

Personagens Virtuais Expressivas

Renato Teixeira

Ana Paula Cláudio

Maria Beatriz Carmo

LabMAg, Faculdade de Ciências, Universidade de Lisboa

1749-016 Lisboa, Portugal

fc37057@alunos.fc.ul.pt, apclaudio@fc.ul.pt, mbcarmo@fc.ul.pt

Augusta Gaspar

CIS- Centro de Investigação e Intervenção Social, ISCTE Instituto Universitário de Lisboa,

1649-026 Lisboa, Portugal

augusta.gaspar@iscte.pt

Resumo

A Ansiedade Social afeta um número significativo de pessoas, limitando-as na vida pessoal e social, e apresentando um nível de comorbilidade elevado com a depressão. Neste artigo descrevemos um projeto que aplica Realidade Virtual, interativa, para a terapia de exposição da Ansiedade Social, recorrendo a Personagens Virtuais exibindo combinações de expressões faciais e corporais. A aplicação descrita neste artigo atualiza e melhora significativamente uma versão anterior, sobretudo no que respeita à qualidade gráfica dos personagens virtuais; é executada em computadores vulgares, sem equipamentos dispendiosos, sendo facilmente utilizada em contextos de aconselhamento e investigação.

A possibilidade de controlar de forma minuciosa as expressões faciais permite o seu uso em contextos de investigação do comportamento não-verbal. Embora se tenha recorrido apenas a modelos gratuitos ou de muito baixo custo, adotaram-se estratégias que permitiram obter a qualidade adequada, facto que conseguimos validar através de testes com um conjunto significativo de observadores.

Keywords

Personagens Virtuais; Fobia Social; Expressões Faciais; Terapia de Exposição; Comunicação não-verbal.

1. INTRODUÇÃO

A fobia social ou Transtorno de Ansiedade Social (TAS) é uma condição caracterizada por intensa ansiedade quando o indivíduo é confrontado com uma situação de exposição pública ou mesmo a antecipação de uma tal situação [APA00]. Esta condição limita extraordinariamente a vida pessoal e social, apresentando-se com uma comorbilidade alta com a depressão [Stein00]. As pessoas que sofrem de TAS temem julgamentos negativos, e acreditam ser julgadas deste modo; a sua perceção de sinais emocionais é semelhante à de outras pessoas quando se tratam de sinais neutros ou característicos de afetos positivos. Mas, são detetores significativamente mais rápidos e eficazes em relação a sinais que podem veicular conteúdo ameaçador ou negativo [Douilliez12] evidenciando assim a sua hipervigilância para a ameaça social.

As terapias usadas para tratar TAS incluem geralmente medicamentos (principalmente anti-depressivos), terapia cognitivo-comportamental, psicoterapia e técnicas de relaxamento. De entre estas, a terapia cognitivo-comportamental é a que mostrou os resultados mais eficientes e duradouros, especialmente quando é aplicada na forma de terapia de exposição [Beidel07], que consiste em expor o paciente à situação que lhe causa a fobia.

A Realidade Virtual (RV) tem vindo a ser usada na terapia de exposição desde o início dos anos 90, designando-se por Realidade Virtual na Terapia de Exposição (RVTE), tendo alguns estudos concluído que produz resultados semelhantes aos da exposição tradicional [Klinger04; Herbelin05]. A RVTE permite um controlo preciso do processo de habituação (e extinção) de medo a um objeto fóbico e, portanto, oferece muitas vantagens adicionais sobre a terapia de exposição clássica, que envolve imagens e contacto posterior com situações da vida real: a) permite a parametrização de cenários e interações para atender às necessidades e progresso de diferentes pacientes, e às de cada paciente durante todo o período de tratamento; b) permite uma otimização da preparação do paciente antes de enfrentar a realidade, evitando assim o risco de uma exposição prematura a uma situação real; c) reduz o risco de retrocessos e reações excessivas e proporciona um ambiente mais estável e progressivo, garantindo um resultado previsível e consolidado; d) e finalmente, protege a privacidade dos pacientes. As principais desvantagens típicas da RVTE são o custo financeiro de aquisição de equipamentos virtuais imersivos (*Head-Mounted Displays*, CAVE) e os efeitos secundários reportados por alguns utilizadores (*cybersickness*) [LaViola00].

Este artigo descreve uma abordagem ao uso da RV no tratamento da Ansiedade Social, em particular, em situações de medo de falar em público na frente de um júri numa situação de avaliação. O paciente enfrenta um júri composto por um a três Personagens Virtuais que exibem expressões faciais e corporais controláveis em tempo real pelo terapeuta. Deste modo, o terapeuta controla o nível de *stress* induzido ao paciente durante a sessão de terapia. O controlo é obtido através de uma interface que (entre outras funções) desencadeia nos personagens um conjunto de movimentos faciais e posturas do tronco que permitem uma vasta gama de composições que podem transmitir conteúdo emocional neutro, positivo ou negativo, bem como simular atenção ou falta de interesse. Antes de uma sessão, o terapeuta configura o cenário: escolhe o aspeto da sala em que encontra o júri, as personagens e a respetiva posição na mesa, as suas roupas e o penteado.

