

Questionário de Personalidade de Eysenck – Forma Revista (EPQ-R): Estudo Normativo da Versão Portuguesa

Eysenck Personality Questionnaire – Revised (EPQ-R): Normative Study of the Portuguese Version

Pedro Armelim Almiro¹ e Mário R. Simões²

Resumo

O *Questionário de Personalidade de Eysenck – Forma Revista (EPQ-R)* avalia as três dimensões fundamentais da personalidade: *Psicoticismo*, *Extroversão*, *Neuroticismo*, e inclui uma escala de *Mentira/Desejabilidade Social*. Neste artigo, apresentamos os resultados do estudo de aferição do EPQ-R para a população portuguesa, obtidos com recurso a uma amostra extensa e representativa recolhida na comunidade. Neste estudo, participaram 1689 sujeitos (783 homens e 906 mulheres, 16-60 anos) provenientes de diversas áreas geográficas de Portugal. As normas de interpretação do EPQ-R foram estabelecidas para as variáveis género e idade, tendo sido encontradas diferenças estatísticas significativas entre os diferentes grupos. No âmbito da Teoria da Resposta ao Item, foi realizada uma análise DIF que confirmou a qualidade dos itens do EPQ-R na medição dos constructos N, E, P e L, para as variáveis género e idade.

Palavras-chave: versão portuguesa do EPQ-R, avaliação da personalidade, aferição nacional, dados normativos, teoria da resposta ao item

Abstract

The *Eysenck Personality Questionnaire – Revised (EPQ-R)* assesses the three fundamental personality dimensions: *Psychoticism*, *Extraversion*, *Neuroticism*, including a *Lie/Social Desirability* scale. This paper presents the results of the normalization study of EPQ-R in the Portuguese population, which were obtained through a large and representative community sample. A sample of 1689 subjects participated in this study (783 men and 906 women, ages 16-60 years old) from several geographic areas in Portugal. The norms of the Portuguese Version of EPQ-R were constructed considering gender and age, and significant statistical differences between several groups were found. Through a model of Item Response Theory, a DIF analysis was performed and showed the psychometric quality of the EPQ-R items in measuring the N, E, P, and L constructs, concerning gender and age variables.

Keywords: EPQ-R Portuguese version, personality assessment, normalization study, normative data, item response theory

¹ Doutor em Avaliação Psicológica. Investigador no Centro de Investigação em Neuropsicologia e Intervenção Cognitivo-Comportamental (CINEICC) e no Laboratório de Avaliação Psicológica e Psicometria (*PsyAssessmentLab*) da Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra (FPCEUC). Professor Auxiliar na Universidade Autónoma de Lisboa (UAL). Investigador no Centro de Investigação em Psicologia (CIP-UAL). Rua de Santa Marta, 47, 1150-293 Lisboa, Portugal. Tel.: 00351 213177600. E-mail: pedroarmelimalmiro@gmail.com

² Professor Catedrático na Universidade de Coimbra. Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra. CINEICC – Centro de Investigação em Neuropsicologia e Intervenção Cognitivo-Comportamental. *PsyAssessmentLab* – Laboratório de Avaliação Psicológica e Psicometria. Rua do Colégio Novo, 3000-115 Coimbra, Portugal. Tel.: 00351 239851450. E-mail: simoesmr@fpce.uc.pt

Introdução

De acordo com H. Eysenck, a personalidade consiste numa organização mais ou menos estável e perdurável do “caráter”, do “temperamento” e dos “aspetos intelectuais” e “físicos” de um sujeito, que determina o seu ajustamento único ao meio ambiente. O “caráter” reflete de modo mais ou menos estável e perdurável o sistema comportamental volitivo (vontade). O “temperamento” reflete o sistema comportamental afetivo (emoção). Os “aspetos intelectuais” refletem o sistema comportamental cognitivo (inteligência). E os “aspetos físicos” refletem a configuração corporal e neuroendócrina do sujeito (H. Eysenck & M. Eysenck, 1985).

Baseado no denominado Modelo P-E-N (ou *Big Three*) (cf. H. Eysenck & M. Eysenck, 1985), o *Questionário de Personalidade de Eysenck – Forma Revista* (EPQ-R; versão original inglesa, 100 itens; S. Eysenck, H. Eysenck, & Barrett, 1985) consiste na versão revista do *Questionário de Personalidade de Eysenck* (EPQ; H. Eysenck & S. Eysenck, 1975; versão portuguesa, Castro Fonseca, S. Eysenck, & Simões, 1991). Este questionário foi construído para avaliar as três dimensões fundamentais da personalidade: *Psicoticismo* (P), *Extroversão* (E), *Neuroticismo* (N). Estas dimensões são examinadas através de três escalas de personalidade – P, E, N – e uma escala de validade, a escala L (*Lie, Mentira/Desejabilidade Social*). As dimensões de personalidade constituem os elementos essenciais da estrutura da personalidade (H. Eysenck & S. Eysenck, 1975, 2008).

No Modelo P-E-N, a dimensão N organiza-se num *continuum* que caracteriza num polo a *personalidade neurótica* (ou instável) – traços de emotividade, ansiedade, depressão, excesso de preocupação, irritabilidade fácil, baixa autoestima, entre outros. No outro polo, N caracteriza a *personalidade estável* – traços de serenidade, autocontrolo, boa disposição, entre outros.

A dimensão E também se organiza num *continuum*, caracterizando num polo a *personalidade extrovertida* – traços de sociabilidade, vivacidade, atividade, assertividade, espontaneidade, espírito de aventura, otimismo, entre outros. No outro polo, a dimensão E caracteriza a *personalidade introvertida* – traços

de introspeção, inibição, baixa sociabilidade, pessimismo, cautela, entre outros.

A dimensão P estabelece um *continuum* entre a psicopatologia e a normalidade. Esta dimensão caracteriza num polo os sujeitos com elevado *psicoticismo* – que são hostis, egocêntricos, desajustados, rígidos, impulsivos, agressivos, desconfiados, pouco empáticos, entre outros traços. E no outro polo, P caracteriza os sujeitos com baixo *psicoticismo* (personalidade ajustada) – que são amáveis, conscienciosos, empáticos, conformistas, cooperativos, entre outros traços (H. Eysenck & S. Eysenck, 2008).

O EPQ-R é um reconhecido instrumento de avaliação da personalidade, que resultou do aperfeiçoamento de vários instrumentos elaborados por H. Eysenck ao longo de mais de cinco décadas de investigação (cf. Barrett & S. Eysenck, 1984; Barrett, Petrides, S. Eysenck, & H. Eysenck, 1998; S. Eysenck & Barrett, 2013; Furnham, S. Eysenck, & Saklofske, 2008). Este instrumento foi estudado em mais de 34 países (e.g., Inglaterra, Espanha, Itália, Alemanha, Holanda, Austrália, Japão) com evidência de boas qualidades psicométricas. No que respeita à validade, a estrutura fatorial é estável e tem sido sucessivamente replicada. Relativamente à precisão, os índices de consistência interna e de estabilidade temporal teste-reteste são na generalidade elevados (ver estudos: versão original, S. Eysenck et al., 1985; versões italianas, San Martini, Mazzotti, & Setaro, 1996; Dazzi, 2011; versão espanhola, Ortet, Ibáñez, Moro, Silva, & Boyle, 1999; versão alemã, Ruch, 1999; ver revisões: Almiro, 2013; Almiro & Simões, 2011; Furnham et al., 2008). A utilidade deste instrumento tem sido demonstrada em diversos contextos (nomeadamente, normativo, clínico, forense, saúde, educacional, organizacional, militar, social, entre outros) (cf. Lynn, 1981; Nyborg, 1997).

Estudo da Versão Portuguesa do EPQ-R

A *Versão Portuguesa do EPQ-R* (Almiro & Simões, 2013, 2014) mantém a mesma estrutura fatorial da versão original do instrumento e é composta por 70 itens distribuídos por quatro escalas: N, E, P e L. Os procedimentos relativos à adaptação cultural do EPQ-R seguiram as normas habitualmente usadas neste tipo de investigação

(cf. Hambleton, Merenda, & Spielberger, 2005; American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 2014; International Test Commission [ITC], 2017).

O estudo das propriedades psicométricas do EPQ-R ($N=1689$ sujeitos, 783 homens e 906 mulheres) revelou, na generalidade, bons índices de precisão e de validade, quer no domínio da Teoria Clássica dos Testes (TCT), quer no domínio da Teoria da Resposta ao Ítem (TRI) (cf. Almiro, 2013).

No que respeita à precisão do EPQ-R, Almiro e Simões (2014) obtiveram os seguintes alfas de Cronbach: .87 para N, .83 para E (coeficientes “muito bons”), .78 para L (“respeitável”) e .55 para P (“inaceitável”) (de acordo com o critério de DeVellis, 2017). Com o método de teste-reteste ($N=124$; estabilidade temporal 4-8 semanas), Almiro e Simões (2014) obtiveram coeficientes adequados (de acordo com o critério de Nunnally & Bernstein, 1994): .89 para E, .86 para N, .86 para L e .72 para P.

