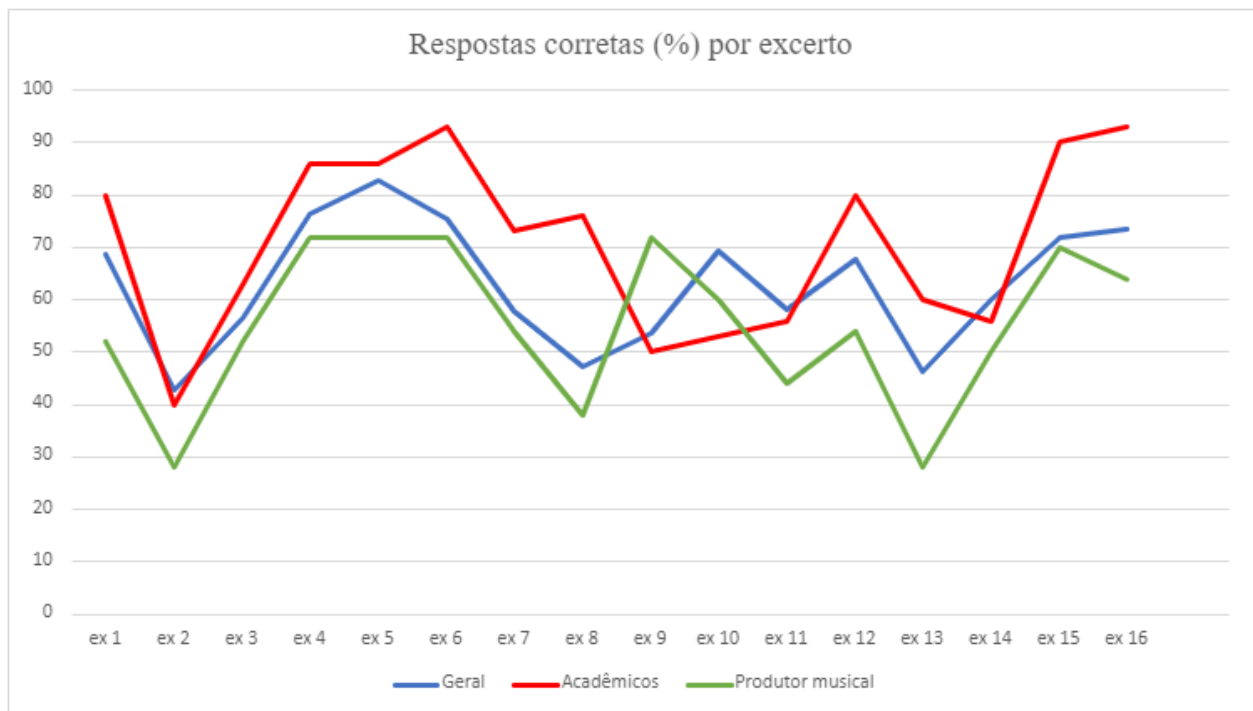
Sendo bem sincero, estas foram algumas das interpretações mais realistas de instrumentos que eu ouvi, todos os excertos soaram reais. Eu classifiquei alguns como virtuais, baseado no que eu diria como sendo um som mais rígido, mas se alguém me falasse que os mesmos trechos eram reais, eu não ficaria surpreso.

Enquanto muitos participantes foram capazes de discernir instrumentos reais de virtuais, a alta incidência de omissões, e a predominância do critério conservador nas respostas, denota que a qualidade geral das imitações foi suficiente para estabelecer um ambiente de incerteza, no qual participantes inclinaram-se para a classificação como virtual. Os valores de sensibilidade mais baixos, obtidos por esta pesquisa, podem ser então, uma decorrência natural dos avanços tecnológicos na criação de instrumentos virtuais nos últimos cinco anos.

Bibliotecas de cordas modernas, como Joshua Bell Solo Violin e Chris Hein solo Cello, utilizam uma mistura de extenso trabalho de gravação com aplicação de técnicas digitais, as quais permitem a manipulação de efeitos que atuam sobre os samples captados. Mesmo participantes que pontuaram acima da média de Turing e foram capazes de identificar a maior parte dos instrumentos virtuais, demonstraram surpresa, tanto com a qualidade sonora dos excertos tocados, como com a capacidade de reproduzir, de forma expressiva, variações de dinâmica e intensidade. É plausível que as diferenças dos resultados desta pesquisa para com suas antecessoras sejam decorrência, em parte, da superioridade sonora e artística dos instrumentos virtuais mais recentes.

