

Funcionamento Diferencial de Itens (DIF) e experiência de afeto: questões de gênero

Differential Item Functioning and affect: gender issues

Pedro Pires* ¹; Alberto Filgueiras**; Rodolfo Ribas* &
Cristina Santana**

*Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil

**Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, Brasil

Resumo

O presente estudo teve por objetivo contemplar o critério de justiça dos itens da Escala de Afeto Positivo e Afeto Negativo (PANAS), considerando contribuições das últimas décadas sobre a evidência de diferenças de gênero na manifestação dos traços em questão. Com esse objetivo, a PANAS foi aplicada em 354 participantes, de ambos os sexos. Os dados foram avaliados sob o paradigma da TRI, segundo o modelo para itens politômicos sugerido por Andrich (1988), e o critério de justiça de acordo com o Funcionamento Diferencial de Itens, conforme os parâmetros sugeridos por Linacre (2011), comparando amostras com sujeitos do sexo feminino e masculino. As evidências do presente estudo apontam para itens que favorecem determinadas características dos grupos em questão.

Palavras-Chave: DIF, Psicologia positiva, Diferenças de gênero, PANAS.

Abstract

This article is an investigation into the justice criterion for the items of the Positive Affect and Negative Affect Schedule (PANAS) considering contributions from the last decades on gender differences as to emotion and affect experience. For this purpose, the PANAS was administered to 354 participants of both genders. Data was evaluated under the paradigm of IRT, according to the model of polytomous responses suggested by Andrich (1988), and the criterion of justice in accordance with the Differential Item Functioning on the parameters suggested by Linacre (2011) comparing male and female samples. Evidence indicates differential item functioning for specific items, favoring groups according to their characteristics.

Key-words: DIF, Positive psychology, Gender differences, PANAS.

¹ Contato: ppires85@gmail.com

O afeto é um construto de marcante relevância na psicologia. Consiste em reações instintivas do organismo, sendo importante na formação de emoções complexas (Zajonc, 1980). Na perspectiva cognitiva, o afeto cumpre um papel importante no julgamento e realização de escolhas dependendo da sua valência (e.g. raiva, medo e tristeza) (Lerner & Keltner, 2000). Ainda assim, Thompson (2007) aponta que o afeto ganhou importância na psicologia após os estudos de Bradburn, em 1969, e o estudo de Zajonc, em 1980, nos quais as dimensões do afeto foram tratadas como positivas ou negativas. Posteriormente, o tema foi investigado em diversos estudos (Diener & Emmons, 1984; Watson & Tellegen, 1985). Diener (1984) aponta também que o afeto cumpre um papel importante ao caracterizar o bem-

-estar subjetivo, principalmente ao conjugar os afetos positivo e negativo com os construtos vitalidade e satisfação com a vida.

A operacionalização do afeto na escala foi realizada por Watson, Clark e Tellegen (1988), que observaram a predominância de dois fatores principais em diversos estudos (Watson & Clark, 1994): afeto positivo e afeto negativo. Os autores realizaram um sumário desses estudos, no qual demonstraram forte evidência da existência uma díade fatorial encontrada a partir de soluções rotacionadas pelo método varimax. Normalmente essa solução empírica define as dimensões como: afeto positivo e afeto negativo. Contudo, em soluções não rotacionadas, foram contempladas dimensões de prazer-desprazer e excitação. Watson, Clark e Te-

llegen (1988) afirmam que até o momento do seu desenvolvimento, os instrumentos disponíveis não apresentavam um grau satisfatório de confiabilidade, sendo alguns deles desenvolvidos por análise fatorial, mas muitos por análises ad hoc.

Watson, Clark e Tellegen (1988) definem “afeto positivo” como o quanto que uma determinada pessoa sente de entusiasmo, atividade e alerta. Altos níveis de “afeto positivo” corresponderiam à energia, à concentração e ao engajamento prazeroso em atividades, em que um baixo nível de “afeto positivo” é caracterizado por tristeza e letargia. Em contraste, “afeto negativo” (Watson, Clark & Tellegen, 1988) é uma dimensão geral para o sofrimento subjetivo e engajamento desagradável, que normalmente reflete estados de raiva, desgosto, culpa, medo e nervosismo, sendo que um baixo nível de “afeto negativo” estaria relacionado a um estado de calma e serenidade.

