

Estudo das Propriedades Psicométricas da Versão Experimental Portuguesa do Questionário de Personalidade de Eysenck – Forma Revista (EPQ-R)

Study of the Psychometric Properties of the Portuguese Experimental Version of the Eysenck Personality Questionnaire – Revised (EPQ-R)

Pedro Armelim Almiro¹ e Mário R. Simões²

Resumo

Com base no Modelo P-E-N, o *Questionário de Personalidade de Eysenck – Forma Revista* (EPQ-R; S. Eysenck, H. Eysenck, & Barrett, 1985) avalia as três dimensões fundamentais da personalidade: *Psicoticismo* (P), *Extroversão* (E) e *Neuroticismo* (N). Este instrumento inclui também a escala L (*Lie*), que consiste numa escala de *Mentira/Desejabilidade Social*. No presente artigo, serão apresentados os resultados dos estudos psicométricos da Versão Experimental Portuguesa do EPQ-R (Almiro & M.R. Simões, 2007) obtidos numa amostra de 671 participantes. Estes estudos incluem: exame da precisão (método da consistência interna e teste-reteste); exame da validade de constructo (análise em componentes principais) e da validade concorrente (utilizando os resultados no BDI, STAI e BSI como critério de validação externo); e exame da qualidade dos itens (ajuste, precisão, validade) com recurso ao Modelo de Rasch. Os resultados evidenciaram a adequabilidade e as boas qualidades psicométricas da Versão Experimental do EPQ-R.

Palavras-chave: versão experimental portuguesa do EPQ-R, Hans Eysenck, avaliação da personalidade, teoria clássica dos testes, teoria da resposta ao item

Abstract

Based on the P-E-N Model, the Eysenck Personality Questionnaire – Revised (EPQ-R; S. Eysenck, H. Eysenck, & Barrett, 1985) assesses the three fundamental personality dimensions: *Psychoticism* (P), *Extraversion* (E), and *Neuroticism* (N). This test also includes the L scale, which is a *Lie/Social Desirability* scale. This paper presents the results of psychometric studies obtained with the Portuguese Experimental Version of the EPQ-R (Almiro & M.R. Simões, 2007) in a sample of 671 participants. These studies include: reliability analysis (internal consistency and test-retest methods); construct validity analysis (principal components analysis) and concurrent validity analysis (using the test scores in BDI, STAI and BSI as an external criterion of validation); and item analysis (item fit, item reliability, item validity) applying the Rasch Model. These results show the good psychometric properties of the Experimental Version of the EPQ-R.

Keywords: portuguese experimental version of the EPQ-R, Hans Eysenck, personality assessment, classical test theory, item response theory

¹ Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra (FPCE-UC); Laboratório de Avaliação Psicológica e Psicometria. Centro de Investigação do Núcleo de Estudos e Intervenção Cognitivo-Comportamental (CINEICC), Portugal. E-mail: psi.armelim22@yahoo.com

² Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra (FPCE-UC); Laboratório de Avaliação Psicológica e Psicometria. Centro de Investigação do Núcleo de Estudos e Intervenção Cognitivo-Comportamental (CINEICC), Portugal.

Introdução

A personalidade é definida por H. Eysenck como uma organização mais ou menos estável e perdurável do “carácter”, do “temperamento” e dos “aspectos intelectuais e físicos” do sujeito, que determinam o seu ajustamento único ao meio ambiente. As dimensões de personalidade são os elementos fundamentais da estrutura da personalidade e consistem, essencialmente, em fatores disposicionais que a determinam de maneira constante e persistente. Para o autor, a organização hierárquica da personalidade comporta (de um modo ascendente): as respostas específicas, as respostas habituais, os traços de personalidade (que resultam das intercorrelações entre respostas habituais) e os tipos de personalidade (que agrupam traços de personalidade intercorrelacionados) (H. Eysenck & M. Eysenck, 1985; H. Eysenck & S. Eysenck, 2008).

O *Questionário de Personalidade de Eysenck – Forma Revista* (EPQ-R; versão original inglesa, 100 itens), que se baseia nestas conceptualizações, foi construído por S. Eysenck, H. Eysenck e Barrett (1985) e avalia os três fatores/dimensões fundamentais da personalidade – Modelo P-E-N: o *Psicoticismo* (P), a *Extroversão* (E) e o *Neuroticismo* (N); este questionário contém também uma escala de validade, a escala L, indicadora de *Mentira/Desejabilidade Social*, que é vulgarmente designada como escala de *Mentira* (*Lie*, no original). Obtidos por meio do método da análise fatorial, estes fatores/dimensões seguem o princípio de ortogonalidade e são designados por fatores de segunda-ordem (ou fatores gerais) (H. Eysenck & S. Eysenck, 1975, 2008).

A dimensão de *Neuroticismo* (N) organiza-se num *continuum* que abrange, num polo, a *personalidade neurótica* (ou instável) – que engloba os traços de emotividade, ansiedade, depressão, excesso de preocupação, irritabilidade fácil, baixa autoestima, entre outros – e, no outro polo, a *personalidade estável* – caracterizada pelos traços (opostos aos primeiros) de serenidade, autocontrolo, boa disposição, entre outros.

A dimensão de *Extroversão* (E) também se organiza num *continuum*, descrevendo num polo a

personalidade extrovertida – representada pelos traços de sociabilidade, vivacidade, assertividade, espontaneidade, espírito de aventura, otimismo, entre outros –, e no outro, a *personalidade introvertida* – pelos traços de introspeção, inibição, baixa sociabilidade, pessimismo, cautela, entre outros.

Por último, a dimensão de *Psicoticismo* (P) estabelece um *continuum* entre a psicopatologia e a normalidade, caracterizando num polo os indivíduos com elevado *psicoticismo* – que são hostis, egocêntricos, desajustados, rígidos, impulsivos, agressivos, desconfiados, pouco empáticos, entre outros traços –, e no outro polo os indivíduos com baixo *psicoticismo* (personalidade ajustada) – que são amáveis, conscienciosos, empáticos, conformistas, cooperativos e socializados (H. Eysenck, 1998; H. Eysenck & M. Eysenck, 1985; H. Eysenck & S. Eysenck, 1976, 2008; Wilson, 1986).

A construção do EPQ-R resultou do desenvolvimento e aperfeiçoamento dos diversos testes de personalidade anteriormente construídos por H. Eysenck e representa mais de cinco décadas de investigação (Barrett & S. Eysenck, 1984; Barrett, Petrides, S. Eysenck, & H. Eysenck, 1998; H. Eysenck & S. Eysenck, 1975, 2008; Furnham, S. Eysenck, & Saklofske, 2008):

- *Maudsley Medical Questionnaire* (MMQ, 40 itens; H. Eysenck, 1952), constituído pelas dimensões N e E;

- *Maudsley Personality Inventory* (MPI, 48 itens; H. Eysenck, 1959; versão portuguesa, Vaz Serra & Allen Gomes, 1973), também constituído pelas dimensões N e E, que resultou da revisão do MMQ;

- *Inventário de Personalidade de Eysenck* (EPI, 57 itens, Forma A e B; H. Eysenck & S. Eysenck, 1964; versão portuguesa, Vaz Serra, Ponciano, & Fidalgo Freitas, 1980), igualmente constituído pelas dimensões N e E, como resultado da revisão dos anteriores questionários (MMQ e MPI); neste teste foi acrescentada uma escala de mentira, a escala L;

- *Questionário de Personalidade de Eysenck* (EPQ, 90 itens; H. Eysenck & S. Eysenck, 1975; versão portuguesa, Fonseca, S. Eysenck, & A. Simões, 1991), constituído pelas dimensões N e E, sendo introduzida uma terceira dimensão da

personalidade, a dimensão P; este questionário inclui a escala L do EPI.

