

Evidências de validade de uma prova informatizada de linguagem oral – Bilo

Maria Cristina Rodrigues Azevedo Joly
Anelise Silva Dias
Universidade São Francisco

Resumo: O presente estudo visou buscar evidências de validade de construto por meio da análise fatorial com informação plena para um instrumento informatizado que avalia compreensão da linguagem oral. Participaram 119 crianças, de 5 a 10 anos que cursavam a pré-escola, a 1ª e 2ª séries do ensino fundamental de escolas paulistas da rede pública de ensino. A aplicação da Bateria Informatizada de Linguagem Oral (Bilo) foi realizada em ambiente escolar, no laboratório de informática com no máximo 10 crianças. Os resultados obtidos por meio da análise fatorial revelaram que os itens de cada subteste indicaram boas cargas fatoriais e a unidimensionalidade de cada prova, constatando a viabilidade do instrumento na análise e compreensão da linguagem oral. Constatou-se a necessidade de retirada de alguns itens, sugerindo uma nova configuração da bateria. Identificaram-se as características dos itens por prova quanto aos seus níveis de dificuldade.

Palavras-chave: análise fatorial; validade; linguagem oral; avaliação psicológica; psicometria.

EVIDENCES OF VALIDITY OF COMPUTERIZED INSTRUMENT OF ORAL LANGUAGE – BILO

Abstract: The present study aimed to search for evidence of validity using the full information factorial analysis for a computerized instrument that evaluates oral language comprehension. 119 children took part in the study, from five to ten year-olds that were going to kindergarten, K1 and K2 from the public schools system. The application of the Computerized Battery of Oral Language (Bilo) was realized in a school environment, in the computer lab with a maximum of 10 children. The results obtained with the factorial analysis revealed that the items of each subtest indicate not only good factorial loads but also the unidimensionality of each test, noting the viability of the instrument in the analysis of oral language comprehension. It was noticed the necessity of the withdrawn of some items, suggesting a new configuration to the battery. It was identified the characteristics of the items per test as to their difficulties levels.

Keywords: factorial analysis; validity; oral language; psychological evaluation; psychometric.

EVIDENCIAS DE VALIDAD DE UNA PRUEBA INFORMATIZADA DE LENGUAJE ORAL – BILO

Resumen: El presente estudio tuvo como objetivo buscar evidencias de validez de construto por medio de análisis factorial con información completa para un instrumento informatizado que evalúa comprensión de lenguaje oral. Participaron 119 niños, de cinco a diez años que cursaban 1ª y 2ª serie de enseñanza primaria de escuelas paulistas de la red pública de enseñanza. La aplicación de la Bateria Informatizada de Lenguaje Oral (Bilo) fue realizada en ambiente escolar, en el laboratorio de informática con lo máximo 10 niños. Los resultados obtenidos por medio del análisis factorial revelaron que los ítems de cada subteste indicaron buenas cargas factoriales bien como la unidimensionalidad de cada prueba, constatando la viabilidad del instrumento en el análisis y comprensión del lenguaje oral. Se constató la necesidad de retirada de algunos ítems, sugiriendo una nueva configuración de la batería. Se identificaron las características de los ítems por prueba en relación a los niveles de dificultad.

Palabras clave: análisis factorial; validez; lenguaje oral; evaluación psicológica; psicometría.

Introdução

A avaliação psicológica, como área de especialidade, ocupa-se dos procedimentos relativos à construção de testes. Isso porque é necessário buscar as qualidades técnicas dos instrumentos utilizados, de forma que seja possível assegurar tanto sua adequação a seus objetivos específicos quanto uma utilização congruente às características daqueles que serão avaliados. Assim, tanto a fidedignidade como a validade são essenciais a qualquer tipo de processo avaliativo, especialmente em se tratando de testes psicológicos e educacionais (ANASTASI; URBINA, 2000).

A fidedignidade ou precisão de um instrumento indica o grau de consistência ou concordância entre dois conjuntos de escores independentes. Para a obtenção da precisão, utilizam-se diferentes métodos, como teste-reteste, forma-alternada, método das metades (*split-half*), coeficiente Kuder-Richardson e coeficiente alfa de Cronbach. O método utilizado neste estudo é o Kuder-Richardson que se dá pela análise de cada item individual do teste, aplicável somente quando a resposta ao item é dicotômica (ANASTASI; URBINA, 2000; PASQUALI, 2003). Para a obtenção da fidedignidade, considera dois fatores: o número de itens do teste e a razão entre a variabilidade no desempenho dos avaliados em todos os itens e a variância total nos escores do teste (URBINA, 2007).

No tocante à validade, pode ser compreendida como a eficácia da mensuração do construto (ou traço latente) que o teste se propõe a medir ou ainda como uma verificação direta da possibilidade de o instrumento satisfazer o seu objetivo (PASQUALI, 2003). Assim, os itens de um teste são elaborados visando representar diretamente o construto de interesse que está relacionado a um modelo teórico. Para tanto, há técnicas clássicas de construção de testes que se baseiam na análise do resultado global do instrumento, como a análise fatorial. Há técnicas voltadas para avaliar parâmetros de cada item em razão das características individuais do avaliado, como a Teoria de Resposta ao Item (TRI). Essas se constituem em procedimentos psicométricos que viabilizam a aferição da validade nos instrumentos de caráter psicológico e educacional (URBINA, 2007).

A validade de construto firma de maneira direta a verificação de hipóteses e legitimidade da representação de um construto utilizando-se modelos psicométricos (PASQUALI, 2001; URBINA, 2007). Uma das análises que podem ser realizadas para a obtenção de evidências de validade de construto é a análise fatorial de informação plena, utilizada no presente estudo (BOCK; AITIKIN, 1981; BOCK, GIBBONS; MURAKI, 1988).

Esse método baseia-se na análise de cada item considerado não apenas pelo pareamento da frequência obtida em função das respostas – certas e erradas –, mas também pela informação hierarquizada de frequências em amostras de escores de itens dicotômicos pelo método da máxima verossimilhança marginal (cf. BOCK; AITIKIN, 1981). Para tanto, usa-se uma matriz de correlação tetracórica para determinar a dimensionalidade do teste e a correlação ponto bisserial para avaliar o poder de discriminação dos itens, sendo classificados como muito bons (índices maiores que 0,40), bons (índices entre 0,30 e 0,40), médios (índices entre 0,20 e 0,30) e fracos (menores do que 0,20) (BOCK; AITIKIN, 1981; BOCK, MURAKI; ENGELHARD, 1985; BOCK; GIBBONS; MURAKI, 1988).

