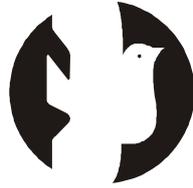


ANELISE SILVA DIAS



UNIVERSIDADE
SÃO FRANCISCO

EVIDÊNCIAS DE VALIDADE DE UMA PROVA DE
COMPREENSÃO EM LEITURA EM ESTATÍSTICA

ITATIBA
2008

ANELISE SILVA DIAS

EVIDÊNCIAS DE VALIDADE DE UMA PROVA DE
COMPREENSÃO EM LEITURA EM ESTATÍSTICA

Dissertação apresentada ao Programa de Pós-
Graduação *Stricto Sensu* em Psicologia da
Universidade São Francisco para obtenção do título
de Mestre.

ORIENTADOR(A): CLAUDETTE MARIA MEDEIROS VENDRAMINI

ITATIBA
2008

UNIVERSIDADE SÃO FRANCISCO
PROGRAMA DE PÓS-GRADUAÇÃO STRICTO SENSU EM PSICOLOGIA
MESTRADO

EVIDÊNCIAS DE VALIDADE DE UMA PROVA DE
COMPREENSÃO EM LEITURA EM ESTATÍSTICA

Autor(a): Anelise Silva Dias

Orientador(a): Claudette Maria Medeiros Vendramini

Este exemplar corresponde à redação final da dissertação de mestrado defendida por Anelise Silva Dias e aprovada pela comissão examinadora.

Itatiba, 27 de fevereiro de 2008.

COMISSÃO EXAMINADORA

Claudette Maria Medeiros Vendramini

Irene Mauricio Cazorla

Maria Cristina Rodrigues Azevedo Joly

ITATIBA
2008

Agradecimentos

Primeiramente à Deus, pela oportunidade de viver, pelas oportunidades de aprendizado e conhecimento contínuo, e pela força por superar todos os obstáculos.

A seguir, as mais importantes entre todas as importantes. A minha família, José Roberto, Maria Luiza, Andreza, Adler e Douglas, que sempre estiveram presentes em minha vida, pelos valiosos conselhos nos momentos angustiantes e difíceis, e pelas vibrações em minhas conquistas e alegrias. Também por sempre contribuírem para o meu eterno aprendizado, sendo grandes incentivadores do meu desenvolvimento e crescimento pessoal e profissional. A quem eu devo uma enorme parte deste trabalho.

À todos os professores do Programa que contribuíram com os seus preciosos ensinamentos, em especial as professoras Dr^a. Claudette Maria Medeiros Vendramini e Dr^a. Maria Cristina Rodrigues Azevedo Joly, por terem apostado em mim, que me ensinaram, principalmente no valor e importância do meu trabalho, e na concretização do meu ideal em fazer pesquisa.

Às professoras Dr^a. Irene Maurício Cazorla e Dr^a. Maria Cristina Rodrigues Azevedo Joly pelas valiosas contribuições para a melhoria do meu trabalho.

Aos meus amigos e amigas que me foram presenteados nesse caminhar, por todos os momentos vividos, sejam nas trocas intelectuais como nos simples atos de companheirismo.

Um agradecimento especial aos funcionários da Universidade São Francisco, a aqueles da Secretaria da Pós-Graduação, Rose, Neisa e Marcela, que, com gentileza e atenção sempre souberam me orientar em todos os procedimentos da universidade durante a realização do meu mestrado.

Aos professores da graduação que me auxiliaram na difícil tarefa de aplicação dos meus instrumentos.

À todos os participantes de minha pesquisa, que sem essa colaboração esta pesquisa não teria sido possível.

À Universidade São Francisco – USF, pela oportunidade em realizar os meus estudos.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES, pelo apoio financeiro para a concretização de mais uma realização profissional.

E agradeço também aqueles que, simplesmente passaram em minha vida e me deram ricas contribuições tanto profissionais, quanto pessoais.

O meu muito obrigada a todos vocês!!!

RESUMO

Dias, A. S. (2008). *Evidências de validade de uma prova de compreensão em leitura em estatística*. Dissertação de Mestrado, Programa de Pós-Graduação Stricto Sensu em Psicologia, Universidade São Francisco, Itatiba.

O objetivo principal deste estudo foi o de verificar as evidências de validade de uma prova de compreensão em leitura de um texto com informações estatísticas em estudantes de Psicologia. Foram participantes 206 estudantes regularmente matriculados no curso de psicologia, sendo 79,1% do gênero feminino, com idades variando de 17 a 54 anos ($M=25,9$; $DP=6,8$), dos turnos diurno (17,5%) e noturno (82,5%). Foram utilizados um questionário sobre conceitos estatísticos, e duas provas de compreensão, uma de um texto geral, e a outra de um texto específico que envolve conceitos estatísticos. Os instrumentos foram aplicados coletivamente pela pesquisadora, em uma sessão, em horários regulares de aula. Os resultados indicaram que houve evidências de validade convergente entre os dois instrumentos de compreensão, uma vez que houve correlação significativa ($r=0,535$; $p<0,001$) entre a prova de compreensão de um texto que envolve conceitos estatísticos e a prova de compreensão de um texto de leitura geral. A análise fatorial com informação completa revelou a predominância unidimensional para os itens da prova de leitura específica, podendo explicar 24,80% da variância entre os itens. A análise pela Teoria de Resposta ao Item mostrou um melhor ajuste pelo modelo de três parâmetros, sendo este, em que se observou o menor resíduo. O coeficiente de fidedignidade avaliado pelo coeficiente de Kuder-Richardson, indicou que a prova apresenta uma boa precisão ($KR-21=0,813$).

Palavras-chave: compreensão em leitura; cloze; educação estatística; teoria de resposta ao item; psicometria.

ABSTRACT

Dias, A. S. (2008). *Validity's evidence of a reading comprehension test in statistic*. Masters Degree Dissertation, Programa de Pós-Graduação Stricto Sensu, Universidade São Francisco, Itatiba.

The main objective of this study was to verify the evidences of validity in reading comprehension test with a text about statistical information on students of psychology. Test's participants were 206 students regularly enrolled in the course of psychology, which 79.1% were female, with ages ranging from 17 to 54 years ($M = 25.9$, $SD = 6.8$), and the coursing daytime's (17.5%) and night's shift (82.5%). It was used a questionnaire on statistical concepts, and two tests of comprehension: which one was a general subject's text and the other was a specific text which involves statistical concepts. The instruments were applied collectively by the researcher, in a session, at the class' regular times. The results indicated that there was evidence of convergent validity between the two instruments of comprehension, since there was significant correlation ($r=0.535$; $p<0.001$) in the comprehension test with a text which involves statistical concepts, with a given external criterion, that was the comprehension test about a general reading text. A factorial analysis with complete information revealed the unidimensional predominance for items of specific reading test, what may explain 24.80% of the variance between the items. The analysis by the Item Response Theory showed up a better fit by the model of three parameters, and this, which it was observed the less resid. The coefficient of reliability evaluated by Kuder-Richardson's coefficient, stated that the text presents good accuracy ($KR-21 = 0,813$).

Keywords: reading comprehension; cloze; statistics education; item response theory; psychometrics.

Sumário

LISTA DE FIGURAS	vii
LISTA DE TABELAS.....	viii
LISTA DE ANEXOS.....	x
INTRODUÇÃO.....	1
AVALIAÇÃO PSICOLÓGICA E PSICOMETRIA	5
PARÂMETROS PSICOMÉTRICOS DE INSTRUMENTOS DE AVALIAÇÃO.....	11
TEORIA DE RESPOSTA AO ÍTEM.....	16
AVALIAÇÃO PSICO-EDUCACIONAL	19
LEITURA E COMPREENSÃO DE TEXTOS	25
A TÉCNICA DE CLOZE.....	32
OS CONCEITOS ESTATÍSTICOS	35
PESQUISAS RELACIONADAS À COMPREENSÃO EM LEITURA	41
OBJETIVOS	47
OBJETIVO GERAL.....	47
OBJETIVOS ESPECÍFICOS	47
MÉTODO	48
PARTICIPANTES.....	48
MATERIAL	49
PROCEDIMENTO	52
PROCEDIMENTO DE ANÁLISE DE DADOS.....	54
RESULTADOS E DISCUSSÃO	55
CONCLUSÃO E CONSIDERAÇÕES FINAIS	95
REFERÊNCIAS.....	99
ANEXOS.....	115

LISTA DE FIGURAS

Figura 1. Esquema de Representação da Leitura.....	26
Figura 2. Diagrama de dispersão e reta de regressão do desempenho em leitura específica em função do desempenho em leitura geral	67
Figura 3. Curvas características de alguns itens da prova de compreensão em leitura geral, modelo de dois parâmetros	88
Figura 4. Curvas características de alguns itens da prova de compreensão em leitura geral, modelo de três parâmetros	89
Figura 5. Curvas características de alguns itens da prova de compreensão em leitura específica, modelo de três parâmetros.....	93

LISTA DE TABELAS

Tabela 1. Distribuição do número de estudantes segundo as faixas etárias	48
Tabela 2. Distribuição do número de estudantes segundo o ano de graduação	49
Tabela 3. Estatísticas descritivas dos itens de múltipla escolha do questionário	56
Tabela 4. Estatísticas descritivas das questões abertas do questionário	58
Tabela 5. Estatísticas referentes à comparação da pontuação total média de compreensão em leitura geral por gênero e turno.....	62
Tabela 6. Estatísticas referentes à comparação da pontuação total média de compreensão em leitura específica por gênero e turno.....	62
Tabela 7. Estatísticas referentes à comparação da pontuação total média de compreensão em leitura geral por faixa etária e ano do curso.....	63
Tabela 8. Estatísticas referentes à comparação da pontuação total média de compreensão em leitura específica por faixa etária e ano do curso.....	63
Tabela 9. Prova de Tukey para a prova de leitura geral	64
Tabela 10. Prova de Tukey para a prova de leitura específica	65
Tabela 11. Prova de Tukey para a prova de leitura específica	66
Tabela 12. Parâmetros descritivos dos itens da prova de leitura geral.....	73
Tabela 13. Parâmetros descritivos dos itens da prova de leitura específica	76
Tabela 14. Cargas não rotacionadas dos fatores principais da prova de compreensão em leitura geral.....	80
Tabela 15. Cargas não rotacionadas dos fatores principais da prova de compreensão em leitura específica.....	82

Tabela 16. Parâmetros dos Itens da prova de compreensão em leitura geral para os modelos de 1, 2 e 3 parâmetros.....	86
Tabela 17. Parâmetros dos Itens da prova de compreensão em leitura específica para os modelos de 1, 2 e 3 parâmetros	91
Tabela 18. Características do componente das provas de compreensão	94

LISTA DE ANEXOS

ANEXO 1 – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO (1ª VIA)	116
ANEXO 2 – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO (2ª VIA)	117
ANEXO 3 – QUESTIONÁRIO	118
ANEXO 4 – CLOZE LEITURA GERAL – TOC	121
ANEXO 5 – CLOZE LEITURA ESPECÍFICA – OPÇÕES	125

INTRODUÇÃO

O ato de ler é uma atividade que se desenvolve ao longo da vida humana, e, tem sido ao longo da história, uma das formas mais importantes de que dispõe o homem, não só para o conhecimento do mundo, mas também para a expressão, criação e recriação desse conhecimento. Dessa maneira, esta é uma habilidade importante, por ser um dos meios de comunicação e manipulação de conhecimento (Witter, 1997, 1999).

Nunes e colaboradores (2006) citado por Gabriel (2006) afirmam que a leitura é uma atividade-meio para que se possam efetivar outras atividades como a comunicação, o acesso a informações, o devaneio, entre as demais que podem ser desempenhadas no mundo contemporâneo. Para que a leitura seja um direito exercido por todos os homens e que esses possam alcançar níveis como o da leitura compreensiva, crítica e criativa identifica-se a necessidade de pesquisar e atuar nesta área.

A leitura é uma atividade complexa que envolve o reconhecimento de palavras impressas, o acesso à palavra na memória permanente e a compreensão da mensagem escrita (Salles & Parente, 2002). A compreensão prevê a integração de elementos da memória com o texto. E, mesmo que o leitor tenha familiaridade com o código escrito, com o gênero textual e conhecimento prévio sobre o assunto, ele deve ter uma atitude ativa de cooperação para construir mentalmente a estrutura do texto, realizar as devidas inferências e aprender por meio da leitura (Gabriel, 2006).

Alguns estudiosos têm enfatizado a compreensão como habilidade fundamental em todos os níveis do desenvolvimento da leitura. Procurando defini-la, preocupando-se em como medi-la, ou buscando estratégias para desenvolvê-la, muitas pesquisas têm sido desenvolvidas com a finalidade de estudar as habilidades complexas da leitura em

estudantes (Centofanti, Ferreira & Tedesco, 1997; Oliveira, 2003; Santos, 1991; Wolff, Polimero, Oliveira, Henriques & Socci, 1997). Ressaltando que, ao averiguarem-se as dificuldades de compreensão em leitura, abre-se o caminho para a busca e descobertas de novas soluções (Witter, 1997).

É importante destacar ainda, que os estudantes ingressam na universidade com deficiências em habilidades básicas como a leitura, e estes déficits são especialmente prejudiciais, pois esta está relacionada com todas as disciplinas curriculares (Gabriel, 2006). Sendo que, o estudante que está inserido neste nível de ensino, visto como um sujeito em formação deve desenvolver tal habilidade, para que adquira padrões de comportamentos que internalize valores úteis para as tomadas de decisões, argumentações e posicionamentos pessoais e profissionais com uma maior segurança (Witter, 1997).

A partir desta perspectiva, a leitura deveria ser uma das preocupações mais importantes na universidade, e embora esta seja uma atividade indispensável na educação de um indivíduo, não tem recebido a atenção merecida desde o ensino fundamental, e principalmente no ensino superior. É evidente que cada vez mais se percebe a necessidade de receber os estudantes no nível superior com treino em leitura e principalmente com o objetivo de habilitá-los na compreensão e análise dos textos indicados, desta forma, não se pode excluir a universidade da tarefa de continuar melhorando o nível de leitura do estudante (Centofanti, e cols., 1997).

Para o universitário, assim como para toda a população, é disponibilizada na mídia impressa textos com diversos tipos de informação, e muitos deles apresentam em seu conteúdo informações resultantes de análises estatísticas de dados, que incluem assim, medidas estatísticas. Desta forma, no mundo de hoje é fundamental que as pessoas desenvolvam habilidades de leitura não só de textos de interesse geral, mas também de

conceitos estatísticos que têm o objetivo de resumir as informações em medidas descritivas, tabulares ou gráficas.

Nesse sentido julgou-se necessário e fundamental a realização desta pesquisa sobre a leitura e compreensão de texto que envolve a leitura de um conteúdo específico que inclui informações estatísticas. Na área de alfabetização estatística em diferentes níveis de instrução, os resultados têm revelado a importância e urgência em pesquisas nessa área. Além desta necessidade, o domínio das habilidades de leitura é crucial para o desempenho do indivíduo nas situações acadêmicas, constituindo-se em um instrumento fundamental, uma vez que, depois de aprender a ler o indivíduo lê para aprender. Vale destacar ainda, que um dos grandes empecilhos para o bom desempenho dos universitários ao longo do curso é a falta de habilidade para compreender o discurso textual.

Assim, o presente trabalho pretende discutir questões relativas à compreensão em leitura de estudantes do curso de psicologia, a partir de um texto de leitura específica com informações estatísticas. Essa escolha surgiu principalmente da necessidade de realizar estudos acerca da compreensão em leitura de universitários em textos técnicos que envolvem conceitos específicos, e também pela necessidade de pesquisas que comprovem, com bases científicas, a validade das técnicas que são empregadas para este fim. Assim como, para auxiliar na compreensão de alguns fatores que poderiam dificultar o processo de aprendizagem de conceitos estatísticos necessários para a pesquisa e prática em psicologia.

Isto posto, o presente estudo teve por objetivo principal verificar evidências de validade de uma prova de compreensão em leitura de um texto com informações estatísticas em estudantes de psicologia. Para introduzir a problemática, inicialmente é apresentada uma revisão da literatura, que aborda conceitos de leitura, compreensão em leitura e a técnica de Cloze, além de relacionar aspectos teóricos com alguns estudos relevantes identificados no levantamento realizado a respeito da temática. Em seqüência, é

apresentado o método, com a descrição dos participantes, dos instrumentos empregados e dos procedimentos utilizados para sua aplicação, bem como o procedimento e a análise de dados. Por último, apresentam-se as referências. Também foi disponibilizado, em anexo, um exemplar dos instrumentos utilizados nesta pesquisa.

AVALIAÇÃO PSICOLÓGICA E PSICOMETRIA

A avaliação psicológica não é uma atividade recente, há registros na literatura de que instrumentos de avaliação foram elaborados já no século XIX. Essa atividade pode ser entendida como um processo de coleta de dados que visa descrever e classificar comportamentos, obtidos por meio da utilização de instrumentos psicológicos a fim de se adquirir um maior conhecimento do indivíduo e conseqüentemente a tomada de decisões (Anastasi & Urbina, 2000; Pasquali, 2001). Para Urbina (2007) avaliação psicológica é um processo flexível cujo objetivo é determinar uma ou mais questões psicológicas por meio da coleta, avaliação e análise de dados coletados com instrumentos psicológicos apropriados ao objetivo em questão.

Pasquali (2001) afirma que a avaliação é um processo integrado que utiliza técnicas mais apropriadas para diagnosticar o problema de um determinado caso, objetivando a intervenção. Uma das mais típicas técnicas de avaliação psicológica é o teste psicológico, que surgiu com o início da ciência psicológica e, de alguma forma, marcou o desenvolvimento desta ciência. Os testes são definidos por Anastasi e Urbina (2000, p. 18) como “uma medida objetiva e padronizada de uma amostra de comportamento”. Os testes vêm ganhando espaço em discussões e eventos promovidos por universidades, entidades, órgãos e associações de classe. São instrumentos exclusivos do psicólogo e podem fornecer importantes informações para a elaboração de um diagnóstico, laudos em qualquer processo de avaliação clínica, educacional, organizacional, entre outros.

Como qualquer instrumento científico ou ferramenta de precisão, os instrumentos precisam ser usados apropriadamente para serem efetivos, caso contrário, podem causar danos. Duas razões para o controle do uso de testes psicológicos, são: a primeira, garantir

que o teste seja aplicado por um examinador qualificado e que os seus escores sejam adequadamente utilizados; e o segundo, evitar uma familiaridade geral com o conteúdo do teste, o que o invalidaria. Um examinador qualificado deve atender a três aspectos principais da situação de testagem: a seleção de teste, a aplicação e a pontuação (Anastasi & Urbina, 2000).

Nos últimos anos a avaliação psicológica tem recebido atenção crescente devido às dificuldades apresentadas pelos profissionais da classe, pois são muitos os problemas encontrados nesta área de conhecimento, tais como, instrumentos desatualizados e sem estudos nacionais, poucos instrumentos para avaliar determinados construtos, má utilização do instrumental por parte dos profissionais, como por exemplo, a falta de domínio e de conhecimento do instrumento, e avaliação incorreta dos resultados obtidos, entre outros, como alguns dos problemas nesta área (Noronha, Primi & Alchieri, 2005; Primi, 2003).

Nesta perspectiva, alguns estudos foram realizados nos últimos anos, como a pesquisa relatada por Almeida, Prieto, Muñiz e Bartram (1998), que indicou dentre os problemas mais frequentes entre os utilizadores de testes: utilização de testes inadequados para algumas situações; não estar em sintonia com as modificações da área; avaliações incorretas; não ter clareza das limitações dos instrumentos, quanto às normas; aplicação de testes por leigos; não adaptar os instrumentos para os determinados países ou regiões; não arquivar os instrumentos e não dar o devido seguimento aos estudos dos testes; e, fazer interpretações que extrapolam o instrumento.

Problemas semelhantes foram detectados por Noronha (2002) quando solicitou aos psicólogos, em uma pesquisa, que determinassem os problemas mais graves e mais frequentes em relação ao uso de testes psicológicos. Os resultados indicaram duas categorias de problemas mais graves presentes na avaliação psicológica, uma relativa aos instrumentos, no que diz respeito à construção, às normas e às características, e a outra

quanto ao seu uso. Os psicólogos também justificaram que esses problemas não estariam ocorrendo se os testes não fossem aplicados indevida, indiscriminada e mecanicamente; por pessoal não qualificado, sem critérios; com instruções erradas; se não houvesse erro de avaliação; supervalorização do quantitativo; interpretação generalizada e o uso de um único instrumento como resultado definitivo. Os participantes da pesquisa consideram que ainda existe preconceito por parte dos profissionais em relação aos testes, que não estão preparados para atuar nessa área e que não há muita divulgação dos instrumentos pela comunidade psicológica.

Ainda nesta perspectiva, Almeida, Prieto, Muniz e Bartram (1998) citado por Noronha (2002) realizaram uma pesquisa considerando a utilização de instrumentos psicológicos. Os autores evidenciaram alguns aspectos, como o de não existir padrões de atuação na área; não haver uma avaliação da competência profissional dos usuários de instrumentos psicológicos; não se determinar uma conduta comum de formação específica; além de não se estabelecerem cursos de formação para aperfeiçoamento da ciência. Destacaram ainda, que os problemas encontrados na formação profissional, podem estar relacionados com os problemas encontrados na utilização dos testes psicológicos.

Considerando tais evidências e urgências, o Conselho Federal de Psicologia (CFP, 2001) publicou uma resolução na qual regulamenta a elaboração, a comercialização e o uso dos instrumentos psicológicos. É grande a preocupação com a qualidade das ações profissionais. A partir desta problemática, existem publicações, como as da Associação Psicológica Americana (APA) que promove aos profissionais, uma discussão do tema, uma articulação dos conteúdos e das práticas profissionais, deixando os psicólogos melhor informados e preparados para os diferentes usos dos testes. A ciência psicológica dispõe de um código de Ética Profissional que traz orientações importantes ao profissional a respeito

da amplitude das possibilidades e das responsabilidades de sua atuação, inclusive no que diz respeito à prática de avaliação.

De acordo com Anastasi e Urbina (2000), para que um teste cumpra a sua função, é essencial uma avaliação de seus méritos técnicos em termos de características como validade, fidedignidade, nível de dificuldade e normas. Assim, os testes psicológicos baseiam-se em fundamentos da medida científica, denominada pela psicologia como psicometria (Vendramini & Noronha, 2003). Só assim, os usuários de testes podem determinar a adequação de algum teste para seus objetivos específicos e sua conveniência para o tipo de pessoas com as quais planejam usá-las, daí a importância de um profissional qualificado para o manejo destes instrumentos.

O uso e o conhecimento de estudantes e profissionais de psicologia acerca dos instrumentos de avaliação psicológica comercializados no Brasil foram investigados por Noronha e Freitas (2005). As autoras verificaram que um pequeno número de profissionais que participou da pesquisa conhece e utiliza os instrumentos de medida. Dos 145 instrumentos, que constavam da lista de consulta da pesquisa, o número máximo de testes conhecido pelos profissionais foi 84 (57,9%). As autoras observaram ainda, uma diferença significativa do número médio de testes conhecidos entre os entrevistados de diferente titulação. De forma que, os psicólogos mestres conhecem mais instrumentos ($M=28,6$) que doutores ($M=20,43$) e graduados ($M=16,83$) e conhecem mais do que especialista ($M=14,93$), sendo este último possui aproximadamente o mesmo conhecimento que um estudante da área ($M=14,15$). Assim, os mestres apresentaram um maior conhecimento de instrumentos, enquanto que o menor foram os estudantes de psicologia juntamente com os especialistas.

Os instrumentos psicológicos devem apresentar fundamentação teórica, informações precisas que evidenciem os estudos de validade e precisão sobre os escores obtidos nos

testes e, ainda, os procedimentos utilizados na investigação. E, no que concerne aos dados estatísticos, os testes psicológicos devem apresentar o sistema de correção e interpretação dos escores, para uma maior especificação da lógica que o fundamenta. Destaca-se o estudo desenvolvido por Lopes e Vendramini (2007a) que objetivaram analisar a leitura realizada por psicólogos, em informações estatísticas e de avaliação psicológica, bem como suas dificuldades apresentadas em manuais de testes psicológicos. Os resultados demonstram que 100% dos entrevistados lêem as informações referentes a normas de aplicação e de correção de testes psicológicos, enquanto que, 46,7% se informam sobre evidências de validade e 40% de precisão, itens detectados como os de menor interesse dos entrevistados. Dos participantes 8,3% afirmam não usarem testes psicológicos em sua prática profissional, a maioria dos participantes informou encontrar alguma dificuldade na correção de testes (65,4%) e interpretação dos escores (52,4%), sendo nos dois casos atribuída à falta de clareza de informações nos manuais.

No que se refere às informações estatísticas lidas em manuais de testes utilizados por estudantes de Psicologia, Lopes e Vendramini (2007b) observaram dificuldades de interpretação de diferentes tópicos associados à avaliação psicológica e à estatística relatadas por estudantes de psicologia e por psicólogos. Dos entrevistados, 100% lêem as informações referentes a normas de aplicação e de correção de testes psicológicos, enquanto que, 38,1% se interessam por evidências empíricas de precisão e 42,9% de validade. Dos participantes, 32,4% não tem dificuldade para corrigir testes psicológicos e 26,5% denotam ter dificuldades na interpretação dos resultados. Quanto ao conhecimento de termos estatísticos, 78% responderam ser capazes de interpretar o que é média e 59,5% reconhecem a importância da estatística para a interpretação de dados de testes psicológicos.

Segundo as autoras, embora o Conselho Federal de Psicologia considere a importância dos fundamentos teóricos e psicométricos dos instrumentos psicológicos, os estudos relatados anteriormente revelaram que nem todos os usuários de testes lêem todas as informações necessárias para o uso seguro e confiável desses instrumentos na prática profissional. Esses resultados indicam a necessidade de preparar melhor os psicólogos para uma leitura mais completa e compreensiva das informações gerais e estatísticas contidas nos manuais de testes psicológicos.

Também se deve levar em consideração o fato de que determinados testes não apresentam em seus manuais os dados mais básicos, quais sejam, os de identificação como nome, autor, data e editor, bem como defasagens referentes às instruções, normas, dados de identificação. Diferentes autores evidenciaram em seus estudos testes que não apresentavam em seus manuais os requisitos necessários, de acordo com a literatura, quanto a elaboração de uma publicação científica, em especial, as de um teste psicológico (Adánéz, 1999; Alves, 2002; Anastasi & Urbina, 2000; CFP, 2001; Noronha, Sartori, Freitas & Ottati, 2001; Noronha, Sartori, Freitas & Ottati, 2002; Ottati, Noronha & Salviati, 2003). Desta maneira, cabe ao profissional adquirir o conhecimento necessário e prudência na utilização de instrumentos, usufruindo apenas daqueles que ofereçam qualidades psicométricas que proporcionem dados confiáveis, que garantam uma maior confiança nos dados obtidos.

A preocupação com a construção dos testes é fundamentada no pressuposto real de que o instrumento fornecerá dados relevantes ao profissional, e que, portanto, é necessário que esses dados sejam verdadeiros e confiáveis (Adánéz, 1999; CFP, 2001; Ottati e cols, 2003). Adánéz (1999) relata algumas etapas de construção de um teste, como a elaboração do manual, considerando que o construtor deverá redigi-lo de maneira que estejam especificadas todas as características do instrumento, entre elas, a fundamentação teórica,

objetivos, possíveis usos, populações a que se destina, recursos necessários para sua administração, procedimentos de administração e correção, precisão, validade, referências bibliográficas e qualquer outra recomendação necessária para o uso correto.

A preocupação com a qualidade da avaliação e dos instrumentos utilizados nesse processo, como as dos pesquisadores citados nos parágrafos anteriores, devem estar presentes tanto na avaliação psicológica quanto na avaliação educacional. Assim, como é o caso da presente pesquisa que visa buscar evidências de validade de uma prova de compreensão em leitura de um texto específico, que poderá contribuir com pesquisadores e professores que atuam nesta área de conhecimento.

Neste sentido, são apresentados a seguir os parâmetros psicométricos de instrumentos de avaliação, suas propriedades e métodos de análise de itens que compõem esses instrumentos, como é o caso da Teoria de Resposta ao Item.

PARÂMETROS PSICOMÉTRICOS DE INSTRUMENTOS DE AVALIAÇÃO

A partir da importância da utilização de instrumentos válidos e confiáveis para inferir sobre o comportamento de uma pessoa ou grupo, faz-se necessário destacar aspectos importantes para construção e utilização de testes psicológicos e educacionais. Para proceder uma avaliação psicológica deve-se não apenas escolher um instrumento adequado para obter informações desejadas, como também garantir a confiabilidade das inferências feitas a partir da aplicação de tal instrumento de pesquisa, o qual deve ter propriedades psicométricas, como fidedignidade e validade (Anastasi & Urbina, 2000).

Pasquali (2003) afirma que os testes são instrumentos de medida na ciência psicológica que devem apresentar certas características que justifiquem a confiança nos dados produzidos. Ainda de acordo com o autor, a psicometria definiu as características

essenciais de um bom instrumento de medida, obtidas por meio de métodos de avaliação objetiva a qual inclui fundamentalmente, a determinação de sua fidedignidade e de sua validade, em situações específicas. Conforme Cronbach (1996), para atribuir ao teste psicológico seus parâmetros fundamentais, precisão e validade, é necessário que se verifiquem os parâmetros dos itens que o compõe.