Mesmo com os resultados gerais e grupais pontuando abaixo das pesquisas prévias, é importante notar que existiu, ainda, uma diferença significativa entre as expertises musicais e seus padrões de resposta. Classes como “músico amador” e “produtor musical” obtiveram taxas de respostas corretas abaixo da média e, destas duas classes, a composta por produtores musicais chama atenção por situar-se entre as piores pontuações, enquanto nos trabalhos anteriores, este grupo normalmente figurava entre as expertises de maior capacidade discriminativa. O gráfico

abaixo mostra, de forma comparativa, como foi o padrão de respostas deste grupo em relação aos valores gerais e de maior pontuação.



Frequentemente fluando abaixo de 50% de respostas corretas, este grupo apresenta também, por consequência, altos valores de omissões e falsos alarmes, resultando em um nível de sensibilidade baixo. Na teoria da detecção de sinais, valores de d' menores ou iguais a zero são considerados frutos de um teste aleatório ou de uma inexistência de padrões perceptíveis, uma vez que se mostra claro a incapacidade dos participantes em estipular fatores racionais para a distinção de dois estímulos. Este acontecimento é representado, de forma concreta, quando as porcentagens de omissões e falsos alarmes são superiores às proporções de acertos e rejeições corretas (Macmillan, N.A. ; Creelman, C.D. 2005).

O elemento da aleatoriedade não se concretizou neste teste, mas os valores de sensibilidade para os dois grupos de menores pontuações figuraram em 0,3, um valor próximo a zero. Este número nos permite analisar, de forma concreta, alguns dos fatores que podem ter influenciado nas tomadas de decisões dessas classes. Um valor baixo de porcentagem de acertos denota que estes grupos provavelmente se guiaram por fatores de diferenciação que não estavam presentes em todos os excertos, uma vez que o padrão de discriminação não se mantém constante ao longo das respostas referentes aos estímulos reais.

Um exemplo disso é a presença de barulhos de fundo, como tosses, espirros, movimentações e ruídos gerais. Tais fatores estavam presentes em alguns dos excertos reais que demonstraram maior índice de acertos entre as duas classes. É importante notar que as duas menores pontuações por excerto, referentes ao número 2 e número 8 no gráfico apresentado acima, correspondem a trechos reais retirados dos quartetos de Janacek e Ravel, os quais apresentavam qualidade sonora limpa e nenhuma interferência de barulhos externos. Destaca-se então, que a opção “presença de ruídos externos” teve maior representatividade entre as duas categorias com menor pontuação, sinalizando que as mesmas não apresentavam as técnicas auditivas e perceptivas necessárias para distinguir, de forma assertiva, a procedência dos diferentes estímulos utilizando critérios universais como timbre e articulação das notas, tendo que recorrer a fatores adjacentes que pudessem servir como um atalho para facilitar a tomada de decisão.

Os baixos valores de especificidade, obtidos por ambos os grupos, sugerem a preferência por um determinado tipo de som, fator que influencia na escolha das respostas e fundamenta a ideia de que o principal critério de diferenciação, utilizado pelo mesmos, não se aplicava a todos os estímulos: excertos reais que não apresentavam ruídos de fundo evidentes foram classificados, muitas vezes, como virtuais, resultando em uma taxa de omissão quase parelha a taxa de acertos para o grupo “produtores musicais”, e perfeitamente igual, para o grupo “músicos amadores”. Ao perceber que essas ocorrências não seriam constantes ao longo dos exemplos, estes grupos desenvolveram a tendência liberal de respostas, tendo uma alta propensão a chamar um instrumento virtual de real, observado pela alta taxa de falsos alarmes (45% e 37%).

A análise dos grupos com maior pontuação revela, no entanto, um caráter diferenciado: a predominância do critério conservador demonstra a tendência desses indivíduos em classificar uma quantia de performances reais como virtuais. Este dado nos permite estipular que os grupos formados por compositores, arranjadores e acadêmicos de música focaram em fatores que estavam presentes em todas as performances, como micro-variações, entonação e articulação. Logo, os quesitos de qualidade sonora e presença de ruídos de fundo, que influenciaram a percepção dos grupos analisados previamente, não constituem aqui, um fator tão proeminente, uma vez que existe uma porcentagem maior de omissões a falsos alarmes.

Uma possível explicação seria a de que estes grupos apresentam uma audição treinada para identificar nuances nas performances, as quais passaram despercebidas por grupos cujas expertises musicais exigem diferentes qualificações e aptidões. Esta proposição pode ser apoiada pelo fato de que os excertos musicais com menores variações de técnicas e dinâmicas obtiveram maiores índices de falsos alarmes, se comparados a exemplos com grande quantidade de

informação e variação. Esses indivíduos, ao focar nas qualidades expressivas e em detalhes das transições entre as notas, acabaram por classificar como virtuais, trechos que naturalmente não apresentavam tais fatores.

