

THAMARA BARRETO PESSANHA DE SOUZA

**PROPRIEDADES PSICOMÉTRICAS DA
VERSÃO PORTUGUESA DA ESCALA DE FADIGA
DE PICHOT E BRUN: UM ESTUDO À LUZ DA
TEORIA DA RESPOSTA AO ITEM NÃO-
PARAMÉTRICA**

Orientador: Professor Doutor Pedro Joel Rosa

Universidade Lusófona de Humanidades e Tecnologias

Escola de Psicologia e Ciências da Vida

Lisboa

2022

THAMARA BARRETO PESSANHA DE SOUZA

**PROPRIEDADES PSICOMÉTRICAS DA
VERSÃO PORTUGUESA DA ESCALA DE FADIGA
DE PICHOT E BRUN: UM ESTUDO À LUZ DA
TEORIA DA RESPOSTA AO ITEM NÃO-
PARAMÉTRICA**

Dissertação defendida em provas públicas para a obtenção do Grau de Mestre em Psicologia Clínica e da Saúde, conferida pela Universidade Lusófona de Humanidades e Tecnologias, no dia 20 de julho de 2022, perante o júri, com o Despacho de Nomeação nº 202/2022, de 13 de maio de 2022, com a seguinte composição:

Presidente: Prof.^a Doutora Patrícia Pascoal

Arguente: Prof.^a. Doutora Carolina da Motta

Orientador: Prof. Doutor Pedro Rosa

Universidade Lusófona de Humanidades e Tecnologias

Escola de Psicologia e Ciências da Vida

Lisboa

2022

Agradecimentos

Foi um caminho longo e com alguns percalços, mas também de superação e aprendizagem, ao que devo um agradecimento especial aqueles que estiveram ao meu lado.

Primeiramente agradeço a Deus pelo sustento e força.

Ao meu marido pelo suporte incondicional e ao meu filho, ainda tão pequenino, por toda inspiração.

Aos meus pais e familiares, que estão tão longe fisicamente, mas não deixam de me apoiar.

Ao meu orientador Professor Doutor Pedro Joel Rosa, pela partilha de conhecimento, disponibilidade e compreensão.

A todos os participantes que responderam e partilharam o questionário, os quais foram fundamentais para a realização deste estudo.

Aos colegas de mestrado pela companhia e partilha de tantos momentos e aprendizagens.

Resumo

A fadiga pode ser definida como uma sensação subjetiva de cansaço físico e/ou mental que limita as atividades do cotidiano e diminuem a qualidade de vida, sendo causadora de grande sofrimento. É um sintoma/fenómeno que acomete grande parte da população e tem origem em diferentes causas. Dado o exposto, é importante estarmos munidos de instrumentos precisos e fiáveis que nos permitam avaliá-la, conduzindo os profissionais a uma intervenção eficaz no tratamento e controlo dos sintomas. Desta forma, este estudo teve como objetivo fazer uma Análise de Escala de Mokken (MSA), utilizando o Modelo de Homogeneidade Monótona (MHM) na Versão Portuguesa da Escala de Fadiga de Pichot e Brun (1984), a fim de avaliar as suas capacidades psicométricas. A amostra recolhida foi não probabilística (snowball) e ficou constituída por 945 participantes. Os resultados da MSA suportam fortemente a estrutura unidimensional da medida e a escalonabilidade dos itens. A análise também apresentou valores altos de fiabilidade. Para além disto, não foram identificados itens com funcionamento diferencial (DIF). Em suma, a versão portuguesa da Escala de Fadiga de Pichot e Brun (1984), apresentou-se como um instrumento fiável e seguro para avaliar a fadiga, atendendo a todos os pressupostos do MHM, indicando capacidades psicométricas sólidas.

Palavras-Chave: Fadiga; Mokken; MSA; TRI; TRIN.

Abstract

Fatigue can be defined as a subjective feeling of physical and/or mental tiredness that limits daily activities and reduces quality of life, causing great suffering. It is a symptom/phenomenon that affects a large part of the population and has its origin in different causes. Given the above, it is important to be equipped with accurate and reliable instruments that allow us to assess it, leading professionals to an effective intervention in the treatment and control of symptoms. Thus, this study aimed to perform a Mokken Scale Analysis (MSA), using the Monotone Homogeneity Model (MHM) in the Portuguese Version of the Pichot and Brun Fatigue Scale (1984), in order to assess their psychometric properties. The collected sample was non-probabilistic (snowball) and composed of 945 participants. The MSA results strongly support the one-dimensional structure of the measure and the scalability of the items. The analysis also showed high reliability values. In addition, items with differential functioning (DIF) were not identified. In short, the Portuguese version of Pichot and Brun's Fatigue Scale (1984) was presented as a reliable and safe instrument to assess fatigue, meeting all MHM assumptions, indicating solid psychometric capabilities.

Keywords: Fatigue; Mokken; MSA; TRI; TRIN.

Lista de abreviaturas e siglas

BFI	The Brief Fatigue Inventory
CEDIC	Comissão de Ética e Deontologia da Investigação Científica
CFS	Chalder Fatigue Scale
DIF	Funcionamento Diferencial do Item
EFPB	Échelle de Fatigue de Pichot e Brun - Escala de Fadiga de Pichot e Brun
EM	Esclerose Múltipla
FAS	Fatigue Assessment Scale - Escala de Avaliação da Fadiga
FSS	Fatigue Severity Scale - Escala de Severidade da Fadiga
LFS	Lee Fatigue Scale
MDM	Modelo de Dupla Monotonicidade
MFIS	Modified Fatigue Impact Scale - Escala de Fadiga Modificada
MHM	Modelo de homogeneidade monótona
MSA	Análise de Escala de Mokken
NIH	National Institute of Health
PFS	Piper Fatigue Scale
SFC	Síndrome da Fadiga Crónica
TCT	Teoria Clássica dos Testes
TRI	Teoria de Resposta ao Item
VAS-F	Visual Analogue Scale for Fatigue

Índice

Propriedades psicométricas da versão Portuguesa da Escala de Fadiga de Pichot e Brun: um estudo à luz da Teoria da Resposta ao Item não-paramétrica	8
Fadiga	8
Avaliação da fadiga	11
Objetivo	14
Método.....	14
Participantes.....	14
Instrumentos.....	14
Procedimentos.....	15
Preparação de dados e análise estatística.....	16
Resultados.....	18
Estatística descritiva e sensibilidade dos itens.....	18
Análise dos pressupostos para o Modelo de Homogeneidade Monótona	19
Unidimensionalidade.....	19
Monotonicidade.....	19
Independência Local	20
Coeficientes de escalonabilidade	21
Fiabilidade da escala EFPB.....	21
Funcionamento diferencial do item (DIF) entre homens e mulheres	21
Discussão.....	25
Conclusão	27
Referências	28
Anexo	i

Índice de tabelas e figuras

Tabela 1: Estatística descritiva dos itens e sensibilidade (N= 945).....	18
Tabela 2: Sumarização da verificação da monotonicidade.....	20
Tabela 3: Matriz da correlação residual entre os itens da EFPB.....	20
Tabela 4: Coeficientes de escalonabilidade.....	21
Tabela 5: Critérios utilizados para detetar o funcionamento diferencial do item entre homens e mulheres na EFPB.....	22
Tabela 6: Versão Portuguesa da escala de fadiga de Pichot e Brun.....	i
Figura 1: Gráfico da análise paralela.....	19
Figura 2: Curvas características dos itens com provável DIF.....	22
Figura 3: Gráfico do impacto de DIF individual cumulativo para a escala de Fadiga de Pichot e Brun.....	23
Figura 4: Impacto dos itens com DIF nas curvas características do teste.....	24
Figura 5: Traço latente entre os géneros.....	24

**Propriedades psicométricas da versão Portuguesa da Escala de Fadiga de Pichot e Brun:
um estudo à luz da Teoria da Resposta ao Item não-paramétrica
Fadiga**

A fadiga é uma experiência física e mental, que abarca consequências a nível do comprometimento funcional, diminuindo a qualidade de vida (Barsevick et al., 2001; Sharpe, 2002). Pode ser definida como uma sensação subjetiva de cansaço, exaustão e desgaste, que ocasiona falta de energia e limita o desempenho em atividades rotineiras (Berrios, 1990; Ream & Richardson, 1996; Williamson et al., 2005; Mota et al., 2005; Mota et al., 2012). Esta definição vai de encontro com a designação de Pichot e Brun (1984), que propõem que a fadiga é “uma sensação de desgaste físico e mental que geralmente ocorre após um esforço sustentado e que conseqüentemente requer um período de descanso”. Já para Aaronson et al., (1999, p. 46), a fadiga é a “consciência de uma diminuição da capacidade de atividade física e/ou mental devido a um desequilíbrio na disponibilidade, utilização e/ou restauração de recursos necessários para realizar a atividade”, se trata da existência de uma complexa interação entre processos biológicos, psicossociais e comportamentais, que afetam a maneira como uma pessoa interpreta ou reage à sensação de fadiga. A fadiga é, portanto, um fenómeno complexo com uma panóplia de definições, porém, em sua maioria, os principais atributos que lhe são designados referem-se ao cansaço, exaustão, desgaste, alteração da capacidade funcional e falta de energia (Mota et al., 2005). Atualmente, a definição de fadiga mais comumente aceite pela área da saúde, é a de um cansaço físico e mental desproporcional ao esforço realizado, e que permanece apesar das tentativas de recuperação de energia, comprometendo a execução das atividades diárias e diminuindo a capacidade de trabalho (Mota & Pimenta, 2002).

