

IVAN SANT'ANA RABELO



MATRIZES PROGRESSIVAS AVANÇADAS DE RAVEN E
TESTE DE RACIOCÍNIO INFERENCIAL: EVIDÊNCIAS DE
VALIDADE

ORIENTADOR: PROF. DR. FERMINO FERNANDES SISTO

ITATIBA — SP

2008

IVAN SANT'ANA RABELO



MATRIZES PROGRESSIVAS AVANÇADAS DE RAVEN E
TESTE DE RACIOCÍNIO INFERENCIAL: EVIDÊNCIAS DE
VALIDADE

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

ORIENTADOR: PROF. DR. FERMINO FERNANDES SISTO

ITATIBA — SP

2008

R114m 157.931 Rabelo, Ivan Sant'Ana.
Matrizes progressivas avançadas de Raven e teste de
raciocínio inferencial: evidências de validade / Ivan
Sant'Ana Rabelo. -- Itatiba, 2008.
88 p.

Dissertação (mestrado) – Programa de Pós-Graduação
Stricto Sensu em Psicologia da Universidade São
Francisco.

Orientação de: Fermino Fernandes Sisto.

1. Avaliação psicológica. 2. Testes psicológicos.
 3. Matrizes progressivas avançadas de Raven.
 4. Raciocínio Inferencial. 5. Inteligência. 6. Fator g.
- I. Sisto, Fermino Fernandes. II. Título.

Ficha catalográfica elaborada pelas bibliotecárias do Setor de
Processamento Técnico da Universidade São Francisco.

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

MATRIZES PROGRESSIVAS AVANÇADAS DE RAVEN E TESTE DE
RACIOCÍNIO INFERENCIAL: EVIDÊNCIAS DE VALIDADE

Autor: Ivan Sant'Ana Rabelo

Orientador: Prof. Dr. Fermino Fernandes Sisto

Este exemplar corresponde à redação final da dissertação de mestrado que será defendida por Ivan Sant'Ana Rabelo e avaliada pela comissão examinadora.

Data: ____ / ____ / _____

COMISSÃO EXAMINADORA

Prof. Dr. Fermino Fernandes Sisto

Profa. Dra. Ana Paula Porto Noronha

Profa. Dra. Selma de Cássia Martinelli

ITATIBA — SP

2008

DEDICATÓRIA

Dedico este meu trabalho a Javé, por me comprovar a cada dia que “tudo concorre para o bem daqueles que amam a Deus” (Rm 8,28), este Deus grandioso que me permitiu concluir este trabalho e me dá imensos motivos a cada dia para confiar em Seu imensurável Amor.

AGRADECIMENTO

Primeiramente, agradeço a meus amados pais e anjos, Vail Rabelo e Iraide Sant'Ana, que me ensinaram a cada dia o maior de todos os aprendizados: o de ter caráter e buscar a sabedoria nas coisas do bem. A minha querida irmã, Renata; meu cunhado, Esequiel; e meus sobrinhos, Thais e Guilherme, minhas razões de existir — minha grande família: A GRANDE FAMÍLIA.

Ao Prof. Dr. e MEU orientador, Fermino Fernandes Sisto, que sempre transmitiu conhecimento, segurança e confiança para buscar aquilo que a mim mesmo parecia quase interminável. Deixo aqui registrado o imenso privilégio que foi ter sido orientado por um profissional como você.

Aos queridos professores do Programa de Pós-Graduação *Stricto Sensu* em Psicologia, que, durante o curso, puderam contribuir imensamente para minha formação. Para não deixar de privilegiar nenhum, amplio meu agradecimento e carinho a todos.

A meus amigos e amigas, que sentiram minha ausência ao longo desta caminhada, torceram por meu sucesso e, muitas vezes, se propuseram a me auxiliar, seja com um palpite em minha dissertação, seja com a colaboração no *input* dos dados, seja com a palavra amiga e confortante, seja com a curtição, o chope nosso de cada dia. Não pensem que minhas noites pouco dormidas durante este período foram apenas de estudo, não, mas também bem empregadas no pequenino, mas maravilhoso tempo ao lado destes amigos especiais. A cada um, meu carinho e agradecimento único.

Não posso deixar de agradecer à instituição que permitiu a realização deste trabalho, da qual hoje sou funcionário: a Casa do Psicólogo. Agradeço aos colegas de trabalho, que, em meio à loucura do dia-a-dia, contribuíram, e muito, para a realização

desta minha pesquisa. Obrigado, Irene Leme, Milena Rossetti, Silvia Pacanaro e demais amigos de trabalho!

Agradeço à Universidade Paulistana e à Universidade Paulista (Unip), pelo apoio e incentivo à minha pesquisa. Um agradecimento especial às Profas. Adele de Miguel, Raphaella De Martini e Dra. Cristina Pellini, pelas importantes contribuições. E, mais uma vez, à amiga Milena Rossetti, que não cansou de contribuir para este trabalho.

