

ESTABELECENDO REGRAS POR MEIO DE MACHINE LEARNING PARA MODULAÇÃO DE PREVISÕES DE AFETO E ESFORÇO

Rodrigo Silveira¹ Raul Prado¹ Douglas Alves¹ Luiz Fernando Santos² Etilene Santos² Milena Delfim² Marcus Vinicius Pinto²
Raul Canestri¹ Gustavo Vasconcelos¹ Paulo Alvarenga¹ Cayque Brietzke¹ Ricardo Asano¹

Resumo: O objetivo do estudo foi identificar por meio de *Machine Learning* regras para a modulação das previsões de afeto (FS) e percepção subjetiva de esforço (PSE) em relação ao exercício, analisando o grau de importância de diferentes fatores biológicos. Cada participante ($n = 34$ [52.9% sexo feminino], Idade = 25.4 ± 6) foi submetido a um teste em que precisou prever o quanto sentiria de FS e PSE em relação a um *Sprint* de 107 metros que estava prestes a realizar. As respectivas respostas foram analisadas através do algoritmo de árvore decisional para estabelecer regras para a modulação da FS e PSE com base nas seguintes características de cada sujeito: percentual de gordura (%G); massa magra (%MM); relação cintura e quadril (RCQ); nível de atividade física autodeclarado (NAF-AUTO); frequência cardíaca em repouso (FCr); sexo e idade. Como resultado, apesar de não significativo, os indivíduos muito ativos previram em média uma resposta afetiva 2.28 (95% IC = 0.01 - 4.55; $p = 0.06$) unidades maior do que aqueles considerados insuficientemente ativos. As principais regras para os participantes preverem os mais baixos valores de FS foram ter as seguintes características: $\%MM \leq 48.9$, $FCr > 73$. Os mais altos valores de FS apresentavam $\%MM > 48.9$. Os mais altos valores de PSE apresentavam $\%G > 16.4$, enquanto os mais baixos apresentavam $\%G \leq 16.4$. Conclui-se que as regras para o perfil que previu o maior desprazer foi apresentar composição corporal e capacidade cardíaca menos saudáveis, o que previu maior prazer apresentava características opostas. No caso da PSE, existem dois perfis iniciais divididos pelo %G, sendo que o com menor valor previu menor esforço e o perfil com maior %G apresentou maior heterogeneidade nas características.

Palavras-chave: Respostas afetivas; Percepção subjetiva de esforço; Composição corporal; Nível de atividade física; Machine learning - aplicação.

¹ Escola de Artes, Ciências e Humanidades da Universidade de São Paulo (EACH-USP); ² Universidade de Mogi das Cruzes (UMC).

ESTABLISHING RULES USING MACHINE LEARNING FOR MODULATING PREDICTIONS OF AFFECT AND EFFORT

Abstract: The aim of this study is to identify through Machine learning rules for predictions modulation of affect (FS) and rating of perceived effort (PSE) related to exercise, analyzing the degree of importance of different biological factors. Each participant ($n = 34$ [52.9% females], Age = 25.4 ± 6) underwent a test in which they had to predict how much FS and PSE would feel for Sprint before perform it. Subsequently, the respective responses were analyzed using a decision tree algorithm to establish rules for modulation of FS and PSE based on the following characteristics of each subject: fat percentage (%G); muscular mass (%MM); waist and hip ratio (RCQ); self-declared physical activity level (NAF-AUTO); resting heart rate (FCr); sex and age. As result, although not significant, very active individuals predicted an average affective response of 2.28 (95% CI = 0.01 - 4.55; $p = 0.06$) units higher than those considered insufficiently active. The main rules for participants to predict the lowest FS values were to have the following characteristics, $\%MM \leq 48.9$ followed by $FCr > 73$. The highest FS values were $\%MM > 48.9$. The highest PSE values were $\%G > 16.4$, while the lowest values were $\%G \leq 16.4$. It was concluded that the rules for profile that predicted the higher displeasure was to present less healthy body composition and cardiac capacity, which predicted higher pleasure had opposite characteristics. In the case of PSE, there are two initial profiles divided by %G, and the one with the lowest value predicted less effort, the profile with the highest %G has greater heterogeneity in characteristics.

Key words: Affective responses; Rating of perceived effort; Body composition; Physical activity level; Machine learning - application.

Introdução

Atualmente uma das maiores preocupações relacionadas a saúde pública é o aumento da prevalência da obesidade, hipertensão e diabetes, que refletem diretamente na incidência de morbidade e mortalidade¹⁻⁴. Em contrapartida, é consolidado que a prática regular de exercício físico e níveis considerados suficientes de atividade física promovem diversas adaptações ao organismo, sendo que, o exercício físico serve como uma estratégia não medicamentosa na prevenção e controle de distúrbios metabólicos e cardiovasculares⁵. Mesmo com os benefícios associados à prática da atividade física, ainda é grande o número de pessoas que são insuficientemente ativas. Dados do IBGE, estimam que cerca de 100,5 milhões de brasileiros, equivalente a 62,1% da população nacional, entre eles jovens e adultos, são insuficientemente ativos⁶, cenário similar aos de outros países, sejam eles desenvolvidos ou em desenvolvimento, segundo a Organização Mundial de Saúde⁷.

Tal paradoxo é reconhecido como um dos fatores capazes de influenciar no aumento de pessoas com obesidade. Por outro lado, estudos vem destacando o oposto, ou seja, a obesidade é uma das causas da prática insuficiente de atividades físicas. Isso é assumido pelo fato de indivíduos com sobrepeso/obesidade possuírem uma percepção negativa ao exercício comparado a pessoas eutróficas⁸. Tais percepções aversivas podem ser expressadas tanto pelo sentimento de desprazer pelo exercício⁹, quanto por uma maior percepção de dificuldade^{10,11} e esforço¹² comparado com sujeitos eutróficos.

Com isso, essas respostas perceptuais tem ganhado espaço nas publicações científicas nos últimos anos, sendo evidenciadas como fatores relacionados ao nível de aderência ao exercício, tornando-se elementos essenciais para prescrição de uma rotina de exercícios em que o indivíduo consiga ter maior aderência¹³. Porém, apesar de respostas de sensação como o afeto serem amplamente estudadas durante e após o exercício, ainda são pouco estudadas antes da prática.

Contudo, em meio desse atual paradigma sobre a relação entre obesidade e prática insuficiente de atividades físicas⁸, surgiu uma teoria denominada como Conflito Regulatório (TCR)^{10,14}, cuja mesma aborda que todo ser-humano sofre um conflito interno quando está diante de uma situação em que precisará realizar esforço físico. Esse conflito interno além de inconsciente, modula a maneira com que o sujeito avalia a dificuldade da situação, sendo-que, quanto maior o Conflito Regulatório, mais extrema e difícil a situação é percebida.