Este argumento também justifica o fato de que, um dos trechos virtuais com maior proporção de falsos alarmes para os grupos de maior pontuação, era um excerto replicado do segundo movimento do quarteto de cordas de Ravel, composto majoritariamente pela escolha estilística da técnica pizzicato. Um dos comentários, deixado por um indivíduo da categoria de bacharel/mestre/PhD, aborda essa questão:

Pizzicatos e articulações breves são mais difíceis de distinguir. Particularmente, eu acredito que [elas] são mais fáceis de programar também. Legato é mais fácil de distinguir, especialmente se existe um efeito de glissando entre as notas. Estes são raramente efetivos

É particularmente interessante e inesperado que a categoria com maior valor de sensibilidade seja a formada por acadêmicos de música. Profissionais de trilha sonora, compositores e arranjadores modernos atuam intrinsecamente com instrumentos virtuais (Morgan, T. 2013), então era esperado que seus resultados fossem significativamente mais expressivos do que os apresentados por outras categorias, que não demonstram a mesma familiaridade com práticas virtuais de orquestração e composição. Ainda assim, enquanto os valores de acertos foram semelhantes entre as 3 categorias no topo, o grupo formado por Bacharéis, Mestres e PhDs em música apresentou as menores taxas de falsos alarmes, significando a maior sensibilidade quanto aos fatores que categorizam um excerto como virtual.

4 CONSIDERAÇÕES FINAIS

Resultados encontrados nesta pesquisa foram significativamente mais baixos do que os previamente verificados por trabalhos anteriores, estipulando uma menor capacidade discriminatória entre performances reais e virtuais. Com uma taxa geral de respostas corretas de 62,75%, abaixo do limite de 70% proposto por Alan Turing no *Jogo da Imitação*, estaria qualificada a capacidade dos instrumentos virtuais em replicar, ao menos em parte, algumas das características da expressividade humana na performance.

Cinco anos separam a confecção das réplicas utilizadas nesta pesquisa, das réplicas utilizadas na pesquisa de Kopiez, em 2017. Uma das questões trazidas por este trabalho é o

quanto a tecnologia envolvendo instrumentos virtuais evoluiu com os últimos produtos disponíveis no mercado. O fato de que os participantes deste teste apresentaram resultados muito menos assertivos sinaliza que as novas tecnologias, empregadas na criação de instrumentos virtuais recentes, como Joshua Bell Solo Violin e Chris Hein solo Cello, realmente proporcionam maior realismo e expressividade musical para composições virtuais.

Em contrapartida, o fato de que grupos de certas expertises foram capazes de atingir altas porcentagens de acertos demonstra que, mesmo com os avanços tecnológicos, ainda existem fatores exclusivos das performances reais. Entre os aspectos mais influentes no processo de distinção desses grupos, encontram-se a realização das articulações, diferenças tímbricas relacionadas a variações de intensidade e expressividade geral de transições, como *legato* e *glissando*. Os comentários deixados pelos indivíduos dessas classes comumente aludiram a alta qualidade sonora das réplicas e a dificuldade geral do teste, explicando que o processo de distinção muitas vezes se resumiu a nuances e trejeitos sonoros sutis.

Os grupos cujas pontuações e níveis de sensibilidade atingiram os menores valores, demonstraram uma tendência geral de discriminação pela sonoridade e presença de ruídos, em detrimento de outros fatores. Comentários deixados por esses indivíduos constantemente remetiam à ideia de que nenhum dos excertos soava virtual, demonstrando que tais grupos não identificavam, de forma clara, os fatores que se mostraram cruciais para a diferenciação atingida pelas expertises de melhor pontuação.

De modo geral, sob a óptica da capacidade de imitação de instrumentos virtuais, os resultados encontrados por esta pesquisa foram promissores, e são mais expressivos levando-se em consideração que o teste foi realizado com instrumentos solo, em contrapartida a formações orquestrais com naipes. As réplicas foram capazes de produzir uma quantidade significativa dos trejeitos humanos, de modo que o ambiente geral do teste e a dificuldade do discernimento levou participantes, mesmo os de expertise focada nessa área, a erroneamente qualificar algumas das performances reais como virtuais.

Enquanto ainda existem trejeitos e particularidades que diferenciam as performances reais das virtuais, o surgimento de novas tecnologias como *Machine Learning* e suas possíveis aplicações à música, algumas já em curso, podem sinalizar um futuro não muito distante em que, instrumentos virtuais, equipados de uma complexa estrutura de programação e síntese, são capazes de replicar, em total extensão, a expressividade da performance humana.

REFERÊNCIAS

BAUDRILLARD, J. **Simulacra and simulacrum**. 1.ed. France: Editions Galilee. 1981.