De acordo com Allen et al. (2008), a fadiga é vista como um constructo bidimensional: a fadiga física e a fadiga mental. A fadiga física corresponde a diminuição do desempenho muscular, no que diz respeito a redução da velocidade e relaxamento, devido à atividade exercida pelos músculos. Geralmente está associada ao consumo excessivo de energia, que gera a diminuição de hormônios, neurotransmissores e substratos essenciais para a função fisiológica (Aaronson et al., 1999), mas que é reversível após um período de descanso (Allen et al., 2008). Já a fadiga mental, envolve aspetos cognitivos, emocionais e motivacionais, expressando-se através da falta de energia, dificuldades de concentração e diminuição da atenção (Neu et al., 2010). À luz dos profissionais de saúde mental, a fadiga pode ser vista como um estado de cansaço resultante da redução de motivação (Lee et al., 1991), como resposta a exigências internas ou externas que excedem os recursos disponíveis (Aaronson et al., 1999). Apesar da

fadiga compreender duas dimensões, as mesmas não podem ser consideradas isoladamente, tendo em vista que se relacionam entre si. Na mesma linha de pensamento, Michielsen et al., (2004) defendem que os aspetos físicos e mentais da fadiga são difíceis de estar dissociados, uma vez que para a realização de tarefas que implicam esforço, os mesmos interagem entre si de forma complexa. Apesar de Chalder et al. (1993) defenderem a multidimensionalidade da fadiga, propondo as dimensões físicas e psicológicas, os mesmos defendem que a fadiga deve ser vista como uma dimensão contínua. Posição similar tem Piper (1998), que advoga que a fadiga deve ser considerada como um acontecimento multidimensional, tendo em consideração a interação entre os fatores físicos e mentais, resultando em uma experiência que engloba o indivíduo de modo geral. No entanto, Michielsen et al. (2004), destacam que existem desvantagens quanto a multidimensionalidade da fadiga, precisamente na sua medição, como por exemplo a construção de questionários extensos e a possibilidade de confusão entre fadiga e doença somática.

Alguns autores diferenciam a fadiga entre aguda/normal ou crónica/patológica (Piper, 1998; Mota & Pimenta, 2002). Esta distinção diz respeito a níveis temporais, nos quais a duração da fadiga implicará na sua evolução e possibilidade de tratamento. A fadiga aguda tem início e duração rápida e maioritariamente está relacionada com a atividade física ou situações cotidianas, como a privação do sono (Chisholm et al., 2001; Van Der Linden & Eling, 2005). Geralmente, esta fadiga acomete os indivíduos saudáveis, e é frequentemente reparada após repouso, dieta, exercício ou gestão de stresse, sendo considerada fadiga normal (Piper 1998). No entanto, a fadiga crónica tem uma duração prolongada, e é de difícil recuperação, persistindo por 6 meses ou mais (National Institute of Health: NIH, 2020). Este tipo de fadiga, que também pode ser considerada como patológica, limita a vida, saúde e felicidade do indivíduo, e necessita de tratamento especializado. A fadiga crónica pode até mesmo ser considerada, em alguns casos, como a Síndrome da Fadiga Crónica (SFC), para isto é necessário que a fadiga seja o sintoma central e persista por 6 meses ou mais, além de não estar associada a nenhuma outra doença (Fukuda et al., 1994; NIH, 2020). A SFC é caracterizada por uma severa fadiga incapacitante, que diminui a capacidade de concentração e memória, e provoca distúrbios do sono e dores músculo-esqueléticas. É um tipo de fadiga persistente e inexplicável, sendo de início recente, não resultante da realização de algum exercício e não podendo ser aliviada pelo descanso, o que por sua vez reduz os níveis precedentes de atividades ocupacionais, educacionais, sociais ou pessoais (Fukuda et al., 1994).

A fadiga é altamente prevalente em diversas populações (Mota et al., 2005; Jason et al., 2010; Williamson et al., 2005). De acordo com Williamson et al. (2005), 5 a 20% da população sofre de fadiga, sendo duas vezes mais provável que ocorra em mulheres do que em homens, acometendo indivíduos de todas as faixas etárias do desenvolvimento (Oliveira et al., 2010). Muitos autores apontam para a multicausalidade da fadiga, que pode ser de origem fisiológica, psicológica ou situacional (Williamson et al., 2005; Jason et al., 2010; Dukes et al., 2021). Alguns fatores predisponentes de fadiga dizem respeito a: “atividade física intensa, esforço prolongado ou excessivo, depressão, gestação e parto, presença de doenças e distúrbios do sono” (Mota et al. 2005, p 290), além de infecções, como na COVID 19, cuja complicação envolve o aparecimento de fadiga crónica na pós-recuperação (Islam et al., 2020). Em concordância, o NIH (2020) aponta para o surgimento da fadiga associado a tratamentos com antidepressivos e anti-histamínicos, procedimentos de quimioterapia e radioterapia, condições de saúde como a diabetes, doença cardíaca, doença da tireoide, anemia e apneia do sono. O mesmo instituto também refere que de igual modo a fadiga pode surgir em detrimento das emoções, isto porque os fatores físicos e psicológicos interagem entre si e se afetam mutuamente, como nos casos de ansiedade, depressão, stresse e luto, geralmente atrelada a sensação de falta de controlo sobre a própria vida, dificuldades em dormir, preocupações relacionadas a vida pessoal ou financeira e a percepção de bem-estar geral. A fadiga também está muitas vezes relacionada aos hábitos, dentre eles: praticar atividade física em excesso, dormir pouco, ingerir muita cafeína, consumir bebidas alcoólicas em demasia e se alimentar de comidas menos nutritivas com regularidade.

No que respeita à sintomatologia da fadiga, esta compreende, ao nível físico, a sensação de fraqueza com uma conseqüente diminuição do desempenho. Já a nível mental, a fadiga está intrincada com os aspetos emocionais e cognitivos, associados a diminuição da motivação, falta de energia, tristeza, dificuldade de concentração, e pensamento (Barsevick et al., 2001). Em concordância, Mota et al., (2015), defendem que comumente, a fadiga tem como conseqüência a diminuição da motivação e atenção, além da sensação de sono e letargia, com necessidade extrema de descanso (Mota et al. 2005). Ream & Richardson (1996), acrescentam que existem sintomas evidentes na fadiga, que são: uma sensação corporal total que variam do cansaço à exaustão, contendo aspetos físicos, cognitivos e emocionais, sendo uma experiência desagradável que causa angústia, assim como um evento crónico, cuja experiência é subjetiva e interfere na funcionalidade do indivíduo.

O facto de a fadiga ser encontrada em diferentes situações e patologias, desperta o interesse de profissionais de diferentes áreas que ao estudá-la dão ênfase ao que vai de encontro a sua atuação, fazendo com que a fadiga seja um conceito complexo e discordante (Mota & Pimenta, 2002). Borges et al., (2018) realçam que a grande quantidade de sinónimos atribuídos a fadiga, e que são utilizados de forma indiscriminada, tal como astenia, letargia, exaustão, sensação de fraqueza, cansaço extremo, e falta de motivação, contribuem para a divergência em sua definição. Apesar de não existir consenso quanto ao conceito de fadiga, os investigadores são unânimes na importância da sua avaliação, visto ser um fenómeno/sintoma encontrado num largo espectro de patologias orgânicas e mentais, muitas vezes identificado não só em situações normais (e.g. após exercício físico). Deste modo é crucial estarmos munidos de instrumentos que permitam avaliar este constructo complexo.

Avaliação da fadiga

A fadiga é difícil de definir e medir (Chalder, et al., 1993; Taylor et al., 2000; Dukes et al., 2021), porém altamente prevalente na população em geral (Williamson et al., 2005), e causa de grande sofrimento físico e mental. Desta maneira, destaca-se a necessidade de que seja avaliada de forma precisa e fiável, de modo que pautem uma intervenção eficaz no tratamento e controlo dos sintomas da fadiga (Mota et al., 2005; Alves, 2017; Dukes et al., 2021). Como a fadiga é subjetiva, é importante que o método utilizado para a avaliar seja de autorrelato, corroborando para uma avaliação feita sob a ótica do paciente e não do profissional (Mota & Pimenta, 2006).

Atualmente, não existe um instrumento-padrão para avaliar a fadiga. De acordo com Aaronson et al. (1999, p. 47), “nenhuma única medida de fadiga captura adequadamente a complexidade do fenómeno”. No entanto, são várias as medidas disponíveis que se propõem a medir a natureza, gravidade e impacto da fadiga em uma variedade de populações (Dittner et al., 2004), e em diferentes línguas, como a inglesa, maioritariamente. Alguns exemplos são: a *Visual Analogue Scale for Fatigue* (VAS-F), a *The Brief Fatigue Inventory* (BFI), a *Chalder Fatigue Scale* (CFS), a *Piper Fatigue Scale* (PFS), a *Lee Fatigue Scale* (LFS), a *Fatigue Severity Scale* (FSS), a *Modified Fatigue Impact Scale* (MFIS) a *Fatigue Assessment Scale* (FAS), a *Échelle de Fatigue de Pichot e Brun* (EFPB). Dentre estas, podemos encontrar algumas traduzidas para Português Europeu, validadas ou em processo de validação como: a FSS, a MFIS, a FAS e a EFPB.

A MFIS (Fisk et al., 1994), diferentemente da FSS, FAS e EFPB (que serão descritas a seguir), é uma escala multidimensional, que avalia a fadiga em aspetos físicos, psicológicos

e psicossociais, e se torna um pouco mais extensa que as demais, compreendendo 21 itens. Porém, em consonância com as outras escalas, também é de autorrelato. Esta é uma escala que tem o objetivo de medir e avaliar o impacto da fadiga na funcionalidade e qualidade de vida do indivíduo, comumente mais utilizada em pacientes com Esclerose Múltipla (EM). No estudo original a escala apresentou um α de Cronbach = .81. Na versão para a população portuguesa (Gomes, 2011), a Escala de Fadiga Modificada (designação portuguesa), apresentou apenas duas dimensões: cognitiva (11 itens) e física (10 itens), com valor geral de α de Cronbach = .96.