No que respeita à validade, Almiro, Moura e Simões (2016) recorreram à Análise Fatorial Confirmatória (AFC) para testar o ajustamento fatorial dos 4 fatores (N, E, P e L) do modelo do EPQ-R. Os autores consideraram quatro indicadores de ajustamento: rácio qui-quadrado/graus de liberdade (χ^2/df), *Comparative Fit Index* (CFI), *Standardized Root Mean Square Residual* (SRMR), *Root Mean Square Error of Approximation* (RMSEA). Os índices obtidos por Almiro e colaboradores (2016) revelaram um bom ajustamento da estrutura fatorial da Versão Portuguesa do EPQ-R [$\chi^2(129)=533.786$, $p<.001$; $\chi^2/df=4.13$; CFI=.961; SRMR=.042; RMSEA=.043] (de acordo com o critério de Hu & Bentler, 1999), demonstrando adequabilidade e robustez.

No âmbito da TRI, Almiro e Simões (2015) aplicaram o *Modelo de Rasch* (modelo logístico de um parâmetro) para examinar a qualidade dos itens do EPQ-R e a sua adequação para avaliar os atributos N, E, P e L nos sujeitos. Nesta análise, todos os itens do EPQ-R apresentam bons índices de ajuste *infit* (no global: mínimo .76 e máximo 1.22; média entre .99 e 1.00) e de ajuste *outfit* (no global: mínimo .58 e máximo 1.47; média entre

.98 e 1.01). De acordo com o critério de Linacre (2009), o ajuste é ótimo quando os índices oscilam entre .50 e 1.50, com média próxima de 1.00. Os dados do referido estudo revelaram a boa qualidade dos itens deste instrumento na avaliação dos constructos N, E, P e L.

Considerando os estudos psicométricos efetuados por Almiro e colaboradores (Almiro, 2013; Almiro & Simões, 2014; Almiro et al., 2016), no presente artigo serão apresentados os resultados do estudo de aferição do EPQ-R para a população portuguesa, com o objetivo de estabelecer as normas de interpretação das pontuações obtidas com o instrumento no contexto português. As normas serão estabelecidas considerando tanto a totalidade da amostra, como as variáveis género e idade.

Com base nas variáveis demográficas consideradas no estudo normativo da Versão Portuguesa do EPQ-R, foram definidas as seguintes hipóteses de estudo: H1 – existem diferenças estatísticas significativas nas pontuações do EPQ-R em relação à variável género; H2 – existem diferenças estatísticas significativas nas pontuações do EPQ-R em relação à variável idade; H3 – o funcionamento diferencial dos itens do EPQ-R cumpre o princípio da invariância da TRI para a variável género; H4 – o funcionamento diferencial dos itens do EPQ-R cumpre o princípio da invariância da TRI para a variável idade.

Método

Participantes

A aferição de um teste psicológico consiste no seu processo de ajustamento para uma população, incluindo: por um lado, os estudos específicos orientados para o exame das características psicométricas do teste (e.g., estudos no âmbito da análise dos itens, precisão e validade); e por outro, os cuidados de natureza metodológica adotados na constituição de uma amostra representativa dessa população que viabilize o estabelecimento de normas de interpretação úteis e válidas (Geisinger, 1992, 1994; cf. AERA, APA & NCME, 2014; ITC, 2017).

No presente estudo de aferição, com o objetivo de estabelecer as normas de interpretação do EPQ-R para as variáveis género e idade, foi

Quadro 1. Distribuição da ANN por NUTS II, localização litoral/interior, áreas metropolitanas e meio

	Amostra Teórica		Amostra Normativa Nacional (ANN)		
	Portugal (%)	N	Total *(%)	Homens (%)	Mulheres (%)
Total	10 562 178 (100.0%)	1600	1689 (100.0%)	783 (46.36%)	906 (53.64%)
Zonas (NUTS)					
Norte Litoral	3 233 571 (30.61%)	490	517 (30.61%)	251 (14.86%)	266 (15.75%)
Norte Interior	456 111 (4.32%)	70	87 (5.15%)	41 (2.43%)	46 (2.72%)
Centro Litoral	1 346 630 (12.75%)	204	208 (12.31%)	87 (5.15%)	121 (7.16%)
Centro Interior	981 125 (9.29%)	149	164 (9.71%)	86 (5.09%)	78 (4.62%)
Lisboa	2 821 876 (26.72%)	428	429 (25.40%)	190 (11.25%)	239 (14.15%)
Alentejo Litoral	345 378 (3.27%)	53	60 (3.55%)	27 (1.60%)	33 (1.95%)
Alentejo Interior	411 924 (3.90%)	63	75 (4.44%)	34 (2.01%)	41 (2.43%)
Algarve	451 006 (4.27%)	69	67 (3.97%)	32 (1.89%)	35 (2.07%)
Açores	246 772 (2.34%)	38	47 (2.78%)	21 (1.24%)	26 (1.54%)
Madeira	267 785 (2.53%)	41	35 (2.07%)	14 (0.83%)	21 (1.24%)
Localização					
Litoral	8 713 018 (82.49%)	1320	1363 (80.70%)	622 (36.83%)	741 (43.87%)
Interior	1 849 160 (17.51%)	281	326 (19.30%)	161 (9.53%)	165 (9.77%)
Áreas Metropolitanas					
Grande Lisboa	2 042 477 (19.34%)	310	242 (14.33%)	113 (6.69%)	129 (7.64%)
Grande Porto	1 287 282 (12.19%)	196	290 (17.17%)	145 (8.58%)	145 (8.58%)
Meio					
APU (Área Predominantemente Urbana)		–	1406 (83.24%)	639 (37.83%)	767 (45.41%)
AMU (Área Medianamente Urbana)		–	152 (9.00%)	80 (4.74%)	72 (4.26%)
APR (Área Predominantemente Rural)		–	131 (7.76%)	64 (3.79%)	67 (3.97%)

Nota. N=participantes, * percentagens calculadas para N=1689.

constituída uma amostra normativa significativa – a Amostra Normativa Nacional (ANN) – a partir de uma amostra teórica de 1600 sujeitos, com base nas Nomenclaturas de Unidades Territoriais para fins Estatísticos (NUTS) que constam nos dados do Censos 2011. Pretendia-se que a ANN fosse uma amostra extensa, abrangente e o mais representativa possível da população portuguesa. Os princípios de significância e de representatividade considerados seguiram as recomendações de Almeida e Freire (2017).

A dimensão da referida amostra teórica foi inicialmente fixada em 1000 sujeitos, por dois motivos. Por um lado, as amostras utilizadas nos estudos das versões inglesa (100 itens; S. Eysenck et al., 1985) e espanhola (83 itens; Ortet et al., 1999) do EPQ-R têm um número aproximado de participantes (EPQ-R inglês, N=902; EPQ-R espanhol, N=1110). Por outro lado, Nunnally e Bernstein (1994) recomendam um mínimo de 10 sujeitos por item/variável (10:1) para aplicar o método da análise fatorial, sendo que os estudos psicométricos da Versão Portuguesa do EPQ-R foram efetuados com recurso à denominada ANN (cf. Almiro, 2013).

Neste sentido, para a realização do presente estudo de aferição, foram consultados os dados mais recentes do recenseamento geral da população portuguesa (Censos 2011, XV Recenseamento Geral da População) disponibilizados pelo Instituto

Nacional de Estatística (INE, 2012). Em seguida, foram calculadas as percentagens para a distribuição da amostra com base nos dados da população residente, considerando as zonas geográficas estratificadas por NUTS. Segundo o INE (2012), as NUTS II abrangem as seguintes grandes zonas geográficas: Norte, Centro, Lisboa, Alentejo, Algarve, Açores, Madeira. Com efeito, na recolha de dados foram seguidas as proporções geográficas definidas pelo INE (Censos 2011), considerando uma amostra teórica de 1600 sujeitos (800 homens e 800 mulheres).

