

ESTUDO CORRELACIONAL DA BATERIA INFORMATIZADA DE LINGUAGEM ORAL (VERSÃO 3) COM O VOCABULÁRIO RECEPTIVO

CORRELATIONAL STUDY OF THE COMPUTER BASED TEST OF ORAL LANGUAGE (VERSION 3) WITH THE RECEPTIVE VOCABULARY

Maria Cristina Rodrigues Azevedo Joly
Aline Christina Istome

Universidade de Brasília

Sobre os autores

Maria Cristina Rodrigues
Azevedo Joly
Universidade de Brasília
cristina@networkedu.com.br
Rua Dona Libania, 2191/22
Campinas/SP- CEP 13015-090

Aline Christina Istome
Mestre em Avaliação Psicológica
pela Universidade São Francisco,
alineistome@hotmail.com
Rua Dona Libania, 2191/22
Campinas/SP- CEP 13015-090

RESUMO

A avaliação e detecção precoces de distúrbios na linguagem oral permitem a introdução eficaz de procedimentos de intervenção, diminuindo a incidência ou a severidade de problemas na alfabetização. Deste modo, o presente estudo teve como objetivo avaliar a compreensão da linguagem oral, buscando evidências de validade para a Bateria Informatizada de Linguagem Oral versão 3 (BILOv3) por meio da correlação com o desempenho em vocabulário, avaliado pelo Teste de Vocabulário por Imagens Peabody (TVIP) computadorizado. Participaram 157 alunos de ambos os sexos com idade entre quatro e sete anos, que cursavam o último nível de ensino da educação infantil e primeiro ano do ensino fundamental de escolas particulares do estado São Paulo (Brasil). A aplicação ocorreu coletivamente em duas sessões com intervalo de uma semana. Os resultados indicaram aumento significativo nos escores da BILOv3 e do TVIP em função da idade e do nível escolar, bem como, a diminuição significativa do tempo de execução, fornecendo à BILO evidência de validade teste-critério. A MANOVA indicou a interação da idade, nível escolar e sexo tendo o vocabulário como covariante; e o coeficiente de correlação de Pearson indicou relação entre todos os escores por subteste da BILOv3 com o TVIP, fornecendo evidência de validade convergente à BILOv3.

Palavras-chave: Psicometria, Avaliação educacional, Compreensão, testes de linguagem.

ABSTRACT

The assessment and early detection of oral language disorders allow the introduction of effective intervention procedures, reducing the incidence or severity of problems in literacy. Thus, the present study aimed to assess the oral language comprehension, seeking evidence for the validity of Computer Based Test of Oral Language version 3 (BILOv3) by correlation with performance in vocabulary, assessed by the Peabody Test of Vocabulary by Images (TVIP) computerized. 157 students of both sexes aged between four and seven years old attending the last grade level of kindergarten and first grade of primary education private schools of São Paulo State (Brazil) took part in the study. The application occurred collectively in two sessions, one week apart. The results showed

significant increase in scores and the TVIP BILOv3 according to age and grade level, as well as the significant decrease of the execution time, providing the BILO evidence of test-criterion. MANOVA indicated the interaction of age, educational level and gender with the vocabulary as covariate, and the Pearson correlation coefficient indicated relationship between all of BILOv3 subtest scores for the TVIP, providing evidence of convergent validity for BILOv3.

Keywords: Psychometrics, Educational measurement, Comprehension, Language tests.

1 - INTRODUÇÃO

A avaliação psicológica tem como objetivos obter informações a respeito de diferentes dimensões psicológicas do indivíduo ou grupo, e, por conseguinte, levar a uma maior compreensão dos mesmos. Entre os instrumentos psicológicos que auxiliam a avaliação estão os testes informatizados, ou seja, aqueles nos quais a organização e a aplicação são realizadas por aplicativos instalados no computador ou por sistemas acessados via internet (ADANÉS, 1999; JOLY; NORONHA, 2006).

Embora o uso de testes de psicológicos informatizados ainda seja restrito (JOLY; MARTINS; ABREU; SOUZA; COZZA, 2004), a área de construção destes testes tende a avançar rapidamente, tornando esta modalidade de testagem psicológica um dos recursos científicos de maior qualidade, considerando as especificidades de precisão e fidedignidade (JOLY; NORONHA, 2006). Além disso, dentre as vantagens possibilitadas por esses testes, está o fato de que os recursos visuais, sonoros e multimídia informatizados permitem a simulação de situações que levam os testandos a apresentarem suas habilidades em condições mais realistas. Tais situações e condições mais autênticas seriam difíceis ou impraticáveis em testes do tipo lápis-papel padronizados (REYNOLDS; LIVINGSTON; WILSON, 2009).

Ressalta-se a necessidade de que os instrumentos de testagem devam passar por

estudos psicométricos que lhe confirmem validade, ou seja, que permitam atestar o grau de representatividade do domínio teórico que o instrumento se baseia, possibilitando a interpretação correta do fenômeno observado. Dentre as formas de se obter validade, destaca-se, para o presente estudo, a validade concorrente que diz respeito à relação dos escores do teste com outras variáveis. Esta validade pode ser realizada por dois procedimentos: pela relação convergente (relação entre escores do teste e outras medidas com a intenção de avaliar construtos similares) ou pela relação teste-critério (busca compreender a eficácia dos escores do teste em prever determinado critério) (AERA; APA; NMCE, 1999; ANASTASI; URBINA, 2000).