O EPQ-R (S. Eysenck et al., 1985), que consiste na versão revista do EPQ, mantém a mesma estrutura fatorial (P, E, N, L) e surge com o objetivo de contornar as principais limitações da anterior escala P – reduzida fiabilidade e baixa amplitude das pontuações. Com efeito, o EPQ-R possui índices de precisão superiores aos que tinham sido obtidos pelo EPQ, bem como pontuações médias mais elevadas e uma curva de distribuição melhorada (na escala P).

Os estudos psicométricos com o EPQ-R, realizados em várias dezenas de países (nomeadamente, Inglaterra, Espanha, Itália, Alemanha, Holanda, Grécia, Canadá, Austrália, entre outros), evidenciam a presença de uma estrutura fatorial estável e sucessivamente replicada. A precisão (avaliada através dos coeficientes alfa) é na generalidade boa, variando, nos diferentes fatores (em amostras distintas, considerando género e idades) entre .71 e .90. Com efeito, este instrumento de avaliação da personalidade é um dos mais estudados e utilizados em todo o mundo (ver estudos: versão original, S. Eysenck et al., 1985; versão italiana, San Martini, Mazzotti, & Setaro, 1996; versão espanhola, Ortet, Ibáñez, Moro, Silva, & Boyle, 1999; versão alemã, Ruch, 1999; ver revisões, Almiro & M.R. Simões, 2008a, 2011; Barrett et al., 1998; Furnham et al., 2008).

Neste sentido, o objetivo do presente artigo consiste na apresentação dos dados relativos à evidência de validade e de precisão da *Versão Experimental Portuguesa do Questionário de Personalidade de Eysenck – Forma Revista* (EPQ-R; Almiro & M.R. Simões, 2007), quer no âmbito da Teoria Clássica dos Testes (TCT), quer no âmbito da Teoria da Resposta ao Item (TRI). Os estudos psicométricos efetuados incluem: estudos de validade (validade de constructo, validade concorrente); estudos de precisão (consistência interna, teste-reteste); e análise da qualidade dos itens (Erro Padrão para cada item, Precisão da Separação dos Itens, Precisão da Separação dos Sujeitos) com recurso ao Modelo de Rasch (TRI).

Método

Amostra

Para estudar as propriedades psicométricas da Versão Experimental Portuguesa do EPQ-R, aplicou-se este instrumento a uma amostra da comunidade (diversificada) de 671 participantes, constituída por 264 homens (39.34%) e 407 mulheres (60.66%), com idades compreendidas entre 16 e 40 anos (Total: $M = 23.20$, $DP = 6.08$; Homens: $M = 24.30$, $DP = 6.92$; Mulheres: $M = 22.49$, $DP = 5.36$). A amostra foi obtida por conveniência.

Os participantes exerciam várias atividades profissionais (estudantes, administrativos, vendedores, empresários, agentes da polícia, eletricitas, auxiliares de ação educativa, professores, educadores de infância, assistentes sociais, psicólogos, auxiliares de ação médica, enfermeiros, técnicos de informática, entre outros) e residiam em diversas áreas geográficas de Portugal Continental (meio urbano e rural), nomeadamente em Aveiro, Coimbra, Lisboa, Porto, Setúbal, Vila Real e Viseu.

Nos estudos de validade concorrente, foram aplicados os seguintes instrumentos: o Inventário Depressivo de Beck (BDI) a uma subamostra de 55 participantes (4 homens e 51 mulheres); o Inventário de Estado-Traço de Ansiedade (STAI) a uma subamostra de 67 participantes (11 homens e 56 mulheres); e o Inventário de Sintomas Psicopatológicos (BSI) a uma subamostra de 222 participantes (37 homens e 185 mulheres).

Instrumentos

A versão original inglesa do *Questionário de Personalidade de Eysenck – Forma Revista* (EPQ-R; S. Eysenck, H. Eysenck, & Barrett, 1985) avalia as três dimensões fundamentais da personalidade na perspetiva de H. Eysenck: Psicoticismo (P), Extroversão (E) e Neuroticismo (N). Este instrumento é constituído por 100 itens (resposta dicotómica, Sim/Não), distribuídos por quatro escalas: 32 itens na escala P; 23 itens na escala E; 24 itens na escala N; 21 itens na escala L (Mentira/Desejabilidade Social).

A *Versão Experimental Portuguesa do Questionário de Personalidade de Eysenck – Forma Revista* (EPQ-R; Almiro & M.R. Simões,

2007) é constituída por 76 itens e pelos mesmos quatro fatores (PENL, ver Almiro & M.R. Simões, 2008a, 2008b): 17 itens na escala P; 19 itens na escala E; 22 itens na escala N; 18 itens na escala L.

No que respeita aos estudos de validade concorrente, utilizou-se como critério de validação externo os resultados obtidos em alguns dos mais importantes instrumentos de avaliação clínica (Maruish, 2004a, 2004b):

- *Inventário Depressivo de Beck* (BDI, 21 itens; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961; versão portuguesa, Vaz Serra & Pio Abreu, 1973), que avalia a intensidade de depressão;

- *Inventário de Estado-Traço de Ansiedade* (STAI, 40 itens; Spielberger, 1983; versão portuguesa, Silva, 2006), que consiste numa medida da ansiedade-estado e da ansiedade-traço (a ansiedade-traço é uma característica da personalidade que remete para as diferenças individuais relativamente estáveis quanto à propensão para experienciar a ansiedade); o instrumento é constituído pelas escalas de Ansiedade-Estado (Ans-E) e de Ansiedade-Traço (Ans-T);

- *Inventário de Sintomas Psicopatológicos* (BSI, 53 itens; Derogatis, 1982; versão portuguesa, Canavarro, 2007), que examina 9 grupos de sintomas (constituindo as 9 subescalas deste instrumento) que caracterizam as seguintes perturbações psicológicas e psiquiátricas: Somatização (I), Obsessões-Compulsões (II), Sensibilidade Interpessoal (III), Depressão (IV), Ansiedade (V), Hostilidade (VI), Ansiedade Fóbica (VII), Ideação Paranóide (VIII), Psicoticismo (IX).

Procedimentos

A Versão Experimental Portuguesa do EPQ-R (Almiro & M.R. Simões, 2007) foi construída com base na versão inglesa (S. Eysenck et al., 1985), na versão espanhola (Ortet et al., 1999) e na versão alemã (Ruch, 1999) do EPQ-R, bem como na versão portuguesa do EPQ (Fonseca et al., 1991). Para a elaboração desta versão experimental, foram mantidos os 100 itens da versão original inglesa do EPQ-R e acrescentaram-se 23 itens experimentais provenientes das referidas versões (espanhola,

alemã e portuguesa), perfazendo um total de 123 itens. A seleção dos itens para a nova versão teve sempre por base a intenção de construir um instrumento com características semelhantes ao original (inglês). Os itens foram alvo de um processo de tradução e retroversão. Todos os procedimentos relativos à construção do teste seguiram as normas habitualmente usadas neste tipo de investigação (cf. American Psychological Association, 1999; Hambleton, 1994, 2001; Hambleton, Merenda, & Spielberger, 2005; International Test Commission, 1999/2003).

Os estudos das várias versões do EPQ-R (cf. S. Eysenck et al., 1985; Fonseca et al., 1991; Ortet et al., 1999; Ruch, 1999) mostraram que as pontuações na escala L, que é uma escala de validade, tendem a aumentar com a idade. Assim, com a finalidade de atenuar a possível influência da desejabilidade social nos resultados, optou-se por reduzir a amplitude das idades consideradas no presente estudo, incluindo apenas os indivíduos com idades compreendidas entre 16 e 40 anos (ver Almiro & M.R. Simões, 2008a, 2008b).