A fidedignidade pode ser denominada por alguns autores como padronização ou confiabilidade, é definida por Anastasi e Urbina (2000, p.84) como “consistência dos escores obtidos pelas mesmas pessoas quando elas são reexaminadas com o mesmo teste em diferentes ocasiões, ou com diferentes conjuntos de itens equivalentes, ou sob outras condições variáveis de exame”. De acordo com as autoras, todos os tipos de fidedignidade preocupam-se com o grau de consistência ou concordância entre dois conjuntos de escores independentes, assim, todos eles podem ser expressos em termos de um coeficiente de correlação (r) que indica o grau de correspondência ou relacionamento entre dois conjuntos de escores.

Normalmente, os métodos que são utilizados para a determinação do grau de fidedignidade de um instrumento de medida são: teste-reteste, forma-alternada, método das metades (*split-half*), coeficiente Kuder-Richardson e coeficiente alfa de Cronbach. O método teste-reteste correlaciona às pontuações obtidas pelos mesmos sujeitos, em duas ocasiões com o mesmo teste, o coeficiente calculado indica em que medida as pontuações obtidas na primeira ocasião mantém sua estabilidade na segunda. Porém, o maior inconveniente técnico é determinar o tempo adequado entre as duas ocasiões para evitar a recordação das respostas e a aprendizagem.

O coeficiente de precisão de um teste obtido pelo método de forma-alternada consiste na correlação entre os escores produzidos pelas respostas dos sujeitos a duas formas paralelas de um mesmo teste construído para satisfazer às mesmas especificações. O

método das metades (*Split-Half*) baseia-se em uma única forma do instrumento de medida aplicada em uma única vez, cujos resultados são divididos em duas metades e estabelecendo-se o índice de correlação existente entre os resultados dessas duas metades.

Dentre outras medidas para obtenção da consistência interna, há utilização de fórmulas que visam as correlações entre os itens do instrumento, em outras palavras, é a correlação do desempenho entre cada item do teste (Urbina, 2007). As duas fórmulas mais utilizadas são o *Kuder-Richardson* que se dá pela análise de cada item individual do teste, aplicável somente quando a resposta ao item é dicotômica; e o *Alfa de Cronbach* que é a divisão racional do teste em subtestes, com aplicação em testes heterogêneos, formada por uma amostra estratificada de itens (Anastasi & Urbina, 2000; Pasquali, 2003).

Essas duas fórmulas de obtenção da fidedignidade consideram em sua função, dois fatores, quais sejam: 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. Desta forma, essas duas fórmulas determinam estimativas de fidedignidade que se igualam à medida de todos os coeficientes das metades possíveis que resultariam de todas as possíveis formas de se dividir o teste (Urbina, 2007).

Como já citado anteriormente, a outra característica importante de um bom instrumento é a *validade*, que pode ser compreendida como o grau em que o instrumento mede aquilo que se propõe a medir e representa uma verificação direta da possibilidade do instrumento satisfazer o seu objetivo (Pasquali, 2003). São diversas as técnicas empregadas para a determinação do índice de validade de um instrumento, de forma geral pode-se agrupá-las em três categorias: validade de construto, validade de conteúdo e validade de critério (Anastasi & Urbina, 2000).

Segundo Anastasi e Urbina (2000) a validade de construto de um teste é a extensão em que se pode dizer que o teste mede um construto teórico ou um traço. Para Pasquali

(2003), a validade de construto de um teste pode ser trabalhada sob duas formas, (1) análise da representação comportamental do construto e (2) análise por hipótese. Para a análise da representação comportamental são utilizadas duas técnicas como demonstração da adequação da representação do construto: análise da consistência interna e a análise fatorial. A primeira consiste essencialmente em verificar a homogeneidade dos itens que compõem o teste e a segunda tem como lógica verificar quantos construtos comuns são necessários para explicar as intercorrelações dos itens. A análise fatorial também postula que um número menor de traços latentes é suficiente para explicar um número maior de variáveis observadas (os itens).

Já a análise por hipótese se refere ao poder que um instrumento psicológico tem em discriminar um determinado critério externo, em outras palavras, discriminar grupos caracterizados como critérios, que deferiram no traço que o teste mede. Esse método é muito eficaz quando o método e as variáveis são suficientemente diferentes entre si. Um exemplo muito utilizado como critério externo é a idade, para a validação de construto quando um teste mede traços que são dependentes de mudanças no desenvolvimento do sujeito (Pasquali, 2003).

De acordo com Anastasi e Urbina (2000), a validade de conteúdo é inserida em um teste desde o início, através da escolha de itens apropriados, também pode se dar pela participação de peritos no assunto. Assim, para Pasquali (2003) este tipo de validade constitui uma amostra representativa de um universo finito de comportamentos. A descrição do conteúdo envolve essencialmente o exame sistemático do conteúdo do teste para determinar se abrange uma amostra representativa do domínio do comportamento a ser medido.

A validade de critério, de acordo com Pasquali (2003) é o grau de eficácia que o teste tem para predizer um desempenho específico de um indivíduo. Podem-se distinguir

dois tipos de critério, a validade preditiva e a validade concorrente. A preditiva refere-se ao grau em que as pontuações do teste predizem medidas do critério realizadas posteriormente (comportamentos futuros do indivíduo no critério). Já a concorrente refere-se ao grau em que as pontuações do teste se correlacionam com as do critério, medidas ao mesmo tempo (comportamentos atuais do indivíduo no critério).

Os padrões para testagem psicológica e educacional propõem que o processo de validação envolva o acúmulo de evidências, com o objetivo de gerar base científica para a interpretação confiáveis. De acordo com estes padrões a validade se refere ao grau de evidência e suporte para as interpretações dos escores do teste, vinculados aos usos propostos pelos testes. As fontes de evidência podem dar subsídios para diferentes aspectos da validade, porém esses dados obtidos não são suficientemente capazes de representar em si, os diferentes tipos de validade existentes.

De acordo com os padrões para testagem psicológica e educacional as evidências, são baseadas, no conteúdo do teste, na relação entre o conteúdo do teste e o que o construto pode medir; no processo de resposta que provém da análise da resposta individual; na estrutura interna que pode indicar o grau das relações entre os itens dos testes e os seus componentes; nas relações com outras variáveis, análise dos escores dos testes com variáveis externas, sendo outra fonte importante de validade; na evidência convergente que consiste nas descrições entre escores do teste e outras medidas pretendidas para avaliar construtos semelhantes; e, na evidência discriminante avaliam ligações entre escores dos testes e medidas propostas de diferentes construtos (Pasquali, 2003).

Os procedimentos utilizados para verificar se os parâmetros de um instrumento de avaliação psicológica ou educacional atendem as propriedades psicométricas de um instrumento de qualidade são procedimentos estatísticos que compõem a Teoria Clássica dos Testes (TCT) ou a Teoria de Resposta ao Item (TRI). Os métodos utilizados pela TCT,

por serem mais antigos, já foram discutidos e apresentados em inúmeros livros e documentos científicos de psicometria. Por esse motivo, optou-se por descrever nesse trabalho apenas os procedimentos mais recentes de análise de instrumentos de medida, os procedimentos da TRI, e que estão descritos a seguir.

TEORIA DE RESPOSTA AO ITEM

A Teoria de Resposta ao Item surgiu com os trabalhos pioneiros de Lord (1952), nos Estados Unidos, e Rasch (1960) na Holanda. A partir da década de 80, a TRI, utilizada em testes de desempenho e aptidão, vem se tornando uma técnica predominante no campo de testes, passando a substituir grande parte da Teoria Clássica dos Testes (TCT) aplicada à avaliação educacional (Soares & Pereira, 2001). De acordo com Fletcher (1994), a TRI oferece recursos mais eficazes e flexíveis na confecção, análise e apresentação dos resultados de provas, em relação a outros recursos equivalentes derivados da teoria clássica.

Segundo Valle (2000), com a utilização dessa técnica é possível demonstrar a relação entre a probabilidade de um estudante responder corretamente a um item e suas habilidades na área de conhecimento que será avaliada como a matemática e a estatística. A TRI trabalha com traços latentes (aptidões ou habilidades) e se baseia em duas premissas, a primeira, fundamenta-se na função em que o desempenho de um examinando em um item do teste é função de um conjunto de traços latentes. A segunda, na relação entre o desempenho e as habilidades e que pode ser descrita por uma equação monotônica crescente, ou seja, a Curva Característica do Item (Pasquali, 2003).

Pode-se citar duas características da TRI, a primeira refere-se a possibilidade de se atribuir em uma mesma escala métrica a dificuldade dos itens, assim como a habilidade dos sujeitos. O modelo matemático que fundamenta a TRI permite prever a probabilidade de

acerto de uma pessoa com determinada habilidade representada pelo resultado da prova (Fletcher, 1994). A segunda é com relação à premissa de fornecer medidas invariantes do desempenho cognitivo, não dependendo dos itens que compõem a prova, nem das pessoas investigadas. Assim, esta teoria permite que se comparem os resultados de provas de dificuldade variável, e ainda relacionar resultados na mesma escala de desempenho cognitivo. Esta metodologia vem sendo utilizada nos processos qualitativos da avaliação psicológica e educacional, em instrumentos de medida, como as escalas de habilidades, para avaliar e acompanhar o conhecimento adquirido pelos estudantes como também o desenvolvimento de habilidades básicas (Vendramini, 2002).

Os objetivos básicos da TRI são a obtenção de estimações do nível dos indivíduos no traço, independente da amostra de itens utilizada a construção de testes cujas propriedades não dependam da amostra de indivíduos selecionada para o estudo. A função que representa a probabilidade de um indivíduo, com um nível de traço conhecido, dar uma resposta correta ou típica ao item é denominada Curva Característica do Item (CCI) e é determinada por meio dos parâmetros dos itens. Na TRI, a seleção dos itens que compõem um instrumento é feita pela função de informação do item que indica o quanto o item é preciso e útil para aquele instrumento (Pasquali, 2003).

Ainda de acordo com o autor, existem modelos matemáticos que são aplicados aos dados, que visam calibrar os itens de teste em relação a um ou mais parâmetros e derivar estimativas de probabilidade de quantificação de habilidade ou nível de traço, necessárias para responder a cada item de certa maneira. Essencialmente, estes modelos colocam pessoas e itens de teste em uma escala comum.

A TRI é um método sofisticado para estimar a fidedignidade, bem como o erro de mensuração, que são abordados do ponto de vista da função de informação de itens individuais do teste. Como o nível de dificuldade e o poder discriminativo de itens

individuais, em relação ao traço avaliado pelo teste, podem ser calibrados mais cuidadosamente pelos métodos da TRI, as informações oferecidas pela resposta de cada testando são mais precisas e, por isso, mais fidedignas (Urbina, 2007).

Assim, justifica a utilização dessa teoria para a obtenção da fidedignidade da prova de compreensão em leitura específica, que é utilizada neste estudo. É a partir da verificação dos parâmetros dos itens que compõem a prova que se pode inferir sobre a qualidade do mesmo, daí a importância de se verificar a qualidade desses parâmetros.

Para a realização da avaliação psicológica, como uma ciência, são utilizados testes psicológicos que como descritos anteriormente tem que possuir qualidades psicométricas. Porém, existe outras formas de avaliação para a obtenção de informações, sobre conhecimentos, habilidades ou funções cognitivas, e até mesmo o desempenho geral dos estudantes, sendo estes testes utilizados na área psico-educacional. Uma breve descrição será realizada no próximo capítulo.

AVALIAÇÃO PSICO-EDUCACIONAL

Anastasi e Urbina já apontavam em 2000 que diferentes tipos de testes de avaliação são utilizados na área educacional, como por exemplo, os testes de habilidades cognitivas e de personalidade. Contudo, certos tipos de testes foram desenvolvidos especificamente para uso em contextos educacionais, que incluem instrumentos planejados para a predição e a classificação em ambientes educacionais específicos e uma enorme variedade de testes de realização e desempenho escolar.

Para alguns autores o desempenho é o comportamento emitido como decorrência do processo de aprendizagem, podendo ser adversamente afetado por falta de motivação ou ansiedade, de modo que a aprendizagem pode ter ocorrido, mas não ser demonstrável (Kaplan, Sadock & Grebb, 1997; Witter & Lomônaco, 1984). O desempenho acadêmico envolve a ação, sendo apresentado na forma do rendimento acadêmico por meio das notas obtidas em atividades específicas, ou seja, é o resultado da realização de atividades acadêmicas que envolvem as habilidades cognitivas e o domínio acadêmico específico mostrado pelo rendimento médio do aluno (Munhoz, 2004).

De acordo com Krutetskii (1976), habilidades são características psicológicas individuais de um sujeito que favorecem um domínio rápido e fácil de uma atividade definida, ou seja, são traços psicológicos dos indivíduos, que dependem de um domínio fácil das destrezas e hábitos em uma atividade, como por exemplo, a atividade matemática e estatística. Para avaliar o desempenho acadêmico, deve-se considerar a relação entre as habilidades cognitivas e o domínio de conteúdos específicos como apontado por Flanagan, Ortiz, Alfonso e Mascolo (2002). Segundo esses autores para se obter um bom desempenho

em qualquer área é preciso que se realize um processo mental, a partir das informações relativamente específicas.

Como relatado por Anastasi e Urbina (2000) os educadores cada vez mais se preocupam com as dificuldades de aprendizagem de seus estudantes. Entre elas, com o baixo nível de desempenho em leitura, escrita e aritmética, essas habilidades básicas destacadas por Flanagan e colaboradores (2002) como sendo as três áreas tradicionais que envolvem o conceito de desempenho, que podem ser avaliadas por testes que certifiquem a obtenção de uma competência mínima.

Destaca-se que há uma preocupação crescente com instrumentos de avaliação, tanto psicológicos, quanto os educacionais. Uma vez que, os testes psicológicos devem apresentar qualidades psicométricas de forma a propiciar resultados confiáveis sobre os construtos psicológicos. Desta forma, avaliações têm sido realizadas para a caracterização dos instrumentos como adequados para a sua utilização. E, no que se refere às avaliações educacionais não há um rigor tão fundamentado para a sua realização, porém, a qualidade se refere na construção de itens que condizem com o objetivo principal do teste.

Assim, de acordo com Anastasi e Urbina (2000), alterações significativas estão sendo realizadas em quase todos os tipos de testes, como pode-se citar a atualização em um programa de testagem que foi desenvolvido para o uso na admissão, colocação e no aconselhamento de alunos universitários nos Estados Unidos. O mais conhecido, e o que mais se destaca, é o Programa de Testes de Avaliação Escolar (*Scholastic Aptitude Test – SAT*) do *College Board*. O SAT é composto por dois subtestes, o SAT-I, um teste de raciocínio, que consiste em perguntas de múltipla escolha destinadas a medir habilidades verbais e matemáticas, utilizado na avaliação da prontidão de um indivíduo para realizar um trabalho de nível universitário; e, o SAT-II, um teste de conhecimento específico, que

tem como objetivo avaliar o conhecimento em áreas específicas, como literatura e química, e pode ser usado para colocação, assim como para admissão.

No Brasil, também existem algumas formas de avaliar o desempenho do estudante em sua vida acadêmica, uma delas é a avaliação do Exame Nacional do Ensino Médio (ENEM), realizado com alunos do Ensino Médio, e cuja pontuação é utilizada no processo seletivo de ingresso a várias instituições de ensino superior. Uma outra forma de avaliar o desempenho do estudante em nível universitário, é Exame Nacional de Desempenho do Estudante (ENADE) que tem por objetivo avaliar o conhecimento adquirido no curso universitário.

O ENADE foi criado com o intuito de obter informações e proceder análises sobre a evolução dos processos educativos em cada área do conhecimento e suas relações com a totalidade, incorporando questões relacionadas ao domínio de aprendizagens e habilidades dos estudantes. Tem como finalidade principal acompanhar o processo de aprendizagem e o próprio desempenho do estudante frente aos conteúdos que compõem os currículos dos cursos, bem como suas habilidades e competências, esperando no final do curso níveis de habilidade mais elevados que no início, indicando ganhos para o aluno, para o curso e para a instituição (Vendramini, 2005).

De acordo com a legislação em vigor (Lei nº 10.861 de 14 de Abril de 2004, art. 5º e seus respectivos parágrafos), o ENADE tem por objetivo aferir “o desempenho dos estudantes em relação aos conteúdos programáticos previstos nas diretrizes curriculares dos respectivos cursos de graduação, suas habilidades para ajustamento às exigências decorrentes da evolução do conhecimento e suas competências para compreender temas exteriores ao âmbito específicos de sua profissão ligados à realidade brasileira e mundial e a outras áreas do conhecimento” (MEC, 2006). O ENADE é um exame em larga escala, que também objetiva colher informações relativas às características sócio-econômicas a

respeito dos estudantes selecionados através de procedimentos de amostragem (Limana & Brito, 2005).

O conjunto de habilidades acadêmicas, competências profissionais e conhecimentos que devem ser avaliados pelo ENADE são aqueles definidos nas diretrizes curriculares de cada área. Portanto, na elaboração das provas de cada curso avaliado pelo ENADE, tenta-se basicamente operacionalizar itens na medida dos construtos que são definidos nas diretrizes de cada curso.

O ENADE, deste modo, é formulado com o objetivo de verificar as competências e habilidades básicas das áreas, os conhecimentos sobre conteúdos básicos e profissionalizantes e, ainda verificar o desempenho em questões transdisciplinares, envolvendo o conhecimento mais geral. O exame é aplicado trienalmente por grupos de áreas, em amostras estratificadas de estudantes do primeiro e do último ano dos cursos avaliados em cada ano.

Os alunos que participam do exame são selecionados aleatoriamente a partir de uma relação completa de ingressantes e concluintes enviadas pelas instituições de ensino ao Ministério da Educação. A amostragem aleatória onde todos os elementos da população têm probabilidade conhecida e não nula de pertencerem à amostra e a sua representatividade permite que sejam feitas inferências confiáveis de características educacionais da população o que possibilita orientar a tomada de decisões relacionadas aos cursos avaliados (Vendramini, 2005).

As provas do ENADE são compostas por 30 questões de conhecimento específico da área e 10 questões de formação geral, idênticas para todas as áreas, totalizando 40 questões. As provas específicas de área são construídas a partir de uma matriz de especificação elaborada inicialmente por uma comissão formada por especialistas da área em questão que descreve detalhadamente os objetivos da prova, os conteúdos, habilidades

acadêmicas e competências profissionais a serem avaliados. Essa especificação é apresentada a um outro grupo de especialistas que elabora as questões. Uma versão paralela é discutida com a primeira comissão para o ajuste final das questões antes da composição da prova final, em caráter sigiloso. Do total das questões da prova, a maioria é criada em formato de múltipla escolha, mas também há uma fração dessas que exigem resposta dissertativa.

Dentre as respostas apresentadas pelos estudantes nessas questões dissertativas, a banca responsável pela correção avaliou que os estudantes, de uma maneira geral, apresentam um baixo domínio da Língua Portuguesa, sendo verificados por meio de inúmeros erros de ortografia e de gramática que acabavam por comprometer a compreensão da resposta dada pelo estudante. Outra dificuldade apresentada foi com relação à leitura e compreensão do enunciado da questão, apresentando-se de diferentes formas. A primeira dificuldade refere-se a confusão na compreensão de um problema de pesquisa, de forma que os estudantes elaboraram a sua resposta como um problema a ser solucionado, quando na verdade deveria apresentar uma proposta de intervenção. A segunda forma apresentada pelos corretores, foi o fato de que os universitários compreendiam o problema apresentado como erro a ser identificado e esclarecido, quando deveriam apresentar uma proposta corretiva à questão abordada pelo item (MEC, 2006).

Uma última forma de dificuldade apresentada foi com relação a um discurso produzido pelos universitários, com respostas estereotipadas que propunham temas pertencentes ao senso-comum, ao invés de trazerem elementos investigados pertencentes à ciência psicológica, devendo apresentar um problema de pesquisa pertinente e criativo. Desta forma, os estudantes não apresentaram o desempenho esperado, identificando muitas dificuldades em suas respostas que reflete, não apenas a falta de conhecimento da ciência,

mas dificuldades básicas na leitura e compreensão das informações que estão sendo lidas (MEC, 2006).

Considerando a importância da leitura e compreensão de textos impressos, e sua contribuição para o desempenho acadêmico de universitários, pretende-se com este estudo, verificar a compreensão que os universitários possuem em dois textos, um de leitura geral, e outro texto técnico-científico. O próximo capítulo expõe sobre leitura e compreensão de textos e pesquisas desenvolvidas nesta área.

LEITURA E COMPREENSÃO DE TEXTOS

A leitura faz parte do conjunto de atividades da vida diária das pessoas, sendo importante para o desenvolvimento e ação do homem moderno, pois é uma das maneiras de obtenção do conhecimento, e, assim, possibilita a reflexão e a opinião sobre diversos aspectos da vida (Silva, 1998). Snowling e Hulme (2005) afirmam que a leitura é o processamento da informação, de uma forma escrita para um significado, conseqüentemente, sob o aspecto cognitivo, esta pode ser vista por meio dos processos mentais associados à transformação.

Segundo Braibant (1997) e Martins e cols. (2003) a leitura pode ser compreendida como uma habilidade, que visa atribuir um significado a um símbolo, e, desta forma, não é apenas a decodificação de registros gráficos, mas a interpretação das letras de modo a compreender a mensagem contida no texto.

Baron (1977) citado por Snowling e Hulme (2005) afirma que existem duas estratégias para a identificação de uma palavra escrita: i) estratégia fonética, que envolve a tradução da palavra de uma representação fonética que é utilizada para buscar o significado da palavra; e, ii) estratégia visual que envolve o uso da informação visual para alcançar o significado, como pode ser visto na Figura 1.

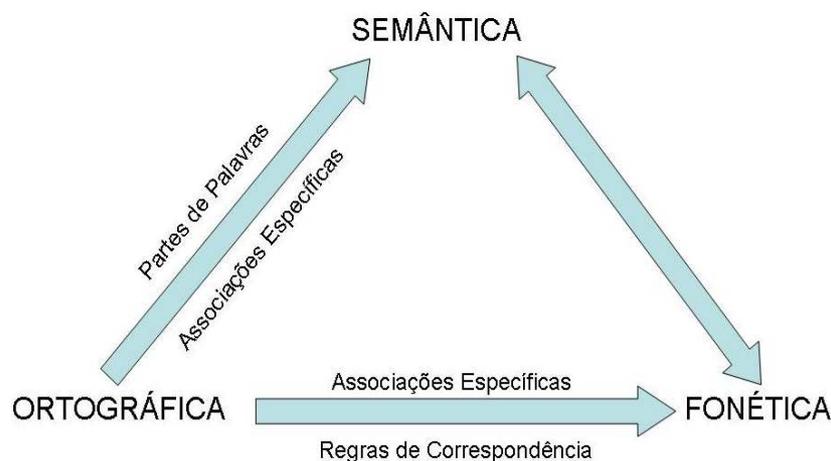


Figura 1. Esquema de Representação da Leitura

Fonte: Snowling e Hulme (2005)

Adicionalmente, a leitura também pode ser compreendida como um processo dinâmico e criativo, com base na relação constituída entre o escritor e o leitor, que se firma por meio do texto, bem como das hipóteses elaboradas pelo próprio leitor. Tais hipóteses são expressas com base na determinação de vinculações intertextuais, ou seja, aquelas que são advindas de conhecimentos adquiridos de leituras anteriores, permitindo assim, uma leitura significativa. Esse processo de relacionamento entre as informações leva à realização de uma análise reflexiva do conteúdo cujas vinculações intertextuais são relacionadas ao indivíduo, à sua estrutura cognitiva, referindo-se tanto ao conhecimento que o sujeito possui quanto da disposição do querer saber (Barbosa, 1992; Martins e cols., 2003; Nassri, 2002; Oliveira, Santos & Primi, 2003).

Segundo Texas Reading Initiative (2006) a leitura envolve: i) interatividade: a relação de trocas entre o texto, o leitor e o contexto; ii) estratégia: necessidades e objetivos de compreensão fazem com que os leitores utilizem diferentes habilidades para construir o significado; e, iii) adaptação: mudanças na estratégia para textos diferentes e diferentes

objetivos de leitura. Para Cappovilla, Joly e Tonelotto (2006), as estratégias de leitura compreendem a análise da estrutura textual, a predição e a adequação do ritmo de leitura à complexidade do texto e da linguagem. Assim, o processo de compreensão em leitura pode ser influenciado pelas características do leitor, tais como, o saber, o conhecimento prévio, o propósito e as metas do leitor, seu nível de motivação, suas crenças e atitudes, e o uso de estratégias (Oliveira, 2003).

Estes aspectos abordam a dinâmica do leitor com o texto e com a sua estrutura cognitiva, contudo, existem outros aspectos relevantes no estudo do desenvolvimento das habilidades de leitura. Salles e Parente (2002) afirmam que a leitura pode ser estudada sob diversos aspectos, tais como, sócio-cultural, afetivo, pedagógico e cognitivo. Diversos autores já realizaram estudos sobre estes diferentes aspectos: Calfee (1977), Gibson e Levin (1975), Perfetti e Lesgold (1978), Pinheiro (1994), Rego (1995) e Smith e Spoehr (1974).

Segundo Salles e Parente (2002) a leitura não envolve apenas o reconhecimento de palavras isoladas, mas a compreensão do material lido, o que requer capacidades cognitivas e lingüísticas, como a elaboração de inferências, conhecimento sobre sintaxe e vocabulário. Para Kintsch (1998) citado por Gabriel (2006), a compreensão textual é a criação de uma estrutura mental que representa o significado da mensagem no texto. Contudo, a estrutura mental pode ser danificada por falta de conhecimento prévio sobre o assunto e falta de familiaridade com a escrita (Gabriel, 2006).

Sob perspectiva da teoria cognitiva de compreensão das informações que são lidas em materiais, são primeiramente extraídas as idéias principais do texto, que são combinadas de forma a elaborar uma representação estruturada e coerente, sendo esta uma vinculação do conhecimento já existente na memória com as novas informações adquiridas. Essas representações são ativadas no momento da leitura, formando-se os esquemas necessários para o entendimento da compreensão na leitura. Dessa forma, a leitura é um processo ativo,

no qual o leitor traz todas as suas experiências para o texto e as utiliza para interpretar e elaborar as novas informações adquiridas (Ellis, 1995).

Assim, uma das principais metas da leitura é a habilidade para compreender a linguagem escrita, envolvendo assim um processo de pensamento que se dá por meio da interação entre o leitor, o texto e o contexto. Para que isto ocorra, o leitor deve estabelecer relações entre os seus conhecimentos prévios e a nova informação cedida pelo texto, assim como fazer inferências, comparações e formulações de questionamentos relacionados ao seu conteúdo. Deve-se levar em consideração o fato de que, quando as relações são estabelecidas, a compreensão do texto é completamente individual, devido à interpretação, que tem um caráter único (Joly, 2001; Alliende & Condemarín, 2005).

Além disso, por fazer parte do processo de compreensão, a estrutura pode ser ou não facilitadora do entendimento da idéia central do texto. Essa estrutura pode ser considerada em termos de *macro*, que se refere aos textos expositivos ou narrativos, ou *micro* estrutura, que avalia os aspectos das sentenças, como os conectivos lingüísticos. (Joly & Paula, 2005; Oliveira, 2003).

Desta forma, um aspecto importante relacionado à leitura é a necessidade de um aprendizado contínuo, pois, a cada leitura o homem torna-se apto no conhecimento de seu mundo e no reconhecimento de novas palavras (Ellis, 1995), transformando-se em um leitor hábil ou proficiente. Gabriel (2006) afirma que um leitor proficiente é aquele para o qual a tradução do código escrito, não é mais um obstáculo e, assim, pode dedicar a atenção ao sentido, e, para o qual a leitura torna-se uma habilidade construída de forma que não são vistas apenas letras, mas sim, significados.

A cada nova leitura que o leitor hábil faz, ele é capaz de atribuir um sentido diferente ao que foi lido anteriormente e, desse modo, quanto mais se lê, mais significativa fica a leitura, e essa significação é resultante da dinâmica social, ativa e crítica. Além disso,

como o nível de escolaridade influencia na habilidade de leitura, um estudante ou profissional que tem essa habilidade desenvolvida, possui um pensamento crítico e criativo com uma maior variedade e qualidade de idéias, a partir de seu próprio contexto e de suas motivações (Witter, 1996, 1999).

Ainda de acordo com esse pressuposto, quando o leitor está no nível universitário, espera-se que possua habilidades próprias de um leitor hábil, sendo capaz de desenvolver as suas próprias estratégias de compreender as informações, captando assim, a micro, macro e superestrutura do texto. Outras habilidades necessárias ao acadêmico, é que possua a habilidade de sintetizar as informações recebidas do texto, e que consiga avaliar a informação recebida, faça inferências e empregue o conhecimento adquirido na resolução de problemas (Joly & Paula, 2005; Vicenteli, 1999; Witter, 1990).

A habilidade de leitura tem uma função importante na vida do sujeito, principalmente no que se refere ao sistema educacional, o qual tem por objetivo a aprendizagem de conceitos específicos por meio de atividades que requerem tal habilidade (Santos, Primi, Taxa & Vendramini, 2002). Destaca-se, ainda, que a leitura é um comportamento essencial para o ensino-aprendizagem nos diversos níveis de ensino, principalmente no ensino superior, já que esse é a última oportunidade para tornar o cidadão um leitor competente que utilize de forma adequada às informações que são adquirias por meio da leitura (Witter, 1999).

