

A Criação de um *Dataset* Audiovisual Brasileiro de Expressões Emocionais

Luisa Medina Fermino Carlos

Paradigma Centro de Ciências e Tecnologia do Comportamento

Mestrado em Análise do Comportamento Aplicado

Profº. Drº. Felipe D'Alessandro F. Corchs

Agosto, 2022

A Criação de um *Dataset* Audiovisual Brasileiro de Expressões Emocionais

Luisa Medina Fermino Carlos

Dissertação apresentada como parte dos requisitos para a obtenção do título de mestre profissional em análise do comportamento aplicada na Associação Paradigma Centro de Ciência e Tecnologia do Comportamento sob orientação do Prof^o. Dr^o. Felipe D'Alessandro F. Corchs

São Paulo

2022

Resumo

As emoções são um processo que envolve múltiplas variáveis, abrangendo alterações psicofisiológicas do indivíduo. Permeiam todas as relações humanas e cumprem uma função essencial na comunicação. A importância do reconhecimento de ocorrências emocionais é notável em inúmeras atividades humanas e há diversas formas de expressar e reconhecer uma emoção, tendo a cultura um importante papel na formação de nossa comunicação emocional. O objetivo do presente trabalho foi desenvolver um *dataset* de brasileiros encenando expressões emocionais com material gravado no formato audiovisual, denominado *Dataset Audiovisual Brasileiro de Expressões Emocionais* (*Brazilian Audiovisual Dataset of Emotions* [BADEM]). Para isso, baseamo-nos majoritariamente em partes de métodos utilizados para criação de *datasets* similares. Descrevemos a construção do *dataset*, contendo cenas de doze atores declarando diferentes frases neutras em sete expressões emocionais. O material desenvolvido foi avaliado por psicólogos por meio de testes de reconhecimento das emoções e julgamento quanto ao grau de naturalidade de suas representações. De acordo com os resultados obtidos, é possível notar a discrepância do reconhecimento entre algumas emoções, atores e avaliadores. Dos 1008 vídeos gravados, 83,82% foram identificados corretamente, e mais de 90% dos vídeos foram avaliados com as emoções encenadas como “moderada” a “completamente” naturais. Os resultados revelam um nível de acurácia nas avaliações das expressões emocionais, como também um grau de naturalidade nas encenações que sugerem que a metodologia aqui proposta produziu um material de qualidade na direção dos objetivos propostos.

Palavras-chave: *dataset* audiovisual, expressões emocionais, reconhecimento de emoções

Abstract

Emotions are a process that involves multiple variables, causing psychophysiological changes in the individual. They permeate all human relationships and play an essential role in communication. The importance of recognizing emotional occurrences is remarkable in numerous human activities, and there are several ways to express and recognize an emotion, with culture having an important role in the formation of our emotional communication. The objective of the present work was to develop a dataset of Brazilians staging emotional expressions with material recorded in audiovisual format, named Brazilian Audiovisual Dataset of Emotions (*Dataset Audiovisual Brasileiro de Expressões Emocionais* [BADEM]). For this, we rely mostly on parts of methods used to create similar datasets. We describe the construction of the dataset containing scenes of twelve actors declaring different neutral phrases in seven emotional expressions. The material created was evaluated by psychologists through tests of emotion recognition and judged for its degree of naturalness. According to the results obtained it is possible to notice the discrepancy of recognition between some emotions, actors and evaluators. Of the 1008 videos recorded, 83.82% were correctly identified, and more than 90% of the videos were rated with the emotions staged as “moderate” to “completely” natural. We believe that the results reveal a level of accuracy in the evaluations of emotional expressions as well as a degree of naturalness in the expressions that suggest that the methodology proposed here produced a quality material towards the proposed objectives.

Keywords: audiovisual dataset, emotional expressions, emotion recognition

Sumário

Resumo	3
Abstract	4
Lista de tabelas e figuras	7
A Criação de um <i>Dataset</i> Audiovisual Brasileiro de Expressões Emocionais	8
Método	13
<i>Atores</i>	14
<i>Avaliadores</i>	15
<i>Frases</i>	15
<i>Procedimentos</i>	16
Gravação dos Vídeos.....	16
Instrumentos e Cenário.....	17
Avaliação do Material Audiovisual.....	19
Resultados	20
Discussão	25
Conclusão	27
Referências	28
Apêndices	34
<i>Apêndice 1</i>	34
Termo de compromisso e responsabilidade – ator / atriz	34
<i>Apêndice 2</i>	36

Termo de compromisso e responsabilidade – avaliadores do banco de dados audiovisuais.....	36
<i>Apêndice 3</i>	37
Termo de compromisso de utilização de dados – Dataset Audiovisual Brasileiro De Expressões Emocionais (BADEM).....	37
Term of commitment for the use of data - Brazilian Audiovisual Dataset of Emotional Expressions (BADEM) (versão em inglês).....	38
<i>Apêndice 4</i>	39
Questionário de avaliação das emoções	39
<i>Apêndice 5</i>	42
Protocolo de gravação dos vídeos de expressões emocionais	42
<i>Apêndice 6</i>	44
Protocolo geral para atores	44
<i>Apêndice 7</i>	46
Roteiro individual (exemplo ator 1).....	46
<i>Apêndice 8</i>	49
Relação de acertos de cada frase e as emoções	49
<i>Apêndice 9</i>	50
Resultados dos testes de qui-quadrado de Pearson com comparação post-hoc por pares (pairwise) e correção de Holm para múltiplos testes de hipóteses	50

Lista de tabelas e figuras

Tabela 1- Frases empregadas nas gravações.....	14
Tabela 2- Resultado da avaliação dos vídeos pelos avaliadores conforme cada emoção.....	19
Tabela 3- Síntese dos vídeos com as emoções identificadas corretamente de cada ator/atriz.....	22
Tabela 4- Relação de acertos entre as emoções representadas em cada vídeo e as frases empregadas.....	23
Tabela 5- Resultado do julgamento dos avaliadores de quão natural cada emoção foi representada nos vídeos.....	24
Figura 1 - Montagem da cena para filmagem.....	17
Figura 2 - Recorte de cenas de vídeos representando as expressões emocionais	21

A Criação de um *Dataset* Audiovisual Brasileiro de Expressões Emocionais

Definir o que é emoção não tem se mostrado uma tarefa simples e há décadas vários teóricos se empenham nessa missão. Muitas teorias diferentes sobre as emoções concordam no entendimento de que as emoções são um processo que envolve múltiplas variáveis e uma condição complexa e momentânea que surge em experiências de caráter afetivo, provocando alterações em várias áreas do funcionamento psicológico e fisiológico do indivíduo (Ekman, 2003; Miguel, 2015; Plutchik, 2003).

No trabalho de Cowen et al. (2019), os autores discorrem sobre as emoções como fenômenos que possibilitam que o indivíduo responda de forma adaptativa a ameaças e oportunidades evolutivamente significativas no ambiente, ou seja, que o preparam para uma ação. Como exemplo, ouvir o grito de um filho, ser ameaçado por um inimigo ou encontrar um parceiro sexual potencialmente disponível podem despertar emoções distintas. Nessa perspectiva, as emoções constituem ocorrências evolutivas com importantes funções na sobrevivência da espécie (Van Kleef, 2017).

As emoções permeiam todas as relações humanas, pois, como a vida em sociedade é considerada inerentemente emocional, são as próprias interações sociais as principais eliciadoras de emoções (Van Kleef & Fischer, 2016). Ainda nessa perspectiva interpessoal, Van Kleef (2017) considera que as pessoas influenciam continuamente umas às outras por meio de suas expressões emocionais (deliberadamente ou não) nas relações familiares, íntimas, no local de trabalho e no domínio político. Ekman (2011) defende também que as emoções estão intimamente envolvidas na qualidade de vida dos indivíduos e ocorrem em cada relacionamento do ser humano.

Diante disso, as emoções podem fornecer uma riqueza de informações não apenas para aqueles que as vivenciam (Frijda, 1986; Schwarz & Clore, 1983), mas também para aqueles que as reconhecem (Van Kleef, 2009). Os estados emocionais do falante são informados ao ouvinte e, cumprindo uma função essencial na comunicação humana, influenciam o comportamento do ouvinte (Levenson, 1999). Scarantino (2017) defende que as expressões emocionais comunicam quatro tipos de informações: (1) o sentimento atual

do indivíduo; (2) o que está acontecendo no presente contexto; (3) os cursos de ação desejados de outras pessoas que percebem a expressão; e (4) a intenção e os planos sobre o que a pessoa pode fazer. Desta forma, de acordo com Andrade et al. (2013), todas as pessoas capazes de compreender as suas emoções e as daqueles que vivem ao seu redor estão propícias a terem melhores interações sociais e qualidade de vida.

Seguindo essa concepção, de acordo com Busso et al. (2004), a comunicação interpessoal inclui não apenas a linguagem falada, mas também dicas não-verbais, como gestos com as mãos, expressões faciais e tom de voz, que por sua vez são usadas também para expressar emoções. Os seres humanos podem reconhecer esses padrões de sinais emocionais, mesmo que sejam exibidos sutilmente, processando simultaneamente as informações adquiridas visual e/ou auditivamente.

Há, portanto, diversas formas de expressar e reconhecer uma emoção. Miguel (2015) discorre sobre três possíveis formas de manifestação: (1) pela expressão vocal, sendo que diferentes estados emocionais implicam em alterações na frequência (altura ou tom), no volume e no ritmo (tempo) da vocalização; (2) pela gesticulação, que se trata de um componente importante da expressão emocional, porém dependente da cultura em que o indivíduo está inserido, já que são entendidos como gírias visuais, normalmente servindo para enfatizar o que está sendo comunicado; e (3) pela expressão facial, que é foco principal da maioria das pesquisas sobre expressão e reconhecimento de emoções, em razão de ser a região corporal humana com mais recursos musculares de expressão.

Para exemplificar como ocorre o reconhecimento de um episódio emocional por um observador, analisaremos a emoção "tristeza". Para Ekman (2011), tal como todas as outras emoções básicas (medo, raiva, nojo, surpresa e alegria), a tristeza também tem expressões faciais características reconhecidas universalmente: as sobrancelhas se juntam; as pálpebras superiores relaxam e as inferiores contraem-se, fazendo um movimento para baixo e na horizontal; as narinas se contraem; a raiz do nariz encorriha para baixo; a boca se fecha e contrai-se; e o queixo fica tenso e pode até franzir (Magalhães, 2007). A partir da identificação dessas características, a expressão da tristeza de um indivíduo pode sinalizar

ao outro que um objetivo ou um estado que desejava alcançar não foi atingido, ou ainda indicar a existência de algum problema que o envolva (Melo, 2005).

