

Carolina Alves Ferreira de Carvalho¹
 Patrícia Silva Lúcio²
 Clara Regina Brandão de Ávila³

Descritores

Testes de Linguagem
 Psicometria
 Dificuldade
 Discriminação
 Fonoaudiologia

Keywords

Language Tests
 Psychometrics
 Difficulty
 Discrimination
 Speech, Language and Hearing Sciences

Endereço para correspondência:

Carolina Alves Ferreira de Carvalho
 Rua Botucatu, 802, São Paulo (SP),
 Brasil, CEP: 04023-900.
 E-mail: carolcarvalho_fono@yahoo.com.br

Recebido em: 03/10/2013

Aceito em: 02/02/2015

CoDAS 2015;27(4):344-9

Equivalência psicométrica da versão brasileira do *Test of Pragmatic Language 2* – TOPL-2

Psychometric equivalence of the Brazilian version of the Test of Pragmatic Language 2 – TOPL-2

RESUMO

Introdução: Tendo em vista a relevância da investigação sobre a função pragmática para o diagnóstico dos transtornos de linguagem e a necessidade da utilização de testes com evidências de validade, o objetivo deste trabalho foi examinar a equivalência psicométrica entre as versões americana e brasileira do *Test of Pragmatic Language – Second Edition* (TOPL-2). **Método:** Um total de 81 escolares do terceiro ao sétimo ano do ensino fundamental público do município de São Paulo (63% meninas; média de idade=9,42, DP=0,93) foi escolhido por seus professores em função da ausência de queixas ou indicadores de baixo rendimento escolar. Responderam à aplicação oral de questões da versão traduzida e adaptada para o português brasileiro do TOPL-2. Replicaram-se alguns dos estudos conduzidos para a padronização americana (versão original) com a análise de dificuldade, consistência interna, discriminação e funcionamento diferencial. **Resultados:** A maioria dos itens (86%) apresentou dificuldade semelhante à da amostra americana. Os valores de consistência interna mostraram-se dentro dos limites aceitáveis (0,70). Poucos itens (26%) tiveram discriminação adequada, porém 49% mostraram-se próximos do desejável. Oito itens apresentaram funcionamento diferencial, cinco favoráveis aos meninos. **Conclusão:** Apesar das limitações de tamanho e variabilidade da amostra, quase metade dos itens demonstrou equivalência psicométrica com a versão americana.

ABSTRACT

Introduction: Considering the relevance of the investigation about the pragmatic function for language disorders diagnosis, as well as the relevance of validity evidence for test interpretation, this paper aimed to explore the psychometric equivalence between the American and Brazilian versions of the Test of Pragmatic Language, Second Edition (TOPL-2). **Methods:** A total of 81 students from the third to seventh year from public elementary schools of São Paulo (63% girls; average age=9.42 years, SD=0.93) with no complaints or indicators of low school performance were selected. Students answered the verbal application of the TOPL-2 Brazilian-Portuguese version. Some of the psychometric inquiries reported in the original version of the test were reproduced focusing on the analyses of difficulty, internal consistency, discrimination, and differential item functioning. **Results:** Most of the items (86%) presented similar difficulties to the American version. The internal consistency index had acceptable values (0.70). Few items (26%) showed adequate discrimination, but 49% were close to the desirable cutoffs. Eight items showed differential functioning; of which, five were favorable to boys. **Conclusion:** Despite sample limitations related to size and variability, almost half of the items showed psychometric equivalence to the American version.

Trabalho realizado na Escola Paulista de Medicina, Universidade Federal de São Paulo – EPM-UNIFESP – São Paulo (SP), Brasil.

(1) Departamento de Fonoaudiologia da Escola Paulista de Medicina, Universidade Federal de São Paulo – EPM-UNIFESP – São Paulo (SP), Brasil.

(2) Departamento de Psicologia e Psicanálise, Universidade Estadual de Londrina – UEL – Londrina (PR), Brasil.

(3) Departamento de Fonoaudiologia da Escola Paulista de Medicina, Universidade Federal de São Paulo – EPM-UNIFESP – São Paulo (SP), Brasil.

Conflito de interesses: nada a declarar.

INTRODUÇÃO

Na área dos distúrbios da comunicação humana, principalmente da Fonoaudiologia, nota-se uma recorrente preocupação dos pesquisadores na busca de identificar as relações entre os processamentos dos diversos subsistemas da linguagem. A análise do ato de linguagem não se dissocia das características de uso da própria linguagem, o que implica reconhecer que a função pragmática da linguagem subjaz aos atos reguladores da comunicação e, em alguma instância, os determina⁽¹⁾.