Os dados obtidos com a aplicação da Versão Experimental do EPQ-R foram analisados nos programas *SPSS Statistics* (versão 17.0; SPSS Inc., Chicago, IL) e *WINSTEPS – Rasch Measurement* (versão 3.69; Linacre, 2009). Todos os questionários com dados omissos foram eliminados.

No âmbito da TCT, para estudar a validade de constructo do instrumento, aplicou-se o método de Análise em Componentes Principais (ACP), e para estudar a precisão, o método da consistência interna (alfa de Cronbach) e o método de teste-reteste (estabilidade temporal). Foram também realizados estudos de validade de critério concorrente (BDI, STAI, BSI).

No âmbito da TRI, foi efetuado um exame da qualidade dos itens (ajuste *infit* e *outfit*, precisão, validade), recorrendo ao Modelo de Rasch (1PL, modelo logístico de 1 parâmetro). Uma vez que este instrumento avalia três dimensões independentes da personalidade (N, E, P) e contém uma escala de validade (L), a análise de cada uma das escalas foi efetuada em separado; os participantes com pontuações extremas (mínimas e máximas) foram eliminados desta análise.

Resultados

Estatística Descritiva

As pontuações médias obtidas pelos 671 indivíduos (264 homens e 407 mulheres) que participaram no presente estudo psicométrico da Versão Experimental do EPQ-R (76 itens) são apresentadas no Quadro 1.

Quadro 1. Pontuações médias obtidas na Versão Experimental Portuguesa do EPQ-R

Escala	Total	Homens	Mulheres
	M (DP)	M (DP)	M (DP)
N	10.79 (5.37)	9.17 (5.41)	11.83 (5.09)
E	13.45 (4.01)	13.56 (4.07)	13.39 (3.98)
P	2.51 (2.16)	3.24 (2.44)	2.03 (1.81)
L	8.44 (3.58)	8.11 (3.58)	8.66 (3.57)

Nota. N=Neuroticismo. E=Extroversão. P=Psicoticismo. L=Mentira/Desejabilidade Social. M=média. DP=desvio-padrão.

Para a analisar a existência de diferenças estatisticamente significativas nas médias das pontuações obtidas em relação ao género, foi aplicado o *Teste t* (para duas amostras independentes) e calculou-se o correspondente coeficiente *d* de Cohen para estimar a dimensão do efeito (cf. Cohen, 1988). Assim, foi possível verificar que existem diferenças ($p < .001$) ao nível da escala N [$t(537.240) = -6.377$; $d = -0.55$] e ao nível da escala P [$t(669) = 7.369$; $d = 0.57$], para as quais a dimensão do efeito é moderada. Estes dados mostram que as mulheres apresentam um nível de *Neuroticismo* superior ao dos homens e que o nível de *Psicoticismo* nos homens é superior ao das mulheres. Em relação às escalas E e L, não foram encontradas diferenças estatisticamente significativas.

Validade de Constructo

No âmbito da TCT, foi aplicado o método de *Análise em Componentes Principais* (ACP), com rotação *direct oblimin* e normalização de Kaiser, por ter sido o método utilizado por H. Eysenck e colaboradores na construção da versão original do EPQ-R (ver S. Eysenck et al., 1985). A Versão Experimental do EPQ-R era inicialmente composta por 123 itens (ver Almiro & M.R. Simões, 2008a).

Tendo em conta a matriz das correlações, os valores obtidos no Teste de Kaiser-Meyer-Olkin

(KMO) – o valor de .823 é considerado “bom” (entre .80 e .90; ver critério, Kaiser, 1974) – e no Teste de Esfericidade de Bartlett – o valor do qui-quadrado [$\chi^2 = 11885.38$; $df = 2850$; $p = .000$] é estatisticamente significativo – confirmam que esta amostra é adequada para fatorialização (cf. Pestana & Gageiro, 2005). Com o objetivo de determinar a estrutura fatorial da Versão Experimental do EPQ-R, efetuaram-se sucessivas análises exploratórias, com três, quatro e cinco fatores, até encontrar a estrutura mais adequada e consistente. A estrutura de quatro fatores foi a que demonstrou ser mais adequada. Tal facto foi também observado através do *scree-plot*. Como critério para a extração dos fatores, consideraram-se os *eigenvalues* ≥ 2.5 , adotando, simultaneamente, o critério de Kaiser-Guttman (*eigenvalues* ≥ 1.0) e o já mencionado *scree test* de Cattell (cf. Nunnally, 1978).

Além da estrutura de quatro fatores ser a que mais se aproxima da versão original do EPQ-R (S. Eysenck et al., 1985), segue o princípio psicométrico de “estrutura simples” definido por Thurstone (1935, 1947; cf. Nunnally, 1978). H. Eysenck (1998) elaborou o seu modelo hierárquico da personalidade com base neste princípio – os quatro tipos de fatores correspondem aos quatro níveis de organização hierárquica do modelo (de modo ascendente): os fatores de erro correspondem às respostas específicas, os fatores específicos às respostas habituais, os fatores de grupo (fatores de primeira-ordem) aos traços de personalidade (correlações de respostas habituais), e os fatores gerais (fatores de segunda-ordem) aos tipos de personalidade (correlações de traços de personalidade). A par disso, de acordo com o critério de Kline (1994), consideraram-se apenas os itens com cargas fatoriais (saturações) $\geq .30$, sendo os restantes eliminados.

A factor loading of .30 indicates that 9 per cent of the variance is accounted for by the factor; this is taken as large enough to indicate that the loading is salient; thus in factor analyses where the sample is at least 100 subjects this is a reasonable criterion; loadings of .30 or larger are regarded as significant (Kline, 1994, p.52).

Com efeito, neste estudo psicométrico, com recurso à ACP, a Versão Experimental do EPQ-R

Quadro 2. Análise em Componentes Principais da Versão Experimental Portuguesa do EPQ-R

Itens	Fatores				Itens	Fatores			
	N	E	L	P		N	E	L	P
E_63	.66	-.09	-.11	.01	E_14	-.12	-.05	.57	-.19
E_17	.63	-.05	-.12	.20	E_21	.05	-.03	.57	-.09
E_19	.62	-.08	-.07	.10	E_47	-.04	-.01	.54	-.04
E_81	.62	-.18	-.09	.15	E_26	-.10	-.04	.53	-.16
E_74	.62	-.04	-.08	.03	E_55	-.12	.00	.46	-.07
E_88	.61	-.06	-.05	-.11	E_106	-.13	-.04	.45	-.26
E_34	.59	-.09	-.08	-.05	E_31	-.07	-.02	.45	-.12
E_3	.57	-.07	-.12	.03	E_69	-.08	-.17	.44	-.06
E_45	.55	.01	-.04	.13	E_35	-.16	.09	.42	-.19
E_108	.53	-.11	-.08	.16	E_64	.06	-.16	.40	-.08
E_30	.52	-.08	-.14	.06	E_79	-.14	-.14	.40	.05
E_8	.52	-.09	-.23	-.06	E_15	-.04	.11	.40	-.07
E_77	.50	-.17	-.03	-.24	E_85	.02	-.04	.40	.14
E_22	.49	-.15	-.02	.05	E_52	-.07	-.03	.39	.06
E_96	.47	.05	-.20	.19	E_121	-.15	-.02	.39	-.29
E_37	.46	-.06	.08	-.27	E_23	-.06	.07	.39	.11
E_87	.46	-.11	.04	.22	E_83	-.08	.11	.38	.08
E_93	.45	-.13	-.16	-.17	E_89	-.05	.05	.36	.08
E_13	.41	-.07	-.01	-.17	E_98	.28	-.12	.06	.49
E_44	.41	-.17	.00	.06	E_27	.03	-.04	-.09	.48
E_73	.41	-.08	-.08	.13	E_57	-.10	.04	-.04	.42
E_122	.38	-.08	-.09	.07	E_51	.05	.03	-.01	.41
E_50	-.10	.64	.04	.13	E_76	.00	.03	.02	.41
E_75	-.13	.63	.03	-.02	E_9	-.04	.03	-.09	.40
E_92	-.03	.62	.00	.07	E_67	-.03	-.01	-.06	.40
E_118	-.11	.61	.09	-.02	E_70	.25	-.17	-.01	.39
E_58	-.08	.60	.11	-.02	E_123	.12	.04	-.16	.39
E_46	-.21	.55	-.07	-.15	E_105	.01	-.08	.04	.38
E_5	.04	.55	-.07	-.03	E_112	.12	.05	-.16	.35
E_11	-.26	.50	.07	-.11	E_114	.15	.04	.09	.34
E_56	-.08	.49	-.02	-.13	E_41	.09	.26	-.25	.34
E_16	-.20	.46	.05	-.13	E_90	.20	.19	-.25	.33
E_28	-.11	.45	-.08	.09	E_101	.01	.10	-.07	.31
E_120	-.09	.43	-.15	-.06	E_12	-.11	-.13	-.03	.31
E_94	-.28	.41	.07	.00	E_29	-.02	-.06	.06	.30
E_65	-.02	.40	.02	.15					
E_24	.01	.38	-.07	.05					
E_4	-.10	.36	.00	-.18					
E_61	-.03	.36	.04	.05					
E_1	-.17	.35	-.05	.06					
E_117	.06	.30	-.23	.02					