Além de se destinar ao tipo de terapia descrita, a aplicação desenvolvida é uma ferramenta útil na investigação do comportamento não-verbal: o efeito da comunicação através do comportamento facial e de outros elementos ou atitudes corporais não é totalmente conhecido, apesar de ser comumente apresentado como tal e é objeto atualmente de intenso debate e diversos estudos [Gaspar12; Gaspar14].

A aplicação permite visualizar individualmente o rosto de cada personagem e compor diferentes constelações de unidades de comportamento facial e corporal, controlando o efeito da mudança de uma única ação individualizada no rosto, enquanto outras ações são mantidas constantes. Avaliar o impacto destas combinações sobre os observadores pode clarificar o papel comunicativo das ações unitárias ou compostas, e auxiliar na definição de parâmetros objetivos para o aconselhamento em tratamentos de ansiedade causada por exposição social, bem como para o avanço da investigação em comportamento não-verbal.

A imagem do cenário deve ser projetada numa tela ou parede, de modo que permita simular o tamanho real das personagens, criando assim um efeito imersivo.

O trabalho aqui descrito é uma evolução de trabalho anterior da mesma equipa. Uma primeira versão da aplicação recria um auditório de personagens virtuais com comportamentos controláveis mas com pouca expressão facial [Cláudio13]; uma segunda versão implementa já uma simulação de um júri com um conjunto de personagens virtuais com expressões faciais e corporais controláveis pelo terapeuta em tempo-real [Cláudio14]. O desenvolvimento de todas as versões tem sido sempre acompanhado de perto por uma psicóloga, que tem tido um papel fundamental na identificação das expressões dos personagens virtuais que são potencialmente indutores de retorno positivo, negativo ou neutro sobre os observadores.

Este artigo debruça-se sobre uma etapa em que se procedeu a melhoramentos na interface do terapeuta, se

investiu na utilização de modelos virtuais de melhor qualidade gráfica, simulando idades variadas e visuais distintos, e em que se efetuou um estudo de validação preliminar destes modelos e das suas expressões faciais, confirmando os pressupostos tidos em conta pela psicóloga da equipa na elaboração do seu menu de ações faciais.

A aplicação recorre a ferramentas de software gratuitas ou a licenças temporárias de produtos comerciais e a modelos virtuais gratuitos ou de baixo custo. Pode ser executada em computadores vulgares, e não recorre a equipamentos de RV imersivos, razões pelas quais pode ser facilmente utilizada no aconselhamento e na investigação.

Este artigo está organizado da seguinte forma: a próxima secção apresenta alguns dos trabalhos relacionados mais relevantes; na secção 3 descrevemos a nossa aplicação; a secção 4 é dedicada à validação do comportamento expressivo da face dos personagens virtuais e finalmente, a secção 5 é dedicada às conclusões e trabalho futuro.

2. TRABALHO RELACIONADO

A primeira aplicação de Realidade Virtual (RV) para tratar o medo de falar em público foi descrita em [North98]. Esta aplicação incluía um cenário de um auditório com até 100 personagens e durante uma sessão de terapia, o terapeuta podia variar o número de personagens e as suas atitudes, utilizando sequências de vídeo pré-gravadas. O paciente utilizava um HMD, ouvindo o eco de sua própria voz. Slater et al. criaram uma sala de seminários virtual com 8 personagens exibindo comportamentos autónomos aleatórios, como balançar a cabeça e piscar os olhos [Slater99]. O estudo inicial realizado com 10 alunos com diferentes níveis de dificuldade de falar em público foi alargado a um conjunto mais amplo de pessoas, incluindo indivíduos fóbicos e não-fóbicos [Pertaub01; Pertaub02; Slater06].

James et al. propuseram um cenário duplo: uma carruagem de metro na qual os personagens expressam comportamentos neutros, o que se configura um ambiente não exigente do ponto de vista de interação social e uma situação com nível de exigência diametralmente oposto, um bar em que os personagens pareciam muito desinteressados [James03]. Os comportamentos dos personagens incluíam movimentação do olhar e frases pré-gravadas.

Klinger et al. conduziram um estudo em que se analisou as mudanças no medo de falar em público em 36 participantes ao longo de 12 sessões [Klinger04]. Para recriar os personagens virtuais usaram imagens de pessoas reais em situações quotidianas. Os participantes foram divididos em dois grupos: um tratado com terapia cognitiva comportamental tradicional e outro com RVTE, tendo-se verificado uma maior redução da ansiedade social no grupo tratado com RVTE. Herbelin publicou um estudo de validação com 200 pacientes, confirmando que a sua plataforma de RV preenchia os requisitos da exposição terapêutica para a fobia social [Herbelin05].