A leitura no âmbito acadêmico pode ser vista como uma das maneiras que levam o estudante a ter acesso às informações textuais, destacando-se os informativos das diversas áreas do conhecimento que o universitário utilizará para formar a base de seu conhecimento. A leitura neste nível de escolaridade, na grande maioria dos cursos, é basicamente fundamentada por textos técnico-científicos, que é indispensável ao desempenho profissional de qualquer área (Witter, 1990, 1997, 1999).

Sabendo que a leitura se estabelece a partir da compreensão entre as relações textuais implícitas ou explícitas dos mais diversos tipos de textos, existem alguns recursos utilizados pelos próprios estudantes para obter a compreensão em leitura. Podem-se destacar alguns destes recursos, como os *intratextuais*, que são informações advindas de fontes presentes no texto; *intertextuais*, são informações presentes em textos lidos previamente; e *extratextuais*, aqueles extraídos de situações externas à leitura (Santos & Oliveira, 2004).

Assim, para a aquisição do conhecimento do estudante e do profissional, a leitura tem um papel fundamental, seja ela feita a partir de qualquer tipo de texto, seja de interesse geral ou de interesse específico. A leitura também se mostra como facilitadora para a participação na sociedade e para a integração com o meio, uma vez que possibilita o desenvolvimento de posturas argumentativas e críticas além de promover aos futuros profissionais condições para a tomada de decisões em sua prática profissional (Joly & Paula, 2005).

Desta forma, a compreensão em leitura está diretamente envolvida ao processo que implica tanto a apreensão, quanto a abstração do significado do que se lê, de forma a compreender a idéia principal do texto trazida pelo autor, promovendo assim, a obtenção de uma nova informação. No atual contexto de desenvolvimento, o universitário deve pesquisar e acessar informações de uma maneira contínua, uma vez que, de maneira geral, o processo de produção do conhecimento se dá de uma maneira rápida e difusa com o constante avanço das ciências. Porém, segundo Behrens (2001), o estudante que não tem o hábito de estar atento às inovações e às mudanças nas informações, tem a propensão a perder espaço profissional para os que são mais estudiosos e criativos.

Os estudos realizados por Witter (1999) indicam que, a habilidade de leitura dos acadêmicos, não estão sendo efetivamente desenvolvidas, o que resulta em leitores não

críticos, com poucas habilidades argumentativas e baixo repertório para a resolução de problemas. De acordo com Oliveira (2003), seria necessário permitir o desenvolvimento da habilidade em leitura nos universitários de forma a consentir a obtenção de comportamentos e valores para a crítica, argumentação, questionamento e tomada de decisões.

Algumas pesquisas realizadas em estudantes brasileiros revelam o hábito limitado de leitura dos universitários e as deficiências na compreensão do que é lido, sendo que esses indícios estão diretamente relacionados com o baixo desempenho acadêmico. Em contraste, o processo de escolarização no qual o universitário está inserido, há uma grande exigência em atividades que envolvam a leitura, bem como um nível mais alto de compreensão e apresentação dos conteúdos complexos, quando comparados com outras etapas escolares (Joly & Paula, 2005; Oliveira, 2003; Oliveira e colaboradores, 2003; Pellegrini, 1996; Santos, Vendramini, Suehiro & Santos, 2006; Silva & Santos, 2004; Vicentelli, 1999; entre outros).

Dentre as pesquisas destacadas por Witter (1997), a de Gonçalves, Garcia e Pereira avaliou a compreensão de leitura em 50 estudantes universitários do curso de pedagogia, de uma universidade da cidade de São Paulo. Foi aplicada a técnica *Cloze* em dois textos, um de abrangência geral, e o outro específico da área, preparados de forma tradicional com 40 lacunas. As autoras apresentaram como resultado principal as dificuldades dos estudantes de compreensão em leitura, mostrando claramente o despreparo dos acadêmicos no que se refere à leitura técnica, específica de um assunto pertinente ao curso.

Dada a importância da capacidade dos estudantes em compreender o que é lido no processo de aprendizagem, nota-se a necessidade de avaliar o nível de compreensão do leitor a fim de identificar como este entende o significado do conteúdo que está sendo lido. Quando há compreensão, o leitor é capaz de julgar, criticar, integrar, transformar e aplicar,

com base nos conhecimentos prévios adquiridos anteriormente, juntamente com a abstração do objetivo do texto lido (Joly, 2001). Desta forma, uma das avaliações que tem sido muito realizada é a aplicação da técnica de cloze, que é consistente com os objetivos das avaliações em compreensão em leitura, tanto em termos teóricos, compreender as habilidades necessárias, quanto na prática, para o uso apropriado dos escores obtidos por um instrumento que possui qualidades psicométricas (Abraham & Chapelle, 1992).

A TÉCNICA DE CLOZE

Destacada a importância do papel da leitura para o estudante universitário, percebe-se a necessidade de avaliar o desempenho na compreensão do material que é lido. Essa avaliação permite verificar como se dá o processo de leitura do estudante de forma a: estimar o nível de leitura que possui; determinar quais habilidades específicas devem aprender; identificar as dificuldades na leitura; desenvolver estratégias apropriadas para lidar com a dificuldade detectada; ajudar o estudante a conhecer suas próprias limitações; e avaliar o progresso obtido (Alliende & Condemarín, 2005; Joly, 1999; Santos & Di Nucci, 1984).

Sob esta perspectiva, considera-se fundamental o diagnóstico da compreensão em leitura dos acadêmicos, já que é principalmente por meio da leitura que ocorre o acesso ao conteúdo técnico-científico das diversas disciplinas. Destaca-se a importância, de primeiramente se realizar um diagnóstico nos universitários, para que posteriormente, possa realizar atividades preventivas e corretivas em função da deficiência detectada, a fim de possibilitar não apenas o desenvolvimento cognitivo, como também o desejável desenvolvimento pessoal e profissional do estudante (Witter, 1997). O diagnóstico das dificuldades de compreensão em leitura é feito por instrumentos de avaliação, como por

exemplo, o teste de cloze, o qual é comumente utilizado, pois pode ser aplicado em todas as etapas de escolarização, desde o fundamental até o nível superior.

A técnica de cloze foi desenvolvida por W. L. Taylor em 1953, que consiste em omitir palavras de um texto, de forma regular, substituindo-as por espaços em branco (Joly & Paula, 2005). Diversas formas são usadas para mensurar a compreensão em leitura, contudo de acordo com Ashby-Davis (1985), pode-se dividir em duas categorias principais, o cloze exato e o cloze modificado. O cloze exato conforme a autora designa como sendo o modelo original desenvolvido por Taylor.

Esse modelo consiste em um texto de 250 palavras que é modificado por omissões em intervalos regulares, suprimindo-se uma a cada cinco, oito ou dez palavras, as omissões são substituídas por linhas, todas do mesmo tamanho. A primeira e última sentenças do texto permanecem intactas. Somente as palavras exatas, iguais as do autor são consideradas como corretas (Alliende & Condemarín, 2005; Ashby-Davis, 1985; Bormuth, 1968; Joly & Paula, 2005; Taylor, 1953). Já o cloze modificado refere-se a qualquer outra técnica de cloze, que pode ser muito diversificada.

Dentre as técnicas modificadas Schoenfeld (1980) sugere omitir as palavras do texto de acordo com alguns critérios, como a supressão em função de determinadas classes gramaticais, como substantivos, verbos, adjetivos e advérbios. Nestes casos, dá-se ao estudante a oportunidade de aprender dedutivamente a função de cada classe gramatical, e de observar o papel significativo que a linguagem descritiva desempenha na modificação do significado de um trecho. Outros procedimentos destacados por Alliende e Condemarín (2005) podem ser adotados, são eles: *lista de palavras*, *indícios fônicos* e *múltipla escolha*, que serão melhores detalhados no próximo parágrafo.

No que se refere à *lista de palavras* as palavras omitidas são dispostas de acordo com a ordem alfabética na margem da folha, e uma maneira de diferenciar esta modalidade

é modificando-a em função de aumentar o nível de dificuldade é a inclusão na lista, vocábulos inadequados. Para os *índices fônicos*, a letra inicial, dígrafo ou a letra final podem apresentadas em cada omissão, porém as palavras são substituídas por uma linha de extensão constante. Com relação à modalidade *múltipla escolha*, são oferecidas três, quatro ou mais palavras que podem ser colocadas na margem próximas ao espaço em branco. Mesmo quando todas elas pertencem à mesma classe gramatical, apenas uma alternativa será semântica e sintaticamente correta (Alliende & Condemarín, 2005). Esta técnica vem sendo aplicada em alguns instrumentos utilizados por Joly (2006) em suas pesquisas, tanto com alunos das séries iniciais do ensino fundamental, quanto para os estudantes do ensino médio.

O teste de Cloze avalia o nível funcional da compreensão da leitura, como um instrumento específico surgido do próprio texto. A vantagem como meio de avaliação da compreensão da leitura, pode ser definida pelo fato do leitor contar com o contexto como único apoio, e não correr o risco de que ele use palavras chaves da pergunta para adivinhar a resposta, evitando a utilização de perguntas cuja compreensão pode ser mais difícil do que a própria leitura. A técnica serve ainda como um meio de desenvolvimento da compreensão da leitura, ao permitir ao estudante descobrir as palavras que se encontram omitidas no texto, graças ao seu domínio das estruturas semântico-sintáticas da linguagem e de seus conhecimentos anteriores acerca do conteúdo do texto (Joly, 2006).

Assim, somente os bons leitores irão conseguir preencher adequadamente as omissões e não dependerão da subjetividade do avaliador em concordar ou não com a interpretação do leitor, recomenda-se que se dê preferência a utilização da forma de correção literal, para o teste Cloze (Bitar, 1989).

Quanto à forma de correção, existem dois métodos de atribuição das pontuações às respostas corretas que são empregados. A primeira é conhecida como correção literal, no

qual se aceita como resposta exata apenas a palavra idêntica à que foi suprimida do texto. A segunda é denominada como correção sinônima, em que se aceita uma palavra que dê sentido equivalente ou uma outra resposta contextualmente aceitável como correta (Brown, 1980; Santos & Oliveira, 2004).

Outras vantagens dessa técnica são, a facilidade na sua elaboração, administração e interpretação, assim, não é necessário ser um especialista para utilizá-la. O teste de Cloze é uma técnica que pode ser utilizada no diagnóstico da compreensão em leitura, em textos de conteúdos diversos, de caráter geral ou específico de qualquer área de conhecimento (Alliende & Condemarín, 2005; Kletzein, 1991).

Essa técnica vem sendo amplamente utilizada para o diagnóstico de compreensão em leitura de textos técnicos de diversas áreas científicas, como biologia, contabilidade, física, psicologia entre outros (Centofanti e cols., 1997; Castro, Fukuda & Gritti, 1997; Costa, 2006; Martins & Andrade, 2004; Silva, Malozze & Leme, 1997). A partir dos estudos mencionados, percebe-se a necessidade de novas pesquisas com estudantes de psicologia, em temas mais específicos, no que se refere a leitura de textos da área psicológica, mais precisamente aqueles que envolvem conceitos estatísticos. Esses conceitos estão presentes em vários textos, sejam de leitura científica ou de atuação prática do psicólogo, como os manuais de testes psicológicos, tornando-se relevante uma vez que são instrumentos utilizados na prática e atuação profissional. Assim, faz-se necessário relatar uma breve explanação sobre conceitos estatísticos.

OS CONCEITOS ESTATÍSTICOS

O conteúdo específico a ser estudado pela pesquisa em questão, é o de Estatística. Assim, pode-se compreender a leitura Estatística como a habilidade de um sujeito em

interpretar criticamente a informação estatística, os argumentos estatísticos ou o fenômeno estocástico, que são encontrados em diversos contextos que os dados estatísticos são utilizados (Gal, 2002). A compreensão em estatística é a base para muitas das decisões tomadas na sociedade atual, os aspectos da literatura de estatística (que se refere à aplicação da compreensão estatística em um contexto). Desta forma, a literatura estatística não se torna importante apenas para a sociedade em geral, mas também para cada sujeito em particular, como parte da sociedade. Pois, as pessoas tomam decisões em suas vidas pessoais baseadas em análises de informações e riscos fornecidos por outros membros da comunidade (Watson & Callingham, 2003).

A literatura em estatística pode ser entendida como a habilidade de compreender e avaliar criticamente os resultados estatísticos que estão presentes na vida diária das pessoas, assim como uma habilidade em apreciar as contribuições que o pensamento estatístico pode proporcionar nas decisões da vida, tanto pessoais, quanto profissionais (Watson & Callingham, 2003). De acordo com Quintino, Guedes e Martins (2001) a Estatística é uma ciência importante para diversas áreas do conhecimento e por isso ela está presente nos cursos de graduação. A sua importância se torna maior principalmente quando se trata de atividades de pesquisa. A utilização de métodos estatísticos vem se expandindo entre os acadêmicos e entre os pesquisadores de todas as áreas de estudo. Ressaltando que a compreensão em estatística pode ser entendida como uma intersecção entre a literatura e a habilidade para aritmética, sendo essencial para o alcance dos objetivos das novas grades curriculares dos cursos de graduação (Watson & Callingham, 2003).

No que se refere à escrita matemática e estatística, há uma especificidade, uma característica própria que envolve uma combinação de sinais, letras e palavras que se organizam de acordo com regras para expressar idéias. A organização desta escrita aritmética é diferenciada, acabando por exigir um processo particular de leitura que

diferencia de outros textos, como por exemplo, os livros de leitura e de interesse geral (Alliende & Condemarín, 2005; Smole & Diniz, 2001).

O domínio da leitura é um fator importante para o sucesso no estudo em qualquer área, assim como para a matemática e para a estatística. Considerando como parte importante do vocabulário geral e das comunicações do dia-a-dia a linguagem matemática e estatística, o estudante pode compreendê-las à medida que entende a linguagem empregada pelas mesmas. Além do vocabulário especializado, o universitário tem que se familiarizar com alguns procedimentos que o levam a considerações lógicas dos elementos que contribuem para a solução de problemas (Alliende & Condemarin, 1987).

Ainda segundo os autores, o nível de compreensibilidade dos textos utilizados para o ensino da aritmética, favorece o sucesso na aprendizagem da disciplina, já que frequentemente o texto constitui a fonte mais importante para a instrução. Geralmente estes textos exigem uma leitura cuidadosa e exata, porque se caracterizam pela abstração, concisão e por estabelecer relações complexas. Assim, textos de cada disciplina exigem tipos especiais de leitura, uma vez que, os textos de matemática e estatística caracterizam-se pela grande quantidade de informação numérica e pelos conteúdos verbais concisos. Já, os textos literários podem ser considerados como construtores de linguagem, quando comparadas com os textos de compreensões textuais específicas.

Uma forma freqüente de representação de dados em artigos científicos ou publicações de abrangência geral são os diagramas, quadros, tabelas e gráficos estatísticos. A leitura e compreensão destas informações exigem do leitor habilidades essenciais ao processo de ler qualquer tipo de textos, além de desenvolver habilidades de questionar, levantar hipóteses, bem como procurar relações entre os dados (Smole & Diniz, 2001). Assim, não é suficiente que o indivíduo desenvolva apenas a capacidade de organização e representação de dados, mas sim uma compreensão das mensagens estatísticas, que requer a

utilização de diversas habilidades na interpretação da leitura dos dados, para obter significado na informação apresentada ao leitor (Bifi, 2006).

Segundo Smole e Diniz (2001), a dificuldade que alguns estudantes apresentam em estatística, pode ser decorrência do fato de que há uma falta de compreensão de um conceito envolvido na escrita da matemática e da estatística, no uso de termos específicos que não fazem parte do cotidiano do aluno. Utilizando palavras em que se atribuem significados diferentes ao cotidiano do indivíduo, a saber, diferença, ímpar, média, probabilidade, entre outras. Além disso, alguns alunos apresentam uma barreira inicial quanto à disciplina Estatística, conforme relatado por Vendramini e Brito (2001), ao considerar que o desempenho acadêmico apresentado, pode estar relacionado a experiências anteriores de aprendizagem. Segundo as autoras, estudantes das mais variadas áreas de conhecimento enfrentam problemas para aprender Estatística por esta lidar com conceitos abstratos, usar notações e terminologias complexas e muitas vezes ambíguas.

Batanero (2001) têm revelado que, mesmo para utilização das ferramentas de análise de dados mais simples, os estudantes mostram ter dificuldades. É o caso das representações em tabelas e gráficos e manuseio de frequências relativas, que envolvem as aplicações matemáticas com números racionais e conceitos relacionados a medidas. Percebe-se que o aluno, durante a sua formação, depara-se com algumas situações-problema, uma vez que, a maioria dos livros utilizados nos cursos para a aprendizagem dos estudantes é organizada didaticamente para o desenvolvimento de um programa, mas não necessariamente, adequada à realidade profissional que os universitários enfrentarão após o término do curso.

Nos cursos de graduação em Ciências Humanas, deve-se enfatizar a importância da disciplina Estatística na análise e interpretação de dados, na ajuda de decisões tomadas na prática profissional (Bifi, 2006). Segundo Gal e Garfield (1997) a instituição deve

desenvolver no estudante as capacidades de compreender e lidar com incerteza, variabilidade e informação estatística; assim como, contribuir ou tomar parte na produção, interpretação e comunicação de dados pertencentes a problemas que ele deve encontrar na sua vida profissional.

O ensino da Estatística ocorre desde o início do ensino fundamental até o superior, neste caso, presente como disciplina na maioria dos cursos de graduação. Na Psicologia, a Estatística é uma ferramenta de análise de dados principalmente no que se refere à avaliação psicológica que necessita de instrumentos de avaliação válidos e fidedignos. Para a construção, avaliação ou utilização desses instrumentos é necessário saber alguns conceitos e técnicas estatísticas.

Alguns dos conceitos estatísticos, que são amplamente utilizados no contexto científico, estão presentes também, em manuais de testes psicológicos muito utilizados na prática da avaliação psicológica por estudantes e profissionais da área, quais sejam os conceitos de média, desvio padrão, percentil, entre outros. Como destaca Anastasi e Urbina (2000), é necessário que um examinador qualificado saiba utilizar um teste psicológico considerando três aspectos principais da situação de testagem, a seleção do teste, a aplicação e pontuação e a interpretação dos escores, aspectos estes que incluem informações estatísticas também presentes na verificação e apresentação das qualidades psicométricas dos instrumentos.

A Associação Psicológica Americana (APA) vem estabelecendo padrões, princípios para a utilização de testes educacionais e diretrizes para o uso competente e ético dos testes psicológicos (Noronha & Vendramini, 2003). Destacando-se desta maneira, a importância dos conhecimentos de tais conceitos, pois os psicólogos, desde a graduação estão em contato com informações estatísticas presentes nas diversas fontes de informação, para atualização e obtenção de conhecimentos específicos, bem como para a aprendizagem dos

conhecimentos científicos para a sua atuação profissional. Dada à importância, observou-se a necessidade de destacar as pesquisas e estudos realizados nesta área, que serão melhores descritas a seguir.

PESQUISAS RELACIONADAS À COMPREENSÃO EM LEITURA

O estudo de Santos e colaboradores (2006) comparou o nível de compreensão em leitura percebido por alunos de Psicologia, as estratégias de aprendizagem utilizadas na superação dos problemas encontrados, além de verificar a real dificuldade de compreensão dos alunos para responder um teste de Cloze, em 178 universitários, sendo 113 do 1º ano e 65 do 4º. O instrumento foi aplicado numa única sessão, consistiu na seguinte forma, primeiramente foi realizada a leitura de um texto científico da área, em seguida, questões sobre as dificuldades percebidas e sobre as estratégias utilizadas para a sua superação e por fim, a aplicação de um teste de Cloze em texto similar ao primeiro, para a avaliação da compreensão em leitura. Os resultados indicaram diferença estatisticamente significativa no nível de compreensão em leitura com superioridade dos alunos do 4º ano em relação aos do 1º. Contudo os alunos do 4º ano perceberam-se com maior dificuldade na compreensão do primeiro texto, embora a diferença entre as turmas não tenha sido significativa. Os tipos de dificuldades de compreensão apontadas e as estratégias utilizadas pelos alunos apresentaram diferenças qualitativas entre as turmas.

Marini (2006) objetivou buscar evidências de validade para uma escala de estratégias metacognitivas de leitura. Participaram do estudo 641 estudantes do ensino médio de escolas particulares e públicas, com idade entre 14 e 17 anos. Os instrumentos utilizados foram a Escala de Estratégias Metacognitivas de Leitura – Ensino Médio e Prova de Compreensão em Leitura elaborada pela técnica de Cloze. Como medida de rendimento acadêmico, foi utilizada a média escolar do participante em Língua Portuguesa. Os resultados revelaram maior uso das estratégias de solução de problemas. Com relação ao momento no qual, os estudantes revelaram usar mais estratégias, foi durante a leitura.

Verificou-se que há diferenças quanto ao gênero, tipo de escola, série e turno frequentado. Foi realizada também uma análise dos itens da escala com a Prova de Compreensão em Leitura, identificando-se correlação significativa para 48,72% dos itens. Quanto ao rendimento acadêmico, a correlação foi positiva e significativa para 18% dos itens.

Joly e Paula (2005) realizaram um estudo com 201 estudantes, com o objetivo de verificar o rendimento em compreensão em leitura de universitários ingressantes, identificar a frequência de uso das estratégias de aprendizagem e analisar a existência de relação entre as estratégias de aprendizagem utilizadas e o rendimento em compreensão em leitura. Os resultados indicaram que o rendimento em compreensão de 69% dos universitários da amostra ficou no nível de frustração. Os resultados não evidenciaram correlações entre o uso de estratégias de aprendizagem e a compreensão em leitura.

Silva e Santos (2004) avaliaram a compreensão em leitura de universitários e sua relação com o rendimento no vestibular na prova de língua portuguesa e o rendimento médio dos estudantes no primeiro ano. Foi aplicado um teste elaborado a partir da Técnica de Cloze em 782 ingressantes de oito cursos de uma universidade particular do interior do estado de São Paulo. As comparações realizadas entre rendimento médio e os escores do Cloze apontaram um índice de correlação positiva e significativa entre as duas variáveis. Dentre os principais resultados destacaram-se a baixa pontuação obtida no Cloze, indicando que os estudantes apresentam um nível de compreensão aquém do esperado para essa etapa de escolaridade. Assim, foi possível concluir que o teste de Cloze mostrou-se um instrumento adequado para avaliar a compreensão em leitura e que apresentava evidências de validade concorrente, tendo como critério o rendimento acadêmico.

O estudo de Saito (2003) objetivou avaliar a validade de construto do teste de Cloze no Exame para obter o Certificado de Proficiência em Inglês (ECPE), avaliado pelo Instituto de Língua Inglesa na Universidade de Michigan. Por meio de uma rigorosa

investigação utilizando o Modelo de Equação Estrutural (SEM) como um procedimento estatístico primário, o modelo composto por dois fatores, sendo estes, a habilidade léxico-gramatical e habilidade em leitura, que foram confirmadas e assumiu-se como a melhor representação dos dados. Além disso, este estudo demonstrou que o teste de cloze na prova do ECPE como medida de gramática. Em outras palavras, os itens do Cloze apresentam-se como uma medida do conhecimento da gramática nas sentenças e níveis de sentenças da proficiência na língua. Este estudo também demonstrou a utilidade do modelo de equação estrutural como medida de validade de construto.

Oliveira (2003) realizou uma pesquisa com 270 estudantes universitários ingressantes dos cursos de Administração, Direito e Psicologia de uma universidade particular. Teve como objetivo verificar a relação entre compreensão em leitura, o rendimento acadêmico e a avaliação da aprendizagem. Foram utilizados dois textos elaborados com a técnica de Cloze, as notas escolares e um questionário para caracterizar os meios de avaliação mais frequentemente utilizados, na opinião dos estudantes. Os resultados demonstraram correlação significativa entre o escore do Cloze e o rendimento acadêmico, confirmando pesquisas anteriores. Quanto ao tipo de avaliação mais usada pelos professores, os resultados evidenciaram a prova dissertativa individual como sendo a mais utilizada e os trabalhos dissertativos em grupo como a melhor estratégia a ser usada sob o ponto de vista do estudante.

Com o objetivo de verificar a relação entre compreensão em leitura e rendimento acadêmico, Oliveira e colaboradores (2003) realizaram uma pesquisa com 412 universitários do primeiro ano dos cursos de Letras, Psicologia, Matemática e Odontologia. O instrumento utilizado foi o Teste de Cloze, que teve os escores correlacionados com as notas médias das disciplinas cursadas pelos estudantes no primeiro semestre. À exceção do

curso de matemática, em todos os demais cursos os resultados obtidos revelaram uma associação significativa entre compreensão e leitura e rendimento acadêmico.

No intuito de avaliar a aplicação de um programa de intervenção em leitura e redação, Sampaio e Santos (2002) realizaram o estudo com 42 estudantes ingressantes de dois cursos na área de negócios de uma universidade particular, do período noturno. Os resultados indicaram que as diferenças de desempenho não foram estatisticamente significativas entre o pré e pós-teste, mas que houve mudanças qualitativas nas atitudes dos universitários em relação aos temas.

Santos e colaboradores (2002) conduziram um estudo com 612 estudantes ingressantes nos cursos de Medicina, Odontologia, Administração e Psicologia de uma universidade particular, tendo como objetivo investigar a aplicação da Teoria de Resposta ao Item na análise de um instrumento construído com a técnica de Cloze para avaliar a compreensão em leitura. Os resultados demonstraram, que a habilidade de compreensão de textos da maneira como é exigida no teste de Cloze, depende de inúmeras variáveis como recuperação da palavra a partir da estrutura sintática das frases nas quais a palavra omitida é inserida, da competência do leitor em estabelecer relações entre os elementos do texto e de sua capacidade em desenvolver associações apropriadas entre o conhecimento prévio e a informação trazida pelo texto. A análise estatística por meio da TRI demonstrou que a escala atingiu um alto índice de consistência interna, embora algumas lacunas não tenham apresentado parâmetros psicométricos aceitáveis. Identificou-se que algumas lacunas apresentaram grau de dificuldade muito elevado e baixa discriminação. O estudo reafirmou a boa qualidade do Cloze como instrumento da avaliação da compreensão em leitura.

Com o objetivo de identificar o uso das estratégias de compreensão m leitura de textos em espanhol e comprar, mediante a Técnica Cloze, com o rendimento em leitura em espanhol básico, Santos (2001) avaliou estudantes do curso de Turismo. Os dados foram

coletados em dois momentos, um pré-teste, em que foi aplicado um questionário para levantamento das estratégias de leitura utilizadas e um teste de Cloze para medir o rendimento acadêmico de compreensão de leitura, no início do semestre; e um pós-teste, sendo aplicado um novo questionário e um segundo teste de cloze no final do semestre letivo. Com base nos resultados, no pré-teste os estudantes indicaram satisfatoriamente o uso das estratégias de compreensão em leitura, contudo, apresentaram um nível de rendimento baixo no conhecimento do idioma espanhol. No pós-teste, verificou-se que os estudantes não sabem que lêem com estratégias de compreensão e não mudaram as estratégias utilizadas. Quanto ao teste de Cloze, houve uma melhora significativa no rendimento superior aos estudantes do diurno. Pelos dados obtidos, verificou-se que a técnica de Cloze mostrou-se um instrumento eficaz para avaliar o rendimento em leitura de uma língua estrangeira.

Vicentelli (1999) aponta a deficiência na compreensão da leitura pelos estudantes do ensino superior. Assim, o objetivo do estudo foi avaliar o nível de compreensão de 115 universitários dos cursos de educação física, matemática, espanhol e filosofia, pertencentes a uma universidade pública da Venezuela. Foi utilizado um instrumento, no qual foi aplicada a técnica de Cloze (modelo original) e utilizada classificação da compreensão dos sujeitos em função da Escala de Bormuth (1968). Os resultados indicam deficiências importantes na compreensão da leitura, dos sujeitos avaliados: 80% dos estudantes de educação física e 75,9% dos alunos de castelhano se encontraram em um nível de frustração, enquanto que os estudantes de filosofia mostraram melhor desempenho na classificação somente neste nível com 52,5%.

Em seu estudo, Pellegrini (1996) identificou o nível de compreensão e as atitudes frente à leitura em estudantes de 3º grau, por meio de comparações entre 100 estudantes iniciantes e concluintes de dois cursos de uma universidade do interior do estado de São

Paulo, a saber, Engenharia Industrial (Mecânica) e Psicologia. Os instrumentos utilizados para a coleta de dados foram, Questionário Informativo, texto programado em Cloze e Escala de Atitudes em Leitura. Com relação à compreensão em leitura, os resultados indicaram que a maioria dos sujeitos (76%) apresenta sérias dificuldades, não havendo diferenças significantes entre as séries e cursos. Com relação às atitudes frente à leitura, observou-se que os estudantes de Psicologia apresentam uma média significativamente superior nas atitudes positivas, do que os sujeitos do curso de Engenharia Industrial, não havendo diferenças entre as séries do mesmo curso.

Dada a necessidade de se utilizar instrumentos válidos e fidedignos para inferir com segurança sobre os construtos estudados, este estudo objetivou verificar as qualidades psicométricas da prova de compreensão em leitura, de um texto de leitura específica. Assim, pretende-se utilizar a validade de construto entre as provas de compreensão em leitura, de um texto de leitura geral, com um texto de leitura específica. Com essas análises, além de se obter as evidências de validade da prova de compreensão em leitura, será verificado, se o estudante que apresenta boa compreensão em leitura de um texto de interesse geral, também apresenta um bom desempenho na prova do texto de leitura específica. Espera-se que este estudo possa auxiliar na elucidação sobre possíveis fatores que interferem no processo de aprendizagem de conceitos estatísticos necessários para a pesquisa e prática em psicologia.