De forma sistemática, estudos na área da linguística postulam que as bases pragmáticas são fundamentais para a efetividade da comunicação humana, tanto mediada pela linguagem oral quanto pela escrita, em seus modos expressivo e compreensivo⁽²⁻⁵⁾. Na avaliação da comunicação humana, sobretudo a fonoaudiológica, deve-se considerar, portanto, que, além dos mecanismos básicos para a expressão e compreensão oral ou escrita, outros relacionados à função pragmática da linguagem são fundamentais. Sob essa óptica, entende-se que a avaliação da função pragmática deve informar sobre as capacidades de realizar inferências e as habilidades de automonitoramento e crítica, as quais são indispensáveis para que atividades discursivas e ligadas à compreensão textual sejam proficientes.

A significativa escassez de instrumentos nacionais que avaliam diferentes funções ou processamentos de linguagem, aliada ao reconhecimento da importância da generalização dos resultados em diferentes amostras, explica o interesse dos pesquisadores brasileiros na condução de estudos que visem adaptar, à nossa realidade, instrumentos linguísticos já consagrados internacionalmente⁽⁶⁻⁷⁾. Entretanto, para que um teste construído em uma determinada cultura possa ser utilizado em uma realidade distinta, com a garantia de manutenção de sua qualidade, além da preocupação com os procedimentos tradicionais de tradução e adaptação, que comprovam a equivalência semântica e cultural dos itens, o pesquisador também deve se ater à equivalência dos aspectos psicométricos entre a versão original e a nova⁽⁸⁾.

Admitir a importância da linguagem pragmática na regulação dos mecanismos de automonitoramento da compreensão e elaboração do discurso determinou a construção de um projeto de pesquisa, que tem investigado a relação entre a função pragmática da linguagem e o desempenho na compreensão da leitura e na elaboração escrita de textos^(1,9-10). Para isso, analisaram-se alguns instrumentos de avaliação da função pragmática. O *Test of Pragmatic Language* (TOPL)⁽¹¹⁾ é um teste americano, atualmente em sua segunda edição (TOPL-2)⁽¹²⁾. Ambas as versões do teste têm como objetivo avaliar a função pragmática ou social da linguagem. O TOPL-2 foi traduzido e adaptado para o Português Brasileiro (PB)⁽¹³⁾. Seus 43 itens, de aplicação oral, provêm importantes informações referentes às capacidades para resolução de conflitos e habilidades sociais. O teste, em suas duas versões americanas, tem se mostrado uma ferramenta útil para a avaliação da função pragmática da linguagem no domínio da comunicação oral. Estudos demonstraram sua efetividade na identificação de problemas de populações específicas, tais como no autismo, na síndrome de Williams e no transtorno específico da linguagem⁽¹⁴⁻¹⁶⁾.

O processo de tradução e adaptação do TOPL-2 para o PB seguiu as recomendações propostas por Beaton et al.⁽¹⁷⁾. A tradução foi realizada por uma tradutora juramentada e aplicada em um grupo de escolares brasileiros do ensino fundamental⁽¹⁸⁾. As respostas coletadas nesse procedimento foram analisadas e aquelas que apresentaram maior número de erros foram reanalisadas por três fonoaudiólogas brasileiras, também fluentes em língua inglesa, as quais revisaram as proposições do teste. Para a revisão e equivalência gramatical e idiomática, o teste foi enviado a um novo tradutor de inglês que desconhecia o original, para que fosse feito o processo inverso, ou seja, sua versão ao inglês, a qual, comparada ao original, mostrou-se fiel.

Tendo em vista a relevância da investigação da função pragmática para o diagnóstico dos transtornos de linguagem, oral e escrita, e a necessidade da utilização dos testes de avaliação de linguagem com validade psicométrica, o objetivo do presente trabalho foi coletar evidências para a equivalência psicométrica entre as versões americana e brasileira do TOPL-2.

Assim, após os procedimentos de tradução e adaptação linguística, o TOPL-2 foi novamente aplicado em outra amostra de escolares típicos, a fim de replicar as principais análises estatísticas reportadas no manual do instrumento, que se referem às propriedades dos itens (fidedignidade, dificuldade, discriminação e funcionamento diferencial). A análise psicométrica das propriedades dos itens permite identificar os que devem ser revisados para aplicação em um futuro estudo normativo. A hipótese deste estudo-piloto foi a de que a maioria dos itens do teste apresentasse semelhança psicométrica com a versão original (americana) do instrumento.