Além disso, provou que é possível melhorar a avaliação clínica com ferramentas de monitorização integradas, como o seguimento do olhar (*eye-tracking*).

Todas as abordagens referidas recorrem a equipamentos do tipo HMD; no estudo descrito por Pertaub et al. metade dos pacientes experimentaram um dos ambientes virtuais usando um HMD enquanto o restante grupo usou um computador de secretária [Pertaub02]. Herbelin e Grillon, além de recorrerem a HMD e a um ecrã de computador simples também usaram uma grande superfície de retro-projeção [Herbelin05; Grillon09].

Haworth e outros autores implementaram cenários virtuais para serem visualizados simultaneamente por pacientes e pelo terapeuta em ecrãs de computador, possivelmente em diferentes locais físicos e comunicação via internet [Haworth12]. Os cenários destinam-se a pacientes com acrofobia (medo das alturas) ou aracnofobia (medo de aranhas). A Kinect¹ é usada para controlar os movimentos do corpo do paciente. Os ainda poucos resultados deste estudo parecem provar que este tipo de solução de baixo custo é eficaz para estas situações fóbicas particulares.

3. VIRTUAL SPECTATORS

O trabalho descrito neste artigo é relativo à aplicação Virtual Spectators que tem vindo a ser desenvolvida pela nossa equipa que inclui uma psicóloga. A situação fóbica escolhida é o medo de falar em público, recorrendo-se à Realidade Virtual no uso da terapia de exposição para o tratamento desta situação fóbica.

No Virtual Spectators, o cenário virtual é, durante uma sessão de terapia, o palco de uma simulação controlada pelo terapeuta e observada pelo paciente- indivíduo que apresenta elevados níveis de ansiedade quando confrontado com uma situação em que tem de falar perante outras pessoas. O paciente, enquanto efetua o seu discurso perante os personagens virtuais, recebe os estímulos provenientes da simulação que observa; o terapeuta, atento ao comportamento e às respostas do paciente aos estímulos, interage com a aplicação de modo a afetar o decurso da simulação em conformidade, quer alterando o comportamento dos personagens, quer desencadeando eventos específicos no cenário.

A instalação adotada para usar a aplicação é composta por um computador, duas colunas de som, um projetor e uma tela ou uma parede como superfície de projeção. A aplicação cria duas janelas disjuntas: a da simulação, que deve ser projetada, e a da interface, no computador do terapeuta. As colunas de som devem estar junto à projeção para aumentar o grau de realismo da simulação.

Assim, temos as vantagens de o equipamento ser relativamente barato e fácil de instalar, de o tamanho da imagem projetada na tela ser de grandes dimensões, o que facilita a sensação de imersão sentida pelo paciente, e de

haver a possibilidade de várias pessoas assistirem em simultâneo a uma sessão de terapia, o que pode ser valioso, por exemplo, em ambientes de ensino. Além disto, eliminam-se os eventuais efeitos secundários desagradáveis que podem advir da utilização de equipamentos de RV, como vertigens e náuseas.

A versão aqui apresentada segue-se a duas outras: a situação simulada na primeira versão é uma audiência com humanos virtuais [Cláudio13] e, na segunda versão, um júri composto por um a três personagens virtuais [Cláudio14].

3.1 Virtual Spectators - versão atual

O principal objectivo desta etapa de desenvolvimento do Virtual Spectator era o de dotar as personagens de maior realismo e consequentemente de maior credibilidade. Tal objectivo teria que ser possível sem comprometer o desempenho da aplicação, a qual deveria continuar a suportar uma resposta em tempo real à interação do terapeuta. Foi assim necessário encontrar um equilíbrio entre o aspeto final das personagens e os aspetos críticos para a obtenção da resposta em tempo-real, nomeadamente, o detalhe da malha poligonal, o nível de resolução das texturas e os métodos de iluminação utilizados.

Numa primeira instância, constatou-se que não seria possível melhorar quaisquer das personagens das versões anteriores pelo se decidiu recorrer a novas personagens. Rapidamente se concluiu, depois de um conjunto de testes, que os modelos obtidos por recurso a ferramentas gratuitas, por exemplo, MakeHuman² ou Mixamo Fuse³, não satisfaziam o nível de qualidade pretendido. Levou-se então a cabo uma tarefa de combinação de modelos ou partes de modelos obtidos no software Poser⁴ e em repositórios *online*. Deste modo produziram-se duas personagens base do género masculino designadas por John e Carl, e uma personagem feminina, Jessi. Os três personagens aparentam idades distintas.