Conforme já afirmado, a importância do reconhecimento de ocorrências emocionais é notável em inúmeras atividades humanas. Como no exemplo do trabalho de enfermeiros estudado por Bachion et al. (2004), geralmente as pessoas enfrentam emoções como medo e ansiedade em um momento pré-operatório. A identificação destas emoções, o encorajamento da verbalização, o saber ouvir, fornecer informações adequadas, entre outras medidas, podem ajudar na assistência e contribuir para uma boa recuperação do paciente. Os autores consideram ser essencial que o enfermeiro esteja familiarizado com as reações emocionais que afetam os pacientes no pré-operatório, de modo que possa manejar adequadamente a situação e contribuir para uma boa recuperação.

A identificação de emoções também se faz presente no processo de psicoterapia, como descrito no estudo de Palhoço (2011). A relação entre terapeuta-cliente é central e imprescindível à qualidade do processo psicoterapêutico e, ainda segundo a autora, para que um terapeuta possa estabelecer uma relação empática com o cliente, é necessário que consiga perceber com facilidade as emoções que este expressa. O fato de o terapeuta dominar esta capacidade de reconhecimento de emoções permite-lhe promover *insights* potenciadores de mudanças terapêuticas.

De forma similar, as emoções desempenham um papel significativo na área comercial de atendimento ao cliente. Galanis et al. (2013) apresentam a detecção de traços emocionais em interações de *call center* e consideram que tal ação pode ser benéfica para a gestão da qualidade dos serviços prestados. De acordo com os autores, o reconhecimento da emoção revela o posicionamento de ambos os falantes: por um lado a satisfação ou frustração do cliente, e, por outro, a detecção de estresse por parte dos operadores, seja por uma possível redução ou falha no fornecimento do serviço solicitado. O estudo de Galanis et al. (2013) descreve uma abordagem computacional para classificar unidades de fala emocional que ocorrem em um conjunto de dados, empregando rótulos relacionados à emoção em unidades positivas ou negativas.

Dada a importância em diversos aspectos na vida dos seres humanos, nos últimos anos houve um aumento significativo de pesquisas sobre o tema de emoções também no âmbito computacional. Souza (2019) aponta que, com a evolução tecnológica, sistemas de software têm a capacidade de inferir e responder à emoção detectada no usuário e tomar decisões considerando esta informação, tais como aplicações médicas, jogos e monitoramento de situações cotidianas.

O aprofundamento do estudo das emoções depende de exemplares da expressão emocional em cada modalidade de estímulo envolvido em seu reconhecimento, como características do som e da imagem de uma pessoa tendo uma determinada emoção. Bancos com múltiplos exemplares de uma pessoa tendo uma determinada emoção gravada em material de áudio e/ou vídeo podem ser usados para treinar pessoas e mesmo computadores a reconhecer aquela emoção.

Uma das formas de contribuir para isso é a preparação de bases de dados (ou *dataset*) que armazenam um conjunto de material áudio e/ou visual de expressões emocionais humanas. Dada a relevância cultural de como uma emoção se expressa, encontramos vários bancos de dados criados em diferentes culturas e línguas. Dentre os principais exemplos encontrados, referenciamos a Base de Dados de Emoções de Berlin (EMO-DB; Burkhardt et al., 2005), SmartKom (Schiel et al., 2002), *Japanese Female Facial Expression* (JAFFE; Shan et al., 2017), *eINTERFACE Audio-Visual Emotion Database* (eINTERFACE; Martin et al., 2006), *CMU Multimodal Opinion Sentiment and Emotion Intensity* (CMU-MOSEI; Zadeh et al., 2016), *Interactive Emotional Dyadic Motion Capture Database* (IEMOCAP; Busso et al., 2008), *Vera am Mittag German Audio-visual Emotional Speech Database* (VAM; Grimm et al., 2008), *Natural Visible and Infrared facial Expression Database* (Wang et al., 2010), *The Danish Emotional Speech Database* (DES; Engberg et al., 1997), *Geneva Multimodal Emotion Portrayals* (GEMEP; Banziger et al., 2007), *Radboud Faces Database* (Langner et al., 2010), *FAU Aibo* (Batliner et al., 2008), *Indian Semi-Acted Facial Expression* (iSAFE; Singh & Benedict, 2019), *VERBO: Voice Emotion Recognition*

dataBase in Portuguese Language (Torres Neto et al., 2018) e *Ryerson Audio-Visual Database of Emotional Speech and Song* (RAVDESS; Livingstone & Russo, 2018).

Pode-se notar que todos esses bancos com exemplares de emoções são estrangeiros e, como já introduzido, há evidências científicas de diferenças culturais importantes no reconhecimento e expressão de emoções, trazendo limitações ao uso de *datasets* de outros países em estudos brasileiros. Os resultados da meta-análise realizada por de Elfenbein & Ambady (2002), por exemplo, indicam que as emoções podem ser compreendidas com mais precisão quando julgadas por membros do mesmo grupo nacional, étnico ou regional de quem a expressou. Esta vantagem dentro do grupo significa que a cultura tem um importante papel na formação de nossa comunicação emocional. A variabilidade cultural na precisão do reconhecimento de emoções também foi atribuída a diferenças no idioma, em que palavras usadas para traduzir conceitos emocionais podem transmitir significados diferentes entre culturas (Harré, 1989; Mesquita & Walker, 2003).

Em outra revisão, Douglas-Cowie et al. (2003) concluíram que a maior parte dos trabalhos sobre desenvolvimento de base de dados de emoções foram feitas em alemão. Apontam que, embora existam conjuntos de dados para espanhol, russo, hebraico, coreano e japonês, a quantidade ainda é limitada. Como resultado, avaliam que dificilmente as relações que a literatura descreve podem ser específicas a um único meio cultural relativamente homogêneo.

Em concordância com tal perspectiva e tendo em vista que até o momento não há uma base de dados audiovisual brasileira reconhecida de expressões emocionais, uma das demandas cruciais necessárias para alcançar grandes progressos no estudo nacional das emoções é o desenvolvimento de novos *datasets* brasileiros. Por esses motivos, o objetivo do presente trabalho foi produzir o *Dataset Audiovisual Brasileiro de Expressões Emocionais* (*Brazilian Audiovisual Dataset of EMotions* - BADEM), um *dataset* de brasileiros encenando expressões emocionais com material no formato audiovisual. Acredita-se que esse produto técnico viabilizará pesquisas em diversas áreas de conhecimento, como na psicologia, na computação, na educação, entre outros. Mais especificamente, tal produto

poderá ser usado para treinar pessoas e algoritmos computacionais a identificar tais ocorrências emocionais registradas em vídeos e/ou áudios.

Método

Para a criação do *Dataset* Audiovisual Brasileiro de Expressões Emocionais (BADEM), baseamo-nos majoritariamente em partes de metodologias utilizadas na criação de *datasets* similares já existentes que são referências nesse tipo de material (Feraru & Zbancioc, 2015; Rosa Jr, 2017; Schuller et al., 2007; Schuller et al., 2009; Tawari & Trivedi, 2010). A Base de Dados de Emoções de Berlin (EMO-DB; Burkhardt et al., 2005) e a *Danish Emotional Speech Database* (DES; Engberg et al., 1997) são as principais referências para o desenvolvimento de material de áudio. Além dessas, alguns aspectos do *Ryerson Audio-Visual Database of Emotional Speech and Song* (RAVDESS; Livingstone & Russo, 2018) e do *eINTERFACE'05 Audio-Visual Emotion Database* (eINTERFACE; Martin, 2006) foram adaptados para a criação do material de áudio e vídeo.

Com o intuito de definir parâmetros experimentais, primeiramente foi desenvolvido uma primeira versão da construção do *dataset*. Testes de cenas com atores expressando emoções ao declarar diferentes frases, somado à avaliação desse primeiro material gravado, foram essenciais para delinear a metodologia mais adequada para então desenvolver o *dataset* audiovisual proposto.

Assim como colocado pelos bancos de dados de referência, a escolha pela gravação de atores sob condições controladas se baseia na observação de que tal formato pode superar limitações técnicas de gravação de situações reais, entre elas a posição dos microfones e câmeras, o léxico, o conteúdo da fala e os fundos visuais e acústicos, além de oferecerem a possibilidade do registro de várias expressões para os mesmos indivíduos (Banziger et al., 2006).

Para o desenvolvimento do BADEM, doze atores gravaram cenas declarando doze diferentes frases em sete expressões emocionais, sendo as emoções: (1) raiva, (2) tristeza, (3) nojo, (4) medo, (5) felicidade, (6) surpresa e (7) expressão neutra. Tais emoções foram escolhidas por serem representadas nos bancos de dados de referência. Ao todo, foram

gravados 1008 vídeos, ou seja, 84 vídeos por ator/atriz. Esse material audiovisual foi analisado por psicólogos clínicos, passando por um teste de reconhecimento de cada emoção e de julgamento quanto ao grau de naturalidade da atuação de cada ator/atriz.

Todos os vídeos produzidos serão disponibilizados abertamente na internet no repositório científico osf.io, junto ao resultado da avaliação feita pelos avaliadores, evidenciando as taxas de identificação correta e o grau de naturalidade da emoção representada em cada vídeo. O *dataset* pode ser conferido através do seguinte endereço: https://osf.io/68g9m/?view_only=a76ca1e33689416c81eb87bc76313ab2. Qualquer pessoa que se comprometa a usar de forma ética e sem fins lucrativos, concordando previamente com o termo de compromisso de utilização de dados (Apêndice 3), poderá ter acesso gratuito aos vídeos.

É válido ressaltar que, em relação à aprovação ética para realização do presente estudo, este trabalho faz parte de um projeto maior intitulado “Pesquisa de processo-resultado em psicoterapia: um estudo computacional”, número CAAE 32025920.9.0000.5473, aprovado pelo Comitê de Ética em Pesquisa (CEP) - IFSP - Rua Pedro Vicente, 625 - Canindé- São Paulo- SP. Tel: (11) 3775-4665, E-mail: cep_ifsp@ifsp.edu.br; CAAE 32025920.9.0000.5473.

Atores

Na gravação do projeto, contamos com doze atores voluntários (6 mulheres e 6 homens). Os atores foram chamados por meio de convites diretos realizados pela mestrandia. Todos os atores são brasileiros e falantes nativos de português.