Isto posto, considerando-se os aspectos envolvidos na avaliação psicológica informatizada, o presente estudo buscou evidências de validade convergente e teste-critério para um instrumento informatizado destinado a avaliar a compreensão da linguagem oral em crianças do Ensino Infantil e Fundamental no Brasil. Capovilla, Joly e Tonelotto (2006) destacam que é de fundamental importância desenvolver, obter dados de precisão, validade e normatização de instrumentos para avaliar a linguagem oral em crianças, pois a avaliação e detecção precoces de distúrbios permitem a introdução eficaz de procedimentos de intervenção, diminuindo a incidência ou a severidade de problemas na alfabetização. Isto porque diversas pesquisas têm relacionado o desenvolvimento da linguagem oral com o posterior desenvolvimento da leitura e escrita, salientando

que as dificuldades de aprendizagem observadas em alunos a partir do ingresso no ensino formal, podem estar associadas a comprometimento da linguagem oral (CAPOVILLA; DIAS, 2008; GÓMEZ et al., 2007, HENAO et al., 2008).

Considerando que a linguagem oral está organizada nos aspectos fonológico, sintático, semântico e pragmático, e sendo ela constituída por diferentes habilidades, como nomeação, repetição, fluência, consciência fonológica e consciência sintática, entre outras, já existem instrumentos brasileiros destinados à avaliação destas habilidades, conforme levantamento realizado por Capovilla, Joly e Tonelotto (2006). Entre estes instrumentos estão a Lista de Avaliação de Vocabulário Expressivo – LAVE, que avalia quais palavras uma criança fala; o Teste de Vocabulário por Imagens *Peabody*, que avalia o vocabulário receptivo; o Teste Infantil de Nomeação, que avalia a capacidade de nomear figuras; a Prova de Consciência Fonológica, que avalia a capacidade de manipular segmentos da fala; e a Prova de Consciência Sintática, que está diretamente correlacionada ao desempenho em consciência fonológica, leitura, escrita e vocabulário.

Pesquisas relacionadas à linguagem oral têm sido realizadas em vários países da América Latina. Exemplo é a pesquisa de Gómez et al (2007) que avaliou características da consciência fonológica (capacidade de analisar e segmentar os componentes da fala e efetuar relações complexas sobre eles) e as habilidades verbais, de crianças com idades entre 7 e 10 anos da cidade de Medellín (Colômbia). Também Henao et al (2008) avaliaram alunos da cidade de Medellín (Colômbia) com idades entre 6 e 16 anos de idade diagnosticadas com Transtorno Específico do Desenvolvimento da Linguagem (TEDL) e confirmaram que alunos com TELD, apresentaram dificuldades tanto nos níveis de compreensão, quanto de expressão da linguagem, e que estas dificuldades afetam o desempenho escolar destes alunos.

Já em 2010, Crespo-Allende e Alfaro-Faccio, visando investigar o desenvolvimento da consciência metapragmática (conhecimento das relações entre contexto e mensagem,

consciência e autocontrole da compreensão e expressão linguística) aplicaram a prova de consciência metapragmática, um instrumento de avaliação informatizado, em 160 alunos com idades de 6, 8, 10 e 12 anos de escolas particulares de Valparaíso e Quilpué (Chile). Os resultados apontaram que, com o aumento da idade, aumentam as respostas metapragmáticas (conscientes) e diminuíram as respostas automáticas, uma vez que, as diferenças foram significativas entre os diferentes grupos etários.

Para o presente estudo considera-se como habilidade oral a ser investigada a linguagem receptiva, que se refere aos canais de entrada da informação, que comportam tanto a audição e a compreensão da linguagem falada, como também a visão (GIL, 2002). É esta habilidade que a Bateria Informatizada de Linguagem Oral – BILOv3 (JOLY, 2006) tem como foco avaliativo, porquanto está pautada no exame do que a criança é capaz de entender, e na avaliação do conteúdo da linguagem (ACOSTA et al 2003).

Deve-se considerar também que a BILO permite a aferição do tempo utilizado em cada item, em cada subteste e no total da bateria. A importância de tal medida cognitiva respalda-se no fator *Rapidez de Decisão* (Gt) descrito pelo modelo multidimensional da inteligência de *Cattell-Horn-Carroll* – CHC. O Gt indica a rapidez em reagir ou tomar decisões envolvendo processamentos mais complexos, posto que quantifica a latência média de resposta, ou seja, o tempo de intervalo entre o aparecimento do estímulo e a resposta (FLANAGAN et al, 2002).

A BILO é um instrumento que se encontra na terceira versão, sendo que os estudos anteriores para a versão 1 registram os estágios de validação no processo de construção e desenvolvimento, pelos quais o instrumento passou até a consolidação da versão atual (BILOv3).

Na versão 1, Joly et al (2008) encontraram evidências de validade de construto para a Bateria Informatizada de Linguagem Oral – BILO. Almeida e Joly (2008) indicaram

validade para a BILOv1, correlacionando-a com o desempenho em vocabulário, verificado por meio do Teste de Vocabulário por Imagens Peabody (TVIP). Igualmente, Joly e Piovezan (2012) encontraram validade para a BILOv1 quando a relacionaram com a habilidade de raciocinar globalmente, avaliada pela Escala de Maturidade Mental Colúmbia. Issa (2008) demonstrou evidências de validade para a BILOv1, correlacionando-a com a Escala de Transtornos de Déficit de Atenção e Hiperatividade para Professores. Joly e Dias (2009), encontraram evidências de validade de construto para a BILOv1 por meio da análise fatorial com informação plena nos escores, no entanto, apesar dos resultados obtidos revelarem que os itens de cada subteste possuíam boas cargas fatoriais e as que as provas possuíam unidimensionalidade, verificou-se a necessidade de retirada de alguns itens, o que determinou a reformulação do instrumento, dando origem a BILOv2.

Em relação à versão 2 da BILO, Joly et al (2009a), buscaram evidências de validade correlacionando-a com o Teste de Vocabulário por Imagens Peabody (TVIP), assim como realizado com a BILOv1 (Joly & Almeida, 2008). Também buscando validade para a BILOv2, Joly et al (2009b), o fizeram por meio da correlação com o Teste de Competência de Leitura Silenciosa (TeCoLeSi). Nos dois estudos relativos a BILOv2, apesar de serem encontradas evidências de validade, foram indicadas alterações relativas às provas Completar Frases e Completar Histórias, definido-se como necessária a versão 3 da BILO.