A TCR é fundamentada em estudos clássicos que observaram diferentes percepções a tarefas físicas por sujeitos com diferentes características físicas e estados corporais^{15,16}. Assim,

o que determina o tamanho do conflito regulatório é o potencial fisiológico, conceituado como um conjunto de características que estimam a saúde de um indivíduo (ex., composição corporal e capacidade cardíaca). Assim, segundo a teoria, quanto pior o potencial fisiológico maior será o conflito regulatório sofrido, e consequentemente, o sujeito irá prever uma situação “talvez” mais extrema do que a realidade.

De acordo a teoria¹⁰, o possível motivo de um sujeito vivenciar o conflito regulatório é em virtude de um instinto natural de economizar energia. Dessa maneira, quanto pior o potencial fisiológico, mais o sujeito inconscientemente terá suas percepções da tarefa alteradas para poder preservar energia.

Entretanto, apesar de estudos mostrarem dados reforçando que características biológicas como sobrepeso/obesidade¹² e/ou a capacidade cardíaca¹⁷ estão associadas à aversão pelo esforço físico conforme descrito na TCR, esses achados são apresentados por meio de métodos estatísticos que não são capazes de levar em conta a multifatorialidade do fenômeno. Ou seja, apesar de sua consistência, respostas psicológicas possuem grande complexidade por serem moduladas por fenômenos multifatoriais.

Devido ao crescimento de perguntas complexas em estudos sobre exercício, estão sendo cada vez mais implementados estudos com *Machine learning*^{18,19} pela sua capacidade de descrever fenômenos complexos de maneira mais simplificada e com alta precisão.

Dessa maneira, o presente estudo levanta as seguintes questões: 1) Será que por meio de *Machine learning* é possível compreender como o afeto e a PSE são modulados antes de iniciar um exercício? 2) Se sim, será que os fatores que modulam as percepções previstas em relação ao exercício apresentarão características similares aos da TCR?

Sendo assim, o objetivo do estudo foi identificar por meio de *Machine Learning* regras para a modulação das previsões de afeto e PSE em relação ao exercício, analisando o grau de importância de diferentes fatores biológicos. Baseando-se na TCR acreditou-se que os sujeitos com maior percentual de gordura corporal, menor percentual de massa muscular, pior capacidade cardíaca, que se autodeclararam insuficientemente ativos e com maior idade iriam prever maior sensação de desprazer e maior PSE ao exercício, e quanto mais opostas fossem essas características, mais opostas as previsões também seriam.

Com base em estudos anteriores^{8,10-12}, acreditou-se que a gordura corporal seria o fator mais importante na modulação dessas percepções.

Método

Participantes

Após a liberação do comitê de ética em pesquisa institucional (CAAE: 56483916.1.0000.5497), foram convidados a participar do estudo 34 alunos do curso de graduação em Educação Física da Universidade de Mogi das Cruzes, com idades entre 18 e 35 anos, sendo 16 do sexo masculino e 18 sexo feminino (Tabela 1), e que tenham concluído pelo menos a disciplina de Atletismo do curso, garantindo que tenham maior similaridade no nível de conhecimento em técnicas de corrida e na experiência com as sensações da prática. O estudo seguiu um delineamento transversal e foi excluído do estudo qualquer indivíduo que apresentasse alguma das seguintes características: ser menor de idade; não ter concluído pelo menos o 1º ano do curso; relatar já ter sido diagnosticado com algum distúrbio psicológico; fazer uso de medicamento para alguma doença cardiovascular ou neurológica; ter algum tipo de deficiência física ou neurológica.

Procedimentos

Desenho experimental

Antes da aplicação dos testes, foi detalhado aos participantes o procedimento experimental, em seguida foi entregue o termo de consentimento livre e esclarecido (TCLE) que foi lido e assinado pelos mesmos.

Ao chegar no local do estudo, cada participante foi submetido à uma série de avaliações (exemplificadas na figura 1), sendo respectivamente:

1) Coleta da composição corporal, do nível de atividade física autodeclarado (NAF-AUTO) - Nessa etapa era coletado o índice de massa corporal (IMC), relação cintura-quadril (RCQ), percentual de gordura corporal (%G), percentual de massa magra (%MM) e era questionado ao participante se o mesmo se considerava um sujeito insuficientemente ativo (pratica alguma atividade física, porém, menos de 150 minutos semanais de atividade física de intensidade moderada, menos de 75 minutos semanais de atividade física de intensidade vigorosa ou uma combinação equivalente), fisicamente ativo (faz o equivalente a 150 minutos a 300 minutos de atividade física de intensidade moderada por semana) ou muito ativo (faz o equivalente a mais de 300 minutos de atividade física de intensidade moderada por semana) de acordo as classificações do *American College Sports Medicine*²⁰;

2) Coleta da frequência cardíaca em repouso (FCr);

3) Coleta das previsões de *Feeling Scale* (FS) e percepção subjetiva de esforço (PSE) antes de realizar um Sprint máximo de 107 metros (espaço máximo e mais exequível do ambiente de coleta para a realização da tarefa).

Para essa última etapa, foram posicionadas uma linha de largada e outra de chegada, medindo 107 metros uma da outra. Ao chegar nessa etapa, o participante era instruído a correr até a linha de chegada realizando um único *Sprint* máximo. Porém, antes de começar a correr, o participante precisava reportar qual o valor predominante de FS e de PSE que o mesmo previa que sentiria durante a execução do *Sprint*. Após a coleta das escalas o participante realizava o *Sprint*.