Procedeu-se à animação destas personagens tendo em conta as características/funcionalidades das principais ferramentas de software envolvidas na concretização da aplicação: o Blender⁵ e o Unity 3D⁶ na sua versão gratuita. Na ferramenta Unity as animações são conseguidas à custa de *rigging animation*, ou seja, pela animação de um esqueleto, uma estrutura hierárquica de ossos ligados entre si. A animação compreende ainda o processo de *skinning*: a associação de um osso a um determinado conjunto de vértices vizinhos. Quando o osso se movimenta, os vértices que lhe estão associados deslocam-se; os vértices mais próximos do osso sofrem maior deslocamento do que os vértices mais afastados.

² <http://www.makehuman.org> (último acesso set 2014)

³ <http://www.mixamo.com/fuse> (último acesso set 2014)

⁴ <http://poser.smithmicro.com/> (último acesso set 2014)

⁵ <http://www.blender.org/> (último acesso set 2014)

⁶ <http://Unity.com/> (último acesso set 2014)

¹ www.microsoft.com/en-us/kinectforwindows/ (último acesso set 2014)

No Blender a associação osso-vértices e os pesos que definem a amplitude do deslocamento é realizada usando a funcionalidade *weight painting*.

De modo a aproveitar tanto quanto possível as animações utilizadas na versão anterior, procedeu-se à integração do esqueleto das antigas personagens usadas pela aplicação nas novas personagens. Contudo, a posição base destes esqueletos não coincide com a posição base da malha poligonal das novas personagens, o que obrigou a realizar um ajustamento manual. Este ajustamento leva à modificação da posição base do esqueleto que resulta no desfaseamento das animações que dependem do(s) osso(s) afectados pelo que cada animação teve também de ser reajustada. Relativamente ao *skinning*, este é realizado de raiz, tanto para o rosto como para o corpo. Dado que as três personagens são anatomicamente diferentes entre si, este processo exaustivo é repetido para cada uma das personagens.

3.2 Expressões e Animações Faciais

A qualidade do modelo da face e as respetivas expressões são elementos fulcrais para a credibilidade da expressão das emoções. Com o propósito de obter estas características foram tomadas as seguintes opções:

- eliminar os ossos da cara não relevantes para as animações pretendidas (por exemplo, ossos associados ao nariz), ao mesmo tempo que se adicionaram outros em pontos específicos a fim de obter efeitos desejados em certas animações (por exemplo, franzir os cantos do olho);
- reduzir o número de polígonos do corpo, roupa e cabelo de cada personagem para poder aumentar, sem perda de desempenho, o número de polígonos da face;
- recorrer a um *shader* adquirido no próprio *site* do Unity para obter melhores resultados gráficos nos modelos.

Como resultado final, verificou-se uma melhoria na aparência das personagens virtuais, tornando possível e interessante o estudo de validação das expressões faciais que realizámos subsequentemente e que se encontra descrito na secção 4. A figura 1 exhibe o personagem John com duas expressões faciais distintas.

3.3 Interface de Configuração

Nesta etapa foram também incorporadas novas funcionalidades na aplicação, o que teve reflexos na interface de configuração do cenário.

As novas funcionalidades implementadas são:

- A possibilidade de alternar entre diferentes cenários. A aplicação está preparada para facilmente incluir outros cenários que se considerem relevantes.
- Um sistema de escolha de personagens em que as personagens estão todas visíveis para o utilizador. O processo de escolha (ou troca) de personagens usa uma interação simples do tipo *drag-and-drop*: o utilizador seleciona a fotografia que representa a personagem escolhida e arrasta-a para a posição que esta ocupará durante a simulação na mesa do cenário virtual; a fotografia passa a ser representada em tons de cinzento

como indicação de que aquela personagem não se encontra mais disponível para escolha;

- Escolher óculos: selecionada uma personagem que figure no cenário virtual é dada a opção de poder utilizar óculos tradicionais ou óculos de sol.
- Escolher um estilo de roupa formal ou informal.
- Escolher um tipo de penteado formal ou informal (exceto na personagem Carl que é calvo). A escolha de cabelo é independente da escolha de roupa, o que permite diversas possibilidades de conjugação. A figura 2 mostra a personagem Jessi com duas aparências distintas.
- Pré-visualizar a personagem: o terapeuta tem a oportunidade de visualizar a personagem selecionada e personalizada com as escolhas realizadas. A aplicação permite visualizar um *close-up* da sua face.



Figura 1: Personagem John com duas expressões faciais distintas.



Figura 2: Personagem Jessi com dois tipos distintos de aparência graças à alteração do cabelo e da roupa.

3.4 A Interface das Expressões Faciais

A figura 3 mostra os três personagens exibindo expressões faciais variadas. A interface que permite ao terapeuta controlar as expressões faciais (figura 4) contém botões que correspondem a unidades de ação faciais ou Action Units (AU) do Facial Action Coding System (FACS) [Ekman02]; em dois AUs (AU4, AU12) existe uma opção para dois níveis de intensidade. Tem havido grande controvérsia em torno da suposição, sem o apoio de estudos de expressão facial espontânea, de que certas expressões faciais transmitem determinadas emoções (ver, por exemplo, [Russell97]). Por este motivo, as AU escolhidas são as mais validadas em estudos de comportamento e percepção [Gaspar12]. Estas AU podem ser combinadas e a sua inserção simultânea no

rosto do humano virtual é o que cria a sua expressão facial.