Por conseguinte, a ANN é constituída por sujeitos provenientes das grandes zonas geográficas de Portugal – Norte, Centro, Lisboa, Alentejo, Algarve, Açores, Madeira – e de todos os Distritos do território português. O Quadro 1 mostra a comparação (N e respetivas percentagens) entre a amostra teórica de 1600 sujeitos e a ANN (Amostra Normativa Nacional, N=1689), que é a amostra real (obtida para o presente estudo de aferição). Neste sentido, consideraram-se os sujeitos oriundos das grandes zonas geográficas e depois a sua estratificação em função das áreas litoral/interior de Portugal e das Áreas Metropolitanas da Grande Lisboa e do Grande Porto. No presente estudo, a amostra real (1689 sujeitos) é mais extensa do que a amostra teórica (1600 sujeitos). No entanto, a ANN não atingiu o número previsto de participantes nas zonas

Quadro 2. Caracterização demográfica da ANN

ANN	T (%)	H (%)	M (%)	Escolaridade	T (%)	H (%)	M (%)
<i>N</i>	1689 (100.0%)	783 (46.36%)	906 (53.64%)	Básico	431 (25.52%)	219 (12.97%)	212 (12.55%)
Idade	T (%)	H (%)	M (%)	Secundário	685 (40.56%)	320 (18.95%)	365 (21.61%)
<i>M</i>	32.34 anos	32.58 anos	32.13 anos	Bacharelato/ Licenciatura	473 (28.00%)	201 (11.90%)	272 (16.10%)
<i>DP</i>	11.22 anos	11.66 anos	10.81 anos	Mestrado/ Doutoramento	100 (5.92%)	43 (2.55%)	57 (3.37%)
Grupo Etário	T (%)	H (%)	M (%)	Estado Civil	T (%)	H (%)	M (%)
16-20 anos	265 (15.69%)	122 (7.22%)	143 (8.47%)	Solteiro	870 (51.51%)	417 (24.69%)	453 (26.82%)
21-30 anos	568 (33.63%)	264 (15.63%)	304 (18.00%)	Casado	633 (37.48%)	296 (17.53%)	337 (19.95%)
31-40 anos	430 (25.46%)	195 (11.55%)	235 (13.91%)	Divorciado	113 (6.69%)	39 (2.31%)	74 (4.38%)
41-50 anos	259 (15.33%)	123 (7.28%)	136 (8.05%)	Viúvo	6 (0.36%)	1 (0.06%)	5 (0.30%)
51-60 anos	167 (9.89%)	79 (4.68%)	88 (5.21%)	União de Facto	67 (3.97%)	30 (1.78%)	37 (2.19%)

Nota. *N*=participantes, T=total, H=homens, M=mulheres, *M*=média, *DP*=desvio-padrão.

do Algarve e da Madeira, e na Área Metropolitana da Grande Lisboa (Quadro 1).

A amostra em estudo, a ANN, é constituída por 1689 sujeitos, 783 homens (46.36%) e 906 mulheres (53.64%), com idades compreendidas entre 16 e 60 anos [total: $M=32.34$ ($DP=11.22$); homens: $M=32.58$ ($DP=11.66$); mulheres: $M=32.13$ ($DP=10.81$)], e corresponde à amostra real, tomando como referência a amostra teórica de 1600 sujeitos (dados do INE, 2012).

A ANN é também constituída pelos seguintes grupos etários: Grupo 1 (16-20 anos), composto por 265 sujeitos (122 homens e 143 mulheres); Grupo 2 (21-30 anos), por 568 sujeitos (264 homens e 304 mulheres); Grupo 3 (31-40 anos), por 430 sujeitos (195 homens e 235 mulheres); Grupo 4 (41-50 anos), por 259 sujeitos (123 homens e 136 mulheres); e Grupo 5 (51-60 anos), por 167 sujeitos (79 homens e 88 mulheres). A definição dos grupos etários e da respetiva amplitude de idades seguiu os critérios adotados por S. Eysenck et al. (1985) (estudo da versão original inglesa do EPQ-R) e por Ruch (1999) (estudo da versão alemã do instrumento).

Os dados do INE (2012) reportam uma relação de masculinidade na população portuguesa de 91.50 homens por cada 100 mulheres [quociente entre os efetivos populacionais do sexo masculino e do sexo feminino (total: 10 562 178 sujeitos, 5 046 600 homens e 5 515 578 mulheres)]. Na ANN, a proporção é de 86.42 homens por 100 mulheres. Em relação à idade, os dados do INE (2012)

referem uma média total de idades de 41.8 anos, 40.3 anos nos homens e 43.2 anos nas mulheres. Na ANN, a média total de idades é de 32.34 anos, 32.58 anos nos homens e 32.13 anos nas mulheres.

Os sujeitos que participaram nesta investigação diferiam quanto ao nível de escolaridade, estado civil, exerciam diferentes atividades profissionais e residiam em diversas áreas geográficas de Portugal (Continental e Ilhas, litoral/interior, meio urbano/rural). A caracterização demográfica da ANN é apresentada no Quadro 2, considerando as variáveis género, grupo etário, escolaridade e estado civil. Esta amostra apresenta, assim, diversidade quanto às variáveis demográficas em estudo, mas foi recolhida por conveniência.

Instrumentos

No presente estudo foi aplicada a *Versão Portuguesa do Questionário de Personalidade de Eysenck – Forma Revista* (EPQ-R; Almiro & Simões, 2014), que é constituída por 70 itens, distribuídos por quatro escalas: *Neuroticismo* (N), que tem 23 itens; *Extroversão* (E), que tem 20 itens; *Psicoticismo* (P), que tem 9 itens; *Mentira/Desejabilidade Social* (L), que tem 18 itens. O EPQ-R é um instrumento de autorresposta. Os itens são de resposta dicotómica (“Sim”/“Não”), alguns cotados de modo direto e outros de modo inverso. O EPQ-R é um teste normativo e pode ser aplicado a sujeitos a partir dos 16 anos (adolescentes, adultos e idosos), para fins de investigação e de avaliação psicológica na

comunidade e em contextos clínicos e forenses.

Antes de responderem ao EPQ-R, os participantes preencheram um questionário sociodemográfico especificamente elaborado para obter todas as informações relevantes para a presente investigação. Neste questionário, constavam os seguintes elementos: data (de preenchimento), sexo, idade, área de residência (Distrito, código postal), profissão, escolaridade [Ensino Básico, Ensino Secundário, Bacharelato/Licenciatura, Mestrado/Doutoramento], estado civil [solteiro(a), casado(a), divorciado(a), viúvo(a), outro], e as questões-critério “Sofre de alguma doença psiquiátrica ou neurológica?” [sim ou não; se sim, qual?] e “Toma alguma medicação?” [sim ou não; se sim, qual?].

Procedimentos

A aplicação do EPQ-R foi efetuada em diversos contextos da comunidade, de modo formal ou informal, incluindo administrações individuais ou coletivas, respeitando sempre os princípios da confidencialidade, mediante o preenchimento e assinatura do consentimento informado.

No modo formal, a aplicação dos questionários envolveu a colaboração de diversas organizações, depois de efetuado um pedido aos respetivos órgãos diretivos: as aplicações individuais foram realizadas pelo investigador ou por meio da entrega dos questionários aos recursos humanos, em envelopes individuais destinados a serem distribuídos pelos participantes, dando a indicação, neste caso, para a sua devolução em envelope lacrado para posterior recolha; nas aplicações coletivas, a recolha dos dados foi realizada pelo investigador em sessões de grupo agendadas pela própria organização. No modo informal, foi solicitada a colaboração voluntária de diversos sujeitos recrutados em locais públicos.

A participação nesta investigação foi voluntária e anónima, tendo sido concretizada mediante o consentimento informado, que incluía a explicação prévia da natureza do estudo, dos seus objetivos e da sua importância.

Com base no questionário sociodemográfico, foi definida a zona geográfica (Freguesia, Concelho e Distrito), a localização geográfica

(litoral/interior) e o meio (urbano/rural) a que pertencia cada sujeito, considerando as NUTS que definem o território português (NUTS Nível I, Nível II, Nível III). No que respeita à localização geográfica, a classificação foi realizada mediante as áreas litorais e interiores. Quanto ao meio, foi utilizada a classificação em Área Predominantemente Urbana (APU), Área Medianamente Urbana (AMU) e Área Predominantemente Rural (APR), com recurso à Tipologia de Áreas Urbanas (TAU, 2009) do INE para a geografia de difusão do Censos 2011.

Neste estudo, pelo facto do EPQ-R ser um instrumento sensível à psicopatologia/saúde mental e com reconhecidas potencialidades na sua avaliação (Almiro, 2013), utilizaram-se as questões-critério do questionário sociodemográfico para detetar e excluir os sujeitos com queixas clínicas de natureza psiquiátrica ou neurológica. Todos os questionários com dados omissos foram eliminados.

Estudo de Aferição e Análise de Dados

No âmbito do estudo de aferição nacional do EPQ-R (com recurso à ANN), foram estabelecidas as normas de interpretação através das médias (M) e dos desvios-padrão (DP), em função das variáveis género e idade. As pontuações médias obtidas, considerando estas variáveis demográficas, foram comparadas através do *teste t-student* e da *One-Way ANOVA* para analisar a existência de diferenças estatísticas significativas.

Com recurso ao Modelo de Rasch, no domínio da TRI, foi realizada uma análise do *Funcionamento Diferencial dos Itens* (DIF) para examinar a qualidade dos itens do EPQ-R na medição dos constructos de N, E, P e L (estudo de validade). Todas as análises estatísticas foram efetuadas através dos programas *SPSS Statistics* (versão 17.0; SPSS Inc., Chicago, IL) e *WINSTEPS – Rasch Measurement* (versão 3.69; Linacre, 2009).

