

Confiabilidade da versão em Português do Inventário de Fobia Social (SPIN) entre adolescentes estudantes do Município do Rio de Janeiro

Reliability of the Portuguese-language version
of the Social Phobia Inventory (SPIN) among
adolescent students in the city of Rio de Janeiro

Liliane Maria Pereira Vilete ¹
Evandro da Silva Freire Coutinho ²
Ivan Luiz de Vasconcellos Figueira ³

Abstract

It is believed that social phobia has its onset during adolescence and precedes other mental disorders; it is thus important to investigate the condition among young people. To date there is no self-reported scale validated for the Brazilian population. The present study investigated the reliability of the Portuguese-language version of the Social Phobia Inventory (SPIN) among adolescent students from public schools in the city of Rio de Janeiro. After SPIN was translated into Portuguese, a test-retest reliability study was carried out with 190 students. Intra-class correlation coefficient (ICC) and weighted kappa (w^2) were estimated, log-linear models were fitted, and Bland & Altman graphs were built. The Portuguese version showed good internal consistency (Cronbach's $\alpha = 0.88$) and good reliability for total score (ICC = 0.78). Reliability for single items was not very good (w^2 between 0.32 and 0.65). The "semi-association" model fit most of the items best. Based on these findings we concluded that the Portuguese-language version of SPIN showed good reliability results, similar to those obtained with the original English-language version.

Phobic Disorders; Psychometrics; Reliability and Validity

Introdução

A fobia social é, atualmente, definida como um medo acentuado e persistente de ser julgado, criticado ou humilhado em situações sociais e de desempenho. Desta forma, o indivíduo acometido por esse transtorno de ansiedade enfrentaria tais situações com grande sofrimento, chegando, muitas vezes, a evitá-las, o que levaria a um conseqüente prejuízo no seu funcionamento acadêmico, social ou ocupacional ¹. Desde a sua inclusão, em 1980, na terceira edição da classificação diagnóstica DSM (*Diagnostic and Statistical Manual*), a fobia social vem sendo descrita como tendo, mais freqüentemente, início na adolescência e curso crônico ^{2,3,4}. Vem também sendo observada uma elevada proporção de co-morbidade com este transtorno mental, dentre elas: agorafobia, fobia simples, transtorno de somatização, depressão maior, transtorno obsessivo-compulsivo, distímia, transtorno afetivo bipolar, abuso de álcool e outras substâncias ⁵. Uma vez que os indivíduos entrevistados costumam referir que a fobia social precederia estes outros transtornos ^{6,7}, questiona-se se a intervenção precoce em uma condição potencialmente tratável poderia ajudar a prevenir estas complicações tardias. No entanto, a maioria desses estudos é do tipo transversal, não sendo possível descartar, nesses casos, um viés de memória.

¹ Instituto Phillippe Pinel, Rio de Janeiro, Brasil.

² Departamento de Epidemiologia e Métodos Quantitativos em Saúde, Escola Nacional de Saúde Pública, Fundação Oswaldo Cruz, Rio de Janeiro, Brasil.

³ Instituto de Psiquiatria, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil.

Correspondência

Liliane Maria Pereira Vilete
Instituto Phillippe Pinel.
Av Venceslau Bráz 65,
Rio de Janeiro, RJ
22290-140, Brasil.
liliane.vilete@hotmail.com

SOMATIZAÇÃO: distúrbio crônico caracterizado pelo aparecimento de vários sintomas físicos para os quais não se encontra nenhuma lesão orgânica objectiva, e que resulta do recalçamento de tendências instintivas.

Apesar de a ansiedade social ser uma experiência normal cuja primeira manifestação ocorre precocemente na infância, em algumas crianças e adolescentes, entretanto, níveis intensos de ansiedade podem provocar um impacto negativo sobre seu desenvolvimento e bem-estar. Os quadros patológicos irão se manifestar pelo grau de dificuldade para lidar com os medos, do prejuízo sobre o funcionamento social ou acadêmico-ocupacional e da gravidade do desconforto subjetivo sentido⁸.

As manifestações clínicas de fobia social basicamente não diferem nas diversas faixas etárias. As diferenças podem apenas refletir as oportunidades para o engajamento em diferentes tipos de situações sociais, e habilidades cognitivas distintas podem limitar a compreensão do problema por parte de uma criança muito pequena⁹. Como as crianças também têm menos oportunidades de evitar as situações temidas, quase sempre sendo obrigadas a enfrentá-las, freqüentemente, apresentam um padrão maior de angústia do que de evitação. Entretanto, a fobia social é uma costumeira causa de evitação da escola entre os adolescentes, o que pode ser interpretado como “vadiagem”^{10,11}.

Nessa faixa etária, também são descritas limitações funcionais decorrentes da fobia social: em um estudo entre pré-adolescentes com diagnóstico de fobia social, 75% relataram que possuíam poucos amigos ou nenhum; 50% relataram que não gostavam da escola; e 10% se recusavam regularmente a freqüentá-la. Além disso, 35% das situações ansiogênicas resultavam em estratégia de evitação⁹.

A presença de fobia social em adolescentes também vem sendo, como nos adultos, associada a uma maior prevalência de transtorno depressivo maior e ideação suicida. Tentativas de suicídio ocorrem mais comumente na vigência do transtorno depressivo, mas estas são ainda mais prevalentes quando existe fobia social associada⁷. Este transtorno também vem apresentando forte associação com outras comorbidades, como transtorno somatoforme e abuso/dependência de álcool. Entretanto, apenas uma pequena parcela dos adolescentes fóbicos sociais procuram tratamento¹².