Os resultados da análise fatorial com informação plena possibilitam a análise das características dos parâmetros de cada um de seus respectivos itens. Os parâmetros analisados são a *média do total de acertos* dos participantes que responderam de maneira correta a um determinado item e *índice de facilidade*, aferido pela proporção de participantes que responderam corretamente ao item. Os itens podem ser classificados em muito fáceis (valores acima de 80%), fáceis (valores entre 60% e 80%), médios (valores entre 40% e 60%), difíceis (valores entre 20% e 40%) e muito difíceis (valores abaixo de 20%), de acordo com Wilson, Wood e Gibbons (1998) e Pasquali (2003).

Cabe destacar que a análise fatorial com informação plena, por resultar em informações completas dos parâmetros do item, é muito recomendada para buscar evidências de validade de construto em estudos iniciais de instrumentos psicológicos e educacionais (MUTHEN, 1978). Considerando, pois, o exposto, o presente estudo visou buscar evidências de validade para um instrumento que avalia a linguagem oral, dada a escassez desse tipo de prova para avaliação psicoeducacional.

Cabe, pois, considerar aspectos relativos à avaliação da linguagem. De acordo com Acosta et al. (2003), essa avaliação se baseia na verificação das bases anatômicas e funcionais (audição e fonação), nas dimensões da linguagem (forma, conteúdo e uso) e nos processos de linguagem (compreensão e produção em consonância com o desenvolvimento cognitivo do indivíduo). Cabe destaque, no presente estudo, à avaliação da linguagem, no que se refere à sua forma, especificamente à *morfologia e sintaxe* (ou seja, a análise de como a criança constrói as palavras por meio da combinação de unidades e como estrutura as frases, relacionando-as aos seus componentes) e *semântica* (vocabulário). Em relação à avaliação dos processos de linguagem, evidencia-se a *compreensão* do que a criança é capaz de entender ao ouvir uma informação ou instrução.

Ao lado disso, a avaliação da compreensão da linguagem oral na criança pequena deve levar em conta, principalmente, suas características cognitivas, seu momento de desenvolvimento e seu processo de aprendizagem. Isso porque *compreender* adquire diferentes significados em função dessas características, bem como do tipo de informação a ser compreendida (símbolos, objetos, instruções, discursos). A compreensão pode ser definida como uma habilidade relacionada àquilo que a criança se lembra da comunicação feita ou da aplicação das informações obtidas do que ouviu, ou ainda ligada às relações causais estabelecidas entre os elementos presentes na comunicação a fim de lhes atribuir coerência. Dessa forma, a compreensão deve ser considerada um conjunto de habilidades relacionadas, requerendo especial atenção a esse aspecto do construto quando de sua avaliação (PARIS; STHAL, 2005).

Assim, segundo Acosta et al. (2003), algumas tarefas para avaliar a compreensão da linguagem apoiam-se na análise das produções gestuais ou gráficas ante estímulos visuais e/ou verbais (por exemplo, a indicação do desenho pertinente à frase dada, em que se pede à criança que indique o desenho que se ajuste à ordem verbal dada; evocação de conceitos mediante imagens; na leitura de imagens ou cenas de uma história; e na tarefa de organização lógico-verbal de cenas de uma história; completar frases). Complemen-

tando, Klein (2004) sugere a investigação do conhecimento de sequências automáticas (datas, meses do ano, dias da semana), da organização de sequência lógica de figuras que representam uma história (quadros com figuras) e de avaliação de vocabulário receptivo e expressivo.

Invariavelmente, a despeito de quais instrumentos ou procedimentos sejam adotados, Broek et al. (2005) consideram que a compreensão implica a interpretação da informação, o uso do conhecimento prévio e a construção de uma representação mental sobre a essência da informação. Nesse sentido, há de se cuidar para que o instrumento de avaliação defina uma resposta adequada que seja sinal inequívoco de que a locução foi compreendida; concretizar o tipo de tarefa a ser apresentada à criança, como apontar a imagem apropriada (diante de duas ou três) ante uma frase-estímulo; especificar os estímulos linguísticos de modo que sejam inteligíveis para a criança, empregando-se itens mais complexos para avaliar aspectos predeterminados, utilizando pares de frases para avaliar o mesmo problema, e as frases devem ser breves e com a mesma extensão (ACOSTA et al., 2003).

Em síntese, existem vários procedimentos e técnicas para avaliar a linguagem oral no que se refere à compreensão ou à evocação. A literatura não referencia testes padronizados para aferir a compreensão da linguagem oral nem das habilidades visuoespaciais, sequenciais e aritméticas (CAPOVILLA, CAPOVILLA; SOARES, 2004; CAPOVILLA, JOLY; TONELOTTO, 2006), a despeito de sua relevância para a aprendizagem durante o período de escolarização e especificamente da linguagem escrita (MINISTÉRIO DA EDUCAÇÃO, 1997).

Considerando os aspectos citados e a relevância educacional e psicológica de uma identificação precoce de habilidades deficitárias de linguagem oral com critérios de investigação válidos e fidedignos (SCHELINI; GOMES; WESCHLER, 2006), o presente estudo visou buscar evidências de validade de construto para uma bateria que avalia linguagem oral em crianças que iniciam a escolarização formal.

Método

Participantes

Participaram deste estudo 119 crianças de ambos os sexos, sendo 47,9% do sexo feminino. As idades variaram de 5 a 10 anos ($M = 6,90$; $DP = 1,31$), e as crianças estavam regularmente matriculados na última etapa da educação infantil ($N = 47$), 1ª série ($N = 37$) e 2ª série ($N = 35$) do ensino fundamental de duas escolas públicas do interior do Estado de São Paulo.