Frente ao exposto e tendo em vista as adaptações da BILO, considera-se pertinente que, para a validação da terceira versão seja realizada a correlação entre a BILOv3 com o Teste de Vocabulário por Imagens Peabody – TVIP. O processo justifica-se pelo fato de que os estudos com as versões anteriores (JOLY; ALMEIDA, 2008; JOLY et al, 2009a) indicaram que o vocabulário receptivo correlaciona-se positiva e significativamente com os escores por provas e total da BILO.

No tocante ao uso do TVIP, para este estudo, foi utilizado em sua versão informatizada, haja vista, esta demandar menos tempo de aplicação, facilidade na tabulação de dados e aferição de resultados. Macedo et al (2006) compararam as pontuações de estudantes nas versões tradicional e computadorizada do TVIP. Foram avaliadas 90 crianças de uma escola particular do município de São Paulo, matriculadas no infantil I, no infantil II e no infantil III. A análise de regressão do escore na versão tradicional em relação à versão computadorizada revelou correlação positiva ($r = 0,646$; $r^2 = 0,418$) significativa entre os desempenhos nas duas versões ($F[1,89] = 63,187$; $p = 0,001$) indicando que crianças com bom desempenho na versão tradicional estão inclinadas a repetir o mesmo resultado na versão computadorizada.

Trevisan et al (2008) visando avaliar a habilidade de vocabulário receptivo em crianças de 1ª à 4ª série do ensino fundamental e verificar a progressão ao longo das faixas de escolaridade, avaliaram 363 crianças de ambos os sexos com idades entre 6 anos e 6 meses e 11 anos e 9 meses de uma escola municipal do interior de São Paulo. Foram utilizados como instrumentos avaliativos as Matrizes Progressivas de Raven e o Teste de vocabulário por Imagens Peabody (TVIP) versão computadorizada. As avaliações ocorreram em duas sessões. A Análise de Covariância (ANCOVA) tendo o percentil do Raven como covariante, e a idade e série como fatores, revelou efeito significativo de idade ($F[3, 251] = 9,05$; $p \leq 0,001$) e do percentil no Raven ($F[1, 249] = 21,38$; $p \leq 0,001$). Também revelou o efeito significativo da série ($F [3, 251] = 20,42$; $p \leq 0,001$) e do percentil no Raven ($F [1, 251] = 17,92$; $p \leq 0,001$). Ou seja, os escores aumentaram com a progressão da idade e da série e a habilidade de vocabulário receptivo esteve relacionada com a inteligência não verbal, com a idade e com a série escolar. Além disso, foi encontrado um bom índice de precisão do TVIP, verificado pelo Alfa de Crombach igual a 0,87.

Destarte, o presente estudo objetivou avaliar a compreensão da linguagem oral e vocabulário receptivo para estudantes do ensino infantil e anos iniciais do ensino fundamental e buscar evidências de validade para a Bateria Informatizada de Linguagem Oral versão 3 (BILOv3), por meio da correlação com o desempenho em vocabulário, avaliado pelo Teste de Vocabulário por Imagens Peabody. Também foram verificadas se o desempenho na BILO revela diferenças relativas à série, idade e gênero.

2. MÉTODO

Participantes

Participaram do presente estudo 157 alunos, sendo 52,2% meninas e 47,8% meninos, com idade entre quatro e sete anos ($M = 5,37$ e $DP = 0,59$). Os participantes estavam regularmente matriculados no último ano da Educação Infantil e primeiro ano do ensino fundamental de escolas particulares do estado de São Paulo, Brasil.

Instrumentos

- Bateria Informatizada de Linguagem Oral-BILOv3 (JOLY, 2009)

Tem por objetivo avaliar a compreensão da linguagem oral em alunos da educação infantil ao ensino fundamental I. Sendo um instrumento informatizado, possui recursos de multimídia e interface com banco de dados para armazenamento das respostas dadas. Em se tratando de uma avaliação voltada para a linguagem oral com respondentes que não adquiriram o código formal de leitura, optou-se por respostas apresentadas como desenhos simples feitos apenas com linhas, sem detalhes. Em um banco de dados são registrados a opção de resposta e o tempo total por item.

É constituída pelas provas: Compreensão Morfosintática (CM), que tem por objetivo avaliar a relação entre um vocábulo e a sua representação gráfica, revelando a compreensão

do significado; Seqüência Lógica (SL), que avalia a organização lógica e temporal do respondente para estímulos visuais organizados em cenas cujo conjunto seqüenciado resulta em uma história; Organização Lógico verbal (OLV), que avalia a compreensão receptiva do conteúdo de uma história apresentada por completo em vídeo e oralmente, por meio de uma organização lógica de cenas que representam esse conteúdo; Interpretação de Histórias (IH), que avalia a compreensão receptiva da história por meio de questões de múltipla escolha; Completar Frases (CF), que avalia a compreensão de palavras organizadas em frases; e Completar Histórias (CH) que avalia a compreensão oral de histórias. Tanto a CF quanto a CH foram baseadas no Sistema Orientado de Cloze (JOLY, 2007), que possibilita a organização de um texto para avaliar compreensão em leitura, com base em critérios específicos relativos a número de palavras, omissões de vocábulos, tamanho de lacunas e opções de resposta, visando determinar níveis de dificuldade de texto em função da habilidade de compreensão dos respondentes. em que a tarefa do respondente é identificar que alternativa, dentre as opções, dá melhor sentido à frase.

A aplicação da BILO é coletiva e orientada por um aplicador e um auxiliar e deve ser realizada em um laboratório de informática para grupos de 15 respondentes, no máximo. O tempo médio de aplicação é de 40 minutos. Cada prova contém 10 itens de múltipla escolha com três opções de resposta. A somatória dos pontos por item determina o escore por prova, sendo que o escore máximo obtido é de 20 pontos por prova. A somatória dos escores por prova resulta no escore total, que perfaz o máximo de 120 pontos.