Nota. Cargas fatoriais $\geq .30$ assinaladas a negrito. N=Neuroticismo. E=Extroversão. P=Psicoticismo. L=Mentira/Desejabilidade Social.

passou a integrar um total de 76 itens distribuídos por quatro fatores – N, E, P, L (dos 123 itens experimentais, foram eliminados 47 itens; em relação à versão original do EPQ-R, foram eliminados 2 itens na escala N, 4 itens na escala E, 3 itens na escala L e 15 itens na escala P). A estrutura de quatro fatores obtida explica 24.88% da variância total (ver Quadro 2):

- o Fator N explica 9.81% da variância (*eigenvalue* = 7.46), é constituído por 22 itens (saturações significativas com valores entre .66 e .38, média de .52);

- o Fator E explica 6.28% da variância (*eigenvalue* = 4.77), é constituído por 19 itens (saturações significativas com valores entre .64 e .30, média de .48);

- o Fator L (*Mentira/Desejabilidade Social*) explica 4.89% da variância (*eigenvalue* = 3.71), é constituído por 18 itens (saturações significativas com valores entre .57 e .36, média de .44);

- o Fator P explica 3.90% da variância (*eigenvalue* = 2.97), é constituído por 17 itens (saturações significativas com valores entre .49 e .30, média de .38).

No âmbito da TCT, com base nos itens da ACP, foram calculadas as correlações item-fator (correlações entre os itens de cada fator e o fator da Versão Experimental do EPQ-R a que correspondem), incluindo o próprio item (para analisar a convergência no fator) e excluindo o próprio item (correlação bisserial-pontual corrigida, para examinar o índice de discriminação). Para o fator N, incluindo o item, os valores oscilam entre .65 e .41 (média de .53), e excluindo o item, variam entre .60 e .34 (média de .46). Para o fator E, incluindo o item, os valores oscilam entre .63 e .32 (média de .48), e excluindo o item, variam entre .55 e .22 (média de .40). Para o fator P, incluindo o item, os valores variam entre .51 e .22 (média de .39), e excluindo o item, oscilam entre .35 e .15 (média de .26). Para o fator L, incluindo o item, os índices variam entre .57 e .34 (média de .45), e excluindo o item, oscilam entre .48 e .25 (média de .34). Todas as correlações são positivas e estatisticamente significativas ($p < .01$; *r de Pearson*), apresentando coeficientes adequados (ver critérios, Cohen, 1988; Kline, 1986; Nunnally, 1978).

Quadro 3. Correlações entre os fatores da Versão Experimental Portuguesa do EPQ-R

Fatores	N	E	P	L
N				
E	-.22**			
P	.17**			
L	-.18**		-.21**	

Nota. ** $p < .01$. Os coeficientes de correlação não significativos foram suprimidos. N=Neuroticismo. E=Extroversão. P=Psicoticismo. L=Mentira/Desejabilidade Social

No âmbito da TCT, foram também calculadas as correlações entre os fatores N, E, P, L (ver Quadro 3). Todos os coeficientes de correlação obtidos são fracos e estatisticamente significativos

($p < .01$; *r de Pearson*) – os valores oscilam entre -.22 e .17. As correlações com E são nulas (ver critério, Cohen, 1988).

No âmbito da TRI (Modelo de Rasch), foi testada a unidimensionalidade das escalas da Versão Experimental do EPQ-R aplicando o método da *Análise de Componentes Principais dos Resíduos* (ACPR). A unidimensionalidade é um princípio fundamental da TRI, no qual se estabelece que os itens de um teste (ou escala) de avaliação psicológica devem medir um único atributo (ou traço latente) e que o desempenho dos indivíduos nessa escala apenas é influenciado por um fator predominante (o constructo medido pela escala) (cf. Embretson & Reise, 2000; Hambleton & Swaminathan, 1985). De acordo com o critério de Linacre (1999), uma escala é unidimensional quando a variância explicada pelo *componente principal* da dimensão examinada é superior a 20% e o *eigenvalue* do *primeiro contraste* (unidade de variância residual) é baixo (o valor não deve exceder 2.0).

Os resultados obtidos na ACPR mostram que o componente principal da dimensão N explica 39.00% da variância, com *eigenvalue* 14.1, e que o primeiro contraste explica 6.70% da variância, com *eigenvalue* 2.4. O componente principal da dimensão E explica 28.10% da variância, com *eigenvalue* 7.4, e o primeiro contraste explica 6.70% da variância, com *eigenvalue* 1.8. O componente principal da dimensão P explica 29.50% da variância, com *eigenvalue* 7.1, e o primeiro contraste explica 5.70% da variância, com *eigenvalue* 1.4. Por último, o componente principal da escala L explica 37.60% da variância, com um *eigenvalue* 10.9, e o primeiro contraste explica 5.50% da variância, com um *eigenvalue* 1.6.

Precisão

No âmbito da TCT, relativamente à precisão, foram calculados os alfas de Cronbach ($N = 671$) e os índices de precisão teste-reteste ($n = 33$, subamostra feminina; estabilidade temporal entre 4 a 6 semanas) para cada escala – N, E, P, L (ver Quadro 4).

No âmbito da TRI (Modelo de Rasch), a precisão foi analisada mediante três índices: o Erro Padrão (EP) (para cada item), a Precisão da

Quadro 4. Índices de precisão da Versão Experimental Portuguesa do EPQ-R

Escala	Consistência Interna	Teste-Retestes	Itens	Total
	α	4-6 semanas	Escala	
N	.88	.91	22	
E	.82	.85	19	
P	.64	.80	17	76
L	.77	.71	18	

Nota. α =alfa de Cronbach. N=Neuroticismo. E=Extroversão. P=Psicoticismo. L=Mentira/Desejabilidade Social

Separação dos Itens (PSI), e a Precisão da Separação dos Sujeitos (PSS). Com efeito, os valores do EP dos itens da escala N oscilam entre 0.09 e 0.12 (média de 0.10), a PSI é .99, e a PSS é .84. Na escala E, os valores do EP dos itens variam entre 0.09 e 0.21 (média de 0.11), a PSI é .99, e a PSS é .71. Na escala P, os valores do EP dos itens oscilam entre 0.09 e 0.29 (média de 0.16), a PSI é .98, e a PSS é .27. Na escala L, os valores do EP dos itens variam entre 0.09 e 0.14

(média de 0.10), a PSI é .99, e a PSS é .77 (ver Quadro 5).