Figura 3 : Três personagens base exibindo expressões faciais variadas (da esquerda para a direita, John, Carl e Jessi).

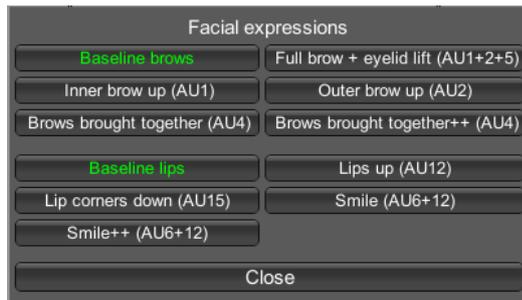


Figure 4. Interface de Controlo das Expressões Faciais. Os botões ficam verdes quando a respetiva expressão está ativa na personagem virtual.

Além disso, estas composições faciais podem também ser associadas com várias opções de postura corporal. As posturas e orientações foram escolhidas de acordo com resultados de conteúdos relevantes na comunicação não-verbal humana [Eibl- Eibbesfeldt89].

O painel de controlo da expressão facial (figura 4) inclui ações para transmitir afeto positivo: AU6+12, com dois níveis de intensidade, ativadas pelos botões "Smile (AU6+12)" e "Smile ++ (AU6+12)"; e só AU12, ativado por "Lips up (AU12)". Incluem-se igualmente ações muitas vezes associadas pelos observadores a retorno negativo, e associadas pelos próprios como atenção (AU4, AU5, AU2, AU1+2) ou afeto negativo (AU4, AU1+4 e AU15). O franzir das sobrancelhas formando rugas entre estas (AU4), produzido em humanos pela contração do corrugador, é particularmente relevante e, portanto, inseriram-se dois níveis de intensidade com os botões "Brows brought together (AU4)" e "Brows brought together (AU4)++". O menu também inclui dois botões para as posições neutras das sobrancelhas e da boca, que permitem obter um personagem quase inexpressivo.

A figura 5 mostra o mesmo modelo com diferentes combinações de AU. O rosto neutro, sem AU ativado (canto superior esquerdo) e uma expressão severa com cenho franzido, AU4 (canto superior direito). Na parte inferior direita, vemos uma combinação de AU4 e AU15,

ambos associados com retorno negativo. No canto inferior esquerdo, vemos uma combinação de AU4 e AU6+12 (o botão de "smile"). A combinação AU6+12 é a característica que define o "*Duchenne smile*", que é associado a afeto positivo, de forma mais honesta do que qualquer outro sorriso. O efeito ameaçador associado ao cenho franzido é mitigado e, possivelmente, este é associado a outro contexto frequente, a atenção.

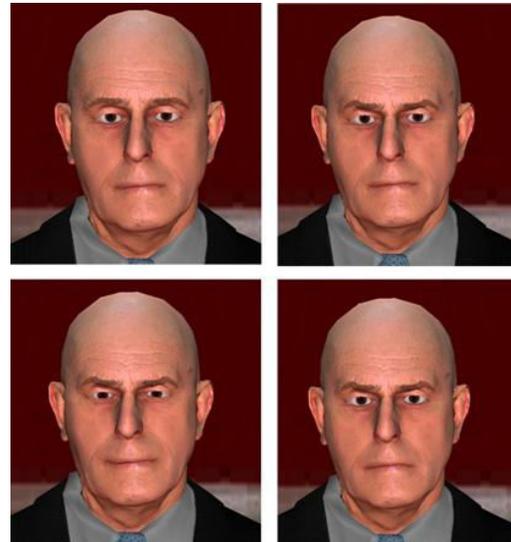


Figure 5- A partir do canto superior esquerdo e no sentido dos ponteiros do relógio: Carl com uma expressão neutra, com as sobrancelhas franzidas (AU4), mostrando uma combinação de AU4 e AU15, mostrando uma combinação de AU4 e AU6+12 (o botão "smile").

4. VALIDAÇÃO DO COMPORTAMENTO EXPRESSIVO DA FACE DOS HUMANOS VIRTUAIS

4.1 Participantes, materiais e procedimentos

Um estudo preliminar de validação dos pressupostos do comportamento facial dos Personagens Virtuais foi efetuado com uma amostra de 38 estudantes universitários (31F;7M), com idades compreendidas entre os 18 e os 25 anos, e que participaram voluntariamente como juízes das imagens após uma breve apresentação do projeto.