Teste de Vocabulário por Imagens Peabody – TVIP (Capovilla & Capovilla, 1997)

Avalia o desenvolvimento lexical no domínio receptivo em pessoas com idades entre 2a6m até 18 anos. Permite uma avaliação objetiva, rápida e precisa do vocabulário receptivo auditivo em ampla variedade de áreas, incluindo pessoas, ações, qualidades, partes do corpo, tempo,

natureza, lugares, objetos, animais, termos matemáticos, ferramentas e instrumentos. É recomendado para indicar o nível de desenvolvimento da linguagem receptiva em pré-escolares, bem como em crianças e adultos incapazes de ler, escrever ou mesmo de falar, já que, a figura corresponde à palavra falada pelo examinador. A criança deve apontar a figura que corresponde à palavra que ouve e marcar um x. O critério de correção corresponde a um ponto para os acertos e zero ponto para os erros. Quanto maior o número de pontos, melhor o desempenho da criança. No presente estudo foi usada a versão computadorizada do TVIP (MACEDO et al, 2006) que foi produzida a partir da adaptação hispano-americana de 125 itens, cada um com quatro desenhos, adaptada e validada e normatizada no Brasil. As mensagens solicitam que a criança escolha uma das figuras são emitidas a partir de arquivos de sons digitalizados e salvos no formato WAV. Uma vantagem da versão computadorizada do teste em relação à versão tradicional é a de registrar todas as respostas do avaliando, bem como o tempo consumido em responder. Além disso, essa versão pode ser configurada rapidamente para apresentar os comandos verbais, escritos ou, ainda, em diferentes línguas.

Equipamentos

Microcomputadores que possuam como requisitos mínimos do sistema processador Pentium 4; memória mínima de 256 MB; espaço em disco de 180 MB; plataforma Windows XP, Vista ou 7; com resolução de tela 1024 por 768 pixels; e fones de ouvido.

Procedimento

Inicialmente foram solicitadas as autorizações das escolas onde foram coletados os dados. O projeto de pesquisa foi submetido à avaliação do Comitê de Ética da Universidade São Francisco e foi aprovado sob Protocolo CAAE nº 0272.0.142.000-09. Após a liberação, buscou-se autorização dos responsáveis pelos alunos por meio do Termo de consentimento Livre e Esclarecido (anexo). A aplicação ocorreu em duas sessões, uma para cada instrumento, tendo uma semana de distância entre a 1ª e a 2ª aplicação. Foi aplicada a BILO no primeiro dia e o TVIP no segundo. As aplicações foram realizadas coletivamente no laboratório de informática das escolas, com turmas equivalentes ao número de computadores disponíveis para comportar os participantes, em horário previamente agendado.

3. RESULTADOS

Os dados referentes ao desempenho em compreensão da linguagem oral e do vocabulário podem ser observados na Tabela 1 que apresenta a análise descritiva dos escores dos subtestes e do total da BILO e do TVIP em função do nível escolar. Observa-se que as médias de acertos aumentaram no 1º ano do ensino fundamental quando comparados com os escores do Ensino Infantil.

Tabela 1: Estatísticas descritivas do desempenho nos subtestes da BILO e no TVIP por série.

Nível de Ensino	Estatística descritiva	Provas						BILO total	TVIP
		CM	SL	OLV	IH	CF	CH		
Ed. Infantil N = 82	Mínimo	6	1	2	7	0	10	50	36
	Máximo	20	16	19	20	20	20	113	77
	Média	17,83	6,11	7,27	15,99	16,62	18,06	81,88	60,26
	<i>DP</i>	2,84	4,09	4,35	3,63	3,37	2,46	14,94	9,41
1º Ano N = 75	Mínimo	9	0	0	8	8	9	61	36
	Máximo	20	20	20	20	20	20	119	87
	Média	18,60	9,41	9,60	17,83	17,85	18,85	92,15	67,55
	<i>DP</i>	2,26	6,37	5,88	2,85	2,47	1,64	15,24	10,23

Com a finalidade de verificar se as diferenças encontradas entre as médias dos alunos da educação infantil e do ensino fundamental foram estatisticamente significativas, utilizou-se o Teste *t* de Student. Exceto pela prova Compreensão Morfossintática, que teve diferença marginalmente significativa ($t[155] = -1,866$; $p \leq 0,064$), todas as provas ($t_{SL}[152] = -3,901$; $p \leq 0,001$; $t_{OLV}[155] = -2,840$; $p = 0,005$; $t_{IH}[155] = -3,504$; $p = 0,001$; $t_{CF}[155] = -2,592$; $p = 0,010$; $t_{CH}[155] = -4,257$; $p \leq 0,001$), inclusive o escore total da BILO ($t[155] = -4,652$; $p \leq 0,000$), diferiam significativamente entre os níveis de ensino, assim como encontrado por Joly, Reppold e Dias (2009) com a BILOv2, tanto na população paulista como na gaúcha. Os estudos anteriores com a BILOv1 (ALMEIDA; JOLY, 2008; ISSA, 2008; JOLY; PIOVEZAN, 2012) e com a BILOv2 (JOLY et al, 2009a; JOLY et al, 2009b) também comprovaram o efeito do nível escolar sobre o desempenho na compreensão da linguagem oral avaliado pela BILOv3. Estes

resultados atribuem à BILOv3 validade de critério por escolaridade (AERA; APA; NMCE, 1999; ANASTASI; URBINA, 2000).

Em relação ao TVIP, a média do desempenho dos alunos também aumentou no 1º ano do ensino fundamental em comparação com a média dos alunos no ensino infantil. Ao aplicar o Teste *t* de Student, foi possível comprovar que as diferenças foram estatisticamente significativas ($t[150,521] = -4,635$; $p \leq 0,000$). Deste modo, pode-se afirmar que o instrumento foi sensível ao desenvolvimento acadêmico, ou seja, que as crianças do 1º ano possuem um vocabulário receptivo maior do que as crianças da educação infantil, conforme encontrado por Almeida e Joly (2008). Trevisan et al (2008) também encontraram aumento da pontuação no TVIP com a progressão do nível escolar.