Material

Bateria Informatizada de Linguagem Oral (Bilo – versão I)

A Bilo (JOLY, 2006) foi construída com base na revisão da literatura (AMERICAN SPEECH-LANGUAGE-HEARING ASSOCIATION, 2007; ELLIS, 1995; KLEIN, 2004; PARIS; STHAL, 2005;

TELLEZ, 2004; BROEK et al., 2005) e na análise de testes equivalentes publicados no exterior (DRP - Degrees of Reading Power -, International Dyslexia Test). Tem por objetivo avaliar a compreensão da linguagem oral em alunos da educação infantil à segunda série do ensino fundamental (ciclo I). A Bilo foi desenvolvida por meio do *software* Run Revolution que tem recursos de multimídia e interface com banco de dados (MySQL) para armazenamento das respostas dadas. Gera um aplicativo que é instalado em cada equipamento.

A Bilo conta com sete provas referentes à avaliação morfofossintática, sequência lógica e compreensão de frases, parágrafos e histórias. As provas relativas à compreensão da linguagem, considerando sua estrutura em frases, parágrafos e histórias, foram elaboradas por meio do Sistema Orientado de Cloze (SOC) (JOLY, 2007). Esse sistema 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 compreensão diferenciados. Os critérios definidos para a prova Completar Histórias da Bilo foram omissões de um substantivo por frase.

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. Estes foram selecionados considerando-se símbolos, significados e contextos familiares à criança.

Cada tela tem apenas um item da prova; após a escolha da opção, o respondente é solicitado a confirmar sua resposta para continuar a prova. Em um banco de dados, registram-se a opção de resposta, o tempo total por prova e o tempo total na bateria de todos os respondentes. A aplicação 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. A Bilo não possui características psicométricas, pois esse é o objetivo da investigação relatada no presente artigo.

Os critérios de correção são por item, por prova e geral. Todos os itens da bateria são de múltipla escolha com três opções de resposta. Duas opções referem-se a conteúdos pertencentes à mesma classe gramatical ou categoria, sendo um deles a resposta correta, e a terceira opção é de outra classe gramatical ou categoria diferente da opção correta. Por exemplo, se a palavra omitida no texto for gato, as opções serão: gato, cão e alegre. O somatório dos pontos por item determina o escore por prova, e o escore máximo obtido é de 20 pontos. O somatório dos escores por prova resulta no escore total.

As instruções gerais e específicas de cada prova e item são apresentadas oralmente pelo programa e estão também impressas usando letra bastão. Há um tutorial interativo indicando quais os recursos de interface do computador são utilizados para responder à Bilo. Nesse tutorial, é solicitado que o respondente complete alguns itens visando exercitar tais habilidades necessárias. As provas, Interpretação de História (IH), Completar Parágrafos (CP) e Completar Histórias (CH), são precedidas por uma apresentação de um vídeo com áudio e imagem sobre o tema e a sequência das informações da história.

Descrição das provas

Prova 1: Compreensão Morfossintática (CM)

Tem por objetivo avaliar a relação entre um vocábulo e a sua representação gráfica, revelando a compreensão do significado. Esta prova é constituída por 12 itens, sendo os dois primeiros de exemplo. Cada item é formado por duas orações com um termo integrante que pode ser uma logomarca, um substantivo, adjetivo, verbo ou advérbio, a ser compreendido pelo respondente. A tarefa do respondente é escolher, dentre as três opções, a que revela a relação correta entre desenho e nome. São atribuídos dois pontos para a resposta correta, um ponto para a opção de mesma categoria da correta e nenhum ponto para a resposta errada. O escore máximo é de 20 pontos, obtido pelo somatório de pontos atribuídos aos itens.

CM EX2. SOU MUITO CARINHOSO E FOFINHO. SOU O GATO JUJU.

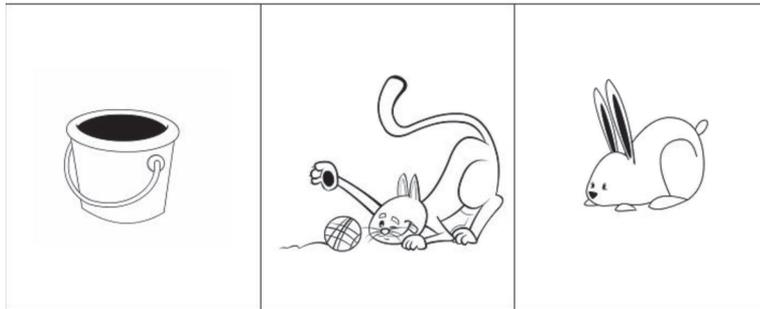


Figura 1. Exercício de CM

Prova 2: Sequência Lógica (SL)

Visa avaliar a organização lógica e temporal do respondente para estímulos visuais organizados em cenas cujo conjunto sequenciado resulta em uma história. Esta prova é constituída por 12 itens, sendo os dois primeiros de exemplo. Cada item é formado por uma série de cenas que, organizadas em uma sequência com começo, meio e fim, adquirem significado de uma história. As séries têm contextos específicos com temas familiares ao respondente. Cada cena da série é composta por uma ação, há três itens da prova com duas cenas, outros três com três cenas, dois itens com quatro cenas e os últimos dois com cinco cenas. A tarefa do respondente é organizar as cenas de cada série de modo que haja relação de causa e efeito e contiguidade entre elas. Orienta-se exclusivamente pelos estímulos visuais. Trata-se da única prova da Bilo que não agrega estímulos orais às imagens para sua execução. São atribuídos dois pontos por série sequenciada corretamente e um ponto por série com, no mínimo, 50% das cenas organizadas corretamente. O escore máximo é de 20 pontos, obtido pelo somatório de pontos atribuídos aos itens.

SL EX1. Praia

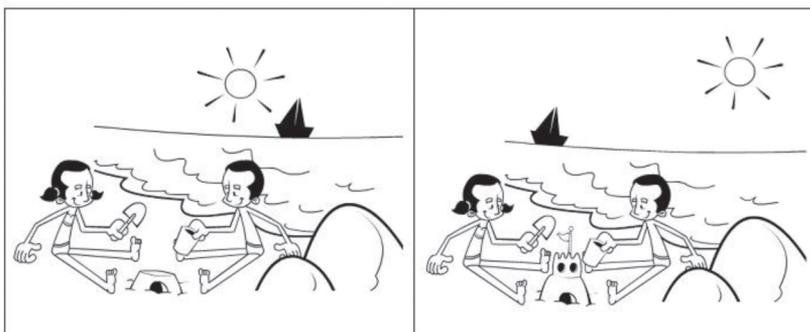


Figura 2. Exercício de SL

Prova 3: Organização Lógico-Verbal (OLV)

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. Essa prova é constituída por 10 itens. Cada item é composto por uma história e por uma série de cenas. Nesta tarefa, o respondente ouve a história e deve organizar as cenas de cada série, de acordo com a sequência lógica do conteúdo visual e verbal. O número de itens por série varia. Há três séries de duas cenas com três histórias diferentes, contendo 10 palavras em cada uma; há três séries de três cenas com três histórias, contendo 15 palavras em cada uma; há dois itens para quatro cenas com duas histórias diferentes, contendo 20 palavras em cada uma; há dois itens para cinco cenas com duas histórias diferentes, contendo 25 palavras em cada uma. São atribuídos dois pontos por série sequenciada corretamente e um ponto por série com, no mínimo, 50% das cenas organizadas corretamente. O escore máximo é de 20 pontos, obtido pelo somatório de pontos atribuídos aos itens.