MÉTODO

Considerações éticas

Este estudo recebeu aprovação do Comitê de Ética em Pesquisa da Escola Paulista de Medicina, Universidade Federal de São Paulo (EPM-UNIFESP, protocolo nº 1.731/08).

Amostra

A amostra foi composta por 81 escolares (63% meninas), entre 8 e 11 anos (média=9,42, DP=0,93), matriculados entre o terceiro e sétimo ano do ensino fundamental de duas escolas da rede pública de ensino do município de São Paulo. Os participantes foram, a princípio, indicados por seus professores, os quais haviam sido instruídos a selecionar escolares sem queixas ou indicadores de déficits de leitura ou baixo rendimento escolar. Os estudantes indicados passaram por triagem e, somente aqueles que alcançaram valores de taxa e acurácia de leitura esperados para a faixa de escolaridade, de acordo com os critérios apresentados por Carvalho⁽⁹⁾, constituíram a amostra. Os responsáveis pelos participantes assinaram o termo de consentimento livre e esclarecido.

Instrumentos

Os participantes responderam à versão traduzida e adaptada para o português brasileiro⁽¹³⁾ do TOPL-2⁽¹²⁾, que é um

teste de aplicação oral composto por questões realizadas pelo examinador, as quais são, em sua maioria, baseadas em figuras do próprio teste. Elas buscam fornecer informações sobre seis subcomponentes da função pragmática da linguagem, a saber:

- Contexto físico – fornece pistas ambientais que sinalizam padrões adequados de comunicação (por exemplo, falar em voz baixa em uma biblioteca);
- Audiência – tem a capacidade de monitorar uma variedade de fatores relacionados ao interlocutor, adaptando o modo de se comunicar a eles (por exemplo, a idade ou o número de interlocutores);
- Tópico – tem a capacidade de gerenciar conteúdo apropriado ao tópico, garantir coerência lógica ao fluxo da conversa e monitorar a evolução do assunto de modo a resolver problemas que dificultem o entendimento;
- Objetivo – congrega características relacionadas ao propósito de uma conversa, bem como as modificações e manipulações linguísticas utilizadas para atingi-lo (por exemplo, fazer perguntas);
- Pistas visuais – atenção a aspectos não verbais da comunicação (por exemplo, linguagem corporal);
- Abstração – percepção das informações transmitidas pela linguagem abstrata, geralmente utilizada para comunicar emoções, imagens e outras mensagens que não podem ou não devem ser transmitidas diretamente (por exemplo, o provérbio “o seguro morreu de velho” passa a mensagem de prudência).

As funções descritas podem ser avaliadas em um ou diferentes itens. O 2 avalia os subcomponentes “Tópico” e “Objetivo”, enquanto que o item 3 analisa apenas “Abstrações”.

Além da função pragmática da linguagem, os participantes foram avaliados quanto à decodificação leitora, closura gramatical, memória e compreensão oral e leitora, cujos resultados não serão reportados no presente artigo.

Procedimentos

Os escolares foram avaliados por uma única examinadora em salas reservadas pela própria escola. As aplicações do TOPL-2 foram individuais e despenderam, em média, 40 minutos.

Análise dos dados

A análise replicou, em amostra-piloto, alguns dos estudos conduzidos para a padronização americana do TOPL-2. Pretendeu-se investigar a correspondência psicométrica das adaptações cultural e linguística do instrumento, com a versão original. Desse modo, foram utilizadas análises clássicas, cujos critérios de ajuste foram definidos pelos autores do teste: consistência interna (alfa de Cronbach) – valores acima de 0,70 foram considerados adequados; dificuldade (proporção de acerto nos itens) – desejáveis valores medianos (entre 0,15 e 0,85); discriminação (correlação ponto-biserial) – a partir de 0,35; funcionamento diferencial do item (DIF) – regressão linear logística, tendo como preditor da precisão no item o escore total (Modelo 1) ou o escore total mais o sexo (Modelo 2). Quando houve variação entre os modelos, atendeu-se para o funcionamento diferencial do item, dado pela diferença entre as variâncias explicadas pelos modelos (R^2), seguindo o critério⁽¹²⁾: (d.1) DIF desprezível – diferença entre os R^2 menor do que 0,035; (d.2) DIF moderado – diferença maior ou igual a 0,035 e menor do que 0,070; (d.3) DIF elevado – diferença maior ou igual a 0,070.

Além das estatísticas relatadas no estudo original, também calculou-se o índice D de discriminação, que reflete a diferença da proporção de acerto no item dos indivíduos com os escores 27% maiores e 27% menores. Valores de D abaixo de 0,28 foram considerados inadequados⁽¹⁹⁾.