A PANAS foi criada para representar os traços mencionados acima e foi inicialmente estruturada como uma lista composta por 60 itens, organizados em 20 categorias diferentes. Cada uma das 10 categorias corresponde às dimensões do afeto: negativa e positiva (Watson, Clark & Tellegen, 1988). Os itens foram estruturados com uma única palavra cada, com o objetivo de refletir os estados afetivos. A versão inicial da PANAS foi denominada como PANAS-X. A versão final do instrumento consiste de 20 itens selecionados a partir das características psicométricas do modelo completo. Watson, Clark e Tellegen (1988) reportam que a escala em sua forma completa possui uma qualidade

marcante a respeito de sua consistência interna. Essa propriedade foi verificada através do alfa de Cronbach, com valores acima de $\alpha = 0.84$ (o menor valor obtido para a escala de “afeto negativo”, usando a instrução de tempo de como a pessoa se sentiu durante o ano) e resultados consistentes na forma de validade por teste-reteste com diferentes instruções temporais, como: “no momento”, “hoje”, “durante os últimos dias”, “durante a última semana”, “ano” e “geral”. Os índices de confiabilidade obtidos não mudaram de forma significativa para cada uma das diferentes instruções oferecidas.

Ainda que a estrutura fatorial ortogonal apoiada na contribuição de Watson, Clark e Tellegen (1988) apresentou uma interpretação sólida, avanços foram realizados recentemente na discussão da sua estrutura fatorial. Leue e Beauducel (2011) colocam que, durante a última década, tem ocorrido um número crescente de estudos, inclusive transculturais, investigando modelos por análise fatorial confirmatória. Desses estudos, não apenas modelos ortogonais foram aceitos, mas também modelos oblíquos, considerando afetos positivo e negativo como correlacionados. Além do modelo de dois fatores, alguns estudos têm sugerido inclusive novos modelos fatoriais para a PANAS (e.g. modelo de três fatores, modelos hierárquicos, entre outros derivados das formas resumidas e expandida da PANAS). Pires, Filgueiras e Ribas (no prelo) apresentam a possibilidade de interpretação dos fatores através de modelos oblíquos exploratórios em amostra brasileira.

Em função de suas fortes propriedades psicométricas, a escala tem sido amplamente utilizada também no Brasil. A versão traduzida para o português brasileiro pode ser observada nos estudos de Pereira, Calvano e Cunha (1992), porém o estudo não contou com uma investigação de suas propriedades psicométricas ou parâmetros. Os autores implementaram modificações às instruções, especialmente incluindo a expressão “sintome”, com o objetivo de ajudar os participantes a alcançar uma resposta a nível pessoal. O tipo de resposta promovido também foi modificado, consistindo em níveis de concordância com o que é expresso no item. A versão original (Watson, Clark & Tellegen, 1988) contém respostas que representam a intensidade do afeto retratado pelo item. As propriedades psicométricas dessa versão foram recentemente investigadas pelo estudo de Pires, Filgueiras e Ribas (no prelo), em que a escala mostrou ser válida e com boas características psicométricas sobre a sua escalabilidade no paradigma de “teoria de resposta ao item”.

O campo de contribuições não se restringe ao plano psicométrico. Uma vez que o afeto é uma resposta emocional expressada ou observada, ele guarda correlatos neurais largamente estudados. Emoções positivas relacionadas a prazer, euforia e outras que corresponderiam à operacionalização proposta em Watson, Clark e Tellegen (1988) seriam moduladas por regiões ligadas ao sistema de recompensa. Gazzaniga e Heaton (2005) colocam, por um lado, que essa ativação positiva estaria associada a um au-

mento de dopamina. Por outro lado emoções como raiva, medo e tristeza teriam maior conotação de acordo com a operacionalização para afeto negativo, cuja modulação seria fortemente associada a um aumento de noradrenalina (Gazzaniga & Heaton, 2005). A questão dos correlatos neurais para o afeto ganha importância conforme contribuições apontam para comparações de gênero na modulação de emoções, mais achados psicométricos. Há ainda pouco consenso tanto nas neurociências, quanto nas investigações em psicometria.

Kring e Albert (1998) realizaram dois estudos para observar essas diferenças. Em um primeiro estudo, foi proposto que estudantes de graduação veriam cenas de filmes com forte teor emocional. Comparadas aos homens, mulheres se mostraram com maior expressividade, apesar de não diferir ao reportar a experiência emocional, demonstrando padrões diferenciados de resposta de condutância da pele. No segundo estudo realizado, Kring e Albert (1998) tomaram por base o mesmo procedimento, mas averiguando escalas de autoavaliação de expressividade emocional (Emotional Expressivity Scale), expressividade familiar (Family Expressiveness Questionnaire) e características de gênero (Bem Sex-Role Inventory). No contexto desses instrumentos, Kring e Albert (1998) apontam que o termo expressividade consiste na exteriorização de comportamentos relacionados ao construto avaliado. Os resultados foram replicados de acordo com o estudo prévio, sendo que características de gênero e expressividade fami-

liar se mostraram moderadores do relacionamento entre gênero e a expressividade emocional.