OBJETIVOS

OBJETIVO GERAL

Esta pesquisa teve como objetivo verificar evidências de validade de uma prova de compreensão em leitura de um texto com informações estatísticas em estudantes de Psicologia.

OBJETIVOS ESPECÍFICOS

- ✓ Verificar evidências de validade de construto a partir da análise fatorial com informação completa baseada na Teoria de Resposta ao Item.
- ✓ Buscar evidências de validade convergente entre compreensão em leitura de um texto com informações estatísticas e um texto de interesse geral.
- ✓ Buscar evidências de validade convergente entre compreensão em leitura de um texto com informações estatísticas com o questionário de estatística.
- ✓ Examinar a precisão da prova de compreensão em leitura específica.
- ✓ Analisar o desempenho nas provas em compreensão em leitura, com as variáveis de interesse, como ano de graduação, sexo, idade e turno.

MÉTODO

PARTICIPANTES

Foram participantes desta pesquisa 206 estudantes do curso de Psicologia de uma universidade do interior do Estado de São Paulo. A amostra foi composta por conveniência, por universitários regularmente matriculados no curso de Psicologia, sendo 79,1% do gênero feminino, dos turnos diurno (17,5%) e noturno (82,5%). As idades variaram de 17 a 54 anos, com uma média de 25,92 anos ($DP=6,80$). Optou-se por organizar as idades por faixa etária devido a sua grande variabilidade, formando-se quatro grupos de idade. A primeira faixa etária corresponde aos participantes de 17 a 22 anos, a segunda de 23 a 28, a terceira de 29 a 34, e a quarta por estudantes de 35 anos ou mais, conforme a Tabela 1.

Tabela 1. Distribuição do número de estudantes segundo as faixas etárias

Faixas etárias	Número de estudantes		Porcentagem	
	N	%	Válida	Acumulada
17-22	77	37,4	37,6	37,6
23-28	79	38,3	38,5	76,1
29-34	24	11,7	11,7	87,8
35 ou mais	25	12,1	12,2	100,0
Sub-total	205	99,5	100,0	
Omissão	1	0,5		
Total	206	100,0		

Amostra também foi composta pelos estudantes dos cinco anos da graduação, tanto aqueles que já cursaram as disciplinas de Estatística e Psicometria, como aqueles que não as cumpriram, de modo a analisar a hipótese de que há diferença entre a compreensão dos universitários que cursaram ou não tais disciplinas. Assim, a quantidade de estudantes

distribuídos entre os cinco anos da graduação está mais bem descrita na Tabela 2. Apenas 23,3% dos universitários não realizaram as disciplinas de estatística e psicometria.

Tabela 2. Distribuição do número de estudantes segundo o ano de graduação

Ano da graduação	Número de estudantes	
	N	%
1º	48	23,3
2º	44	21,4
3º	34	16,5
4º	37	18,0
5º	43	20,9
Total	206	100,0

MATERIAL

Questionário

Para a avaliação do conhecimento que o estudante possui de alguns conceitos estatísticos, elaborou-se um questionário (Anexo 3) que contém alguns itens de identificação, para a caracterização do sujeito, e outros itens que correspondem à percepção do universitário frente à disciplina estatística, bem como itens que envolvem diretamente alguns conceitos estatísticos, como probabilidade, porcentagem, média, mediana e moda. Assim, os primeiros itens do questionário envolvem a caracterização do sujeito, como semestre da graduação, período, idade, gênero e reprovação em disciplinas do curso.

Também compõem o questionário itens como: conceitos estatísticos que o estudante julga capaz de interpretar; significado, importância e utilização da Estatística; em seu cotidiano; leitura de materiais que envolvem conceitos estatísticos; e, a porcentagem de professores que incluem a Estatística em suas aulas.

O questionário inclui ainda, questões e problemas de Estatística distribuídos em duas partes, a primeira com quatro itens de múltipla escolha, que envolvem conceitos como média e probabilidade e a segunda com cinco questões abertas envolvendo conceitos de porcentagem, percentil, moda, mediana, média e variabilidade. Os problemas da questão treze deste questionário, referentes aos itens I, II, III, IV, foram elaborados a partir de um instrumento utilizado por Roca (2002) em sua tese de doutorado, sendo traduzidos e adaptados para a amostra brasileira. No que se refere à aplicação, pode ser realizada de forma coletiva. O tempo de aplicação foi de aproximadamente 20 minutos por turma.

É importante destacar que, este questionário foi elaborado para a realização de uma pesquisa maior. Porém, para este estudo utilizou-se apenas as questões que se referem aos conceitos estatísticos, permitindo assim, avaliar o desempenho geral do universitário sobre o conhecimento estatístico.

Provas de Compreensão em Leitura

Para a avaliação da compreensão em leitura, foram utilizados dois textos, um de leitura geral e outro de leitura específica, que inclui conceitos estatísticos. O texto de leitura geral é composto por um texto editorial já utilizado em outras pesquisas, como pode ser citado o estudo realizado por Marini em 2006. O texto específico foi elaborado a partir do manual de teste psicológico, sendo este a Escala de Inteligência Wechsler para Adultos (WAIS-III, Nascimento, 2004), buscando incluir alguns dos conceitos básicos de Estatística que freqüentemente estão presentes nos manuais de testes psicológicos. Em ambos os textos foram aplicados o Sistema Orientado de Cloze (SOC) de Joly (2006).

O Sistema Orientado de Cloze (Joly, 2006) é baseado na Técnica de Cloze descrita por Taylor (1953), porém com alternativas de múltipla escolha. As omissões foram substituídas por lacunas, sendo todas do mesmo tamanho. Estas são identificadas por um

número, e cada lacuna enumerada possui cinco alternativas de respostas, que são dispostas em um quadro ao lado direito do texto (Anexo 4 e 5).

O primeiro parágrafo foi mantido em sua forma original, e o apagamento iniciou-se a partir do quinto vocábulo do parágrafo seguinte. Após a quadragésima lacuna, o texto manteve sua forma original, com aproximadamente 300 palavras. Assim, foram deixadas 40 lacunas, passíveis de preenchimento.

Os sujeitos escolhem entre as cinco alternativas de múltipla escolha, a palavra que preenche corretamente cada lacuna, dando sentido à frase. As alternativas de múltipla escolha possuem uma única resposta correta, uma palavra da mesma classe gramatical da palavra que foi omitida, duas palavras relacionadas à classe gramatical e uma palavra de classe gramatical diferente da omitida. São consideradas corretas as palavras que completam a lacuna de forma exatamente igual ao texto original. As respostas em branco são computadas como erro, para cada acerto é atribuído um ponto. O rendimento dos sujeitos é registrado considerando acertos e erros. As orientações para o participante estão descritas no instrumento, antecedendo o texto. O tempo de aplicação médio foi de 40 minutos.

Prova de compreensão em leitura geral (Joly, 2005 – pesquisa em desenvolvimento)

Foi realizado um estudo de validade do instrumento com estudantes do Ensino Médio (Joly, 2005), sendo 704 estudantes regularmente matriculados, 53,8% do sexo feminino e 46,2% do sexo masculino de escolas públicas e particulares, com idades variando de 14 a 23 anos. O estudo revelou um índice de fidedignidade global de $\alpha = 0,90$. Outros estudos com este instrumento ainda estão sendo realizados.

Uma segunda análise realizada por Joly e Dias (2007a; 2007b), evidenciou um bom coeficiente de Kuder-Richardson com um valor igual a 0,91. Para a validade da prova, foi

realizada a análise fatorial, que indicou uma predominância unidimensional, com uma correlação tetracórica moderada ($r_{tet} = 0,44$) e cargas fatoriais superiores a 0,395.

Prova de compreensão em leitura específica

Esta prova de compreensão foi construída especificamente para este estudo e pretendeu-se buscar as qualidades psicométricas desta prova de leitura específica. A escolha do texto foi em função de um material que o estudante tem contato, e que inclui dados estatísticos que é o objetivo do trabalho.

Desta forma, a prova é composta por uma seleção de algumas partes do manual do teste WAIS (Nascimento, 2004), formando um texto coeso. As partes do manual foram escolhidas de forma a criar um texto que permita ao estudante o entendimento do objetivo do instrumento selecionado, para aqueles que não o conhecem. Assim, o texto é composto por um parágrafo de esclarecimentos gerais sobre o instrumento para dar seqüência ao parágrafo que contém informações estatísticas que é o foco do trabalho, destacando a forma de avaliação do teste.

PROCEDIMENTO

Após a aprovação do projeto pelo Comitê de Ética em Pesquisa da Universidade São Francisco, foi feito o contato com a universidade para a aplicação dos instrumentos. Com a autorização dos professores para a realização da pesquisa em todas as turmas já selecionadas e autorizadas, foram aplicadas as provas em horários pré-estabelecidos.

Aplicou-se coletivamente em uma única sessão as duas provas de compreensão em leitura, na própria universidade, nas salas de aulas com no máximo 40 estudantes. O tempo de aplicação total dos três instrumentos foi de aproximadamente uma hora e quarenta

minutos. Primeiramente, os participantes foram convidados a lerem e assinarem o Termo de Consentimento Livre e Esclarecido (TCLE). Foram incluídos no grupo de participantes apenas aqueles que aceitaram participar da pesquisa.

A aplicação foi realizada em duas ordens de distribuição dos instrumentos de forma que fossem distribuídas quantidades iguais das duas ordens. A primeira obedeceu a seguinte ordem de aplicação, Termo de Consentimento Livre e Esclarecido (TCLE) e o questionário de estatística, a prova de compreensão em leitura do texto de leitura geral (Cloze 1), seguida da prova de compreensão em leitura do texto de leitura específica (Cloze 2). A segunda ordem de aplicação iniciou-se com a distribuição do TCLE, do questionário, seguido do Cloze 2 e do Cloze 1.

As instruções de preenchimento das provas de compreensão em leitura foram lidas pelo aplicador para os sujeitos (Anexo 3 e 4). As dúvidas que ocorreram foram esclarecidas e após, iniciou-se a aplicação das provas. Foi recomendado que primeiramente o participante realizasse uma leitura silenciosa do texto, e após retornasse ao início do texto e começasse o preenchimento das lacunas.

PROCEDIMENTO DE ANÁLISE DE DADOS

A evidência de validade de construto foi obtida por meio da análise da dimensionalidade da prova de compreensão em leitura que envolve conceitos estatísticos, sendo verificada pela análise fatorial com informação completa, que é baseada na Teoria de Resposta ao Item. Também foram verificadas evidências de validade convergente por meio de correlações entre os resultados obtidos nas provas de compreensão em leitura e com o desempenho no questionário de conteúdo estatístico. Verificou-se a fidedignidade da prova por meio do coeficiente de Kuder-Richardson. Uma análise exploratória das variáveis de interesse, abordadas pelas questões de identificação foi realizada, com o auxílio da estatística descritiva.

RESULTADOS E DISCUSSÃO

Inicialmente foi realizada uma análise descritiva para os desempenhos no questionário de conceitos estatísticos e nas provas de compreensão em leitura. Posteriormente, foi realizada uma análise dos itens a fim de buscar as evidências de validade da prova de compreensão em leitura específica, além da confiabilidade da mesma, bem como, evidências de validade para os estudantes do ensino superior para a prova de compreensão de um texto geral. As análises descritivas foram realizadas com o auxílio do programa estatístico SPSS for *Windows* (versão 13) e as evidências de validade foram verificadas pela Teoria Clássica dos Testes (com o auxílio do programa SPSS for *Windows*) e pela Teoria de Resposta ao Item (com o auxílio dos programas XCALIBRE e WINSTEPS).

ANÁLISE DESCRITIVA DOS RESULTADOS

Questionário de conceitos de estatística

Com a finalidade de verificar se o conhecimento de conceitos básicos de estatística contribui para uma melhor compreensão de textos que possuem esses conceitos, examinou-se apenas a parte do questionário geral de interesse para a presente pesquisa.

As questões selecionadas para análise do questionário (Anexo 3) foram apresentadas sob duas formas diferentes, quatro de múltipla escolha e seis abertas. As questões de múltipla escolha referem-se aos conceitos de média (questões I e IV), porcentagem (questão II) e probabilidade (questão III). Já as questões abertas correspondem aos conceitos de média (I), variabilidade (II), moda (III), porcentagem (IV), mediana (V) e percentil (VI). Nos parágrafos seguintes será exposto o desempenho dos estudantes em

cada questão individualmente, e após uma breve descrição do desempenho geral em todas as questões conjuntamente.

Tabela 3. Estatísticas descritivas dos itens de múltipla escolha do questionário

Alternativas	Porcentagem de estudantes			
	Item 1	Item 2	Item 3	Item 4
a	25,0	8,8	33,5	8,6
b	2,6	1,5	7,6	5,6
c	40,8	7,4	11,2	36,0
d	25,5	66,7	42,1	4,6
e	6,1	15,7	5,6	37,6
Total	100,0	100,0	100,0	100,0

A primeira questão refere-se à presença de erros de medida sobre os efeitos de valores atípicos no cálculo da média do peso de um pequeno objeto, bem como, a confusão estabelecida entre média e moda. Apesar da média ser considerada um conceito estatístico elementar, mais da metade dos universitários erraram a questão, e apenas 25,5% dos participantes acertaram o item, como pode ser visto na Tabela 3. Porém, deve-se destacar que houve uma porcentagem importante de universitários (66,3% – somando-se as alternativas *c* e *d*) que reconhecem a média como a solução ao problema de estimação de uma quantidade equitativa na presença de erros de medida. Um erro importante é a confusão entre dois conceitos de medida de tendência central, a saber, média e moda, que se torna um problema estabelecido em 25,0% dos universitários, que optaram pela escolha da alternativa *a*.

A segunda questão envolve a interpretação correta da probabilidade e frequência relativa, por meio da interpretação de uma advertência escrita em um frasco de medicamento. Como mostra a Tabela 3, nesse item uma grande porcentagem dos

universitários (66,7%) acertou o item. Embora, pode-se considerar como parcialmente correta a alternativa *e*, sendo esta a segunda com maior porcentagem de estudantes que assinalaram tal opção de resposta (15,7%), deve-se considerar que tal alternativa é menos precisa que a anterior, por ter uma conclusão unicamente qualitativa e não quantitativa da probabilidade.

Para o terceiro item, buscou-se avaliar a compreensão dos conceitos de frequência e probabilidade, bem como a sua relação em um item que envolvia um problema com relação a previsão da probabilidade do tempo. Quase a metade dos estudantes acertou o item (42,1%), e a outra metade cometeu o erro de atribuir uma estimacão diferente da frequência, em que os estudantes optaram, pela maior frequência da probabilidade, destacando-se o primeiro intervalo (95-100), conforme a Tabela 3.

Quanto ao quarto item, objetivou-se avaliar a compreensão dos conceitos de média, por meio de um problema que se refere o número mais comum de filhos nas famílias existentes em uma pequena cidade. Como indica a Tabela 3, os resultados não são tão favoráveis, pois apenas 36,0% da amostra acertaram o item. E, uma porcentagem considerável assinalou a moda (37,6%), o que pode indicar uma confusão da posição relativa da média, mediana e moda em distribuições assimétricas, assim como o efeito dos valores atípicos sobre o valor da média. Isso pode ser a representação de um conhecimento estatístico puramente algorítmico, apresentando dessa forma, uma dificuldade na interpretação e aplicação desses conhecimentos a problemas reais.

Resultados similares são encontrados na segunda parte do questionário, que referem-se às questões abertas, sobre média, variabilidade, moda, porcentagem, mediana e percentil. Essas questões são referentes a dois grupos de idades de 10 profissionais e de 10 estudantes, onde foram elaboradas questões que avaliam o conhecimento do estudante em

cada um dos conceitos citados. Sendo que, na Tabela 4, são mostrados os resultados corrigidos dos estudantes que responderam ao questionário.

Tabela 4. Estatísticas descritivas das questões abertas do questionário

Itens	Porcentagem		
	Correta	Errada	Não respondentes
5	87,9	4,4	7,7
6	62,3	29,0	8,7
7 (profissionais)	93,7	2,4	3,9
7 (estudantes)	91,3	5,3	3,4
8	53,1	31,4	15,5
9	70,4	3,4	26,2
10	60,7	19,4	19,9

Na quinta questão avaliou-se o conceito de média, em que a maioria dos estudantes (87,9%) acertou o item, sendo esta uma questão de aplicação do algoritmo de cálculo da média. Vale ressaltar que, 7,7% dos participantes não responderam tal questão.

Referente à sexta questão, procurou-se analisar o conceito de variabilidade, desta forma, 62,3% dos universitários acertaram tal item, porém, 19,8% destacaram que ambos os grupos (profissionais e estudantes) possuíam a mesma variabilidade, respondendo erroneamente. Neste item, 8,7% dos estudantes deixaram de responder a questão.

Na sétima questão, analisou-se o conceito de moda, para o grupo de profissionais, com 93,7% das respostas da amostra estudada acertou o item, dentre as outras idades que foram citadas as que tiveram 0,5% de respondentes foram, 30 e 33 anos. Quando se analisou a resposta da moda referente ao grupo de estudantes, 91,3% acertaram o item, porém, outras idades foram citadas, a saber, 27, 28, 34 e 35 anos, com 0,5% de respondentes para cada idade.

No que se refere ao conceito de porcentagem avaliado na oitava questão, 53,1% da amostra acertou o item, porém houve uma variabilidade entre as respostas erradas dadas pelos estudantes, variou de 0% a 98%. Assim, percebeu-se uma dificuldade quanto ao conceito de porcentagem. Também, deve-se destacar que 15,5% não responderam tal item.

Quanto a mediana, analisada na questão nove, 70,4% dos estudantes acertaram tal item, embora, uma porcentagem considerável de universitários (26,2%) não respondeu tal item. O conceito analisado na questão 10 é referente ao percentil, percebeu-se que muitos não responderam tal item (19,9%) e 60,7% respondeu corretamente o item.

Quanto ao desempenho geral dos estudantes que participaram da amostra, as pontuações poderiam variar de zero a dez pontos, sendo esta variabilidade alcançada pelos participantes da pesquisa. Apenas 0,5% dos estudantes acertaram todas as questões, e 1,0% erraram todos os itens do questionário. Cerca de 22,2% obtiveram seis pontos no questionário. Porém, cabe destacar que a maioria (79,6%) dos estudantes teve notas entre 5 e 9. A média obtida pela amostra foi de 5,89, com um desvio padrão igual a 1,99 pontos. Os resultados indicam que, apesar das dificuldades apresentadas pelos estudantes, eles possuem algum conhecimento a respeito dos conceitos estatísticos abordados pelo questionário, que estão inseridos no texto em que se avalia a compreensão em leitura de tais estudantes.

Algumas das questões elaboradas para este questionário foram baseadas na tese de Roca (2002). Nestes itens, os resultados encontrados nesta pesquisa, são similares aos relatados por Roca, quando avaliou o conhecimento básico da estatística tanto em professores em formação, quanto àqueles que já exercem a profissão. Assim, percebe-se que, os estudantes de psicologia avaliados nesta pesquisa, apresentam as mesmas dificuldades que professores das séries primárias.

Prova de Compreensão em leitura

A análise descritiva inicial dos itens fornecidos pela amostra dos participantes, indicou que, no que se refere à pontuação total da prova de compreensão, a soma de acertos nas 40 lacunas do texto, foi superior à metade. Os resultados obtidos pelos universitários na prova de leitura geral atingiu uma média de 32,74 acertos, e desvio padrão de 5,44, representando a variabilidade destes escores em torno dessa média. Observou-se uma tendência de pontuação total acima da metade da prova. Essa tendência ocorreu para as pontuações 35 (15,5%), 36 (15,0%), 37 (12,6%) e 38 (6,8%), representando 49,9% do total da amostra, evidenciando que quase a metade dos participantes obteve notas entre 35 a 38 pontos.

Já para a prova de compreensão em leitura específica, a média foi um pouco inferior, igual a 28,74 acertos, com um desvio padrão de 5,53. Desta forma, a compreensão dos estudantes de psicologia neste texto, embora ligeiramente acima da metade do total de lacunas disponíveis para preenchimento, está abaixo do número médio de acertos obtido pelos universitários na prova de compreensão de um texto de leitura geral. As maiores frequências de participantes foram obtidas para as pontuações, acima da metade da prova, a saber, 29 (8,7%), 30 (9,7%), 31 (7,8%), 32 (8,3%), 33 (7,8%) e 34 (8,7%) pontos, representando 51% dos estudantes.

Assim, mais da metade dos estudantes, possui uma compreensão das informações que estão sendo lidas, porém, deve-se levar em consideração que é exigido do universitário, uma grande demanda de atividades que envolvem a leitura, sendo importante atentar para as dificuldades da outra metade de estudantes no que se refere a compreensão do material lido. Outro aspecto importante a se destacar, é o fato de que, nenhum estudante, em ambas as provas, apresentou a pontuação máxima, diferenciando-se da amostra de estudantes do

ensino médio, em que foi aplicada a mesma prova de compreensão em leitura geral, onde 5% dos sujeitos acertaram todas as alternativas de resposta (Marini, 2006).

Assim, esses resultados evidenciam a necessidade de uma maior atenção para a compreensão dos universitários em relação às informações que estão sendo adquiridas. Os resultados obtidos nesta pesquisa, com relação à dificuldade de compreensão em leitura dos universitários, foram reveladas em pesquisas anteriores como a de Santos e cols. (2006), Silva e Santos (2004), Oliveira (2003), Vicenteli (1999), entre outros.

ANÁLISE DE DIFERENÇAS DE COMPREENSÃO EM LEITURA ENTRE GRUPOS DE ESTUDANTES

Foi verificada a existência de possíveis diferenças significativas entre gênero e turno, quanto à compreensão em leitura de um texto geral dos universitários, pelo teste *t* de *student*. O agrupamento dos participantes de acordo com o gênero revelou que não houve diferença significativa entre a compreensão em leitura dos estudantes entre o gênero masculino e feminino (Tabela 5). Esses resultados são semelhantes às pesquisas realizadas anteriormente, como a de Oliveira (2003). Da mesma forma, não foram encontradas diferenças significativas de pontuação média na prova de compreensão entre estudantes dos turnos matutino e noturno.

Tabela 5. Estatísticas referentes à comparação da pontuação total média de compreensão em leitura geral por gênero e turno

Variáveis	Grupo	Pontuação média	Estatística <i>t de Student</i>	Significância <i>p</i>
Gênero	Masculino	33,02	0,380	0,705
	Feminino	32,67		
Turno	Matutino	32,14	-0,661	0,510
	Noturno	32,81		

Também foi realizada a verificação de possíveis diferenças entre gênero e turno, quanto à compreensão em leitura de um texto específico, por meio do teste *t de student*. Novamente, o agrupamento dos participantes de acordo com o gênero e com o turno revelou não haver diferença significativa de compreensão em leitura entre os grupos (Tabela 6).

Tabela 6. Estatísticas referentes à comparação da pontuação total média de compreensão em leitura específica por gênero e turno

Variáveis	Grupo	Pontuação média	Estatística <i>t de Student</i>	Significância <i>p</i>
Gênero	Masculino	29,33	0,776	0,438
	Feminino	28,59		
Turno	Matutino	29,56	0,954	0,342
	Noturno	28,58		

A seguir foi utilizada a análise de variância (ANOVA) para verificar a existência de diferenças significativas da compreensão entre estudantes de diferentes grupos de variáveis, como faixa etária e ano de matrícula na graduação, que estão apresentadas na Tabela 7.

Tabela 7. Estatísticas referentes à comparação da pontuação total média de compreensão em leitura geral por faixa etária e ano do curso

Variáveis	Teste F (ANOVA)			Significância <i>p</i>
	<i>F</i>	Graus de liberdade		
		entre	Dentro	
Faixa Etária	0,259	3	201	0,855
Ano do curso	4,645	4	201	0,001

Os resultados indicaram não haver diferenças significativas na compreensão em leitura tanto no texto geral como no específico (Tabelas 7 e 8) no que se refere às faixas etárias, tais resultados são semelhantes aos encontrados por Oliveira em 2003. Deve-se levar em consideração o fato de que as faixas etárias, consideradas para este estudo, correspondem ao agrupamentos, 17 a 22 anos, de 23 a 28 anos, de 29 a 34 e de 35 ou mais. Assim, a hipótese de que conforme aumenta a idade, aumenta o nível de compreensão do universitário, não foi confirmada a partir da amostra testada.

Tabela 8. Estatísticas referentes à comparação da pontuação total média de compreensão em leitura específica por faixa etária e ano do curso

Variáveis	Teste F (ANOVA)			Significância <i>p</i>
	<i>F</i>	Graus de liberdade		
		entre	Dentro	
Faixa Etária	1,773	3	201	0,154
Ano do curso	5,527	4	201	0,000

Porém, quando a compreensão em leitura dos textos geral e específico é comparada pelo ano do curso de graduação em que o estudante está matriculado, verificou-se uma diferença significativa de médias entre os cinco anos de graduação (com uma tendência das

médias menores nos primeiros e mais elevadas nos últimos). O teste de Tukey *post hoc* (HSD), conforme Tabela 9, revelou diferenças na compreensão em leitura entre os universitários do primeiro ano de graduação em relação aos demais. Desta forma, são necessários somente dois grupos referentes ao ano de graduação para explicar a variância do escore total na prova de leitura geral na amostra total. Nesses termos, foram necessários um grupo para o primeiro ($M=30,13$), quarto ($M=32,81$) e segundo ano ($M=32,84$) da graduação; e outro para o quarto, segundo, quinto ($M=34,28$) e terceiro ano ($M=34,29$). Embora os do segundo e terceiro ano tenham ficado congregados a dois grupos, as médias aumentaram com o decorrer do curso, demonstrando que essa prova identificou o desenvolvimento da compreensão em leitura do universitário, uma vez que diferenciou os universitários do primeiro em relação aos demais ano da graduação. Esses resultados são similares aos encontrados por Santos e colaboradores (2006), Marini (2006) entre outros, embora outras pesquisas não tenham apontado a diferença entre as séries do curso (Pellegrini, 1996).

Tabela 9. Prova de Tukey para a prova de leitura geral

Ano de graduação	N	Subgrupo para alfa=0,05	
		1	2
1	48	30,13	
4	37	32,81	32,81
2	44	32,84	32,84
5	43		34,28
3	34		34,29
Sig.		0,140	0,709

O mesmo princípio ocorre para as diferenças significativas apresentadas pela amostra entre as médias na prova de compreensão específica. De forma que, o do teste de

Tukey, novamente revelou a formação de dois grupos principais para explicar a variância do escore total na prova de leitura específica, um grupo engloba o primeiro ano ($M=25,94$), e o segundo ano ($M=28,43$); o grupo dois contém o segundo, quarto ($M=29,19$), quinto ($M=30,21$) e terceiro ano ($M=30,76$), conforme melhor visualizado na Tabela 10. Esses resultados indicam que há diferenças entre os estudantes do primeiro e do quinto ano da graduação, o que é esperado, de forma que há uma melhora na compreensão das informações que são lidas, na medida em que o estudante adquire novos conhecimentos.

Tabela 10. Prova de Tukey para a prova de leitura específica

Ano de graduação	N	Subgrupo para alfa=0,05	
		1	2
1	48	25,94	
2	44	28,43	28,43
4	37		29,19
5	43		30,21
3	34		30,76
	Sig.	0,216	0,278

Para a investigação de possíveis diferenças significativas entre o desempenho na prova de compreensão em leitura de um texto específico entre os universitários que realizaram ou não as disciplinas de estatística e psicometria, a análise de variância indicou haver diferenças entre os estudantes ($F[4, 201]=155,212$; $p<0,001$). O teste de Tukey revelou a formação de dois grupos principais, o de menor média, que compreende o primeiro e segundo ano, e o de maiores médias, abrangendo o terceiro, quarto e quinto ano. Porém, deve-se levar em consideração o fato de que, houve um aumento da compreensão no decorrer dos anos da graduação, por este motivo, optou-se por realizar uma nova análise, por meio do teste *t de student*, separando a amostra nos dois grupos principais.

Com o intuito de verificar se o estudante que já cursou as disciplinas de Estatística e Psicometria possui diferenças nas médias da prova de compreensão em leitura, foi realizado um agrupamento dos sujeitos, segundo aos anos de graduação. De acordo com o seguinte critério, grupo um refere-se aos universitários do primeiro ano que não concluiu nenhuma das duas disciplinas; grupo dois, formado por aqueles que estão matriculados no segundo ano, que concluíram a disciplina de estatística e ainda não concluíram a de psicometria; e, o grupo três, que são aqueles que já concluíram as disciplinas citadas. Os resultados revelaram que novamente houve diferenças entre as médias dos três grupos ($F[2,203]=287,488$; $p<0,001$). Ao analisar o teste de Tukey, pode-se observar, conforme a Tabela 11, que são necessários apenas dois grupos para explicar a variância do escore total. Assim, o primeiro grupo é formado apenas pelos estudantes que não concluíram as disciplinas, que obtiveram a menor média; já os que obtiveram as maiores médias, compreendem os grupos dois, que são aqueles que já concluíram a disciplina de estatística, mas não a psicometria, e pelo grupo três, que são aqueles que já concluíram as duas disciplinas.