Resultados

Distribuição das pontuações obtidas no EPQ-R

As pontuações médias obtidas nas escalas N, E, P e L do EPQ-R para a ANN ($N=1689$, 783 homens e 906 mulheres; idades entre 16 e 60 anos; grupos etários: 16-20 anos, 21-30 anos, 31-40 anos,

Quadro 3. Dados normativos do EPQ-R para a população portuguesa (ANN), total e grupo etário

Total	Total (N=1689)			
	Amplitude*	M	DP	Md z, [z]
N	[0, 23]	10.44	5.53	-0.08 [-1.89; 2.27]
E	[0, 20]	12.61	4.43	0.09 [-2.85; 1.67]
P	[0, 9]	1.01	1.31	-0.01 [-0.78; 5.35]
L	[0, 18]	9.68	3.72	0.08 [-2.60; 2.23]
G1	Total (N=265)			
16-20	Amplitude*	M	DP	Md z, [z]
N	[0, 23]	11.60	5.45	0.10 [-1.89; 2.09]
E	[0, 20]	13.41	4.43	0.31 [-2.62; 1.67]
P	[0, 9]	1.45	1.51	-0.01 [-0.78; 3.82]
L	[0, 18]	8.12	3.54	-0.45 [-2.60; 1.70]
G2	Total (N=568)			
21-30	Amplitude*	M	DP	Md z, [z]
N	[0, 23]	10.61	5.46	-0.08 [-1.89; 2.27]
E	[0, 20]	12.86	4.46	0.31 [-2.62; 1.67]
P	[0, 9]	1.07	1.37	-0.01 [-0.78; 5.35]
L	[0, 18]	8.82	3.50	-0.18 [-2.60; 2.23]
G3	Total (N=430)			
31-40	Amplitude*	M	DP	Md z, [z]
N	[0, 23]	9.81	5.34	-0.26 [-1.89; 2.27]
E	[0, 20]	12.20	4.29	0.09 [-2.85; 1.67]
P	[0, 9]	0.93	1.24	-0.39 [-0.78; 3.82]
L	[0, 18]	9.86	3.62	0.08 [-2.60; 2.23]
G4	Total (N=259)			
41-50	Amplitude*	M	DP	Md z, [z]
N	[0, 23]	10.10	5.70	-0.08 [-1.89; 2.27]
E	[0, 20]	12.53	4.41	0.09 [-2.62; 1.67]
P	[0, 9]	0.85	1.17	-0.77 [-0.78; 3.82]
L	[0, 18]	11.10	3.26	0.35 [-2.33; 2.23]
G5	Total (N=167)			
51-60	Amplitude*	M	DP	Md z, [z]
N	[0, 23]	10.13	5.82	-0.08 [-1.89; 2.09]
E	[0, 20]	11.65	4.48	-0.36 [-2.62; 1.44]
P	[0, 9]	0.63	0.83	-0.77 [-0.78; 1.52]
L	[0, 18]	12.43	3.41	0.89 [-2.33; 2.23]

Nota. N=participantes, M=média, DP=desvio-padrão. G1=16-20 anos, G2=21-30 anos, G3=31-40 anos, G4=41-50 anos, G5=51-60 anos. Md z=mediana dos z-scores, [z]=amplitude dos z-scores, * amplitude de resultados possíveis para cada escala.

41-50 anos, e 51-60 anos) são apresentadas no Quadro 3 (total e grupo etário), no Quadro 4 (homens e grupo etário), e no Quadro 5 (mulheres e grupo etário). Com efeito, estes constituem os dados normativos para a população portuguesa, que são apresentados em função das médias (M) e dos desvios-padrão (DP), para a totalidade da amostra e para as variáveis género e grupo etário, conforme o estudo normativo original do EPQ-R (cf. S. Eysenck et al., 1985). Nos referidos quadros, são também apresentados os respetivos z-scores (mediana e amplitude de valores), considerando as pontuações obtidas pelos participantes da ANN.

Comparação das pontuações obtidas no EPQ-R em função do género e da idade

Tendo em conta a representatividade da ANN, foram comparadas as pontuações médias obtidas nas escalas N, E, P e L do EPQ-R para verificar se existem diferenças estatísticas significativas e analisar a sua importância na caracterização da população portuguesa. Foram consideradas as variáveis género e grupo etário. Com efeito, recorreu-se à análise da *One-Way ANOVA*, seguida do *teste de Bonferroni* para analisar contrastes, e ao *teste t-student* (ambas para amostras independentes) para estudar as diferenças.

Quadro 4. Dados normativos do EPQ-R para a população portuguesa (ANN), homens e grupo etário

Total	Homens (n=783)			
	Amplitude*	M	DP	Md z, [z]
N	[0, 23]	9.35	5.31	-0.26 [-1.89; 2.27]
E	[0, 20]	12.66	4.64	0.09 [-2.85; 1.67]
P	[0, 9]	1.26	1.46	-0.01 [-0.78; 5.35]
L	[0, 18]	9.22	3.76	-0.18 [-2.60; 2.23]
G1	Homens (n=122)			
16-20	Amplitude*	M	DP	Md z, [z]
N	[0, 23]	10.71	5.56	0.10 [-1.89; 2.09]
E	[0, 20]	13.73	4.43	0.31 [-2.62; 1.67]
P	[0, 9]	1.98	1.63	0.75 [-0.78; 3.82]
L	[0, 18]	7.58	3.48	-0.72 [-2.60; 1.70]
G2	Homens (n=264)			
21-30	Amplitude*	M	DP	Md z, [z]
N	[0, 23]	9.41	5.08	-0.26 [-1.89; 2.27]
E	[0, 20]	13.08	4.73	0.31 [-2.62; 1.67]
P	[0, 9]	1.33	1.50	-0.01 [-0.78; 5.35]
L	[0, 18]	8.45	3.38	-0.45 [-2.60; 1.96]
G3	Homens (n=195)			
31-40	Amplitude*	M	DP	Md z, [z]
N	[0, 23]	8.80	4.96	-0.44 [-1.89; 1.91]
E	[0, 20]	11.64	4.60	-0.14 [-2.85; 1.67]
P	[0, 9]	1.11	1.42	-0.01 [-0.78; 3.82]
L	[0, 18]	9.06	3.88	-0.18 [-2.60; 2.23]
G4	Homens (n=123)			
41-50	Amplitude*	M	DP	Md z, [z]
N	[0, 23]	8.93	5.48	-0.44 [-1.89; 2.27]
E	[0, 20]	12.44	4.35	0.09 [-2.62; 1.67]
P	[0, 9]	0.97	1.27	-0.77 [-0.78; 3.82]
L	[0, 18]	10.72	3.37	0.08 [-2.33; 2.23]
G5	Homens (n=79)			
51-60	Amplitude*	M	DP	Md z, [z]
N	[0, 23]	9.11	5.97	-0.44 [-1.89; 1.73]
E	[0, 20]	12.44	4.72	0.09 [-2.62; 1.44]
P	[0, 9]	0.70	0.85	-0.77 [-0.78; 1.52]
L	[0, 18]	12.38	3.10	0.89 [-1.53; 2.23]

Nota. N=participantes, M=média, DP=desvio-padrão. G1=16-20 anos, G2=21-30 anos, G3=31-40 anos, G4=41-50 anos, G5=51-60 anos. Md z=mediana dos z-scores, [z]=amplitude dos z-scores, * amplitude de resultados possíveis para cada escala.

Calculou-se o correspondente coeficiente *d* de Cohen para estimar a dimensão do efeito dos testes *t* (Cohen, 1988).

Na variável género, os resultados da ANOVA foram significativos ($p<.001$) para as escalas N [$F(1688)=57.847$], P [$F(1688)=51.589$] e L [$F(1688)=22.973$], permitindo rejeitar a hipótese nula (H_0) e assumir que existem diferenças significativas entre os grupos comparados (H_1). Na escala E não foram registadas diferenças. Através do teste *t* foram encontradas diferenças significativas ($p<.001$) nas escalas N [$t(1562.271)=-6.982$; $d=$

-0.35], P [$t(1564)=7.367$; $d=0.37$], e L [$t(1562.303) = -5.241$; $d=-0.27$] (diferenças de efeito moderado) entre os homens (H) e as mulheres (M).