Os achados concordam em parte com a investigação de Fujita, Diener e Sandvik (1991), em que foi observado também maior intensidade de experiência de afeto em mulheres, mas, ainda assim, para o construto felicidade, gênero teria um impacto inferior a 1% se considerando a variância do objeto, enquanto que 13% quando se trata de intensidade do afeto. Os resultados em Fujita, Diener e Sandvik (1991) indicam que a experiência de felicidade é minimamente explicada pelas diferenças de gênero, enquanto que a magnitude da experiência de afeto seria mais intensamente regulada por essas diferenças. Fujita, Diener e Sandvik (1991) afirmam também que suas observações não entram em conflito com os achados acerca do impacto da depressão em mulheres, em que foi visto que elas possuem uma experiência mais intensa de afeto negativo. Para os autores, a capacidade averiguada das mulheres em sentir com maior magnitude emoções positivas serviria para equilibrar a sensação contrária.

McRae, Ochsner, Mauss, Gabrieli e Gross (2008) observam as dificuldades metodológicas e a escassez de contribuições, o que motivou seu estudo com uso de ressonância magnética. McRae, Ochsner, Mauss, Gabrieli e Gross (2008) investigaram, por uso de imagem em ressonância magnética funcional, a regulação emocional de homens e mulheres: foi pedido a participantes de ambos os gêneros que utilizassem

regulação cognitiva das emoções através da estratégia de reestruturação para figuras apresentadas com conotação negativa. A regulação cognitiva abordada na tarefa consiste na reapresentação ou recontextualização de um estímulo negativo de uma forma menos emotiva. Foi observado durante o estudo realizado que tanto homens, quanto mulheres foram capazes de reduzir a emoção negativa sentida, porém neurologicamente as diferenças de gênero emergiram. McRae, Ochsner, Mauss, Gabrieli e Gross (2008) constataram que, em comparação às mulheres, os homens mostraram um aumento menor na atividade de regiões pré-frontais, que são associadas à reestruturação cognitiva, redução significativa nas amígdalas, que estariam associadas à resposta emocional e menor envolvimento de regiões ventrais estriadas, associadas ao processamento de recompensas. McRae, Ochsner, Mauss, Gabrieli e Gross (2008) oferecem duas explicações possíveis para isso: a primeira seria que homens se engajariam menos na regulação cognitiva das emoções, utilizando-se provavelmente de processos de cunho mais automático, o que o estudo não pôde verificar de forma mais específica em função de limitações metodológicas (os processos automáticos não foram operacionalizados nem mensurados no estudo); a segunda explicação, que não seria concorrente, seria que as mulheres talvez usassem mais emoções positivas para reavaliar emoções negativas em sua maior parte.

Tendo em vista esse campo de contribuições, resta compreender se diferenças de gênero poderiam interferir com o critério de

justiça no funcionamento dos itens. O critério de justiça remete ao fato de que um determinado item deveria a priori ter um funcionamento que não favoreça populações específicas. Isso significa que aspectos culturais possam interagir com o construto, produzindo um resultado que reflete o traço latente, mas também outras variáveis (Sisto, 2006). Everson e Osterlind (2009) indicam que o papel central do *Differential Item Functioning* (DIF) é identificar quando um participante oferece respostas a determinados itens que estão conectados de forma sistemática a características pessoais, sem estar relacionados ao construto do teste. No presente estudo, tomam-se as diferenças biológicas e de desenvolvimento emocional como passíveis de interação com os construtos presentes na escala PANAS (Watson, Clark & Tellegen, 1988).

A importância de estudos como o presente seria em prevenir que, ao ser utilizado um instrumento psicológico, ele subestime ou superestime diferenças entre grupos conforme características que não seriam relevantes para o traço sendo avaliado. Esse esforço permite que futuros instrumentos incluam construtos que possam favorecer determinados grupos para então colaborar para a tomada de decisão em pesquisas de elaboração de medidas, para que não sejam elaborados itens com os mesmos problemas em questão (Osterlind & Everson, 2009; Sisto, 2006).