Em termos de unidimensionalidade tanto a FAS, quanto a FSS e a EFPB se igualam, assim como também são todas escalas breves e de autorrelato. A FAS (Michelsen et al., 2004), tem como objetivo avaliar a fadiga crônica. Possui 10 itens que são avaliados numa escala de tipo *Likert* que varia de 1 a 5. Quanto maior for a pontuação, maiores serão os níveis de fadiga. O α de Cronbach apresentado no estudo original foi de .90. Na versão para a população portuguesa (Alves, 2017), a Escala de Avaliação da Fadiga (designação portuguesa), foi adaptada focando a sua aplicação em homens e mulheres após o nascimento de um filho. Todos os itens da versão original foram mantidos. O α de Cronbach apresentado foi bom (.87).

A FSS (Krupp et al., 1989), avalia o impacto da fadiga em vários aspectos funcionais e comportamentais do cotidiano de pacientes com variadas condições de saúde. E assim como a MFIS foi originalmente desenvolvida para pacientes com EM, acrescentando também pacientes com Lúpus Eritematoso. É uma escala de autorrelato, unidimensional e composta por 9 itens, os quais tem resposta de tipo *Likert* que variam de 1 a 7, e o resultado total aponta a intensidade da fadiga. O instrumento apresenta uma alta consistência interna (α de Cronbach = .88). Na versão portuguesa (Gomes, 2011), a Escala de Severidade da Fadiga (designação portuguesa), apresenta 8 itens, pois o primeiro item da escala original foi excluído. O valor de alfa de Cronbach foi de .88, assim como na versão original.

Algo que difere a EFPB das demais escalas é o fato de a sua aplicação ser destinada a população geral, sem restrições a um contexto ou patologia específica. Esta escala também se destaca devido a seu preenchimento ser fácil e intuitivo, dado os seus 8 itens compreenderem valores de 0 a 4, que correspondem a “nada” e “extremamente”, respetivamente. A EFPB (Pichot e Brun, 1984), é um instrumento breve e de autorrelato que avalia a fadiga como um construto unidimensional, verificando a influência da fadiga na capacidade funcional do indivíduo (Barata, 2015). O valor final é obtido através da soma do valor das respostas, podendo variar entre 0 e 32, sendo considerado um nível de fadiga excessivo quando a pontuação é igual

ou superior a 22. Um estudo realizado por Brito (2020), teve como objetivo, através do estudo das características psicométricas, com base na Teoria Clássica dos Testes, validar a Escala de Fadiga de Pichot e Brun (1984), na sua versão traduzida por Rosa e Paiva (2012), para a população portuguesa. O seu estudo manteve os 8 itens da versão original da escala e apresentou um α de Cronbach considerado excelente ($\alpha = .93$). A EFPB é amplamente utilizada na língua inglesa e francesa, e embora ainda não validada em contexto português, a sua versão traduzida, tem apresentado em contexto de investigação evidências de fiabilidade. Em um estudo publicado por Rosa et al. (2015), a versão traduzida da EFPB, apresentou uma boa consistência interna com um α de Cronbach de .81. Em outros estudos teve valores de α de Cronbach entre .78 e .93 (Barata, 2015; Gonçalves, 2018; Neves, 2019; Brito, 2020; Caetano 2020).

A validação da EFPB justifica-se ao considerarmos que são poucos os instrumentos para o auxílio do diagnóstico da fadiga validados em Português Europeu, além de que, como pudemos observar, são, em sua maioria, designados a pacientes com alguma patologia específica, diferentemente da EFPB que tem objetivo de medir o nível geral de fadiga, indiferentemente de terem ou não uma patologia associada. Esta justificativa concerne para a relevância do presente estudo que tem como objetivo fazer uma Análise de Escala de Mokken (MSA) na Versão Portuguesa da Escala de Fadiga de Pichot & Brun (1984), complementando desta forma o estudo iniciado por Brito (2020) à luz Teoria Clássica dos Testes (TCT), fornecendo um instrumento que possa ser utilizado, de forma segura, pelos profissionais de saúde em contexto clínico ou de investigação.

A escolha da realização de uma análise psicométrica pautada na Teoria de Resposta ao Item (TRI) se dá pela maioria dos instrumentos validados em contexto português para o diagnóstico da fadiga terem sido validados à luz da TCT. Esta última tem como objetivo avaliar a resposta final de um indivíduo em um teste, isto porque de acordo com esta teoria, a probabilidade de um teste medir aquilo a que se propõe é avaliado pela soma dos itens (Pasquali, 2003). E embora seja a análise psicométrica mais utilizada, apresenta limitações que podem ser colmatadas através da TRI. Andrade et. al., (2000, p. 7), define a TRI como um “conjunto de modelos matemáticos que procura representar a probabilidade de um indivíduo dar uma certa resposta a um item como função dos parâmetros do item e da habilidade [ou habilidades] do respondente”. Em detrimento das limitações apresentadas pela TCT, a TRI apresenta algumas vantagens (e.g., Lopes et al., 2018). Pasquali e Primi (2003), apontam que diferentemente do que ocorre na psicométrica clássica, na TRI os instrumentos não são dependentes de seus itens, fazendo com que o resultado de um sujeito não dependa do teste que foi utilizado. Da mesma

maneira, as características de um item não dependem da amostra de sujeitos (Embretson & Reise, 2013). Na TRI o que importa é a análise de cada item e não a do teste como um todo (Andrade et al., 2000), além de que para avaliar a fidedignidade do instrumento, não é exigido formas rigorosamente paralelas (Pasquali & Primi, 2003).

Dentro da TRI, existe o modelo paramétrico e não-paramétrico, e estes podem ser representados, pela Análise de Rasch e pela MSA, respetivamente. No modelo paramétrico existe uma função algébrica bem definida, em que o traço latente do indivíduo e as características do item funcionam como parâmetro. Já na formulação não-paramétrica, não é imposto uma forma algébrica para a função de resposta ao item e não é exigido qualquer suposição a respeito da distribuição do traço latente, o que a torna menos conservadora, enquanto mantém as propriedades de medição, evitando com que os pesquisadores removam itens desnecessariamente (Aleo et al., 2019; Palmgren et al. 2018).

Objetivo

Pretendemos com este estudo fazer uma Análise de Escala de Mokken na Versão Portuguesa da Escala de Fadiga de Pichot e Brun (1984) analisando as qualidades psicométricas da escala, que possa ser utilizado, de forma segura, pelos profissionais de saúde em contexto clínico ou de investigação.

Método

Participantes

O presente estudo foi constituído por uma amostra inicial, não probabilística (snowball) de 1129 participantes. Os critérios de inclusão estabelecidos foram: a) possuir nacionalidade portuguesa b) ter idade igual ou superior a 18 anos e c) responder de forma individual. Foram excluídos 184 participantes que não concluíram o questionário, o que resultou em uma amostra final de 945 participantes, com idades compreendidas entre os 18 e 75 anos ($M=45.9$; $DP=10.36$), sendo a maioria do sexo feminino (72%, $n=684$). Relativamente à escolaridade, 74% ($n=699$) dos participantes tinham o ensino superior, 25.1% (237) tinham o ensino secundário, e 1% ($n=9$) o ensino básico. No que diz respeito à variável etnia, 97.1% ($n=918$) dos respondentes são de etnia caucasiana, 1.5% ($n=14$) de etnia negra, 0.4% ($n=4$) de etnia asiática, 0.2% ($n=2$) de etnia hispânica, e 0.7% ($n=7$) declararam pertencer a outra etnia.

Instrumentos

O formulário de avaliação utilizado no presente estudo incluiu um breve questionário sociodemográfico com o objetivo de caracterizar a população, e conferir a adequabilidade da

amostra aos critérios de inclusão. Este questionário é constituído pelos itens: sexo; idade; etnia; nacionalidade e escolaridade. De seguida, foi utilizada a versão portuguesa da Escala de Fadiga de Pichot e Brun descrita a seguir.

Escala de Fadiga de Pichot e Brun (EFP). A EFP (Pichot & Brun, 1984) adaptada por Rosa e Paiva (2012) é uma escala breve e de autorrelato, que tem o objetivo de medir o nível geral de fadiga. É constituída por 8 questões relativamente à fadiga sentida nos últimos meses. A resposta de cada um dos itens tem formato tipo *Likert* e variam entre 0 e 4, na qual o valor “0” corresponde a “nada”; “1” a “um pouco”; “2” a “moderadamente”; “3” a “muito” e “4” a “extremamente”. O valor final é obtido através da soma do valor das respostas, e varia entre 0 e 32, considerando um nível de fadiga excessivo quando a pontuação é igual ou superior a 22. É um instrumento que tem evidenciado boas qualidades psicométricas, com um alfa de Cronbach entre .78 e .93 (Brito, 2020; Caetano 2020; Gamito & Rosa, 2014; Gonçalves, 2018; Neves, 2019; Rosa et al., 2014, 2015, 2017, 2019).

Procedimentos

Primeiramente, foi solicitado o parecer para a realização do estudo à Comissão de Ética e Deontologia da Investigação Científica (CEDIC) da Escola de Psicologia e Ciências da Vida da Universidade Lusófona de Humanidades e Tecnologias, o qual foi aprovado. Trata-se de um estudo de natureza psicométrica e transversal.

O passo seguinte consistiu na recolha de dados *online* através da disponibilização de um *link* para um formulário via *Qualtrics*, que foi divulgado nas redes sociais e enviado por email., seguindo o método de snowball. Este formulário de avaliação esteve disponível de dezembro de 2018 até fevereiro de 2021.

O formulário disponibilizado iniciou com a apresentação dos critérios de inclusão, e a explicação, em linhas gerais, da natureza do estudo e de que forma o participante poderia colaborar. Em seguida, no mesmo formulário, foi apresentado o consentimento informado, garantindo o sigilo e a confidencialidade do participante. Posteriormente, um breve questionário sociodemográfico e a versão traduzida da Escala de Fadiga de Pichot e Brun foi exposta, para que os participantes pudessem preencher.

O protocolo de avaliação teve um tempo médio de 15 minutos para ser preenchido, porém cada participante teve o tempo livre para preenchimento. Todas as questões presentes no respetivo protocolo foram de natureza obrigatória, não havendo possibilidade de não resposta.