Os dados foram analisados por meio do programa *SPSS for Windows*, versão 18.0. Todas as comparações que apresentaram $p < 0,05$ foram significativas.

RESULTADOS

A Tabela 1 apresenta as análises descritivas do estudo. A média geral de acertos foi de pouco mais de 50% (Média=22,54; DP=5,16). Apesar de os valores de assimetria e curtose se aproximarem de zero (respectivamente, 0,39 e -0,21), não foi observada distribuição normal dos itens (*Kolmogorov-Smirnov*=0,131, $df=81$, $p < 0,001$). De maneira semelhante ao estudo original, houve correlação entre a idade e os escores dos participantes. Ao contrário do estudo original, que utilizou a correlação de

Tabela 1. Estatísticas descritivas por faixa etária, sexo e total e valores de alfa de Cronbach (α) e da correlação para a amostra do presente estudo e para o estudo original

Idade (anos)	n	Mínimo	Máximo	Média	DP	α	α – estudo original
8	13	15	36	22,62	6,44	0,80	0,86
9	33	11	32	20,91	4,93	0,69	0,91
10	23	16	33	24,22	4,55	0,61	0,92
11	12	19	31	23,75	4,58	0,63	0,93
Total	81	11	36	22,54	5,16	0,71	–
Correlação com a idade em meses (Spearman)						0,263*	0,64
Sexo							
Feminino	63	14	36	22,47	5,52	0,75	–
Masculino	35	14	32	22,69	4,41	0,58	–
Correlação com o sexo (Spearman)						0,04 (NS)	–

* $p < 0,05$

Legenda: NS = não significante; DP = desvio padrão

Pearson, neste, relata-se a correlação de Spearman devido à característica de não normalidade dos dados. Entretanto, no presente estudo, esta correlação mostrou-se baixa e não moderada como na amostra americana.

Em relação à análise da consistência interna, o conjunto de itens do TOPL-2 obteve valor de 0,71, o que está dentro do limite desejável⁽²⁰⁾. Nas faixas etárias das crianças mais jovens (8 e 9 anos), observou-se tendência a valores de alfa mais robustos (ao contrário do estudo original, em que eles permaneceram elevados e mais ou menos constantes). A Figura 1 mostra a distribuição de frequência dos itens do TOPL-2 na amostra brasileira.

A Tabela 2 exhibe os valores dos índices psicométricos do estudo. Quanto à dificuldade (coluna p), apenas 6 itens (14%) tiveram índices muito baixos (1 item) ou muito elevados (5 itens) de acerto. Para as estatísticas do alfa de Cronbach, foi possível observar que os valores praticamente não se alteraram com as análises conduzidas com os itens removidos e sempre oscilaram entre 0,68 e 0,72 (coluna α na Tabela 2). Apenas 11 itens (26,0%) apresentaram correlação ponto-biserial dentro do ponto de corte estipulado pelo estudo americano ($>0,35$). Entretanto, outros 16 (37,2%) receberam correlações significativas, cujos valores oscilaram entre 0,23 e 0,33 (coluna r_{pb}). Quanto à discriminação (coluna D), 21 itens (48,8%) apresentaram valores adequados ($\geq 0,28$). Apenas três desses não mostraram uma correlação ponto-biserial.

Consideraram-se psicometricamente adaptados os itens que, simultaneamente, apresentaram correlação ponto-biserial, valores medianos de dificuldade e índices D maiores do que 0,28. Dos 43 itens do TOPL-2, apenas 19 (44,2%) mostraram estatísticas adequadas para todas as medidas utilizadas (marcados como A na coluna de “Equivalência” da Tabela 2).