Análise dos Itens

No âmbito da TRI, procedeu-se ao exame da qualidade dos itens da Versão Experimental do EPQ-R mediante a análise do seu índice de dificuldade (“acertos” e “erros”) e da sua adequabilidade para avaliar o nível do atributo nos sujeitos (medição conjunta de itens e sujeitos na mesma escala *logit*). No contexto da avaliação da personalidade, o índice de dificuldade deve ser interpretado do seguinte modo: no atributo medido, o sujeito “acerta” o item quando está de acordo com seu conteúdo (indica que o traço de personalidade está presente no indivíduo) ou “erra” quando está em desacordo (indica que o traço de personalidade não está presente no indivíduo).

No Modelo de Rasch, existem dois indicadores fundamentais na estatística de ajuste

Quadro 5. Resumo dos valores da calibração entre itens/sujeitos e índices de precisão da Versão Experimental do EPQ-R (TRI)

N	Itens		Sujeitos		Precisão		
	<i>M (DP)</i>	[Min, Max]	<i>M (DP)</i>	[Min, Max]	PSI	PSS	α
Medida	0.00 (1.21)	[-2.17, 2.05]	-0.07 (1.53)	[-3.69, 3.60]			
<i>Infit</i>	1.00 (0.11)	[0.80, 1.24]	1.00 (0.23)	[0.45, 1.97]	.99	.84	.88
<i>Outfit</i>	0.98 (0.22)	[0.65, 1.58]	0.98 (0.48)	[0.25, 4.19]			
EP	0.10 (0.01)	[0.09, 0.12]	0.57 (0.13)	[0.49, 1.06]			
E	Itens		Sujeitos		Precisão		
	<i>M (DP)</i>	[Min, Max]	<i>M (DP)</i>	[Min, Max]	PSI	PSS	α
Medida	0.00 (0.98)	[-2.72, 1.42]	1.15 (1.24)	[-3.41, 3.21]			
<i>Infit</i>	0.99 (0.12)	[0.82, 1.23]	1.00 (0.16)	[0.50, 1.70]	.99	.71	.82
<i>Outfit</i>	0.96 (0.23)	[0.60, 1.43]	0.96 (0.42)	[0.14, 6.16]			
EP	0.11 (0.03)	[0.09, 0.21]	0.63 (0.16)	[0.50, 1.09]			
P	Itens		Sujeitos		Precisão		
	<i>M (DP)</i>	[Min, Max]	<i>M (DP)</i>	[Min, Max]	PSI	PSS	α
Medida	0.00 (1.34)	[-2.71, 2.05]	-2.32 (1.04)	[-3.53, 2.59]			
<i>Infit</i>	0.99 (0.05)	[0.90, 1.10]	1.00 (0.30)	[0.49, 2.09]	.98	.27	.64
<i>Outfit</i>	0.95 (0.22)	[0.41, 1.35]	0.94 (1.04)	[0.18, 9.90]			
EP	0.16 (0.06)	[0.09, 0.29]	0.83 (0.18)	[0.57, 1.09]			
L	Itens		Sujeitos		Precisão		
	<i>M (DP)</i>	[Min, Max]	<i>M (DP)</i>	[Min, Max]	PSI	PSS	α
Medida	0.00 (1.38)	[-2.02, 2.62]	-0.20 (1.25)	[-3.53, 3.61]			
<i>Infit</i>	1.00 (0.07)	[0.85, 1.10]	1.00 (0.29)	[0.45, 2.21]	.99	.73	.77
<i>Outfit</i>	0.99 (0.12)	[0.73, 1.16]	0.99 (0.66)	[0.19, 7.38]			
EP	0.10 (0.01)	[0.09, 0.14]	0.61 (0.08)	[0.56, 1.08]			

Nota. N=Neuroticismo. E=Extroversão. P=Psicoticismo. L=Mentira/Desejabilidade Social. *M*=média. *DP*= desvio-padrão. Min=valor mínimo. Max=valor máximo. EP=erro padrão. PSI=precisão da separação dos itens. PSS=precisão da separação dos sujeitos. α =alfa de Cronbach

dos itens e dos sujeitos ao modelo: o índice de *infit*, que representa a média dos resíduos quadráticos ponderados com a sua variância e que reflecte as respostas inesperadas dadas aos itens próximos do nível do sujeito no atributo; e o índice de *outfit*, que equivale à média dos resíduos quadráticos não ponderados, manifestando os comportamentos inesperados em relação aos itens distantes do nível do sujeito no atributo (ou seja, quando o sujeito, em função do seu nível, erra nos itens muito fáceis ou acerta nos itens mais difíceis). Neste sentido, enquanto que o ajuste *outfit* é mais sensível aos *outliers*, podendo alcançar valores elevados em consequência de respostas inesperadas, o ajuste *infit* é mais robusto e constitui, por isso, o melhor indicador das qualidades psicométricas dos itens (para ambos, os valores assumem uma distribuição entre 0.00 e infinito, com um valor médio esperado de 1.00) (Bond, & Fox, 2007; Hambleton & Swaminathan, 1985).

Neste sentido, o exame da qualidade dos itens, realizado através da estatística *infit* e *outfit*, baseia-se em dois princípios fundamentais: o princípio da *unidimensionalidade* (já mencionado) e o princípio da *independência local*, o qual estabelece que para um mesmo nível de atributo, as respostas dos indivíduos aos itens do teste (ou escala) de avaliação psicológica são estatisticamente independentes, isto é, a resposta a um item não é influenciada pelas respostas dadas aos restantes itens da escala (Embretson & Reise, 2000; Hambleton & Swaminathan, 1985). De acordo com o critério de Linacre (1999), considera-se que: existe ajuste (itens/sujeitos) quando os índices *infit* e *outfit* oscilam entre 0.50 e 1.50 (média 1.00), sendo que os índices próximos de 1.00 revelam um ajuste perfeito; acima de 1.50, existe um desajuste moderado; acima de 2.00, existe um desajuste severo; abaixo de 0.50, existe um “sobreajuste”; o índice *infit* constitui o melhor indicador de ajuste.

Neste estudo psicométrico, os índices de ajuste *infit* e *outfit* para os itens são adequados em todas as escalas da Versão Experimental do EPQ-R (N, E, P, L): os índices *infit* oscilam entre o valor mínimo 0.80 e o valor máximo 1.24 (média entre 0.99 e 1.00) e os índices *outfit* variam entre o valor mínimo 0.41 e o valor máximo 1.58

(média entre 0.95 e 0.99). O valor mínimo 0.41 de ajuste *outfit* corresponde à escala P (apenas 2 itens apresentam um índice inferior a 0.50, indicador de “sobreajuste”); o valor máximo 1.58 de ajuste *outfit* corresponde à escala N (apenas 1 item apresenta um índice superior a 1.50, indicador de desajuste moderado) (ver Quadro 5).