Preparação de dados e análise estatística

Primeiro, os dados recolhidos foram exportados do *Qualtrics* para o programa de tratamento e análise estatística IBM SPSS 25. Seguidamente, foi realizada uma análise descrita das respostas dadas através da média, desvio-padrão, mínimo-máximo e amplitude, assim como a curtose e assimetria dos itens. A sensibilidade dos itens foi verificada através dos valores de assimetria e curtose, que devem estar compreendidos entre -2 e 2 (Almeida & Freire, 2008). Não se verificaram efeitos chão e teto para os itens da escala de fadiga de Pichot- PT, isto é, percentagens de resposta nos extremos da escala de resposta abaixo dos 80% (Hilari, Byng, Lamping, & Smith, 2003).

A Análise de Mokken, que é uma análise não-paramétrica baseada na TRI, pode ser utilizada a partir de dois modelos, o modelo de homogeneidade monótona (MHM), que tem como pressupostos a unidimensionalidade, independência local e monotonicidade, e o modelo de dupla monotonicidade (MDM), que assume, além dos três pressupostos anteriormente referidos, um quarto pressuposto, na qual as funções de resposta ao item não se devem cruzar. Para o presente estudo assumimos o MHM, tendo em vista ser um modelo menos restritivo e suficiente para muitas aplicações da TRI não-paramétrica (Stoch et al., 2012).

A unidimensionalidade é um pressuposto que assume que todos os itens que formam uma escala medem o mesmo traço latente. Esta propriedade de medida é importante tendo em consideração que favorece a interpretação das respostas aos itens e garante que a pontuação total dos itens expressa as mesmas características (Drasgow & Parsons, 1983). Para avaliarmos a unidimensionalidade da escala realizámos uma análise paralela, na qual foi considerado como critério para retenção de fatores *eigenvalues* observados superiores a 1, que fossem, ao mesmo tempo, superiores ao seu respetivo *eigenvalue* estimado para o percentil 95% (Crawford et al., 2010).

Já o pressuposto de independência local assume que a probabilidade de um indivíduo responder a um item em uma escala não é influenciada ou afetada por qualquer outro item, dependendo unicamente do seu nível no traço latente (Friedrich et al., 2015). Para testar a independência local, utilizamos a estatística Q3 de Yen (1984), que teve o objetivo de calcular a correlação entre os resíduos para cada par de itens, na qual valores de Q3 $>.36$ são considerados como dependentes locais. Contudo é importante perceber que a independência local total é praticamente inatingível (Stoch et al., 2012), pois se um item variar no mesmo sentido de outro item, significa que estão relacionados, não havendo independência local, e esta

variação no mesmo sentido é expectável ao se tratar de itens que fazem parte de um mesmo instrumento.

A monotonicidade é o terceiro pressuposto e supõe que para cada item a probabilidade condicional de acerto seja não decrescente em função do traço latente. A avaliação da monotonicidade é dada por meio do número de violações da premissa. Na existência de alguma violação, é importante que seja verificada a sua gravidade considerando a estatística crítica (crit). Itens que apresentem valores crit abaixo de .40 não violam gravemente os critérios de monotonicidade e podem, desta forma, serem incluídos na escala de Mokken (Molenaar & Sijtsma, 2000).

Posteriormente foram calculados os coeficientes de escalonabilidade com o objetivo de avaliarmos a homogeneidade da escala. Deste modo, consideramos os índices de escalonabilidade H de Loevinger para cada par de itens (H_{ij}), para um único item (H_i) e para o conjunto completo dos itens (H). Para efeitos de interpretação, os valores aceitáveis para os coeficientes de escalonabilidade são os acima de 0.30 (Van der Ark, 2007; van der Linden, 2016). De acordo com Sijtsma e Molenaar (2002), um valor de $H < 0.30$ implica que os itens não podem ser escalados, $0.30 \leq H < 0.40$ indica que a escala é fraca, $0.40 \leq H < 0.50$ a escala é média, e quando $H \geq 0.50$, a escala é forte. A delimitação de 0.30 para H foi inicialmente proposto por Mokken (1971), tendo em vista que valores baixos de H , mesmo que positivos, não seriam suficientes para verificarem o modelo de homogeneidade monótona, assim, também, como poderiam contribuir para um ordenamento inadequado ou incorreto dos respondentes.

Por fim, examinamos o funcionamento diferencial do item (DIF), com o objetivo de avaliar a invariância da EFPB. Desta forma testamos as diferenças entre o género (masculino e feminino) através de uma estrutura híbrida e interativa utilizando a Regressão Logística e a Teoria de Resposta ao Item. Este procedimento incluiu três modelos logísticos ordinais, no qual o modelo 1 avaliou a pontuação no traço latente, o modelo 2 avaliou a score do traço latente e a associação ao grupo, e o modelo 3 avaliou a interação entre o score do traço latente e a associação ao grupo. Para detetarmos a presença do DIF utilizamos como critério a razão de verossimilhança do teste χ^2 com uma taxa de erro assumida de α de .01. Havendo diferença significativa dos valores entre os modelos 1 e 2, assumimos um DIF uniforme, enquanto uma diferença significativa entre os modelos 3 e 2 indica DIF não uniforme. Já para a magnitude do DIF nos baseamos no Pseudo- R^2 de McFadden, no qual consideramos valores $< .13$ como insignificante, $.13 < R^2 < .26$ como moderado, $> .26$ como grande (Zumbo, 1999). Ainda foram

consideradas as diferenças entre os β s resultantes dos modelos 1 e 2 da regressão ($\Delta\beta_1$), no qual assumimos valores $<.1$ como indicativo de ausência de DIF (Crane et al., 2004).

A Análise de Mokken foi realizada por meio do software R versão 4.0.5, com os packages *irtGUI*, *IRTShiny* e *irtreliability* (unidimensionalidade e independência local), *mokken* (unidimensionalidade e monotonicidade) e *lordif* (funcionamento diferencial dos itens).

Resultados

Estatística descritiva e sensibilidade dos itens

A média, desvio-padrão, mínimo-máximo e amplitude, assim como a análise da curtose e assimetria dos itens estão detalhados na Tabela 1.

Tabela 1

Estatística descritiva dos itens e sensibilidade (N= 945)

EFPB	<i>M (DP)</i>	Mín-Máx (Amp)	Efeito- chão (%)	Efeito-teto (%)	Assimetria	Curtose
Item 1	1. (1.05)	0-4 (4)	14.3	4.9	0.153	-0.479
Item 2	1.59 (1.03)	0-4 (4)	13.0	5.1	0.458	-0.227
Item 3	1.53 (1.14)	0-4 (4)	20.7	5.0	0.341	-0.733
Item 4	1.32 (1.21)	0-4 (4)	32.3	4.8	0.548	-0.785
Item 5	1.37 (1.19)	0-4 (4)	30.2	4.6	0.447	-0.822
Item 6	1.83 (1.30)	0-4 (4)	18.9	12.2	0.135	-1.122
Item 7	1.58 (1.16)	0-4 (4)	18.7	6.8	0.405	-0.711
Item 8	1.49 (1.24)	0-4 (4)	25.9	7.4	0.454	-0.842

A maior classificação média ($M=1.83$), foi apresentada pelo item 6 “Tenho vontade de me deitar e descansar”, tendo igualmente apresentado a maior dispersão de valores ($DP=1.30$).

Já o item 4 “Sinto as minhas pernas e os meus braços exaustos” apresentou a menor classificação média ($M=1.32$). E o item 2 “Parece que é necessário esforço em tudo que faço” apresentou a menor dispersão de valores ($DP=1.03$). A adequada sensibilidade dos itens pode ser conferida através dos valores de assimetria e curtose que se mantiveram entre -2 e 2 (Almeida & Freire, 2008). Também não foi verificado efeito-chão ou efeito-teto em nenhum dos itens, tendo em vista que todos apresentaram valores menores que 80% (Hilari, Byng, Lamping, & Smith, 2003).

Análise dos pressupostos para o Modelo de Homogeneidade Monótona

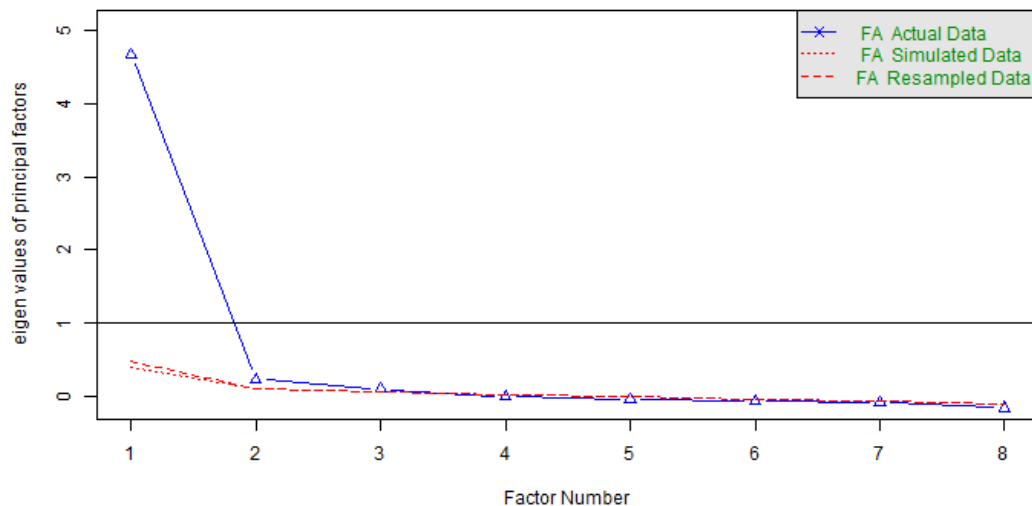
A qualidade do ajuste do MHM depende do grau com que os dados aderem aos pressupostos, pelo que a unidimensionalidade, independência local e monotonicidade da EFPB foram primeiramente avaliadas.