Método

Participantes

A presente pesquisa contou com 354 parti-

cipantes, sendo 239 do sexo feminino e 115 do sexo masculino, idade média de 29,5 (DP = 10,18) anos. Considerando o nível educacional, 148 participantes (41,8%) possuem até o ensino médio, 121 (34,18%) tem nível universitário e 85 (24,01%) tem pós-graduação em andamento. A amostra contou com ampla diversidade de sujeitos de localidades distribuídas pelo território brasileiro em função de sua aplicação online. Não ocorreram exclusões por formulários em branco, uma vez que a ferramenta de pesquisa do Google Docs desconsidera esses casos. Por outro lado, foi empregado procedimento de conversão dos escores da escala em escores Z, com o objetivo de detectar e eliminar outliers.

Instrumento

O instrumento é antecedido por um termo de consentimento, contemplando aspectos éticos da pesquisa, como o conhecimento sobre o conteúdo da pesquisa, anonimato, liberdade para interromper o procedimento, ausência de riscos e ônus, assim como benefícios imediatos ou retorno financeiro. Seguindo ao termo de consentimento, há uma lista de variável sociodemográfica pertinente ao estudo. A versão utilizada da PANAS para o presente artigo foi obtida da tradução e ajustes realizados por Pereira, Calvano e Cunha (1993), contando com respostas na forma de uma escala do tipo likert, com cinco graus de concordância, devendo os participantes reportar o estado afetivo no momento em que respondem. O uso repetido da expressão “sinto-me” foi mantido de acordo com a sugestão de Pereira, Calvano e

Cunha (1993) no sentido de reforçar ao participante a percepção do próprio afeto como uma sensação.

Procedimento

A aplicação foi realizada pela internet, utilizando a ferramenta de criação de formulários do Google Docs, com divulgação do link através de comunidades virtuais de sites de relacionamento (Orkut e Facebook), divulgação presencial na Universidade Federal do Rio de Janeiro e por listas de e-mail. Segundo Gosling, Vazire e Srivastav (2004), o rápido crescimento da internet promoveu novas oportunidades para pesquisadores, com uma nova forma de acesso a respondentes para pesquisas do tipo survey. Os autores apontam, tendo por base uma amostra total de 361.703 participantes, obtida em 510 publicações, que há uma diversidade amostral de gênero, status socioeconômico, região geográfica e idade no ambiente virtual. Além disso, os resultados contemplados em Gosling, Vazire e Srivastav (2004) indicam uma concordância entre os achados de estudos realizados com survey pela internet com os respectivos dados obtidos de forma convencional.

Análise dos dados

O modelo de análise do modelo de créditos parciais (PCM) foi aplicado, uma vez que as respostas contam com um total de cinco categorias (politômica) sugerido por Andrich (1988), derivado do modelo Rasch (1960). Esse modelo possui uma vantagem de manter as propriedades dos modelos Rasch (1960) de estimativa individual de parâme-

tros, sendo menos população-dependente se comparado aos demais modelos. Este realiza uma avaliação diferenciada para cada categoria de resposta dos itens, contando com um parâmetro próprio, em que a resposta se torna provável a partir de um determinado nível de aptidão ou traço. Para esse fim, o modelo de PCM dicotomiza as categorias para compor as respectivas curvas de probabilidade da categoria (CPC) (Ostini & Nering, 2006).

Para a detecção do funcionamento diferencial do item (DIF), a análise foi executada pelo programa Winsteps. Os critérios utilizados foram: DIF Medida, que fornece a dificuldade do item para as diferentes amostras avaliadas; DIF S.E., que indica a possibilidade de se observar a ocorrência de DIF com os dados obtidos; DIF Contraste, o qual reflete a diferença entre a dificuldade do item para cada amostra; ao passo que a prova de Welch t oferece uma estatística de significância para o DIF.

Resultados

Primeiramente, os pressupostos de unidimensionalidade (Ayala, 2009) foram verificados para cada escala separadamente, através dos índices de uma análise fatorial confirmatória. Respeitado o pressuposto de unidimensionalidade do modelo utilizado, procedeu-se com uma análise da precisão das escalas.

O índice de precisão do instrumento na subescala de “afeto positivo” obtido foi de 0,88 para as pessoas e 0,97 para os itens (índices item e person reliability do Winsteps). O erro de medida médio para as pessoas foi de 0,51 (DP = 0,10), enquanto que para os itens foi de 0,08 (DP = 0,01). Por sua vez, para a subescala de “afeto negativo”, foram obtidos índices de precisão 0,87 para as pessoas e 0,99 para os itens. O erro de medida médio encontrado para a dimensão de “afeto negativo” foi de 0,41 (DP = 0,14) para as pessoas, enquanto que para os itens foi de 0,06 (DP = 0,01). A Tabela 1 apresenta o sumário de medidas de ajuste e parâmetros quanto aos itens para ambas as escalas.