Tabela 11. Prova de Tukey para a prova de leitura específica

Ano de graduação	N	Subgrupo para alfa=0,05	
		1	2
1	48	25,94	
2	44		28,43
3	114		30,04
	Sig.	1,000	0,235

ANÁLISE DE CORRELAÇÃO DA PROVA DE COMPREENSÃO EM LEITURA DE UM TEXTO ESPECÍFICO COM OUTRAS VARIÁVEIS

As evidências de validade convergente foram investigadas com base na relação entre variáveis, de forma que o desempenho na prova de compreensão em leitura específica foi relacionado ao desempenho na prova de compreensão em leitura geral. Observou-se que existe uma correlação positiva significativamente diferente de zero entre as duas provas de compreensão em leitura ($r=0,535$; $p<0,001$). Esses valores indicam que quanto maior a compreensão em um texto de leitura geral, maior tende a ser a compreensão em um texto de leitura que envolve conceitos específicos, sugerindo uma associação entre a compreensão de ambos os textos, como mostra a Figura 2.

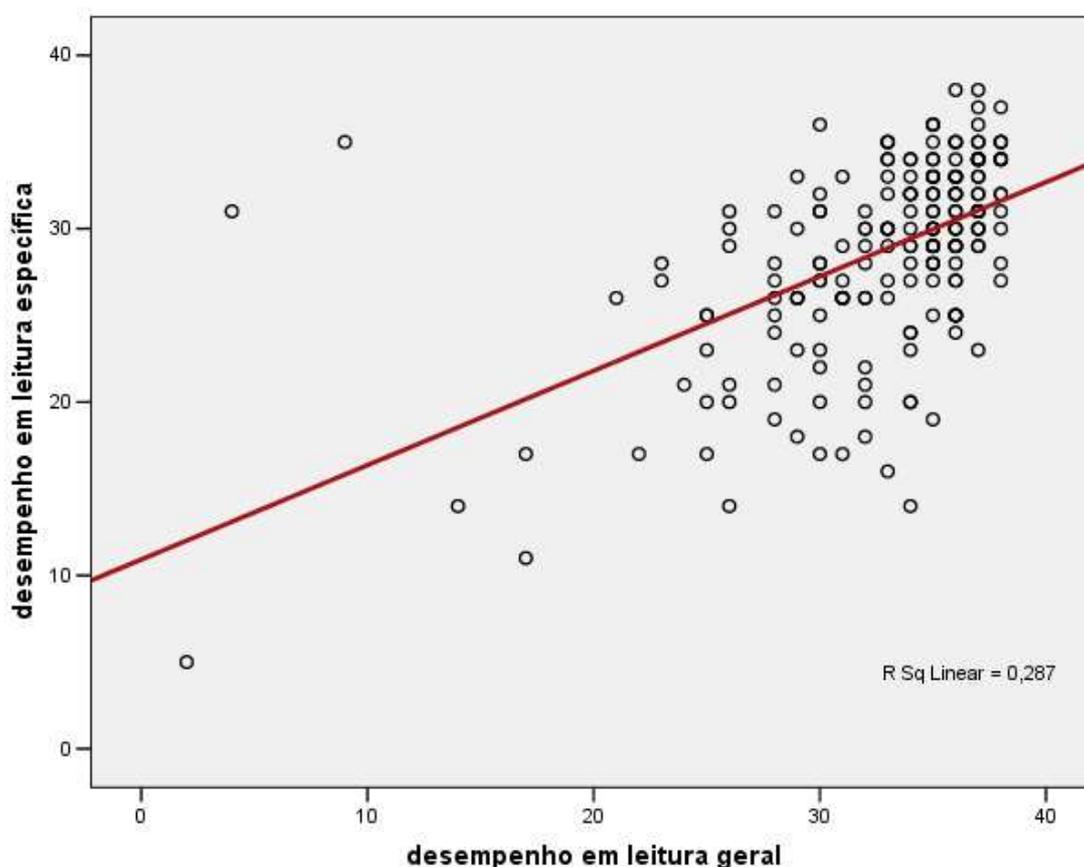


Figura 2. Diagrama de dispersão e reta de regressão do desempenho em leitura específica em função do desempenho em leitura geral

Uma nova análise foi realizada para verificar a existência de possíveis efeitos nos desempenhos nas provas de compreensão em leitura foi realizada, quando considera uma terceira variável, no caso a realização das disciplinas Estatística e Psicometria nos desempenhos em leitura. Assim, pode-se perceber por meio da correlação obtida que, ainda permanece moderada e significativamente diferente de zero. Pode-se notar também, na Figura 3, que os estudantes que não concluíram nenhuma das duas disciplinas (grupo 1) apresentam um relacionamento condicional, de forma que não apresentam uma relação forte entre os desempenhos. O mesmo não ocorre para o grupo 2 e 3, que apresentam uma associação mais evidente entre os mesmos e entre os desempenhos em leitura.

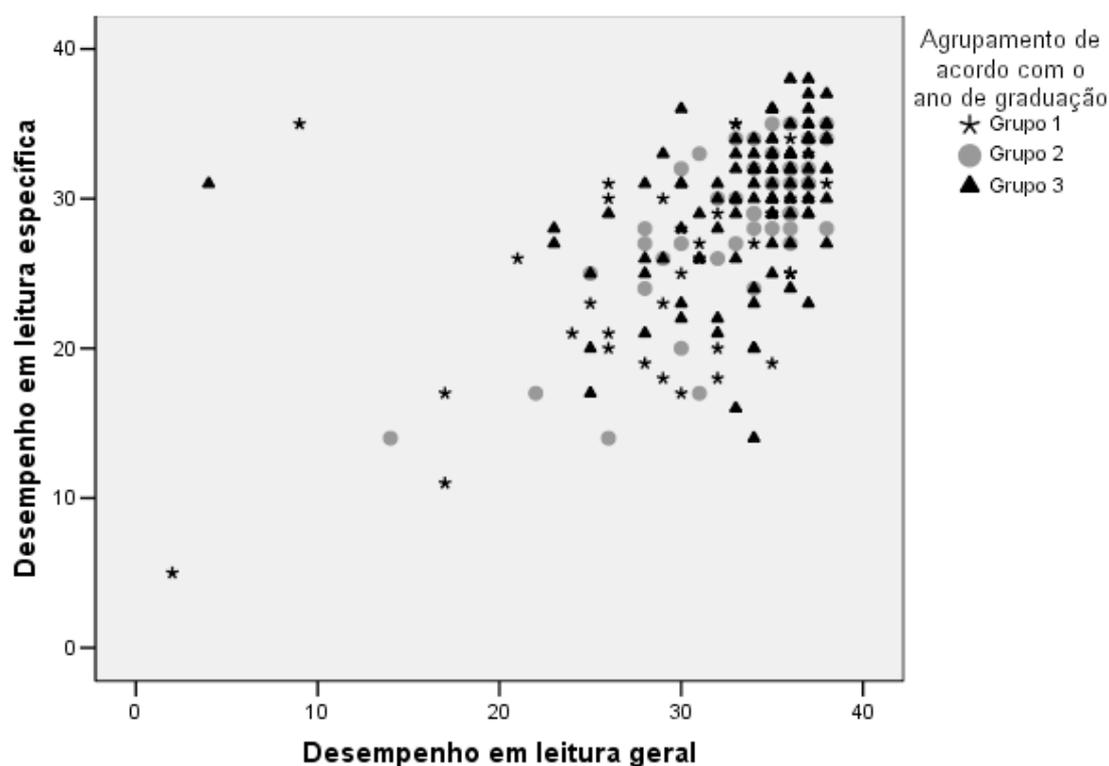


Figura 3. Diagrama de dispersão do desempenho em leitura específica em função do desempenho em leitura geral pela realização da estatística e psicometria

Buscou-se também evidências de validade com base em um outro critério externo, o desempenho do estudante no questionário que contém problemas estatísticos. Procedeu-se a análise correlacional entre as notas dos estudantes no questionário com as notas na prova de compreensão em leitura específica. Com base nestes dados, verificou-se uma correlação positiva e significativamente diferente de zero entre a prova de compreensão específica com o desempenho no questionário ($r=0,257$; $p<0,001$), assim, pode-se compreender que, quanto mais o sujeito conhece os conceitos estatísticos, melhor será a sua compreensão.

ANÁLISE DA DIMENSIONALIDADE DA PROVA

Para obter a evidência de validade de construto para a prova de leitura específica, foi realizada uma análise fatorial. Conforme Pestana e Gageiro (2003) para poder aplicar o modelo fatorial deve-se observar correlação entre as variáveis, para a verificação de fatores comuns entre as mesmas. Assim, faz-se necessário realizar uma análise de adequação da amostra para a análise fatorial. Os resultados indicaram que, para a amostra dos 40 itens de cada uma das duas provas foram adequadas para avaliar a compreensão em leitura, tanto de um texto geral, quanto a do texto de leitura específica, conferindo validade aos resultados encontrados. O valor encontrado para a medida de adequação da amostra de Kaiser-Meyer-Olkin foi igual a 0,826 indicando um resultado satisfatório para a prova de leitura geral. O mesmo resultado satisfatório foi encontrado para a prova de leitura específica o KMO foi igual a 0,721.

O teste de esfericidade de Bartlett que permiti avaliar a hipótese de igualdade de variância-covariância no grupo estudado, isto é, que a matriz de correlação é uma matriz identidade, mostrando, portanto, que existe correlação entre as variáveis estudadas. Os valores para a prova de leitura geral são $\chi^2(780)=3345,539$, com um índice de significância

de $p < 0,001$, e para a prova de leitura específica é $\chi^2(780) = 1694,160$, com o nível de significância de $p < 0,001$ (Hair, Anderson, Tatham & Black, 2005; Pestana & Gageiro, 2003). A partir desses dados, realizou-se a análise fatorial, considerando o valor satisfatório do KMO e do teste de esfericidade de Bartlett, que favorecem a possibilidade de se realizar a análise de dimensionalidade das provas de compreensão.

Para a análise de itens, é essencial que todos os itens avaliem o mesmo construto, tanto para a Teoria Clássica dos Testes (TCT) quanto para a Teoria de Resposta ao Item (TRI). Na TCT, esta é uma propriedade necessária, uma vez que trabalha com o escore total e cada item é avaliado em função deste escore total. Assim, há a necessidade de que os itens sejam somáveis e isto somente faz sentido, se os itens do instrumento medem o mesmo construto. Já, no caso da TRI, todos os itens devem se referir a um único traço latente, de forma que esses são avaliados apenas em função desse traço latente. A unidimensionalidade de um teste é uma condição necessária para analisar qualquer característica subsequente de um item, tais como a dificuldade e discriminação do mesmo (Pasquali, 2003).

Ainda de acordo com o autor, uma análise que vem sendo muito utilizada é a análise fatorial *full information*, que pode ser realizada pelo programa TESTFACT. Esse novo método de análise de itens é baseado na TRI, esta não requer o cálculo das intercorrelações entre os itens, porém, trabalha com as informações completas substituindo a utilização dos métodos de informação sumarizada, como as correlações.

Partindo desse pressuposto, a análise clássica dos itens de um teste baseia-se nos parâmetros descritivos dos itens, que auxiliam na interpretação da distribuição das respostas em cada alternativa do item. Realizou-se então a análise fatorial *full information* para as provas de compreensão em leitura, tanto de leitura geral, como de leitura específica; de

forma que serão apresentados os resultados da prova de leitura geral, conjuntamente com os resultados encontrados na prova de leitura específica.

Primeiramente, os resultados indicaram as propriedades psicométricas dos itens da prova (Tabela 12 e 13) apresentados por meio dos seguintes parâmetros: (i) *média do total de acertos* dos participantes que acertaram um determinado item (lacunas); (ii) *índice de facilidade*, referindo-se a proporção de participantes que responderam corretamente ao item; (iii) *índice de dificuldade*, que corresponde ao índice de facilidade calculado em uma escala *delta* que possui uma média 13 e desvio padrão 4; e (iv) *correlação ponto bisserial*, que mede o poder de discriminação dos itens. Cada um dos quatro parâmetros descritos é apresentado nos próximos parágrafos em função dos resultados obtidos.

O número médio de acertos para cada um dos itens da prova de leitura geral apresentado na Tabela 12, na segunda coluna, corresponde à média do número total de acertos dos estudantes que acertaram cada um dos itens. Observa-se nesta tabela que, para o item 12 a pontuação total na prova, os universitários que acertaram este item, é superior a todas as outras médias, indicando que este item está entre os mais difíceis na prova. Esse mesmo resultado também pode ser verificado pelo índice de facilidade, que é o menor de todos para o item 12 (0,549).

Ainda com relação à média de acertos que os participantes tiveram na prova de compreensão em leitura geral, pode-se observar que os estudantes acertaram mais que 50% dos itens ($M=34,63$; $DP=5,65$). Dos quarenta itens dessa prova, a média de acertos para os participantes que acertaram um determinado item variou de 35,02 a 36,68. Para a prova de compreensão de leitura específica, a média de acertos novamente, foi superior a 50% das questões ($M=28,74$; $DP=5,53$), variando de 29,12 a 31,31, contudo esses valores obtidos na prova de compreensão específica, foram inferiores aos da prova de leitura geral.

O índice de facilidade indica a proporção de acertos em cada item. Assim, os valores podem variar em um intervalo de 0 (nenhum participante acerta) a 1 (todos os participantes acertam). Quanto mais o índice estiver próximo de 1 maior a facilidade dos estudantes para resolver a questão (Cronbach, 1994).

Os resultados apresentados na Tabela 12 revelam que a prova de leitura geral apresentam em sua maioria itens muito fáceis (com valores acima de 80% de acerto – como, por exemplo, os itens 1, 2, 3, entre outros), itens fáceis (com valores entre 60% e 80% – itens 6, 8, 26, 27 e 35) e um item de nível médio (com valores entre 40% e 60% – item 12). Pode-se destacar que nenhuma questão foi muito difícil, uma vez que, não se observa nenhum valor abaixo de 20%. Porém, pode-se observar apenas um item difícil (com valores entre 20% e 40% – item 33), esses resultados acabam por evidenciar que não há uma distribuição equilibrada dos itens em relação à sua facilidade.

Tabela 12. Parâmetros descritivos dos itens da prova de leitura geral

Item	Média de acertos	Índice de facilidade	Dificuldade	Correlação Ponto-bisserial
1	35,24	0,947	6,55	0,458
2	35,20	0,951	6,37	0,449
3	35,21	0,966	5,72	0,552
4	35,13	0,942	6,73	0,358
5	35,44	0,898	7,92	0,426
6	36,25	0,757	10,21	0,507
7	35,27	0,961	5,96	0,569
8	35,80	0,791	9,76	0,406
9	35,60	0,835	9,10	0,387
10	35,10	0,956	6,17	0,393
11	35,18	0,976	5,14	0,620
12	36,68	0,549	12,51	0,401
13	35,92	0,806	9,55	0,465
14	35,95	0,864	8,60	0,590
15	35,23	0,922	7,32	0,369
16	35,90	0,840	9,03	0,517
17	35,43	0,917	7,45	0,473
18	35,92	0,845	8,94	0,534
19	35,54	0,932	7,04	0,596
20	35,19	0,961	5,96	0,498
21	35,65	0,898	7,92	0,540
22	35,21	0,956	6,17	0,486
23	35,58	0,893	8,02	0,489
24	35,29	0,951	6,37	0,517
25	35,29	0,937	6,89	0,453
26	36,25	0,772	10,02	0,529
27	36,21	0,660	11,35	0,391
28	35,16	0,932	7,04	0,351
29	35,75	0,874	8,42	0,523
30	35,45	0,942	6,73	0,589
31	35,80	0,845	8,94	0,486
32	35,41	0,951	6,37	0,612
33	36,33	0,364	14,39	0,229
34	35,94	0,748	10,33	0,400
35	36,25	0,612	11,84	0,361
36	35,02	0,966	5,72	0,367
37	35,60	0,908	7,69	0,543
38	35,23	0,937	6,89	0,615
39	35,40	0,937	6,89	0,530
40	35,52	0,927	7,18	0,563
Total	34,63			
DP	5,65			

Os resultados revelados na Tabela 13 mostram que a prova de leitura específica, quase que em sua maioria itens muito fáceis (como os itens 2, 5, 6, 10, 13, 16, 17, 18, 21, 24, 26, 27, 28, 31, 32, 36, 38, 39 e 40), itens fáceis (itens 1, 3, 7, 11, 12, 14, 15, 19, 25, 33 e 35) e itens de nível médio (itens 4, 8, 20, 30, 34 e 37). Pode ser observado ainda, que apenas uma questão foi considerada muito difícil (item 22), e também, apenas três itens foram difíceis (itens 9, 23 e 29), e, novamente, nesta prova esses resultados indicaram que não houve uma distribuição equilibrada dos itens em relação à sua facilidade. Ainda com relação ao número médio de acertos para cada um dos itens da prova de leitura específica apresentado na Tabela 13, observa-se que, para o item 9 a pontuação total na prova, os universitários que acertaram este item, é superior a todas as outras médias, indicando que este item está entre os mais difíceis na prova.

O índice de dificuldade dos itens é apresentado em uma escala Delta (Δ), representado por uma escala intervalar que varia de 1 a 25, com média 13 e desvio padrão igual a 4. Esta escala é uma transformação da escala dos escores z da curva normal com a utilização da fórmula ($\Delta = 13 + 4z$), facilitando desta forma, a sua leitura, uma vez que para a escala z os valores variam em uma escala que vai de -3,00 (item mais fácil) a +3,00 (item mais difícil) (Pasquali, 2003).

Assim, os resultados da Tabela 12 revelam que os itens da prova de leitura geral variam de 5,14 a 14,39. A prova apresentou cinco itens (12,5%) quase sem dificuldade (com valores entre 1 a 5) como, por exemplo, 3, 7, 10, entre outros. Também possui itens com pouca dificuldade (com valores entre 6 e 10), representando 77,5% de toda a prova (31 itens), citando o 1, 2, 4, entre outros; além dos quatro itens (10%) com dificuldade média (valores entre 11 e 15), são eles, 12, 27, 33 e 35.

Já, para a prova de leitura específica (Tabela 13), os níveis de dificuldade variam entre 5,96 a 16,74. Desta forma, possui apenas um item quase sem dificuldade (item 10 –

índice igual a 5,96), itens pouco difíceis, representando 62,5% da prova (25 itens), como pode-se citar os itens 2, 3, 5, entre outros. Também há, itens com dificuldade média correspondendo a 32,5% (13 itens) do total de itens, como o 1, 4, 8, 9, entre outros. Possui ainda, apenas um item com certa dificuldade, sendo este o item 22, com um índice de dificuldade igual a 16,74. Vale ressaltar ainda, que em ambas as provas, não possuíram itens muito difíceis com índices de dificuldade entre 21 a 25.

Tabela 13. Parâmetros descritivos dos itens da prova de leitura específica

Item	Média de acertos	Índice de facilidade	Dificuldade	Correlação Ponto-bisserial
1	30,28	0,617	11,81	0,352
2	29,20	0,942	6,73	0,333
3	29,66	0,748	10,33	0,286
4	30,76	0,500	13,00	0,364
5	29,69	0,864	8,60	0,432
6	29,27	0,942	6,73	0,386
7	30,04	0,743	10,39	0,398
8	30,16	0,456	13,44	0,235
9	31,31	0,359	14,44	0,348
10	29,12	0,961	5,96	0,341
11	29,57	0,743	10,39	0,254
12	29,96	0,786	9,82	0,424
13	29,51	0,917	7,45	0,465
14	30,43	0,757	10,21	0,539
15	30,92	0,607	11,92	0,489
16	29,57	0,806	9,55	0,303
17	29,51	0,859	8,69	0,345
18	29,78	0,883	8,23	0,517
19	30,44	0,602	11,97	0,378
20	29,76	0,432	13,68	0,161
21	29,19	0,937	6,89	0,313
22	31,11	0,175	16,74	0,197
23	29,44	0,301	15,09	0,082
24	30,34	0,806	9,55	0,588
25	29,90	0,660	11,35	0,293
26	29,47	0,874	8,42	0,345
27	29,46	0,879	8,33	0,351
28	30,16	0,835	9,10	0,575
29	31,24	0,267	15,49	0,272
30	30,40	0,553	12,46	0,334
31	29,35	0,922	7,32	0,380
32	29,34	0,942	6,73	0,431
33	30,15	0,791	9,76	0,497
34	30,50	0,510	12,90	0,325
35	30,08	0,743	10,39	0,411
36	29,24	0,917	7,45	0,299
37	31,17	0,500	13,00	0,438
38	29,61	0,806	9,55	0,319
39	29,38	0,942	6,73	0,461
40	29,12	0,859	8,69	0,171
Total	28,74			
DP	5,53			

No que se refere ao índice de discriminação, estimado pela correlação ponto-bisserial, que mensura o grau com que um item consegue diferenciar participantes com pontuações altas dos que obtiveram pontuações baixas pontuações nas provas. Desta forma, itens que são classificados como muito fáceis (aqueles que a maioria dos participantes acerta) ou muito difíceis (a maioria dos participantes erram), apresentam baixo índice de discriminação (Pasquali, 2003).

Novamente na Tabela 12, observa-se que 77,5% dos itens da prova de leitura geral são classificados como tendo índice de discriminação muito bom (correlação ponto bisserial acima de 0,40) e oito itens classificados como tendo índice de discriminação bom (correlação ponto bisserial entre 0,30 e 0,40 – itens 4, 9, 10, 15, 27, 28,35 e 36), indicando que 39 itens (97,5% da prova) foram capazes de separar os estudantes com melhor pontuação daqueles com piores pontuações. Apenas um item apresentou um índice médio (correlação ponto bisserial entre 0,20 e 0,30 – item 33), mas ainda sim, este é um item capaz de diferenciar os participantes, aqueles com maiores daqueles com menores pontuações na prova, mesmo sendo fácil.

Para os itens da prova de leitura específica (Tabela 13) os resultados revelaram que 32,5% dos itens possuem índices de discriminação muito bons (itens 5, 12, 13, 14, 15, 18, 24, 28, 32, 33, 35, 37 e 39), cerca de 42,5% da prova apresentou um bom índice (itens 1, 2, 4, 6, 7, 9, 10, 16, 17, 19, 21, 26, 27, 30, 31, 34 e 38), 15% dos itens possuem índices médios (itens 3, 8, 11, 25, 29 e 36) e 10% índices fracos (correlação ponto bisserial menor que 0,20 – itens 20, 22, 23, e 40). Desta forma, 75% dos itens da prova, 30 dos 40 itens, foram capazes de separar os participantes que possuem melhores pontuações daqueles que possuem menores pontos na prova, sendo necessário rever os itens que apresentaram um fraco índice de discriminação. Pois, os itens 22 e 23 apresentaram-se como sendo muito

difícil e difícil respectivamente, e os itens 20 e 40 como de dificuldade mediana e fácil, respectivamente.

No que se refere à consistência interna, a análise revelou um bom índice na prova de compreensão em leitura geral, sendo constatada por meio da utilização da técnica de *Kuder-Richardson*, que permite verificar a fidedignidade do teste ($KR-20 = 0,897$). Para os itens da prova de leitura específica apresentaram uma consistência interna boa, obtida por meio da técnica de *Kuder-Richardson* ($KR-20 = 0,814$). A partir desses resultados pode-se verificar que a prova de leitura geral apresentou um índice ligeiramente superior a prova de leitura específica, bem como melhores índices de discriminação, isso pode ser explicado pelo fato da prova de leitura específica englobar um conteúdo específico, no caso conceitos estatísticos, já a prova de leitura geral abrange um conhecimento mais diversificado.

Para a análise pela TRI, é necessário que se cumpra um dos seus pressupostos, o de que a prova seja unidimensional, ou que, se possa assumir um fator predominante para a utilização de modelos unidimensionais. A análise fatorial com informação completa realizada neste estudo, por meio do TESTFACT, inclui índices que indicam a contribuição de um fator para a explicação das correlações entre os itens, possibilitando fazer previsões das respostas dos participantes aos itens a partir das curvas dos itens.

Por meio da análise com informação completa pode-se perceber ainda que, diferentes padrões de resposta podem ser esperados quando os itens são completamente independentes mediando mais de um fator ou quando medem um único fator. A verificação da adequação de um modelo ser unidimensional ou multidimensional aos padrões de respostas dos participantes é por meio do teste Qui-quadrado (Pasquali, 2003). Nas Tabelas 14 e 15 estão apresentadas às cargas fatoriais dos itens nos três fatores considerados para a análise fatorial dos dados, bem como as comunalidades, que se referem à porcentagem de variância explicada de cada item em relação à prova como um todo.

Foi realizada a rotação PROMAX para os itens da prova de compreensão em leitura geral, por meio da análise fatorial com informação completa, os resultados evidenciaram que, considerando o modelo unidimensional, 43,41% da variância total pode ser explicada por esse modelo. Incluindo-se um segundo fator, modelo bidimensional, a variância explicada aumenta para 49,9%.

Pode-se observar, na Tabela 14, que não houve nenhum item com carga fatorial menor que 0,30. A análise fatorial com informação completa pelo método de rotação PROMAX, considerando-se 3 fatores na análise indicou que os mesmos são correlacionados entre si ($r_{12}=0,711$; $r_{13}=0,551$; $r_{23}=0,506$). A análise efetuada também indicou uma correlação média tetracórica moderada ($r_{tet}=0,448$) entre os 774 pares de combinações, dois a dois, dos itens. Porém, optou-se por realizar uma segunda extração para a verificação da predominância unidimensional da prova, observando-se assim, que os itens apresentam cargas fatoriais maiores que 0,38, exceto o item 33, em que foi observada uma tendência de escolha da opção errada pelos participantes que obtiveram os escores mais altos no teste. Destaca-se ainda que a correlação tetracórica manteve-se igual.

Os resultados referentes a prova de leitura geral, em sua maioria não diferem dos apresentados pela análise realizada com uma amostra de estudantes do ensino médio (Joly & Dias, 2007a). Porém deve-se levar em consideração, que a amostra formada neste estudo é uma amostra de estudantes do nível universitário, do curso de psicologia, de apenas uma universidade do interior do estado de São Paulo.

Tabela 14. Cargas não rotacionadas dos fatores principais da prova de compreensão em leitura geral

Item	Primeira extração			Segunda	
	Comunalidade	Fator 1	Fator 2	Fator 3	Fator 1
1	0,538	0,598	0,191	0,260	0,582
2	0,213	0,142	0,364	0,525	0,475
3	0,572	0,417	0,615	0,270	0,733
4	0,171	0,386	0,311	0,137	0,469
5	0,147	0,285	0,437	0,219	0,513
6	0,375	0,516	0,402	0,210	0,625
7	0,003	0,172	0,872	0,277	0,757
8	0,101	0,398	0,174	0,369	0,498
9	0,036	0,226	0,301	0,424	0,482
10	0,735	-0,111	0,784	0,361	0,475
11	0,126	0,265	0,891	0,273	0,859
12	0,161	0,309	0,501	0,126	0,501
13	0,082	0,221	0,658	0,263	0,599
14	0,026	0,416	0,578	0,447	0,777
15	0,680	0,343	0,350	0,027	0,412
16	-0,072	0,688	0,169	0,234	0,620
17	-0,038	0,494	0,061	0,550	0,596
18	-0,264	0,565	0,453	0,213	0,668
19	0,121	0,280	0,555	0,583	0,736
20	0,116	0,213	0,477	0,507	0,606
21	0,528	0,404	0,276	0,571	0,664
22	0,195	0,327	0,318	0,662	0,677
23	0,171	0,344	0,339	0,397	0,562
24	0,041	0,238	0,652	0,382	0,676
25	0,165	0,358	0,258	0,604	0,632
26	0,162	0,470	0,310	0,410	0,651
27	-0,086	0,179	0,566	0,035	0,414
28	0,132	0,259	0,157	0,359	0,388
29	-0,062	0,292	0,449	0,461	0,631
30	-0,068	0,344	0,324	0,725	0,707
31	0,014	0,408	0,335	0,350	0,587
32	-0,061	0,397	0,568	0,459	0,774
33	0,142	0,293	0,458	-0,478	0,188
34	0,159	0,493	0,196	0,136	0,466
35	0,012	0,366	0,290	0,107	0,411
36	0,675	0,382	0,082	0,588	0,504
37	-0,096	0,730	0,203	0,285	0,669
38	0,083	0,857	0,244	0,406	0,843
39	-0,056	0,820	0,022	0,437	0,728
40	0,205	0,780	0,303	0,296	0,760

O modelo unidimensional para os dados da prova de compreensão em leitura de um texto específico, pode explicar 24,80% da variância entre os itens e incluindo um segundo fator, modelo bidimensional, a variância explicada aumenta para 30,94%. Ao se considerar um modelo tridimensional, a variância total explicada aumenta muito pouco (35,85%). A análise fatorial com informação completa pelo método de rotação PROMAX indicou que os 3 fatores considerados na análise (Tabela 15) possuem correlações significativas entre eles ($r_{12}=-0,482$; $r_{13}=0,540$; $r_{23}=-0,467$). A análise fatorial com informação completa efetuada para a prova de leitura específica indicou uma baixa correlação média tetracórica ($r_{tet}=0,225$) entre os 776 pares de combinações, dois a dois, dos itens. Lembrando que a correlação tetracórica é utilizada quando duas variáveis contínuas e normalmente distribuídas foram artificialmente reduzidas a duas categorias (dicotomizadas), ou seja, é utilizada para relacionar duas medidas binárias.