Em razão da dificuldade de realização de estudos epidemiológicos baseados nas entrevistas diagnósticas disponíveis – por serem muito extensas, consumirem muito tempo e representarem um custo elevado – vários instrumentos de autopreenchimento para a detecção da fobia social vêm sendo desenvolvidos, na tentativa de viabilizarem um maior número de estudos. No entanto, para que estes instrumentos, na maior parte das vezes elaborados em

países ocidentais desenvolvidos, sejam utilizados em realidades sócio-culturais tão distintas, é necessário um processo de tradução e adaptação abrangentes, na tentativa de alcançar uma equivalência cultural, com posterior estudo de sua validade nesta nova população¹³. Até o momento, todavia, nenhum instrumento de autopreenchimento para a detecção da fobia social foi validado para a população brasileira.

O *Inventário de Fobia Social* (SPIN – *Social Phobia Inventory*) é um instrumento que consiste de apenas 17 itens, que abarcam três importantes critérios que definem a fobia social: o medo, a esquiva das situações e os sintomas de desconforto físico. Engloba tanto situações de desempenho quanto de interação social. Para cada item do questionário, solicita-se ao indivíduo que indique o quanto as situações ou sintomas descritos o incomodaram na última semana, devendo este marcar uma entre as cinco opções, que variam de “Not at all” a “Extremely”. A pontuação para cada uma das graduações varia, portanto, de 0 a 4, e a pontuação total irá variar de 0 a 68. Foi realizada, pelos autores do instrumento, uma avaliação psicométrica com 353 indivíduos com média de idade de 36 anos, sendo os fóbicos sociais obtidos de três ensaios clínicos e os controles de dois grupos (voluntários saudáveis de um estudo de pressão arterial – I, e pacientes psiquiátricos ambulatoriais com outro diagnóstico do que fobia social – II). O instrumento demonstrou boa consistência interna (de Cronbach variando de 0,82 a 0,94) e boa confiabilidade teste-reteste, com coeficiente de correlação de Spearman variando de 0,78 a 0,89. Também pelo método desenvolvido por Bland & Altman¹⁴, demonstrou boa confiabilidade¹⁵.

Neste artigo apresentamos o processo de tradução e adaptação cultural do SPIN em nosso meio, bem como a avaliação de sua confiabilidade em uma população de adolescentes.

Material e métodos

Equivalência semântica

O processo de equivalência semântica foi baseado em metodologia utilizada por Reichenheim et al.¹⁶.

Inicialmente, foram feitas duas traduções do SPIN para o Português, de forma independente, por dois psiquiatras com experiência em fobia social.

As traduções foram retrovertidas, também de forma independente, por dois tradutores bilíngües que não conheciam a versão original em

obs:
Será que o irmão mais velho
tbn teve fobia social?
E passou pra outro estágio
de transtorno?

inglês. Essas novas versões foram comparadas com o texto original, para a correção de discrepâncias e elaboração de uma versão final de consenso. Priorizou-se a utilização de um vocabulário coloquial, com o qual a população alvo do estudo pudesse estar familiarizada.

Posteriormente, realizamos um pré-teste com vinte adolescentes de uma escola da rede pública do Município do Rio de Janeiro e de um ambulatório de psiquiatria infanto-juvenil, com idades entre 13 e 18 anos, para a avaliação da compreensão do instrumento e correções necessárias. Com base em suas sugestões, foi elaborada a versão final em português do questionário SPIN, incluindo algumas explicações entre parênteses para aquelas expressões consideradas de difícil compreensão. Tal procedimento foi consentido pelos autores do questionário original, que aprovaram a versão em português, fundamentando-se na similaridade entre a versão original e a retrotradução.

População do estudo

A versão final do questionário SPIN foi aplicada em estudantes de 5ª a 8ª série de duas escolas públicas do **Município do Rio de Janeiro**, situadas no bairro de Ramos e que possuíam, ao todo, cerca de mil alunos matriculados. A participação destes foi condicionada à devolução do termo de consentimento livre e esclarecido assinado pelo responsável e pelo aluno, ou somente pelo último, quando maior do que 18 anos. **Os motivos apresentados pelos os alunos para a não-participação da pesquisa foram em sua maioria: perda do termo de consentimento, esquecimento de entregá-lo ao responsável, ou de trazê-lo preenchido de volta para a escola. Poucos alunos se recusaram explicitamente a participar e alguns referiram não terem obtido a permissão dos responsáveis.**

Os questionários foram distribuídos nas salas de aula, com esclarecimento por parte da autora de quaisquer dúvidas apresentadas pelos alunos durante o seu preenchimento. Trezentos e noventa e oito alunos devolveram o termo de consentimento assinado pelo responsável (ou pelo próprio, no caso de maior de 18 anos) e preencheram o SPIN pelo menos uma vez. Desses 398, 190 participaram do estudo de confiabilidade, preenchendo-o uma segunda vez. Embora não tenha sido possível repetir o teste em toda população inicial pelo limite de tempo disponível para o estudo, selecionamos 190 alunos na tentativa de ter, no mínimo, trinta indivíduos em cada um dos estratos a serem comparados na análise. Participaram do reteste, portanto, todos os alunos que

entregaram o termo de consentimento em tempo hábil para a realização do estudo.

Desenho teste-reteste

A confiabilidade teste-reteste da tradução do SPIN foi avaliada por intermédio da aplicação do questionário em dois momentos, com intervalo de 15 dias na primeira escola e de 13 dias na segunda. Vinte questionários foram excluídos por apresentarem algum erro de preenchimento. A confiabilidade foi investigada tanto para cada uma das 17 perguntas do SPIN quanto para o seu escore total.

O projeto foi aprovado pelo Comitê de Ética em Pesquisa da Escola Nacional de Saúde Pública da Fundação Oswaldo Cruz.

Análise dos dados

Análise descritiva

Os escores do teste e do reteste foram calculados e estratificados segundo as características dos estudantes. As médias foram comparadas pela utilização do teste de Mann-Whitney, já que as distribuições não eram simétricas.