Unidimensionalidade

A unidimensionalidade da EFPB foi evidenciada através da análise paralela, que revelou a presença de um único fator latente. Ao observarmos a figura 1 verificamos que apenas um *eigenvalue* observado foi maior que 1 e, ao mesmo tempo, superior ao *eigenvalue* simulado e remostrado.

Figura 1

Gráfico da análise paralela



Nota. É considerada a presença de um fator latente sempre que o *eigenvalue* observado (a azul) for superior ao *eigenvalue* simulado e reamostrado (a vermelho).

Monotonicidade

Verificamos que o item 1 apresentava 4 violações ($\#v_i = 4$), porém a análise do valor da estatística crítica (valor crit = 11), indicou que embora existissem violações na

monotonicidade, estas não eram graves ($\text{crit} < 40$). Sendo assim a monotonicidade da escala foi conferida (Tabela 2).

Tabela 2

Sumarização da verificação da monotonicidade

EFPB	#ac	#vi	#zsig	crit
Item 1	105	4	0	11
Item 2	73	0	0	0
Item 3	100	0	0	0
Item 4	56	0	0	0
Item 5	72	0	0	0
Item 6	92	0	0	0
Item 7	73	0	0	0
Item 8	58	0	0	0

Nota. #ac = número de pares positivos; #vi= número de violações entre itens; #zsig = violações significativas; crit = valor crit.

Independência Local

O item 4 e 8 apresentaram uma correlação através de um valor de $Q3 = .37$. Este valor está muito próximo do limite ($.36$) (Yen,1984), e, tendo em consideração que, de fato é normal que os itens partilhem variância entre si, pois fazem parte de um mesmo instrumento, optou-se por manter os itens na escala. Os itens restantes da EFPB mostraram independência local, apresentando valores abaixo de $.36$.

Tabela 3

Matriz da correlação residual entre os itens da EFPB

	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8
Item 1	-							
Item 2	0.07	-						
Item 3	0.02	0.09	-					
Item 4	-0.07	-0.12	0.29	-				
Item 5	-0.14	-0.22	-0.19	-0.07	-			
Item 6	-0.14	-0.10	-0.17	-0.17	-0.19	-		
Item 7	-0.13	-0.08	-0.27	-0.27	-0.07	-0.11	-	
Item 8	-0.19	-0.20	-0.31	-0.37	-0.16	-0.18	-0.03	-

Coefficientes de escalonabilidade

Apontando para a homogeneidade, a EFPB apresentou um forte índice de escalonabilidade ($H=0.62$) (Van der Ark, 2007). Os coeficientes de escalonabilidade inter-item (H_{ij}) variaram de 0.49 (item 3 e 7) a 0.78 (item 6 e 8). Já os coeficientes de escalonabilidade dos itens (H_i) variaram de 0.54 (item 1) a 0.70 (item 8). Todos os coeficientes foram significativamente maiores do que o valor limite inferior convencional de 0.30.

Tabela 4

Coefficientes de escalonabilidade

	H_i	H_{ij}							
		Item	Item	Item	Item	Item	Item	Item	Item
		1	2	3	4	5	6	7	8
Item 1	0.54	-							
Item 2	0.61	0.58	-						
Item 3	0.60	0.55	0.57	-					
Item 4	0.63	0.53	0.60	0.76	-				
Item 5	0.63	0.52	0.59	0.58	0.64	-			
Item 6	0.65	0.53	0.65	0.60	0.64	0.68	-		
Item 7	0.60	0.50	0.60	0.49	0.54	0.65	0.65	-	
Item 8	0.70	0.59	0.70	0.64	0.67	0.75	0.78	0.73	-
Escala									0.62
(H)									

Nota. H_i = escalabilidade do item; H_{ij} = escalonabilidade inter-item.

Fiabilidade da escala EFPB

Relativamente ao índice de consistência interna da medida, pudemos verificar através do valor obtido do coeficiente de fiabilidade Rho de Mokken (.92), um elevado nível de fiabilidade.

Funcionamento diferencial do item (DIF) entre homens e mulheres

Através da razão da verossimilhança foram identificados dois itens com provável DIF (p -value <.01), sendo um item para DIF uniforme (item 1) e outro item para DIF não uniforme (item 8) (Figura 2). No entanto, ao avaliarmos a magnitude dos efeitos das regressões, através da estatística Pseudo- R^2 de *McFadden*, pudemos verificar que os valores encontrados foram

insignificantes, pois permaneceram abaixo de 0.13 (Zumbo, 1999), variando de $R^2_{(2;3)}=0.0005$ a $R^2_{(1;2)}=0.0066$. As diferenças entre os β s resultantes dos modelos 1 e 2 da regressão ($\Delta\beta_1$), também foram baixos para os dois itens ($< .1$) (Crane et al., 2004), pelo que o item 1 apresentou $\Delta\beta_1 = 0.0171$ e o item 8 $\Delta\beta_1 = 0.0416$.

Tabela 5

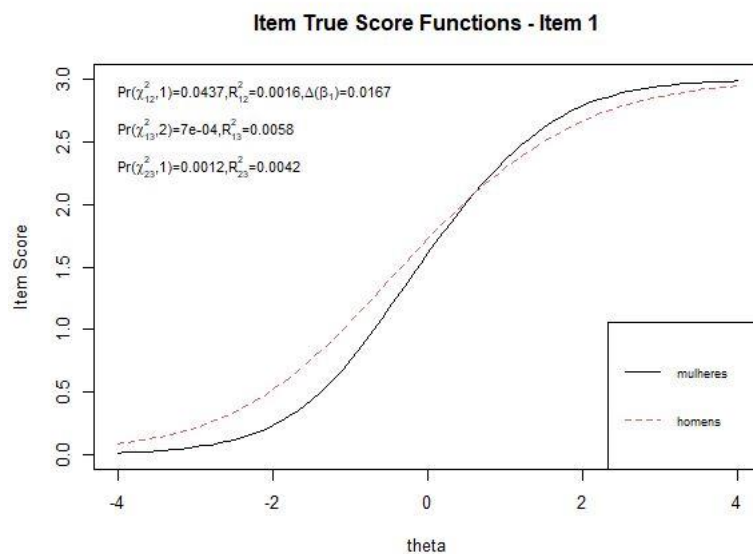
Critérios utilizados para detetar o funcionamento diferencial do item entre homens e mulheres na EFPB

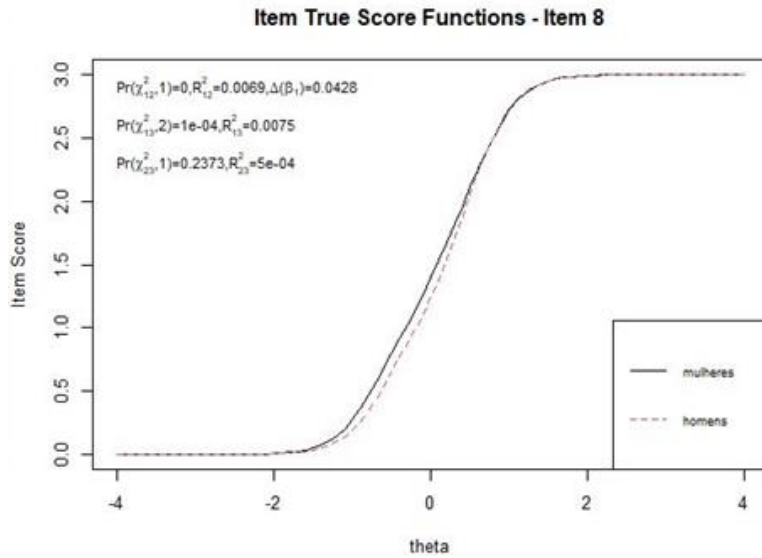
	Testes de Razão de Verossimilhança (p)			Diferença entre os Pseudo-R ²			$\Delta(\beta_1)$
	$Pr(\chi^2_{(1;2)})$	$Pr(\chi^2_{(1;3)})$	$Pr(\chi^2_{(2;3)})$	$R^2_{(1;2)}$	$R^2_{(1;3)}$	$R^2_{(2;3)}$	
Item 1	0.0397	0.0004	0.0007	0.0017	0.0063	0.0046	0.0171
Item 8	0.0000	0.0001	0.2636	0.0066	0.0071	0.0005	0.0416

Nota. $Pr(\chi^2_{(1;2)})$ = comparação da razão de verossimilhança entre os passos um e dois da regressão. $Pr(\chi^2_{(1;3)})$ = comparação da razão de verossimilhança entre os passos um e três da regressão. $Pr(\chi^2_{(2;3)})$ = comparação da razão de verossimilhança entre os passos dois e três da regressão. $R^2_{(1;2)}$ = diferença entre os Pseudo-R² dos passos um e dois da regressão. $R^2_{(1;3)}$ = diferença entre os Pseudo-R² dos passos um e três da regressão. $R^2_{(2;3)}$ = diferença entre os Pseudo-R² dos passos dois e três da regressão. $\Delta(\beta_1)$ = diferenças entre os β s resultantes dos modelos 1 e 2 da regressão.