Para as experiências, os participantes foram reunidos em grupos de aproximadamente 10, instruídos sobre a tarefa, tendo-lhes sido fornecidas folhas de resposta, bem como formulário de consentimento informado. As experiências decorreram em sala semi-obscura, com a apresentação dos estímulos numa tela de projeção, encontrando-se os participantes em duas filas e a 2-3m de distância da tela, consoante a fila. Os participantes fizeram 3 ensaios iniciais das tarefas e esclareceram todas as dúvidas antes da experiência começar.

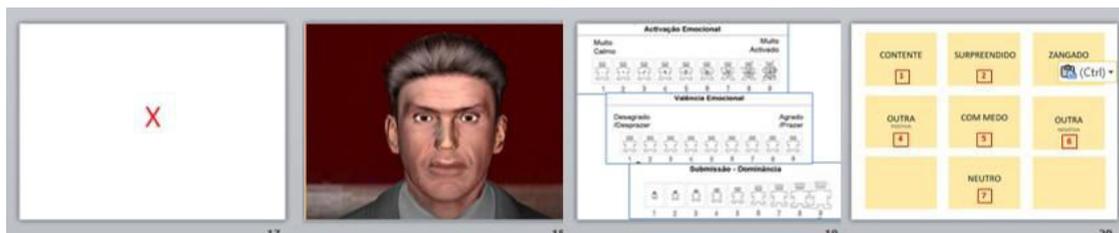


Figura 6. Tipos de ecrãs na experiência.

A experiência consistiu na apresentação de 28 *clips* animados de 3 segundos, contendo cada um grande plano de um dos avatares apresentando uma animação a partir da expressão neutra e evoluindo até uma outra expressão que corresponda a uma Unidade de Ação Facial (AU) ou a uma combinação de AUs. A lista de 14 AUs e combinações testadas foi a seguinte:

- Sobrancelhas em Baseline + Lábios em Baseline
- Aproximação das sobrancelhas AU4
- Aproximação (intensa) das sobrancelhas ++ (AU4)
- Levantamento das sobrancelhas + Levantamento da palpebral superior (AU1+2+5)
- Levantamento dos cantos internos das sobrancelhas (AU1)
- Levantamento dos cantos externos das sobrancelhas (AU2)
- Levantamento dos cantos dos lábios (AU12)
- Sorriso (AU6+12)
- Sorriso intenso ++(AU6+12)
- Abaixamento dos cantos dos lábios (AU15)
- Aproximação (intensa) das sobrancelhas ++ (AU4) + Levantamento dos cantos dos lábios (AU12)
- Aproximação (intensa) das sobrancelhas ++ (AU4)+ Abaixamento dos cantos dos lábios (AU15)
- Levantamento dos cantos internos das sobrancelhas (AU1)+ Levantamento dos cantos dos lábios (AU12)
- Levantamento dos cantos internos das sobrancelhas (AU1)+ Abaixamento dos cantos dos lábios (AU15)

Os *clips* foram apresentados por ordem aleatória. A apresentação de cada *clip* foi precedida por um ecrã de fixação com 1 segundo de duração e seguida da apresentação de escalas pictóricas de ativação, valência e dominância - o SAM (Self Assessment Manikin) [Bradley94] que estão internacional e amplamente validadas para avaliarem a percepção que os participantes têm relativamente ao modo como os estímulos os fazem sentir em termos de cada uma das dimensões inativo-ativo, emoção negativa-positiva e submissão – dominância. Os participantes dispuseram de 15 segundos para esta tarefa. Finalmente surgia um ecrã contendo rótulos com nomes de possíveis estados emocionais (contente, zangado, surpreendido, com medo, outro positivo, outro negativo ou neutro) expressados na face do personagem (Figura 6). Esta tarefa repetiu-se 28 vezes, pois cada combinação de AUs era replicada em 2 personagens virtuais. Os participantes foram encorajados a sugerir novos rótulos sempre que escolhessem as opções “Outro positivo” ou “Outro negativo”.

4.2 Resultados

No que concerne ao conteúdo atribuído pelos participantes às diferentes composições de unidades de ação facial, uma análise de X^2 de Independência mostrou existir associação altamente significativa entre as imagens e as frequências de conteúdos atribuídos ($X^2=2088,07$; $p<=.001$; $N=978$), sendo que 12 dos 28 *clips* obtiveram níveis de convergência acima dos 75%, todos eles representando associações significativas entre a imagem e o conteúdo ($p<=0,05$). As composições faciais com níveis maiores de concordância entre os vários observadores participantes foram as que o quadro 1, em anexo, ilustra. Este quadro mostra as imagens que obtiveram níveis de concordância mais elevados (acima dos 75%) na atribuição de conteúdo. As percentagens indicam a percentagem de participantes que atribuiu aquele conteúdo à imagem.