Na variável grupo etário, os resultados da ANOVA foram significativos ($p<.001$) para as escalas N [$F(1688) = 4.858$], E [$F(1688)=5.602$], P [$F(1688)=13.123$] e L [$F(1688)=58.766$], permitindo rejeitar a hipótese nula (H_0) e assumir que existem diferenças significativas entre os grupos comparados (H_1). Com o teste *t* [G1 (16-20 anos), G2 (21-30 anos), G3 (31-40 anos), G4 (41-50 anos), G5 (51-60 anos)], registaram-se diferenças

Quadro 5. Dados normativos do EPQ-R para a população portuguesa (ANN), mulheres e grupo etário

Total	Mulheres (n=906)			
	Amplitude*	M	DP	Md z, [z]
N	[0, 23]	11.37	5.55	0.10 [-1.89; 2.27]
E	[0, 20]	12.57	4.24	0.09 [-2.62; 1.67]
P	[0, 9]	0.81	1.12	-0.77 [-0.78; 5.35]
L	[0, 18]	10.08	3.64	0.08 [-2.60; 2.23]
G1	Mulheres (n=143)			
16-20	Amplitude*	M	DP	Md z, [z]
N	[0, 23]	12.36	5.26	0.28 [-1.71; 2.09]
E	[0, 20]	13.14	4.42	0.09 [-2.17; 1.67]
P	[0, 9]	1.00	1.23	-0.01 [-0.78; 3.82]
L	[0, 18]	8.57	3.54	-0.18 [-2.60; 1.70]
G2	Mulheres (n=304)			
21-30	Amplitude*	M	DP	Md z, [z]
N	[0, 23]	11.65	5.58	0.10 [-1.89; 2.27]
E	[0, 20]	12.67	4.20	0.09 [-2.62; 1.67]
P	[0, 9]	0.84	1.20	-0.77 [-0.78; 5.35]
L	[0, 18]	9.15	3.57	-0.18 [-2.60; 2.23]
G3	Mulheres (n=235)			
31-40	Amplitude*	M	DP	Md z, [z]
N	[0, 23]	10.66	5.51	0.10 [-1.89; 2.27]
E	[0, 20]	12.66	3.96	0.09 [-1.94; 1.67]
P	[0, 9]	0.77	1.04	-0.77 [-0.78; 3.05]
L	[0, 18]	10.53	3.25	0.35 [-2.33; 2.23]
G4	Mulheres (n=136)			
41-50	Amplitude*	M	DP	Md z, [z]
N	[0, 23]	11.17	5.70	0.10 [-1.89; 2.27]
E	[0, 20]	12.61	4.48	0.09 [-2.62; 1.67]
P	[0, 9]	0.74	1.06	-0.77 [-0.78; 3.82]
L	[0, 18]	11.45	3.13	0.35 [-1.79; 2.23]
G5	Mulheres (n=88)			
51-60	Amplitude*	M	DP	Md z, [z]
N	[0, 23]	11.03	5.56	-0.08 [-1.71; 2.09]
E	[0, 20]	10.94	4.15	-0.36 [-2.62; 1.44]
P	[0, 9]	0.57	0.80	-0.77 [-0.78; 1.52]
L	[0, 18]	12.48	3.69	0.89 [-2.33; 2.23]

Nota. N=participantes, M=média, DP=desvio-padrão. G1=16-20 anos, G2=21-30 anos, G3=31-40 anos, G4=41-50 anos, G5=51-60 anos. Md z=mediana dos z-scores, [z]=amplitude dos z-scores, * amplitude de resultados possíveis para cada escala.

significativas ($p<.001$) na escala N na comparação do G1 com: G3 [$t(526.807)=4.606$; $d=0.40$], G4 [$t(519.710)=3.077$; $d=0.27$], G5 [$t(326.575)=3.157$; $d = 0.35$] (efeito moderado); e entre G2 e G3 [$t(857.151)=2.701$; $d=0.18$] (efeito pequeno). Na escala E, registaram-se diferenças significativas ($p<.001$) na comparação do G1 com: G3 [$t(527.882)=2.630$; $d=0.23$], G5 [$t(331.985)=2.885$; $d=0.32$]; e entre G2 e G5 [$t(331.673)=2.778$; $d=0.31$] (efeito moderado). Na escala P, registaram-se diferenças significativas ($p<.001$) na comparação do G1 com: G2 [$t(528)=3.020$; $d=0.26$], G3 [$t(528)=$

4.559; $d=0.40$], G4 [$t(522)=5.080$; $d=0.44$] (efeito moderado), G5 [$t(332)=6.529$; $d=0.72$] (efeito elevado); e entre G2 e G5 [$t(332)=3.426$; $d=0.38$] (efeito moderado). Na escala L, registaram-se diferenças significativas ($p<.001$) na comparação entre todos os grupos etários, exceto entre G1 e G2. No G1, existem diferenças em relação ao: G3 [$t(527.806)=-5.532$; $d=-0.48$] (efeito moderado), G4 [$t(520.221)=-10.052$; $d=-0.88$] (efeito elevado), G5 [$t(331.979)=-12.052$; $d=-1.32$] (efeito muito elevado). No G2, existem diferenças em relação ao: G3 [$t(852.838)=-4.485$; $d=-0.31$]

(efeito moderado), G4 [$t(515.137)=-8.457$; $d=-0.75$] (efeito elevado), G5 [$t(331.799)=-9.544$; $d=-1.05$] (efeito muito elevado). No G3, existem diferenças em relação ao: G4 [$t(511.065)=-4.107$; $d=-0.36$] (efeito moderado), G5 [$t(331.380)=-6.302$; $d=-0.69$] (efeito elevado). Entre G4 e G5 também existem diferenças [$t(330.641)=-4.050$; $d=-0.45$] (efeito moderado). Nos restantes grupos, para N, E, P e L, não foram encontradas diferenças.

No que respeita ainda à variável idade, foram constituídos dois grandes grupos para averiguar a existência de diferenças estatisticamente significativas. Um grupo de sujeitos mais jovens (GJ, incluindo adolescentes e jovens adultos), 16-30 anos ($n=833$, 386 homens e 447 mulheres, 49.32% da ANN) [dimensão N, $M=10.92$ ($DP=5.48$); dimensão E, $M=13.04$ ($DP=4.45$); dimensão P, $M=1.19$ ($DP=1.43$); escala L, $M=8.60$ ($DP=3.52$)]. E um grupo de sujeitos adultos (GA, incluindo adultos e adultos de meia-idade), 31-60 anos ($n=856$, 397 homens e 459 mulheres, 50.68% da ANN) [dimensão N, $M=9.96$ ($DP=5.54$); dimensão E, $M=12.19$ ($DP=4.37$); dimensão P, $M=0.84$ ($DP=1.15$); escala L, $M=10.74$ ($DP=3.61$)]. Os resultados obtidos mostram que existem diferenças estatísticas significativas ($p<.001$) em todas as dimensões (N, E, P e L) entre os sujeitos mais jovens e os sujeitos adultos: [$t(1686.593)=3.583$; $d=0.17$] para N, [$t(1683.418)=3.945$; $d=0.19$] para E (efeito pequeno), [$t(1687)=5.475$; $d=0.27$] para P (efeito moderado), [$t(1686.992)=-12.331$; $d=-0.60$] para L (efeito elevado).

Nesta pesquisa, foi ainda calculado o coeficiente de correlação (r de Pearson) entre a idade e o constructo de desajustabilidade social medido pela escala L do EPQ-R. A correlação obtida é positiva, moderada e estatisticamente significativa ($.35$, $p<.01$).

Análise do Funcionamento Diferencial dos Itens (DIF) do EPQ-R

No âmbito da TRI, foi efetuada uma análise DIF em relação às variáveis género (homens e mulheres) e idade (GJ, grupo de sujeitos mais jovens, incluindo adolescentes e jovens adultos, e GA, grupo de adultos, incluindo adultos e adultos de meia-idade) para verificar se o princípio da invariância é cumprido na avaliação dos constructos N, E, P e L. De acordo com este princípio fundamental, que se baseia no índice de

dificuldade dos itens (medido em unidades *logit*), a qualidade da medida deve ser a mesma independentemente das características do grupo a que pertence o sujeito, como o género ou a idade. Neste sentido, os sujeitos com o mesmo nível no atributo (os constructos N, E, P e L, nos níveis baixo, médio, e elevado) devem ter a mesma probabilidade de “responder acertadamente” a um item, seja homem ou mulher, seja jovem ou adulto.

No contexto da avaliação da personalidade, a análise em termos de “acertos” e “erros” inerente à estimação do índice de dificuldade dos itens (modelo logístico de um parâmetro), deve ser interpretada do seguinte modo: quando um item é cotado com 1 ponto (“acerto”), significa que o sujeito está de acordo com seu conteúdo, isto é, que o traço de personalidade avaliado pelo item está presente naquele sujeito; ao invés, quando um item é cotado com 0 pontos (“erro”), indica que o sujeito está em desacordo com o seu conteúdo, ou seja, que o traço avaliado pelo item não está presente naquele sujeito.

Os enviesamentos produzidos pelo DIF remetem para um conjunto de erros que colocam em causa a validade de constructo da medida. Portanto, a qualidade da medida é estimada pela ausência de DIF (Aliste, 1996; Hambleton & Swaminathan, 1985).