No tocante às diferenças por idade, a análise de variância das médias foi obtida por meio da comparação de dois grupos etários: alunos com cinco anos ou menos e alunos com

seis anos ou mais, por meio do Teste *t* de Student. Na BILOv3 foi possível observar que somente a prova CM mostrou-se marginalmente significativa ($t[152,025] = -1,885$; $p \leq 0,061$), sendo que, tanto as demais provas quanto o escore total da BILO ($t[153,164] = -4,257$; $p \leq 0,000$) tiveram diferenças significativas ($p \leq 0,05$) ou muito significativas ($p \leq 0,001$), indicando que a diferença encontrada entre a média de pontuação das crianças com cinco anos ou menos e das crianças com seis anos ou mais não pode ser atribuída ao acaso, como

indica a Tabela 2. Este resultado revela evidências de validade da BILOv3 para o critério relativo ao desenvolvimento (AERA; APA; NMCE, 1999; ANASTASI; URBINA, 2000), corroborando os estudos de Almeida e Joly (2008); Issa (2008) e Joly e Piovezan (2012) com a BILOv1, e os estudos de Joly, Reppold e Dias (2009) e Joly et al (2009a), com a BILOv2.

Tabela 2: Estatísticas descritiva e inferencial da BILO por subtestes e total em relação ao grupo etário ≤ 5 (58,6%) e ≥ 6 (41,4%).

Prova	Faixa etária	Número de participantes	Média	Desvio padrão	<i>gl</i>	<i>t</i>	<i>p</i> ≤
CM	≤5	92	17,91	2,65	152,025	-1,885	0,061
	≥6	65	18,60	2,51			
SL	≤5	92	6,51	4,69	155	-3,901	0,000
	≥6	65	9,35	6,21			
OLV	≤5	92	7,59	4,76	155	-2,840	0,005
	≥6	65	9,51	5,73			
IH	≤5	92	16,18	3,60	155	-3,504	0,001
	≥6	65	17,83	2,85			
CF	≤5	92	16,91	3,25	155	-2,592	0,010
	≥6	65	17,63	2,65			
CH	≤5	92	18,17	2,37	155	-2,350	0,020
	≥6	65	18,82	1,72			
BILO total	≤5	92	83,28	15,44	153,164	-4,257	0,000
	≥6	65	91,74	15,29			

Como forma de verificar possíveis diferenças de desempenho na BILO em função do sexo, foi realizado o Teste *t* de Student para comparar a pontuação dos meninos e das meninas. Observou-se que as meninas obtiveram médias mais altas nas provas CM, SL, OLV, CH e na pontuação total da BILO ($t[155] = 0,549$;

$p \leq 0,584$), enquanto que as médias dos meninos superaram as das meninas nas provas IH e CF.

Entretanto, tais diferenças encontradas não podem ser consideradas, haja vista não serem estatisticamente significativas, sendo atribuídas, portanto, ao acaso. Esses achados corroboram os estudos anteriores de Almeida e Joly (2008); Issa (2008) e Joly e Piovezan (2012) com a

BILOv1, e os estudos de Joly, Reppold e Dias (2009) e Joly et al, 2009a), com a BILOv2, nos quais também não foram encontradas diferenças de sexo.

Igualmente no TVIP, não houve diferença significativa ($t[155] = -0,559$; $p \leq 0,577$) entre a média em vocabulário dos meninos ($M = 64,23$; $DP \leq 9,12$) e das meninas ($M = 63,29$; $DP = 11,55$). Também no estudo de Macedo et al (2006) não foi evidenciado o efeito do gênero sobre o desempenho no TVIP.

Visando analisar as médias da BILO tendo como covariante o desempenho no TVIP, realizou-se uma Análise de Covariância Multivariada (MANCOVA). Para tanto, definiu-se como variáveis dependentes os escores da BILO (por prova e total) covariando com os escores do TVIP e como variáveis independentes o sexo, faixa etária, nível escolar e escola frequentada. A MANCOVA indicou diferenças significativas nos escores da BILO, tendo como covariante significativo o desempenho em vocabulário receptivo (TVIP) somente para a variável nível escolar nas provas SL, OLV e na BILO total ($F[1] = 8,124$; $p = 0,005$).

Ao lado disso, constatou-se que houve interação significativa entre as variáveis sexo e escola para o desempenho no escore total da BILO ($F[2] = 4,492$; $p \leq 0,013$) indicando que o sexo é uma variável que se comporta de modo diferenciado em função de cada instituição de ensino onde ocorreu a coleta de dados. Os meninos de uma das escolas particulares, não-confessional, revelaram desempenho menor e inferior ao das meninas. Houve também interação significativa entre nível escolar, idade e sexo na diferença de médias das provas CM ($F[1] = 4,843$; $p \leq 0,030$); SL ($F[1] = 4,731$; $p \leq 0,031$) e no desempenho total da BILO ($F[1] = 7,532$; $p \leq 0,007$), indicando que estas

três variáveis funcionam de modo interdependente para a amostra investigada. As meninas mais velhas e que frequentavam as duas escolas confessionais particulares apresentaram o desempenho superior ao dos meninos menores e da escola particular, não-confessional. Nesse sentido, a interação entre as variáveis pode indicar a necessidade de investigar variáveis relativas a currículo, princípios do projeto político-pedagógico das escolas envolvidas a fim de buscar elementos que justifiquem as interações constatadas, para além das anteriormente observadas em estudos preliminares com a BILO.