Tabela 15. Cargas não rotacionadas dos fatores principais da prova de compreensão em leitura específica

Item	Primeira extração			Segunda	
	Comunalidade	Fator 1	Fator 2	Fator 3	Fator 1
1	0,855	0,210	-0,094	0,304	0,372
2	0,095	0,292	0,221	0,627	0,540
3	0,473	-0,025	0,017	0,532	0,349
4	0,050	0,339	-0,166	0,228	0,408
5	-0,019	0,790	0,114	0,235	0,534
6	0,716	0,759	0,073	0,253	0,595
7	-0,147	0,262	0,005	0,483	0,464
8	0,007	0,236	-0,243	-0,044	0,192
9	0,003	-0,001	-0,667	0,186	0,373
10	0,805	0,669	-0,278	0,104	0,585
11	0,096	0,147	-0,049	0,215	0,247
12	0,014	0,467	-0,165	0,269	0,508
13	0,116	0,396	-0,309	0,476	0,695
14	-0,145	0,229	-0,132	0,685	0,665
15	0,308	0,156	-0,259	0,560	0,582
16	-0,004	0,288	-0,090	0,214	0,349
17	0,033	0,070	-0,106	0,526	0,439
18	0,100	0,197	-0,176	0,763	0,722
19	0,018	0,211	-0,164	0,417	0,476
20	-0,051	-0,095	0,142	0,355	0,132
21	0,355	-0,033	-0,319	0,584	0,513
22	-0,101	0,096	-0,638	-0,003	0,275
23	0,023	-0,121	-0,007	0,033	-0,033
24	-0,015	0,544	-0,140	0,567	0,756
25	0,005	0,216	-0,136	0,141	0,275
26	0,060	0,315	-0,165	0,309	0,457
27	0,202	0,263	-0,191	0,248	0,396
28	0,698	0,370	-0,712	0,360	0,743
29	-0,084	-0,083	-0,698	0,092	0,283
30	0,056	0,444	-0,269	0,045	0,386
31	-0,192	0,509	-0,313	0,192	0,564
32	0,108	0,429	-0,552	0,316	0,690
33	-0,049	0,225	-0,341	0,481	0,613
34	-0,105	0,291	-0,204	0,188	0,369
35	0,017	-0,045	-0,322	0,521	0,477
36	0,883	0,265	-0,194	0,225	0,390
37	0,018	0,334	-0,114	0,461	0,552
38	0,065	0,635	0,255	0,209	0,381
39	-0,100	0,520	-0,323	0,417	0,733
40	0,003	0,348	-0,009	-0,015	0,192

Foram realizadas outras duas análises para a verificação da prova como bidimensional ou unidimensional, os resultados evidenciaram que para as duas extrações posteriores os itens que apresentam carga fatorial inferior a 0,30 são os mesmos tanto para a análise considerando um ou dois fatores. Vale ressaltar que somente dois itens, o número 22 e 25 apresentam cargas fatoriais elevadas para o primeiro fator. Para essa nova análise, de apenas um fator, a análise indicou novamente uma baixa correlação média tetracórica ($r_{tet}=0,225$) entre os 776 pares de combinações.

Assim, se objetivasse aumentar a consistência interna entre os itens, sugere-se então, que sejam retirados os itens com cargas fatoriais inferiores a 0,30 na prova de leitura específica. Porém, para as análises posteriores desse estudo, considerar-se-á a prova como predominantemente unidimensional, dado que a cada dimensão pouca variância explicada foi acrescida e a prova possui boa consistência interna.

Com base nos resultados da análise fatorial realizadas nas provas, posteriormente foram analisados os ajustes dos dados aos modelos da TRI considerando os 40 itens de cada prova de compreensão separadamente. Os resultados das estatísticas gerais são calculados a partir do agrupamento dos escores de habilidades semelhantes, da probabilidade de acerto teórica para cada subgrupo, da curva característica do item e dos resíduos relativos à probabilidade real observada, serão apresentados a seguir.

ANÁLISE DE ITENS DA PROVA DE COMPREENSÃO EM LEITURA PELA TEORIA DE RESPOSTA AO ITEM

Após atingir os pressupostos de unidimensionalidade e independência local realizou-se a análise dos itens pela TRI com o auxílio dos programas *XCALIBRE* e *WINSTEPS*. Foram analisados os ajustes dos dados aos três modelos da TRI considerando

os 40 itens das duas provas de compreensão em leitura. Os resultados das estatísticas gerais dos itens indicam a adequabilidade ou não dos itens aos modelos matemáticos da TRI propostos. Essas estatísticas são calculadas a partir do agrupamento dos escores de habilidades semelhantes, da probabilidade de acerto teórica para cada subgrupo, da curva característica do item e dos resíduos relativos à probabilidade real observada.

Lembrando que na TRI a probabilidade de resposta correta aos itens depende tanto da habilidade do sujeito, que é representado por um valor que pode variar de $-\infty$ a $+\infty$, quanto dos parâmetros dos itens. Esses parâmetros são valores estimados que melhor explicam as respostas dos sujeitos aos itens. São três os parâmetros dos itens estimados na TRI, a saber, *índice de discriminação* (a), que indica o quanto indivíduos de diferentes habilidades diferem em relação à probabilidade de acertar um item; *índice de dificuldade* (b), que se refere à habilidade necessária para uma dada probabilidade de acertar o item, calculada a partir da probabilidade de acertá-lo por acaso; e probabilidade de *acerto ao acaso* (c), ou seja, a probabilidade de um sujeito de baixa habilidade dar uma resposta correta a um item difícil (Vendramini & Dias, 2005).

A partir da amostra de universitários do curso de Psicologia, foi realizada uma estimação dos parâmetros dos itens por meio dos modelos logísticos de um parâmetro, também conhecido como modelo de Rasch, além dos modelos de 2 e de 3 parâmetros, que estão descritos nas Tabelas 16 e 17. Uma breve descrição das tabelas faz-se necessário, assim, a primeira coluna apresenta os itens, que corresponde às lacunas da prova; a segunda coluna apresenta o índice de dificuldade b calculado usando o modelo de Rasch; a terceira e quarta, os índices de ajuste do modelo de Rasch (*infit* e *outfit*); a quinta e sexta coluna, apresentam os parâmetros de discriminação a e dificuldade b calculados usando o modelo de 2-parâmetros; a sétima coluna apresenta o índice de ajuste ao modelo de 2-parâmetros (Resíduo); a oitava, a nona e a décima colunas apresentam os parâmetros de discriminação

a , dificuldade b e probabilidade de acerto ao acaso c calculados usando o modelo de 3-parâmetros; a décima primeira coluna apresenta o índice de ajuste ao modelo de 3-parâmetros (Resíduo); a décima segunda coluna mostra a proporção de estudantes na amostra que responderam ao item corretamente (PC); a décima terceira coluna apresenta a correlação item-total pelo produto momento de Pearson (PBs) e a décima quarta coluna apresenta uma medida analógica, a correlação item-total na qual se substitui o escore total pelo traço latente (θ) subjacente (PBt) (Oliveira, 2006).

Tabela 16. Parâmetros dos Itens da prova de compreensão em leitura geral para os modelos de 1, 2 e 3 parâmetros

Item	Modelo de Rasch			2 parâmetros			3 parâmetros						
	<i>b</i>	<i>infit</i>	<i>outfit</i>	<i>a</i>	<i>b</i>	resíduo	<i>a</i>	<i>b</i>	<i>c</i>	resíduo	PC	PBs	PBt
1	-0,91	0,99	0,76	0,91	-2,55	0,36	0,88	-2,37	0,25	0,36	0,95	0,46	0,38
2	-1,05	0,94	1,99	0,88	-2,67	1,24	0,84	-2,53	0,25	1,12	0,95	0,45	0,35
3	-1,54	0,80	0,44	0,98	-2,78	0,29	0,96	-2,69	0,25	0,25	0,97	0,55	0,43
4	-0,79	1,16	0,91	0,84	-2,61	0,46	0,83	-2,32	0,25	0,32	0,94	0,36	0,31
5	0,00	1,08	1,26	0,82	-2,13	0,37	0,82	-1,79	0,25	0,28	0,90	0,43	0,39
6	1,37	0,91	0,79	0,89	-1,12	0,66	1,02	-0,68	0,24	0,54	0,76	0,51	0,54
7	-1,36	0,81	0,38	0,99	-2,67	0,34	0,97	-2,56	0,24	0,35	0,96	0,57	0,45
8	1,11	1,09	1,11	0,74	-1,45	0,44	0,85	-0,91	0,25	0,25	0,79	0,41	0,43
9	0,73	1,12	1,25	0,74	-1,73	0,55	0,82	-1,22	0,25	0,42	0,83	0,39	0,40
10	-1,19	1,07	1,22	0,87	-2,76	0,86	0,85	-2,57	0,25	0,84	0,96	0,39	0,32
11	-2,01	0,65	0,17	1,03	-2,94	0,49	1,02	-2,90	0,25	0,40	0,98	0,62	0,46
12	2,66	1,03	1,00	0,69	-0,24	0,85	0,94	0,34	0,24	0,56	0,55	0,40	0,48
13	0,99	0,99	0,97	0,85	-1,43	0,49	0,94	-0,96	0,25	0,31	0,81	0,47	0,49
14	0,43	0,85	0,54	1,02	-1,63	0,55	1,02	-1,37	0,24	0,61	0,86	0,59	0,57
15	-0,39	1,12	1,44	0,80	-2,44	0,75	0,78	-2,12	0,25	0,63	0,92	0,37	0,31
16	0,68	0,95	0,73	0,88	-1,59	0,36	0,90	-1,25	0,24	0,27	0,84	0,52	0,50
17	-0,30	1,01	1,25	0,89	-2,22	0,55	0,93	-1,89	0,25	0,46	0,92	0,47	0,44
18	0,63	0,93	0,74	0,90	-1,60	0,32	0,92	-1,27	0,24	0,31	0,84	0,53	0,51
19	-0,58	0,80	0,80	0,99	-2,26	0,54	0,96	-2,12	0,24	0,46	0,93	0,60	0,50
20	-1,36	0,89	1,03	0,94	-2,74	0,61	0,92	-2,60	0,25	0,58	0,96	0,50	0,39
21	0,00	0,91	0,78	0,93	-1,97	0,53	0,90	-1,77	0,24	0,46	0,90	0,54	0,49
22	-1,19	0,98	0,50	0,95	-2,64	0,32	0,95	-2,44	0,25	0,35	0,96	0,49	0,42
23	0,07	0,98	1,07	0,86	-2,02	0,50	0,85	-1,75	0,24	0,47	0,89	0,49	0,44
24	-1,05	0,90	0,55	0,95	-2,56	0,52	0,93	-2,42	0,24	0,61	0,95	0,52	0,43
25	-0,68	1,06	0,69	0,91	-2,41	0,60	0,94	-2,12	0,25	0,46	0,94	0,45	0,43
26	1,26	0,93	0,92	0,87	-1,20	0,56	0,90	-0,85	0,24	0,55	0,77	0,53	0,54
27	2,01	1,13	1,09	0,65	-0,80	0,58	0,79	-0,19	0,25	0,46	0,66	0,39	0,41
28	-0,58	1,15	2,13	0,79	-2,57	0,83	0,80	-2,22	0,25	0,78	0,93	0,35	0,31
29	0,31	0,94	0,72	0,92	-1,79	0,51	0,92	-1,50	0,24	0,37	0,87	0,52	0,49
30	-0,79	0,78	0,88	0,98	-2,39	0,75	0,95	-2,27	0,24	0,68	0,94	0,59	0,49
31	0,63	0,99	0,98	0,85	-1,66	0,51	0,87	-1,30	0,24	0,39	0,84	0,49	0,47
32	-1,05	0,76	0,47	1,01	-2,49	0,33	0,99	-2,38	0,24	0,28	0,95	0,61	0,49
33	3,73	1,31	1,75	0,55	0,68	1,14	0,88	1,51	0,25	0,36	0,36	0,23	0,23
34	1,44	1,12	1,02	0,70	-1,23	0,52	0,80	-0,67	0,25	0,37	0,75	0,40	0,42
35	2,30	1,12	1,30	0,64	-0,56	0,58	0,82	0,06	0,25	0,49	0,61	0,36	0,41
36	-1,54	1,18	2,71	0,89	-2,92	0,76	0,88	-2,69	0,25	0,72	0,97	0,37	0,30
37	-0,14	0,91	0,75	0,95	-2,05	0,16	0,92	-1,82	0,24	0,21	0,91	0,54	0,48
38	-0,68	0,81	0,33	1,05	-2,26	0,69	1,03	-2,09	0,24	0,64	0,94	0,62	0,53
39	-0,68	0,91	0,61	0,96	-2,35	0,58	0,95	-2,15	0,24	0,54	0,94	0,53	0,46
40	-0,48	0,89	0,44	0,98	-2,21	0,51	0,98	-2,00	0,24	0,48	0,93	0,56	0,50

É importante destacar que, a partir da estimação dos parâmetros dos itens, podem ocorrer os seguintes problemas: (i) o índice de discriminação a ter um valor abaixo do valor crítico 0,30; (ii) o índice de dificuldade b apresentar um valor crítico acima de 2,95 ou abaixo de -2,95; (iii) o valor da probabilidade de acerto casual estar acima de 0,40; (iv) apresentar um erro de chaveamento, que representa a correlação entre uma das opções de resposta incorreta mais marcada pela amostra e o escore total, sendo esta correlação mais alta que a correlação entre a resposta correta e o escore total; e (v) os resíduos padronizados do ajuste do modelo excederem o valor crítico 2,0 (Vendramini & Dias, 2005).

Com base nessas condições, ao observar os resultados apresentados na Tabela 16 percebe-se que os dados dos itens da prova de compreensão em leitura de um texto geral, apresentam certo desajuste no modelo de Rasch, não sendo, portanto observados tais desajustes nos modelos de 2 e 3 parâmetros. Assim, na estimação dos parâmetros pelo modelo de Rasch observou-se que 9 dos 40 itens não se ajustam ao modelo, desrespeitando assim, os critérios assumidos. A análise do *infit* e *outfit* apresenta alguns resultados fora do valor crítico esperado, entre 0,50 e 1,50, a saber, os itens 3, 7, 11, 15, 28, 32, 33, 36, 38 e 40.

Desta forma, após a constatação de que o modelo logístico de um parâmetro implicava a eliminação de 9 itens da prova (22,5% da prova), por não atenderem às condições exigidas, concluiu-se que, para não comprometer a representatividade do domínio avaliado, os modelos de dois e três parâmetros se justam melhor aos dados. Os 40 itens se ajustam aos modelos de dois e três parâmetros, não tendo ocorrido nenhum dos problemas citados anteriormente. E, deve-se destacar ainda que, apesar de não se observar discrepância significativa dos modelos ajustados, com relação a valores dos resíduos menores que 2,0, o modelo que apresentou menor resíduo foi o de três parâmetros (Tabela 16).

Ainda na Tabela 16 pode-se observar na décima segunda coluna a proporção de estudantes que responderam corretamente o item, permitindo concluir que os itens 12 e 33 foram os itens mais difíceis para essa amostra com apenas 55% e 36% dos universitários acertaram tais itens, respectivamente. Tais dados são confirmados a partir da análise de dificuldade do item apresentado nos três modelos matemáticos. Porém, deve-se destacar que o item 33 apresentou um desajuste no modelo de um parâmetro, que pode ser decorrência da alta dificuldade do item para a amostra avaliada. E, o item 11 o que os estudantes mais acertaram com 98% dos universitários que responderam corretamente o item, confirmando também o desajuste apresentado no modelo de Rasch.

Os resultados apresentados na Tabela 16 podem ser mais bem visualizados pela Curvas Características dos Itens (CCI) apresentadas nas Figuras 3 e 4, para os modelos de 2 e 3 parâmetros, respectivamente.

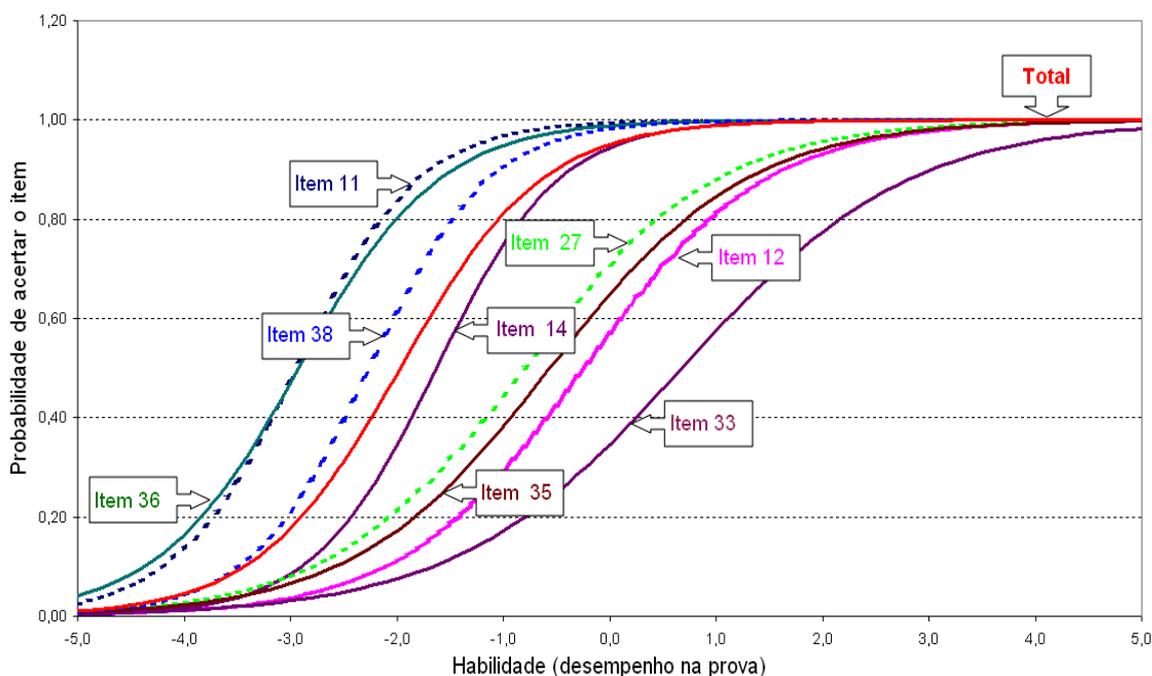


Figura 3. Curvas características de alguns itens da prova de compreensão em leitura geral, modelo de dois parâmetros

Ao observar a Figura 4, o item 11 foi o mais fácil da prova e o item 33 foi o mais difícil. O item 15 o de menor discriminação e maior valor possível na probabilidade de acerto ao acaso, enquanto o item 38 o de maior discriminação e menor valor possível na probabilidade de acerto ao acaso. Com relação à probabilidade de acerto ao acaso (c), a prova apresenta um valor médio igual a 0,25, um pouco acima do esperado, uma vez que os itens da prova possuem cinco alternativas possíveis de resposta cada, podendo-se afirmar matematicamente que a probabilidade de um estudante de baixa habilidade acertar o item é aproximadamente igual a 0,20. Resultados similares a este foram observados em outra pesquisa realizada com a mesma prova, porém a amostra utilizada foi em estudantes do ensino médio (Joly & Dias, 2007b).

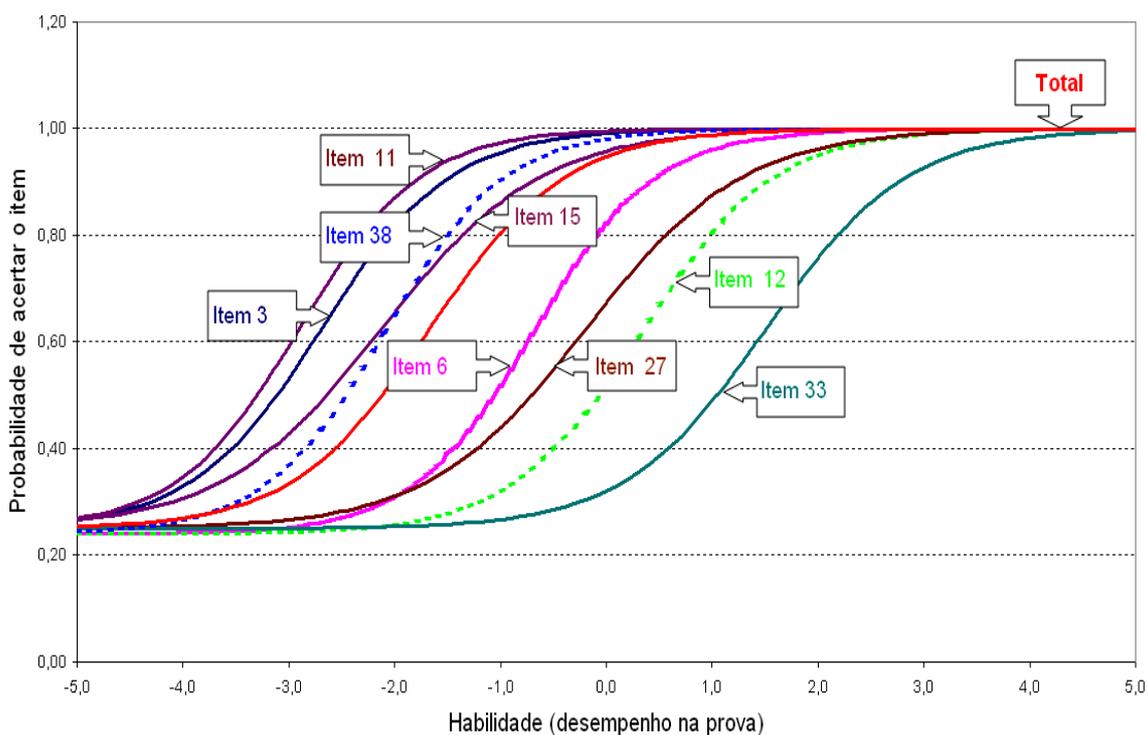


Figura 4. Curvas características de alguns itens da prova de compreensão em leitura geral, modelo de três parâmetros

Ao analisar os dados da prova de compreensão em leitura de um texto específico (Tabela 17) nota-se que os itens não apresentaram ajuste adequado a todos os três modelos propostos pela TRI. Uma vez que, no modelo de três parâmetros, um item (número 23) apresentou um problema de chaveamento, que refere-se a escolha por uma alternativa errada pelos universitários. No modelo de dois parâmetros, três itens, a saber, 2, 10 e 21 apresentam valores de b acima do valor crítico igual a -2,95. Para o modelo de Rasch os itens 39 e 40 apresentam valores de *infit* e *outfit* fora do valor crítico esperado, entre 0,50 e 1,50. Porém, a estimação realizada para os itens pelo modelo de 3 parâmetros encontra-se maior número de itens ajustados podendo considerar que este modelo é o que apresenta o ajuste mais adequado.

Tabela 17. Parâmetros dos Itens da prova de compreensão em leitura específica para os modelos de 1, 2 e 3 parâmetros

Item	Modelo de Rasch			2 parâmetros			3 parâmetros						
	<i>b</i>	<i>infit</i>	<i>outfit</i>	<i>a</i>	<i>b</i>	resíduo	<i>a</i>	<i>b</i>	<i>c</i>	resíduo	PC	PBs	PBt
1	0,83	1,04	1,02	0,53	-0,62	0,77	0,67	0,09	0,25	0,37	0,62	0,35	0,34
2	-1,83	0,95	0,72	0,71	-3,00	0,43	0,77	-2,38	0,25	0,24	0,94	0,33	0,35
3	0,10	1,10	1,08	0,54	-1,52	0,94	0,64	-0,73	0,25	0,59	0,75	0,29	0,31
4	1,39	1,01	0,99	0,55	0,07	0,60	0,73	0,70	0,24	0,25	0,50	0,37	0,36
5	-0,79	0,90	0,97	0,68	-2,14	0,79	0,73	-1,59	0,25	0,53	0,86	0,43	0,43
6	-1,83	0,88	0,73	0,73	-2,90	0,41	0,79	-2,39	0,25	0,41	0,94	0,39	0,39
7	0,13	0,99	0,98	0,60	-1,36	0,56	0,69	-0,71	0,25	0,47	0,74	0,40	0,41
8	1,60	1,13	1,20	0,47	0,31	1,19	0,66	1,19	0,26	0,71	0,46	0,24	0,20
9	2,07	0,98	1,08	0,56	0,79	0,59	0,74	1,45	0,23	0,74	0,36	0,35	0,31
10	-2,30	0,88	0,74	0,73	-3,00	0,44	0,79	-2,75	0,25	0,27	0,96	0,34	0,32
11	0,13	1,13	1,22	0,51	-1,55	1,05	0,59	-0,73	0,25	0,78	0,74	0,25	0,26
12	-0,15	0,96	0,92	0,64	-1,62	0,60	0,71	-0,99	0,25	0,27	0,79	0,42	0,42
13	-1,42	0,88	0,52	0,78	-2,54	0,69	0,82	-2,03	0,25	0,47	0,92	0,47	0,45
14	0,04	0,85	0,76	0,76	-1,25	0,33	0,82	-0,79	0,24	0,49	0,76	0,54	0,54
15	0,88	0,90	0,85	0,68	-0,45	0,38	0,79	0,03	0,24	0,39	0,61	0,49	0,49
16	-0,29	1,06	1,05	0,56	-1,92	1,24	0,62	-1,19	0,25	0,92	0,81	0,30	0,30
17	-0,74	1,03	0,86	0,62	-2,22	0,65	0,71	-1,48	0,25	0,35	0,86	0,35	0,34
18	-0,99	0,84	0,60	0,80	-2,13	0,36	0,84	-1,68	0,25	0,37	0,88	0,52	0,52
19	0,90	1,00	1,09	0,57	-0,49	0,50	0,75	0,13	0,25	0,19	0,60	0,38	0,40
20	1,71	1,17	1,41	0,44	0,48	1,27	0,68	1,38	0,26	0,60	0,43	0,16	0,15
21	-1,74	0,98	0,70	0,69	-2,95	0,73	0,74	-2,34	0,25	0,42	0,94	0,31	0,31
22	3,17	1,04	1,18	0,56	1,97	1,21	0,84	2,66	0,21	1, 10	0,17	0,20	0,16
23	2,37	1,25	1,50	0,42	1,33	1,52	0,75	2,66	0,25	0,85	0,30	0,08	0,01
24	-0,29	0,80	0,70	0,83	-1,48	0,90	0,88	-1,08	0,24	0,85	0,81	0,59	0,59
25	0,60	1,10	1,12	0,50	-0,95	1,11	0,60	-0,18	0,25	0,90	0,66	0,29	0,29
26	-0,88	1,00	0,91	0,63	-2,34	0,67	0,68	-1,70	0,25	0,41	0,87	0,35	0,35
27	-0,93	0,98	0,97	0,63	-2,40	0,65	0,68	-1,78	0,25	0,50	0,88	0,35	0,34
28	-0,52	0,81	0,60	0,83	-1,69	0,57	0,86	-1,26	0,24	0,46	0,83	0,58	0,55
29	2,56	1,04	1,10	0,54	1,37	0,89	0,76	2,20	0,23	0,73	0,27	0,27	0,21
30	1,14	1,04	1,03	0,53	-0,23	0,83	0,72	0,47	0,25	0,44	0,55	0,34	0,34
31	-1,49	0,95	0,70	0,71	-2,70	0,48	0,78	-2,10	0,25	0,33	0,92	0,38	0,39
32	-1,83	0,91	0,50	0,76	-2,84	0,52	0,85	-2,31	0,25	0,43	0,94	0,43	0,42
33	-0,19	0,89	0,77	0,73	-1,52	0,47	0,78	-1,01	0,24	0,52	0,79	0,50	0,50
34	1,34	1,03	1,26	0,53	0,03	0,58	0,72	0,67	0,25	0,39	0,51	0,33	0,33
35	0,13	0,98	0,96	0,61	-1,35	0,53	0,70	-0,71	0,25	0,43	0,74	0,41	0,41
36	-1,42	1,01	1,10	0,64	-2,80	0,74	0,69	-2,15	0,25	0,51	0,92	0,30	0,29
37	1,39	0,93	0,91	0,63	0,07	0,46	0,80	0,58	0,24	0,37	0,50	0,44	0,46
38	-0,29	1,05	1,00	0,57	-1,89	1,07	0,66	-1,17	0,25	0,73	0,81	0,32	0,35
39	-1,83	0,85	0,49	0,78	-2,79	0,52	0,85	-2,33	0,25	0,37	0,94	0,46	0,44
40	-0,74	1,14	1,52	0,52	-2,49	1,46	0,58	-1,69	0,25	1,15	0,86	0,17	0,18

Na prova de leitura específica, pode-se observar alguns dos problemas citados com relação a análise pela TRI, sendo assim, encontrados 2 itens desajustados quando estimado pelo modelo de Rasch, três itens para o modelo de dois parâmetros e apenas um no de três parâmetros. Mesmo a prova não ter atendido algumas das condições exigidas, concluiu-se que, para não comprometer a representatividade do domínio avaliado, os modelos de três parâmetros se ajustam melhor aos dados, por apresentar apenas um problema de chaveamento, além de que não se observou discrepância significativa dos modelos ajustados, com relação a valores dos resíduos menores que 2,0, e sendo este, o modelo que apresentou menor resíduo (Tabela 17).

Pode-se observar ainda, na Tabela 17, na coluna que refere-se a proporção de estudantes que responderam corretamente o item (PC), que o item 22 foi o mais difícil para essa amostra com apenas 17% dos universitários que acertou tal item. Esse resultado também é confirmado ao analisar a dificuldade do item (*b*) apresentado nos três modelos matemáticos. Já o item 10 foi o que os estudantes mais acertaram com 96% dos universitários que responderam corretamente o item. Os resultados das análises podem ser melhores visualizados na figura a seguir.