Em relação aos participantes, na escala N ($N = 659$), para *infit* (índices entre 0.45 e 1.97, média 1.00), 2.43% dos participantes revelam desajuste moderado, e para *outfit* (entre 0.25 e 4.19, média 0.98) 8.50% mostram desajuste moderado e 4.10% desajuste severo. Na escala E ($N = 629$), para *infit* (índices entre 0.50 e 1.70, média 1.00), 1.11% dos participantes revelam desajuste moderado, e para *outfit* (entre 0.14 e 6.16, média 0.96) 6.36% mostram desajuste moderado e 2.23% desajuste severo. Na escala P ($N = 579$), para *infit* (índices entre 0.49 e 2.09, média 1.00), 4.32% dos participantes revelam desajuste moderado e 0.35% desajuste severo, e para *outfit* (entre 0.18 e 9.90, média 0.94), 5.35% mostram desajuste moderado e 8.12% desajuste severo. Por último, na escala L ($N = 664$), para *infit* (índices entre 0.45 e 2.21, média 1.00), 5.87% dos participantes revelam desajuste moderado e 0.30% desajuste severo, e para *outfit* (entre 0.19 e 7.38, média 0.99), 6.78% mostram desajuste moderado e 5.87% desajuste severo.

A par disso, como é desejável, os valores da medida (localização na escala *logit*) para os itens assumem o ponto médio 0.00 em todas as escalas da Versão Experimental do EPQ-R (N, E, P, L). Na calibração entre itens/sujeitos, as escalas N (0.00/-0.07) e L (0.00/-0.20) apresentam um ajuste muito bom entre o índice de dificuldade dos itens e o nível do atributo nos sujeitos, medindo convenientemente os níveis baixos, médios e elevados do atributo. Na escala E (0.00/1.15), os itens são relativamente “fáceis” para os sujeitos avaliados, sendo que estes medem melhor os níveis médios e elevados do atributo do que os níveis baixos (os sujeitos avaliados são tendencialmente extrovertidos; a pontuação média é 13.45 e o desvio-padrão é 4.01, numa escala com 19 itens). Na escala P (0.00/-2.32), os itens são muito “difíceis” para os sujeitos, o que significa que a sua medição só se ajusta aos níveis baixos do atributo (a pontuação média é 2.51 e o

Quadro 6. Correlações entre Versão Experimental Portuguesa do EPQ-R e BDI, STAI, BSI

		EPQ-R			
		N	E	P	L
BDI		.71**	-.39**		
STAI	Ansiedade-Estado (Ans-E)	.67**	-.28*		
	Ansiedade-Traço (Ans-T)	.72**	-.41**		
Índice Geral de Sintomas (IGS)		.66**		.19**	
BSI	I. Somatização	.47**		.15*	
	II. Obsessões-Compulsões	.52**		.20**	
	III. Sensibilidade Interpessoal	.54**	-.17**	.20**	-.16*
	IV. Depressão	.66**	-.15*	.15*	
	V. Ansiedade	.66**			
	VI. Hostilidade	.58**		.23**	-.20**
	VII. Ansiedade Fóbica	.40**			
	VIII. Ideação Paranóide	.49**		.26**	
	IX. Psicoticismo	.58**		.15*	

Nota. ** $p < .01$, * $p < .05$. Os coeficientes de correlação não significativos foram suprimidos. N=Neuroticismo. E=Extroversão. P=Psicoticismo. L=Mentira/Desejabilidade Social

desvio-padrão é 2.16, numa escala com 17 itens) (ver Quadro 5).

Validade de Critério Concorrente

No âmbito da TCT, relativamente aos estudos de validade concorrente realizados com a Versão Experimental do EPQ-R, foram considerados os resultados obtidos no BDI ($n = 55$), no STAI ($n = 67$) e no BSI ($n = 222$) como critério de validação externo. No BDI, a média das pontuações é 6.15 ($DP = 5.55$). No STAI, a média das pontuações é 35.36 ($DP = 9.54$) na Ans-E e 38.45 ($DP = 10.55$) na Ans-T. No BSI (Índice Geral de Sintomas), a média das pontuações é 45.31 ($DP = 29.20$). As correlações obtidas (r de Pearson) entre o EPQ-R e o BDI, o EPQ-R e o STAI, e o EPQ-R e o BSI são apresentadas no Quadro 6.

Discussão

Os resultados dos estudos psicométricos da Versão Experimental Portuguesa do EPQ-R (Almiro & M.R. Simões, 2007), efetuados tanto no âmbito da TCT como no âmbito da TRI, sugerem na generalidade bons índices de validade e de precisão.

Na TCT (estudo da validade de constructo), a estrutura fatorial da versão original do EPQ-R (S. Eysenck et al., 1985) é replicada na presente investigação através da ACP, contemplando os

quatro fatores do instrumento: Neuroticismo (N), Extroversão (E), Psicoticismo (P), Mentira/Desejabilidade Social (L). Mas, os fatores desta versão apenas explicam 24.88% da variância total; todavia, vários resultados semelhantes foram reportados (com valores entre 20% e 30%), nomeadamente nos estudos da versão reduzida do EPQ-R (Alexopoulos & Kalaitzidis, 2004; Aluja, Ó. García, & L. García, 2003; Glicksohn & Abulafia, 1998) e do *Inventário de Personalidade NEO – Revisto* (NEO-PI-R; versão portuguesa, Lima & A. Simões, 1997), que é constituído pelas dimensões de Neuroticismo (N), Extroversão (E), Abertura à Experiência (O), Amabilidade (A) e Conscienciosidade (C).

A par disso, como é desejável, as correlações item-fator são adequadas e as correlações entre os fatores da Versão Experimental do EPQ-R são fracas (ou nulas). As correlações item-fator moderadas e elevadas demonstram, por um lado, que todos os itens convergem nos fatores a que pertencem (em coerência com a versão original do instrumento) e, por outro lado, que os itens são discriminativos; segundo Nunnally (1978) e Kline (1986), os índices devem oscilar entre .20 e .40.

Na TRI (estudo da validade de constructo), por meio da ACPR, verifica-se também que as percentagens de variância explicada que correspondem aos componentes principais das escalas N, E, P, L são elevadas e que os

eigenvalues dos respetivos primeiros contrastes são baixos (como é desejável). Deste modo, as escalas E, P e L preenchem na totalidade o critério de Linacre (1999) (variância explicada pelo componente principal > 20%, *eigenvalue* do primeiro contraste < 2.0). Contudo, na escala N, apesar da variância explicada pelo componente principal N cumprir o critério, o *eigenvalue* do primeiro contraste excede ligeiramente 2.0; ainda assim, de acordo com Fisher (2007), se a variância explicada pelo primeiro contraste for baixa, entre 5% e 10%, considera-se que existe unidimensionalidade.

Ainda no âmbito da TRI, o exame da qualidade dos itens da Versão Experimental do EPQ-R (N, E, P, L) revela bons índices de ajuste *infit* e *outfit* (ver critério, Linacre, 1999), o que comprova a sua adequabilidade para avaliar os níveis baixo, médio e elevado do atributo N, E, L nos indivíduos. No entanto, apesar da boa qualidade dos itens da escala P, demonstrada quer pelo ajuste *infit* e *outfit*, quer pelo valor PSI, os dados da calibração entre itens/sujeitos (0.00/-2.32) mostram que o seu índice de dificuldade é muito elevado, o que indica que a sua medição apenas é adequada para os níveis baixos do atributo P nos indivíduos. A par disso, os princípios de unidimensionalidade e de independência local foram igualmente confirmados (cf. Embretson & Reise, 2000; Hambleton & Swaminathan, 1985).

Com efeito, os dados da TCT (ACP, correlações item-fator, correlações entre fatores) e da TRI (ACPR, ajuste *infit* e *outfit*) são comprovativos da validade de constructo da Versão Experimental do EPQ-R, evidenciando, por um lado, a ortogonalidade dos fatores (que são relativamente independentes), e por outro lado, a sua unidimensionalidade (que medem apenas um constructo).