5. CONCLUSÕES E TRABALHO FUTURO

Os resultados observados na validação preliminar e situada exclusivamente na interpretação das expressões faciais são extremamente encorajadores: por um lado, as composições com elevadas concordâncias vão ao encontro dos pressupostos baseados nas observações de humanos, segundo as quais a interpretação universal tende a restringir-se a sinais de ameaça (como a aproximação de sobrancelhas) ou de disposição amigável/brincalhona (como o sorriso) e, por outro lado, os valores desta concordância são elevados relativamente aos reportados em estudos com humanos [Gaspar13; Gaspar14] o que aponta para as vantagens do controlo de todas as variáveis (fisionómicas, dimensão, luz, vestuário) com os personagens virtuais. É também de salientar como elemento importante nesta validação que os indivíduos identificaram corretamente a face neutra, um controlo importante dado que as imagens que não obtiveram elevadas concordâncias não o terão sido por outras razões que não o facto daquilo que comunicam depender de variáveis do próprio observador, sendo portanto ideais para comparar populações com características diferentes em termos de ansiedade social.

Pretende-se prosseguir com o desenvolvimento da aplicação com vista a um produto mais completo. Nomeadamente, através da concretização de um módulo de comportamentos baseados em personalidade com vista a facilitar o papel desempenhado pelo terapeuta de controlar o comportamento de cada personagem individualmente. Deste modo, este optaria por definir um tipo de personalidade para cada personagem que, por sua vez, adotaria durante a simulação, uma atitude

correspondente à escolha efetuada. Por exemplo, se uma personagem fosse escolhida como sendo “pouco amistosa”, passaria a adotar, automaticamente, comportamentos característicos (por exemplo, franzir sobrancelhas, postura corporal agressiva). Esta funcionalidade permitirá aumentar o número de personagens da cena sem sobrecarregar demasiado o terapeuta, contudo, sem nunca esquecer os aspetos relativos ao desempenho.

De momento ultimamos um módulo para prover as personagens de fala com o intuito de enriquecer a interação entre estas e o paciente.

Por fim, e acima de tudo, a validação da aplicação deverá compreender um processo de tratamento concreto de fobia social a fim de aferir quais as mais-valias da utilização desta aplicação quando comparada aos métodos tradicionais. Dado o interesse acentuado observado pelos terapeutas justificado, em parte, pelas limitações da tipicamente utilizada terapia de exposição *in imagino*, expectam-se resultados positivos e animadores e que justificam *de per se* a continuidade do projeto.

6. AGRADECIMENTOS

Agradecemos à Fundação para a Ciência e Tecnologia (FCT) e ao LabMAg o apoio financeiro ao abrigo do projeto PEst-OE/EEI/UI0434/2014.