Tabela 1: *Parâmetros de ajuste de itens (N = 354)*

Parâmetros	Afeto negativo				Afeto positivo			
	Dificuldade	Infit	Outfit	Erro	Dificuldade	Infit	Outfit	Erro
Média	0,01	1,02	1,05	0,06	0,01	0,99	1,01	0,08
DP	0,54	0,18	0,20	0,01	0,49	0,18	0,19	0,01
Máximo	0,87	1,33	1,40	0,07	0,67	1,24	1,35	0,09
Mínimo	-0,87	0,70	0,71	0,06	-1,10	0,75	0,74	0,07

A média de infit foi de 1,02 (DP = 0,18) para afeto negativo e 0,99 (DP = 0,18) para afeto positivo, indicando que, num aspecto geral, grande parte dos itens foi respondida

dentro do padrão esperado para um ideal de 1,00 (Ayala, 2009). Os resultados apontam um intervalo de infit entre 0,70 e 1,33 para afeto negativo, enquanto que, para a subes-

cala de afeto positivo, o intervalo foi entre 0,75 e 1,24, sendo possível constatar que todos os itens se enquadraram no parâmetro entre 0,50 e 1,50 (Ayala, 2009; Linacre, 2011), indicativo de bom ajuste. Vale ressaltar que, em análises prévias, em estudo ainda não publicado por Pires, Filgueiras e Ribas (no prelo), o item “sinto-me alerta” foi excluído da análise em função de problemas de ajuste, indicando ambos os problemas dimensionais e nas estatísticas de fit, não se comportando como item da dimensão de afeto positivo. Para o presente artigo, o item foi excluído das análises por apresentar um viés de dimensionalidade, o que foge ao escopo do presente estudo.

As estatísticas de outfit, por sua vez, indicam uma média de 1,05 (DP = 0,20) para a subescala de afeto negativo e 1,01 (DP = 0,19) para afeto positivo, apontando boa adequação dos itens ao MCP. Os valores máximos de outfit foram maiores se comparados aos de infit, apesar de ainda localizados dentro dos parâmetros aceitos de 0,5 a 1,50. O intervalo obtido para o outfit foi de 0,71 a 1,40 para a subescala de afeto negativo, sendo que para afeto positivo o intervalo foi de 0,74 a 1,35.

Acerca do grau de dificuldade para os itens das escalas, o item com maior facilidade foi “sinto-me tenso” na subescala de afeto negativo para um parâmetro de dificuldade (b) de -0,87, enquanto que, na subescala de afeto positivo, o item “sinto-me interessado” apresentou $b = -1,01$. Os itens com maior b para as respectivas escalas foram “sinto-me transtornado” com $b = 0,87$ e “sinto-me seguro” em que apresentou $b =$

0,67. Os resultados apontam que traços de sensação de transtorno e segurança foram os mais difíceis de serem expressados, havendo maior expressão de interesse e tensão.

Os resultados para as análises de DIF apontam que os dados se comportam de forma a tornar possível a investigação de ocorrência do funcionamento diferencial para todos os itens no índice DIF S.E. acima de 0,00 como sugerido por Linacre (2011). Para prosseguimento na análise de DIF, serão observados ambos o contraste e a significância de acordo com o teste t de Welch (Linacre, 2011). Preocupa-se em não só abordar critérios de significância ($p < 0,05$), mas também o tamanho do DIF como alertado na bibliografia sobre o tema (Linacre, 2011; Ayala, 2009). Zwick, Thayer e Lewis (1999) classificam os contrastes de DIF em uma escala em que valores acima de logit $> 0,64$ indicam um DIF moderada à intensa, valores acima de logit $> 0,43$ indicam DIF leve a moderada, enquanto que, para valores abaixo, seriam irrelevantes.