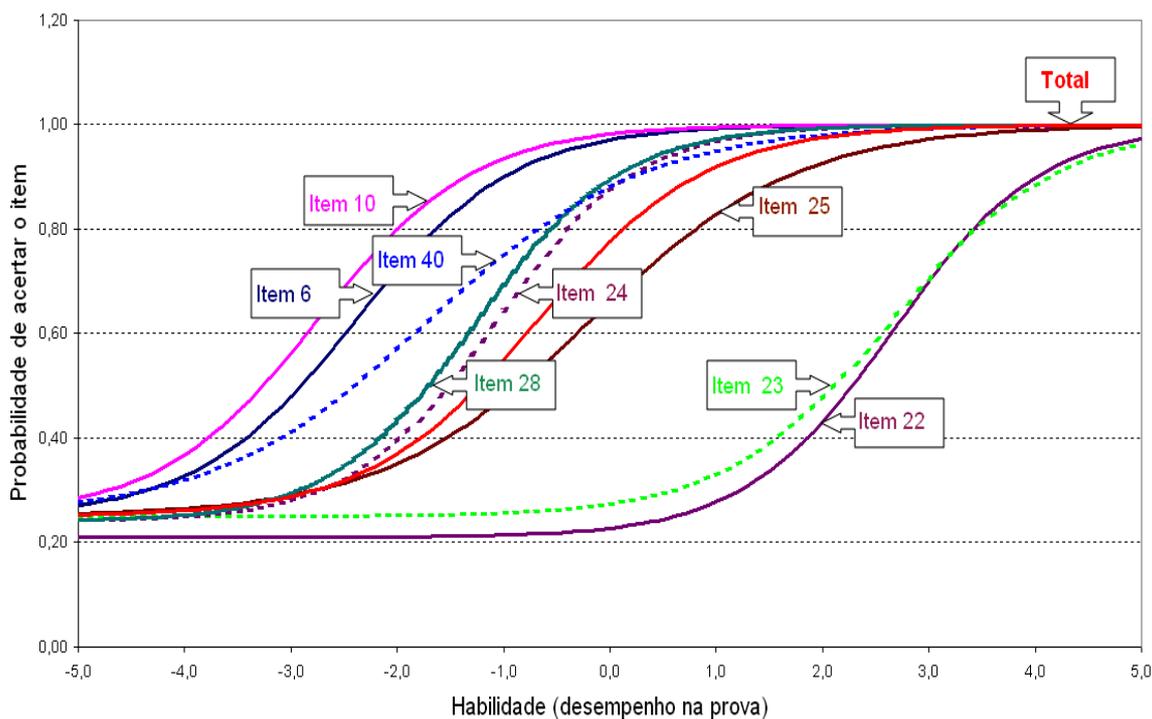


Figura 5. Curvas características de alguns itens da prova de compreensão em leitura específica, modelo de três parâmetros

A Figura 5 evidencia que o item 10 foi o mais fácil da prova e os itens 22 e 23 foram os mais difíceis. Com relação a discriminação, o item 40 apresentou o menor valor de a e maior valor possível na probabilidade de acerto ao acaso, enquanto o item 24 o de maior a e menor valor possível na probabilidade de acerto ao acaso. No que se refere à probabilidade de acerto ao acaso (c), essa prova também apresenta um valor médio igual a 0,25.

Também, foi realizado o exame dos índices de fidedignidade para a prova de compreensão em leitura geral estimado pelos modelos de 1, 2 e 3 parâmetros. A Tabela 18 mostra as informações sobre os valores da fidedignidade da prova de leitura geral e específica calculados pela fórmula $KR-21$, juntamente com valores máximo de informação estimada para os universitários. Pode-se observar que de modo geral a fidedignidade é

mais alta para os modelos de 2 e 3 parâmetros, e com relação a informação esperada e a média de informação é mais alta para o modelo de dois parâmetros.

Tabela 18. Características do componente das provas de compreensão

Modelo	Leitura geral			Leitura específica		
	Fidedignidade KR-21	Informação Esperada	Média de Informação	Fidedignidade KR-21	Informação Esperada	Média de Informação
Rasch	0,720	-	-	0,770	-	-
2-parâmetros	0,897	6,030	7,554	0,813	5,367	5,228
3-parâmetros	0,897	5,233	5,034	0,813	4,717	3,955

De uma maneira geral, as duas provas apresentaram um bom índice de fidedignidade obtidos por meio da técnica de Kuder-Richardson. Resultados semelhantes aos encontrados na precisão da prova de leitura geral, foram obtidos em pesquisa realizada por Joly e Dias em 2007. Tais evidências reforçam a utilização da técnica de cloze como um procedimento adequado na avaliação da compreensão em leitura, como apontado por pesquisas realizadas anteriormente com a aplicação dessa técnica que foi desenvolvida por Taylor em 1953, a saber, Santos e cols. (2002), Saito (2003), Santos (2001), entre outros.

CONCLUSÃO E CONSIDERAÇÕES FINAIS

A leitura tem sido considerada por diversos profissionais, como psicólogos e educadores, como uma atividade fundamental no processo de aprendizagem e aquisição de novos conhecimentos. Deve-se destacar também, a relevância em avaliar as dificuldades relacionadas ao processo de ler, sendo que essas não compreendem apenas o ato de ler, mas a capacidade de compreensão, a atitude crítica, a reorganização, interpretação, assim como a reflexão sobre as mesmas (Santos & Di Nucci, 1984).

As pesquisas têm revelado cada vez mais, que os universitários estão ingressando com muitas dificuldades em leitura (Centofanti e cols., 1997; Nassri, 2002; Oliveira, 2003; Pellegrini, 1996; Silva e cols., 1997). No grau de escolaridade em que o universitário está inserido, há uma necessidade grande em atividades que envolvem a leitura, principalmente de textos mais complexos, em que há uma necessidade do universitário apresentar capacidade de adaptação aos diferentes conteúdos, vocabulário e características das áreas afins (Castro e cols., 1997).

Para a avaliação do nível de compreensão que os universitários estão apresentando, uma técnica que vem sendo muito utilizada em pesquisas psicológicas ou educacionais é a técnica de cloze. Embora essa técnica tenha sido elaborada inicialmente para avaliar a legibilidade de um texto, foi posteriormente utilizada para desenvolver a habilidade do leitor, bem como melhorar a sua motivação para a leitura. Essa técnica também produz mudanças nas estratégias de compreensão, aumentando assim, a capacidade de compreensão das informações que estão sendo lidas (Santos & Di Nucci, 1984).

Considerando a importância da avaliação da compreensão das informações de diversas áreas de conhecimento que são lidas pelos universitários, e com o propósito de

auxiliar os estudantes para a superação das dificuldades em atividades de ensino-aprendizagem nas diferentes disciplinas, este estudo se propôs verificar evidências de validade de uma prova de compreensão em leitura de um texto com informações estatísticas em estudantes de Psicologia.

Os resultados obtidos evidenciaram que na prova de compreensão em leitura específica, que envolve conceitos estatísticos, não houve o ajuste na segunda e terceira dimensão devido o reduzido número dos itens que compunham tais dimensões, optando-se desta maneira, apenas pela análise da dimensão predominante. Considerando-se assim, a prova como sendo predominantemente unidimensional, apresentando 24,80% da variância que pode ser explicada por esse modelo.

Para a prova de leitura geral, os resultados novamente indicaram uma predominância unidimensional, explicando 43,41% da variância que pode ser explicada por este modelo. As análises realizadas para a obtenção das qualidades psicométricas desta prova, garantiu a utilização desta, para a obtenção de evidências de validade convergente para a prova de compreensão de um texto estatístico.

Após, realizou-se a análise pela Teoria de Resposta ao Item (TRI), porém para tanto, devem ser obedecidos dois pressupostos dessa teoria: o primeiro é que se atinja a unidimensionalidade, e de acordo com Backer (2001), pelo fato destes serem equivalentes, obtendo-se o primeiro, matematicamente cumpri-se a independência local entre os itens da prova, que é o segundo pressuposto. Por meio da análise fatorial realizada nas duas provas de compreensão, observou-se a unidimensionalidade, sendo possível a realização da TRI.

De acordo com os resultados demonstrados, a prova de leitura específica se ajustou melhor ao modelo de três parâmetros para os 40 itens da prova, apresentando um índice médio de discriminação igual a 0,74, índice médio de dificuldade igual a -0,68, e probabilidade de acerto ao acaso de 0,25. Pode-se verificar também que a prova apresentou

um bom índice de fidedignidade pelo método de Kuder-Richardson ($KR-20=0,813$). A prova de leitura geral também se adequou melhor ao modelo de três parâmetros, com um índice médio de discriminação igual a 0,91, índice médio de dificuldade igual a -1,67 , e probabilidade de acerto ao acaso de 0,25, com um bom índice de fidedignidade ($KR-20=0,897$).

Com relação às evidências de validade convergente os resultados indicaram uma correlação positiva significativamente diferente de zero entre as duas provas de compreensão em leitura ($r=0,535$; $p<0,001$). Bem como uma correlação positiva significativa entre a prova de compreensão específica com o desempenho no questionário que envolve conceitos estatísticos ($r=0,257$; $p<0,001$). Assim, quanto maior o conhecimento que o universitário possui em estatística e na compreensão de textos de leitura geral, maior será a sua compreensão em um texto de leitura de conceitos estatísticos.

Também foi realizada uma análise do desempenho geral do universitário nas provas de compreensão, o que evidenciou um desempenho médio para este nível de ensino, porém a média apresentada na prova de leitura geral foi ligeiramente maior do que a de leitura específica. Verificaram-se ainda possíveis efeitos das variáveis, gênero, faixa etária, turno e ano de graduação. Apenas comprovaram-se diferenças entre os anos de graduação para as duas provas. Esses resultados mostram as dificuldades apresentadas pelos estudantes de psicologia com relação aos conceitos matemáticos e estatísticos que estão inseridos em determinadas leituras, como a de textos científicos e até os próprios manuais de testes psicológicos, que são utilizados na atuação prática. Sendo que essa dificuldade já vem sendo investigada em outros estudos, como o de Vendramini e Dias (2005) com relação à dificuldade de leitura de dados estatísticos.

Considerando que, muitas das informações contidas nos manuais de testes psicológicos são derivadas de conceitos estatísticos, e se o estudante de psicologia não

apresentar um domínio de alguns conceitos de estatística, muitos resultados de pesquisa podem ficar sem significado, dificultando assim, a análise e interpretação dos dados. Além do fato de que os instrumentos que são utilizados na prática devem apresentar certas qualidades psicométricas, que garantam a utilização de dados confiáveis, como validação e normatização. Sendo que, tais características dos testes envolvem conceitos estatísticos, evidenciando ainda mais, a necessidade da compreensão de textos que engloba tal conteúdo específico.

Outra evidência importante é que, uma das limitações deste estudo foi em relação a procedência da amostra obtida, sendo apenas de uma universidade, e com uma discrepância entre a quantidade de estudantes do turno diurno e noturno. Sendo decorrência do número de turmas oferecidas pelo turno noturno ser maior que o diurno.

Cabe ressaltar ainda, uma outra limitação neste estudo, no que se refere à quantidade de instrumentos aplicados em uma única sessão nos estudantes. Uma vez que, a realização do questionário seguido das duas provas de compreensão em leitura pode ter levado a um efeito fadiga. Mesmo quando se optou pela inversão das provas de compreensão em leitura, o cansaço pode ter interferido nos resultados encontrados, pois os estudantes responderam a muitos instrumentos em um único momento.

O objetivo deste trabalho foi o de fornecer subsídios para uma complementação nos estudos dessa área, inteirando assim, análises sobre provas que avaliem conceitos específicos que fazem parte da leitura científica dos universitários para a aquisição do conhecimento. Embora, tenha-se optado por analisar a qualidade de uma prova específica com uma amostra limitada de estudantes de psicologia, outros estudos podem ser realizados para verificar se existiriam diferenças nos parâmetros ajustados, quando se considera uma amostra maior de diferentes localidades, incluindo universidade públicas.

REFERÊNCIAS

Abraham, R. G., & Chapelle, C. A. (1992). The meaning of cloze test scores: an item difficulty perspective. *The Modern Language Journal*, 76(4), 468-479.

Adanéz, G. A. (1999). Procedimientos de construcción y análisis de tests psicométricos. Em S. M. Wechsler & R. S. L. Guzzo (Orgs.), *Avaliação Psicológica: Perspectiva Internacional* (pp. 57-100). São Paulo: Casa do Psicólogo.

Alliende, F., & Condemarin, M. (1987). *A leitura: teoria, avaliação e desenvolvimento*. Porto Alegre: Artmed.

Alliende, F., & Condemarín, M. (2005). *A leitura: teoria, avaliação e desenvolvimento*. (8. ed.). Porto Alegre: Artmed.

Alves, I. C. B. (2002). Instrumentos disponíveis no Brasil para avaliação da inteligência. Em R. Primi (Org.), *Temas em Avaliação Psicológica* (pp. 80-102). São Paulo: Casa do Psicólogo.

Anastasi, A., & Urbina, S. (2000). *Testagem psicológica*. (7ª ed.). Porto Alegre: Artmed.

Ashby-Davis, C. (1985). Cloze and comprehension: a qualitative analysis and critique. *Journal of Reading*, 28(7), 585-589.

Assessment System Corporation (1995). *User's manual for the XCALIBRE-Marginal Maximum-Likelihood IRT Parameter Estimation Program*. (2ª ed.). Windows 3.x/95/NT, St. Paul, MN: Author.

Backer, F. B. (2001). *The Basics of Item Response Theory*. USA: ERIC Clearinghouse on Assessment and Evaluation.

Barbosa, J. J. (1992). *Alfabetização e leitura*. (2ª ed.). São Paulo: Cortez.

Batanero, C. (2001). *Didáctica de la Estadística, Granada*: Grupo de Investigación e Educación Estadística. ISBN 84-699-4295-6.

Behrens, M. A. (2001). Projetos de aprendizagem colaborativa num paradigma emergente. Em B. R. Moran. *Novas tecnologias e mediação pedagógica* (pp. 67-132). São Paulo: Papirus.

Bifi, C. R. (2006). *Estatística em um curso de administração de empresas - mobilização dos conceitos estatísticos de base*. Dissertação de Mestrado, Pontifícia Universidade Católica de São Paulo, São Paulo.

Bitar, M. L. (1989). Redundancy and the cohesion Cloze. *Journal of research in reading*, 13(1), 18-37.

Bormuth, J. R. (1968). Cloze Test readability: Criterion references scores. *Journal of Education Measurement*, 5, 189-196.

Braibant, J. M. (1997). A decodificação e a compreensão: Dois componentes essenciais da leitura no 2º ano primário. Em J. Grégoire & B. Piérart (Orgs.), *Avaliação dos problemas de leitura: Os novos modelos teóricos e suas implicações diagnósticas* (pp. 166-187). Porto Alegre: Artes Médicas.

Brown, J. D. (1980). Relative merits of methods for scoring cloze tests. *The Modern Language Journal*, 64(3), 311-317.

Calfee, R. (1977). Assessment of independent reading skills: basic research and practical implications. Em A. Reber & D. Scarborough (Eds.), *Toward a psychology of reading*. Hillsdale, N. J.: Erlbaum.

Capovilla, A. S. G., Joly, M. C. R. A., & Tonelotto, J. F. (2006). Avaliação neuropsicológica e aprendizagem. Em A. P. P. Noronha, A. A. A. Santos & F. F. Sisto (Orgs.), *Facetas do fazer em avaliação psicológica* (pp. 141-162, 1ª ed.). São Paulo: Vetor.

Castro, C. A. S., Fukuda, M., & Gritti, N. H. S. (1997). Compreensão em leitura entre universitários de curso de letras. Em G. P. Witter (Org.), *Psicologia: leitura e universidade* (pp. 111-122). Campinas: Alínea.

Centofanti, E. M., Ferreira, S. M., & Tedesco, T. D. (1997). Compreensão em leitura por universitários de psicologia. Em G. P. Witter (Org.), *Psicologia: leitura e universidade* (pp. 33-60). Campinas: Alínea.

Conselho Federal de Psicologia – CFP (2001). Resolução 25/01. Disponível: www.pol.org.br.

Costa, P. S. (2006). O Teste *Cloze* na Avaliação de Aprendizagem: O Caso dos Alunos do Curso de Ciências Contábeis da Universidade Federal do Rio Grande do Norte. Em 7º Congresso USP de Controladoria e Contabilidade, São Paulo. Anais do Congresso. [Online]. Disponível: <http://www.congressoeac.locaweb.com.br/artigos62006/573.pdf>. Recuperado: 25 ago 2007.

Cronbach, J. L. (1996). *Fundamentos da testagem psicológica*. Porto Alegre: Artmed.

Ellis, A. W. (1995). *Leitura escrita e dislexia: uma análise cognitiva*. (2ª ed.). Porto Alegre: Artmed.

Flanagan, D. P., Ortiz, S. O, Alfonso, V. C, & Mascolo, J. T. (2002). *The achievement test desk reference (ATDR): Comprehensive Assessment and Learning Disabilities*. Needham Heights: Allyn and Bacon.

Fletcher, P. (1994). A teoria da resposta ao item: medidas invariantes do desempenho escolar. *Ensaio: Avaliação e Políticas Públicas em Educação*, 1(2), 21-28.

Gabriel, R. (2006). A compreensão em leitura enquanto processo cognitivo. *Revista Signo*, 31, 73-83.

Gal, I. (2002). Adults' statistical literacy: meanings, components, responsibilities. *International Statistical Review*, 70, 1-51.

Gal, I., & Garfield, J. (1997). Curricular goals and assessment challenges in Statistics Education. Em I. Gal & J. B. Garfield (Eds.), *The assessment challenge in Statistics Education*. International Statistical Institute, Amsterdam: IOS Press.

Gibson, E. J., & Levin, H. (1975). *The psychology of reading*. Cambridge: The MIT Press.

Hair, J. J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (2005). *Análise Multivariada de dados*. (5ª ed.). Porto Alegre: Bookman.

Joly, M. C. R. A. (1999). Leitura: o que sabemos, o que precisamos saber (influência da família na alfabetização). Em G. P. Witter (Org.), *Leitura: textos e pesquisas*, (pp. 23-36, 1ª ed.). Campinas: Alínea.

Joly, M. C. R. A. (2001). Leitura no contexto educacional: avaliando estratégias para aquisição de habilidades criativas. Em F. F. Sisto, E. T. B. Sbasdelini & R. Primi (Orgs.), *Contextos e questões da avaliação psicológica* (pp. 99-116). São Paulo: Casa do Psicólogo.

Joly, M. C. R. A. (2005). *Prova de Compreensão em Leitura – Ensino Médio*. Pesquisa em desenvolvimento. Universidade São Francisco, Itatiba.

Joly, M. C. R. A. (2006). Cloze Oriented System (COS) in an electronic comprehension program and reading attitude in Brazil. Em A. Mendéz-Villas, B. G. Pereira, J. M. González & J. A. M. González (Ed.), *Current Developments in Technology-Assisted Education* (pp. 1669-1674). Badajoz: INDUGRAFIC.

Joly, M. C. R. A., & Dias, A. S. (2007a, Outubro). Análise Fatorial em uma prova de compreensão em leitura para o ensino médio. Resumo apresentado no *I Congresso Multidisciplinar “Ciência Cidadania e Diversidade”*, Universidade São Francisco, São Paulo.

Joly, M. C. R. A., & Dias, A. S. (2007b, Outubro). Aplicação da Teoria de Resposta ao Item em uma prova de compreensão em leitura do ensino médio. Resumo apresentado no *I Congresso Multidisciplinar “Ciência Cidadania e Diversidade”*, Universidade São Francisco, São Paulo.

Joly, M. C. R. A., & Paula, L. M. (2005). Avaliação do uso de estratégias de aprendizagem e a compreensão em leitura de universitários. Em M. C. R. Joly, A. A. A. Santos & F. F. Sisto (Orgs.), *Questões do cotidiano universitário* (pp. 33-57). São Paulo: Casa do Psicólogo.

Kaplan, H., Sadock, B., & Grebb, J. A. (1997). *Compêndio de Psiquiatria: ciências do comportamento e psiquiatria clínica*. (7ª ed.). Porto Alegre: Artmed.

Kletzien, S. B. (1991). Strategy use by good and poor comprehenders reading expository text of differing levels. *Reading research quarterly: International Reading Association*, XXVI(1), 67-86.

Krutetskii, V. A. (1976). *The Psychology of mathematical abilities en scholchindren*. Chicago: University of Chicago Press.

Limana, A., & Brito, M. R. (2005). O modelo de avaliação dinâmica e o desenvolvimento de competências: algumas considerações a respeito do ENADE. *Revista Avaliação*, 10(2), 9-32.

Lopes, F. L., & Vendramini, C. M. M. (2007a, Julho). Leitura de manuais de testes psicológicos por profissionais de psicologia. Resumo apresentado no *III Congresso Brasileiro de Avaliação Psicológica e XII Conferência Internacional da Avaliação Psicológica: Formas e Contextos*, João Pessoa, Paraíba.

Lopes, F. L., & Vendramini, C. M. M. (2007b, Julho). Importância e contribuição da Estatística para a Psicologia. Resumo apresentado no *I Congresso Multidisciplinar "Ciência Cidadania e Diversidade"*, Universidade São Francisco, São Paulo.

Lord, F. M. (1952). *A theory of test scores: Psychometric Monograph*. Iowa: Psychometric Society, 7.

Marini, J. A. S. (2006). *Escala de estratégias metacognitivas de leitura para o ensino médio: evidências de validade*. Dissertação de Mestrado, Universidade São Francisco, Itatiba.

Martins, G. A., & Andrade, J. X. (2004). Compreensibilidade de livros-texto de contabilidade: um estudo empírico com a utilização da técnica cloze. Em *4º Congresso USP de Controladoria e Contabilidade*, São Paulo. Anais do Congresso (CD_ROM), 2004, v. 1, pp. 1-20.

Martins, M. C., Capovilla, F., Gombert, J. E., Oliveira, J. B. A., Morais, J. C. J., Adams, M. J., & Beard R. (2003). *Alfabetização do Brasil: Políticas e Práticas*. Brasília: Ministério da Educação. Comissão de educação e cultura. Relatório de Pesquisa. (pp. 50-53).

MEC – Ministério da Educação (2006). *Relatório Síntese do curso de Psicologia*. Brasília-DF: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira.

Munhoz, A. M. H. (2004). *Uma análise multidisciplinar da relação entre inteligência e desempenho acadêmico em universitários ingressantes*. Tese de Doutorado, Faculdade de Educação, Universidade Estadual de Campinas, Campinas.

Nascimento, E. (2004). *Escala de Inteligência Wechsler para Adultos – III: Adaptação e Padronização de uma amostra Brasileira* (1ª ed.). São Paulo: Casa do Psicólogo. Original: Wechsler, D. (1997).

Nassri, R. C. B. M. (2002). Compreensão em leitura de universitários de Direito e Medicina. Em G. P. Witter (Org.), *Psicologia: tópicos gerais* (pp. 180-191). Campinas: Alínea.

Noronha, A. P. P. (2002). Os problemas mais graves e mais freqüentes no uso dos testes psicológicos. *Psicologia: Reflexão e Crítica*, 1(15), 135-142.

Noronha, A. P. P., & Vendramini, C. M. M. (2003). Parâmetros psicométricos: estudo comparativo entre testes de inteligência e de personalidade. *Psicologia: Reflexão e Crítica*, 16(1), 177-182.

Noronha, A. P. P., & Freitas, F. A. (2005). Testes psicológicos, usos e conhecimento. *Psico*, 36(1), 21-28.

Noronha, A. P. P., Primi, R., & Alchieri, J. C. (2005). Instrumentos de avaliação mais conhecidos/utilizados por psicólogos e estudantes de psicologia. *Psicologia Reflexão e Crítica*, 18(3), 390-401.

Noronha, A. P. P., Sartori, F. A., Freitas, F. A., & Ottati, F. (2001). Informações contidas nos manuais de testes de inteligência publicados no Brasil. *Psicologia em Estudo*, 2(6), 101-106.

Noronha, A. P. P., Sartori, F. A., Freitas, F. A., & Ottati, F. (2002). Informações contidas nos manuais de testes de personalidade. *Psicologia em Estudo*, 7(1), 143-149.

Oliveira, K. L. (2003). *Compreensão de leitura, desempenho acadêmico e avaliação de aprendizagem em universitários*. Dissertação de Mestrado, Universidade São Francisco, Itatiba.

Oliveira, K. S. (2006). *Avaliação do Exame Nacional de Desempenho do Estudante pela Teoria de Resposta ao Item*. Dissertação de Mestrado, Universidade São Francisco, Itatiba, São Paulo.

Oliveira, K. L., Santos, A. A. A., & Primi, R. (2003). Estudo das relações entre compreensão em leitura e desempenho acadêmico na universidade. *Interação em Psicologia*, 7(1), 19-25.

Ottati, F., Noronha, A. P. P., & Salviati, M. (2003). Testes psicológicos: qualidade de instrumentos de interesse profissional. *Interação em Psicologia*, 7(1), 65-71.

Pasquali, L. (Org.). (2001) *Técnicas de Exame Psicológico – TEP: manual*. (vol. I). São Paulo: Casa do Psicólogo, Conselho Federal de Psicologia.

Pasquali, L. (2003). *Psicometria: teoria dos testes na psicologia e na educação*. Petrópolis: Vozes.

Pellegrini, M. C. K. (1996). *Avaliação dos níveis de compreensão e atitudes frente à leitura em universitários*. Dissertação de Mestrado, Universidade São Francisco, Itatiba.

Perfetti, C., & Lesgold, A. (1978). Discourse comprehension and sources of individual differences. Em M. Just & P. Carpenter (Eds.), *Cognitive processes in comprehension*. Hillsdale, N. J.: Erlbaum.

Pestana, M. H., & Gageiro, J. G. (2003). *Análise de dados para ciências sociais: a complementaridade do SPSS* (3ª ed.). Lisboa: Silabo.

Pinheiro, A. M. V. (1994). *Leitura e escrita: Uma abordagem cognitiva*. Campinas: Editorial Psy.

Primi, R. (2003). Inteligência: Avanços nos Modelos Teóricos e nos Instrumentos de Medida. *Avaliação Psicológica*, 2(1), 67-77.

Quintino, C. A. A., Guedes, T. A., & Martins, A. B. T. (2001). Análise estatística das atitudes dos alunos de iniciação científica da Universidade Estadual de Maringá, em relação à disciplina Estatística – 2000. *Maringá*, 23(6), 1523-1529.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research.

Rego, L. L. B. (1995). Diferenças individuais na aprendizagem inicial da leitura: Papel desempenhado por fatores metalingüísticos. *Psicologia: Teoria e Pesquisa*, 11(1), 51-60.

Roca, M. A. E. (2002). *Analisis de las actitudes y conocimientos estadísticos elementares en la formación del profesorado*. Tese de Doutorado, Universitat Autònoma de Barcelona, Belleterra.

Saito, Y. (2003). Investigating the construct validity of the cloze section in the examination for the certificate of proficiency in English. *Spain Fellow Working Papers in Second or Foreign Language Assessment*, 1, 39-82.

Salles, J. F., & Parente, M. A. M. P. (2002). Processos cognitivos na leitura de palavras em crianças: relações com compreensão e tempo de leitura. *Psicologia: Reflexão e Crítica*, 15(2), 321-331.

Sampaio, I. S., & Santos, A. A. A. (2002). Leitura e redação entre universitários: avaliação de um programa de intervenção. *Psicologia em Estudo*, 7(1), 31-38.

Santos, A. A. A. (1991). Desempenho em leitura: um estudo diagnóstico da compreensão e hábitos de leitura entre universitários. *Estudos de Psicologia*, 8(1), 6-19.

Santos, E. C. P. (2001). *Compreensão em leitura de textos em espanhol: um estudo com universitários*. Dissertação de Mestrado, Universidade São Francisco, Itatiba.

Santos, A. A. A., & Di Nucci, S. H. P. (1984). A Utilização de Questões de Estudo Cloze como Técnicas de Desenvolvimento da Compreensão em Leitura com Estudantes Universitários. *Psico-USF*, 2(5), 11-30.

Santos, A. A. A., Primi, R., Taxa, F. O. S., & Vendramini, C. M. M. (2002). O Teste de Cloze na avaliação da compreensão em leitura. *Psicologia: Reflexão e Crítica*, 15(3), 549-560.

Santos, A. A. A., & Oliveira, K. L. (2004). A importância da compreensão em leitura para a aprendizagem de universitários. Em E. Boruchovitch & J. A. Bzuneck (Orgs.), *Aprendizagem: processos psicológicos e o contexto social na escola* (pp. 119-148). Petrópolis: Vozes.

Santos, A. A. A., Vendramini, C. M. M., Suehiro, A. C. B., & Santos, L. A. D. (2006). Leitura compreensiva e utilização de estratégias de aprendizagem em alunos de Psicologia. *Estudos de Psicologia*, 23(1), 83-91.

Schoenfeld, F. G. (1980). Instructional uses of cloze. *The Reading Teacher*, 34(2), 147-151.

Silva, E. M. T. (1998). *Compreensão de leitura em estudantes de direito*. Dissertação de Mestrado, Instituto de Psicologia, Pontifícia Universidade Católica de Campinas, Campinas.