Confiabilidade do escore do questionário

A confiabilidade da pontuação total da tradução do SPIN foi avaliada mediante o coeficiente de correlação intraclasse (CCIC) e dos gráficos de Bland & Altman¹⁴. No caso do CCIC, o cálculo foi realizado pela análise de variância com dois fatores e efeitos aleatórios (*two-way random*)¹⁷.

Confiabilidade das perguntas do questionário

A confiabilidade dos 17 itens do questionário foi avaliada pelo kappa ponderado, uma vez que estes itens têm o formato de variáveis ordinais. Esta estatística leva em conta não apenas a concordância perfeita entre as duas avaliações, que excede aquela esperada ao acaso (como no kappa simples). O kappa ponderado avalia a concordância considerando não só o grau de concordância perfeita (diagonal principal da tabela de contingência), mas também a magnitude da discordância, ao atribuir pesos diferenciados de acordo com a maior ou menor proximidade entre as categorias da variável¹⁸. A ponderação foi feita pelo método do quadrado dos erros¹⁹.

No entanto, embora amplamente utilizada, a estatística kappa, por ser uma medida resu-

mo, não permite investigar a estrutura da concordância dos dados. A investigação da estrutura de concordância e discordância consiste em estudar, na tabela de contingência, as frequências observadas e esperadas ao acaso, tanto da diagonal principal quanto fora desta. Para superar esta limitação foram ajustados modelos log-lineares conforme a metodologia descrita em Agresti ²⁰, Graham & Jackson ²¹, May ²², Silva & Pereira ²³. Neste estudo foram **ajustados sete modelos:**

a) **Independência:** A concordância observada entre o teste e o reteste é igual àquela esperada ao acaso.

$$\log m_{ij} = \mu + \frac{A}{i} + \frac{B}{j}$$

b) **Concordância diagonal:** Nesse modelo, só existe associação entre as respostas do teste e do reteste quando a concordância é exata. Essa concordância é maior do que aquela esperada ao acaso, isto é, excede aquela que seria obtida caso os resultados do teste e do reteste fossem independentes. No caso de discordância entre as respostas (células fora da diagonal da tabela), o modelo assume que as avaliações do teste e do reteste são independentes.

$$\log m_{ij} = \mu + \frac{A}{i} + \frac{B}{j} + \gamma_{ij}$$

c) **Quasi-independência:** Assim como no modelo anterior, não há associação fora da diagonal da tabela de contingência, além daquela esperada ao acaso. A diferença entre esses modelos é que, no caso do diagonal, assume-se um parâmetro único para todas as células da diagonal da tabela. No modelo de *quasi*-independência, assumem-se parâmetros distintos para as células da diagonal, ou seja, a concordância varia entre as células da diagonal. Em outras palavras, certas categorias da variável que está sendo avaliada têm maior ou menor propensão à concordância.

$$\log m_{ij} = \mu + \frac{A}{i} + \frac{B}{j} + \gamma_{ij}$$

onde $\gamma_{ij} = \gamma_i$ caso $i = j$; caso contrário, $\gamma_{ij} = 0$

d) **Associação linear por linear:** Este modelo assume, ao contrário dos anteriores, que existe algum grau de associação entre as observações que estão situadas nas células fora da diagonal principal da tabela. Neste caso, embora as respostas nos dois momentos não sejam idênticas, há uma tendência dos escores baixos (altos) no teste corresponderem a escores baixos (altos) no reteste. O modelo considera ainda que os intervalos entre as categorias da variável ordinal são uniformes.

$$\log m_{ij} = \mu + \frac{A}{i} + \frac{B}{j} + \gamma_{ij}$$

e) **Concordância diagonal + associação linear por linear:** Combina os padrões “b” e “d”, vale dizer, a distribuição dos dados na tabela pode ser explicada por um padrão de concordância (dados na diagonal da tabela) e por um padrão de discordância em que pontuações baixas (altas) no teste correspondam a pontuações baixas (altas) no reteste.

$$\log m_{ij} = \mu + \frac{A}{i} + \frac{B}{j} + \gamma_{ij} + \delta_{ij}$$

f) **Semi-associação:** Combina os padrões “c” e “d”, a saber, a distribuição dos dados na tabela pode ser explicada por um padrão de concordância diferente para cada categoria da variável, e um padrão de discordância onde existe uma tendência em que pontuações baixas (altas) no teste correspondam a pontuações baixas (altas) no reteste.

$$\log m_{ij} = \mu + \frac{A}{i} + \frac{B}{j} + \gamma_{ij} + \delta_{ij}$$

onde $\gamma_{ij} = \gamma_i$ caso $i = j$; caso contrário, $\gamma_{ij} = 0$

g) **Associação triangular:** Respostas do teste têm pontuação sistematicamente mais altas que as do reteste, e vice-versa.

$$\log m_{ij} = \mu + \frac{A}{i} + \frac{B}{j} + \gamma_{ij}$$

onde $\gamma_{ij} = \gamma_i$ caso $i > j$; caso contrário, $\gamma_{ij} = 0$

O ajuste de cada modelo – isto é, a comparação entre as frequências observadas e as estimadas por este – foi avaliado por meio da estatística ² da razão de verossimilhança (G^2). Consideramos que o modelo se ajustou bem aos dados, quando o p-valor foi maior que 0,10 (ou seja, quando os valores estimados pelo ajuste do modelo não diferiram significativamente dos valores observados, para um nível de significância de 10%). Entre os modelos que atenderam a essa exigência, aquele considerado de melhor ajuste aos dados foi o mais simples (pelo critério de parcimônia) e com menor valor de G^2 (para os respectivos graus de liberdade). Os programas estatísticos utilizados para a análise dos dados foram o **SPSS** (para cálculos dos coeficientes de correlação intraclass, para gerar gráficos Bland & Altman e para o ajuste de modelos log-lineares) e o Stata (para o cálculo das estatísticas kappa).