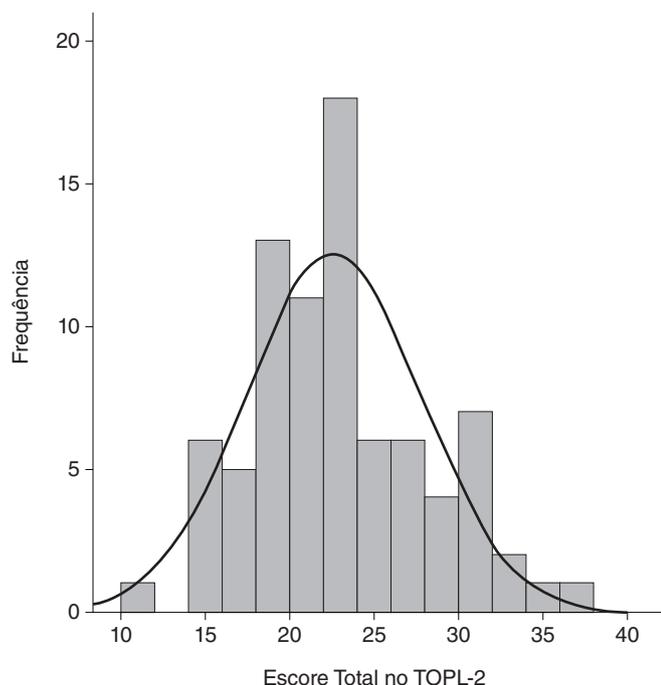


Figura 1. Distribuição dos escores do TOPL-2 para a amostra do estudo

Destes 19 itens, 7 (36,8%) tiveram a dificuldade mais elevada ($0,15 \leq p \leq 0,40$). Os itens com dificuldade mediana e correlação ponto-biserial significativa ou índice D acima de 0,28 (classificados como B na Tabela 2) foram definidos como parcialmente equivalentes, totalizando 10. Itens com precisão muito elevada ou demasiadamente baixa, índice D muito baixo e/ou correlação ponto-biserial não significativa (itens marcados como C

Tabela 2. Índices de dificuldade (p), valores de Alfa de Cronbach (α), correlação ponto-biserial (r_{pb}), índices de discriminação (D) e equivalência (Eq.) para os itens TOPL-2

Item	p	α^a	r_{pb}	D	Eq.
1	0,74	0,72	0,00	0,05	C
2	0,90	0,71	-0,09	-0,08	C
3	0,88	0,70	0,26*	0,16	C
4	0,69	0,70	0,24*	0,13	B
5	0,93	0,70	0,25*	0,16	C
6	0,73	0,70	0,20	0,28	B
7	0,35	0,71	0,20	0,34	B
8	0,95	0,71	0,01	-0,04	C
9	0,80	0,70	0,19	0,08	C
10	0,89	0,70	0,31**	0,20	C
11	0,84	0,70	0,19	0,16	C
12	0,36	0,70	0,27*	0,26	B
13	0,27	0,69	0,40**	0,34	A
14	0,65	0,70	0,33**	0,29	A
15	0,35	0,69	0,41**	0,42	A
16	0,52	0,69	0,37**	0,45	A
17	0,23	0,69	0,52**	0,50	A
18	0,44	0,70	0,21	0,22	C
19	0,73	0,70	0,29**	0,26	B
20	0,69	0,70	0,34**	0,40	A
21	0,23	0,69	0,38**	0,38	A
22	0,44	0,70	0,32**	0,29	A
23	0,64	0,69	0,39**	0,49	A
24	0,59	0,70	0,32**	0,45	A
25	0,60	0,70	0,25*	0,21	B
26	0,27	0,70	0,18	0,19	C
27	0,64	0,70	0,29**	0,37	A
28	0,70	0,70	0,30**	0,33	A
29	0,06	0,70	0,28*	0,12	C
30	0,37	0,70	0,24*	0,26	B
31	0,41	0,69	0,37**	0,41	A
32	0,56	0,70	0,35**	0,49	A
33	0,53	0,70	0,33**	0,33	A
34	0,72	0,70	0,15	0,17	C
35	0,49	0,71	0,11	0,14	C
36	0,26	0,71	0,08	0,07	C
37	0,57	0,70	0,21	0,29	B
38	0,17	0,70	0,26*	0,19	B
39	0,57	0,69	0,38**	0,37	A
40	0,20	0,70	0,32**	0,38	A
41	0,21	0,70	0,23*	0,22	B
42	0,21	0,68	0,61**	0,58	A
43	0,15	0,69	0,39**	0,31	A

^avalores para o item excluído da análise; * $p < 0,05$; ** $p < 0,01$

Legenda: A = itens adequados psicometricamente ao estudo original; B = itens com valores muito próximos aos desejáveis; C = itens a serem revisados

na coluna de “Equivalência” da Tabela 2) são não equivalentes, constituindo um total de 14. Eles necessitarão de revisões futuras mais significativas. Com esse resultado, pode-se dizer que mais da metade dos itens (67,44%) apresentou qualidade psicométrica na direção esperada, seja total ou parcial.