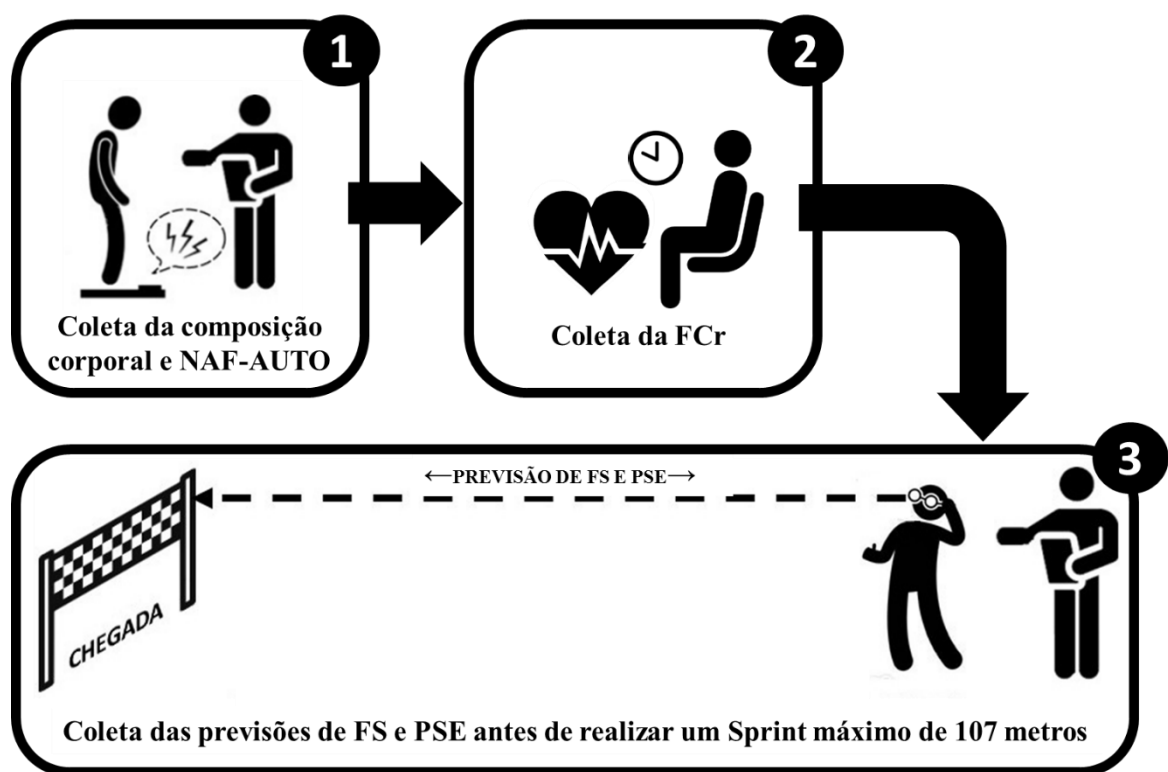


Figura 1 - Desenho experimental.

NAF-AUTO = nível de atividade física autodeclarado / FCr = frequência cardíaca em repouso / FS = *Feeling Scale* / PSE = percepção subjetiva de esforço.

Composição corporal

Para a análise da composição corporal foram mensurados estatura em centímetros (cm), massa em quilogramas (kg), circunferência da cintura e do quadril em cm. Para o cálculo do

IMC divide-se a massa do sujeito pela estatura ao quadrado, para o cálculo do RCQ deve-se dividir a circunferência da cintura pela do quadril²¹. Para o %G e %MM foi utilizada uma balança de bioimpedância (Balança Digital BalgPro G-TECH®).

Respostas psicológicas ao exercício

Para que o participante estimasse como seriam suas percepções quanto ao esforço físico, antes mesmo de realizá-lo, foi adotado o protocolo utilizado por Cole; Balcetis; Zhang (2013)¹⁰, no qual o participante sem ter nenhum tipo de familiarização, precisa reportar a sua previsão no momento em que descobre que realizará o *sprint* e imediatamente antes de executá-lo.

Assim, para a previsão de PSE foi utilizado a escala de percepção subjetiva de esforço²², cujo consiste em uma escala de 15 pontos, sendo 6 o menor esforço percebido e 20 o maior esforço percebido. Ao reportar a PSE, o participante foi orientado a considerar falta de ar, trabalho cardíaco, temperatura corporal e o desconforto na musculatura dos membros inferiores induzido pela corrida. Após receber as orientações sobre como reportar a PSE, o participante respondia a seguinte pergunta “Quanto esforço que você sentirá correndo daqui até a linha de chegada?”.

Para a mensuração da valência afetiva prevista foi utilizada a FS²³, uma escala bipolar de prazer e desprazer, sendo que -5 representa o maior desprazer percebido, 0 representa um valor neutro, e +5 representa o maior prazer percebido. Após receber as orientações sobre como reportar a FS, o participante respondia a seguinte pergunta “Como você se sentirá correndo na maior velocidade possível até a linha de chegada?”.

Capacidade cardíaca

Para avaliar a capacidade cardíaca foi realizada a coleta da FCr, que para ser mensurada o participante se mantém sentado confortavelmente em uma cadeira por 5 minutos, sendo que, o mesmo é orientado a não se movimentar, dentro desse período o valor mais baixo de FC é considerado a FCr, seguindo assim um padrão já utilizado anteriormente²⁴. Para a aquisição do sinal foi utilizado um aparelho *Microlife BP 3BT0-A®*.

Análise Estatística

Para a caracterização da amostra foram utilizados modelos estatísticos inferenciais, sendo que, inicialmente foram utilizados histogramas e gráficos quantil-quantil para analisar a distribuição dos dados. Feito isso, os dados contínuos foram apresentados por média e desvio padrão. Assim, para comparar a FCr entre os diferentes níveis de NAF-AUTO, foi realizado o teste Kruskal-Wallis com o tamanho de efeito observado pelo coeficiente de Cohen (d), assumindo $d < 0.2$ para não efeito, $d < 0.5$ para efeito pequeno, $d < 0.8$ para efeito moderado e $d \geq 0.8$ para efeito grande²⁵.

Para observar a relação entre PSE e FS, controlando-a pelo NAF-AUTO dos indivíduos, optou-se por um Modelo Linear Generalizado (GzLM), sendo assim testado um modelo com distribuição de probabilidade do tipo linear (simétrica) e outro com uma distribuição do tipo Gamma (assimétrica) e posteriormente ambos foram comparados para se analisar qual deles aderiu melhor aos dados. Para comparar qual distribuição aderiu mais aos dados, utilizou-se o critério de informação de Akaike (AIC), sendo que, considera-se que quanto menor o valor de AIC, mais os dados estão aderentes ao modelo. Para todas as análises inferenciais o alfa adotado foi de $p < 0.05$, e foram processadas no *software* SPSS 23.

Posteriormente, os dados de PSE e FS foram analisados por meio do algoritmo de *Machine learning* Árvore de Decisão, cujo mesmo, através dos cálculos de entropia ($Entropia(S) = \sum_{i=1}^c -\rho_i \log_2 \rho_i$) e ganho de informação ($Ganho(S, A) = Entropia(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} Entropia(S_v)$) determina regras para descrever determinados fenômenos por meio de uma representação gráfica hierárquica com um formato que se assemelha a uma árvore. Dessa maneira, as variáveis que apresentarem maiores valores de ganho de informação são as mais importantes para a modulação do fenômeno observado, sendo consequentemente expressas no topo da árvore. Sendo assim, o resultado da Árvore de Decisão expresso no topo é a variável de maior importância para o fenômeno observado, ramificando através das regras estabelecidas pelo próprio algoritmo até chegar na variável com menor importância.