Resultados

Descrição da população do estudo

Como já referido anteriormente, dos 190 alunos que participaram do estudo teste-reteste, 20 foram excluídos do estudo da confiabilidade global do SPIN, por não ter sido possível calcular o escore final em virtude de algum erro de preenchimento dos questionários, como, por exemplo, alguma resposta em branco ou duas respostas para um mesmo item.

Dos 170 alunos restantes, 139 eram da Escola Municipal João Barbalho (Escola I) e 31 da Escola Municipal Padre Manoel (Escola II). A maioria dos participantes foi do sexo feminino (65,9%), com média e mediana de idade em torno dos 13 anos, com um mínimo de 10 e um máximo de 21 anos. Setenta e cinco por cento dos alunos tinham menos de 14 anos e 67% estudavam até a 6ª série. Cerca de 40% dos alunos haviam sido reprovados em pelo menos uma série, sendo que 92% destes foram reprovados até duas vezes. A descrição desta população pode ser observada na Tabela 1.

Os escores obtidos variaram de 1 a 52 no teste (primeira aplicação do questionário) e de 0 a 52 no reteste (segunda aplicação). A média de escores do teste foi de 18,95 e a mediana, 17. A média de escores do reteste foi muito próxima: 18,08; e a mediana foi a mesma do teste anterior. A média obtida entre as adolescentes foi significativamente maior do que a média obtida entre os estudantes do sexo masculino – 20,77 e 15,45, nos testes ($p = 0,001$); 19,79 e 14,78 nos retestes ($p = 0,002$). Os estudantes que apresentaram menor pontuação foram aqueles menores de 12 anos de idade.

Consistência interna

A consistência interna do questionário, vale dizer, a consistência entre os seus itens, estimada pelo α de Cronbach foi igual a 0,88.

Confiabilidade da pontuação total do questionário

• Método Bland & Altman

Observando a Figura 1, entre os estudantes do sexo feminino, quanto maiores os escores do teste, aparentemente maior foi a discordância entre a pontuação total das duas aplicações do questionário. Já entre os estudantes do sexo masculino (Figura 2), não podemos observar nenhum padrão de discordância, embora essa conclusão fique prejudicada pelo menor nú-

Tabela 1

Descrição da população do estudo de confiabilidade.

	Escola I		Escola II		Total	
	n	%	n	%	n	%
Sexo						
Masculino	47	33,8	11	35,5	58	34,1
Feminino	92	66,2	20	64,5	112	65,9
Escolaridade						
5ª série	54	38,8	2	6,5	56	32,9
6ª série	39	28,1	19	61,3	58	34,1
7ª série	24	17,3	0	0,0	24	14,1
8ª série	22	15,8	10	32,3	32	18,8
Reprovação						
Sim	50	36,0	17	54,8	67	39,4
Não	89	64,0	14	45,2	103	60,6
Grupo etário (anos)						
10-11	21	15,1	0	0,0	21	12,4
12-13	69	49,6	9	29,0	78	45,9
14-15	35	25,2	13	41,9	48	28,2
16-17	10	7,2	7	22,6	17	10,0
18	4	2,9	2	6,5	6	3,5
Total	139	81,8	31	18,2	170	100,0

mero de indivíduos desse grupo e a rarefação dos pontos no gráfico a partir do escore 20.

• Coeficiente de Correlação Intraclass

Foram calculados os CCIC para o escore global do SPIN. Este coeficiente foi estimado para o conjunto dos alunos, assim como para subgrupos estratificados por sexo, grupo etário, escolaridade, reprovação e escola. A estratificação foi realizada a fim de que pudéssemos analisar o efeito de cada um desses fatores na confiabilidade do questionário. O coeficiente de correlação intraclass global estimado foi 0,78 (com intervalo de confiança de 95% de 0,72 a 0,83), sugerindo uma boa confiabilidade da versão do questionário SPIN, nessa população.

A diferença observada entre o sexo feminino (CCIC = 0,79) e masculino (CCIC = 0,72) não foi estatisticamente significativa para um nível de 5%, uma vez que os intervalos de confiança do sexo feminino incluem a estimativa pontual do sexo masculino, e vice-versa. Também não foi observada diferença estatisticamente significativa entre os diferentes grupos etários, ou quando comparamos alunos que haviam sido reprovados com aqueles sem nenhuma reprovação.

Figura 1

Gráfico de Bland & Altman para os escores do teste e reteste (sexo feminino).