As divergências culturais que existem entre as regiões do Brasil foram consideradas relevantes quanto às possíveis diferenças na expressão emocional de cada um. Por isso, optamos por convidar atores de vários estados do país para realizarem as gravações remotamente. Nenhum ator/artiz foi pago(a) por sua participação e todos assinaram um termo de responsabilidade (Apêndice 1).

Os atores tinham entre 21 e 31 anos de idade, todos com curso ou formação em teatro, e mínimo de dois anos de experiência. Ao todo, oito atores residiam e realizaram a

gravação presencialmente em Londrina-PR. Desses, cinco eram londrinenses, um nascido em Maringá-PR, um em Ibiporã-PR e um ator natural de Alta Floresta D'oeste-RO. Os outros quatro atores realizaram as gravações dos vídeos de forma remota, sendo que dois atores nasceram, residiam e gravaram em Brasília-DF, e os outros dois nasceram, residiam e gravaram em São Luis-MA.

Avaliadores

Oito psicólogos clínicos participaram como avaliadores na pesquisa, realizando testes de reconhecimento das emoções a partir dos vídeos produzidos e, além disso, julgaram quão natural consideravam a atuação do(a) ator/atriz. Na seção "Avaliação do material audiovisual" está descrito de que forma realizaram as avaliações. Essa categoria de profissionais foi escolhida por considerarmos o contato que apresentam com a temática de expressão emocional no seu campo de atuação. Os avaliadores tinham de 3 e 20 anos de experiência profissional e, entre as áreas de atuação citadas, está a clínica e a docência. Além de assinarem o termo de responsabilidade (Apêndice 2), os avaliadores preencheram um formulário de dados gerais antes de iniciarem a participação na pesquisa.

Frases

Seguindo o padrão usado pelo EMO-DB (Burkhardt et al., 2005), foram utilizadas frases com conteúdos neutros, que não possuíssem nenhum significado que influenciasse no reconhecimento da emoção. Foram elaboradas doze frases apresentadas na Tabela 1.

Tabela 1

Frases empregadas nas gravações

Número	Frases
1	Aquela música me lembra de muitas coisas.
2	Hoje à noite eu posso conversar com ela sobre isso.
3	Você foi capaz de fazer isso sozinho.
4	Ele vai conseguir terminar isso a tempo?
5	Vocês estão fazendo isso há muitos anos.

Número	Frases
6	Nós vamos viajar no final de semana com aquele grupo.
7	Se eu pudesse prever o futuro, com certeza eu diria isso.
8	Toda vez que nos encontramos é desse jeito.
9	Aos poucos eu fui entendendo cada parte do plano.
10	Nunca cheguei a conhecer alguém assim antes.
11	Há vários animais naquela casa.
12	Como pode eles serem tão parecidos?

Ressaltamos que, em princípio, não nos parece relevante termos utilizado diferentes frases, uma vez que nosso material foca em criar exemplares de áudio e vídeo de cada emoção e não do conteúdo representado. Deste modo, foi importante garantir que o conteúdo não tivesse potencial para influenciar o reconhecimento na direção de uma das emoções e nem que fosse demasiadamente complexo e/ou sem sentido para a nossa cultura, o que poderia distrair o avaliador ao tentar entender o conteúdo e não identificar a emoção.

Procedimentos

Gravação dos Vídeos

Para realizar a filmagem, diretrizes de gravação do trabalho do RAVDESS foram adotadas como referência, no qual os atores foram informados de que teriam o tempo que fosse necessário para que eles entrassem no estado emocional desejado e que, uma vez alcançado, eles o sinalizariam sua prontidão para iniciar a gravação (Livingstone & Russo, 2018). Enfatizou-se que os atores deveriam fornecer expressões genuínas de emoção e que poderiam repetir um determinado teste até se sentirem confortáveis com sua produção. Nenhuma instrução foi dada sobre como uma emoção deveria ser expressada.

Para as gravações do BADEM, desenvolvemos um Protocolo de gravação dos vídeos de expressões emocionais (Apêndice 5), o qual foi prontamente seguido nas

gravações presenciais e remotas. Duas semanas antes de iniciarem as gravações, os atores receberam um Protocolo Geral (Apêndice 6) com informações sobre o modo de gravar, bem como seu respectivo Roteiro Individual (Apêndice 7). Cada ator foi instruído a encenar as doze frases, cada uma em sete emoções distintas, conforme descrito pelo roteiro individual, contendo variações na ordem das frases e emoções. A sequência das emoções e frases foi feita randomicamente para cada um deles.

Para realizar a filmagem presencialmente, a pesquisadora executante do projeto conduziu sozinha as gravações de oito atores individualmente (ver descrição em "Instrumentos e Cenário"). Houve uma única sessão de gravação com cada ator/atriz, com duração de aproximadamente 2 horas. Foram gravados sete vídeos com cada ator/atriz, sendo um vídeo por emoção, encenando todas as suas respectivas frases. De forma remota, os outros quatro atores realizaram suas gravações seguindo o mesmo protocolo de gravação e, em seguida, fizeram o *upload* dos seus vídeos para uma pasta privada do *Google Drive* na conta de e-mail da pesquisa.

Após o término das gravações presenciais e o recebimento dos vídeos gravados remotamente, a mestranda realizou a edição de todo material de áudio e imagem, segmentando cada vídeo em doze recortes de cenas, garantindo que cada vídeo contivesse um(a) ator/atriz dizendo uma frase em uma emoção, com aproximadamente quatro segundos de duração. Todo o material audiovisual produzido foi salvo na mesma pasta do *Google Drive* da pesquisa.

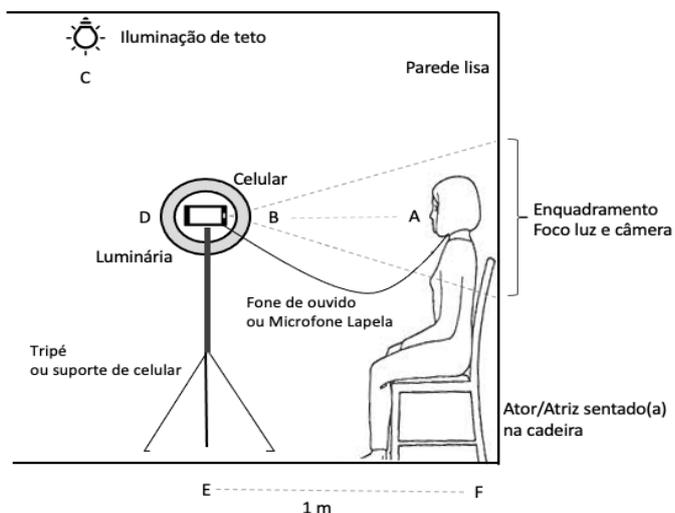
Instrumentos e Cenário

Como referência de instrumentos e montagem do cenário para a filmagem dos vídeos, nos baseamos no trabalho de desenvolvimento do eNTERFACE (Martin et al., 2006). Tal banco de dados foi gravado usando uma câmera de vídeo digital, enquanto a gravação da fala foi realizada através de um microfone separado. O fundo consistiu em um plano monocromático e a iluminação constante foi feita através de um refletor posicionado de frente para o usuário (Martin et al., 2006).

Foi desenvolvido o Protocolo de gravação dos vídeos de expressões emocionais (Apêndice 6) com o intuito de padronizar o cenário e instrumentos utilizados pela mestrande e pelos atores de forma remota. Para garantir a qualidade da imagem e do som dos vídeos, instrui-se o uso de aparelho celular com câmera de vídeo (*webcam*) com capacidade de gravação em alta definição para a filmagem e um microfone lapela fixado no(a) ator/atriz conectado ao celular utilizado para a gravação do vídeo. O celular deveria estar acoplado a um Tripé ou a um suporte de mesa, a 1 metro de distância do(a) ator/atriz, este(a) sentado(a) em uma cadeira com uma parede de cor clara e lisa ao fundo. Além disso, foi orientado que a filmagem ocorresse em uma sala com boa iluminação e acústica. De frente para o(a) ator/atriz, foi posicionado um ponto de luz de uma luminária circular led, como representado na Figura 1.

Figura 1

Cenário de gravação



Para executar a filmagem presencialmente dos oito atores, a mestrande utilizou um celular modelo Iphone Xs acoplado a um Tripé. As gravações ocorreram em uma sala de atendimento psicológico na cidade de Londrina-PR. Os atores foram gravados sentados em

uma cadeira voltada para frente da câmera, com uma parede branca ao fundo. Uma luminária circular led foi posicionada frontalmente ao(à) ator/atriz.

Os quatro atores que realizaram as gravações em suas respectivas cidades seguiram o mesmo protocolo de gravação, porém com algumas adaptações. O modelo do aparelho celular de cada um variou e apenas uma atriz utilizou o microfone lapela, os outros três recorreram ao próprio fone de ouvido com microfone do celular. Todos usaram uma luminária e suporte de mesa para estabilizar o celular e buscaram ambientes com boa iluminação e paredes claras.

Avaliação do Material Audiovisual

Os 1008 vídeos produzidos foram adicionados de forma privada no *Google Drive* na conta de e-mail criada para a presente pesquisa. Por se tratar de uma grande quantidade de material audiovisual, a avaliação dos vídeos foi realizada da seguinte forma: foram criados oito *Google Formulários*, um para cada avaliador (Apêndice 4). Cada formulário continha 126 vídeos a serem avaliados uma única vez e individualmente por meio de duas perguntas (X representando o número do vídeo): (1) Qual emoção você julga ser expressa no vídeo X? (2) Quão natural você julga ter ficado a emoção encenada no vídeo X pelo(a) ator/atriz? Após selecionarem qual emoção estava sendo encenada nos vídeos, os avaliadores julgaram conforme os seguintes níveis de naturalidade: “nada natural”, “pouco natural”, “moderadamente natural”, “muito natural” ou “completamente natural”. Cada uma das perguntas continha o *link* de acesso automático ao referido vídeo (X) no *Google Drive* da pesquisa. A ordem dos vídeos em cada formulário foi definida randomicamente.

Tal qual no procedimento de avaliação do DES, os colaboradores assistiram a cada vídeo quantas vezes julgaram necessárias para avaliar o material audiovisual e não tiveram nenhum treinamento prévio. Os formulários foram enviados por e-mail aos avaliadores, que tiveram um mês para enviar as respostas à pesquisadora.

As respostas dos avaliadores foram organizadas em um banco de dados e aplicado o teste estatístico do qui-quadrado de Pearson para comparar os acertos das representações em relação a cada ator/atriz, emoção, frase e avaliador(a). Quando os

resultados dos testes se mostraram significativos, foi realizada uma comparação post-hoc por pares (*pairwise*) com correção de Holm para múltiplos testes de hipóteses.