Tendo em vista os critérios abordados, para afeto positivo podem ser contemplados dois itens que apresentam DIF leve à moderada (Tabela 2), sendo esses os itens “sinto-me seguro” (contraste logit = -0,58) e “sinto-me dinâmico” (contraste logit = 0,46), sendo o primeiro com DIF moderada, em função de sua proximidade com o limite superior da classe. Os respectivos valores para o teste t de Welch foram $t = 3,68$ ($p = 0,001$, significativo para $p < 0,05$) e $t = 2,91$ ($p = 0,003$, significativo para $p < 0,05$). A sensação de segurança foi mais facilmente manifestada por participantes do grupo

masculino, em que o item apresentou um grau de dificuldade de $b = 0,28$ em contraste ao grupo feminino, em que o item obteve $b = 0,86$. Por outro lado, o item “sinto-me dinâmico” foi expressado com favorecimento ao grupo feminino para um $b = 0,05$, o que no grupo masculino foi de $b = 0,51$.

Em afeto negativo, pode ser observado que há um total de três itens com DIF irre-

levante, entretanto se verifica a presença de ao menos um item com DIF leve à moderada (Tabela 3). O item que apresentou o fenômeno foi “sinto-me transtornado” (contraste logit = $-0,57$). O teste t de Welch aponta para um contraste significativo para $t = -3,80$ ($p = 0,001$).

Tabela 2: DIF para a subescala de afeto positivo ($N = 354$)

Grupo	DIF Medida	DIF S.E.	DIF Contraste	Welch		Itens
				t	p	
1	-1,26	0,16	-0,23	-1,21	0,228	interessado
1	-0,04	0,14	0,26	1,55	0,122	animado
1	0,33	0,13	-0,30	-1,87	0,062	inspirado
1	0,28	0,13	-0,58	-3,68	0,001	seguro*
1	-0,08	0,14	0,33	1,95	0,052	determinado
1	0,09	0,13	0,01	0,01	1,001	entusiasmado
1	0,51	0,13	0,46	2,91	0,003	dinâmico*
1	0,18	0,13	0,11	0,66	0,511	orgulhoso
1	0,01	0,13	0,01	0,01	1,001	atento
2	-1,03	0,11	0,23	1,21	0,228	interessado
2	-0,30	0,10	-0,26	-1,55	0,122	animado
2	0,63	0,09	0,30	1,87	0,062	inspirado
2	0,86	0,09	0,58	3,68	0,001	seguro*
2	-0,40	0,10	-0,33	-1,95	0,052	determinado
2	0,09	0,09	0,01	0,01	1,001	entusiasmado
2	0,05	0,09	-0,46	-2,91	0,003	dinâmico*
2	0,07	0,09	-0,11	-0,66	0,511	orgulhoso
2	0,01	0,09	0,01	0,01	1,001	atento

Nota: * DIF de contraste $>0,43$ e significativa para $p < 0,05$.

** DIF apenas significativa, mas de contraste irrelevante significativo para $p < 0,05$.

Os resultados indicam favorecimento para a manifestação do traço no grupo masculino, com $b = 0,50$ de dificuldade no item para a categoria, ao passo que o grupo feminino apresentou $b = 1,07$.

Apontados os problemas de funcionamento dos itens, o presente estudo prossegue com a discussão sobre as implicações dos presentes achados, assim como possíveis evidências na bibliografia comentada.

Tabela 3: DIF para a subescala de Afeto Negativo ($N = 354$)

Grupo	DIF Medi-da	DIF S.E.	DIF Contraste	Welch		Itens
				t	p	
1	-0,20	0,11	0,27	2,09	0,038	irritado**
1	-0,52	0,10	0,29	2,24	0,026	angustiado**
1	0,35	0,12	-0,23	-1,63	0,104	envergonhado
1	0,50	0,12	-0,57	-3,80	0,001	transtornado*
1	-0,52	0,10	-0,14	-1,11	0,269	nervoso
1	0,24	0,11	0,01	0,01	1,001	culpado
1	0,18	0,11	-0,02	-0,18	0,856	assustado
1	0,31	0,11	-0,12	-0,83	0,409	hostil
1	-0,84	0,10	0,06	0,45	0,650	tenso
1	0,39	0,12	0,28	1,99	0,047	amedrontado**
2	-0,47	0,07	-0,27	-2,09	0,038	irritado**
2	-0,81	0,07	-0,29	-2,24	0,026	angustiado**
2	0,58	0,08	0,23	1,63	0,104	envergonhado
2	1,07	0,09	0,57	3,80	0,001	transtornado*
2	-0,38	0,07	0,14	1,11	0,269	nervoso
2	0,24	0,08	-0,01	0,01	1,001	culpado
2	0,20	0,08	0,02	0,18	0,856	assustado
2	0,42	0,08	0,12	0,83	0,409	hostil
2	-0,89	0,07	-0,06	-0,45	0,650	tenso
2	0,11	0,08	-0,28	-1,99	0,047	amedrontado**

Nota: * DIF de contraste $>0,43$ e significativa para $p < 0,05$.