No que diz respeito ao DIF, conforme assinalado no método, foram comparados dois modelos distintos de regressão

Tabela 3. Estatísticas descritivas dos itens por sexo, diferença entre as médias dos grupos de meninas e meninos, diferença entre os valores de R² dos modelos (R² de Nagelkerke) e classificação do DIF dos itens

Item	Feminino		Masculino		Diferença de média	R ² F - R ² M	DIF
	Média	DP	Média	DP			
1	0,76	0,43	0,69	0,47	0,07	0,01	Desprezível
2	0,89	0,32	0,92	0,27	-0,03	0,01	Desprezível
3	0,87	0,34	0,88	0,33	-0,01	0,00	Desprezível
4	0,71	0,46	0,65	0,49	0,06	0,01	Desprezível
5	0,95	0,23	0,88	0,33	0,07	0,05	Moderado
6	0,73	0,45	0,73	0,45	0,00	0,00	Desprezível
7	0,31	0,47	0,42	0,50	-0,11	0,02	Desprezível
8	0,95	0,23	0,96	0,20	-0,01	0,00	Desprezível
9	0,85	0,36	0,69	0,47	0,16	0,06	Moderado
10	0,89	0,32	0,88	0,33	0,01	0,00	Desprezível
11	0,82	0,39	0,88	0,33	-0,06	0,01	Desprezível
12	0,38	0,49	0,31	0,47	0,07	0,01	Desprezível
13	0,35	0,48	0,12	0,33	0,23	0,09	Elevado
14	0,65	0,48	0,65	0,49	0,00	0,00	Desprezível
15	0,33	0,47	0,38	0,50	-0,05	0,01	Desprezível
16	0,51	0,51	0,54	0,51	-0,03	0,00	Desprezível
17	0,22	0,42	0,27	0,45	-0,05	0,01	Desprezível
18	0,38	0,49	0,58	0,50	-0,20	0,04	Moderado
19	0,71	0,46	0,77	0,43	-0,06	0,00	Desprezível
20	0,69	0,47	0,69	0,47	0,00	0,00	Desprezível
21	0,22	0,42	0,27	0,45	-0,05	0,01	Desprezível
22	0,47	0,50	0,38	0,50	0,09	0,01	Desprezível
23	0,55	0,50	0,85	0,37	-0,30	0,10	Elevado
24	0,58	0,50	0,62	0,50	-0,04	0,00	Desprezível
25	0,58	0,50	0,65	0,49	-0,07	0,00	Desprezível
26	0,29	0,46	0,23	0,43	0,06	0,01	Desprezível
27	0,67	0,47	0,58	0,50	0,09	0,02	Desprezível
28	0,75	0,44	0,62	0,50	0,13	0,03	Desprezível
29	0,04	0,19	0,12	0,33	-0,08	0,08	Elevado
30	0,33	0,47	0,46	0,51	-0,13	0,02	Desprezível
31	0,44	0,50	0,35	0,49	0,09	0,01	Desprezível
32	0,58	0,50	0,50	0,51	0,08	0,01	Desprezível
33	0,51	0,51	0,58	0,50	-0,07	0,00	Desprezível
34	0,67	0,47	0,81	0,40	-0,14	0,03	Desprezível
35	0,51	0,51	0,46	0,51	0,05	0,00	Desprezível
36	0,29	0,46	0,19	0,40	0,10	0,02	Desprezível
37	0,55	0,50	0,62	0,50	-0,07	0,01	Desprezível
38	0,16	0,37	0,19	0,40	-0,03	0,00	Desprezível
39	0,65	0,48	0,38	0,50	0,27	0,09	Elevado
40	0,16	0,37	0,27	0,45	-0,11	0,03	Desprezível
41	0,18	0,39	0,27	0,45	-0,09	0,02	Desprezível
42	0,25	0,44	0,12	0,33	0,13	0,03	Desprezível
43	0,09	0,29	0,27	0,45	-0,18	0,11	Elevado

logística. Utilizou-se como desfecho o escore no item e como preditores o escore total do participante (Modelo 1) e o total mais o sexo (Modelo 2). A Tabela 1 apresenta a média geral de acertos dos participantes por sexo e a correlação entre o sexo e o escore total. Já a Tabela 3 mostra a média e o desvio padrão de acerto por item em função do sexo dos participantes. A maioria dos itens (81,4%) teve DIF desprezível (média=0,01). Três itens (5, 9 e 18) apresentaram DIF moderado (média=0,05) e cinco deles (13, 23, 29, 39 e 43) estavam com DIF elevado (média=0,09). A maior parte dos itens com DIF elevado (60%) foi favorável aos meninos, enquanto que os itens com DIF moderado tenderam para as meninas (66,7%). Daqueles com DIF, quatro haviam sido considerados psicometricamente adequados (13, 23, 39 e 43). Na versão original, não houve DIF por sexo, apenas um item com DIF moderado, favorável aos afro-americanos em relação aos americanos de origem europeia. A Tabela 3 sintetiza tais resultados.