OLV EX1. O MENINO GANHOU UMA PIPA. ELE EMPINOU NO PARQUE.

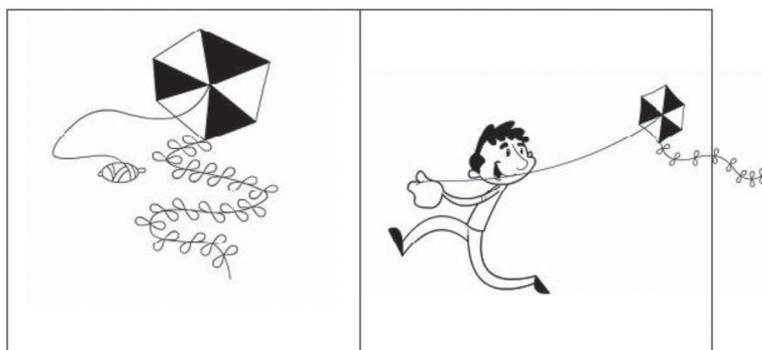


Figura 3. Exercício de OLV

Prova 4: Interpretação de História (IH)

Avalia a compreensão receptiva da história por meio de questões de múltipla escolha. Esta prova é composta por três descrições apropriadas à faixa etária dos respondentes, totalizando nove itens. A primeira história tem 25 palavras e dois itens, a segunda 50 palavras com três itens e a terceira 100 palavras com cinco itens. Todos os itens referem-se à compreensão literal da história. A tarefa do respondente é assistir ao vídeo e ouvir a história que é contada duas vezes e escolher uma opção de resposta por item. Há três opções de resposta por item. As opções contam com uma alternativa com a mesma classificação da resposta correta e a outra com uma palavra qualquer, desde que seja classificada como diferente das outras duas opções. Serão atribuídos dois pontos para a resposta correta, um ponto para a resposta com mesma classificação da opção certa e zero ponto para a errada. O escore máximo é de 18 pontos, obtido pelo somatório de pontos atribuídos aos itens.

O GATO DE MARIANA

EU TENHO UMA GATA. SEU NOME É MIMI. GANHEI MIMI DO MEU AVÔ. MIMI É MUITO BRINCALHONA. ELA É A MINHA MELHOR AMIGA. ADORO MIMI

IH1. QUAL É O BICHO DE MARIANA?

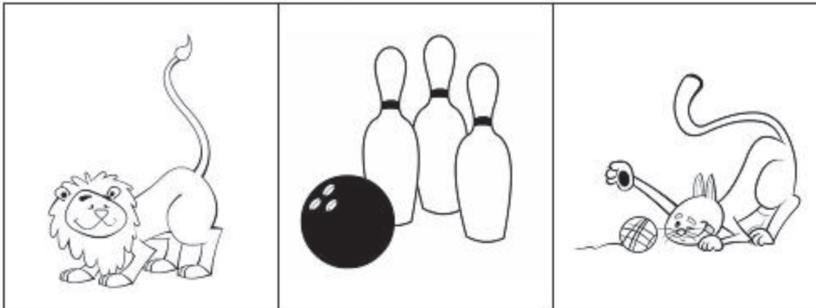


Figura 4. Exercício de IH

Prova 5: Completar Frases (CF)

Avalia a compreensão de palavras organizadas em frases. Deve-se destacar que um dos substantivos é omitido. Esta prova é constituída por dez itens. A tarefa do respondente é identificar que alternativa, dentre as opções, dá melhor sentido à frase. As opções de resposta por item contam com uma alternativa com a mesma classificação gramatical da resposta correta e a outra com uma palavra qualquer, desde que seja classificada como diferente das outras duas opções. Cada item será apresentado duas vezes ao respondente. Serão atribuídos dois pontos para a resposta correta, um ponto para a resposta com mesma classificação da opção certa e zero ponto para a errada. O escore máximo é de 20 pontos, obtido pelo somatório de pontos atribuídos aos itens.

CP2. O _____ COMEU CENOURA.



Figura 5. Exercício de CF

Prova 6: Completar Parágrafos (CP)

Avalia a compreensão de palavras organizadas em três frases que compõem um parágrafo, e o substantivo é o critério de omissão. Essa prova é constituída por quatro parágrafos com temas diferentes e conta com nove omissões no total, sendo duas para cada um dos dois primeiros itens e três para os demais. Cada omissão é considerada um item. A tarefa do respondente é identificar que alternativa, dentre as opções, dá melhor sentido à frase. As opções de resposta por item são três e contam com uma alternativa com a mesma classificação gramatical da resposta correta e a outra com uma palavra qualquer, desde que seja classificada como diferente das outras duas opções. Serão atribuídos dois pontos para a resposta correta, um ponto para a resposta com mesma classificação da opção certa e zero ponto para a errada. O score máximo é de 20 pontos, obtido pelo somatório de pontos atribuídos aos itens.

O VENTO

CP1. O VENTO LEVANTA A _____ DE DONA SÔNIA.



CP2. BATE A _____ COM FORÇA.