Com o objetivo de analisar o tempo utilizado pelas crianças na execução das tarefas de cada instrumento, uma vez que esta medida possibilita inferir acerca do processamento cognitivo (FLANAGAN et al, 2002), são apresentadas a seguir as estatísticas descritivas do tempo. É importante destacar que o tempo na BILO começa a ser registrado a partir do momento em que se finaliza o áudio contendo a consigna de cada item e termina quando a criança confirma a resposta. Por isso, o tempo apresentado nas análises não é o tempo total em que a criança manteve-se em atividade, pois este não contempla o tempo usado para o áudio. O tempo de cada prova da BILO é verificado a partir da soma do tempo gasto em cada item, bem como o tempo total da BILO é a somatória do tempo gasto nos itens de todas as provas. Ressalta-se que o tempo no TVIP é originariamente registrado em milésimos de segundos e na BILO o tempo é registrado em segundos. Na exposição dos dados na Tabela 3, para que ficassem equivalentes as unidades de medida, de modo a facilitar a visualização e comparação, o tempo do TVIP foi convertido para segundos.

Tabela 3: Estatísticas descritivas do tempo em segundos para a realização da BILO do TVIP.

Prova	Mínimo	Máximo	Média	Desvio padrão
CM	2	265	31,55	32,887
SL	73	1209	233,90	146,956
OLV	54	979	182,53	106,180
IH	2	149	27,36	24,748
CF	0	107	20,41	15,511
CH	0	93	13,66	14,025
Total BILO	197	1868	509,42	272,038
TVIP	312,91	1592,86	562,68	168,908

Observa-se na Tabela 3 que, de todas as provas da BILO, a SL e a OLV foram as provas que demandaram maior tempo de execução. Isto provavelmente se dá devido ao fato de que a tarefa do respondente é organizar as cenas de cada série de modo que haja relação de causa e efeito e contigüidade entre elas, o que requer a capacidade de memória para seqüência, orientação direita-esquerda e orientação temporal, além da maior habilidade de manuseio do mouse. Nas demais provas a tarefa exigia que se escolhesse uma opção dentre três, portanto, o aluno deveria clicar uma vez sobre a resposta que julgasse correta, o que justifica os menores tempos na execução da tarefa. O aparecimento de tempos ínfimos como zero e dois segundos para a execução de uma prova inteira (IH, CF, CH) se deu porque quando o aluno clica em sua escolha imediatamente depois de terminado o áudio (pois as figuras já estão expostas), o programa não é sensível para registrar o tempo, uma vez que sua unidade de medida é o segundo, diferentemente do TVIP que se utiliza de milissegundos.

Visando verificar se houve diferença no tempo de execução das provas em função do sexo, com teste *t* de Student comparou-se as médias dos meninos ($M = 444,09$; $DP = 178,30$) e das meninas ($M = 569,17$; $DP = 325,45$) na BILOv3. Observou-se que as meninas utilizaram mais tempo na execução da BILOv3 e que esta diferença foi significativa ($t[155] = 2,948$; $p = 0,004$). Já no TVIP, apesar da média de tempo das meninas ($M = 574,48$;

$DP = 170,31$) ter sido superior à dos meninos ($M = 549,77$; $DP = 167,53$), esta não se mostrou significativa ($t[155] = 0,915$; $p \leq 0,361$), podendo ser atribuída ao acaso; conforme encontrado no estudo de Macedo e cols. (2006).

Em relação à diferença de tempo em função do nível de ensino, exceto pela prova CM, todas as provas da BILO e o tempo da BILO total tiveram diferenças significativas ($t[155] = 3,623$; $p \leq 0,000$) indicando que as crianças do ensino infantil levaram mais tempo para realizar as atividades do que as crianças do primeiro ano do ensino fundamental. No TVIP também foram encontradas diferenças significativas ($t[155] = 4,688$; $p \leq 0,000$).

A análise da diferença de tempo para realização dos testes em função da idade, mostrou que houve diferença significativa entre as crianças com cinco anos (ou menos) e as de seis anos ou mais para na execução de todas as provas e no total da BILO ($t[155] = 3,627$; $p \leq 0,000$) e também no TVIP ($t[155] = 4,740$; $p \leq 0,000$). Macedo et al. (2006) também comprovaram o efeito da idade sobre o tempo na realização do TVIP, contudo somente entre crianças de quatro e seis anos.

Por fim, como forma de verificar a existência de consistência na relação entre os dois testes, obtendo assim, a evidência de validade convergente (AERA; APA; NMCE, 1999; ANASTASI; URBINA, 2000) para a BILO, a Tabela 4 apresenta a análise de correlacional entre os escores da BILO por prova e total e o desempenho no TVIP. Observam-se correlações

positivas moderadas entre as provas SL, OLV, IH e CH da BILO e o desempenho em vocabulário receptivo (TVIP), destacando-se a correlação entre escore total da BILO e o TVIP ($r = 487$; $p \leq 0,000$). Somente nas provas CM e CF obteve-se correlação positiva fraca, porém não menos significativa ($p \leq 0,000$). As correlações positivas e estatisticamente significativas atribuem validade convergente para a BILO, conforme encontrado nos estudos de Almeida e Joly (2008) para a versão 1 e Joly et al (2009a) para a versão 2, por meio da relação com o vocabulário receptivo.