DISCUSSÃO

O presente estudo replicou, em amostra brasileira, as principais investigações psicométricas conduzidas na versão americana do TOPL-2. De um modo geral, pode-se afirmar que os resultados são promissores no que diz respeito à equivalência psicométrica das duas versões dos instrumentos, mas aponta a necessidade de ajustar alguns itens de modo a garantir a qualidade geral da versão adaptada.

Em geral, os valores de alfa de Cronbach encontrados no presente estudo mostraram-se satisfatórios, apesar de um pouco abaixo daqueles apresentados na versão original do instrumento. Um fato que pode explicar este resultado é o tamanho da amostra deste estudo, que é inferior à versão americana, a qual foi constituída por 1.136 pessoas. Sabe-se que o coeficiente alfa é sensível ao tamanho da amostra de sujeitos, assim como de itens⁽²¹⁾. Esta hipótese é corroborada quando se verifica a diferença entre os valores de alfa para a amostra de meninas (n=63) e meninos (n=35), respectivamente, de 0,75 e 0,58 (Tabela 1). Acredita-se que com o aumento do tamanho da amostra, no estudo de padronização, resultados mais semelhantes ao da versão em inglês sejam encontrados.

Com relação ao índice de dificuldade, pode-se afirmar que, para este parâmetro, os itens apresentaram maior semelhança com os da amostra americana. Assim, 86% dos itens obtiveram índices de dificuldade dentro da faixa estipulada pelos autores (entre 0,15 e 0,85), conforme os dados do manual. Daqueles que alcançaram índices fora dos limites esperados, a maioria foi classificada como muito fácil (apenas um item necessita de revisão, por ser muito difícil). Acredita-se que, de maneira semelhante ao alfa de Cronbach, com amostra mais ampla e heterogênea, seja possível que esses itens possam exibir níveis de dificuldade dentro do esperado.

Do ponto de vista da análise da discriminação por meio da correlação ponto-bisserial, apenas 26% dos itens alcançou o ponto de corte de 0,35 indicado pelos autores do teste. No entanto, a maioria deles mostrou coeficientes de correlação ponto-bisserial significativos (30 itens ou 70% do total), apesar

de não terem cumprido o critério estipulado. Desse modo, é possível dizer que a maioria dos itens está na direção esperada, em termos de discriminação. Isso é corroborado pelos índices D no teste: apenas 11 itens com correlação ponto-bisserial no nível de $p < 0,05$ conseguiram $D < 0,28$.

Com os resultados obtidos, seria esperado que 14 necessitariam de revisões mais significativas, por não serem considerados equivalentes culturalmente. No entanto, a análise de funcionamento diferencial dos itens reportou que, dos oito itens com DIF, quatro haviam sido classificados como psicometricamente adequados. Esses resultados mostram que uma análise mais detalhada do conteúdo semântico desses itens deve ser realizada a fim de observar as possíveis causas desse efeito. Assim, 18 itens necessitarão de revisões mais sistemáticas antes de serem aplicados na amostra mais ampla para a padronização. Os demais (classificados como A ou B) podem ser considerados psicometricamente ajustados na direção esperada. Nesse sentido, pode-se afirmar que 58% dos itens do presente estudo demonstraram equiparação psicométrica com os itens da versão americana.

É importante ressaltar que as etapas preconizadas para a tradução e adaptação dos testes estrangeiros, normalmente utilizadas nas pesquisas nacionais, foram seguidas no presente estudo. Entretanto, os resultados mostram que os procedimentos seguidos para a adaptação do TOPL-2 não foram suficientes para a adequação de sua aplicação na população brasileira. Questões linguísticas e culturais precisam ser revisadas, assim como o aumento da amostra em estudos futuros.

Os próximos passos da pesquisa preveem a revisão de aspectos semânticos dos itens que se mostraram psicometricamente inadequados quanto aos parâmetros dificuldade, discriminação e funcionamento diferencial. É importante levar em conta que o manual do teste reporta apenas estatísticas gerais e por faixa etária. Assim, a ausência de dados sobre a qualidade psicométrica dos itens, analisados individualmente, pode dificultar as adequações necessárias. Além disso, após a adequação psicométrica dos itens, devem-se conduzir estudos que verifiquem a equivalência do construto por eles avaliado, a serem realizados por meio da investigação de populações especiais, tais como a de indivíduos com diferentes distúrbios da comunicação, oral ou escrita, que envolvam alterações da função pragmática da linguagem. A expectativa é que em breve seja possível realizar o estudo para a normatização da versão brasileira do instrumento, com equivalências cultural, linguística e psicométrica comprovadas.