Figura 2

Curvas características dos itens com provável DIF

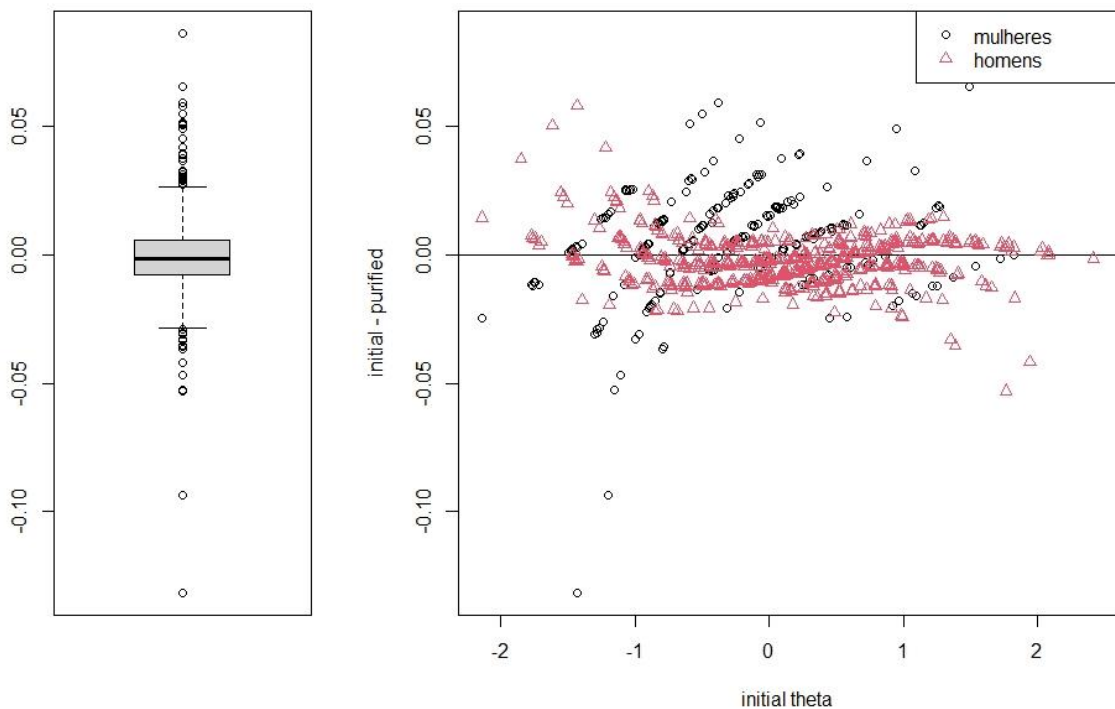




Nota. O quadro superior ilustra a curva característica do item 1 (DIF uniforme) e o quadro inferior ilustra a curva característica do item 8 (DIF não uniforme).

Figura 3

Gráfico do impacto de DIF individual cumulativo para a escala de Fadiga de Pichot e Brun



Nota. À esquerda, o gráfico mostra as diferenças entre as pontuações sem DIF e as pontuações com DIF. À direita, o gráfico de dispersão mostra as pontuações da diferença (inicial-purificada), em função das pontuações iniciais sem DIF (teta inicial).

Figura 4

Impacto dos itens com DIF nas curvas características do teste

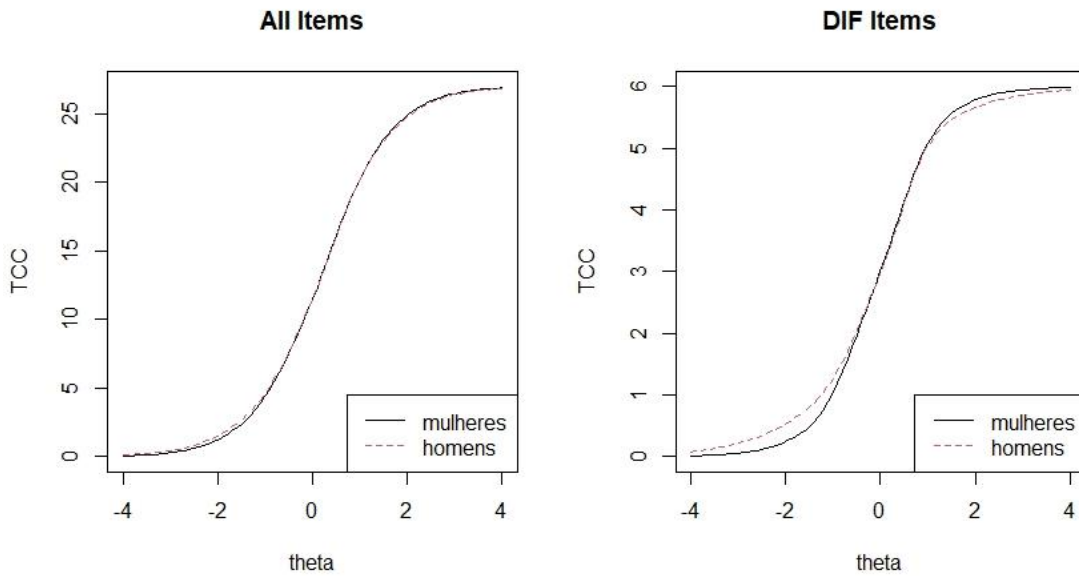
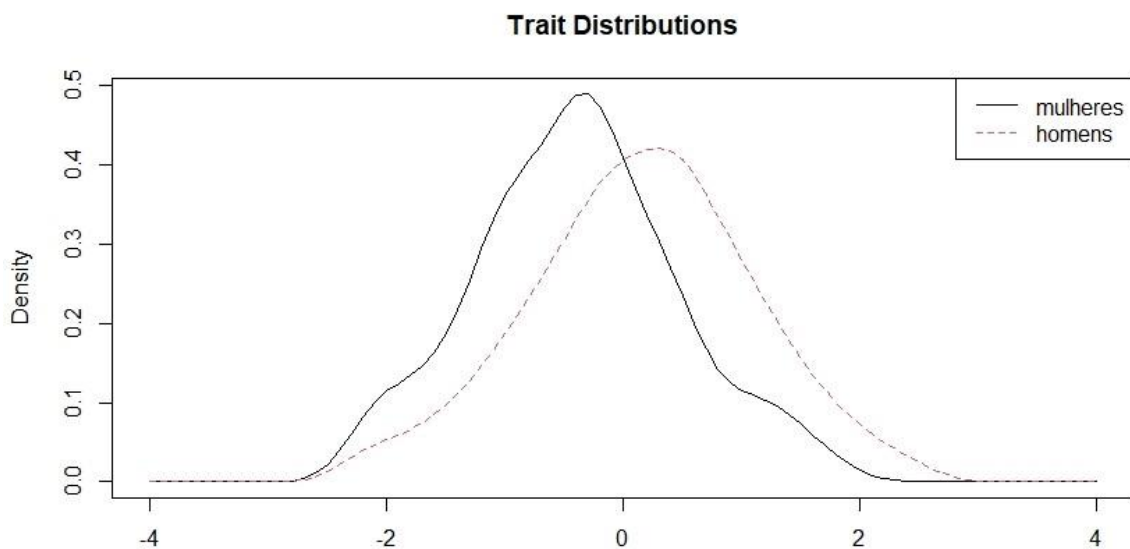


Figura 5

Traço latente entre os géneros



Nota. A figura ilustra a média de homens e mulheres no traço latente.

Desta forma, considerando os três critérios avaliados e que podem ser visualizados na Tabela 5 e a análise gráfica, podemos constatar a ausência de DIF entre homens e mulheres na EFPB.

Discussão

A fadiga atinge uma variedade de indivíduos e pode ter origem a partir de fatores físicos e/ou psicológicos, e situacionais (Mota et al. 2005; NIH, 2020). A sua presença causa muitas vezes limitações ao cotidiano do indivíduo, afetando a sua vida, saúde e felicidade, o que implica na necessidade de um tratamento. Desta forma, se faz necessário um instrumento fiável que avalie a fadiga, auxiliando a conduta dos profissionais frente ao controlo de sintomas (Mota et al., 2005; Alves, 2017; Dukes et al., 2021).

Em contexto português ainda são poucos os instrumentos validados que avaliam a fadiga, e destes a maioria é designada a uma patologia específica. Já a EFPB é um instrumento que se propõe a avaliar um nível de fadiga geral, de forma breve, fácil e intuitiva. Esta é uma medida amplamente utilizada em inglês e francês, mas a sua versão traduzida para português europeu vem apresentando, em contexto de investigação, evidências de fiabilidade (Rosa et al., 2015).

A fim de complementar o estudo realizado por Brito (2020), em que foi realizada uma análise psicométrica da EFPB baseada na TCT, este estudo teve como objetivo realizar uma análise baseada na TRI não-paramétrica, mais precisamente através do MHM da MSA. No qual os resultados revelaram que a versão portuguesa da EFPB possui propriedades psicométricas sólidas. A sua unidimensionalidade foi confirmada, assim como a sua monotonicidade e independência local. Também não foram detetados funcionamento diferencial dos itens entre homens e mulheres.

A confirmação da unidimensionalidade da escala, torna os seus itens seguros para a deteção de fadiga na população portuguesa. Através da MSA, pudemos confirmar de forma clara a estrutura da escala que apresentou apenas um fator latente, justificando a soma da pontuação dos seus oito itens como indicador de fadiga. A escalonabilidade, que indica a homogeneidade dos itens, obteve ótimos valores, tendo sido todos os coeficientes significativamente maiores do que o valor limite inferior convencional de 0.30, indicando precisão na discriminação da variável latente em estudo. Isto vai de encontro ao estudo de Brito (2020) baseado na TCT, no qual utilizou a análise fatorial exploratória e a análise confirmatória, que apontaram para uma medida unifatorial. Este resultado também aponta para a mesma estrutura fatorial da versão original de Pichot e Brun (1984).

Através da verificação da monotonicidade a medida revelou que a probabilidade de acerto de um item aumentou de forma correspondente ao aumento da fadiga, não havendo nenhuma violação significativa. Já com relação a independência local da medida, conclui-se que apesar de termos encontrado correlação entre dois itens, esta era ligeira e refletia a possibilidade de itens de um mesmo instrumento partilharem alguma variância (Stoch et al., 2012). Desta forma, optamos por manter os dois itens na escala e confirmamos a independência local da medida, indicando que a possibilidade de resposta de um indivíduo a um item não é influenciada por outro item, e sim pelo traço latente, neste caso a fadiga (Friedrich et al., 2015). A escala apresentou uma consistência interna excelente, tendo apresentado um valor de *Rho* de Mokken de .92, evidenciando fiabilidade (Van der Ark, 2007; van der Linden, 2016).