Tabela 4: Correlação entre a Bateria Informatizada de Linguagem Oral e o Teste de Vocabulário por Imagens Peabody - TVIP

Provas	Correlação Pearson	Peabody
CM	<i>R</i>	0,294(**)
	<i>P</i>	0,000
SL	<i>R</i>	0,333(**)
	<i>P</i>	0,000
OLV	<i>R</i>	0,394(**)
	<i>P</i>	0,000
IH	<i>R</i>	0,381(**)
	<i>P</i>	0,000
CF	<i>R</i>	0,272(**)
	<i>P</i>	0,001
CH	<i>R</i>	0,372(**)
	<i>P</i>	0,000
Total	<i>R</i>	0,478(**)
	<i>P</i>	0,000

** $p = 0,001$

4. CONSIDERAÇÕES FINAIS

O presente estudo objetivou avaliar a compreensão da linguagem oral de crianças matriculadas no último ano da educação infantil e no 1º ano da educação fundamental, buscando evidências de validade para a Bateria Informatizada de Linguagem Oral versão 3 (BILOv3). Constatou-se que a terceira versão da BILO também tem forte influência do

vocabulário, conforme observado nos resultados apontados pela MANOVA dos escores da BILO por prova e total com o TVIP covariando. Ao lado disso, as correlações moderadas entre os escores da BILOv3 e o TVIP indicaram que os dois instrumentos requerem uma capacidade cognitiva em comum, ratificando, portanto, que o vocabulário está fortemente correlacionado com a habilidade de compreensão. Considera-se, pois, que a baixa compreensão oral tende a estar, em certa medida, correlacionada à falta de conhecimento sobre a palavra. Estes resultados já eram esperados uma vez que nas versões anteriores da BILO, estudos de Almeida e Joly (2008) e Joly et al (2009a) já haviam demonstrado tal correlação. Assim, evidenciou-se validade por meio da relação convergente para a BILOv3 por meio da correlação com o TVIP (AERA; APA; NMCE, 1999; ANASTASI; URBINA, 2000).

Neste estudo também se verificou que os escores da BILOv3 aumentaram significativamente com a progressão da idade e da série, indicando assim, validade para o instrumento por mudanças desenvolvimentais (teste-critério). Tais influências do desenvolvimento já haviam sido encontradas nas versões anteriores da BILOv1 e da BILOv2, conforme já mencionado. Já no que tange à influência do gênero sobre o desempenho em compreensão da linguagem oral, novamente não foram encontradas diferenças significativas entre os escores das meninas e dos meninos, conforme as pesquisas anteriores.

Considera-se que a presente investigação tem a verificação do tempo de execução das provas como uma contribuição adicional aos estudos realizados com as versões anteriores da BILO. Sobre este aspecto, observou-se diferença significativa no tempo utilizado para a realização das provas da BILO em função da idade, da série, indicando que com a progressão da idade e da série há uma diminuição do tempo de execução. Também foi observado que houve diferença significativa em função do gênero no tempo nas provas SL e OLV e no tempo total da BILOv3, indicando que as meninas usaram mais tempo do que os meninos para executar as

tarefas. Contudo, destaca-se que a diferença do tempo entre as meninas e os meninos não influenciou a diferença de pontuação, ou seja o maior tempo não foi determinante de maior pontuação, pois conforme já mencionado, não houve diferença significativa em função do gênero sobre os escores da BILO. Sugere-se que os posteriores estudos atentem-se para as diferenças de tempo. Há que se considerar também que o tempo medido em segundos não aferiu acuradamente a rapidez de decisão (fator Gt do Modelo CHC descrito por FLANAGAN et al, 2002), haja vista alguns itens das provas CF e CH terem apresentado tempo zero. Considerando a importância desta medida enquanto habilidade cognitiva (fator Gt), sugere-se que seja alterado o programa afim de que registre o tempo em milésimos de segundos, garantindo assim, maior aferição.

Faz-se necessário que novos estudos contemplem a análise fatorial com informação plena, bem como uma análise pela Teoria de Resposta ao Item para avaliar a terceira versão da BILO. Também se recomenda a ampliação deste estudo para todo o primeiro ciclo do ensino fundamental, ou seja, até o 5º ano, para que contemple todas as idades e séries que a BILO se propõe a avaliar. Igualmente, julga-se necessário que o instrumento seja correlacionado com instrumentos que avaliem habilidades cognitivas como memória, atenção e inteligência cristalizada, entre outras, uma vez que a linguagem é dependente da interação de tais habilidades que estão envolvidas no processamento de informações.

É importante destacar que a BILO têm se mostrado, desde os estudos com as primeiras versões, como um instrumento diferenciado de avaliação, porquanto os recursos multimídia empregados na sua estrutura permitem, por um lado mensuração acurada das habilidades cognitivas, e por outro, maior motivação e interesse das crianças em executar as tarefas, tornando a testagem uma atividade gratificante tanto para o avaliador quanto para o avaliando. Constatou-se por meio das análises realizadas no presente estudo que a BILO^{v3} possui características psicométricas que a validam e

qualificam como um instrumento psicoeducacional que contribui para a compreensão dos processos cognitivos envolvidos na compreensão da linguagem oral. Apresenta-se, assim, de forma responsável e eficiente como recurso de avaliação e detecção precoce de distúrbios na linguagem receptiva que podem vir a influenciar a aquisição da leitura e escrita.

REFERÊNCIAS

ACOSTA, V. M. ET AL. **Avaliação da linguagem: teoria e prática do processamento de avaliação do comportamento lingüístico infantil**. São Paulo: Livraria Santos, 2003

ADÁNEZ, G. A. Procedimientos de construcción y análisis de tests psicométricos. In: WECHESLER S. M.; GUZZO R. S. L. (Orgs.) **Avaliação Psicológica: Perspectiva Internacional**. São Paulo: Casa do Psicólogo, 1999. p.57-100.

ALMEIDA, A. R.; JOLY, M. C. R. A. Estudo correlacional entre a Bateria Informatizada de Linguagem Oral (BILO) e Peabody. In: XIII CONFERÊNCIA INTERNACIONAL EM AVALIAÇÃO PSICOLÓGICA: FORMAS E CONTEXTOS, BRAGA. **Actas do XIII Conferência Internacional em Avaliação Psicológica: Formas e Contextos**. Braga: Psiquilíbrios, 2008.

AERA, APA; NMCE – American Educational Research Association, American Psychological Association and National Council on Measurement in Education. **Standards for educational and psychological testing**. New York: American Educational Research Association, 1999.

ANASTASI, A.; URBINA, S. **Testagem Psicológica**. 7ª ed. Porto Alegre: Artes Médicas, 2000.