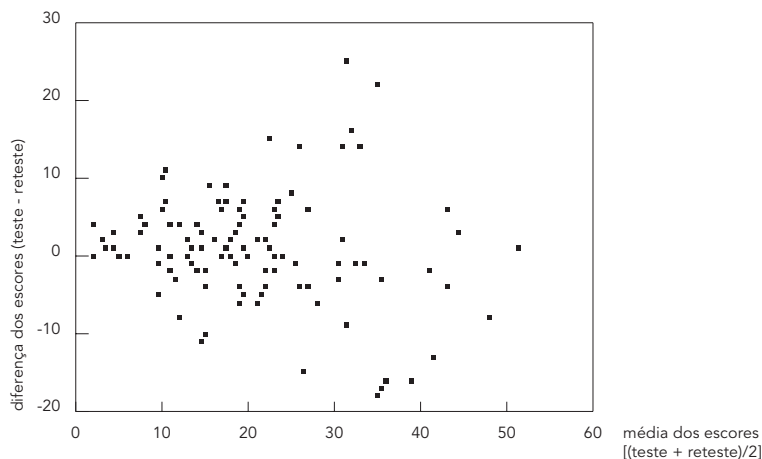
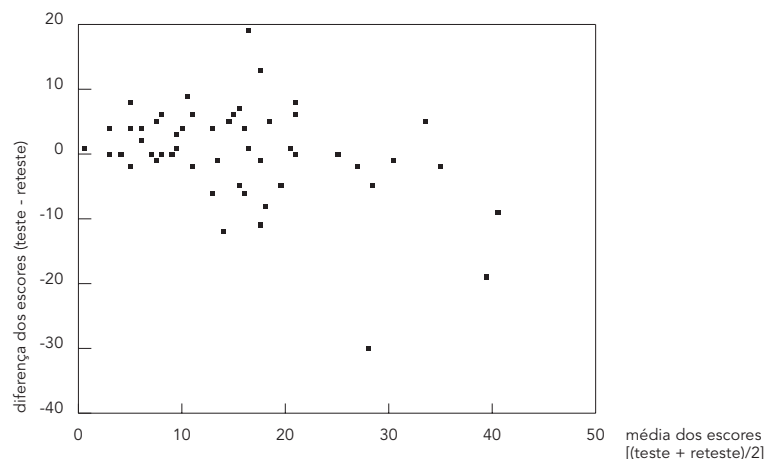


Figura 2

Gráfico de Bland & Altman para os escores do teste e reteste (sexo masculino).



Já na estratificação por escolaridade, pudemos verificar diferença entre os coeficientes de alguns estratos (0,87; 0,76; 0,53 e 0,82, respectivamente para as 5ª, 6ª, 7ª e 8ª séries), com CCIC mais elevados nos extremos. A estimativa do coeficiente do terceiro estrato apresentou amplo intervalo de confiança, não sendo significativa essa sua diferença dos demais.

Uma diferença maior pode ser notada quando comparamos os coeficientes de correlação

intraclasse das duas escolas (0,76 para a escola I e 0,91 para a escola II). Esta tendência foi mantida mesmo com nova estratificação para as demais variáveis, apesar da perda de significância estatística, possivelmente causada pela redução da amostra nos estratos.

Confiabilidade das perguntas do questionário

No cálculo da confiabilidade para cada um dos 17 itens da versão do questionário SPIN, foram obtidos valores de kappa ponderado variando de 0,32 a 0,65 (Tabela 2). Os itens 1 e 4 do questionário (*“Eu tenho medo de autoridades”* e *“Eu evito falar com pessoas que eu não conheço”*, respectivamente) apresentaram as menores concordâncias entre as respostas do teste e reteste, com valor de kappa considerado sofrível ou pobre^{24,25}. Os itens que apresentaram as maiores concordâncias entre as respostas da primeira e da segunda aplicação do questionário foram: o item 7 (*“Suar na frente dos outros me causa mal-estar”*); o item 11 (*“Eu evito ter que fazer discursos – como falar na frente da turma ou para uma platéia”*); o item 13 (*“Palpitações – batidas fortes ou rápidas – do coração me incomodam quando eu estou perto dos outros”*) e o item 17 (*“Tremer na frente dos outros me causa mal-estar”*), com uma substancial (ou boa) estimativa do kappa. Para os demais itens, foram estimados valores situados na faixa de concordância considerada razoável ou moderada.

A análise da confiabilidade para cada item do questionário também foi estratificada pelas demais variáveis.

Não foi observada diferença estatisticamente significativa entre os sexos, entre as escolas, entre os grupos etários, escolaridade ou na presença ou ausência de reprovação. Com a estratificação, e conseqüente redução da amostra em cada um dos estratos, também houve aumento da amplitude dos intervalos de confiança, com diminuição da precisão dos valores encontrados.

Os valores estimados de kappa não ponderado foram ainda menores do que aqueles estimados a partir da ponderação. Este coeficiente é uma medida da exata concordância entre as respostas dos testes e dos retestes, que excede a concordância esperada ao acaso.

Com o intuito de identificar a estrutura de concordância dos dados, foram ajustados diferentes modelos log-lineares (Tabela 3).

Em nenhum dos itens do questionário SPIN, o modelo de “independência” se ajustou aos dados, indicando que a concordância observada entre as respostas do teste e reteste não po-

Tabela 2

Confiabilidade teste-reteste. Valores de kappa ponderado para cada item do questionário SPIN.

	Global	
	w2	IC95%
1. Eu tenho medo de autoridades (p.ex, professores, instrutores, diretor etc.).	0,36	0,22-0,50
2. Eu fico incomodado de corar (ficar vermelho) na frente dos outros.	0,45	0,31-0,59
3. Festas e eventos sociais me assustam.	0,45	0,31-0,59
4. Eu evito falar com pessoas que eu não conheço.	0,32	0,18-0,46
5. Ser criticado me assusta muito.	0,50	0,36-0,64
6. O medo de ficar constrangido me faz evitar fazer coisas ou falar com outras pessoas.	0,45	0,31-0,59
7. Suar na frente dos outros me causa mal-estar.	0,61	0,47-0,75
8. Eu evito ir à festas.	0,55	0,41-0,69
9. Eu evito atividades nas quais sou o centro das atenções.	0,52	0,38-0,66
10. Falar com estranhos me assusta.	0,43	0,29-0,57
11. Eu evito ter que fazer discursos (como falar na frente da turma ou para uma platéia).	0,65	0,51-0,79
12. Eu faria qualquer coisa para evitar ser o centro das atenções.	0,45	0,31-0,59
13. Palpitações (batidas fortes ou rápidas) do coração me incomodam quando eu estou perto dos outros.	0,65	0,51-0,79
14. Eu tenho medo de fazer coisas quando as pessoas possam estar olhando.	0,44	0,30-0,58
15. Ficar constrangido ou parecer estúpido estão entre meus piores medos.	0,48	0,34-0,62
16. Eu evito falar com qualquer autoridade (p.ex, professores, instrutores, diretor etc.).	0,53	0,39-0,67
17. Tremer na frente dos outros me causa mal-estar.	0,63	0,49-0,77

de ser atribuída somente ao acaso, e que algum tipo de relação deve existir entre as duas medidas.