BERGSTRÖM, I.; AZEVEDO, S.; PAPIOTIS, P.; SALDANHA, N.; SLATER, M. The plausibility of a string quartet performance in virtual reality. **IEEE transactions on visualization and computer graphics**. v. 23, n. 4, p.1332, 2017.

BODEN, M.A. creativity and artificial intelligence. **artificial intelligence**, Brighton, v. 103, n. 1, p. 347-365, 1998.

CLARKE, E.F. Imitating and evaluating real and transformed musical performances. **Music perception**, California, v. 10, n.3, p. 317 c- 341, 1993.

CLARKE, E.F.; WINDSOR, W.L. Real and simulated expression: a listening study. **Music perception**, California, v. 17, n. 3, p. 277-313, 2000.

D' ERRICO, M.A. **Behind the beat: technical and practical aspects of instrumental hip-hop composition**, tese (mestrado em artes e música) - TUFTs university, Massachussets, 2011.

EXARCHOS, M. Boom bap ex machina: hip-hop aesthetics and the Akai MPC. In: **Producing Music. Perspectives on Music Production**. New York, 13 de junho de 2018. Disponível em: <https://repository.uwl.ac.uk/id/eprint/5150/>

GLOWINSKI, D.; MANCINI, M.; COWIE, R.; CAMURRI, A.; CHIORRI, A.; DOHERTY, C. the movements made by performers in a skilled quartet: a distinctive pattern, and the function that it serves. **Frontiers in psychology**. v. 4, 13 de novembro de 2013. Disponível em: 10.3389/fpsyg.2013.00841

GODLOVITCH, S. innovation and conservatism in performance practice. **The journal of aesthetic and art criticism**. v. 55, n. 2, p. 151-168, 1997.

GROGONO, P. Musys: software for an electronic music studio. **Software: practice and experience**, v. 3, n. 4, p. 369-383, 1973.

HÄHNEL, T. Musical plausibility, estimation and synchronization. **Musicae Scientiae**, v. 17, n. 4, 8 de agosto de 2013, disponível em: <https://doi.org/10.1177/1029864913492531>

KELLEN, D.; NUNES, L.; GARCIA-MARQUES, L. Sensibilidade e bom senso: princípios fundamentais da teoria de detecção de sinal na investigação em psicologia. Lisboa, 2008. disponível em: <https://doi.org/10.14417/lp.694>

KLEIN, E. Feigning humanity. **IASPM@journal**. Queensland, v. 6, n. 2, 2016. disponível em: [10.5429/2079-3871\(2016\)v6i2.3en](https://doi.org/10.5429/2079-3871(2016)v6i2.3en)

KOPIEZ, R.; WOLF, A.; PLATZ, F.; MONS, J. Replacing the Orchestra? – The Discernibility of Sample Library and Live Orchestra Sounds. **PLoS ONE**, Estados Unidos, v.11, 6 de julho de 2016. Disponível em: [e0158324. doi:10.1371/journal.pone.0158324](https://doi.org/10.1371/journal.pone.0158324)

MACMILLAN, N.A.; CREELMAN, C.D. **Detection theory -- a user's guide**. 2.ed. New Jersey: Lawrence Erlbawn, 2005.

OEHLER, M.; REUTER, C.; vibrato experiments with bassoon sounds by means of the digital pulse forming synthesis and analysis framework. In: AUDIO ENGINEERING SOCIETY CONVENTION, 123., 2007, New York. **Anais ... audio engineering society convention paper, new york, 2007**.

PEJROLO, A.; DEROSA, R. **Acoustic and MIDI orchestration for the contemporary composer**. 1.ed. Oxford: Focal Press, 2007.

PALMER, C. music performance: music and coordination. In: DEUSTCH, D. **The psychology of music**. 3.ed. California: Elsevier, 2013.

PALMER, C. Mapping Musical Thought to Musical Performance. **Journal of Experimental Psychology: Human Perception and Performance**, New York, v. 15, n. 12, p. 331-346, 1989.

WINDSOR, W.L.; DESAIN, P.; PENEL, A.; BORKENT, M. A structurally guided method for the decomposition of expression in music performance. **The journal of the acoustical society of america**, v. 119, n. 1182, 31 de janeiro de 2006. Disponível em: [doi-org.ez22.periodicos.capes.gov.br/10.1121/1.2146091](https://doi.org/10.1121/1.2146091)

WOODY, R.H. The Relationship Between Explicit Planning and Expressive Performance of Dynamic Variations in an Aural Modeling Task. **Journal of Research in Music Education**, v.47, n.4, p. 331-342, 1999.

WOODY, R.H. Learning Expressivity in Music Performance: An Exploratory Study. **Research Studies in Music Education**, v. 14, n. 1, p. 14-23, 2000.