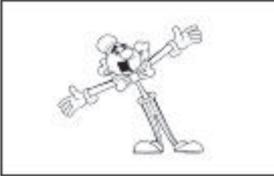
Figura 6. Exercício de CP

Prova 7: Completar Histórias (CH)

Avalia a compreensão oral de histórias contando com 13 itens. A primeira história tem 50 vocábulos, e a segunda, 100. É eliminado um substantivo a cada dez vocábulos. A história um possui quatro omissões, e a história dois, seis. As opções de resposta por item contam com uma alternativa com a mesma classificação da resposta correta e a outra com uma palavra qualquer, desde que seja classificada como diferente das outras duas opções. Cada história completa é, primeiro, apresentada em forma de vídeo (imagens com áudio) para, em seguida, ser apresentada na tela na forma impressa, acompanhada do respectivo áudio com as opções de resposta (figuras) por item. A tarefa do respondente é assistir ao vídeo e ouvir a história que é contada duas vezes e escolher uma opção por item. São atribuídos dois pontos para a resposta correta, um ponto para a resposta com mesma classificação da opção certa e zero ponto para a errada, por item. O escore máximo total é de 26 pontos.

O CIRCO

CH1. O CIRCO É MUITO DIVERTIDO. TEM A S BRINCADEIRAS E CAMBALHOTAS DO _____.

		
-----------------------------------------------------------------------------------	-----------------------------------------------------------------------------------	------------------------------------------------------------------------------------

CH2. TEM O MÁGICO QUE FAZ TRUQUES COM A S CARTAS DE _____.

		
-------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------

CH3. OS BICHOS SÃO AINDA MAIS ENGRAÇADOS. O _____ ENCHE A TROMBA DE ÁGUA.

		
-------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------

COM UM JATO MOLHA A JUBA DO FERAZ LEÃO. ELE FICA MUITO BRAVO.

Figura 7. Exercício de CH

Equipamentos

Foram utilizados microcomputadores com capacidade de *hardware* básica (1 GB de memória; 100 GB de disco; Windows XP) para suportar os aplicativos de vídeo e imagem ne-

cessários para a Bilo e o armazenamento dos resultados em banco de dados local (MySQL). Cada participante contou com um fone de ouvido conectado ao microcomputador no qual respondeu à Bilo.

Procedimento

Após a aprovação do projeto pelo Comitê de Ética, realizou-se contato com as direções das instituições de ensino para solicitar autorização. Após, enviou-se aos pais ou responsáveis pelos alunos o Termo de Consentimento Livre e Esclarecido solicitando autorização para a participação de seus filhos na pesquisa. Participaram do estudo somente os alunos que entregaram o termo de consentimento assinado.

A aplicação do instrumento foi realizada coletivamente no laboratório de informática da escola, com no máximo dez crianças, em horário escolar, por três aplicadores treinados. Esclareceram-se os procedimentos gerais para cada criança individualmente, e, após, a criança iniciou a realização do instrumento. O tempo total médio de aplicação foi de 45 minutos.

Resultados

Considerando que a Bilo é uma bateria cujas provas avaliam construtos específicos relacionados à linguagem oral (PARIS; STHAL, 2005), para obter a evidência de validade de construto para tal instrumento de avaliação, foi realizada uma análise fatorial com informação plena para cada prova da bateria, a qual contou com o auxílio do programa Testfact (WILSON, WOOD; GIBBONS, 1998), a fim de se buscar a unidimensionalidade de cada componente do instrumento. Avaliou-se especificamente o construto de cada prova com base na análise dos parâmetros e das cargas fatoriais de cada item, de acordo com modelos de Bock e Aitikin (1981), Muraki e Engelhard (1985) e Bock, Gibbons e Muraki (1988).

Tabela 1. Parâmetros descritivos dos itens dos subtestes Compreensão Morfossintática e Sequência Lógica

Prova	Item	Média de acertos	Índice de facilidade	Correlação ponto bisserial
Compreensão Morfossintática	1	7,97	0,882	0,469
	2	6,50	0,017	-0,074
	3	8,41	0,790	0,764
	4	8,07	0,874	0,578
	5	8,38	0,790	0,733
	6	8,25	0,857	0,769
	7	8,20	0,849	0,676

(continua)

Tabela 1. Parâmetros descritivos dos itens dos subtestes Compreensão Morfossintática e Sequência Lógica (continuação)

Prova	Item	Média de acertos	Índice de facilidade	Correlação ponto bisserial
Compreensão Morfossintática	8	8,20	0,849	0,676
	9	7,85	0,908	0,349
	10	8,06	0,815	0,455
Sequência Lógica	1	6,18	0,708	0,575
	2	6,02	0,741	0,533
	3	6,45	0,652	0,605
	4	7,93	0,482	0,849
	5	8,29	0,429	0,856
	6	7,51	0,527	0,797
	7	8,29	0,277	0,611
	8	7,21	0,420	0,568
	9	8,90	0,277	0,724
	10	8,10	0,455	0,851

Conforme os dados da Tabela 1 resultantes da análise dos parâmetros dos itens pela análise fatorial plena, no que se refere à média total de acertos da prova de Compreensão Morfossintática (CM), o item CM3 (*VIVO NO MAR, COM AS ESTRELAS, PLANTAS, CONCHAS E CARAMUJOS. EU SOU UM TUBARÃO, tendo por opção de respostas as figuras de tubarão, bala e tartaruga*) foi o que apresentou maior média e, portanto, o que contou com maior frequência de acertos por parte dos respondentes. Constatou-se, também, que a maioria dos itens mostrou-se muito fáceis (CM1, CM4, CM6, CM7, CM8, CM9 e CM10), por apresentarem índices de 80%, e dois itens (CM3 e CM5) fáceis, com valores entre 60% e 80%. Há apenas um item (CM2) considerado muito difícil. Além disso, a correlação ponto bisserial indicou que, dos 10 itens, 8 são muito bons (CM1, CM3, CM4, CM5, CM6, CM7, CM8 e CM10), um item (CM9) bom e apenas um item (CM2) fraco, por possuir um índice menor do que 0,20.

Com relação à prova Sequência Lógica (SL), a *média do total de acertos* dos participantes foi superior a todas as outras médias para o item SL9, indicando que este item está entre os mais difíceis na prova, uma vez que esses participantes atingiram o melhor desempenho na prova como um todo. Com relação ao *índice de facilidade*, três itens apresentaram-se como fáceis (itens SL1, SL2 e SL3), cinco itens (SL4, SL5, SL6, SL8 e SL10) tendo facilidade média, e apenas dois itens (SL7 e SL9) considerados difíceis, por obterem índices entre 20% e 40%. A *correlação ponto bisserial* indicou que todos os 10 itens são muito bons por possuírem índices maiores que 0,40 (Tabela 1).