Nos itens 1, 2, 5, 7, 9, 10, 11, 12, 15, 16 e 17, o modelo que melhor se ajustou aos dados foi o de “semi-associação”. Na diagonal principal, a concordância das respostas entre o teste e reteste foi maior para as categorias “Nada” e “Um pouquinho”. Quando não houve concordância exata das respostas, ainda assim se verifica um padrão de discordância em que pontuações baixas (altas) no item do teste correspondam a pontuações baixas (altas) no mesmo item do reteste.

Já nos itens 3, 4, 6 e 8, o modelo que melhor se ajustou aos dados foi o de “quasi-independência”. Na diagonal principal, a concordância das respostas entre o teste e reteste também foi maior para as categorias “Nada” e “Um pouquinho”. Nas demais células fora da diagonal principal, não se observou associação além daquela esperada ao acaso.

No item 13, o modelo que melhor se ajustou aos dados foi o modelo de “concordância diagonal + associação linear por linear”. Este indica que, além da concordância exata entre as respostas e da concordância por conta do acaso, pode também ser observada uma tendência em que altas (baixas) pontuações no item do teste estejam associadas a altas (baixas) pon-

tuações no mesmo item do reteste, quando as respostas não concordam exatamente.

No item 14, nenhum dos modelos investigados foi capaz de explicar adequadamente a distribuição dos dados, ainda que o modelo de semi-associação tenha alcançado um nível limítrofe de significância estatística (p-valor = 0,08).

Discussão

Assim como na avaliação psicométrica do instrumento original, a versão do SPIN para o português também apresentou um grau de consistência interna das respostas considerado bom, quando utilizado para pesquisa, e com valor muito próximo ao que Bland & Altman²⁶ consideram necessário para aplicação clínica (de Cronbach > 0,90). Essa boa consistência interna sugere que existe correlação entre os itens da escala, que representam um mesmo domínio – no caso, a fobia social.

Foi observado, também, um bom desempenho do instrumento no que diz respeito à confiabilidade teste-reteste do seu escore geral, sugerindo boa reprodutibilidade e estabilidade dos dados, **que parece não depender do sexo, da idade ou do histórico de reprovação dos alunos – indicando que o instrumento pode ser utilizado nesses diferentes subgrupos.** Contudo,

Tabela 3

Modelos Log-lineares. Valores da razão de verossimilhança (G^2) e graus de liberdade (df) para cada um dos 17 itens do questionário SPIN. Os modelos que melhor se ajustaram aos dados se apresentam sombreados na tabela.

Itens do SPIN	Independência		Concordância diagonal		Associação Linear por Linear		Descrição do modelo Concordância diagonal + Associação Linear por Linear		Parâmetros Triangulares		Semi-Associação		Quasi-Independência	
	G^2	df	G^2	df	G^2	df	G^2	df	G^2	df	G^2	df	G^2	df
1.	91,59	16	27,27	15	69,74	15	26,85	14	26,37	14	10,32	10	15,64	11
2.	62,16	16	26,82	15	28,13	15	15,65	14	26,58	14	8,18	10	12,52	11
3.	58,32	16	29,51	15	36,57	15	23,75	14	29,28	14	15,38	10	16,09	11
4.	46,94	16	11,70	15	26,87	15	9,53	14	11,13	14	5,75	10	5,85	11
5.	87,47	16	39,96	15	36,05	15	23,63	14	39,80	14	13,07	10	18,31	11
6.	77,35	16	31,74	15	38,30	15	21,76	14	31,23	14	12,41	10	12,46	11
7.	77,69	16	46,73	15	21,05	15	17,34	14	45,32	14	9,10	10	30,16	11
8.	72,65	16	28,45	15	38,68	15	21,97	14	28,11	14	15,91	10	16,32	11
9.	83,96	16	32,65	15	33,46	15	18,06	14	32,35	14	10,71	10	13,07	11
10.	71,60	16	27,77	15	36,80	15	19,69	14	27,48	14	14,93	10	17,23	11
11.	131,94	16	36,69	15	34,72	15	12,95	14	35,84	14	7,19	10	20,88	11
12.	76,27	16	27,61	15	34,83	15	18,48	14	27,19	14	13,52	10	16,90	11
13.	103,62	16	43,69	15	20,61	15	11,38	14	41,80	14	10,63	10	33,40	11
14.	71,96	16	43,94	15	33,91	15	27,77	14	43,14	14	16,82*	10	23,49	11
15.	76,55	16	27,81	15	28,43	15	13,50	14	27,63	14	5,10	10	8,52	11
16.	88,51	16	44,19	15	43,50	15	30,11	14	44,18	14	12,23	10	33,31	11
17.	108,83	16	61,34	15	35,63	15	28,95	14	58,45	14	15,28	10	35,82	11

* $p = 0,08$

do, os alunos que cursavam a 5ª e a 8ª série apresentaram maiores coeficientes de correlação intraclasse, mesmo quando os dados foram estratificados por escola. Este achado é de difícil explicação com os dados disponíveis no estudo.