Para detetar DIF, existem dois critérios a considerar em simultâneo (Linacre, 2009): 1) um item tem DIF quando o valor absoluto do *contraste DIF* entre os grupos avaliados $|D1-D2| \geq 0.50$ e o valor *t de Welch* é significativo, aplicando o *critério de Bonferroni* [e.g., a escala N tem 23 itens, então *t de Welch* é significativo se $p<(0.05/23)=0.0022$]; 2) um item tem DIF quando o valor do *teste de Mantel-Haenszel* (MH) é significativo para o Tipo C, aplicando o *critério de Bonferroni*, ou seja, quando o Delta-MH é superior a 1.5 ($|DIF| > |1.5|$) [este valor é obtido pela multiplicação de $-2.35 \ln$ pelo valor absoluto do *size* (de DIF) do teste MH; 1 unidade *logit* no Modelo de Rasch corresponde a 2.35 unidades *delta* no sistema do Educational Testing Service (ETS)]. Então, de acordo com o ETS, existem três tipos de DIF (Linacre, 2009; Zwick, & Ercikan, 1989; Zwick, Thayer, & Lewis, 1999): 1) DIF Tipo A, sem significado, quando $|DIF| < |1|$; 2) DIF Tipo B, ligeiro, também sem significado,

quando $1 \leq |DIF| \leq |1.5|$, o que corresponde a $MH=0.43$; 3) DIF Tipo C, severo, quando $|DIF| > |1.5|$, o que corresponde a $MH=0.64$.

No que concerne à variável género (t de Welch e Delta-MH, $p < .001$), foram registados os seguintes itens com DIF Tipo C (severo). Na dimensão N (um em 23 itens), o item 9 (“*Preocupa-se frequentemente com coisas que não deveria ter feito ou dito?*”) [$|DIF|=0.73$, $t=5.96$, $MH=0.66$], mais “difícil” para as mulheres. Na dimensão E (três em 20 itens), o item 4 (“*É uma pessoa faladora?*”) [$|DIF|=0.68$, $t=5.51$, $MH=0.74$] e o item 45 (“*É capaz de organizar e animar uma festa?*”) [$|DIF|=0.64$, $t=5.23$, $MH=0.71$], ambos mais “difíceis” para os homens, e o item 1 (“*Tem muitos passatempos diferentes?*”) [$|DIF|=1.03$, $t=8.80$, $MH=0.92$], mais “difícil” para as mulheres. Na escala L (um em 18 itens), o item 30 (“*Alguma vez disse mal de alguém?*”) [$|DIF|=0.63$, $t=4.17$, $MH=0.72$], mais “difícil” para as mulheres. Com DIF Tipo B (ligeiro), surgem: na dimensão N, dois itens, um mais “difícil” para os homens e outro mais “difícil” para as mulheres; na dimensão E, dois itens, ambos mais “difíceis” para as mulheres; na dimensão P, dois itens, um mais “difícil” para os homens e outro mais “difícil” para as mulheres; na escala L, três itens, dois mais “difíceis” para os homens e um mais “difícil” para as mulheres.

No que respeita à variável idade (t de Welch e Delta-MH, $p < .001$), foram registados os seguintes itens com DIF Tipo C (severo). Na dimensão N (um em 23 itens), o item 21 (“*É uma pessoa preocupada?*”) [$|DIF|=0.92$, $t=6.69$, $MH=0.79$], mais “difícil” para o grupo de sujeitos mais jovens. Na dimensão E (dois em 20 itens), o item 16 (“*Em encontros sociais prefere ficar em segundo plano?*”) [$|DIF|=0.73$, $t=6.20$, $MH=0.71$] e o item 64 (“*Gosta de muita excitação e animação à sua volta?*”) [$|DIF|=0.65$, $t=5.62$, $MH=0.67$], ambos mais “difíceis” para o grupo de adultos. Na dimensão P (um em 9 itens), item 35 (“*Costuma sentir prazer em ver alguém com quem não simpatiza ser humilhado perante outras pessoas?*”) [$|DIF|=0.83$, $t=5.22$, $MH=0.85$], mais “difícil” para o grupo de adultos. Na escala L (um em 18 itens), o item 46 (“*Já alguma vez insistiu em impor a sua vontade?*”) [$|DIF|=0.89$, $t=7.17$, $MH=0.80$], mais “difícil” para o grupo de adultos. Com DIF Tipo B (ligeiro), surgem: na dimensão

N, dois itens, ambos mais “difíceis” para o grupo de adultos; na dimensão E, um item, mais “difícil” para o grupo de sujeitos mais jovens; na escala L, três itens, mais “difíceis” para o grupo de sujeitos mais jovens.

Discussão

A importância do EPQ-R na avaliação da personalidade é reconhecida internacionalmente. Este instrumento apresenta boas qualidades psicométricas, um vasto suporte empírico, que foi desenvolvido ao longo de mais de cinco décadas de investigação (H. Eysenck e colaboradores), e uma comprovada aplicabilidade nos diversos contextos de intervenção psicológica (nomeadamente, nos contextos da comunidade e nos contextos clínicos e forenses) (cf. Furnham et al., 2008; Nyborg, 1997). No que respeita às qualidades psicométricas, o EPQ-R tem evidenciado índices de precisão e validade adequados em mais de três dezenas de países (cf. Barrett et al., 1998; Furnham et al., 2008), incluindo Portugal (Almiro et al., 2016). Os estudos psicométricos do EPQ-R sugerem que o instrumento é adequado e robusto na avaliação dos constructos de Neuroticismo, Extroversão, Psicoticismo, e Mentira/Desejabilidade Social no contexto português (cf. Almiro, 2013; Almiro & Simões, 2014).

No presente estudo de aferição, com recurso a uma amostra extensa e representativa da população portuguesa (ANN, $N = 1689$), foram estabelecidas as normas para a interpretação das pontuações do EPQ-R, considerando as médias (M) e os desvios-padrão (DP), em função das variáveis género e idade. Neste âmbito, relativamente à variável género, verificou-se que existem diferenças estatísticas significativas (*teste t*) nas pontuações médias obtidas no EPQ-R entre homens ($n=783$) e mulheres ($n=906$). Por um lado, as mulheres tendem a pontuar mais na dimensão N e na escala L do que os homens. Ou seja, de uma forma geral, as mulheres têm tendência para apresentar uma personalidade mais instável (personalidade neurótica), evidenciando traços mais intensos de emotividade, ansiedade, depressão, excesso de preocupação, irritabilidade fácil, baixa autoestima, entre outros traços. De um modo geral, as mulheres têm também mais

tendência para se comportarem de acordo com o que consideram ser socialmente aceite ou tido como mais correto (traço de desejabilidade social). Por outro lado, os homens tendem a pontuar mais na dimensão P do que as mulheres. Isto é, de uma forma geral, os homens têm mais tendência para apresentar uma personalidade com características de elevado psicoticismo, evidenciando traços mais intensos de hostilidade, agressividade, impulsividade, desconfiança, baixa empatia, entre outros traços. Nas diferenças encontradas relativamente ao género, a dimensão do efeito (d de Cohen) é apenas moderada. Estes resultados confirmam parcialmente a hipótese H1 (existem diferenças estatísticas significativas nas pontuações do EPQ-R em relação à variável género), uma vez que foram encontradas diferenças entre homens e mulheres nas escalas N, P e L, mas não na escala E. Os dados do presente estudo são similares aos encontrados por H. Eysenck e colaboradores na amostra inglesa de referência (S. Eysenck et al., 1985).

No que se refere à variável idade (entre 16 e 60 anos) foram reportadas diferenças estatísticas significativas (*teste t*) nas pontuações médias obtidas no EPQ-R entre os diversos grupos etários (16-20 anos, 21-30 anos, 31-40 anos, 41-50 anos, 51-60 anos) e entre os dois grandes grupos de jovens e adultos (GJ, 16-30 anos, $n=833$; GA, 31-60 anos, $n = 856$). Estes resultados mostram, de um modo global, que os sujeitos mais jovens tendem a pontuar mais nas dimensões N, E e P, evidenciando traços mais intensos nessas mesmas dimensões. Em N, traços mais intensos de emotividade, ansiedade, depressão, entre outros traços (personalidade instável). Em E, traços mais intensos de sociabilidade, atividade, assertividade, espontaneidade, espírito de aventura, otimismo, entre outros traços (personalidade extrovertida). Em P, traços mais intensos de hostilidade, impulsividade, baixa empatia, entre outros traços (elevado psicoticismo). Por seu turno, os sujeitos adultos tendem, de uma forma geral, a pontuar mais na escala L, manifestando assim uma maior propensão para se comportarem de acordo com o que consideram ser socialmente aceite ou tido como mais correto (traço de desejabilidade social). Ao nível dos grupos etários, a intensidade da desejabilidade social aumenta sistematicamente em função da idade (com uma diferença cada vez

mais acentuada). O grupo com menor desejabilidade social é o G1 (16-20 anos) e o grupo com maior desejabilidade social é o G5 (51-60 anos). A correlação de .35 ($p<.01$) obtida entre a idade e a desejabilidade social reflete a associação positiva e moderada que existe entre as duas variáveis. Nas diferenças encontradas relativamente à idade, a dimensão do efeito (d de Cohen) é geralmente pequena ou moderada nas dimensões N, E e P, e oscila entre moderada e muito elevada na escala L. Estes resultados confirmam a hipótese H2 (existem diferenças estatísticas significativas nas pontuações do EPQ-R em relação à variável idade). Os dados do presente estudo vão no mesmo sentido dos dados reportados no estudo original inglês (S. Eysenck et al., 1985) e mostram a importância acrescida da utilização de normas específicas para a idade na interpretação das pontuações no EPQ-R.