Resultados

Os resultados obtidos com a avaliação dos 1008 vídeos serão apresentados a seguir. Considerando que todas as emoções foram representadas por cada um dos doze atores e que oito avaliadores fizeram a identificação de uma parcela do total de vídeos (126 vídeos para cada avaliador), cada emoção teve um total de 144 vídeos avaliados.

A soma total de vídeos que foram identificados corretamente pelos avaliadores foi 845, ou seja, 83,82% dos 1008 vídeos gravados. Para ser possível analisar o desempenho dos avaliadores e a identificação correta de cada emoção em função de cada avaliador, encontra-se descrita na Tabela 2 a relação de acertos dos avaliadores por cada emoção. Nota-se que a quantidade de vídeos por emoção variou conforme o avaliador. O teste de qui-quadrado das variáveis avaliador e acertos mostrou diferenças significativas ($X^2 = 74,8$; $p < 0,001$). Com a comparação *pairwise* realizada a *posteriori* foi possível identificar discrepância entre os avaliadores “7” e “2”, que tiveram mais acertos (respectivamente 98,41% e 96,82% de acertos), em relação ao “3”, que teve menos acertos (70,63%). Já o teste de qui-quadrado considerando avaliador e emoção não indicou diferenças significativas ($X^2 = 48,3$; $p = 0,232$). Não houve diferenças significativas comparando as variáveis avaliador, acertos e ator ($X^2 = 91,4$; $p = 0,126$). As tabelas com os resultados estatísticos mais detalhados podem ser conferidas no Apêndice 9.

Tabela 2

Resultado da avaliação dos vídeos pelos avaliadores conforme cada emoção.

Avaliador	Acertos/Quantidade total de vídeos avaliados por emoção							Total de acertos
	Nojo	Neutro	Felicidade	Raiva	Tristeza	Surpresa	Medo	
1	20/22	15/15	13/14	22/22	18/19	13/15	13/19	90,00%
2	18/18	20/20	17/17	24/24	17/17	14/17	12/13	96,82%

Avaliador	Acertos/Quantidade total de vídeos avaliados por emoção							Total de acertos
	Nojo	Neutro	Felicidade	Raiva	Tristeza	Surpresa	Medo	
3	12/14	11/15	17/18	17/17	12/22	12/23	8/17	70,63%
4	24/25	23/24	10/14	9/11	14/22	12/15	8/15	79,36%
5	13/14	15/15	19/25	12/18	14/19	15/20	9/15	76,98%
6	12/16	13/17	14/18	13/18	17/18	12/15	10/24	72,22%
7	25/26	16/17	23/23	18/18	11/11	15/15	16/16	98,41%
8	9/9	20/21	15/15	13/16	13/16	22/24	16/25	85,71%
Total de acertos por emoção	92%	92%	88%	88%	80%	79%	63%	83,82%

Comentado [LMI]: Passei TOTAL para % nas tabelas

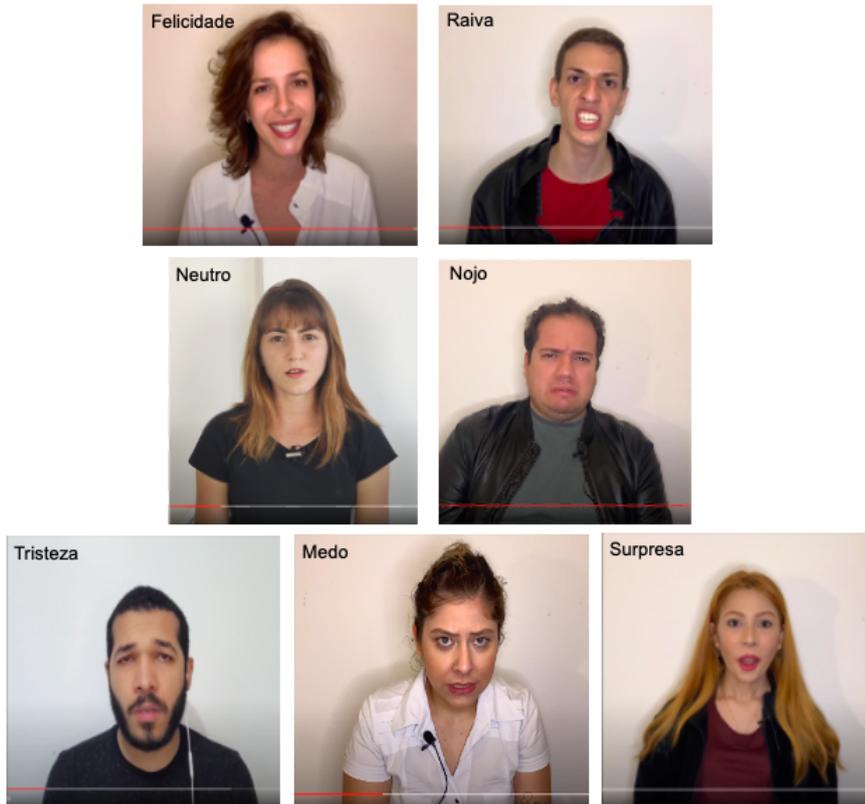
Na Tabela 2 também é possível visualizar o número de vídeos que foram avaliados corretamente por emoção, por avaliador e pela combinação dessas duas variáveis, isto é, o quanto os avaliadores conseguiram identificar cada emoção representada. O teste de qui-quadrado comparando as variáveis emoção, avaliador e acertos não expressou diferenças estatisticamente significativas ($X^2 = 43,5$; $p = 0,409$).

Realizando a avaliação das emoções de forma independente, o teste de qui-quadrado mostrou diferenças significativas na comparação entre as variáveis emoção e acertos ($X^2 = 66$; $p < 0,001$). Os testes *post hoc* por pares revelaram que a expressão “neutra” e a emoção “nojo” foram identificadas corretamente um maior número de vezes, ambas com 92,36% de acerto na avaliação. A emoção “medo” mostrou-se com número de acertos inferior (63,88%) em comparação com todas as outras emoções, exceto “surpresa”.

Para ilustrar as expressões emocionais, na Figura 2 há recortes de cenas de vídeos em que os atores estavam representando as emoções “neutro”, “nojo”, “medo”, “felicidade”, “raiva”, “tristeza” e “surpresa”.

Figura 2

Representações de cada emoção



Para avaliar a performance dos atores, foi considerado o total de acertos que os avaliadores atingiram ao identificar a emoção encenada. Na comparação das variáveis ator e acertos foram apresentadas diferenças significativas ($X^2 = 27,7$; $p = 0,003$) em seus desempenhos. A comparação *post hoc pairwise* aponta para uma tendência de significância entre os atores “8” e “3” ($p = 0,072$), respectivamente 92,8% e 72,6%, na quantidade de acertos em seus vídeos. Em contrapartida, comparando as variáveis ator, acertos e emoção não há indicação de diferenças significativas ($X^2 = 26,1$; $p > 0,999$) conforme o teste de qui-quadrado. Apresenta-se na Tabela 3 a síntese da relação de vídeos identificados corretamente por emoção conforme cada ator/atriz.

Tabela 3

Síntese dos vídeos com as emoções identificadas corretamente de cada ator/atriz

Ator/Atriz	Acertos/Quantidade total de vídeos por emoção							Total de acertos por ator/atriz
	Felicidade	Nojo	Medo	Neutro	Raiva	Surpresa	Tristeza	
1	10/12	11/12	10/12	12/12	12/12	9/12	11/12	89,2%
2	12/12	12/12	8/12	11/12	12/12	10/12	11/12	90,4%
3	11/12	10/12	10/12	11/12	12/12	12/12	12/12	92,8%
4	12/12	12/12	6/12	12/12	10/12	10/12	11/12	86,9%
5	10/12	12/12	5/12	11/12	12/12	8/12	10/12	80,9%
6	10/12	12/12	10/12	12/12	8/12	12/12	9/12	86,9%
7	11/12	10/12	9/12	12/12	9/12	10/12	11/12	85,7%
8	9/12	12/12	1/12	10/12	11/12	7/12	11/12	72,6%
9	12/12	12/12	11/12	12/12	12/12	8/12	5/12	85,7%
10	10/12	10/12	9/12	7/12	12/12	11/12	10/12	82,1%
11	11/12	11/12	6/12	12/12	12/12	6/12	8/12	78,5%
12	10/12	9/12	7/12	11/12	6/12	12/12	7/12	73,8%
Total de acertos por emoção	88%	92%	63%	92%	88%	79%	80%	83%

Encontra-se sintetizada na Tabela 4 a relação das frases empregadas nos vídeos que tiveram a emoção corretamente identificada. Os testes de qui-quadrado sugerem não haver diferenças estatisticamente significativas entre as frases em termos do total de acertos das emoções representadas ($X^2 = 10,2$; $p = 0,512$); de acertos considerando cada avaliador ($X^2 = 89,3$; $p = 0,159$); de acertos ao se considerar cada emoção ($X^2 = 14,5$; $p > 0,999$); nem de acertos considerando cada ator ($X^2 = 15,7$; $p > 0,999$). A relação de acertos considerando cada frase por emoção está reportada na Tabela 6 (Apêndice 8).

Tabela 4

Relação de acertos entre as emoções representadas em cada vídeo e as frases empregadas.

Número	Frase	Acertos
1	Aquela música me lembra de muitas coisas.	86,90%
2	Hoje à noite eu posso conversar com ela sobre isso.	79,76%
3	Você foi capaz de fazer isso sozinho.	82,14%
4	Ele vai conseguir terminar isso a tempo?	79,76%

5	Vocês estão fazendo isso há muitos anos.	76,19%
6	Nós vamos viajar no final de semana com aquele grupo.	85,71%
7	Se eu pudesse prever o futuro, com certeza eu diria isso.	88,09%
8	Toda vez que nos encontramos é desse jeito.	83,33%
9	Aos poucos eu fui entendendo cada parte do plano.	88,09%
10	Nunca cheguei a conhecer alguém assim antes.	90,47%
11	Há vários animais naquela casa.	80,95%
12	Como pode eles serem tão parecidos?	84,52%

Por fim, outro aspecto avaliado nos vídeos se refere a quão natural os avaliadores julgaram as encenações de cada vídeo, conforme apresentado na Tabela 5. Considerando a totalidade de 1008 vídeos, as emoções encenadas foram avaliadas como "pouco" ou "nada" naturais em 9,83% das vezes. Mais de 90% dos vídeos foram avaliados com as emoções encenadas como "moderada" a "completamente" naturais.