Para o presente estudo, foram adotados dois modelos de Árvore de Decisão para dados contínuos, sendo um para identificar regras para a modulação da previsão de FS e outro de PSE, cada um recebeu como variáveis preditoras a idade, %G, %MM, RCQ, sexo, NAF-AUTO e FCr. A arquitetura adotada para o modelo foi determinada por meio de testes de validação cruzada, cujas estratificações foram configuradas com o número de repartições de 10 *folds*, assumindo como indicadores de desempenho, valores de erro absoluto médio (MAE) e a

proporção da variância explicada pelo o modelo (R^2).

A interpretação do valor de R^2 consiste em quanto o modelo é capaz de explicar a variância dos dados, no caso do presente estudo, as variações da PSE e FS nos diferentes participantes. Assim, um modelo que não explica nenhuma variância teria um R^2 de 0. Um modelo com um R^2 de 1 explicaria toda a variância.

A interpretação do valor de MAE consiste em uma média de quanto os valores previstos pelo modelo foram diferentes da média dos valores reais. Como exemplo, no presente estudo, se o valor de MAE é igual a 0 significa que a média de todas as previsões do modelo foram iguais ao valor médio dos valores reais da PSE ou FS. Entretanto, se o valor de MAE for igual a 3, significa que o erro do modelo em média foi de 3 pontos na respectiva escala.

A arquitetura que apresentou melhor desempenho para o modelo preditor de FS tinha os seguintes parâmetros: sem indução binária; número mínimo de instâncias por folha da árvore = 2; não realizar divisões de subconjuntos < 7; sem limite de profundidade da árvore; a classificação para quando a maioria dos nós de decisão alcançarem 95%. O modelo preditor de PSE com melhor desempenho tinha os seguintes parâmetros: sem indução binária; número mínimo de instâncias por folha da árvore = 2; não realizar divisões de subconjuntos < 6; sem limite de profundidade da árvore; a classificação para quando a maioria dos nós de decisão alcançarem 95%. Todas as análises foram realizadas pelo *software Orange Data Mining 3.23.1*.

Resultados

Dos 34 indivíduos do estudo, 17.6% ($n = 6$) se autodeclararam ser insuficientemente ativos, 61.8% ($n = 21$) fisicamente ativos e 20.6% ($n = 7$) muito ativos e suas respectivas FCr eram 76 (51-95), 78 (56-114) e 68 (50-79) batimentos por minuto (Kruskal-Wallis = 4.579; $d = 0.6$; $p = 0.10$; Poder $[1 - \beta] = 0.32$).

Tabela 1 - Caracterização da amostra

	Média	DP
Idade	25.4	6
%G	25.0	10
%MM	48.4	11.5
IMC	26.6	4.94
RCQ	0.85	0.08

Legendas: DP = Desvio padrão / IMC = Índice de massa corporal / RCQ = Relação cintura-quadril/ %G = Percentual de gordura / %MM = Percentual de massa magra.

A partir da análise da aderência dos dados ao modelo, pode-se observar que o modelo com distribuição linear apresentou um menor valor de AIC (152.08) quando comparado ao modelo Gamma (162.11), sendo assim optado pela utilização do GzLM com distribuição linear. Por meio do mesmo, não se foi possível observar efeito significativo da PSE ($\chi^2(1) = 2.06$; $p = 0.15$; Poder $[1 - \beta] = 0.60$) nem mesmo do NAF-AUTO ($\chi^2(2) = 4.14$; $p = 0.13$) sobre os valores de afeto previstos pelos indivíduos. Entretanto, pôde-se observar que, a cada aumento na previsão de PSE, foi-se observado em média uma redução de 0.15 (95% IC = -0,36 - 0.05; $p = 0.16$) unidades na resposta afetiva prevista pelos indivíduos. Assim como, como expresso na figura 2, pôde-se observar que, em média os indivíduos considerados fisicamente ativos previram uma resposta afetiva 0.85 (95% IC = -1,04 - 2,74; $p = 0.38$) unidades maior do que os indivíduos considerados insuficientemente ativos, assim como os indivíduos considerados muito ativos previram em média uma resposta afetiva 2.28 (95% IC = 0.01 - 4.55; $p = 0.06$) unidades maior do que aqueles considerados insuficientemente ativos.

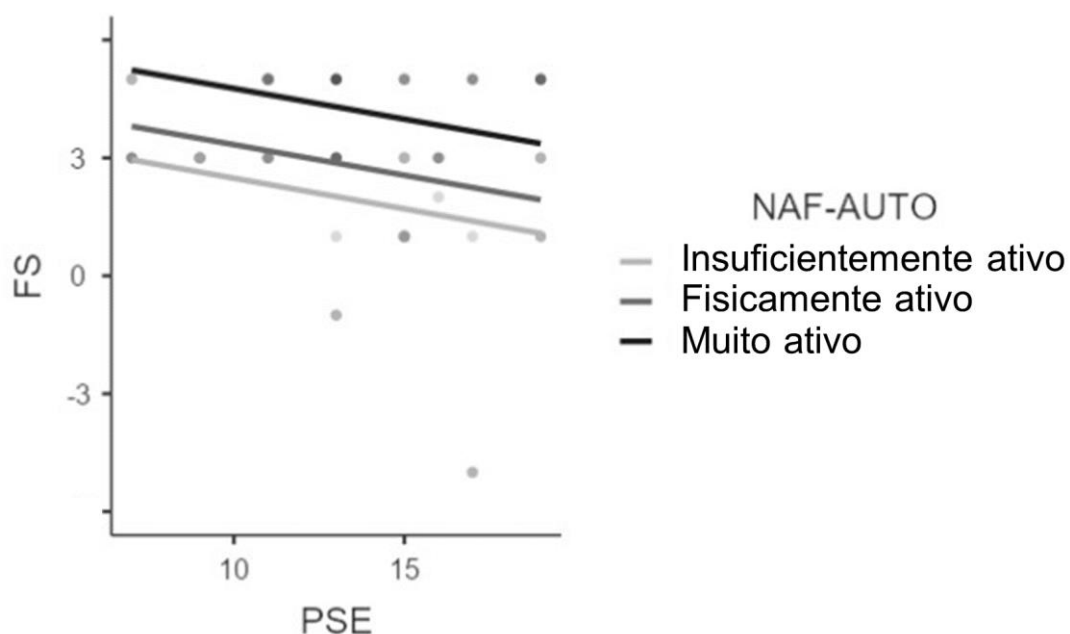


Figura 2 - Relação entre percepção subjetiva de esforço e afeto, controlado pelo nível de atividade física autodeclarado pelos indivíduos.

FS = *Feeling Scale* / PSE = Percepção subjetiva de esforço / NAF-AUTO = Nível de atividade física autodeclarado.