A análise de DIF realizada em função de género, revelou que o funcionamento dos itens da EFPB foi igual em ambos os sexos. O que implica que os itens são confiáveis para avaliar a fadiga independente do género do indivíduo que respondeu, sendo as respostas apenas como competência do indivíduo, medida através da dificuldade dos itens e não oriunda da classe de género a que pertence. A análise gráfica foi concordante com a ausência de DIF, tendo ilustrado pequenas sobreposições entre os géneros, por vezes insignificantes. Esta verificação supre uma limitação indicada no estudo de Brito (2020), o qual refletiu sobre se uma amostra composta por maioritariamente mulheres não enviesaria os resultados, tendo em vista que a fadiga pode ser sentida de forma diferente entre homens e mulheres, e a sua amostra era mais composta por mulheres.

Apesar do estudo ter tido resultados satisfatórios que indicam a sua utilização em contexto português de forma segura, podemos verificar algumas limitações, a exemplo da amostra e do modelo utilizado. A amostra foi não probabilística, o que significa que não é representativa da população portuguesa, isto é, os resultados não podem ser generalizados à população portuguesa. Desta forma é importante pensar em estudos futuros com amostras probabilísticas, o qual permita uma seleção aleatória dos participantes, representando bem a população de interesse. Já com relação ao modelo utilizado neste estudo (MHM), apesar de a princípio ter ido de encontro ao que pretendíamos, se trata de um modelo menos rígido. Estudos futuros poderiam utilizar o modelo de dupla monotonicidade, que para além dos três pressupostos assumidos no MHM, assume um quarto pressuposto, que diz respeito ao não cruzamento de funções de resposta ao item, complementando as análises já realizadas neste estudo.

Conclusão

Este estudo foi o primeiro que abordou a estrutura dimensional da EFPB com a MSA e confirmou que a medida cumpriu com todos os pressupostos do MHM, o que a constitui como uma medida com propriedades psicométricas sólidas, sendo valiosa no contexto português para a avaliação de fadiga a nível clínico ou de investigação.

Referências

- Aaronson, L. S., Teel, C. S., Cassmeyer, V., Neuberger, G. B., Pallikkathayil, L., Pierce, J., Press, A. N., Williams, P. D., & Wingate, A. (1999). Defining and Measuring Fatigue. *Image: The Journal of Nursing Scholarship*, 31(1), 45–50. <https://doi.org/10.1111/j.1547-5069.1999.tb00420.x>
- Aleo, G., Bagnasco, A., Watson, R., Dyson, J., Cowdell, F., Catania, G., Zanini, M. P., Cozzani, E., Parodi, A., & Sasso, L. (2019). Comparing questionnaires across cultures: Using Mokken scaling to compare the Italian and English versions of the MOLES index. *Nursing Open*. <https://doi.org/10.1002/nop2.297>
- Allen, D. G., Lamb, G. D., & Westerblad, H. (2008). Skeletal Muscle Fatigue: Cellular Mechanisms. *Physiological Reviews*, 88(1), 287–332. <https://doi.org/10.1152/physrev.00015.2007>
- Almeida, L., & Freire, T. (2008). *Metodologia da investigação em psicologia e educação* (5ª ed.). Braga: Psiquilibrios.
- Andrade, D. F., Tavares, H. R., & da Cunha Valle, R. (2000). *Teoria da Resposta ao Item: conceitos e aplicações*. São Paulo: ABE.
- Barata, A. F. S. T. (2015). Tristeza positiva? O efeito das emoções induzidas por excertos musicais na resposta comportamental e pupilar numa tarefa de memória de reconhecimento. Dissertação apresentada ao Departamento de Psicologia da Universidade Lusófona de Humanidades e Tecnologias para obtenção do grau de mestre, orientada por Joana Brites Rosa, Lisboa. Retirado a 06 de dezembro, 2020, em <http://hdl.handle.net/10437/6872>
- Barsevick, A. M., Whitmer, K., & Walker, L. (2001). In their own words: using the common sense model to analyze patient descriptions of cancer-related fatigue. *Oncology Nursing Forum*, 28(9), 1363–1369. <https://pubmed.ncbi.nlm.nih.gov/11683307/>
- Benaicha, N., Diarra, A., Diakite, O. D., SY, O., Najdi, A., & Amine, M. (2016). Moroccan Taxi Drivers Fatigue Using Pichot Questionnaire: A Cross-Sectional Survey. *Journal of Life Sciences*, 10, 114-120.
- Berrios, G. E. (1990). Feelings of fatigue and psychopathology: A conceptual history. *Comprehensive Psychiatry*, 31(2), 140–151. [https://doi.org/10.1016/0010-440x\(90\)90018-n](https://doi.org/10.1016/0010-440x(90)90018-n)
- Borges, J. A., Quintão, M. M. P., Chermont, S. S. M. C., Mendonça Filho, H. T. F. de, & Mesquita, E. T. (2018). Fatigue: A Complex Symptom and its Impact on Cancer and

- Heart Failure. *International Journal of Cardiovascular Sciences*, 31, 433–442. <https://doi.org/10.5935/2359-4802.20180027>
- Brito, L. F. M. (2020). Estudo de Validação Psicométrica da Versão Portuguesa da Escala de Fadiga de Pichot. Dissertação apresentada ao Departamento de Psicologia da Universidade Lusófona de Humanidades e Tecnologias para obtenção do grau de mestre, orientada por Pedro Joel Rosa, Lisboa.
- Cattell, R. B. (1966). The Scree Test For The Number Of Factors. *Multivariate Behavioral Research*, 1(2), 245–276. https://doi.org/10.1207/s15327906mbr0102_10
- Chalder, T., Berelowitz, G., Pawlikowska, T., Watts, L., Wessely, S., Wright, D., & Wallace, E. P. (1993). Development of a fatigue scale. *Journal of Psychosomatic Research*, 37(2), 147–153. [https://doi.org/10.1016/0022-3999\(93\)90081-P](https://doi.org/10.1016/0022-3999(93)90081-P)
- Chisholm, D., Godfrey, E., Ridsdale, L., Chalder, T., King, M., Seed, P., Wallace, P., Wessely, S., & Fatigue Trialists' Group. (2001). Chronic fatigue in general practice: economic evaluation of counselling versus cognitive behaviour therapy. *The British Journal of General Practice: The Journal of the Royal College of General Practitioners*, 51(462), 15–18. <https://pubmed.ncbi.nlm.nih.gov/11271867/>
- Crane, P. K., Belle, G. van, & Larson, E. B. (2004). Test bias in a cognitive test: differential item functioning in the CASI. *Statistics in Medicine*, 23(2), 241–256. <https://doi.org/10.1002/sim.1713>
- Crawford, A. V., Green, S. B., Levy, R., Lo, W.-J., Scott, L., Svetina, D., & Thompson, M. S. (2010). Evaluation of Parallel Analysis Methods for Determining the Number of Factors. *Educational and Psychological Measurement*, 70(6), 885–901. <https://doi.org/10.1177/0013164410379332>
- Dittner, A. J., Wessely, S. C., & Brown, R. G. (2004). The assessment of fatigue. *Journal of Psychosomatic Research*, 56(2), 157–170. [https://doi.org/10.1016/s0022-3999\(03\)00371-4](https://doi.org/10.1016/s0022-3999(03)00371-4)
- Dragow, F., & Parsons, C. K. (1983). Application of Unidimensional Item Response Theory Models to Multidimensional Data. *Applied Psychological Measurement*, 7(2), 189–199. <https://doi.org/10.1177/014662168300700207>
- Dukes, J. C., Chakan, M., Mills, A., & Marcaurd, M. (2021). Approach to Fatigue. *Medical Clinics of North America*, 105(1), 137–148. <https://doi.org/10.1016/j.mcna.2020.09.007>
- Embretson, S. E., & Reise, S. P. (2013). *Item response theory*. Nova York: Psychology Press.

- Fisk, J. D., & Doble, S. E. (2002). Construction and validation of a fatigue impact scale for daily administration (D-FIS). *Quality of Life Research, 11*(3), 263–272. <https://doi.org/10.1023/a:1015295106602>
- Freitag, R. M. K. (2018). Amostras sociolinguísticas: probabilísticas ou por conveniência? *Revista de estudos da linguagem, 26*(2), 667–686. <https://doi.org/10.17851/2237-2083.26.2.667-686>
- Friedrich, O., Sipötz, J., Benzer, W., Kunschitz, E., & Höfer, S. (2015). The dimensional structure of the MacNew Health Related Quality of Life questionnaire: A Mokken Scale Analysis. *Journal of Psychosomatic Research, 79*(1), 43–48. <https://doi.org/10.1016/j.jpsychores.2015.04.007>
- Fukuda, K. (1994). The Chronic Fatigue Syndrome: A Comprehensive Approach to Its Definition and Study. *Annals of Internal Medicine, 121*(12), 953. <https://doi.org/10.7326/0003-4819-121-12-199412150-00009>
- Gamito, P. & Rosa, P. J. (2014). *I see me, you see me: inferring cognitive and emotional processes from gazing behavior*. Newcastle upon Tyne: Cambridge Scholars Publishing. ISBN:1443854603
- Gonçalves, A. J. P. (2018). *Deteção e identificação de ameaça biologicamente relevante: um estudo com image morphing*. Dissertação apresentada ao Departamento de Psicologia da Universidade Lusófona de Humanidades e Tecnologias para obtenção do grau de mestre, orientada por Pedro Joel Rosa, Lisboa. Retirado a 20 de julho, 2021, em <https://recil.ensinolusofona.pt/bitstream/10437/9471/1/Disserta%C3%A7%C3%A3o%20Vers%C3%A3o%20definitiva%20Antonio%20Gon%C3%A7alves.pdf>
- Hilari, K., Byng, S., Lamping, D. L., & Smith, S. C. (2003). Stroke and Aphasia Quality of Life Scale-39 (SAQOL-39). *Stroke, 34*(8), 1944–1950. <https://doi.org/10.1161/01.str.0000081987.46660.ed>
- Islam, M. F., Cotler, J., & Jason, L. A. (2020). Post-viral fatigue and COVID-19: lessons from past epidemics. *Fatigue: Biomedicine, Health & Behavior, 8*(2), 61–69. <https://doi.org/10.1080/21641846.2020.1778227>
- Jason, L. A., Evans, M., Brown, M., & Porter, N. (2010). What is Fatigue? Pathological and Nonpathological Fatigue. *PM&R, 2*(5), 327–331. <https://doi.org/10.1016/j.pmrj.2010.03.028>