Tabela 5

Resultado do julgamento dos avaliadores sobre o quão natural cada emoção foi representada nos vídeos

Avaliador	Nível de naturalidade				
	Nada natural	Pouco natural	Moderadamente natural	Muito natural	Completamente natural
1	0	5	21	51	49
2	0	2	53	71	0
3	1	6	21	48	50
4	7	19	46	26	28
5	0	22	59	44	1
6	1	11	54	46	14
7	0	5	26	63	32
8	0	20	38	62	6
Total	0,89%	8,92%	31,53%	40,87%	17,85%

Discussão

O presente documento apresenta a criação de um *dataset* de exemplares audiovisuais de emoções com base em alguns trabalhos de referência na área (Burkhardt et al., 2005; Engberg et al., 1997; Livingstone & Russo 2018; Martin 2006). De acordo com os resultados obtidos, é possível notar a discrepância do reconhecimento a depender das emoções, atores e avaliadores.

A identificação correta das emoções nos estudos de referência foi acima de 80% na EMO-DB, baseado apenas em áudio no idioma alemão, enquanto no DES foi de 67,3%, também com base exclusiva em áudio, no idioma dinamarquês. No RAVDESS, os avaliadores identificaram corretamente as emoções pretendidas em 80% para áudio e vídeo, já no eINTERFACE a validação foi de 83%, ambos os datasets em inglês.

Os materiais de vídeo e de áudio produzidos pelas bases de dados de referência foram divulgados com as porcentagens de acertos reportadas e, aparentemente, sem exclusão de nenhum material. A partir disso, será disponibilizado em um repositório científico todo o material audiovisual produzido e divulgado as taxas de acertos por parte dos avaliadores, com o grau de reconhecimento e naturalidade da emoção em cada vídeo.

Na construção do BADEM foi identificada uma diferença significativa entre os avaliadores em termos de desempenho nos acertos das emoções representadas nos vídeos. Dos oito psicólogos que participaram da avaliação, três se destacaram, sendo os avaliadores "7" e "2" com as maiores pontuações e o avaliador "3" com uma quantidade inferior de acertos.

De acordo com os resultados da avaliação de cada emoção representada nos vídeos, a emoção "medo" mostrou-se com número de acertos inferior em comparação com todas as outras emoções. A partir da análise das respostas dos formulários, verificou-se que tal diferença parece ter ocorrido pelos avaliadores terem confundido "medo" com a expressão "neutra" ou "tristeza". Supõe-se que tal confusão pode ser decorrente das formas

de manifestação dessas emoções apresentarem aspectos semelhantes, como alguma expressão facial ou tom de voz encenado pelo(a) ator/atriz.

Vale ainda notar que, quanto aos atores, ficou perceptível a diferença entre seus desempenhos, considerando as taxas de acertos pelos avaliadores no reconhecimento das emoções encenadas. Cada ator/atriz teve pelo menos uma das emoções corretamente identificadas em todos os doze vídeos filmados. Destaca-se a atriz “3” que, de acordo com a avaliação, teve as emoções identificadas corretamente de forma mais expressiva em relação ao ator “8”, o qual obteve menor taxa de acertos das emoções encenadas.

Quanto aos resultados apresentados pelas doze frases empregadas nas gravações, os testes estatísticos aplicados indicam que não há influência do conteúdo das frases sobre os acertos.

Em relação ao nível de naturalidade dos vídeos, os avaliadores julgaram que as emoções encenadas foram de “moderada” a “completamente” naturais em mais de 90% dos vídeos. Tal resultado evidencia que a maioria dos vídeos desenvolvidos foram ao encontro de um dos objetivos do projeto: produzir exemplares de expressões o mais naturalmente representadas possíveis. No entanto, pontua-se ainda uma limitação encontrada quanto a qual seria o cálculo estatístico mais adequado a ser aplicado para analisar o nível de naturalidade dos vídeos.

Outro aspecto da metodologia discutido é o número de representantes de atores, emoções e avaliadores que foram usados para alcançar uma quantidade suficiente de vídeos para compor o *dataset*. A intenção foi desenvolver um conjunto de materiais audiovisuais suficiente para viabilizar procedimentos e pesquisas que exigissem exemplares e, mais especificamente, que tal produto técnico pudesse ser usado para treinar pessoas e algoritmos computacionais no reconhecimento de emoções. Nas bases de dados de referência foram desenvolvidos de 800 a 7000 exemplares entre materiais de vídeo, áudio e/ou audiovisual, variando também o número de atores, avaliadores e quantidade de avaliações. Percebe-se que não há um padrão definido na elaboração de um banco de dados de emoções, o que leva a alguns questionamentos, como: (1) há um número mínimo

de quantidade de vídeos para compor um *dataset*? (2) deverá haver outros critérios de seleção para divulgar os vídeos produzidos? (3) quantos atores são necessários e qual nível de formação precisam ter?

Com o intuito de obter uma melhora mínima e continuada no *dataset*, considera-se a possibilidade de otimização do material produzido como, por exemplo, refazer as avaliações dos avaliadores que apresentaram desempenho inferior. Outra estratégia seria também substituir os avaliadores por novos que consigam atingir, segundo essa nova avaliação realizada, desempenhos superiores semelhantes, e então, definir um perfil de avaliadores mais homogêneos entre si. Em seguida, seria interessante refazer todos os outros cálculos estatísticos dos atores e emoções para verificar se houve ou não diferença nas taxas de acertos atingidos. É válido considerar que as variáveis empregadas para a construção do *dataset* se influenciam mutuamente, ou seja, o desempenho dos avaliadores pode ter sido influenciado pelo desempenho dos atores. Sugere-se que, ao mudar os resultados dos acertos quando realizados por novos avaliadores, talvez as diferenças de performances dos atores e emoções também sejam outras. A partir desse ponto, seria possível acrescentar e/ou substituir os atores que apresentarem desempenhos significativamente inferiores, segundo essa nova avaliação. Esse processo pode ser repetido até que se alcance uma quantidade de vídeos, por diversos atores com performances equivalentes e por emoção, que seja suficiente para servir de amostra de expressões emocionais.

De modo geral, acredita-se que os resultados aqui apresentados revelam um nível de acurácia nas avaliações das expressões emocionais, o que sugere que a metodologia proposta parece ter produzido um material de qualidade na direção dos objetivos propostos.

Conclusão

O presente trabalho propôs descrever a construção do *Dataset* Audiovisual Brasileiro de Expressões Emocionais (BADEM). A metodologia empregada produziu um material audiovisual de expressões emocionais de qualidade, considerando que os vídeos

alcançaram altas taxas de acertos nas emoções representadas em comparação com os *datasets* de referência. Além disso, em sua maioria, o BADEM atingiu um nível de naturalidade considerado de “moderada” a “completamente” naturais.

O *dataset* desenvolvido acrescenta um relevante material para viabilizar o desenvolvimento de diversas áreas de conhecimento envolvendo o tema das emoções e, mais especificamente, tal produto poderá ser usado para treinar pessoas e algoritmos computacionais a identificar tais ocorrências emocionais registradas em vídeos e/ou áudios.

Por fim, considerando a possibilidade de atualização continuada, o *dataset* poderá sofrer melhorias com o tempo, como implementações no número de atores, frases e emoções que podem ampliar a quantidade de exemplares emocionais. Ademais, há ainda a possibilidade de substituição de atores e avaliadores que não atingiram performances satisfatórias por outros melhor qualificados, visando garantir um alto nível de reconhecimento e naturalidade das expressões emocionais gravadas no material audiovisual.

Referências

- Andrade, N. C., Abreu, N. S., Duran, V. R., Veloso, T. J., & Moreira, N. A. (2013). Reconhecimento de expressões faciais de emoções: padronização de imagens do teste de conhecimento emocional. *Psico*, 44(3), 382-390
- Bachion, M. M., Magalhães, F. G. S., Munari, D. B., Almeida, S. P. , & Lima, M. L. (2004). Identificação do “medo” no período pré-operatório de cirurgia cardíaca. *Acta Paul Enferm.*, 17(3), 298-304
- Bänziger, T., & Scherer, K. R. (2007). Using actor portrayals to systematically study multimodal emotion expression: The GEMEP corpus. In *International conference on affective computing and intelligent interaction*. Springer, Berlin, pp. 476-487
- Batliner, A., Steidl, S., & Nöth, E. (2008). Releasing a thoroughly annotated and processed spontaneous emotional database: the FAU Aibo Emotion Corpus
- Britto, Ilma A. Goulart de Souza, & Elias, Paula Virgínia Oliveira. (2009). Análise comportamental das emoções. *Psicologia para América Latina*, (16) Recuperado em

25 de outubro de 2020, de

http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1870-

[350X2009000100004&lng=pt&tling=pt](http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1870-350X2009000100004&lng=pt&tling=pt)

- Burkhardt, F., Paeschke, A., Rolfes, M., Sendlmeier, W. F., & Weiss, B. (2005). A database of German emotional speech. In *Ninth European Conference on Speech Communication and Technology*
- Busso, C., Deng, Z., Yildirim, S., Bulut, M., Lee, C. M., Kazemzadeh, A., Narayanan, S. (2004). Analysis of emotion recognition using facial expressions, speech and multimodal information. Proceedings of the 6th International Conference on Multimodal Interfaces - ICMI '04. doi:10.1145/1027933.1027968
- Busso, C., Bulut, M., Lee, C. C., Kazemzadeh, A., Mower, E., Kim, S., & Narayanan, S. S. (2008). IEMOCAP: Interactive emotional dyadic motion capture database. *Language resources and evaluation*, 42(4), 335-359
- Cheshin, A., Amit, A., & Van Kleef, G. A. (2018). The interpersonal effects of emotion intensity in customer service: Perceived appropriateness and authenticity of attendants' emotional displays shape customer trust and satisfaction. *Organizational Behavior and Human Decision Processes*, 144, 97-111
- Cowen, A., Sauter, D., Tracy, J. L., & Keltner, D. (2019). Mapping the passions: Toward a high-dimensional taxonomy of emotional experience and expression. *Psychological Science in the Public Interest*, 20(1), 69-90
- Douglas-Cowie, E., Campbell, N., Cowie, R., & Roach, P. (2003). Emotional speech: Towards a new generation of databases. *Speech Communication*, 40(1-2), 33-60. doi:10.1016/s0167-6393(02)00070-5
- Ekman, P. (1993). Facial expression and emotion. *American Psychologist*, 48(4), 384-392. <https://doi.org/10.1037/0003-066X.48.4.384>
- Ekman, P. (2003). *Emotions revealed: Recognizing faces and feelings to improve communication and emotional life*. Times Books/Henry Holt and Co
- Ekman, P. (2011). *A linguagem das emoções*. São Paulo: Lua de Papel

- Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin*, 128(2), 203–235.
doi:10.1037/0033-2909.128.2.203
- Engberg, I. S., Hansen, A. V., Andersen, O., & Dalsgaard, P. (1997). Design, recording and verification of a Danish emotional speech database. In *Fifth European conference on speech communication and technology*
- Feraru, S. M., & Zbancioc, M. D. (2015). Comparative analysis between SROL - Romanian database and Emo - German database. 2015 International Symposium on Signals, Circuits and Systems (ISSCS). doi:10.1109/isscs.2015.7204015
- Frijda, N. H. (1986). *The emotions*. Cambridge University Press
- Grimm, M., Kroschel, K., & Narayanan, S. (2008). The Vera am Mittag German audio-visual emotional speech database. 2008 IEEE International Conference on Multimedia and Expo. doi:10.1109/icme.2008.4607572
- Harré, R. (1989). Language and the science of psychology. *Journal of Social Behavior and Personality*, 4(3), 165
- Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D.H.J., Hawk, S.T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition & Emotion*, 24(8), 1377–1388. DOI: 10.1080/0269993090348507
- Levenson, R. W. (1999). The intrapersonal functions of emotion. *Cognition & Emotion*, 13(5), 481-504
- Livingstone, S. R., & Russo, F. A. (2018). The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS): A dynamic, multimodal set of facial and vocal expressions in North American English. *PloS one*, 13(5), e0196391
- Magalhães, A. (2007). *A Psicologia das emoções: O fascínio do rosto humano*. Porto: Edições Universidade Fernando Pessoa
- Martin, O., Kotsia, I., Macq, B., & Pitas, I. (2006). The eNTERFACE’05 Audio-Visual Emotion Database. 22nd International Conference on Data Engineering Workshops (ICDEW'06). doi:10.1109/icdew.2006.145

- Melo, A. (2005). Emoções no período escolar: estratégias parentais face à expressão emocional e sintomas de internalização e externalização da criança. Dissertação de mestrado, Universidade do Minho, Porto
- Mesquita, B., & Walker, R. (2003). Cultural differences in emotions: A context for interpreting emotional experiences. *Behaviour research and therapy*, 41(7), 777-793
- Miguel, Fabiano Koich. (2015). Psicologia das emoções: uma proposta integrativa para compreender a expressão emocional. *Psico-USF*, 20(1), 153-162.
<https://doi.org/10.1590/1413-82712015200114>
- Palhoco, Ana Rita de Mendonça Santos (2011). Estudo da empatia e da percepção de emoções em psicoterapeutas e estudantes de psicologia. Tese de mestrado, Psicologia (Secção de Psicologia Clínica e da Saúde - Núcleo de Psicoterapia Cognitiva-Comportamental e Integrativa), Universidade de Lisboa, Faculdade de Psicologia
- Plutchik, R. (2003). *Emotions and life: Perspectives from psychology, biology, and evolution*. American Psychological Association
- Rosa Jr, J. (2017). Reconhecimento automático de emoções através da Voz. Monografia UFSC
- Scarantino, A. (2017). How to do things with emotional expressions: The theory of affective pragmatics. *Psychological Inquiry*, 28(2-3), 165-185
- Schiel, F., Steininger, S., & Türk, U. (2002). The SmartKom multimodal corpus at BAS. In LREC
- Schuller, B., Seppi, D., Batliner, A., Maier, A., & Steidl, S. (2007). Towards more reality in the recognition of emotional speech. 2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07.
doi:10.1109/icassp.2007.367226
- Schuller, B., Vlasenko, B., Eyben, F., Rigoll, G., & Wendemuth, A. (2009). Acoustic emotion recognition: A benchmark comparison of performances. 2009 IEEE Workshop on Automatic Speech Recognition & Understanding. doi:10.1109/asru.2009.5372886

- Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: informative and directive functions of affective states. *Journal of personality and social psychology*, 45(3), 513
- Shan, K., Guo, J., You, W., Lu, D., & Bie, R. (2017). Automatic facial expression recognition based on a deep convolutional-neural-network structure. In 2017 IEEE 15th International Conference on Software Engineering Research, Management and Applications (SERA) (pp. 123-128)
- Singh, S., & Benedict, S. (2019). Indian Semi-Acted Facial Expression (iSAFE) dataset for human emotions recognition. in *International symposium on signal processing and intelligent recognition systems* (pp. 150-162). Springer, Singapore
- Souza, Isaque Elcio de (2019). Classificação de sinais fisiológicos para inferência do estado emocional de usuários. Universidade Federal de São Carlos. Dissertação Programa de Pós-Graduação em Ciência da Computação - PPGCC
- Tawari, A., & Trivedi, M. M. (2010). Speech emotion analysis: Exploring the role of context. *IEEE Transactions on Multimedia*, 12(6), 502–509. doi:10.1109/tmm.2010.2058095
- Torres Neto, J. R., Filho, G. P., Mano, L. Y. & Ueyama, J. (2018). VERBO: Voice Emotion Recognition dataBase in Portuguese Language. *Journal of Computer Science*, 14(11), 1420-1430. <https://doi.org/10.3844/jcssp.2018.1420.1430>
- Van Kleef, G. A. (2009). How emotions regulate social life: The emotions as social information (EASI) model. *Current directions in psychological science*, 18(3), 184-188
- Van Kleef, G. (2017). The social effects of emotions are functionally equivalent across expressive modalities. *Psychological Inquiry*, 28(2-3), 211-216
- Van Kleef, G. A., & Fischer, A. H. (2016). Emotional collectives: How groups shape emotions and emotions shape groups. *Cognition and Emotion*, 30(1), 3-19
- Zadeh, A., Zellers R., Eli Pincus, and LouisPhilippe Morency (2016). Multimodal sentiment intensity analysis in videos: Facial gestures and verbal messages. *IEEE Intelligent Systems* 31(6):82–88

Wang, S., Liu, Z., Lv, S., Lv, Y., Wu, G., Peng, P., ... Wang, X. (2010). A natural visible and infrared facial expression database for expression recognition and emotion inference. *IEEE Transactions on Multimedia*, 12(7), 682–691. doi:10.1109/tmm.2010.2060716

Apêndices

Apêndice 1

Termo de compromisso e responsabilidade – ator / atriz

Eu, _____, portador do RG _____ e CPF _____ concordo em ajudar a produzir um material técnico que será usado primariamente em pesquisas, mas eventualmente também para outros fins, conforme especificado adiante no presente termo. Declaro que tenho conhecimento dos termos abaixo, com os quais estou de acordo e comprometo-me a respeitar:

- Participar da gravação de cenas para a produção de um banco de dados audiovisuais de emoções que faz parte do projeto “Pesquisa de processo-resultado em psicoterapia: um estudo computacional”.
- Entendo que não sou participante de um projeto de pesquisa, mas que estou oferecendo minhas habilidades em atuação para ajudar a produzir um material técnico que será usado primariamente em pesquisas, mas eventualmente também para outros fins.
- Participarei com gravação de cenas nas quais usarei diferentes frases neutras para encenar nove expressões emocionais, sendo elas: (1) raiva, (2) nojo, (3) medo, (4) felicidade, (5) tristeza, (6) surpresa, além da expressão neutra.
- Após a gravação do material, autorizo que os vídeos produzidos sejam avaliados por cinco psicólogos, colaboradores do projeto, com o objetivo de discriminarem quais são as expressões emocionais presentes em cada vídeo e formalizarem um teste de percepção e concordância das mesmas.
- Cedo minha imagem e voz para a criação do banco de dados e, renuncio todo e qualquer direito autoral patrimonial decorrente deste.
- Tenho conhecimento de que o material audiovisual produzido para o banco de dados será disponibilizado abertamente na internet, e para qualquer pessoa que se comprometa a usar de forma ética e sem fins lucrativos.

- O banco de dados não será usado com fins lucrativos pelos pesquisadores e responsáveis pelo projeto.
- Os benefícios envolvidos na minha contribuição incluem, entre outros possíveis, a convicção íntima e / ou social de que é importante contribuir para o desenvolvimento de projetos de relevância para o desenvolvimento de conhecimento científico, bem como de outros possíveis usos éticos do material produzido.
- Exceto pelo material audiovisual produzido, meus dados pessoais serão preservados durante o projeto e após o término do mesmo.
- Minha participação no projeto é voluntária, portanto, não terei nenhum custo ou compensação financeira.

_____, ____ de ____ de ____
(cidade) (dia) (mês) (ano)

assinatura do participante da pesquisa

assinatura do pesquisador

Apêndice 2***Termo de compromisso e responsabilidade – avaliadores do banco de dados audiovisuais***

Eu, _____, portador do RG _____ e CPF _____ realizando a atividade de avaliador do banco de dados audiovisuais para o projeto “Pesquisa de processo-resultado em psicoterapia: um estudo computacional”, comprometo-me a respeitar as seguintes condições: colaborar com a avaliação da percepção e concordância das expressões emocionais apresentadas em cada vídeo da base de dados, não divulgar o material sem autorização dos pesquisadores, não efetuar nenhum tipo de cópia ou duplicação do material, não realizar qualquer tipo de violação do material original e não fazer uso do material, enquanto tiver acesso ao mesmo, para qualquer outro fim que não os solicitados pelo pesquisador responsável.

_____, _____ de _____ de _____
(cidade) (dia) (mês) (ano)

assinatura do participante da pesquisa

assinatura do pesquisador

Apêndice 3

Termo de compromisso de utilização de dados – Dataset Audiovisual Brasileiro De Expressões Emocionais (BADEM)

O *Dataset* Audiovisual Brasileiro de Expressões Emocionais (BADEM) foi desenvolvido como produto técnico do Mestrado em Análise do Comportamento Aplicada no Paradigma - Centro de Ciência e Tecnologia do Comportamento, pela mestranda Luisa Medina e sob orientação do Prof. Dr. Felipe Corchs. Confira a dissertação em: <https://mestrado.paradigmaac.org/alunos-egressos/>.