Já a diferença observada entre as duas escolas, com melhor desempenho da segunda em comparação à primeira, não pode ser explicada pela diferença entre as variâncias dos escores de cada escola (variância da escola I = 107,21; variância da escola II = 88,16). Se a diferença entre as variâncias fosse responsável por este achado, seria esperada, ao contrário do que se observou, uma maior confiabilidade no grupo com maior variância dos escores, no caso a escola I. Assim, é possível que a diferença no CCIC entre as escolas tenha sido causada pela diferença sistemática com que foram selecionados os alunos para participarem do estudo de confiabilidade (viés de seleção). Na segunda escola, só foram “retestados” os primeiros alunos a entregarem o termo de consentimento preenchido, o que poderia indicar um maior interesse e atenção desse grupo, quando comparado com os demais. Além disso, o ab-

senteísmo aparentemente maior na segunda escola pode também ter contribuído com esse viés, com uma maior concentração de alunos interessados entre os presentes na sala de aula. A diferença entre os intervalos de aplicação do teste e reteste (15 dias, para a primeira escola e 13 dias, para a segunda) não nos parece significativa para explicar totalmente esse desempenho diferencial entre as duas escolas. Por mais que seja recomendado que o intervalo entre as duas medidas seja longo o suficiente para reduzir artefatos de memória, mas também próximo o suficiente para diminuir a probabilidade de alterações sistemáticas, a definição exata desse intervalo é arbitrária 27.

A observação das diferenças entre os escores no teste e reteste (gráficos de Bland & Altman) sugere, pelo menos no caso das alunas, que a magnitude da discordância cresce conforme se passa dos escores menores para os maiores; porém, é provável que tal fato não tenha grande influência sobre a validade do questionário, visto que este padrão parece ocorrer apenas nos escores acima de 20-25, faixa esta que já considera o entrevistado como caso suspeito de fobia social.

Ao contrário do resultado observado para a pontuação total do instrumento, **a confiabilidade de cada uma das 17 perguntas que formam o SPIN não foi boa**. Embora em nenhum dos itens tenha sido estimada concordância igual ou inferior àquela esperada ao acaso, alguns itens mostraram confiabilidade baixa. Essa baixa confiabilidade não parece ser explicada por possíveis dificuldades na compreensão dos diferentes itens. Durante a aplicação do questionário, alguns alunos demonstraram dificuldade em compreender certas expressões, como nos itens 6 e 15 (*"ficar constrangido"*), e no item 12 (*"ser o centro das atenções"*). Entretanto, apesar da baixa confiabilidade observada para esses itens (0,45 e 0,48), essa não foi menor do que aquela para itens em que, aparentemente, não teria havido dificuldades de compreensão, como por exemplo, no quarto item (*"evitar falar com pessoas estranhas"*).

O contraste entre boa confiabilidade do escore global e confiabilidade apenas razoável ou boa observada ao analisarmos cada pergunta em particular parece sugerir que a falta de consistência nas respostas seja assistemática. Desse modo, uma menor pontuação no reteste em determinada questão estaria sendo compensada por uma maior pontuação em outro item. Como o questionário solicita que seja respondido com que intensidade os problemas da lista incomodaram na última semana, a menor estabilidade desses itens pode ser decorrente das diferentes oportunidades de exposição a determinadas situações nas semanas anteriores ao teste e reteste. De fato, três das quatro perguntas que apresentaram maiores valores de kappa descrevem o incômodo com sintomas físicos (itens 7, 13 e 17), que podem ser mais constantes do que as diferentes situações às quais o indivíduo possa ser exposto. A quarta pergunta com maior valor de kappa (item 11) descreve a esquivas de situações em que seja necessário falar em público. Vários estudos apontam o medo de falar em público como o sintoma mais prevalente de ansiedade social^{28,29} e essa medida (kappa) pode ser sensível a variações na prevalência, fato que talvez possa justificar esse maior valor.

Quanto ao estudo da estrutura da concordância por intermédio do ajuste de modelos log-lineares, era esperado que o modelo de independência não se ajustasse bem aos dados, como deve ocorrer em estudos de confiabilidade de teste-reteste.

Em 12 dos 17 itens do questionário, o padrão de concordância das respostas dos testes e retestes apresentava dois componentes: a concordância exata e a tendência a que respostas altas (baixas) nos itens do teste correspondessem a respostas altas (baixas) nos mesmos itens do reteste. Este padrão sugere um certo grau de coerência nas respostas dadas pelos alunos, uma vez que, mesmo quando havia discordância entre as respostas do teste e do reteste, os alunos, em geral, assinalavam categorias próximas. No que diz respeito à concordância exata, em onze desses doze itens esta era maior nas categorias em que o sintoma estava ausente ou estava presente de forma branda. Os itens que apresentaram maiores valores de kappa ponderado corresponderam à estrutura de concordância acima descrita.

Conclusão

A versão em português do SPIN exibiu propriedades psicométricas aceitáveis, com resultados de confiabilidade consistentes com aqueles obtidos na avaliação da versão original em inglês. Na análise da estrutura da concordância, o modelo log-linear que melhor se ajustou aos dados na maior parte dos itens foi o modelo de "semi-associação", sugerindo um padrão de concordância exata diferente para cada categoria da variável (maior para as categorias 0 e 1), e um padrão de discordância no qual existe uma tendência em que pontuações baixas (altas) no teste correspondam a pontuações baixas (altas) no reteste.

Esses achados, aliados a um estudo de validade, poderão possibilitar a utilização da versão em português do SPIN em estudos de adolescentes da rede pública de ensino.