Quando se analisa a Tabela 2, pode-se observar que, no que concerne à *média do total de acertos* dos participantes para a prova Organização Lógico-Verbal (OLV), o item OLV10 teve média superior a todas as outras, revelando que está entre os mais difíceis na prova. O *índice de facilidade* revelou que dois itens apresentaram-se como muito fáceis (itens OLV1 e OLV2), dois itens (OLV6 e OLV9) como fáceis, e seis itens (OLV3, OLV4, OLV5, OLV7, OLV8 e OLV10) como muito difíceis, uma vez que os valores apresentados foram abaixo de 20%. A *correlação ponto bisserial* revelou que, dos 10 itens, oito são muito bons (OLV3, OLV4, OLV5, OLV6, OLV7, OLV8, OLV9 e OLV10), e apenas dois itens (OLV1 e OLV2) são classificados como fracos, uma vez que o valor dos índices apresentados foi inferior a 0,20.

Tabela 2. Parâmetros descritivos dos itens dos subtestes Organização Lógico-Verbal e Interpretação de História

Prova	Item	Média de acertos	Índice de facilidade	Correlação ponto bisserial
Organização Lógico-Verbal	1	5,21	0,864	0,089
	2	5,28	0,932	0,227
	3	7,38	0,449	0,698
	4	7,69	0,441	0,781
	5	7,62	0,457	0,769
	6	7,23	0,552	0,784
	7	7,91	0,374	0,729
	8	7,84	0,371	0,700
	9	7,46	0,513	0,795
	10	8,65	0,200	0,590
Interpretação de História	1	4,17	0,878	0,572
	2	4,28	0,843	0,694
	3	0,00	0,000	0,000
	4	4,33	0,783	0,649
	5	4,56	0,078	0,150
	6	0,00	0,000	0,000
	7	4,63	0,522	0,588
	8	0,00	0,000	0,000
	9	4,26	0,814	0,582

No que se refere à prova de Interpretação de Histórias (IH), o item IH7 foi o mais difícil, de acordo com a *média total de acertos*. Com relação ao *índice de facilidade*, a maioria dos itens apresentou-se como muito fácil (itens IH1, IH2, e IH9), com valores acima de 80%, um item (IH4) como fácil, um item (IH7) apresentou um índice de facilidade média, e quatro itens muito difíceis (IH3, IH5, IH6, e IH8), abaixo do valor crítico de 20%. A *correlação ponto bisserial* indicou que, dos 10 itens, cinco são muito bons (IH1, IH2, IH4, IH7 e IH9), e quatro itens (IH3, IH5, IH6 e IH8) fracos por apresentarem índices inferiores a 0,20 (Tabela 2).

Tabela 3. Parâmetros descritivos dos itens dos subtestes Completar Frases e Completar Parágrafos

Prova	Item	Média de acertos	Índice de facilidade	Correlação ponto bisserial
Completar Frases	1	7,50	0,752	0,726
	2	7,11	0,906	0,624
	3	5,33	0,026	-0,128
	4	7,34	0,778	0,613
	5	7,07	0,915	0,571
	6	7,50	0,752	0,726
	7	7,11	0,906	0,624
	8	5,33	0,026	-0,128
	9	7,34	0,778	0,613
	10	7,07	0,915	0,571
Completar Parágrafos	1	6,58	0,974	0,151
	2	6,66	0,948	0,459
	3	6,71	0,931	0,582
	4	7,78	0,078	0,367
	5	6,56	1,000	0,000
	6	0,00	0,000	0,000
	7	6,94	0,741	0,671
	8	6,68	0,957	0,564
	9	6,74	0,931	0,689

Quando se analisa a Tabela 3, pode-se observar que para a prova Completar Frases (CF), no que concerne à *média total de acertos*, o item CF6 foi o que obteve média superior, revelando que está entre os mais difíceis na prova. Quatro itens apresentaram-se

como muito fáceis (itens CF2, CF5, CF7 e CF10), que obtiveram valores acima de 80%, quatro itens (CF1, CF4, CF6 e CF9) como fáceis, e apenas dois itens (CF3 e CF8) como muito difíceis, uma vez que o valor dos índices apresentou-se entre 20% e 40%. A *correlação ponto bisserial* revelou que, dos 10 itens, oito são muito bons (CF1, CF2, CF4, CF5, CF6, CF7, CF9 e CF10), e apenas dois itens (CF3 e CF8) foram fracos por obterem valores inferiores a 0,20.

Identificou-se pela *média do total de acertos* dos participantes que o item CP4, da prova Completar Parágrafos, apresentou média superior a todas as outras, indicando que está entre os mais difíceis na prova. Com relação ao *índice de facilidade*, alguns itens apresentaram-se como muito fáceis (itens CP1, CP2, CP3, CP5, CP8 e CP9), apenas um item (CP7) fácil e dois itens (CP4 e CP6) como muito difíceis, com valores abaixo de 20%. Vale ressaltar que, nos itens CP5 e CP6, não houve variabilidade nas respostas, uma vez que todas as crianças acertaram o item CP5 e erraram o item CP6. A *correlação ponto bisserial* indicou que, dos 10 itens, nove são muito bons (CP2, CP3, CP7, CP8, e CP9), apresentando valores superiores a 0,40, um item (CP4) bom, com valor entre 0,30 e 0,40 e três itens fracos (CP1, CP5, e CP6), por possuírem valores abaixo de 0,20.

Conforme a média total de acertos, a prova Completar Histórias (CH) teve o item CH7 como um dos mais difíceis da prova. Outros dois itens (CH3 e CH5) foram identificados também como muito difíceis. Com relação ao *índice de facilidade*, a maioria dos itens apresentou-se como muito fácil (itens CH1, CH2, CH4, CH6, CH7, CH8, CH9, CH10, CH11, CH12 e CH13). A *correlação ponto bisserial* indicou que, dos 13 itens, nove são muito bons (CH4, CH6, CH7, CH8, CH9, CH10, CH11, CH12 e CH13), um item (CH1) bom, dois itens médios (CH2 e CH3) e apenas um item (CH5) fraco.