CONCLUSÃO

Apesar das limitações do presente estudo em relação ao tamanho e à variabilidade da amostra, pode-se dizer que os resultados apresentados são promissores no que se refere à adequação psicométrica da versão brasileira do TOPL-2, uma vez que quase a metade dos itens do presente estudo mostrou equivalência com aqueles da versão americana.

**CAFC foi responsável pela elaboração da pesquisa, coleta e análise dos dados e redação do artigo; PSL realizou a análise estatística e a redação do artigo; CRBA contribuiu com a elaboração da pesquisa, análise dos dados e redação do artigo.*

REFERÊNCIAS

1. Ávila CR, Carvalho CA, Paolucci-Bigarelli JF, Kida AS. Relação entre a função pragmática da linguagem, competências leitoras e de elaboração e produção de textos. In: Melo LE, organizadora. *Cognição e linguagem: perspectivas interdisciplinares*. Curitiba: CRV, 2011. p. 255-74.
2. Kintsch W, Van Dijk TA. Toward a model of text comprehension and production. *Psychol Rev*. 1978;85(5):363-94.
3. Armengaud F. *A pragmática*. São Paulo: Parábola; 2009.
4. Fávero LL, Koch IG. *Linguística textual: introdução*. 8ª ed. São Paulo: Cortez; 2007.
5. Koch IG, Elias VM. *Ler e compreender: os sentidos do texto*. 2ª ed. São Paulo: Cortez; 2007.
6. American Speech-Language-Hearing Association (ASHA) (2005) [Internet]. Evidence-Based Practice in Communication Disorders [Position Statement]. [cited 2013 Ago 10]. Available from: <http://www.asha.org/docs/html/PS2005-00221.html>
7. Giusti E, Befi-Lopes DM. Translation and cross-cultural adaptation of instruments to the Brazilian Portuguese language (PB). *Pró-Fono Rev Atualização Cient*. 2008;20(3):207-10.
8. Borsa JC, Damásio BF, Bandeira DR. Adaptação e validação de instrumentos psicológicos entre culturas: algumas considerações. *Paidéia (Ribeirão Preto)*. 2012;22(53):423-32.
9. Carvalho CA. Relação entre a função pragmática da linguagem e compreensão de texto [dissertação]. São Paulo: Universidade Federal de São Paulo; 2008.
10. Paolucci JF. Estudo das correlações entre a função pragmática da linguagem e a produção textual escrita [dissertação]. São Paulo: Universidade Federal de São Paulo; 2009.
11. Phelps-Teraski D, Phelps-Gunn T. *Test of Pragmatic Language: Examiner's manual*. Austin, TX: PRO-ED, Inc; 1992.
12. Phelps-Teraski D, Phelps-Gunn T. *Test of Pragmatic Language: Examiner's manual*. 2ª ed. Austin, TX: PRO-ED, Inc; 2007.
13. Carvalho CA. Função pragmática da linguagem e compreensão leitora na dislexia [tese]. São Paulo: Universidade Federal de São Paulo; 2013.
14. Hoffmann A, Martens MA, Fox R, Rabidou P, Andridge R. Pragmatic language assessment in Williams syndrome: a comparison of the Test of Pragmatic Language-2 and the Children's Communication Checklist-2. *Am J Speech Lang Pathol*. 2013;22(2):198-204.
15. Volden J, Philips L. Measuring pragmatic language in speakers with autism spectrum disorders: comparing the children's communication checklist-2 and the Test of Pragmatic Language. *Am J Speech Lang Pathol*. 2010;19(3):204-12.
16. Young EC, Diehl JJ, Morris D, Hyman SL, Bennetto L. The use of two language tests to identify pragmatic language problems in children with autism spectrum disorders. *Lang Speech Hear Serv Sch*. 2005;36(1):62-72.
17. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures. *Spine (Phila Pa 1976)*. 2000;25(24):3186-91.
18. Carvalho CA, Ávila CR. Adaptação linguística e cultural para o Português Brasileiro da segunda versão do Test of Pragmatic Language: projeto piloto In: XIX Congresso de Iniciação Científica PIBIC-UNIFESP – Anais. São Paulo; 2011.
19. Chadha NK. *Applied psychometry*. Delhi, India: Sage; 2009.
20. Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Medical Educ*. 2011;2:53-5.
21. Yurdugül H. Minimum sample size for Cronbach's coefficient alpha: a Monte-Carlo study. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi (H. U. Journal of Education)*. 2008;35:397-405.