- Kaiser, H. F. (1960). The Application of Electronic Computers to Factor Analysis. *Educational and Psychological Measurement*, 20(1), 141–151. <https://doi.org/10.1177/001316446002000116>
- Krupp, L. B., LaRocca, N. G., Muir-Nash, J., & Steinberg, A. D. (1989). The Fatigue Severity Scale. *Archives of Neurology*, 46(10), 1121. <https://doi.org/10.1001/archneur.1989.00520460115022>
- Lee, K. A., Hicks, G., & Nino-Murcia, G. (1991). Validity and reliability of a scale to assess fatigue. *Psychiatry Research*, 36(3), 291–298. [https://doi.org/10.1016/0165-1781\(91\)90027-m](https://doi.org/10.1016/0165-1781(91)90027-m)
- Lopes, P., Silva, R., Oliveira, J., Ambrósio, I., Ferreira, D., Crespo, C., Feiteira, F. & Rosa, P. J. (2018). Rasch Analysis on the Academic Motivation Scale in Portuguese University Students. *NeuroQuantology*, 16(3),41-46 doi: 10.14704/nq.2018.16.3.1062
- Maroco, J., & Garcia-Marques, T. (2006). Qual a fiabilidade do alfa de Cronbach? Questões antigas e soluções modernas? *Laboratório de Psicologia*, 4(1). <https://doi.org/10.14417/lp.763>
- Michielsen, H. J., De Vries, J., Van Heck, G. L., Van de Vijver, F. J. R., & Sijtsma, K. (2004). Examination of the Dimensionality of Fatigue. *European Journal of Psychological Assessment*, 20(1), 39–48. <https://doi.org/10.1027/1015-5759.20.1.39>
- Molenaar, W., & Sijtsma, K. (2000). *MSP5 for Windows User's Manual*. Iec ProGAMMA.
- Mokken, R. J. (1971). A theory and procedure of scale analysis: with applications in political research.
- Mota, D. D. C. F., & Pimenta, C. A. de M. (2002). Fadiga em pacientes com câncer avançado: conceito, avaliação e intervenção. *Revista Brasileira de Cancerologia*, 48(4), 577–583. <https://doi.org/10.32635/2176-9745.rbc.2002v48n4.2172>
- Mota, D. D. C. F., Cruz, D. de A. L. M. da, & Pimenta, C. A. de M. (2005). Fadiga: uma análise do conceito. *Acta Paulista de Enfermagem*, 18(3), 285–293. <https://doi.org/10.1590/s0103-21002005000300009>
- Mota, D. D. C. F., & Pimenta, C. A. M. (2006). Self-Report Instruments for Fatigue Assessment: A Systematic Review. *Research and Theory for Nursing Practice*, 20(1), 49–78. <https://doi.org/10.1891/rtnp.20.1.49>
- Mota, D. D. C. F., Pimenta, C. A. de M., & Caponero, R. (2012). Fatigue in colorectal cancer patients: prevalence and associated factors. *Revista Latino-Americana de Enfermagem*, 20(3), 495–503. <https://doi.org/10.1590/s0104-11692012000300010>

- National Institutes of Health [NIH]. (2020). <https://www.nih.gov/mecfs/about-mecfs>. Acedido a 21 de novembro, 2020 de NIH em www.nlm.nih.gov/.
- Neu, D., Linkowski, P., & le Bon, O. (2010). Clinical complaints of daytime sleepiness and fatigue: how to distinguish and treat them, especially when they become “excessive” or “chronic”?. *Acta Neurologica Belgica*, 110(1), 15–25. <https://pubmed.ncbi.nlm.nih.gov/20514923/>
- Neves, R. S. M. C. (2019). Stresso, logo dilato: um estudo pupilométrico com exposição às imagens da Lusophone Technostress Image Database (LTID). Dissertação apresentada ao Departamento de Psicologia da Universidade Lusófona de Humanidades e Tecnologias para obtenção do grau de mestre, orientada por Pedro Joel Rosa, Lisboa. Retirado a 14 de setembro, 2021, <http://hdl.handle.net/10437/10087>
- Palmgren, P. J., Brodin, U., Nilsson, G. H., Watson, R., & Stenfors, T. (2018). Investigating psychometric properties and dimensional structure of an educational environment measure (DREEM) using Mokken scale analysis – a pragmatic approach. *BMC Medical Education*, 18(1). <https://doi.org/10.1186/s12909-018-1334-8>
- Pasqual, L., (2009). *Psicometria teoria dos testes na psicologia e na educação*. Petrópolis: Vozes.
- Pasquali, L., & Primi, R. (2003). Fundamentos da teoria da resposta ao item: TRI. *Avaliação Psicológica*, 2(2),99–110.
- Pichot, P., & Brun, J. P. (1984). Brief self-evaluation questionnaire for depressive, asthenic and anxious dimensions. *Annales Medico-Psychologiques*, 142(6), 862–865.
- Piper, B. F., Dibble, S. L., Dodd, M. J., Weiss, M. C., Slaughter, R. E., & Paul, S. M. (1998). The revised Piper Fatigue Scale: psychometric evaluation in women with breast cancer. *Oncology Nursing Forum*, 25(4), 677–684.
- Ream, E., & Richardson, A. (1996). Fatigue: a concept analysis. *International Journal of Nursing Studies*, 33(5), 519–529. [https://doi.org/10.1016/0020-7489\(96\)00004-1](https://doi.org/10.1016/0020-7489(96)00004-1)
- Rosa, P. J. M. (2012). Efeitos da exposição repetida subliminar a estímulos biologicamente relevantes nos estados emocionais. Tese apresentada ao Departamento de psicologia do Instituto Universitário de Lisboa para obtenção do grau de doutor, orientada por Francisco Gomes Esteves, Lisboa. Retirado a 18 de janeiro, 2021, em <https://repositorio.iscte-iul.pt/bitstream/10071/7434/1/Tese%20de%20doutoramento%20Pedro%20Rosa.pdf>

- Rosa, P. J., Caires, C., Costa, L., Rodelo, L., & Pinto, L. (2014). Affective And Psychophysiological Responses To Erotic Stimuli: Does Color Matter? In: P. Gamito, P. J. Rosa (Eds.), *I see me, you see me: inferring cognitive and emotional processes from gazing behavior* (pp. 171-190). Newcastle upon Tyne: Cambridge Scholars Publishing.
- Rosa, P. J., Esteves, F. & Arriaga (2015). Beyond traditional clinical measurements for screening fears and phobias. *IEEE Transactions on Instrumentation & Measurement*, 64 (12), 3396-3404. doi 10.1109/TIM.2015.2450292
- Rosa P.J., Lopes P., Oliveira J., & Pascoal, P. (2019). Does Length Really Matter? Effects of Number of Pages in the Informed Consent on Reading Behavior: An Eye-Tracking Study. In: Fardoun H., Hassan A., de la Guía M. (eds) *New Technologies to Improve Patient Rehabilitation* (pp.116-125). Springer, Cham
- Rosa, P. J., Oliveira, J., Alghazzawi, D, Fardoun, H. & Gamito, P. (2017). Affective and Physiological Correlates of Perception of Unimodal and Bimodal Emotional Stimuli. *Psicothema*, 29(3), 364-369.
- Sharpe, M. (2002). ABC of psychological medicine: Fatigue. *BMJ*, 325(7362), 480–483. <https://doi.org/10.1136/bmj.325.7362.480>
- Sijtsma, K., & Molenaar, I. W. (2002). *Introduction to nonparametric item response theory*. Sage Publications.
- Stochl, J., Jones, P. B., & Croudace, T. J. (2012). Mokken scale analysis of mental health and well-being questionnaire item responses: a non-parametric IRT method in empirical research for applied health researchers. *BMC Medical Research Methodology*, 12(1). <https://doi.org/10.1186/1471-2288-12-74>
- Taylor, R. R., Jason, L. A., & Torres, A. (2000). Fatigue rating scales: an empirical comparison. *Psychological Medicine*, 30(4), 849–856. <https://doi.org/10.1017/s0033291799002500>
- Van der Ark, L. A. (2007). Mokken scale analysis in R. *Journal of statistical software*, 20(11). <https://doi.org/10.18637/jss.v020.i11>
- van der Linden, D., & Eling, P. (2005). Mental fatigue disturbs local processing more than global processing. *Psychological Research Psychologische Forschung*, 70(5), 395–402. <https://doi.org/10.1007/s00426-005-0228-7>
- van der Linden, W. J., (2016). *Handbook of modern item response theory: Volume 2, statistical tools*. Crc Press.

- Williamson, R. J., Purcell, S., Sterne, A., Wessely, S., Hotopf, M., Farmer, A., & Sham, P. C. (2005). The relationship of fatigue to mental and physical health in a community sample. *Social Psychiatry and Psychiatric Epidemiology*, *40*(2), 126–132. <https://doi.org/10.1007/s00127-005-0858-5>
- Zumbo, B. D., (1999). A Handbook on the Theory and Methods of Differential Item Functioning (DIF) logistic regression modeling as a unitary framework for binary and likert-type (ordinal) item scores. *Theory and Methods of DIF*, *1*. <https://faculty.educ.ubc.ca/zumbo/DIF/handbook.pdf>

Anexo

Tabela 6

Versão portuguesa da escala de fadiga de Pichot e Brun

Itens

1. Não tenho energia
 2. Parece que é necessário esforço em tudo que faço
 3. Sinto-me fraco em determinadas partes do meu corpo
 4. Sinto as minhas pernas e os meus braços exaustos
 5. Sinto-me cansado sem motivo aparente
 6. Tenho vontade de me deitar e descansar
 7. Tenho dificuldades em concentrar-me
 8. Sinto-me exausto, stressado e pesado
-