Tabela 4. Parâmetros descritivos dos itens do subteste Completar Histórias (CH)

Item	Média de acertos	Índice de facilidade	Correlação ponto bisserial
CH1	10,05	0,813	0,378
CH2	9,82	0,925	0,205
CH3	8,71	0,131	-0,209
CH4	10,04	0,897	0,528
CH5	0,00	0,000	0,000
CH6	10,17	0,849	0,536
CH7	10,34	0,811	0,683
CH8	10,30	0,849	0,728
CH9	10,08	0,915	0,575
CH10	9,97	0,971	0,439
CH11	10,07	0,913	0,450
CH12	9,96	0,971	0,404
CH13	10,14	0,835	0,402

Em seguida, foi realizada a análise fatorial com informação completa por meio da rotação Promax para cada uma das provas da Bilo. Esse procedimento permite identificar a contribuição que cada item fornece para a prova como um todo. Os resultados evidenciaram que, considerando o modelo unidimensional, a maior variância total que pode ser explicada por esse modelo foi de 70,1% na prova Completar Frases (CF). Com base na análise feita, percebeu-se que todos os itens da CF contribuíram para a formação de apenas um fator. Cabe destacar que nenhum item apresentou carga fatorial menor do que 0,60.

Na prova Sequência Lógica (SL), 59,2% da variância total pode ser explicada pelo modelo de um fator, já que os itens não apresentaram cargas fatoriais abaixo de 0,50. Com relação à prova Organização Lógico-Verbal (OLV), os resultados revelaram que 53,4% da variância total pode ser explicada pelo modelo de um fator, já que apenas dois itens não apresentaram cargas fatoriais satisfatórias, abaixo de 0,30; os demais revelaram cargas acima de 0,70. Compreensão Morfossintática (CM) também foi considerada predominantemente unidimensional, sendo 44,4% da variância total explicada por esse modelo. Observou-se também que todos os itens contribuíram para a formação de apenas um fator, de modo que nenhum item apresentou carga fatorial menor do que 0,30.

Contudo, para as provas Interpretação de Histórias (IH), Completar Histórias (CH) e Completar Parágrafos (CP), houve a necessidade de retirada de alguns itens, para a verificação da unidimensionalidade de cada uma delas. Eles foram retirados porque não contribuíam para aferir o construto específico das provas, pois eram itens que todas as crianças acertaram ou que todas elas erraram. No que se refere ao IH, retiraram-se três itens, a saber, IH3, IH6 e IH8, sendo que esses não tiveram qualquer acerto. Após a retirada dos itens da prova IH constatou-se que 38,4% da variância total pode ser explicada pelo modelo de um fator. Destaca-se que, a resposta para IH3 implicava inferência; o IH6 possuía imagens que não eram fiéis à resposta literal esperada e o IH8 contava com um vocabulário tanto do item quanto das opções de resposta pouco frequente para a amostra investigada. Para essa segunda análise, apenas um item (IH5) não apresentou carga fatorial satisfatória, abaixo de 0,30, e os demais revelaram cargas acima de 0,30. Esse item exigia que o respondente deduzisse a resposta considerando o estado emocional do personagem da história.

Para os itens da prova Completar Histórias (CH), dos que compunham a prova, foi retirado apenas o CH5, porque não apresentou nenhum acerto, revelando problema com os desenhos que compunham as opções de resposta, os quais não revelavam a interpretação solicitada pelo item. Com a retirada de CH5, o instrumento apresentou 48,9% da variância total, explicada pelo modelo de um fator. Na segunda análise realizada após a retirada do CH5, apenas o CH2 não apresentou carga fatorial satisfatória (abaixo de 0,30). Identificou-se que o vocabulário utilizado era pouco comum para a amostra investigada. Os demais itens revelaram cargas acima de 0,40, consideradas adequadas à prova.

Já na prova Completar Parágrafos (CP), houve a necessidade de retirada de dois itens (CP5 e CP6). A retirada pode ser explicada no item CP5 pela dificuldade de os respondentes ser em função das imagens e do vocabulário por elas representado nas opções de resposta e para o CP6 cuja resposta exigia inferência. Nessa segunda retirada de itens, mais dois (CP1 e CP4) não apresentaram cargas fatoriais satisfatórias, abaixo de 0,30, tan-

to o CP1 quanto o CP4 que implicavam respostas por inferência; os demais itens revelaram cargas acima de 0,50. Na ocasião final, CP revelou que 40,1% da variância total pode ser explicada pelo modelo de um fator.

Discussão

A validação da Bilo pela análise fatorial com informação plena permitiu identificar os diferentes níveis de dificuldade dos itens por prova, bem como verificar que são aferidas habilidades específicas de linguagem em cada uma das provas que se mostraram unidimensionais (WILSON; WOOD; GIBBONS, 1998; PASQUALI, 2001; URBINA, 2007). Esses resultados revelam que a identificação de características psicométricas para instrumentos de avaliação psicoeducacional, como proposto por Schelini, Gomes e Weschler (2006), possibilita uma melhor análise do desenvolvimento e aprendizagem da linguagem oral para crianças como um conjunto de habilidades relacionadas (PARIS; STHAL, 2005).

A análise fatorial com informação plena revelou que os itens das provas atendem aos critérios de avaliação da linguagem oral propostos por Acosta et al. (2003) e Broek et al. (2005). Assim, de um lado, podem-se considerar a organização lógico-verbal e o vocabulário receptivo e expressivo e, de outro, a compreensão de instruções e a representação da linguagem em símbolos gráficos com interpretação de elementos sintáticos e semânticos como habilidades que antecedem a aquisição da linguagem escrita. Essas habilidades são relevantes para a aquisição e desenvolvimento da linguagem, como afirmam Capovilla, Capovilla e Soares (2004) e Capovilla, Joly e Tonelotto (2006).

Considerando os resultados da análise de parâmetros dos itens e o seu poder de discriminação, faz-se necessário rever a estrutura das provas IH, CP e CH. Propõe-se que para IH sejam alteradas as imagens dos itens IH3, IH5, IH6 e IH8, a fim de possibilitar uma melhor compreensão das opções de resposta por parte do respondente. Sugere-se que CP e CH sejam reduzidas a uma única prova, uma vez que CP não avaliou efetivamente um trecho de uma história como se propunha. Assim, os itens CP7, CP8 e CP9, CH1, CH2, CH3, CH4, CH5, CH6, CH7, CH8, CH9, CH10 E CH11 passam a compor a prova Completar Histórias para a versão 3 da Bilo.