Resumo

Acredita-se que a fobia social teria início na adolescência e que precederia outros transtornos mentais, sendo importante a sua investigação nos jovens. Até o momento, não há nenhuma escala de fobia social validada para a nossa população. Neste estudo investigou-se a confiabilidade da versão em Português do Social Phobia Inventory (SPIN) entre escolares adolescentes da rede pública do Município do Rio de Janeiro. Após a versão do SPIN para o Português, conduziu-se um estudo de confiabilidade teste-reteste com 190 adolescentes de duas escolas, estimando-se os coeficientes de correlação intraclasse (CCIC), kappa ponderado (κ^w) e ajuste de modelos log-lineares. Foram ainda construídos gráficos Bland & Altman. Observou-se uma boa consistência interna (α de Cronbach = 0,88) e boa confiabilidade da pontuação total do instrumento (CCIC = 0,78). A confiabilidade das perguntas isoladas não foi tão boa (κ^w variando de 0,32 a 0,65). O modelo log-linear de melhor ajuste aos dados na maior parte dos itens foi o de "semi-associação". Esses achados nos permitiram concluir que a versão em português do SPIN exibiu resultados de boa confiabilidade, semelhantes aos da versão original em Inglês.

Fobia Social; Psicometria; Confiabilidade e Validade

Colaboradores

L. M. P. Vilete colaborou na coleta de dados, análise de dados e preparo do manuscrito. E. S. F. Coutinho e I. L. V. Figueira contribuíram na análise de dados e preparo do manuscrito

Referências

1. American Psychological Association. Manual diagnóstico e estatístico de transtornos mentais – DSM – IV – TR. 4ª Ed. v. 1. Porto Alegre: Artes Médicas; 1995.
2. Chaleby KS, Raslan A. Delineation of social phobia in Saudian Arabians. Soc Psychiatry Psychiatr Epidemiol 1990; 25:324-7.
3. Pine DS, Cohen P, Gurley D, Brook J, Ma Y. The risk for early-adulthood anxiety and depressive disorders in adolescents with anxiety and depressive disorders. Arch Gen Psychiatry 1998; 5:56-64.
5. Schatzberg AF, Samson JA, Rothschild AJ, Bond TC, Regier DA. McLean Hospital depression research facility: early-onset phobic disorders and adult-onset major depression. Br J Psychiatry 1998; 173 Suppl 34:29-34.
6. DeWit DJ, Macdonald K, Offord DR. Childhood stress and symptoms of drug dependence in adolescence and early adulthood: social phobia as a mediator. Am J Orthopsychiatric 1999; 69:61-72.
7. Nelson EC, Grant JD, Bucholz KK, Glowinski A, Madden PAF, Reich W, et al. Social phobia in a population-based female adolescent twin sample: comorbidity and associated suicide-related symptoms. Psychol Med 2000; 30:797-804.
8. Albano AM, Detweiler MF. The developmental and clinical impact of social anxiety and social phobia in children and adolescents. In: Hofmann SG, Dibartolo PM, editors. From social anxiety to social phobia – multiple perspectives. Boston: Allyn & Bacon; 2001. p. 162-78.
9. Beidel DC, Turner SM, Morris TL. Psychopathology of childhood social phobia. J Am Acad Child Adolesc Psychiatry 1999; 38:643-50.
10. Barlow DH, Liebowitz MR. Specific phobia and social phobia. In: Sadock HIKBJ, editor. Comprehensive textbook of psychiatry. 6th Ed. v. 1. New York: Williams & Wilkins; 1995. p. 1204-18.
11. Beidel DC. Social anxiety disorder: etiology and early clinical presentation. J Clin Psychiatry 1998;

- 59 Suppl 17:27-31.
12. Essau CA, Conradt J, Petermann F. Frequency and comorbidity of social phobia and social fears in adolescents. *Behav Res Ther* 1999; 37:831-43.
13. Szklo M, Nieto FJ. Quality assurance and control. In: Szklo M, Nieto FJ, editors. *Epidemiology – beyond the basics*. Maryland: Aspen Publishers; 2000. p. 343-404.
14. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 1:307-10.
15. Connor KM, Davidson JRT, Churchill LE, Sherwood A, Foa E, Weisler RH. Psychometric properties of the social phobia inventory (SPIN). *Br J Psychiatry* 2000; 176:379-86.
16. Reichenheim ME, Moraes CL, Hasselmann MH. Equivalência semântica da versão em português do instrumento abuse assessment screen para rastrear violência contra a mulher grávida. *Rev Saúde Pública* 2000; 34:610-6.
17. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979; 86:420-8.
18. Fleiss JL. *Statistical methods for proportional*. New York: John Wiley & Sons; 1981.
19. Maclure M, Willett WC. Misinterpretation and misuse of the Kappa statistic. *Am J Epidemiol* 1987; 126:161-9.
20. Agresti A. A model for agreement between ratings on an ordinal scale. *Biometrics* 1988; 44:539-48.
21. Graham P, Jackson R. The analysis of ordinal agreement data: beyond weighted kappa. *J Clin Epidemiol* 1993; 46:1055-62.
22. May SM. Modelling observer agreement – an alternative to Kappa. *J Clin Epidemiol* 1994; 47:1315-24.
23. Silva EFD, Pereira MG. Avaliação das estruturas de concordância e discordância nos estudos de confiabilidade. *Rev Saúde Pública* 1998; 32:383-93.
24. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33:159-74.
25. Morgenstern H. *Epidemiologic methods*. Los Angeles: UCLA School of Public Health; 1989.
26. Bland JM, Altman DG. Statistics notes: Cronbach's alpha. *Br Med J* 1997; 314:572.
27. Shrout PE. Reliability. In: Zahner TTA, editor. *Textbook in psychiatry epidemiology*. New York: Wiley-Liss; 1995. p. 213-27.
28. Lipsitz JD, Schneier FR. Social phobia – epidemiology and cost of illness. *Pharmacoeconomics* 2000; 18:23-32.
29. Stein MB, Torgrud LJ, Walker JR. Social phobia symptoms, subtypes, and severity. *Arch Gen Psychiatry* 2000; 57:1046-52.

Recebido em 13/Mar/2003

Versão final reapresentada em 3/Jun/2003

Aprovado em 27/Ago/2003