Na análise DIF, para a variável género, foram encontrados cinco itens com DIF severo e nove com DIF ligeiro (sem significado). Em relação à dimensão N, no item 9 (“*Preocupa-se frequentemente com coisas que não deveria ter feito ou dito?*”), mais “difícil” para as mulheres), com DIF severo, os homens com um nível menor de Neuroticismo (personalidade estável) tendem mais facilmente a estar de acordo com o conteúdo deste item do que as mulheres. Todavia, as mulheres tendem a pontuar mais na dimensão N (23 itens, *teste t*) do que os homens, por isso, o enviesamento produzido não prejudica a qualidade da medida (dos dois itens com DIF ligeiro, um é mais “difícil” para os homens e outro mais “difícil” para as mulheres, o que também contribui para o equilíbrio deste enviesamento). Contudo, a diferença no atributo entre grupos (*teste t*) e o viés na avaliação do atributo em função da pertença a esses grupos (DIF) não devem ser confundidos. Na dimensão E, nos itens 4 (“*É uma pessoa faladora?*”), mais “difícil” para os homens) e 45 (“*É capaz de organizar e animar uma festa?*”), mais “difícil” para os homens), ambos com DIF severo, as mulheres com um nível menor de Extroversão (personalidade introvertida) tendem mais facilmente a estar de acordo com o conteúdo destes itens do que os homens, e no item 1 (“*Tem muitos passatempos diferentes?*”), mais “difícil” para as mulheres), também com DIF severo, os homens com um

nível menor de Extroversão (personalidade introvertida) tendem mais facilmente a estar de acordo com o conteúdo deste item do que as mulheres. Relativamente a esta dimensão, não há diferenças significativas entre as pontuações dos homens e das mulheres (*teste t*), pelo que o enviesamento produzido por estes itens na avaliação dos sujeitos é maior em relação aos homens (dois itens) do que às mulheres (um item). No entanto, os dois itens com DIF ligeiro são ambos mais “difíceis” para as mulheres, o que de certa forma compensa esta proporção e equilibra o enviesamento produzido, não prejudicando a qualidade da medida. Na dimensão P, dos dois itens com DIF ligeiro, sem significado, um é mais “difícil” para os homens e outro é mais “difícil” para as mulheres, ficando assim equilibrado o pequeno enviesamento da medida. Na escala L, no item 30 (“*Alguma vez disse mal de alguém?*”, mais “difícil” para as mulheres), com DIF severo, os homens com menor desejabilidade social tendem mais facilmente a pontuar neste item (respondendo “Não”, item de cotação inversa) do que as mulheres. Contudo, as mulheres tendem a pontuar mais nesta escala (18 itens, *teste t*) do que os homens, por isso, o enviesamento produzido não prejudica a qualidade da medida (dos três itens com DIF ligeiro, dois são mais “difíceis” para os homens e um é mais “difícil” para as mulheres, o que também contribui para o equilíbrio deste enviesamento).

Ainda no âmbito da análise DIF, no que se refere à variável idade, foram encontrados cinco itens com DIF severo e seis com DIF ligeiro (sem significado). Na dimensão N, no item 21 (“*É uma pessoa preocupada?*”, mais “difícil” para o grupo de sujeitos mais jovens), com DIF severo, os sujeitos do grupo de adultos com um nível menor de Neuroticismo (personalidade estável) tendem mais facilmente a estar de acordo com o conteúdo deste item do que os sujeitos mais jovens. No entanto, os sujeitos mais jovens tendem a pontuar mais na dimensão N (23 itens, *teste t*) do que os sujeitos adultos, por isso, o enviesamento produzido não prejudica a qualidade da medida (os dois itens com DIF ligeiro são ambos mais “difíceis” para o grupo de adultos, o que contribui igualmente para o equilíbrio deste enviesamento). Na dimensão E, o item 16 (“*Em encontros sociais prefere ficar em segundo plano?*”, mais “difícil”

para o grupo de adultos) e o item 64 (“*Gosta de muita excitação e animação à sua volta?*”, mais “difícil” para o grupo de adultos) possuem ambos DIF severo. Com efeito, os sujeitos mais jovens com um nível menor de Extroversão (personalidade introvertida) tendem mais facilmente a pontuar no item 16 (respondendo “Não”, item de cotação inversa) e a estar de acordo com o conteúdo do item 64 do que os sujeitos adultos. A par disso, os sujeitos mais jovens tendem a pontuar mais nesta dimensão do que os sujeitos adultos (*teste t*), pelo que o DIF destes itens produz um enviesamento a considerar na avaliação dos sujeitos adultos, ainda que a dimensão E seja composta por 20 itens (o único item com DIF ligeiro é mais “difícil” para o grupo de sujeitos mais jovens, o que contribui para reduzir este enviesamento). Na dimensão P, no item 35 (“*Costuma sentir prazer em ver alguém com quem não simpatiza ser humilhado perante outras pessoas?*”, mais “difícil” para o grupo de adultos), com DIF severo, os sujeitos mais jovens com um menor nível de Psicoticismo (personalidade normal) tendem mais facilmente a estar de acordo com o conteúdo deste item do que os sujeitos adultos. Além disso, os sujeitos mais jovens tendem a pontuar mais na dimensão P do que os sujeitos adultos (*teste t*), pelo que o enviesamento produzido por estes itens no exame dos sujeitos adultos é importante, sobretudo porque esta dimensão possui apenas 9 itens. Na escala L, no item 46 (“*Já alguma vez insistiu em impor a sua vontade?*”, mais “difícil” para o grupo de sujeitos adultos), com DIF severo, os sujeitos mais jovens com menor desejabilidade social tendem mais facilmente a pontuar neste item (respondendo “Não”, item de cotação inversa) do que os sujeitos adultos. Contudo, os sujeitos adultos tendem a pontuar mais na escala L do que os sujeitos mais jovens (*teste t*), pelo que o enviesamento produzido não prejudica a qualidade da medida (os três itens com DIF ligeiro são todos mais “difíceis” para o grupo de sujeitos mais jovens, o que equilibra este enviesamento).

Por conseguinte, os resultados obtidos na análise DIF evidenciam apenas cinco itens com DIF severo em relação à variável género e cinco itens com DIF severo em relação à variável idade (itens mais “difíceis” para um determinado grupo, alternadamente) num total de 70 itens do

instrumento, cujo enviesamento é na generalidade compensado pelo DIF ligeiro encontrado noutros itens da mesma escala (exceção para um item na escala P). De um modo geral, os resultados permitem confirmar as hipóteses H3 (o funcionamento diferencial dos itens do EPQ-R cumpre o princípio da invariância para a variável género) e H4 (o funcionamento diferencial dos itens do EPQ-R cumpre o princípio da invariância para a variável idade). Neste sentido, os dados da análise DIF são demonstrativos da qualidade dos itens do EPQ-R na medição dos constructos N, E, P e L, tanto em homens como em mulheres, tanto em sujeitos mais jovens como em sujeitos adultos, e da adequação deste instrumento para a avaliação das características emocionais e comportamentais da população portuguesa. Ainda assim, a relação entre a idade e a desejabilidade social deve ser tida em consideração na interpretação das suas pontuações.

O presente estudo apresenta algumas limitações a apontar. Nomeadamente, não foram considerados grupos etários acima dos 60 anos, o que permitiria estabelecer normas de interpretação para os resultados do EPQ-R adaptadas às pessoas mais velhas. Nesta pesquisa, também não foi estudada a influência da variável escolaridade na personalidade. Por último, embora a ANN seja extensa e representativa da população portuguesa, a Área Metropolitana da Grande Lisboa e as zonas do Algarve e da Madeira não atingiram o número de participantes previsto para a amostra teórica definida.