Silva, M. J. M., & Santos, A. A. A. (2004). A avaliação da compreensão em leitura e o rendimento acadêmico de universitários. *Psicologia em estudo*, 9(3), 459-467.

Silva, E. M. T. S., Malozze, G. L. M., & Leme, M. L. C. S. (1997). Compreensão de leitura entre universitário do primeiro e quinto ano de Direito. Em G. P. Witter (Org.), *Psicologia: leitura e universidade* (pp. 101-110). Campinas: Alínea.

Smole, K. S., & Diniz, M. I. (2001). Ler e Aprender Matemática. Em K. C. S. Smole & M. I. Diniz (Orgs.), *Ler, Escrever e resolver problemas: habilidades básicas para aprender matemática* (pp. 69-86). São Paulo: Artmed.

Smith, E. E., & Spoehr, K. (1974). The perception of English: a theoretical perspective. Em B. H. Kantowitz (Ed.), *Human information processing*. Hillsdale, N. J.: Erlbaum.

Snowling, M., & Hulme, C. (2005). *The Science of Reading: a Handbook*. Oxford: Blackwell.

Soares, T. M., & Pereira, D. R. M. (2001). *Estudo de critérios de ajuste para modelos da Teoria de Resposta ao Item (TRI)*. Trabalho apresentado no Simpósio “A pesquisa operacional e o meio ambiente” no XXXIII SBPO. Campos do Jordão, São Paulo, Brasil.

Statistical Package for Social Science (2004). *SPSS for Windows*: standard version, release 13. Chicago, IL: SPSS Inc.

Taylor, W. I. (1953). Cloze procedure: a new tool for measuring readability. *Journalism Quarterly*, 30, 415-433.

Texas Reading Initiative (2006). *Comprehension Instruction – On line Revise Edition*. Texas Education Agency. Austin, Texas.

Urbina, S. (2007). *Fundamentos da testagem psicológica*. Porto Alegre: Artmed.

Valle, R. C. (2000). Teoria de Resposta ao Item. *Estudos em Avaliação educacional*, 21, 7-91.

Vendramini, C. M. M. (2002). Aplicação da teoria de resposta ao item na avaliação educacional. Em R. Primi (Org.), *Temas em avaliação psicológica* (pp. 116-130). Campinas: Instituto Brasileiro de Avaliação Psicológica.

Vendramini, C. M. M. (2005). O Uso de um Modelo Multidimensional para Análise do ENADE. *Revista de Avaliação Institucional*, 10(3), 27-40.

Vendramini, C. M. M., & Brito, M. R. F. (2001). Relações entre Atitude, Conceito e Utilidade da Estatística. *Psicologia Escolar Educacional*, 5(1), 59-73.

Vendramini, C. M. M., & Noronha, A. P. P. (2003). Análise de itens na construção de testes de inteligência. *Psicologia: teoria, investigação e prática*, 1, 73-81.

Vendramini, C. M. M., & Dias, A. S. (2005). Teoria de resposta ao item na análise de uma prova de estatística em universitários. *Psico-USF, 10*(2), 201-210.

Vicentelli, H. (1999). Problemática de la lectura en estudiantes universitarios. *Psicología Escolar e Educacional, 3*(3), 195-202.

Watson, J., & Callingham, R. (2003). Statistical literacy: a complex hierarchical construct. *Statistics Education Research Journal, 2*(2), 3-46.

Wilson, D. T., Wood, R., & Gibbons, R. (1998). *TESTFACT 2 – test scoring, item statistics, and item factor*. Chicago: Scientific Software International.

WINSTEPS (2006). *Rasch Measurement Software for persons & items*, release 3.61.1. Chicago, IL.

Witter, G. P. (1990). Pesquisa documental, pesquisa bibliográfica e busca de informação. *Estudos de Psicologia, 7*(1), 5-30.

Witter, G. P. (1996). Avaliação da produção científica sobre leitura na universidade. *Psicologia Escolar e Educacional, 1*(1), 31-37.

Witter, G. P. (1997). Leitura e Universidade. Em G. P. Witter (Org.), *Psicologia: leitura e universidade* (pp. 09-18). Campinas: Alínea.

Witter, G. P. (1999). Metaciência e leitura. Em G. P. Witter (Org.), *Leitura: textos e pesquisas* (pp. 13-22). Campinas: Alínea.

Witter, G. P., & Lomônaco, J. F. B. (1984). *Psicologia da aprendizagem*. São Paulo: EPU.

Wolff, A. C. R., Polimero, E., Olioveira, E. T. A., Henriques, T., & Socci, V. (1997). Compreensão da leitura em estudantes de comunicação social, educação física e fonoaudiologia. Em G. P. Witter (Org.), *Leitura e Universidade* (pp. 19-32). Campinas: Alínea.

ANEXOS

ANEXO 1 – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO (1ª VIA)

Pesquisa

Evidências de validade de uma Prova de Compreensão em Leitura em Estatística

Eu, _____ com _____ anos de idade, R.G. _____ residente à rua _____ nº _____, da cidade de _____ estado de _____ abaixo assinado, dou meu consentimento livre e esclarecido para participar como voluntário da pesquisa supracitada, sob a responsabilidade da pesquisadora Anelise Silva Dias, aluna de Mestrado em Psicologia e Profª Dra. Claudette Maria Medeiros Vendramini, orientadora, ambas do Programa de Pós-Graduação *Stricto Sensu* em Psicologia da Universidade São Francisco.

Assinando este Termo de Consentimento estou ciente de que:

1. O objetivo da pesquisa é verificar evidências de validade de uma prova de compreensão em leitura de um texto com informações estatísticas em estudantes de psicologia.
2. Durante o estudo serão aplicadas duas provas de compreensão em leitura, uma de leitura geral e a outra de leitura específica.
3. A pesquisa não apresenta riscos conhecidos ou constrangimentos para os participantes.
4. Obtive todas as informações necessárias para poder decidir conscientemente sobre a minha participação na referida pesquisa.
5. Estou livre para interromper a qualquer momento minha participação na pesquisa.
6. A interrupção da minha participação não causará prejuízo ao meu eventual desempenho na disciplina, cuidado e tratamento pela equipe responsável.
7. Meus dados pessoais serão mantidos em sigilo e os resultados gerais obtidos por meio da pesquisa serão utilizados apenas para alcançar os objetivos do trabalho, expostos acima, incluída sua publicação na literatura científica especializada.
8. Poderei contatar o Comitê de Ética em Pesquisa da Universidade São Francisco para apresentar recursos ou reclamações em relação à pesquisa através do telefone: (11) 4534-8117.
9. Poderei entrar em contato com a responsável pelo estudo, Anelise Silva Dias sempre que julgar necessário.
10. Este Termo de Consentimento é feito em duas vias, sendo que uma permanecerá em meu poder e outra com o pesquisador responsável.

Itatiba, _____ de _____ de 2007.

Assinatura do voluntário

Anelise Silva Dias
Pesquisadora responsável pelo estudo

ANEXO 2 – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO (2ª VIA)

Pesquisa

Evidências de validade de uma Prova de Compreensão em Leitura em Estatística

Eu, _____ com _____ anos de idade, R.G. _____ residente à rua _____ nº _____, da cidade de _____ estado de _____ abaixo assinado, dou meu consentimento livre e esclarecido para participar como voluntário da pesquisa supracitada, sob a responsabilidade da pesquisadora Anelise Silva Dias, aluna de Mestrado em Psicologia e Profª Dra. Claudette Maria Medeiros Vendramini, orientadora, ambas do Programa de Pós-Graduação *Stricto Sensu* em Psicologia da Universidade São Francisco.

Assinando este Termo de Consentimento estou ciente de que:

1. O objetivo da pesquisa é verificar evidências de validade de uma prova de compreensão em leitura de um texto com informações estatísticas em estudantes de psicologia.
2. Durante o estudo serão aplicadas duas provas de compreensão em leitura, uma de leitura geral e a outra de leitura específica.
3. A pesquisa não apresenta riscos conhecidos ou constrangimentos para os participantes.
4. Obtive todas as informações necessárias para poder decidir conscientemente sobre a minha participação na referida pesquisa.
5. Estou livre para interromper a qualquer momento minha participação na pesquisa.
6. A interrupção da minha participação não causará prejuízo ao meu eventual desempenho na disciplina, cuidado e tratamento pela equipe responsável.
7. Meus dados pessoais serão mantidos em sigilo e os resultados gerais obtidos por meio da pesquisa serão utilizados apenas para alcançar os objetivos do trabalho, expostos acima, incluída sua publicação na literatura científica especializada.
8. Poderei contatar o Comitê de Ética em Pesquisa da Universidade São Francisco para apresentar recursos ou reclamações em relação à pesquisa através do telefone: (11) 4534-8117.
9. Poderei entrar em contato com a responsável pelo estudo, Anelise Silva Dias sempre que julgar necessário.
10. Este Termo de Consentimento é feito em duas vias, sendo que uma permanecerá em meu poder e outra com o pesquisador responsável.

Itatiba, _____ de _____ de 2007.

Assinatura do voluntário

Anelise Silva Dias
Pesquisadora responsável pelo estudo

ANEXO 3 – QUESTIONÁRIO

Nome: _____ R.A.: _____

1. Curso: _____
2. Semestre em que está matriculado: _____
3. Período: () Diurno () Noturno
4. Idade: _____ anos
5. Sexo: () Masculino () Feminino
6. Quais disciplinas e quantas vezes cada você já reprovou (DP) até agora? _____

7. Dos termos abaixo, utilizados em Estatística, quais você conhece e julga-se capaz de interpretar:

- | | | |
|-------------------|----------------|-------------------|
| () Amostra | () Frequência | () Porcentagem |
| () Amostragem | () Média | () Probabilidade |
| () Amplitude | () Mediana | () Proporção |
| () Correlação | () Mensuração | () Variância |
| () Curva Normal | () Moda | () Variáveis |
| () Desvio padrão | () Percentil | () Significância |
| () Escore bruto | () População | |
| () Outros: _____ | | |

8. Para você, o que lembra a palavra Estatística?

9. Qual importância da estatística para cada um dos itens a seguir, atribua um valor em uma escala que varia de 1 (sem importância) a 5 (muito importante).

- (a) seu cotidiano: _____
- (b) sua área de formação: _____
- (c) as suas disciplinas: _____

10. Quais conteúdos de Estatística que você utiliza em seu dia-a-dia ou em sua atuação acadêmica? _____

11. Você se interessa por textos que envolvem informações estatísticas?

Se **SIM**, que tipo de texto você costuma ler? _____

Se **NÃO**, explique o motivo. _____

12. Em suas aulas da graduação qual é a porcentagem aproximada de professores que utilizam conceitos estatísticos em suas aulas? Quais conceitos? _____

13. Responda as questões abaixo, assinalando a resposta que julgar adequada.

Itens I, II, III e IV (Roca, 2002)

I) Nove estudantes pesaram um pequeno objeto com um mesmo instrumento em uma aula de ciências. Os pesos registrados por cada estudante (em gramas) se mostram a seguir:

6,3 6,0 6,0 15,3 6,1 6,3 6,2 6,15 6,3

Os estudantes querem determinar com a maior precisão possível o peso real do objeto. Qual dos seguintes métodos é recomendado que utilizem?

- (a) Usar o número mais comum, que é 6,3.
- (b) Usar 6,15, posto que é o peso mais preciso.
- (c) Somar os 9 números e dividir a soma por 9.
- (d) Excluir o valor 15,3, somar os outros 8 números e dividir por 8.
- (e) Nenhuma das anteriores.

II) Em um frasco de medicamento está impresso a seguinte mensagem:

“**Advertência:** ao aplicar em superfícies cutâneas há uma possibilidade de 15% de que se produza uma erupção. Se aparecer uma erupção, consulte o seu médico.”

Qual das seguintes é a melhor interpretação para esta advertência?

- (a) Não usar esse medicamento sobre a pele; há grande possibilidade de que se produza uma erupção.
- (b) Em aplicações sobre a pele, usar somente 15% da dose recomendada.
- (c) Se aparecer uma erupção, ela provavelmente afetará somente 15% da pele.
- (d) Aproximadamente 15 de cada 100 pessoas que usam a medicação apresentarão uma erupção.
- (e) Há poucas possibilidades de produzir uma erupção usando esse medicamento.

III) O Centro de Previsão do Tempo e Estudos Climáticos (CPTEC) quis avaliar a precisão das previsões do tempo. Buscaram em seus arquivos aqueles dias em que os meteorologistas haviam informado que havia uma possibilidade de 70% de chuva. Compararam estas previsões com os registros que indicaram se choveu ou não nesses dias em particular.

A previsão de 70% de possibilidades de chuva pode considerar muito precisa, se chover:

- (a) Entre 95% e 100% desses dias
- (b) Entre 85% e 94% desses dias
- (c) Entre 75% e 84% desses dias
- (d) Entre 65% e 74% desses dias
- (e) Entre 55% e 64% desses dias

IV) O comitê escolar de uma pequena cidade quis determinar o número médio de crianças por família em sua cidade. Dividiram o número total de crianças da cidade por 50, que é o número total de famílias. Qual das seguintes frases deve estar correta se o número real de crianças por família é 2,2?

- (a) A metade das famílias da cidade tem mais de 2 filhos.
- (b) Na cidade há mais famílias com 3 filhos do que com 2 filhos.
- (c) Há um total de 110 crianças na cidade.
- (d) Há 2,2 crianças por adulto na cidade.
- (e) O número mais comum de filhos em uma família é 2.
- (f) Nenhuma das anteriores

V) Dez psicólogos e 10 estudantes do curso de psicologia participaram de uma pesquisa sobre utilização de testes psicológicos. As idades dos participantes da pesquisas estão anotadas a seguir:

Profissionais: 25 28 29 35 35 35 35 38 39 40

Estudantes: 18 19 23 23 23 28 30 30 35 40

Com base nas idades dos participantes da pesquisa, responda as seguintes questões.

Responda as questões abaixo.	Explique como chegou a sua resposta.
Qual dos grupos tem a maior idade média? Resposta:	
Qual dos grupos (estudantes ou profissionais) tem maior variabilidade de idades?	
Qual idade mais comum para os: - profissionais: - estudantes:	
Qual a porcentagem de estudantes com idade menor ou igual a 30 anos?	
Qual dos grupos tem a menor idade mediana?	
50% dos profissionais tem idades menores ou iguais a quanto?	

ANEXO 4 – CLOZE LEITURA GERAL – TOC

Nome: _____ RA: _____

No Editorial abaixo, cada espaço numerado só pode ser completado com uma das opções ao lado. Primeiro leia o texto todo e depois retorne ao início e complete cada lacuna. Quando você terminar a tarefa, releia o texto e veja se está com sentido.

VAGA-LUMES LITERÁRIOS

Dr^a Maria Cristina Rodrigues Azevedo Joly

Expedição Vaga-Lume é o simpático nome do projeto de três jovens que deverá levar, nos próximos meses, cerca de 7.000 mil livros a 22 comunidades carentes na região amazônica. A idéia, já contida no próprio nome da iniciativa, é levar partículas de luz à floresta.

Os 7.000 volumes distribuídos em 22 bibliotecas podem parecer pouco, mas, quando se considera o deserto de livros que é o Brasil e, em especial, a região amazônica, esses micropontos de luz ganham as proporções que tem um farol para o naufrago.

Amazônia Legal, que <u>1</u> a	1 (a) corresponde (b) possui (c) atualmente (d) infelizmente (e) sem
61% do território <u>2</u> , não tem	2 (a) europeu (b) doze (c) brasileiro (d) cônsul (e) América
mais do <u>3</u> 400 bibliotecas	3 (a) Qual (b) Que (c) Teclado (d) A (e) Com
públicas. Um <u>4</u> do porte do	4 (a) Sábio (b) País (c) Queimou (d) Estado (e) Admirado
Amazonas <u>5</u> com apenas oito.	5 (a) Dificilmente (b) Menino (c) Conta (d) Sequer (e) Tem
Roraima, <u>6</u> menos: quatro.	6 (a) Já (b) Pesa (c) Ainda (d) Aguarda (e) Três
É verdade <u>7</u> a situação	7 (a) Que (b) Em (c) De (d) Matemática (e) Se
global brasileira <u>8</u> é muito mais	8 (a) Algum (b) Deseja (c) Necessita (d) Não (e) Jamais
confortável. <u>9</u> país possui pouco	9 (a) Brasil (b) O (c) Um (d) França (e) Oba!
mais <u>10</u> 4.000 bibliotecas.	10 (a) De (b) Desde (c) E (d) Como (e) Livros

É difícil <u>11</u> o que é causa	<p style="text-align: center;">11</p> <p>(a) Determinar (b) Ocupar (c) lentamente (d) Aquela (e) Raramente</p>
<u>12</u> o que é efeito, <u>13</u> o	<p style="text-align: center;">12</p> <p>(a) Para (b) Nem (c) Sobre (d) E (e) Fácil</p>
brasileiro quase não <u>14</u> . Existe	<p style="text-align: center;">13</p> <p>(a) economia (b) desde (c) senão (d) após (e) mas</p>
aqui uma livraria <u>15</u> cada 84,4	<p style="text-align: center;">14</p> <p>(a) hoje (b) trabalha (c) ainda bem (d) estudo (e) lê</p>
mil habitantes. <u>16</u> vizinha	<p style="text-align: center;">15</p> <p>(a) apesar (b) com (c) para (d) entretanto (e) leitura</p>
Argentina tem (ou <u>17</u> , antes da	<p style="text-align: center;">16</p> <p>(a) uma (b) surpresa (c) a (d) contra (e) livro</p>
crise) uma <u>18</u> cada 6.200. O	<p style="text-align: center;">17</p> <p>(a) nenhuma (b) claramente (c) fazia (d) distintamente (e) tinha</p>
brasileiro <u>19</u> em média 2,5	<p style="text-align: center;">18</p> <p>(a) mas (b) ainda que (c) literatura (d) para (e) após</p>
livros <u>20</u> ano, aí incluídos os	<p style="text-align: center;">19</p> <p>(a) ruim (b) sabe (c) efetivamente (d) aqui (e) adquire</p>
<u>21</u> , que são distribuídos a	<p style="text-align: center;">20</p> <p>(a) ou (b) por (c) que (d) causa (e) de</p>
<u>22</u> da rede pública. O <u>23</u>	<p style="text-align: center;">21</p> <p>(a) encantados (b) prováveis (c) venderem (d) didáticos (e) lápis</p>
compra mais de sete <u>24</u> por ano.	<p style="text-align: center;">22</p> <p>(a) alunos (b) inspetores (c) paulistas (d) vivamente (e) famosos</p>
	<p style="text-align: center;">23</p> <p>(a) apenas (b) dócil (c) francês (d) menino (e) ousado</p>
	<p style="text-align: center;">24</p> <p>(a) jornais (b) livros (c) novos (d) lerem (e) nobres</p>

E a <u>25</u> Vaga-Lume, ao	<p style="text-align: center;">25</p> <p>(a) imperdoável (b) sem (c) formiga (d) expedição (e) fria</p>
assumir que <u>26</u> um trabalho	<p style="text-align: center;">26</p> <p>(a) realização (b) impede (c) felizmente (d) faz (e) finalmente</p>
pequeno, como <u>27</u> de inseto,	<p style="text-align: center;">27</p> <p>(a) perante (b) ante (c) assim como (d) agradável (e) que</p>
ganha relevância <u>28</u> não se	<p style="text-align: center;">28</p> <p>(a) sem (b) porque (c) sobre (d) nenhuma (e) quanto mais</p>
limita a <u>29</u> os livros nas	<p style="text-align: center;">29</p> <p>(a) despejar (b) colocação (c) embora (d) dizer (e) melhor</p>
comunidades. <u>30</u> também	<p style="text-align: center;">30</p> <p>(a) aparece (b) procura (c) alto (d) realmente (e) menos</p>
trabalhar com alunos <u>31</u>	<p style="text-align: center;">31</p> <p>(a) desde (b) borboletas (c) e (d) a (e) nem</p>
professores, para que o <u>32</u> lhes	<p style="text-align: center;">32</p> <p>(a) livro (b) célebre (c) aluno (d) sagrado (e) apesar</p>
pareça um objeto <u>33</u> exótico.	<p style="text-align: center;">33</p> <p>(a) cem (b) relendo (c) chegado (d) mais (e) menos</p>

Um dos aspectos <u>34</u>	34 (a) era (b) comodamente (c) estava (d) pouco (e) tudo
mencionados do ciclo de <u>35</u> da	35 (a) evasão (b) saber (c) perpetuação (d) agrário (e) melhor
miséria é o <u>36</u> de que o filho	36 (a) estudioso (b) fato (c) realmente (d) diferente (e) conjunto
<u>37</u> pais analfabetos, mesmo	37 (a) contra (b) de (c) entretanto (d) atender (e) porque
quando <u>38</u> à escola, já entra	38 (a) facilmente (b) quatorze (c) nunca (d) chega (e) espera
<u>39</u> desvantagem. Alguns	39 (a) em (b) visto que (c) à (d) logo (e) falando
educadores acreditam <u>40</u> apenas	40 (a) miseráveis (b) até (c) sob (d) se (e) que
ver os pais lendo desde a primeira	
infância já constitui elemento do	
processo de aprendizagem.	

O Brasil não se tornará uma França da noite para o dia. Mas, para que a população venha a ter hábitos de leitura é preciso começar a criar familiaridade com a escrita e os livros, ainda que a ritmo de vaga-lume.

(Folha de São Paulo, 21/02/2002)

ANEXO 5 – CLOZE LEITURA ESPECÍFICA – OPÇÕES

Nome: _____ RA: _____

No texto abaixo, cada espaço numerado só pode ser completado com uma das opções ao lado. Primeiro leia o texto todo e depois retorne ao início e complete cada lacuna. Quando você terminar a tarefa, releia o texto e veja se está com sentido.

WAIS

Anelise Silva Dias

Dr^a Claudette Maria Medeiros Vendramini

Na prática clínica, as medidas de funcionamento intelectual e memória frequentemente são obtidas para se investigar uma ampla gama de capacidades cognitivas. Nesta perspectiva, a Escala de Inteligência Wechsler para Adultos – WAIS-III (3^a ed.) e a Escala de Memória Wechsler – WMS-III (3^a ed.) foram desenvolvidas com metodologias de pesquisa, amostras de padronização, e procedimentos de validação clínica similares. Assim, os dois instrumentos constituem meios de avaliar capacidades cognitivas. Os resultados no WAIS-III e no WSM-III fornecem informações importantes sobre o funcionamento neurocognitivo do indivíduo, mas jamais devem ser interpretados isoladamente.

Os escores brutos de <u> 1 </u> os	1 (a) Dobro (b) Todos (c) Ambos (d) Raros (e) Cujos
testes são transformados <u> 2 </u> escores	2 (a) Logo (b) Em (c) Para (d) Entre (e) Primeiros
ponderados corrigidos segundo <u> 3 </u>	3 (a) A (b) Escala (c) Uma (d) Cada (e) Medida
idade, com média 10 <u> 4 </u> desvio	4 (a) Sobre (b) Nem (c) Menos (d) E (e) Para
padrão 3 (algumas <u> 5 </u> do WMS-III	5 (a) Pequenas (b) Permitida (c) Grandes (d) Instrumentos (e) Instâncias
têm os <u> 6 </u> brutos transformados em	6 (a) Singulares (b) Sobre (c) Escores (d) Diminutos (e) Soma
uma <u> 7 </u> de percentis).	7 (a) Escala (b) Média (c) Maior (d) Ponderada (e) Individual

Escore ponderado <u>8</u> valor 10	<p style="text-align: center;">8</p> <p>(a) Do (b) Entretanto (c) Embora (d) No (e) Mesmo</p>
indicam um <u>9</u> médio em relação	<p style="text-align: center;">9</p> <p>(a) Senão (b) Resultado (c) Significativo (d) Expressivo (e) Desempenho</p>
a <u>10</u> determinado grupo etário	<p style="text-align: center;">10</p> <p>(a) Caso (b) Um (c) O (d) Estava (e) Fato</p>
(ou <u>11</u> de referência). Escores	<p style="text-align: center;">11</p> <p>(a) Pequeno (b) Faixa (c) Por (d) Grupo (e) Quebrado</p>
ponderados <u>12</u> 7 e 13	<p style="text-align: center;">12</p> <p>(a) De (b) Todavia (c) Entre (d) Mas (e) Valores</p>
correspondem <u>13</u> 1 desvio	<p style="text-align: center;">13</p> <p>(a) A (b) Triplo (c) Categoria (d) Uma (e) Percentil</p>
padrão abaixo <u>14</u> acima da	<p style="text-align: center;">14</p> <p>(a) Que (b) Figura (c) Por (d) Sem (e) E</p>
média, respectivamente, <u>15</u>	<p style="text-align: center;">15</p> <p>(a) Porcentagem (b) E (c) Após (d) Se (e) Desde</p>
escores ponderados 4 e 16 <u>16</u>	<p style="text-align: center;">16</p> <p>(a) Muito (b) Décimos (c) Montam (d) Representam (e) Pouco</p>
desvios de 2 desvio <u>17</u> em	<p style="text-align: center;">17</p> <p>(a) Padrão (b) Curtos (c) Mais (d) Implícitos (e) Nível</p>
relação à média. <u>18</u> escores de	<p style="text-align: center;">18</p> <p>(a) Uns (b) Correlações (c) Os (d) Dividir (e) Medianas</p>
QI e <u>19</u> Índices Fatoriais dos	<p style="text-align: center;">19</p> <p>(a) Procedimentos (b) Desde (c) Ou (d) Para Que (e) De</p>
instrumentos, <u>20</u> como escores	<p style="text-align: center;">20</p> <p>(a) Mais (b) Calculados (c) O (d) Provavelmente (e) Definidos</p>
padrão, também <u>21</u> a mesma	<p style="text-align: center;">21</p> <p>(a) Adotam (b) Acaso (c) Dificilmente (d) Ajustam (e) Segundo</p>
escala para <u>22</u> avaliação do nível	<p style="text-align: center;">22</p> <p>(a) Uma (b) Ousada (c) Falta (d) Inteira (e) A</p>
de <u>23</u> .	<p style="text-align: center;">23</p> <p>(a) Apazível (b) Desempenho (c) Onde (d) Inteligência (e) Perspicaz</p>

Os escores de QI <u>24</u> de	<p style="text-align: center;">24</p> <p>(a) Se (b) Desde (c) Nunca (d) E (e) Até</p>
índices fatoriais são <u>25</u> em uma	<p style="text-align: center;">25</p> <p>(a) Expressos (b) Meios (c) Conceitos (d) Importância (e) Relacionados</p>
escala de <u>26</u> 100 e desvio padrão	<p style="text-align: center;">26</p> <p>(a) Nem (b) Semelhante (c) Média (d) Moda (e) Exceto</p>
15 para <u>27</u> grupo etário, um escore	<p style="text-align: center;">27</p> <p>(a) Categoria (b) Cada (c) Alguma (d) Definido (e) Um</p>
<u>28</u> a 100 nestas medidas define	<p style="text-align: center;">28</p> <p>(a) De (b) Puro (c) Erro (d) Média (e) Igual</p>
<u>29</u> desempenho médio de	<p style="text-align: center;">29</p> <p>(a) Um (b) Modo (c) O (d) Faz (e) Papel</p>
indivíduos <u>30</u> faixa etária	<p style="text-align: center;">30</p> <p>(a) Como (b) E (c) Marca (d) Na (e) De</p>
correspondente. Escores <u>31</u> a 85 e	<p style="text-align: center;">31</p> <p>(a) Iguais (b) Além (c) Medida (d) Coeficientes (e) Próprios</p>
115 correspondem a 1 <u>32</u> padrão	<p style="text-align: center;">32</p> <p>(a) Genuíno (b) Direção (c) Desvio (d) Extra (e) O</p>
abaixo e acima <u>33</u> média,	<p style="text-align: center;">33</p> <p>(a) Conforme (b) Terço (c) Daqueles (d) Logo (e) Da</p>
respectivamente, enquanto os <u>34</u> a	<p style="text-align: center;">34</p> <p>(a) Diversos (b) Iguais (c) Escores (d) Contam (e) Erros</p>
70 e 130 estão 2 desvios <u>35</u> abaixo	<p style="text-align: center;">35</p> <p>(a) Semelhantes (b) Molde (c) Fáceis (d) Padrão (e) Os</p>
e acima da <u>36</u> , respectivamente.	<p style="text-align: center;">36</p> <p>(a) Mesma (b) Média (c) Qual (d) Mediana (e) Maior</p>

Cerca de 68% dos <u>37</u> obtêm	37 (a) Resultados (b) Válidos (c) Raros (d) Facilmente (e) Sujeitos
escores entre 85 e 115, <u>38</u> de 95%	38 (a) Cerca (b) Calcula (c) Dobro (d) Chega (e) Mais
situam-se na faixa <u>39</u> 70 a 130, e	39 (a) Quando (b) Por (c) De (d) Pois (e) Procura
aproximadamente todos <u>40</u> escores	40 (a) Obtém (b) Sem (c) Vêm (d) Quanto (e) Demais

entre 55 e 145 (3 desvios padrão positivos

ou negativos em relação à média).

O WMS-III gera quatro escores compostos que representam diferentes aspectos da aprendizagem auditiva, sendo estes apresentados como equivalentes em percentil, que representam a posição do indivíduo no grupo normativo.

Fonte: Nascimento, E. (2004). Escala de Inteligência Wechsler para Adultos – III: Adaptação e Padronização de uma amostra Brasileira (1ª ed.). São Paulo: Casa do Psicólogo.
Original: Wechsler, D. (1997).