** DIF apenas significativa, mas de contraste irrelevante.

Discussão

Os dados encontrados no presente estudo apontam que há itens na escala PANAS (Watson, Clark & Tellegen, 1988) que favorecem características próprias de cada gênero, levando-se em consideração o mé-

todo de detecção de DIF por contraste (Zwick, Thayer, & Lewis, 1999) e os níveis de significância oferecidos pelo teste t de Welch (Linacre, 2011). A presente discussão se fez necessária uma vez que a bibliografia apoia que há diferenças de gênero em mani-

festações afetivas e emocionais (McRae, Ochsner, Mauss, Gabrieli, & Gross, 2008; Kring & Albert, 1998; Fujita, Diener, & Sandvik, 1991).

Para a subescala de afeto positivo, foi observado que os homens apresentam uma maior sensação de segurança, provavelmente em função do locus de controle desenvolvido ao longo da infância sobre o desenvolvimento de papéis sociais, apesar das correntes transformações no paradigma do mundo do trabalho (Shaffer & Kipp, 2009). Haveria uma cobrança sobre as dificuldades da inserção da mulher no mercado, o que poderia provocar uma menor sensação de segurança. Por outro lado, as mulheres teriam um melhor desempenho na sensação de dinamismo. A facilidade na reestruturação cognitiva de eventos poderia também ser a razão dessa sensação de dinamismo (McRae, Ochsner, Mauss, Gabrieli, & Gross, 2008), além de elementos relacionais.

Considerando a dimensão de afeto negativo, para os resultados obtidos, foi verificado que há uma probabilidade maior da sensação de transtorno para participantes do sexo masculino. Conforme já comentado acerca dos achados de McRae, Ochsner, Mauss, Gabrieli, & Gross (2008), novamente a possibilidade de reestruturação cognitiva poderia contribuir para a apresentação do traço no público masculino.

As diferenças aproximadas também poderiam ser explicadas em um paradigma desenvolvimentista. Marini (1988) observou que, tendo em vista o processo de socialização, há claras diferenças no desenvolvimento humano em relação ao gênero. Cria-

ções do sexo masculino seriam mais encorajados a expressar comportamentos tradicionalmente relacionados ao modelo masculino, reforçando agressividade e competitividade, que seriam barreiras para o desenvolvimento de intimidade, mas cumprem um papel forte em questão de assertividade – remete-se ao sentimento de segurança. Já crianças do sexo feminino seriam mais encorajadas a uma competitividade e agressividade mais restritas, sendo estimuladas a exibir comportamentos mais relacionados com aspectos sentimentais e de amparo (Marini, 1988), o que justificaria uma manifestação maior de dinamismo interacional como observado. Gilligan (1982) também apoia que a socialização dentro dos papéis de gênero leva as mulheres a desenvolverem um senso moral mais intenso na direção de questões ligadas a cuidados com as pessoas. Vale ressaltar que a generalização dos dados obtidos no presente estudo é restrita em função da metodologia de aplicação online. Como pode ser observado na sessão Participantes, o nível educacional deles não é composta de forma a retratar a população brasileira. Apesar dos resultados apontados por Gosling, Vazire, Srivastava, & John (2004), o uso de questionários online torna impossível o controle adequado de variáveis que possam intervir com as respostas.

Por fim, para Everson e Osterlind (2009), quando uma quantidade razoável de DIF é encontrada em um instrumento, é possível que haja dimensões diferenciadas em interação com os traços sendo investigados. O presente estudo encontrou DIF para três itens, o que representa 15% dos itens da

escala (para um total de vinte itens). Em estudo a ser realizado, pretende-se verificar propriedades de dimensionalidade tendo em vista as questões apresentadas sobre a influência de gênero em uma quantidade razoável de itens. Conforme os achados recorrentes, é importante que a elaboração de novas escalas leve em consideração as dimensões de gênero para itens que corroborem melhor com a problemática na instrumentação voltada para avaliação do afeto não apenas com respaldo em achados psicológicos, mas também levando em consideração toda a bibliografia que aponta para esses contrastes sociais existentes.

Considerando a PANAS, recomenda-se cautela no emprego da escala em pesquisas futuras em função de os resultados obtidos neste estudo apontarem para a sensibilidade do instrumento para a variável gênero, o que é desfavorável para pesquisas considerando esse tipo de corte amostral. Ainda assim é necessária uma investigação com um desenho de pesquisa que permita uma maior generalidade dos resultados, especialmente considerando o método e a amostra abordada, para definir a adequação ou inadequação da PANAS para pesquisas de gênero.