Cabe destacar a organização informatizada do instrumento, contemplando recursos multimídias e possibilitando uma interatividade independente do respondente com equipamento, além de a pessoa responder ao teste respeitando seu ritmo de trabalho. Os itens e os recursos multimídias utilizados mostraram-se adequados e válidos para aferir as habilidades pretendidas em cada prova e na bateria como um todo. Nesse sentido, vale reforçar a necessidade de se realizem avaliações referentes às qualidades técnicas em termos de características, como validade e fidedignidade, para assegurar a adequação de um instrumento para seus objetivos específicos, além da utilização congruente com o tipo de pessoa a ser avaliada (ANASTASI; URBINA, 2000).

Considerações finais

A despeito de este estudo exploratório revelar evidências de validade de construto para a bateria, outras pesquisas serão necessárias para ampliar as análises, conferindo à

BILO (versão 3) mais evidências de validade que permitam ampliar sua utilização e as generalizações acerca dos resultados obtidos por meio de sua aplicação.

Ampliar estudos para outros estados brasileiros, com a participação de estudantes do ensino fundamental e infantil de escolas particulares e públicas (municipais e estaduais), contribuirá não somente para obtenção de evidências de validade para contextos específicos e para a elaboração de normas brasileiras acerca da avaliação da compreensão da linguagem oral, mas também viabilizará uma aferição precoce de dificuldades específicas da linguagem, tanto oral como escrita, bastante comuns nas etapas iniciais da escolarização formal (CAPOVILLA; JOLY; TONELOTTO, 2006). Assim, constata-se um avanço para a avaliação da linguagem oral, uma vez que há poucos estudos dessa natureza na área.

Referências

- ACOSTA, V. M. et al. **Avaliação da linguagem**. São Paulo: Livraria Santos, 2003.
- AMERICAN SPEECH-LANGUAGE-HEARING ASSOCIATION (ASHA). **What is language? What is speech?** 2007. Disponível em: http://www.asha.org/public/speech/development/language_speech.htm. Acesso em: abr. 2007.
- ANASTASI, A.; URBINA, S. **Testagem psicológica**. 7. ed. Porto Alegre: Artmed, 2000.
- BOCK, R. D.; AITIKIN, M. Marginal maximum likelihood estimation of item parameters: application of an EM algorithm. **Psychometrika**, New York, v. 35, n. 4, p. 443-459, 1981.
- BOCK, R. D.; GIBBONS, R. D.; MURAKI, E. Full-information factor analysis. **Applied Psychological Measurement**, Thousand Oaks, v. 12, n. 3, p. 261-280, 1988.
- BROEK, P. van den et al. (Org.). **Children's reading comprehension and assessment**. Mahwah: LEA, 2005.
- CAPOVILLA, G. S.; CAPOVILLA, F. C.; SOARES, J. V. T. Consciência sintática no ensino fundamental, correlações com consciência fonológica, vocabulário, leitura e escrita. **Psico-USF**, São Paulo, v. 9, n. 1, p. 39-47, 2004.
- CAPOVILLA, A. S. G.; JOLY, M. C. R. A.; TONELLOTO, J. M. F. Avaliação neuropsicológica e aprendizagem. In: NORONHA, A. P. P.; SANTOS, A. A. A.; SISTO, F. F. (Org.). **Facetas do fazer em avaliação psicológica**. São Paulo: Vetor, 2006. p. 141-162.
- ELLIS, A. W. **Leitura escrita e dislexia: uma análise cognitiva**. 2. ed. Porto Alegre: Artmed, 1995.
- JOLY, M. C. R. A. **Bateria Informatizada de Linguagem Oral**. Teste em desenvolvimento. Itatiba: Universidade São Francisco, 2006.
- _____. The validity of Cloze Oriented System (COS): a correlation study with an electronic comprehension test and a reading attitude survey. **Psicologia Escolar e Educacional**, São Paulo, v. 11, p. 49-59, 2007.
- KLEIN, H. Avaliação das dificuldades de linguagem em crianças e em adolescentes. In: SNOWLING, M.; STACKHOUSE, J. (Org.). **Dislexia, fala e linguagem**. Tradução M. F. Lopes. Porto Alegre: Artmed, 2004.

MINISTÉRIO DA EDUCAÇÃO (MEC). **Parâmetros Curriculares Nacionais – Língua Portuguesa**. Brasília: Gráfica do Governo, 1997.

MURAKI, E.; ENGELHARD, G. Full information item factor analysis: applications of EAP scores. **Applied Psychological Measurement**, Thousand Oaks, v. 9, n. 4, p. 417-430, 1985.

MUTHEN, B. Contributions to factor analysis of dichotomous variables. **Psychometrika**, New York, v. 43, n. 4, p. 551-560, 1978.

PARIS, S. G.; STAHL, S. A. **Children's reading comprehension and assessment**. Mahwah: LEA, 2005.

PASQUALI, L. **Técnicas de exame psicológico**. Brasília: Editora UnB, 2001.

_____. **Psicometria: teoria dos testes na psicologia e na educação**. Petrópolis: Vozes, 2003.

SCHELINI, P. W.; GOMES, V. L. T.; WESCHLER, S. M. Avaliação psicológica infantil: aspectos cognitivos e neuropsicológicos. In: NORONHA, A. P. P.; SANTOS, A. A. A.; SISTO, F. F. (Org.). **Facetas do fazer em avaliação psicológica**. São Paulo: Vetor, 2006. p. 81-94.

TELLEZ, J. A. **La comprensión de los textos escritos y la psicología cognitiva**. Madrid: Dykinson, 2004.

URBINA, S. **Fundamentos da testagem psicológica**. Porto Alegre: Artmed, 2007.

WILSON, D. T.; WOOD, R.; GIBBONS, R. **Testfact: test scoring and item factor analysis**. Lincolnwood: Scientific Software Internacional Inc., 1998.

Contato

Maria Cristina Rodrigues Azevedo Joly
Rua Alexandre Rodrigues Barbosa, 45
Centro – Itatiba – SP
CEP 13251-900
e-mail: cristina@e-teste.com.br

Tramitação

Recebido em setembro de 2008

Aceito em agosto de 2009