Referências

- Aliste, A. M. (1996). Funcionamiento diferencial de los items. In J. Muñoz (Ed.), *Psicometría* (pp.371-455). Madrid: Editorial Universitas.
- Almeida, L.S., & Freire, T. (2017). *Metodologia da investigação em psicologia e educação* (5ª ed.). Braga: Psiquilibrios.
- Almiro, P. A. (2013). *Adaptação, validação e aferição do EPQ-R para a população portuguesa: Estudos em contextos clínico, forense e na comunidade*. Dissertação de Doutoramento, Faculdade de Psicologia e Ciências da Educação da Universidade de Coimbra, Coimbra.
- Almiro, P. A., Marques-Costa, C., & Simões, M. R. (2015). Questionário de Personalidade de Eysenck – Forma Revista (EPQ-R). In M. R. Simões, I. Santana, & Grupo de Estudos de Envelhecimento Cerebral e Demência (Eds.), *Escala e testes na demência* (3ª ed., pp.206-213). Porto Salvo: Novartis.
- Almiro, P. A., Moura, O., & Simões, M. R. (2016). Psychometric properties of the European Portuguese version of the Eysenck Personality Questionnaire – Revised (EPQ-R). *Personality and Individual Differences*, 88, 88-93. doi:10.1016/j.paid.2015.08.050
- Almiro, P. A., & Simões, M. R. (2011). Questionário de Personalidade de Eysenck – Forma Revista (EPQ-R): Breve revisão dos estudos de validade concorrente. *Revista Portuguesa de Psicologia*, 42, 101-120. doi:10.21631/rpp42_101
- Almiro, P. A., & Simões, M. R. (2013). *Manual da versão portuguesa do Questionário de Personalidade de Eysenck – Forma Revista (EPQ-R)*. Coimbra: Laboratório de Avaliação Psicológica e Psicometria / Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra.
- Almiro, P. A., & Simões, M. R. (2014). Questionário de Personalidade de Eysenck – Forma Revista (EPQ-R). In L.S. Almeida, M. R. Simões, & M. M. Gonçalves (Eds.), *Instrumentos e contextos de avaliação psicológica – Vol. II* (pp.211-229). Coimbra: Edições Almedina.
- Almiro, P. A., & Simões, M. R. (2015). Estudo da versão portuguesa do EPQ-R: Uma aplicação do Modelo de Rasch. In H.F. Golino, C.M.A., Gomes, A.A.N. Ribeiro, & G. Coelho (Eds.), *Psicometria contemporânea: Compreendendo os Modelos Rasch* (pp.241-281). São Paulo: Editora Casa do Psicólogo/Pearson.
- Almiro, P. A., & Simões, M. R. (2016). Estudo das propriedades psicométricas da Versão Experimental Portuguesa do Questionário de Personalidade de Eysenck – Forma Revista (EPQ-R). *Revista Iberoamericana de Diagnóstico y Evaluación – e Avaliação Psicológica*, 41, 159-173.
- American Educational Research Association, American Psychological Association, National Council on Measurement in

- Education (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Barrett, P. T., & Eysenck, S. B. (1984). The assessment of personality factors across 25 countries. *Personality and Individual Differences*, 5, 615-632. doi:10.1016/0191-8869(84)90110-7
- Barrett, P. T., Petrides, K. V., Eysenck, S. B., & Eysenck, H. J. (1998). The Eysenck Personality Questionnaire: An examination of the factorial similarity of P, E, N, and L across 34 countries. *Personality and Individual Differences*, 26, 805-819. doi:10.1016/S0191-8869(98)00026-9
- Castro Fonseca, A., Eysenck, S. B., & Simões, A. (1991). Um estudo intercultural da personalidade: Comparação de adultos portugueses e ingleses no EPQ. *Revista Portuguesa de Pedagogia*, 25, 187-203.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.) Hillsdale, NJ: Lawrence Erlbaum Associates.
- Dazzi, C. (2011). The Eysenck Personality Questionnaire – Revised (EPQ-R): A confirmation of the factorial structure in the Italian context. *Personality and Individual Differences*, 50, 790-794. doi:10.1016/j.paid.2010.12.032
- DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). Thousand Oaks, CA: SAGE Publications.
- Eysenck, H. J., & Eysenck, M.W. (1985). *Personality and individual differences: A natural science approach*. New York: Plenum Press.
- Eysenck, H. J., & Eysenck, S. B. (1975). *Manual of the Eysenck Personality Questionnaire (Junior & Adult)*. London: Hodder & Stoughton.
- Eysenck, H. J., & Eysenck, S. B. (2008). *EPQ-R: Cuestionario Revisado de Personalidad de Eysenck: Versiones completa (EPQ-R) y abreviada (EPQ-RS)* (3^a ed.). Madrid: TEA.
- Eysenck, S. B., & Barrett, P. (2013). Re-introduction to cross-cultural studies of the EPQ. *Personality and Individual Differences*, 54, 485-489. doi:10.1016/j.paid.2012.09.022
- Eysenck, S. B., Eysenck, H. J., & Barrett, P. T. (1985). A revised version of the Psychoticism scale. *Personality and Individual Differences*, 6, 21-29. doi:10.1016/0191-8869(85)90026-1
- Furnham, A., Eysenck, S.B., & Saklofske, D.H. (2008). The Eysenck personality measures: Fifty years of scale development. In G.J. Boyle, G. Matthews, & D. H. Saklofske (Eds.), *The SAGE handbook of personality theory and assessment – Vol. 2: Personality measurement and testing* (pp.199-218). London: SAGE.
- Geisinger, K. F. (1992). The metamorphosis of test validation. *Educational Psychologist*, 27, 197-222. doi:10.1207/s15326985ep2702_5
- Geisinger, K. F. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological Assessment*, 6(4), 304-312. doi:10.1037/1040-3590.6.4.304
- Hambleton, R. J., & Swaminathan, H. (1985). *Item Response Theory: Principles and applications*. Boston: Kluwer-Nijhoff Publishing.
- Hambleton, R. J., Merenda, P. F., & Spielberger, C.D. (Eds.) (2005). *Adapting educational and psychological tests for cross-cultural assessment*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55. doi:10.1080/10705519909540118
- Instituto Nacional de Estatística (2012). *Censos 2011: Resultados definitivos do XV Recenseamento Geral da População e do V Recenseamento Geral da Habitação* (8 Vols.). Lisboa: Instituto Nacional de Estatística.
- International Test Commission (2017). ITC guidelines for translating and adapting tests (2nd ed.). Retrieved from: https://www.intestcom.org/files/guideline_test_adaptation_2ed.pdf
- Linacre, J. M. (2009). *A user's guide to WINSTEPS 3.69.0: Rasch-Model computer program*. Chicago: Winsteps.com.
- Lynn, R. (Ed.) (1981). *Dimensions of personality: Papers in honour of H.J. Eysenck*. Oxford: Pergamon Press.

- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*, 741-749.
doi:10.1037/0003-066X.50.9.741
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
- Nyborg, H. (Ed.) (1997). *The scientific study of human nature: Tribute to Hans J. Eysenck at eighty*. Oxford: Pergamon Press.
- Ortet, G., Ibáñez, M. I., Moro, M., Silva, F., & Boyle, G. J. (1999). Psychometric appraisal of Eysenck's revised Psychoticism scale: A cross-cultural study. *Personality and Individual Differences, 27*, 1209-1219.
doi:10.1016/S0191-8869(99)00065-3
- Pires, R., Ferreira, A. S., Guedes, D., Gonçalves, B., & Henriques-Calado, J. (2018). Estudo das propriedades psicométricas – Formas Longa, Reduzida e Breve – da Versão Portuguesa do Inventário da Personalidade para o DSM-5 (PID-5). *Revista Iberoamericana de Diagnóstico y Evaluación – e Avaliação Psicológica, 47*(2), 197-212.
doi:10.21865/RIDEP47.2.14
- Primi, R. (2004). Avanços na interpretação de escalas com a aplicação da Teoria de Resposta ao Item. *Avaliação Psicológica, 3*(1), 53-58.
- Ruch, W. (1999). Die revidierte Fassung des Eysenck Personality Questionnaire und die Konstruktion des deutschen EPQ-R bzw. EPQ-RK. *Zeitschrift für Differentielle und Diagnostische Psychologie, 20*, 1-24.
doi:10.1024//0170-1789.20.1.1
- San Martini, P., Mazzotti, E., & Setaro, S. (1996). Factor structure and psychometric features of Italian version for the EPQ-R. *Personality and Individual Differences, 21*, 877-882.
doi:10.1016/S0191-8869(96)00147-X
- Zwick, R., & Ercikan, K. (1989). Analysis of differential item functioning in the NAEP history assessment. *Journal of Educational Measurement, 26*, 55-66.
doi:10.1111/j.1745-3984.1989.tb00318.x
- Zwick, R., Thayer, D. T., & Lewis, C. (1999). An empirical Bayes approach to Mantel-Haenszel DIF analysis. *Journal of Educational Measurement, 36*, 1-28.