Os modelos de árvore decisional aplicados apresentaram diferentes características quanto o desempenho (Tabela 2), pois, o modelo para FS apresentou um valor de MAE equivalente a 18.2% de erro dentro de uma escala de 11 pontos com um R^2 que explica 60% da variância do modelo. Enquanto o modelo para PSE, apresentou um valor de MAE equivalente a 26.6% de erro dentro de uma escala de 15 pontos, porém com um maior R^2 que explica 90% da variância modelo.

Tabela 2 - Desempenho dos modelos de Árvore de Decisão

	MAE	R^2
Modelo para FS	2	-0.6
Modelo para PSE	4	-0.9

FS = *Feeling Scale* / PSE = Percepção Subjetiva de esforço / MAE = Erro absoluto médio / R^2 = Proporção da variância explicada pelo modelo.

Quanto aos moduladores da FS, pôde-se observar (Figura 3) que a variável de maior importância foi o %MM, sendo que essa variável divide a amostra em sujeitos que previram maior FS (%MM > 48.9) e sujeitos que previram menor valor de FS (%MM ≤ 48.9). Na segunda camada da árvore, o que determinou a previsão de FS na ramificação de sujeitos com menor %MM foi a FCr, sendo que, os sujeitos com menor FCr apresentaram maior FS, enquanto dentre os sujeitos com maior FCr estão os que apresentam os mais baixos valores de FS.

Os sujeitos das ramificações com maior %MM além de apresentarem os valores mais altos de FS, os mesmos apresentam menor heterogeneidade em suas características. Assim nesta ramificação com maior %MM, na segunda camada da árvore o RCQ foi a variável que apresentou maior importância na modulação da FS, ou seja, maiores valores de RCQ apresentaram os menores valores de FS e vice versa.

Na terceira camada da árvore o %G foi a variável que apresentou maior importância na modulação da FS. Ao todo foram 3 cenários em que o %G influenciou os valores de FS, sendo 1 cenário na ramificação com maior %MM (cujo ponto de corte no %G é 10.6) e 2 cenários nas ramificações com menor %MM (cujos pontos de corte no %G foram de 22.8 e de 33.1).

Apesar dos pontos de corte de cada um dos 3 cenários apresentar a tendência de um maior valor de %G estar associado à um maior valor de FS. Foi possível evidenciar que o menor ponto de corte (10.6) possui de maneira menos heterogênea um maior valor de FS. Enquanto nos outros cenários (pontos de corte 22.8 e 33.1) houveram maior variabilidade nos valores de

FS, sendo que a única ramificação que apresentou valor de FS similar ao do cenário com menor ponto de corte de %G era dependente dos sujeitos apresentarem menor FCr ($\%MM \leq 48.9$, $FCr \leq 73$, $\%G > 33.1$, $n = 2$).

Por fim, para a ramificação com $\%MM \leq 48.9$, $FCr > 73$, $\%G > 22.8$, a árvore considerou a idade na quarta camada, sendo que, idades acima 29 anos previram maior FS.

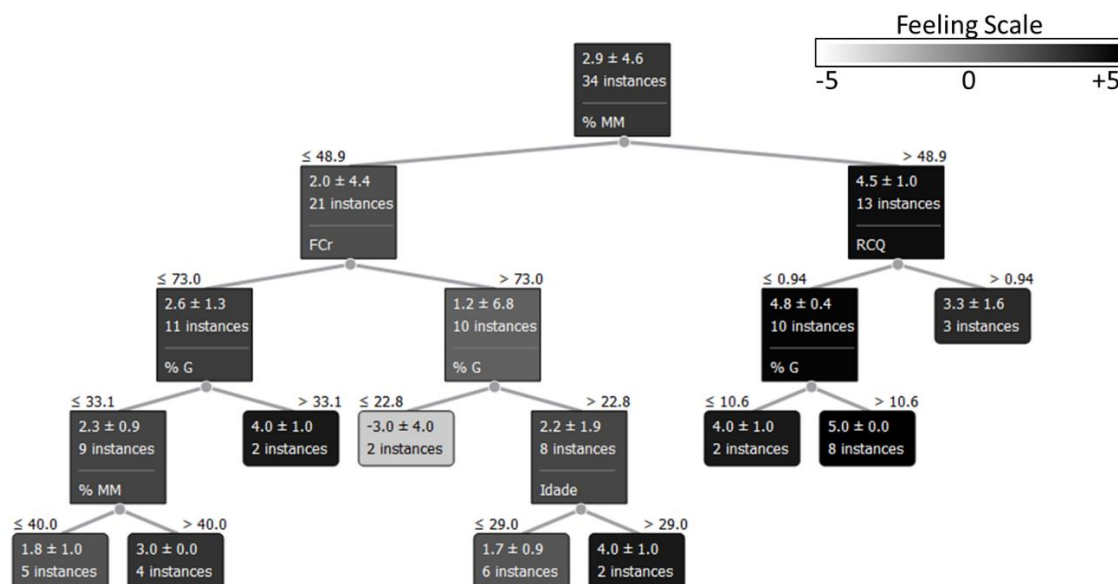


Figura 3 - Regras da modulação da *Feeling Scale*.

Instances = Número de sujeitos / RCQ = Relação cintura-quadril/ $\%G$ = Porcentagem de gordura / $\%MM$ = Porcentagem de massa magra / FCr = Frequência cardíaca em repouso.

Diferente da árvore para FS (Figura 3), a árvore para PSE prevista (Figura 4) apresentou características com maior complexidade quanto aos seus moduladores. Assim, a média das previsões de PSE em toda amostra ($n = 34$) é 13.6 ± 10.9 .

Os sujeitos que previram valores de PSE abaixo da média da amostra inteira apresentaram perfis diferentes, sendo que a característica mais importante desses sujeitos foi ter $\%G \leq 16.4$. Os sujeitos com $\%G > 16.4$ apresentaram maior variabilidade nas previsões de PSE, ou seja, os outros perfis que previram valores de PSE abaixo da média da amostra total apresentaram as seguintes características (em ordem de importância da árvore): 1º) $\%G > 16.4$ e $RCQ > 0.91$ (apresentando menores valores se $FCr \leq 77$); 2º) $\%G > 16.4$, $RCQ \leq 0.91$ e $\%MM \leq 37.8$; 3º) $\%G > 23.8$, $RCQ \leq 0.91$, $\%MM$ entre 37.9 e 48.9 e idade > 29 .

Todas as ramificações da árvore com características diferentes dos padrões acima apresentaram valores de PSE acima da média de toda amostra.