7. REFERÊNCIAS

- [APA00] APA- American Psychiatric Association, 2000. DSM-IV-TR. American Psychiatric Publishing, Inc..
- [Beidel07] Beidel, D. C., Turner, S., 2007. *Shy Children, Phobic Adults: Nature and Treatment of Social Anxiety Disorder*. Washington, DC: APA, 2nd edition.
- [Bradley94] Bradley, M.M. and LANG, P.J., 1994. Measuring Emotion: the Self-Assessment Manikin and the Semantic Differential. *J. Behav. Therap. & Exp. Psvchyatry*, 25, 1, 49-59.
- [Cláudio13] Cláudio, A.P., Carmo, M.B., Pinheiro, T., Esteves, F. (2013), A Virtual Reality Solution to Handle Social Anxiety. *Int. J. of Creative Interfaces and Computer Graphics*, 4(2):57-72.
- [Cláudio14] Cláudio, A.P., Gaspar, A., Lopes, E., Carmo, M.B. (2014), Virtual Characters with Affective Facial Behavior. In Proc. GRAPP 2014, pp 348-355, SciTePress.
- [Douilliez12] Douilliez, C., et al., 2012. Social anxiety biases the evaluation of facial displays: Evidence from single face and multi-facial stimuli. *Cog & Emo*, 26(6): 1107-1115.
- [Eibl-Eibesfeldt89] Eibl-Eibesfeldt, I., 1989. *Human Ethology*. NY: Aldine de Gruyter.
- [Ekman02] Ekman, P., Friesen, W.V., Hager, J.C., 2002. *Facial action coding system*. Salt Lake City, UT: Research Nexus.
- [Gaspar12] Gaspar, A., Esteves, F., 2012. Preschoolers faces in spontaneous emotional contexts – how well do they match adult facial expression prototypes? *Int J Behav Dev*. 36(5), 348–357.
- [Gaspar14] Gaspar, A., Esteves, F., Arriaga, P. (2014) On prototypical facial expressions vs variation in facial behavior: lessons learned on the “visibility” of emotions from measuring facial actions in humans and apes. In M. Pina and N. Gontier (Eds), *The Evol. of Social Comm. in Primates: A Multidisciplinary Approach, Interdisciplinary Evolution Research*. NY: Springer.
- [Grillon09] Grillon, H., 2009. Simulating interactions with virtual characters for the treatment of social phobia. Doctoral dissertation, EPFL.
- [Haworth12] Haworth, M. B., Baljko, M., Faloutsos, P., 2012. PhoVR: a virtual reality system to treat phobias. In Proc. of the 11th ACM SIGGRAPH Int. Conf. VR Continuum and Applic in Industry, 171-174.
- [Herbelin05] Herbelin, B., 2005. Virtual reality exposure therapy for social phobia. Doctoral dissertation, EPFL.
- [James03] James, L. K., et al., 2003. Social anxiety in virtual environments: Results of a pilot study. *Cyberpsychology & behavior*, 6(3), 237-243.
- [Klinger05] Klinger, E., et al., 2005. Virtual reality therapy versus cognitive behavior therapy for social phobia: A preliminary controlled study. *Cyberpsychology & behavior*, 8(1), 76-88.
- [Klinger04] Klinger, E., et al., 2004. Virtual reality exposure in the treatment of social phobia. *Studies in health tech. and informatics*, 99, 91.
- [LaViola00] LaViola Jr, J. J., 2000. A discussion of cybersickness in virtual environments. *ACM SIGCHI Bulletin*, 32(1), 47-56.
- [North98] North, M. M., et al., 1998. Virtual reality therapy: an effective treatment for the fear of public speaking. *The Int. Journal of Virtual Reality* 3, 1-6.
- [Pertaub01] Pertaub, D. P., Slater, M., Barker, C., 2001. An experiment on fear of public speaking in virtual reality. *Studies in health tech. and inf.*, 372-378.
- [Pertaub02] Pertaub, D. P., Slater, M., Barker, C., 2002. An experiment on public speaking anxiety in response to three different types of virtual audience. *Presence: Teleoperators & Virtual Environments*, 11(1), 68-78.
- [Russell97] Russell, J. A., Fernandez-Dols, J. M., 1997. What does a facial expression mean? In *The psychology of facial expression*, pp. 3–30. New York, NY: Cambridge University Press.
- [Slater06] Slater, M., et al., 2006. A virtual reprise of the Stanley Milgram obedience experiments. *PloS one*, 1(1), e39
- [Slater99] Slater, M., Pertaub, D. P., Steed, A., 1999. Public speaking in virtual reality: Facing an audience of avatars. *IEEE CG&A*, 19(2), 6-9.
- [Stein00] Stein, M.B. and Kean, Y.M., 2000. Disability and Quality of Life in Social Phobia: Epidemiologic Findings. *Am J Psychiatry* 157: 1606–1613.

Quadro 1

		Branco (não sei)	Com Medo	Conten- te	Neutro	Outro	outro negativo: pena/tristeza	Outro +	Surpreen- dido	Zanga- do
	10 Jessi Sorriso intenso (AU6+12)++	5,7%	0,0%	91,4%	0,0%	0,0%	0,0%	2,9%	0,0%	0,0%
	12 Jessi Aproximação (intensa) das sobrancelhas (AU4) ++	11,4%	0,0%	0,0%	0,0%	5,7%	0,0%	0,0%	2,9%	80,0%
	16 Jessi Levant das sobrancelhas e abertura da pálpebra AU1+2+5	5,7%	11%	0,0%	0,0%	2,9%	0,0%	2,9%	77,1%	0,0%
	19 Jessi Aproxim. das sobrancelhas (AU4)	8,6%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	2,9%	88,6%
	22 Jessi sorriso (AU6+12)	5,7%	0,0%	80,0%	0,0%	5,7%	0,0%	5,7%	0,0%	2,9%
	24 Carl aproxim.intensa das sobrancelhas (AU4) ++ e abaixamentos dos cantos dos lábios (AU15)	5,7%	0,0%	0,0%	2,9%	5,7%	0,0%	0,0%	5,7%	80,0%
	28 Jessi Levantamento dos cantos dos lábios (AU12)	5,7%	0,0%	80,0%	0,0%	0,0%	0,0%	14,3%	0,0%	0,0%
	4 Jessi (AU4)	11,4%	0,0%	0,0%	2,9%	0,0%	0,0%	0,0%	2,9%	82,9%
	6 Jessi Baseline (sem expressão)	5,7%	2,9%	0,0%	85,7%	0,0%	0,0%	2,9%	2,9%	0,0%
	7 Carl Sorriso (AU6+12)	8,8%	0,0%	76,5%	5,9%	0,0%	0,0%	8,8%	0,0%	0,0%
	8 Jessi aproxim.intensa das sobrancelhas (AU4 ++) e abaixamentos dos cantos dos lábios (AU15)	8,6%	2,9%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	88,6%
	9 Carl Aproximação das sobrancelhas (AU4)	5,7%	0,0%	0,0%	11,4%	5,7%	0,0%	0,0%	0,0%	77,1%