BADEM comporta 1008 vídeos com áudio, de cenas de doze atores brasileiros declarando doze diferentes frases em sete expressões emocionais, sendo estas: (1) raiva, (2) tristeza, (3) nojo, (4) medo, (5) felicidade, (6) expressão neutra e (7) surpresa. O material audiovisual foi analisado por oito psicólogos clínicos, passando por um teste de reconhecimento de cada emoção e de julgamento quanto ao grau de naturalidade da atuação de cada ator/atriz.

Para ter acesso ao *dataset*, preencha os campos abaixo com seus dados:

Nome:

CPF:

Data de nascimento:

Endereço:

Profissão:

E-Mail:

Telefone:

Motivo de acesso a base de dados e finalidade do seu uso:

Comprometo-me a respeitar as seguintes condições:

- Usar o material audiovisual de forma ética e sem fins lucrativos.
- Não divulgar o material, sendo a acessibilidade individual.
- Não repassar os dados em sua íntegra, ou parte dele a outra pessoa.
- Não efetuar nenhum tipo de cópia ou duplicação do material.
- Manter a integridade e confidencialidade dos dados, bem como com a privacidade de seus conteúdos.
- Atribuir o devido crédito em publicações de qualquer natureza e quaisquer outras divulgações que incluam os dados obtidos deste *dataset*, ou suas derivações, utilizando a seguinte referência:

Medina, L. (2022) *A criação de um Dataset Audiovisual Brasileiro de Expressões Emocionais (BADEM)*. Dissertação de mestrado. Paradigma centro de ciências e tecnologia do comportamento. São Paulo, SP.

Ao preencher o presente termo você concorda com as condições acima descritas e declara serem verdadeiras as informações pessoais.

Data:

Assinatura

O presente documento deve ser preenchido, assinado e enviado para e-mail datasetemoco.es.badem@gmail.com. Apenas após o aceite da pesquisadora você terá acesso ao *dataset*. O e-mail cadastrado no termo deverá ser o mesmo que terá acesso ao *dataset*.

Term of commitment for the use of data - Brazilian Audiovisual Dataset of Emotional Expressions (BADEM) (versão em inglês)

The Brazilian Audiovisual Dataset of Emotional Expressions (BADEM) was developed as a technical product of the Master's Degree in Applied Behavior Analysis in Paradigm - Center for Behavioral Science and Technology, by master student Luisa Medina and under the guidance of Prof. doctor Felipe Corchs. Check out the dissertation at: <https://mestrado.paradigmaac.org/alunos-egressos/>.

BADEM contains 1008 videos with audio, of scenes of twelve Brazilian actors declaring twelve different phrases in seven emotional expressions: (1) anger, (2) sadness, (3) disgust, (4) fear, (5) happiness, (6) neutral expression, and (7) surprise. The audiovisual material was analyzed by eight clinical psychologists, passing through a test of recognition of each emotion and of judgment as to the degree of naturalness of the performance of each actor/actress.

To access the dataset, fill in the fields below with your data:

Name:

CPF:

Birth date:

Address:

Profession:

Email:

Telephone:

Reason for accessing the database and purpose of its use:

I undertake to respect the following conditions:

- Use audiovisual material in an ethical and non-profit manner.
- Not to divulge the material, being the individual accessibility.
- Do not pass on the data in its entirety, or part of it to another person.
- Do not make any type of copy or duplication of the material.
- Maintain the integrity and confidentiality of the data, as well as the privacy of its contents.
- Attribute due credit in publications of any nature and any other disclosures that include data obtained from this dataset, or its derivations, using the following reference:

Medina, L. (2022) The creation of a Brazilian Audiovisual Dataset of Emotional Expressions (BADEM). Masters dissertation. Paradigm Center for Behavioral Sciences and Technology. Sao Paulo-SP.

By completing this term, you agree with the conditions described above and declare that the personal information is true.

Date:

Signature

This document must be completed, signed and sent to the email datasetemocoos.badem@gmail.com. Only after acceptance by the researcher will you have access to the dataset. The e-mail registered in the term must be the same that will have access to the dataset.

Apêndice 4

Questionário de avaliação das emoções

O formulário de avaliação das emoções conta com duas seções, sendo a primeira de coleta de dados dos avaliadores, a segunda com questões de 126 vídeos. No presente documento está exemplificado o formulário 1, com o vídeo de número 1.



Avaliação vídeos expressões emocionais

A seguir está a avaliação final dos vídeos gravados na fase definitiva do projeto "Criação de um Dataset Audiovisual Brasileiro de Expressões Emocionais".

Para desenvolver o dataset de expressões emocionais contamos com 12 atores que encenaram 12 frases em 7 emoções distintas. Ao todo foram gravados 1008 vídeos.

Por se tratar de uma grande quantidade de material audiovisual, os vídeos foram divididos randomicamente em 8 formulários, portanto em cada formulário terão 126 vídeos a serem avaliados.

Para iniciar a avaliação, preencha os campos abaixo com suas informações pessoais.

@gmail.com [Alternar conta](#) 

***Obrigatório**

E-mail *

Seu e-mail _____

Avaliação vídeos expressões emocionais

@gmail.com [Alternar conta](#)



*Obrigatório

Instruções

Para ter acesso aos vídeos clique nos links anexos a cada uma das perguntas.
Assista a cada vídeo com bastante atenção e quantas vezes quiser. Em seguida responda às duas perguntas em sequência, primeiro identificando a emoção e depois julgando quão natural aquela emoção foi encenada pelo(a) ator/atriz no vídeo.
Responda as perguntas na ordem estipulada, não pule questões e nem volte para alterar respostas anteriores.
Você poderá assistir os vídeos quantas vezes for necessário. Caso deseje interromper a avaliação é possível retomar em outro momento do ponto em que parou.

Primeiro formulário

Vídeos 1-126

Qual emoção você julga ser expressa no vídeo 1?

1 ponto

<https://drive.google.com/file/d/1Z98YvC4ZqETELsqo4QI3n0td97mPg8WP/view?usp=sharing> *

- Raiva
- Felicidade
- Medo
- Surpresa
- Nojo
- Tristeza
- Neutra

Quão natural você julga ter ficado a emoção encenada no vídeo 1 pelo(a) ator/atriz?

<https://drive.google.com/file/d/1Z98YvC4ZqETELsqo4Ql3n0td97mPg8WP/view?usp=sharing> *

- Nada natural
- Pouco natural
- Moderadamente natural
- Muito natural
- Completamente natural

Apêndice 5

Protocolo de gravação dos vídeos de expressões emocionais

Objetivo: Estabelecer padrões e procedimentos para aquisição de material em vídeo.

1) Lista de equipamentos necessários

- Aparelho celular com câmera de vídeo (*webcam*) com capacidade de gravação em alta definição (mínimo Full HD 1080p).
- Microfone Lapela (com cabo de no mínimo 1,5m) ou fone de ouvido (com microfone).
- Tripé (ou outro suporte de celular, de piso ou mesa) para fixar e posicionar o celular.
- Fonte de luz branca (ex: módulo led circular)
- Trena para medição (caso necessário).

2) Montagem dos equipamentos e cenário

A montagem da gravação deve ser realizada seguindo os passos sugeridos a seguir:

- Procurar um local com boa iluminação e reservado, com o mínimo de barulho possível.
- Colocar uma cadeira encostada em uma parede de fundo claro e liso, de preferência branco. O ator ficará sentado nesta cadeira (F) durante a filmagem.
- Alinhar a cadeira de frente com a iluminação de teto (C) do ambiente.
- Fixar o tripé (E) a uma distância de 1 metro da parede, de frente para a cadeira (F).
- O celular (B) deve estar posicionado (encaixado no suporte do tripé de mesa ou piso) na altura dos olhos do ator (A) sentado na cadeira.
- Uma fonte de luz branca (D) deverá ser posicionada na mesma altura do celular (B), de maneira que não gere sombra no rosto do ator (A).
- A câmera traseira do celular deve estar direcionada frontalmente para a cadeira. O celular deve ficar posicionado na horizontal.
- O enquadramento da imagem deve capturar do tórax para cima do ator sentado na cadeira, incluindo todo o ombro (Figura 2).
- No caso do microfone lapela: conecte o cabo do microfone no celular e prenda a lapela na gola do ator.
No caso do fone de ouvido: conecte o cabo no celular e ajuste os fones nos ouvidos.
No caso de fone de ouvido sem fio: conecte o fone via bluetooth e ajuste os fones nos ouvidos.
- A câmera deve permanecer fixa a partir do posicionamento estabelecido e durante toda a gravação. Não utilizar zoom digital.
- Confira na Figura 1 o cenário de gravação.

3) Ligar equipamentos e iniciar/finalizar a gravação

- Acenda a luz de teto e a fonte de luz branca frontal.
- Selecione a opção vídeo no celular e inicie as gravações.
- Cada ator deverá seguir seu próprio roteiro de cena, fornecido pela pesquisadora.
- O ator deverá encenar todas as 12 frases por emoção, gravando vídeos separados por emoção.
- Entre uma frase e outra, o ator deverá fazer uma pequena pausa (de aproximadamente 5 segundos) para sinalizar o fim de uma frase e o início da próxima, mantendo o olhar para a câmera.
- Se necessário, o ator poderá ter em mãos seu roteiro individual impresso. Poderá consultar o roteiro entre uma frase e outra, desde que faça uma pequena pausa antes e depois olhando para a câmera. Evitar ao máximo tocar na folha enquanto encena as frases, para não gerar nenhum tipo de barulho.
- Ao gravar todas as frases de uma emoção, fará a mesma coisa encenando as frases com a próxima emoção. Ao todo são 7 vídeos.
- Considerando que há 7 emoções e 12 frases, ao todo, cada ator gravará 84 cenas.
- Se ocorrer qualquer intercorrência ou falha na gravação de algum vídeo, o ator poderá descartar o mesmo e regravar por completo.
- Caso o ator precise encenar novamente alguma cena, é imprescindível que façam uma pausa (de aproximadamente 5 segundos) para sinalizar o início e fim da cena daquela frase.

4) Armazenamento dos vídeos

- Os arquivos de vídeo devem ser armazenados em uma pasta individual do Google Drive que será criada e compartilhada pela pesquisadora.
- Cada vídeo deverá ser salvo com o seguinte código: A_E
A= número do ator (fornecido pela pesquisadora)
E= emoção representada no vídeo

Exemplo: O ator 1 irá salvar seu vídeo da emoção Raiva.

Código correspondente do vídeo = 1_R

Apêndice 6

Protocolo geral para atores

Cada ator terá de memorizar 12 frases, para encenar em 7 emoções distintas.