Comparado com a árvore da previsão de FS, a árvore da PSE apresentou maior

heterogeneidade nas características, pois, além dos sujeitos com menores valores de PSE serem divididos em perfis com diferentes características conforme já descrito acima, aqueles com maiores valores de PSE percorrem até 8 camadas na árvore, sendo que, dentro dessas ramificações existem 3 perfis diferentes de sujeitos que reportaram maior PSE, sendo um na quinta camada da árvore ($PSE = 18 \pm 0.9$, $n = 3$), outro na sexta camada ($PSE = 17 \pm 1.6$, $n = 3$) e outro na oitava ($PSE = 18 \pm 0.1$, $n = 2$).

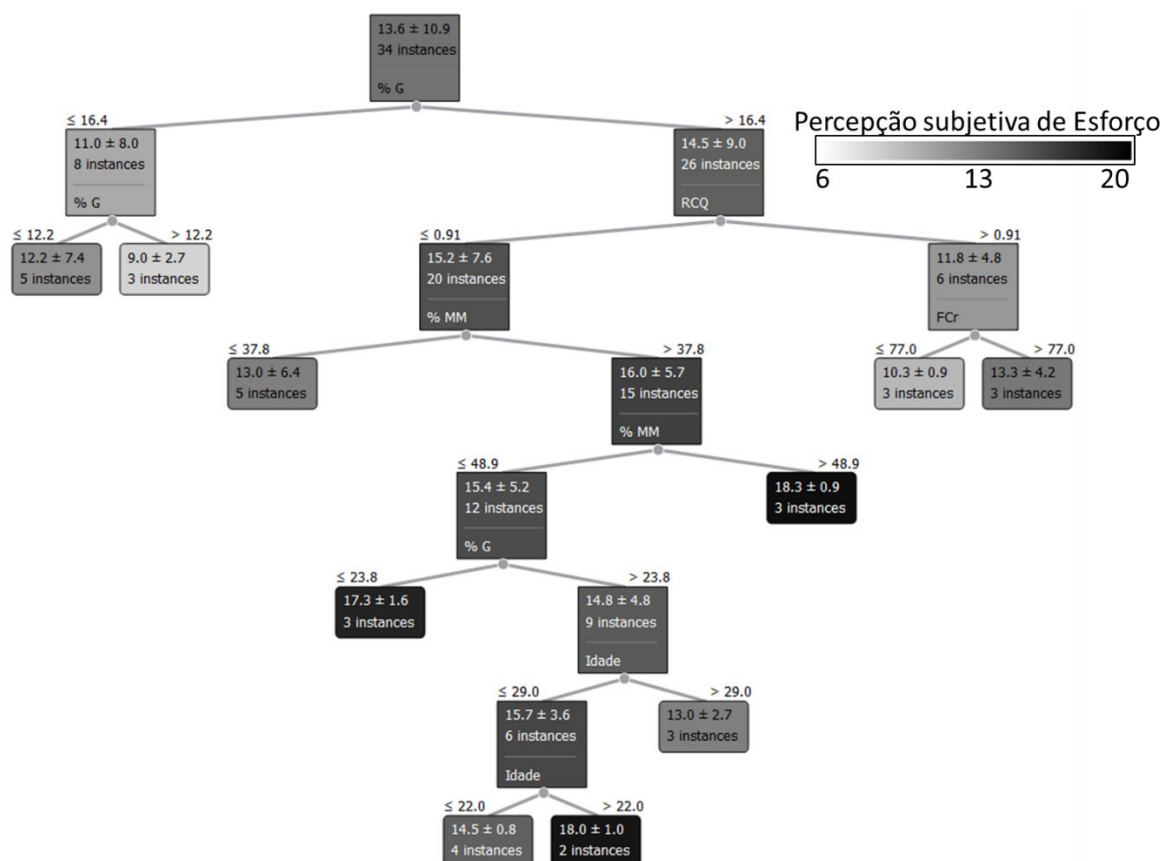


Figura 4 - Regras da modulação da percepção subjetiva de esforço.

Instances = Número de sujeitos / RCQ = Relação cintura-quadril / %G = Porcentagem de gordura / %MM = Porcentagem de massa magra / FCr = Frequência cardíaca em repouso.

Discussão

O presente estudo teve como objetivo identificar regras por meio de *Machine Learning* para a modulação das previsões de afeto e PSE em relação ao exercício, analisando o grau de importância de diferentes fatores biológicos. Como resultado houveram tanto sujeitos que previram que o exercício seria mais desprazeroso (valor de FS mais próximo de -5) quanto sujeitos que previram que seria mais prazeroso (valor de FS mais próximo de 5), assim como

houveram indivíduos que mesmo sabendo que seria um *Sprint* máximo previram que a PSE seria baixa (próximo de 6) quanto outros previram que seria alta (próximo de 20).

Obtida essa variabilidade de respostas, pôde-se estabelecer regras para esses diferentes padrões de percepção e sensação.

Antes de aplicar o algoritmo de árvore decisional, a amostra foi subdividida pelo NAF-AUTO para avaliar se os dados se assemelham aos achados da literatura. Como resultado, primeiramente pôde-se observar que sujeitos insuficientemente ativos previram um menor valor de FS mesmo antes de realizá-lo. Além disso, apesar de nenhuma significância estatística ($p > 0.05$), o comportamento da amostra foi similar ao que a literatura defende¹³, ou seja, em sujeitos que se declararam insuficientemente ativos quanto maior a PSE, apresentaram menor valor de FS. Por outro lado, quanto mais ativo o indivíduo for, mais o mesmo tende a apresentar maior prazer por altas intensidades.

Assim, tendo em vista que a amostra apresenta um comportamento similar ao observado pela literatura, foi aplicada a árvore decisional, onde pôde-se observar que a primeira característica que divide os sujeitos que previram maiores e menores valores de FS foi o %MM, quem apresentou maior %MM reportava maior FS e vice versa. Os sujeitos que apresentaram maiores valores de FS tinham melhor composição corporal em diferentes aspectos, ou seja, além de ter maior %MM, tinham também perfis predominantemente eutróficos (baixos valores de RCQ e %G).

Nos sujeitos com menor %MM, apesar de haver sujeitos com altos valores de FS, esse padrão comportamental era dependente da capacidade cardíaca dos indivíduos para preverem um maior valor de FS, pois, os que apresentaram maior FCr, indicando capacidade cardíaca inferior²⁶, consequentemente apresentaram os mais baixos valores de FS. Além disso, a árvore também mostrou que a idade abaixo de 30 anos influenciou negativamente a FS em alguns dos indivíduos com menor %MM.

Entretanto, no cenário da PSE apesar de haver uma divisão inicial mostrando que os sujeitos com menor %G previram menor PSE, essa árvore apresentou grande heterogeneidade nas características que determinam uma alta ou baixa previsão de PSE. Esse fenômeno pode ser explicado pela figura 2 e pela literatura²⁷, ou seja, a PSE muda de acordo o NAF-AUTO, sendo que, quando o sujeito é insuficientemente ativo, o mesmo tende a associar um valor alto de PSE como algo desprazeroso enquanto os muito ativos associam uma alta PSE como algo prazeroso. Com base nessa inversão de valores dos participantes do presente estudo torna-se nítido o papel de cada instrumento, pois, a associação da FS com PSE depende do nível de

atividade física da amostra estudada.