Na TCT, o estudo da precisão deste instrumento, através do método da consistência interna, revela índices “muito bons” (“*very good*”) para as escalas N (coeficiente alfa de .88) e E (.82), um índice “respeitável” (“*respectable*”) para a escala L (.77) – todos superiores a .70 (ver critério, DeVellis, 1991) –, e um índice “indesejável” (“*undesirable*”) para a escala P (.64). Este valor pode dever-se à extensão reduzida

desta escala (17 itens), em comparação com a versão original inglesa (32 itens) do questionário (cf. Nunnally, 1978). Com a exceção do coeficiente alfa de P, os dados são similares aos que foram obtidos com as principais versões do instrumento (versão original inglesa e versões espanhola, italiana e alemã, cf. S. Eysenck et al., 1985; Ortet et al., 1999; Ruch, 1999; San Martini et al., 1996).

A precisão avaliada pelo método de teste-reteste (4-6 semanas) é elevada para as quatro escalas (.91 para N, .85 para E, .71 para L), nomeadamente para a escala P, que apresenta o valor de .80.

Os índices de precisão estimados no âmbito da TRI são na generalidade elevados. De facto, as escalas N, E, L atingem ótimos valores PSI e bons valores PSS, associados aos baixos valores EP (como é desejável). Todavia, a escala P, apesar de possuir um elevado valor PSI e valores EP dentro dos parâmetros aceitáveis, apresenta um valor PSS muito baixo (.27); este valor está em consonância com o elevado índice de dificuldade da escala P, que se repercute na sua inadequação para avaliar as características dos indivíduos desta amostra (nestes indivíduos, o nível do atributo P é muito baixo).

As correlações obtidas nos estudos de validade concorrente efetuados no âmbito da TCT mostram que existe uma forte associação entre o constructo de depressão medido pelo BDI e o constructo N avaliado pelo EPQ-R (.71, correlação elevada). Por seu turno, o constructo E avaliado pelo EPQ-R e o constructo de depressão examinado pelo BDI encontram-se moderada e inversamente associados (-.39, correlação negativa e moderada). Isto é, o constructo de depressão está intimamente relacionado com N (personalidade neurótica) e de modo moderado com a personalidade introvertida. Estes resultados estão de acordo com dados obtidos por Cheng e Furnham (2003) e Clarke (2004).

Com o STAI, verifica-se que existe uma forte associação entre o constructo de ansiedade-traço e o constructo N avaliado pelo EPQ-R (.72, correlação elevada). No que se refere ao constructo E medido pelo EPQ-R, este assume uma relação moderada e inversa com o constructo de ansiedade-traço do STAI (-.41, correlação

negativa e moderada). Isto é, o constructo de ansiedade-traço está simultânea e fortemente relacionado com N (personalidade neurótica) e com a personalidade introvertida de forma moderada. Relativamente ao constructo de ansiedade-estado, os resultados são similares: existe uma forte associação com o constructo N (personalidade neurótica) examinado pelo EPQ-R (.67, correlação elevada). Mas, o mesmo não se verifica no constructo E, cuja relação com a ansiedade-estado é inversa e de baixa intensidade (correlação negativa e fraca). Stewart, Ebmeier e Deary (2005) e Ortet, Roglà e Ibáñez (in H. Eysenck & S. Eysenck, 2008) obtiveram resultados similares.

No que concerne ao BSI, é possível constatar que existe uma forte associação entre o constructo N medido pelo EPQ-R e os sintomas psicopatológicos avaliados pelo BSI, quer em relação ao IGS (.66), quer em relação aos sintomas de Depressão (.66), Ansiedade (.66), Sensibilidade Interpessoal (.54), Obsessões-Compulsões (.52), Hostilidade (.58) e Psicoticismo (.58) (correlações elevadas). O constructo N encontra-se também moderadamente relacionado com os sintomas de Somatização (.47), Ansiedade Fóbica (.40) e Ideação Paranóide (.49) (correlações moderadas). Estes dados corroboram com os que foram obtidos por Ruipérez, Ibáñez, Lorente, Moro e Ortet (2001) na validação espanhola do BSI.

Os resultados menos esperados, pela sua afinidade do ponto de vista teórico (cf. Claridge, 1997; H. Eysenck & M. Eysenck, 1985; Nias, 1986), dizem respeito às relações entre P e os sintomas de Hostilidade, Ideação Paranóide e Psicoticismo do BSI, cujos coeficientes são de baixa intensidade, apesar de serem positivos e significativos (todavia, esta tendência foi igualmente observada no estudo de Ruipérez et al., 2001). Neste sentido, o constructo P definido por H. Eysenck parece ser diferente do constructo de Psicoticismo do BSI. Contudo, importa reconhecer que a subamostra avaliada não é clínica e isso pode ter influenciado a expressividade da dimensão P neste contexto.

As correlações elevadas entre o constructo N medido pelo EPQ-R e os sintomas de Depressão e Ansiedade examinados pelo BSI vão no mesmo

sentido dos resultados obtidos com o BDI, que é uma medida de depressão, e com o STAI, que é uma medida de ansiedade.

Em conclusão, os bons índices de validade e de precisão evidenciados nos estudos psicométricos da Versão Experimental Portuguesa do EPQ-R (Almiro & M.R. Simões, 2007) recenseados no presente artigo (TCT e TRI) sugerem que este é um instrumento de avaliação da personalidade adequado para ser utilizado no contexto português. No entanto, apesar da amostra avaliada no âmbito da presente investigação ser extensa, existem algumas limitações a apontar; designadamente, o facto da amostra ser maioritariamente constituída por participantes do sexo feminino, o facto de apenas integrar participantes com idades compreendidas entre os 16 e os 40 anos, e o facto de não contemplar grupos clínicos para examinar e compreender melhor as relações verificadas entre o constructo N e os constructos de depressão (BDI), de ansiedade (STAI) e os sintomas psicopatológicos (BSI).

Agradecimentos

Este trabalho foi desenvolvido com o apoio da Fundação para a Ciência e a Tecnologia através de uma Bolsa de Doutoramento (SFRH/BD/37970/2007) atribuída ao primeiro autor.

Referências

- Alexopoulos, D.S., & Kalaitzidis, I. (2004). Psychometric properties of Eysenck Personality Questionnaire Revised (EPQ-R) Short Scale in Greece. *Personality and Individual Differences*, 37, 1205-1220.
- Almiro, P.A., & Simões, M.R. (2007). Questionário de Personalidade de Eysenck – Forma Revista (EPQ-R): Estudos para uma nova adaptação portuguesa. Poster apresentado no X Congresso Internacional e Multidisciplinar do Centro de Psicopedagogia da Universidade de Coimbra “Igualdade para todos: Contributos da Educação e da Psicologia”, 29-30 Maio, Coimbra.
- Almiro, P.A., & Simões, M.R. (2008a). Estudo exploratório com a Versão Experimental