Referências

- Ayala, R. J. (2009). *The theory and practice of Item Response Theory*. Guilford Press, New York.
- Andrich, D. (1988). *Rasch models for measurement*. London: Sage Publications.
- Diener, E. (1984). Subjective well-being. *Psychological Bulletin*, 95(3), 542-575.
- Diener, E., & Emmons, R. A. (1984). The independence of positive and negative affect. *Journal of Personality and Social Psychology*, 47(5), 1105-1117.
- Everson, H., & Osterlind, S. (2009). *Differential Item Functioning*. London: Sage.
- Fujita, F., Diener, E., & Sandvik, E. (1991). Gender differences in negative affect and well-being: the case for emotional intensity. *Journal of Personality and Social Psychology*, 61, 427-34.
- Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should we trust web-based studies? A comparative analysis of six preconceptions about Internet questionnaires. *American Psychologist*, 59(2), 93-104.
- Gazzaniga, M. S., & Heatherton, T. F. (2005). *Ciência psicológica: mente, cérebro e comportamento*. Porto Alegre: Artmed.
- Gilligan, C. (1982). *A different voice*. Massachusetts: Harvard University Press.
- Kring, A. M., & Albert, H.G. (1998). Sex differences in emotion: expression, experience, and physiology. *Journal of Personality and Social Psychology*, 74(3), 686-703.
- Lawrence, R. H., & Liang, J. (1988). Structural integration of the affect balance scale and the life satisfaction index-a-race, sex, and age-differences. *Psychology and Aging*, 3(4), 375-384.
- Lerner, J. S., & Keltner, D. (2000). Beyond valence: toward a model of emotion-specific influences on judgment and choice. *Cognition and Emotion*, 14(4), 473-493.
- Leue, A., & Beaducel, A. (2011). The PANAS structure revisited: on the validity of a bifactor model in community and forensic samples. *Psychological Assessment*, 23, 215-225.
- Linacre, J. M. (2011, February). *A user's guide to winsteps: rasch-model computer programs*. John M. Linacre. Retrieved February 12, 2013, from www.winsteps.com
- Marini, M. M. (1988). Sociology of gender. In E. F. Borgatta, & K. S. Cook (Orgs.). *The future of sociology* (pp. 374-93). Beverly Hills, CA: Sage.
- McRae, K., Ochsner, K. N., Mauss, I. B., Gabrieli, J. D. E., & Gross, J. J. (2008). Gender differences in emotion regulation: an fmri study of cognitive reappraisal. *Group Processes and Intergroup Relations*, 11, 145-162.
- Ostini, R., & Nering, M. L. (2006). *Polytomous item response theory models*. London: Sage.
- Pereira, C. A. A., Calvano, N., & Cunha, V. C. (1992). Estados de ânimo e bem-estar subjetivo: um estudo com LEP, PANAS-S e BES. Reunião

- Annual de Psicologia da Sociedade Brasileira de Psicologia*, 22 (p.123). Ribeirão Preto, São Paulo.
- Pires, P., Filgueiras, F., Ribas, R. (no prelo). Positive and negative affect schedule: psychometric properties for the Brazilian Portuguese version. *Spanish Journal of Psychology*.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research.
- Shaffer, D. R., Kipp, K. (2009). *Developmental Psychology: childhood and adolescence*. Wadsworth: Cengage Learning.
- Shapiro, E. G., McPhee, J. T., Abbott, A. A., & Sulzbacher, S. I. (1994). Minnesota preschool affect rating scales: development, reliability, and validity. *Journal of Pediatric Psychology*, 19(3), 325-345.
- Sisto, F. F. (2006). O funcionamento diferencial do item. *Psico-USF*, 11, 35-43.
- Thompson, E. R. (2007). Development and validation of an internationally reliable short-form of the positive and negative affect schedule (Panas). *Journal of Cross-Cultural Psychology*, 38(2), 227-242.
- Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin*, 98(2), 219-235.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect – the PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1.063-1.070.
- Zajonc, R. B. (1980). Feelings and thinking: preferences need no inferences. *American Psychologist*, 35(2), 151-175.
- Zwick, R., Thayer, D. T., & Lewis, C. (1999). An empirical bayes approach to mantel-haenszel DIF analysis. *Journal of Educational Measurement*, 36, 1-28.

Recebido: 06/12/2011

Aceito: 19/12/2012