Os dados do presente estudo se assemelham ao obtido por Cole et al.¹⁰, cujo compararam as percepções de dificuldade de sujeitos com sobrepeso/obesidade com eutróficos antes de realizar um *Sprint*. No caso do estudo de Cole et al.¹⁰, enquanto sujeitos eutróficos percebiam o exercício mais fácil, as percepções de sujeitos com sobrepeso/obesidade eram dependentes de fatores motivacionais, e por isso apresentavam maior heterogeneidade. No presente estudo, com um desenho experimental similar ao de Cole et al.¹⁰, sujeitos com a composição corporal mais próxima da obesidade apresentaram maior heterogeneidade nas respostas afetivas ao exercício; sendo dependentes da capacidade cardíaca e da idade para prever maior prazer.

Por meio do algoritmo de árvore decisional pôde-se avaliar não somente a consistência da TCR, como também o grau de importância de cada modulador da percepção avaliada. Ou seja, apesar de existirem estudos na literatura^{8,10-12} que defendem que a quantidade de gordura corporal é capaz determinar a aversão pelo exercício, os dados do presente estudo evidenciam que a relação entre composição corporal e percepções do exercício podem ser mais complexas, envolvendo mais outros fatores como massa magra e capacidade cardíaca do que a própria obesidade.

Com isso, acredita-se também que possa existir uma hipótese alternativa, ou seja, a possibilidade das associações entre sobrepeso/obesidade e características aversivas encontradas em outros estudos serem devido a uma consequência de outras características das respectivas amostras estudadas, ou seja, não é que o sobrepeso/obesidade sejam fatores determinantes de percepções aversivas ao exercício, mas sim que são uma consequência dos verdadeiros fatores determinantes, como pouca massa magra e baixa capacidade cardíaca, resultando na aversão por esforços físicos que vão além das capacidades físicas do sujeito.

Além de questões de natureza psicobiológicas, acredita-se também, com base na influência da idade nas percepções de alguns sujeitos, que haja influência de fatores sociais nas percepções pelo exercício. Por exemplo, de acordo a literatura²⁸, sujeitos com características de idades mais avançadas apresentam motivações com foco em manutenção da saúde para se exercitarem e maior aderência pelo exercício, enquanto idades mais joviais apresentam motivações com foco em estética para se exercitarem e menor aderência pelo exercício.

O presente estudo teve algumas limitações, sendo uma delas o tamanho amostral, pois, apesar do algoritmo de árvore decisional ter apresentado bom desempenho, o *n* era baixo para caracterizar os sujeitos pelo NAF-AUTO (Figura 2). Contudo, mesmo com resultados sem apresentar significância estatística ($p > 0.05$) e com um poder estatístico de 60%, acredita-se

que com um maior tamanho amostral, alguns dos resultados não significantes poderiam vir a tornar-se significantes. Pois, a diferença na previsão de afeto observada entre indivíduos insuficientemente e muito ativos tiveram um valor de p baixo mesmo não sendo significantes, e foi-se possível observar uma diferença média considerável entre estes grupos mesmo com uma amostra considerada pequena, e além disso as características dos sujeitos do presente estudo vão ao encontro da literatura. Outra limitação foi não utilizar um instrumento validado para a avaliação do nível de atividade física, contudo, para o presente estudo, acredita-se que o NAF-AUTO não apresente diferença significativa do nível de atividade física real pelos seguintes motivos: 1) a FCr corresponde ao nível de atividade física reportado pelos participantes, seguindo o conceito da literatura de que quanto mais ativo mais a FCr tende a ser menor²⁶, em outras palavras, os que se denominaram muito ativos apresentaram a menor FCr mediana (68 [50-79]), enquanto os demais apresentaram maior FCr, e mesmo com os moderadamente ativos apresentando FCr (78 [56-114]) ligeiramente maior que os insuficientemente ativos (76 [51-95]), os moderadamente ativos também apresentam maior intervalo entre os quartis, indicando maior variabilidade; 2) a amostra foi composta por estudantes de graduação em Educação Física que estivessem pelo menos no segundo ano do curso, garantindo que já tenham cursado a disciplina que conceitua nível de atividade física.

Assim, pelo fato da complexidade das respostas psicológicas como afeto e PSE, recomenda-se para futuros estudos análises de dados com aplicações de *Machine Learning*, como também que sejam exploradas as influências de fatores sociais nesses fenômenos observados.

Concluimos que os algoritmos de árvores decisionais foram capazes de estabelecer regras que determinam como as previsões de afeto e PSE são moduladas. Sendo que, o perfil que previu o maior desprazer apresentava composição corporal e capacidade cardíaca mais próximas de padrões prejudiciais a saúde, enquanto o que previu maior prazer tinha características opostas.

No caso da PSE, apesar de existirem dois perfis iniciais divididos pelo %G, sendo que, o com menor valor previu menor esforço, o perfil com maior %G possui maior heterogeneidade nas características.

Referências

1. Malta DC, França E, Abreu DMX, Perillo RD, Salmen MC, Teixeira RA, et al. Mortality due to noncommunicable diseases in Brazil, 1990 to 2015, according to estimates from the Global Burden of Disease study. *Sao Paulo Med J* [Internet]. 2017 Jun [cited 2018 Feb 22];135(3):213–21. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28746656>
2. Wajner A, Zuchinali P, Olsen V, Polanczyk CA, Rohde LE. Causes and Predictors of In-Hospital Mortality in Patients Admitted with or for Heart Failure at a Tertiary Hospital in Brazil. *Arq Bras Cardiol* [Internet]. 2017 Oct [cited 2018 Feb 22];109(4):321–30. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28977049>
3. Watte G, Fernandes Manenti ER, Anselmo Olinto MT, Henn RL, Vieira Paniz VM, Matata BM, et al. The Impact of Diabetes on Mortality Among Elderly Patients Admitted for Treatment at a Hospital for Cardiovascular Disease in Southern Brazil. *Curr Diabetes Rev* [Internet]. 2016 [cited 2018 Feb 22];12(4):454–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26206093>
4. David CN, de Mello RB, Bruscatto NM, Moriguchi EH. Overweight and abdominal obesity association with all-cause and cardiovascular mortality in the elderly aged 80 and over: A cohort study. *J Nutr Health Aging* [Internet]. 2017 May 8 [cited 2018 Feb 22];21(5):597–603. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28448093>
5. Garber CE, Blissmer B, Deschenes MR, Franklin BA, Lamonte MJ, Lee I-M, et al. Quantity and Quality of Exercise for Developing and Maintaining Cardiorespiratory, Musculoskeletal, and Neuromotor Fitness in Apparently Healthy Adults. *Med Sci Sport Exerc* [Internet]. 2011 Jul [cited 2018 Feb 7];43(7):1334–59. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21694556>
6. IBGE. Práticas de Esporte e Atividade Física em 2015 [Internet]. Rio de Janeiro; 2017 [cited 2018 Feb 19]. Available from: <https://biblioteca.ibge.gov.br/visualizacao/livros/liv100364.pdf>