- Portuguesa do Questionário de Personalidade de Eysenck – Forma Revista (EPQ-R). In *Actas da XIII Conferência Internacional “Avaliação Psicológica: Formas e Contextos”* (artigo, 13 páginas). Braga: Psiquilibrios. ISBN: 978-989-95522-6-5.
- Almiro, P.A., & Simões, M.R. (2008b). Questionário de Personalidade de Eysenck – Forma Revista (EPQ-R): Estudos preliminares de adaptação e de validação. Comunicação apresentada no XIII Congresso Multidisciplinar do Centro de Psicopedagogia da Universidade de Coimbra “Psicologia e Educação: Tendências actuais”, 11-12 Junho, Coimbra.
- 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.
- Aluja, A., García, Ó., & García, L.F. (2003). A psychometric analysis of the Revised Eysenck Personality Questionnaire Short Scale. *Personality and Individual Differences*, 35, 449-460.
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington D.C.: 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.
- 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.
- Bond, T.G., & Fox, C.M. (2007). *Applying the Rasch Model: Fundamental measurement in the human sciences* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Canavarro, M.C. (2007). Inventário da Sintomas Psicopatológicos (BSI). In M.R. Simões, L.S. Almeida, C. Machado, & M.M. Gonçalves (Eds.), *Avaliação psicológica: Instrumentos validados para a população portuguesa* (Vol. III, pp.305-331). Coimbra: Quarteto.
- Castañeiras, C., & Posada, M.C. (2005). Estilos de personalidad y su relación con medidas de ansiedad y depresión: Datos normativos para el Inventario MIPS en adultos marplatenses. *Revista Iberoamericana de Diagnóstico y Evaluación - e Avaliação Psicológica*, 20(2), 79-103.
- Cattell, R.B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1, 245-276.
- Cheng, H., & Furnham, A. (2003). Personality, self-esteem, and demographic predictions of happiness and depression. *Personality and Individual Differences*, 34, 921-942.
- Claridge, G. (1997). Eysenck’s contribution to understanding psychopathology. In H. Nyborg (Ed.), *The scientific study of human nature: Tribute to Hans J. Eysenck at eighty* (pp.364-387). Oxford: Pergamon Press.
- Clarke, D. (2004). Neuroticism: Moderator or mediator in the relation between locus of control and depression?. *Personality and Individual Differences*, 37, 245–258.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.) Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cronbach, L.J., & Meehl, P.E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281-302.
- DeVellis, R.F. (1991). *Scale development: Theory and applications*. London: SAGE.
- Embretson, S.E., & Reise, S.P. (2000). *Item Response Theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Eysenck, H.J. (1991). Dimensions of personality: 16, 5 or 3? – Criteria for a taxonomic paradigm. *Personality and Individual Differences*, 12, 773-790.
- Eysenck, H.J. (1998). *Dimensions of personality*. New Brunswick, NJ: Transaction Publishers.
- 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. (1976). *Psychoticism as a dimension of personality*. London: Hodder and Stoughton.
- Eysenck, H.J., & Eysenck, S.B. (2008). *EPQ-R: Cuestionario Revisado de Personalidad de Eysenck: Versiones completa (EPQ-R) y abreviada (EPQ-RS)* (3ª ed.). Madrid: TEA Ediciones.
- Eysenck, S.B., Eysenck, H.J., & Barrett, P.T. (1985). A revised version of the Psychoticism scale. *Personality and Individual Differences*, 6, 21-29.
- Fisher, W.P. (2007). Rating scale instrument quality criteria. *Rasch Measurement Transactions*, 21, 1095.
- 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.
- 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.
- Glicksohn, J., & Abulafia, J. (1998). Embedding sensation seeking within the Big Three. *Personality and Individual Differences*, 25, 1085-1099.
- Hambleton, R.J. (1994). Guidelines for adapting educational and psychological tests. *European Journal of Psychological Assessment*, 10, 229-224.
- Hambleton, R.J. (2001). Guidelines for test translation/adaptation. *European Journal of Psychological Assessment*, 17, 164-172.
- 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.
- International Test Commission (ITC) (1999/2003). *Adaptação Portuguesa das Diretrizes Internacionais para a Utilização de Testes*. Lisboa: CEGOC.
- Kaiser, H.F. (1974). An index of factorial simplicity. *Psychometrika*, 39, 281-302.
- Kline, P. (1986). *A handbook of test construction: Introduction to psychometric design*. London: Methuen & Co.
- Kline, P. (1993). *Personality: The psychometric view*. London: Routledge.
- Kline, P. (1994). *An easy guide to factor analysis*. London: Routledge.
- Lima, M.P., & Simões, A. (1997). O Inventário da Personalidade NEO-PI-R: Resultados da aferição portuguesa. *Psychologica*, 18, 25-46.
- Linacre, J.M. (2009). *WINSTEPS 3.68.0: Rasch measurement computer program*. Chicago: Winsteps.com.
- Maruish, M.E. (Ed.) (2004a). *The use of psychological testing for treatment planning and outcomes assessment: Volume 1 – General considerations* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Maruish, M.E. (Ed.) (2004b). *The use of psychological testing for treatment planning and outcomes assessment: Volume 3 – Instruments for adults* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Messick, S. (1980). Test validity and the ethics of assessment. *American Psychologist*, 35, 1012-1027.
- 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.
- Moosbrugger, H., & Hartig, J. (2002). Factor analysis in personality research: Some artefacts and their consequences for psychological assessment. *Psychologische Beiträge*, 44, 136-158.
- Nias, D.K. (1986). Variedades do comportamento anormal. In H.J. Eysenck & G.D. Wilson (Eds.), *Manual de psicologia humana* (pp.321-338). Coimbra: Livraria Almedina.

- Nunnally, J.C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.
- 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.
- Pasquali, L. (2003). *Psicometria: Teoria dos testes na psicologia e na educação*. Petrópolis: Editora Vozes.
- Pasquali, L. (Ed.) (1999). *Instrumentos psicológicos: Manual prático de elaboração*. Brasília: LabPAM / IBAPP.
- Pestana, M.H., & Gageiro, J.N. (2005). *Análise de dados para ciências sociais: A complementaridade do SPSS* (4^a ed.). Lisboa: Edições Sílabo.
- Ruch, W. (1999). *Eine Untersuchung der revidierte Fassung des Eysenck Personality Questionnaire und die Konstruktion einer deutschen Standard- und Kurzfassung (EPQ-R bzw. EPQ-RK)*. Institut für Physiologische Psychologie, Heinrich-Heine-Universität Düsseldorf: Düsseldorf.
- Ruipérez, M.A., Ibáñez, M.I., Lorente, E., Moro, M., & Ortet, G. (2001). Psychometric properties of the Spanish version of the BSI: Contributions to the relationship between personality and psychopathology. *European Journal of Psychological Assessment*, 17, 241-250.
- 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.
- Silva, D.R. (2006). O Inventário de Estado-Traço de Ansiedade (STAI). In M.M. Gonçalves, M.R. Simões, L.S. Almeida, & C. Machado (Eds.), *Avaliação psicológica: Instrumentos validados para a população portuguesa* (Vol. I, 2.^a ed., pp.45-60). Coimbra: Quarteto Editora.
- Simões, M.R. (2012). Instrumentos de avaliação psicológica de pessoas idosas: Investigação e estudos de validação em Portugal. *Revista Iberoamericana de Diagnóstico y Evaluación - e Avaliação Psicológica*, 34(1), 9-33.
- Stewart, M.E., Ebmeier, K.P., & Deary, I.J. (2005). Personality correlates of happiness and sadness: EPQ-R e TPQ compared. *Personality and Individual Differences*, 38, 1085-1096.
- Vaz Serra, A., & Allen Gomes, F. (1973). Resultados da aplicação do *Maudsley Personality Inventory* a uma amostra portuguesa de indivíduos normais. *Coimbra Médica*, 20, 859-873.
- Vaz Serra, A., & Pio Abreu, J. (1973). Aferição dos quadros clínicos depressivos I: Ensaio de aplicação do Inventário Depressivo de Beck a uma amostra portuguesa de doentes deprimidos. *Coimbra Médica*, 20, 623-644.
- Vaz Serra, A., Ponciano, E., & Fidalgo Freitas, J. (1980). Resultados da aplicação do *Eysenck Personality Inventory* a uma amostra de população portuguesa. *Psiquiatria Clínica*, 1, 127-132.
- Wilson, G.D. (1986). Personalidade. In H.J. Eysenck, & G.D. Wilson (Eds.), *Manual de psicologia humana* (pp.181-200). Coimbra: Livraria Almedina.