Portanto, ao final, cada ator irá gravar 84 cenas. A seguir estão as frases e as emoções:

Nº	Frase	Código	Emoção
1	Aquela música me lembra de muitas coisas.	N	Neutro
2	Hoje à noite eu posso conversar com ela sobre isso.	F	Felicidade
3	Você foi capaz de fazer isso sozinho.	T	Tristeza
4	Ele vai conseguir terminar isso a tempo?	R	Raiva
5	Vocês estão fazendo isso há muitos anos.	S	Surpresa
6	Nós vamos viajar no final de semana com aquele grupo.	J	Nojo
7	Se eu pudesse prever o futuro, com certeza eu diria isso.	M	Medo
8	Toda vez que nos encontramos é desse jeito.		
9	Aos poucos eu fui entendendo cada parte do plano		
10	Nunca cheguei a conhecer alguém assim antes		
11	Há vários animais naquela casa		
12	Como pode eles serem tão parecidos?		

Orientação para a gravação:

- Não há regras quanto à vestimenta para a gravação.
- O uso de maquiagem, óculos e outros acessórios não serão permitidos, pois pode prejudicar a captura da imagem facial.

- A câmara filmará apenas a parte superior do corpo, do tórax para cima.
- No ensaio das frases recomenda-se que os atores tentem lembrar de alguma situação que tenha vivido em que despertou tal emoção solicitada. O objetivo é o ator conseguir encenar a emoção da forma mais **natural** possível.
- Os vídeos deverão ser gravados por ordem de emoção, portanto o ator encenará uma frase de cada vez em uma emoção. Após gravar todas as frases naquela emoção, passará para a próxima emoção, encenando as mesmas frases novamente.
- A ordem das frases e emoções será definida pela pesquisadora e irão variar por ator. Por este motivo, cada ator deverá se atentar e ensaiar seu respectivo roteiro individual.
- Atenção para o número e código correspondente de cada frase e emoção, esta informação é importante para depois salvar corretamente o vídeo. Consulte detalhes no Protocolo de gravação.
- Durante a gravação, o ator poderá ter em mãos o seu roteiro individual impresso para consultar as frases se necessário.
- Evitar ao máximo mexer na folha do roteiro ou gerar qualquer outro tipo de barulho durante as gravações das frases. Ruídos externos prejudicam a qualidade do som do vídeo.
- Para facilitar as filmagens dos vídeos, recomendamos ter o auxílio de uma outra pessoa na hora da gravação para manusear o celular.
- O tempo médio para ensaiar e fazer as gravações dos vídeos pode demorar por volta de 1h e meia, a depender da velocidade do ator.
- Recomendamos que ensaiem na frente de um espelho antes de realizarem as gravações.

Apêndice 7**Roteiro individual (exemplo ator 1)**

As suas cenas deverão ser gravadas seguindo necessariamente esta ordem por emoção:

Emoção: Neutro (N)

Frases:

- 1 Aquela música me lembra de muitas coisas.
- 2 Hoje à noite eu posso conversar com ela sobre isso.
- 3 Você foi capaz de fazer isso sozinho.
- 4 Ele vai conseguir terminar isso a tempo?
- 5 Vocês estão fazendo isso há muitos anos.
- 6 Nós vamos viajar no final de semana com aquele grupo.
- 7 Se eu pudesse prever o futuro, com certeza eu diria isso.
- 8 Toda vez que nos encontramos é desse jeito.
- 9 Aos poucos eu fui entendendo cada parte do plano.
- 10 Nunca cheguei a conhecer alguém assim antes.
- 11 Há vários animais naquela casa.
- 12 Como pode eles serem tão parecidos?

Emoção: Felicidade (F)

Frases:

- 3 Você foi capaz de fazer isso sozinho.
- 4 Ele vai conseguir terminar isso a tempo?
- 12 Como pode eles serem tão parecidos?
- 1 Aquela música me lembra de muitas coisas.
- 6 Nós vamos viajar no final de semana com aquele grupo.
- 7 Se eu pudesse prever o futuro, com certeza eu diria isso.
- 8 Toda vez que nos encontramos é desse jeito.
- 9 Aos poucos eu fui entendendo cada parte do plano.
- 11 Há vários animais naquela casa.
- 2 Hoje à noite eu posso conversar com ela sobre isso.
- 5 Vocês estão fazendo isso há muitos anos.
- 10 Nunca cheguei a conhecer alguém assim antes.

Emoção: Tristeza (T)

Frases:

- 5 Vocês estão fazendo isso há muitos anos.
- 6 Nós vamos viajar no final de semana com aquele grupo.
- 11 Há vários animais naquela casa.
- 2 Hoje à noite eu posso conversar com ela sobre isso.
- 7 Se eu pudesse prever o futuro, com certeza eu diria isso.
- 8 Toda vez que nos encontramos é desse jeito.
- 9 Aos poucos eu fui entendendo cada parte do plano.
- 10 Nunca cheguei a conhecer alguém assim antes.
- 12 Como pode eles serem tão parecidos?
- 3 Você foi capaz de fazer isso sozinho.

- 1 Aquela música me lembra de muitas coisas.
- 4 Ele vai conseguir terminar isso a tempo?

Emoção: Raiva (R)

Frases:

- 7 Se eu pudesse prever o futuro, com certeza eu diria isso.
- 8 Toda vez que nos encontramos é desse jeito.
- 10 Nunca cheguei a conhecer alguém assim antes.
- 3 Você foi capaz de fazer isso sozinho.
- 9 Aos poucos eu fui entendendo cada parte do plano.
- 5 Vocês estão fazendo isso há muitos anos.
- 11 Há vários animais naquela casa.
- 12 Como pode eles serem tão parecidos?
- 6 Nós vamos viajar no final de semana com aquele grupo.
- 4 Ele vai conseguir terminar isso a tempo?
- 2 Hoje à noite eu posso conversar com ela sobre isso.
- 1 Aquela música me lembra de muitas coisas.

Emoção: Surpresa (S)

Frases:

- 9 Aos poucos eu fui entendendo cada parte do plano.
- 10 Nunca cheguei a conhecer alguém assim antes.
- 8 Toda vez que nos encontramos é desse jeito.
- 5 Vocês estão fazendo isso há muitos anos.
- 7 Se eu pudesse prever o futuro, com certeza eu diria isso.
- 11 Há vários animais naquela casa.
- 12 Como pode eles serem tão parecidos?
- 6 Nós vamos viajar no final de semana com aquele grupo.
- 1 Aquela música me lembra de muitas coisas.
- 2 Hoje à noite eu posso conversar com ela sobre isso.
- 4 Ele vai conseguir terminar isso a tempo?
- 3 Você foi capaz de fazer isso sozinho.

Emoção: Nojo (J)

Frases:

- 11 Há vários animais naquela casa.
- 12 Como pode eles serem tão parecidos?
- 9 Aos poucos eu fui entendendo cada parte do plano.
- 6 Nós vamos viajar no final de semana com aquele grupo.
- 10 Nunca cheguei a conhecer alguém assim antes.
- 4 Ele vai conseguir terminar isso a tempo?
- 1 Aquela música me lembra de muitas coisas.
- 2 Hoje à noite eu posso conversar com ela sobre isso.
- 3 Você foi capaz de fazer isso sozinho.
- 5 Vocês estão fazendo isso há muitos anos.
- 8 Toda vez que nos encontramos é desse jeito.
- 7 Se eu pudesse prever o futuro, com certeza eu diria isso.

Emoção: Medo (M)

Frases:

2 Hoje à noite eu posso conversar com ela sobre isso.

1 Aquela música me lembra de muitas coisas.

7 Se eu pudesse prever o futuro, com certeza eu diria isso.

8 Toda vez que nos encontramos é desse jeito.

4 Ele vai conseguir terminar isso a tempo?

12 Como pode eles serem tão parecidos?

3 Você foi capaz de fazer isso sozinho.

5 Vocês estão fazendo isso há muitos anos.

10 Nunca cheguei a conhecer alguém assim antes.

6 Nós vamos viajar no final de semana com aquele grupo.

9 Aos poucos eu fui entendendo cada parte do plano.

11 Há vários animais naquela casa.

Apêndice 8**Relação de acertos de cada frase e as emoções****Tabela 6***Relação de acertos de cada frase e as emoções*

Frase	Emoção						
	Felicidade	Nojo	Medo	Neutro	Raiva	Surpresa	Tristeza
Aquela música me lembra de muitas coisas.	11	11	11	10	9	10	11
Hoje à noite eu posso conversar com ela sobre isso.	12	10	7	10	11	10	7
Você foi capaz de fazer isso sozinho.	10	12	5	11	10	11	10
Ele vai conseguir terminar isso a tempo?	9	12	9	11	9	11	6
Vocês estão fazendo isso há muitos anos.	8	11	6	11	10	8	10
Nós vamos viajar no final de semana com aquele grupo.	12	11	7	12	12	8	10
Se eu pudesse prever o futuro, com certeza eu diria isso.	12	9	9	11	12	10	11
Toda vez que nos encontramos é desse jeito.	11	12	6	12	12	6	11
Aos poucos eu fui entendendo cada parte do plano.	10	12	10	11	10	10	11
Nunca cheguei a conhecer alguém assim antes.	12	11	9	11	11	12	10
Há vários animais naquela casa.	12	11	6	12	11	8	8
Como pode eles serem tão parecidos?	9	11	7	11	11	11	11

Variáveis: Emoção e Acertos ($X^2=66$; $p=0,0000000000027549$)

	1	2	3	4	5	6
2	1,000	-	-	-	-	-
3	1,000	1,000	-	-	-	-
4	1,000	1,000	1,000	-	-	-
5	0,076	0,076	0,644	0,644	-	-
6	0,060	0,060	0,566	0,566	1,000	-
7	0,000	0,000	0,000	0,000	0,042	0,060

Variáveis: Avaliador e Acertos ($X^2=74,8$; $p=0,0000000000001563$)

	1	2	3	4	5	6	7
2	0,776	-	-	-	-	-	-
3	0,003	0,000	-	-	-	-	-
4	0,265	0,001	1,000	-	-	-	-
5	0,097	0,000	1,000	1,000	-	-	-
6	0,007	0,000	1,000	1,000	1,000	-	-
7	0,186	1,000	0,000	0,000	0,000	0,000	-
8	1,000	0,063	0,097	1,000	1,000	0,186	0,009

Variáveis: Naturalidade e Acertos ($X^2=21,5$; $p=0,0002546165652291$)

	1	2	3	4
2	1,000	-	-	-
3	1,000	0,105	-	-
4	1,000	0,000	0,305	-
5	1,000	0,005	0,842	1,000