7. WHO. Physical activity [Internet]. WHO. World Health Organization; 2018 [cited 2018 Feb 22]. Available from: <http://www.who.int/mediacentre/factsheets/fs385/en/>
8. Ekkekakis P, Vazou S, Bixby W, Georgiadis E. The mysterious case of the public health guideline that is (almost) entirely ignored: Call for a research agenda on the causes of the extreme avoidance of physical activity in obesity. *Obes Rev*. 2016;17(4):313–29.
9. Ekkekakis P, Lind E, Vazou S. Affective Responses to Increasing Levels of Exercise Intensity in Normal-weight, Overweight, and Obese Middle-aged Women. *Obesity* [Internet]. 2010;18(1):79–85. Available from: <http://doi.wiley.com/10.1038/oby.2009.204>
10. Cole S, Balcetis E, Zhang S. Visual perception and regulatory conflict: Motivation and physiology influence distance perception. *J Exp Psychol Gen* [Internet]. 2013;142(1):18–22. Available from: <http://doi.apa.org/getdoi.cfm?doi=10.1037/a0027882>
11. Yáñez-Sepúlveda R, Barraza-Gómez F, Martin EBS, Araneda OF, Zavala JP, Hecht GK, et al. Differences in energy expenditure, amount of physical activity and physical exertion level during a Zumba fitness class among adult women who are normal weight, overweight and obese. *J Sports Med Phys Fitness*. 2018 Jan 1;58(1–2):113–9.
12. Elsangedy HM, Nascimento PHD, Machado DGS, Krinski K, Hardcastle SJ, DaSilva SG. Poorer positive affect in response to self-paced exercise among the obese. *Physiol Behav* [Internet]. 2018 Feb 16 [cited 2018 Feb 22]; Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29458113>
13. Ladwig MA, Matthew A, Hartman M, Ekkekakis P. Affect-based exercise prescription. *ACSM's Heal Fit J*. 2017;21(5):10–5.
14. Baumeister RF, Heatherton TF, Tice DM. Losing Control: How and Why People Fail at Self-Regulation. In 1994.
15. Loewenstein G. Out of Control: Visceral Influences on Behavior. *Organ Behav Hum Decis Process* [Internet]. 1996;65(3):272–92. Available from:

<https://www.sciencedirect.com/science/article/pii/S074959789690028X>

16. Trope Y, Fishbach A. Counteractive self-control in overcoming temptation. Vol. 79, *Journal of Personality and Social Psychology*. US: American Psychological Association; 2000. p. 493–506.
17. Alvarez-Alvarado S, Chow GM, Gabana NT, Hickner RC, Tenenbaum G. Interplay between workload and functional perceptual-cognitive- affective responses: An inclusive model. *J Sport Exerc Psychol*. 2019 Apr 1;41(2):107–18.
18. Gurchiek RD, Rupasinghe Arachchige Don HS, Pelawa Watagoda LCR, McGinnis RS, Van Werkhoven H, Needle AR, et al. Sprint assessment using machine learning and a wearable accelerometer. *J Appl Biomech*. 2019;35(2):164–9.
19. Musa RM, Abdul Majeed APP, Taha Z, Chang SW, Nasir AFA, Abdullah MR. A machine learning approach of predicting high potential archers by means of physical fitness indicators. *PLoS One*. 2019 Jan 1;14(1).
20. Piercy KL, Troiano RP, Ballard RM, Carlson SA, Fulton JE, Galuska DA, et al. The physical activity guidelines for Americans. *JAMA - J Am Med Assoc* [Internet]. 2018 Nov 20 [cited 2020 Nov 21];320(19):2020–8. Available from: <https://jamanetwork.com/>
21. WORLD HEALTH ORGANIZATION. Waist circumference and waist-hip ratio: report of a WHO expert consultation. Geneva; 2008.
22. Borg Gunnar NB. Perceived Exertion. : *Exercise and Sport Sciences Reviews* [Internet]. *Exercise & Sport Sciences Reviews*: January 1974 - Volume 2 - Issue 1. 1974 [cited 2017 Jun 15]. p. 131–54. Available from: http://journals.lww.com/acsm-essr/Citation/1974/00020/Perceived_Exertion_.6.aspx
23. Hardy CJ, Rejeski WJ. Not What, but How One Feels: The Measurement of Affect during Exercise. *J Sport Exerc Psychol* [Internet]. 1989 Sep 1 [cited 2017 Jun 15];11(3):304–17. Available from: <http://journals.humankinetics.com/doi/10.1123/jsep.11.3.304>

24. Fiuza Regaçone S, Baptista de Lima DD, Engrácia Valenti V, Figueiredo Frizzo AC. Resting Heart Rate and Auditory Evoked Potential. *Biomed Res Int* [Internet]. 2015 [cited 2019 Feb 21];2015:1–6. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26504838>
25. Cohen J. Statistical Power Analysis for the Behavioral Sciences [Internet]. Statistical Power Analysis for the Behavioral Sciences. Routledge; 2013 [cited 2021 Feb 18]. Available from: <https://www.taylorfrancis.com/books/statistical-power-analysis-behavioral-sciences-jacob-cohen/10.4324/9780203771587>
26. Kokkinos P. Physical Activity and Cardiovascular Disease Prevention: Current Recommendations. *Angiology* [Internet]. 2008 Aug [cited 2017 Jun 30];59(2_suppl):26S-29S. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/18508850>
27. Frazão DT, de Farias Junior LF, Dantas TCB, Krinski K, Elsangedy HM, Prestes J, et al. Feeling of Pleasure to High-Intensity Interval Exercise Is Dependent of the Number of Work Bouts and Physical Activity Status. Piacentini MF, editor. *PLoS One* [Internet]. 2016 Mar 30 [cited 2017 Oct 21];11(3):e0152752. Available from: <http://dx.plos.org/10.1371/journal.pone.0152752>
28. Sperandei S, Vieira MC, Reis AC. Adherence to physical activity in an unsupervised setting: Explanatory variables for high attrition rates among fitness center members. *J Sci Med Sport*. 2016 Nov 1;19(11):916–20.