CAPOVILLA, A. G. S.; DIAS, N. M. Habilidades de linguagem oral e sua contribuição para a posterior aquisição de leitura. **Psic**, v. 9, n. 2, 2008. p. 135-144.

- CAPOVILLA, A. S. G.; JOLY, M. C. R. A.; TONELLOTO, J. M. de F. Avaliação neuropsicológica e aprendizagem. In: NORONHA, A. P. P., SANTOS, A. A. A. DOS; SISTO F. F. (Orgs). **Facetas do fazer em Avaliação Psicológica**. São Paulo: Vetor Editora, p. 141 - 162, 2006.
- CRESPO-ALLENDE, N.; ALFARO-FACCIO, P. Desarrollo tardío del lenguaje: la conciencia metapragmática en la edad escolar. **Universitas Psychologica**, Bogotá, v.9. n.1, 229-240, 2010.
- FLANAGAN, D. P. et al. **The achievement test desk reference (ATDR)**. Boston, MA: Allyn and Bacon, 2002.
- GIL, R. **Neuropsicologia**. (tradução: Doria M. P. A. S). São Paulo: Livraria Santos, 2002.
- GÓMEZ, L. A.; DUARTE, A. M.; MERCHÁN, V.; AGUIRRE D. C.; PINEDA, D. A. Conciencia fonológica y comportamiento Verbal en niños con dificultades de Aprendizaje. **Universitas Psychologica**, Bogotá, v. 6., n.3, 571-580, 2007.
- HENAO, L. H., PRIETO, M. G., SALAZAR, D. P., REBOLLEDO, R. C., VÁSQUEZ, J. P. L., VILLAMIZAR, N. M., PÉREZ, A. J. & RAMÍREZ, A. A. (2008). Trastorno Específico del Desarrollo del Lenguaje en una población infantil colombiana. **Universitas Psychologica**, Bogotá, v.7. n.2, maio – agosto, 2008.
- ISSA, G. M. P. **Estudos de Evidências de Validade da Bateria Informatizada da Linguagem Oral – BILO**, 2008. 120p. Dissertação de Mestrado, Programa de Pós-Graduação Stricto Sensu em Psicologia, Universidade São Francisco, Itatiba, São Paulo.
- JOLY, M. C. R. A. A validade do Sistema Orientado de Cloze: um estudo correlacional. **Psicologia Escolar Educacional**, v.11, p. 49-57, 2007.
- JOLY, M. C. R. A. (2009). **Bateria Informatizada de Linguagem Oral – BILOv3**. Teste em desenvolvimento.
- JOLY, M. C. R. A.; DIAS, A. S. Evidências de Validade de uma prova informatizada de Linguagem Oral – BILOv1. **Psicologia: Teoria e Prática**, São Paulo, v. 11, n. 2, p. 83-93, 2009.
- JOLY, M. C. R. A.; NORONHA, A. P. P. (2006). Reflexões sobre a construção de instrumentos psicológicos informatizados. In: NORONHA, A. P. P.; SANTOS, A. A. A. DOS; SISTO, F. F. (Orgs.). **Facetas do fazer em avaliação psicológica**. São Paulo: Vetor, 2006. p.95-105.
- JOLY, M. C. R. A.; PIOVEZAN, N. M. Estudo de validade correlacional e de critério da Bateria Informatizada de Linguagem Oral (BILO) com prova de raciocínio. **Estudos de Psicologia**, Campinas, v.29, p.499-508, 2012.
- JOLY, M. C. R. A. ET AL Estudo Correlacional entre a Bateria Informatizada de Linguagem Oral (BILOv2) e Peabody. Pôster apresentado no XXXII **Congresso Iberoamericano de Diagnóstico e Avaliação Psicológica** – Buenos Aires - Argentina – Buenos Aires – Argentina, 2009a.
- JOLY, M. C. R. A. ET AL. Análise da produção científica em avaliação psicológica informatizada. **Avaliação Psicológica**, São Paulo, v 3, n.2, p 121-129, 2004.
- JOLY, M. C. R. A., ET AL. Bateria Informatizada de Linguagem Oral (BILO): validade de construto. In: XIII Conferência Internacional em Avaliação Psicológica: Formas e Contextos, Braga. **Actas do XIII Conferência Internacional em Avaliação Psicológica: Formas e Contextos**. Braga: Psiquilíbrios, 2008.
- JOLY, M. C. R. A. ET AL. Avaliação das características psicométricas da bateria informatizada de linguagem oral - BILOv2. Pôster apresentado no III **Congresso Latino Americano de Psicologia** – Ulapsi, Cidade do México, México, 2009b
- JOLY, M. C. R. A.; REPPOLD, C. T.; DIAS, A. S. Avaliação da linguagem oral de crianças paulistas e gaúchas pela Bateria Informatizada de Linguagem Oral (BILOv2). In: Hutz, C. (Org.). **Avaliação psicológica de crianças e**

adolescentes. São Paulo: Casa do Psicólogo, 2009. p. 175-208.

MACEDO, E. C. ET AL. Avaliando linguagem receptiva via teste de vocabulário por imagens *peabody*: versão tradicional *versus* computadorizada. **Psicologia: Teoria e Prática**, São Paulo, v. 8, n. 2, p. 40-5, 2006.

REYNOLDS, C. R.; LIVINGSTON, R. B.; WILSON, V. **Introduction to educational assessment.** In *Measurement an assessment in Education*, ed. 2. New Jersey: Person, 2009.

SPEARMAN, C. (1927). **Las Habilidades del Hombre: su naturaleza y medición.** Buenos Aires: Paidós.

TREVISAN, B. T ET AL. Avaliação do vocabulário receptivo em crianças do ensino fundamental: Teste de Vocabulário por Imagens Peabody (TVIP). In: SENNYEY, A. L.; CAPOVILLA, F. C.; MONTIEL J. M. (Orgs). **Transtornos de aprendizagem da avaliação à reabilitação.** São Paulo: Artes Médicas, 2008. p. 361-368.