Um agradecimento mais do que especial ao Sr. Ingo Güntert — não tenho palavras para descrever quanto o admiro, como profissional e, acima de tudo, por seu caráter, um grande e verdadeiro amigo — e também a minha amiga Camila, que acreditou em mim desde o início e me deu um empurrão para esta grande conquista. Obrigado por acreditarem em mim!

A todos aqueles que, de alguma forma, me ajudaram na realização deste projeto, o meu muito obrigado. E um agradecimento peculiar a uma pessoa mais do que especial.

RESUMO

Rabelo, I. S. (2008). *Matrizes Progressivas Avançadas de Raven e Teste de Raciocínio Inferencial: Evidências de validade*. Dissertação de mestrado. Programa de Pós-Graduação *Strictu Sensu*. Universidade São Francisco: Itatiba.

O objetivo do presente estudo foi verificar evidência de validade para as Matrizes Progressivas Avançadas de Raven por meio do Teste de Raciocínio Inferencial (RIn), sabendo-se que ambos os testes se propõem a medir a inteligência por intermédio de um fator geral, denominado por Spearman de fator *g*. Para tal, averiguou-se a relação entre estas medidas em alunos do ensino superior. Participaram da pesquisa 291 alunos entre 17 e 63 anos ($M = 28$ e $DP = 9,42$), cursando ensino superior, oriundos de duas universidades privadas do estado de São Paulo. A média de pontuação obtida no teste Raven foi de 14,94 acertos e desvio-padrão de 6,52, enquanto, para o teste RIn, verificou-se uma média de 23,03 acertos, com desvio-padrão de 5,79. O estudo de correlação entre os resultados dos instrumentos apresentou coeficientes de correlação moderados e positivos entre a pontuação total do teste Raven e as três primeiras séries do teste RIn e apenas uma correlação baixa positiva com a Série D; apresentou também uma correlação alta e positiva entre o total de acertos dos dois testes, o que corresponde ao esperado e evidencia que ambas as medidas foram sensíveis para captar o construto em que os instrumentos se propõem a medir, a inteligência geral. Verificou-se que os resultados do teste RIn não apresentaram diferenças significativas em relação à variável sexo, ocorrendo o contrário com as Matrizes Progressivas Avançadas de Raven, que apresentou médias maiores para o sexo masculino. No que tange aos estudos relacionados com possíveis diferenças entre os participantes em relação à inteligência geral, não se evidenciaram também diferenças por meio da variável idade. Sugere-se a necessidade de que novos estudos sejam feitos para elucidar tais resultados, congruentes com a literatura na área.

Palavras-chave: avaliação psicológica; Matrizes Progressivas Avançadas de Raven; Raciocínio Inferencial (RIn); inteligência; fator *g*.

ABSTRACT

Rabelo, I. S. (2008). *Advanced Progressive Matrices of Raven e Teste de Raciocínio Inferencial: Evidence of validity*. Master's Thesis. Graduate Program in Psychology. University of São Francisco: Itatiba.

The present study aimed to verify evidence of concurrent validity for the Advanced Progressive Matrices of Raven through the Teste de Raciocínio Inferencial (RIn), when it is known what both tests propose to measure the intelligence through a general factor called by Spearman of factor g . The aim was to verify a relationship between those measures in students from universities. The participants were 291 students between 17 and 63 years ($M = 28$ and $DP = 9,42$), originating from two private universities of the state of São Paulo. The average of score obtained in the test Raven was of 14,94 points and standard deviation of 6,52, while for the test RIn there was an average of 23,03 points, with standard deviation 5,79. The study of correlation between the results of the instruments presented coefficients of moderate and positive correlation between the total score of the test Raven and the first three series of test RIn and only a low positive correlation with the Series D; it presented also a high and positive correlation between the total of points of two tests, which corresponds to an expected one and shows up that both measures were sensitive to catch the theory in which the instruments are proposed to measure, the general intelligence. It was found that the test results RIn showed no significant differences in relation to variable sex, the opposite occurs with the Advanced Progressive Matrices of Raven, who presented averages higher for males. As regards, the studies made a list to possible differences between the participants regarding the general intelligence, one did not show differences up also through the variable age. Suggest the necessity of what new studies they are done to elucidate such results, congruent with the literature in the area.

Key words: psychological evaluation; Advanced Progressive Matrices of Raven; Teste de Raciocínio Inferencial (RIn); intelligence, factor g .

SUMÁRIO

DEDICATÓRIA.....	v
AGRADECIMENTO	vi
RESUMO	viii
ABSTRACT	ix
LISTA DE TABELAS	xi
LISTA DE FIGURAS	xii
APRESENTAÇÃO.....	1
AVALIAÇÃO PSICOLÓGICA.....	4
AVALIAÇÃO PSICOLÓGICA NO CONTEXTO BRASILEIRO	9
EVOLUÇÃO DO ESTUDO DA INTELIGÊNCIA.....	13
INSTRUMENTOS DE AVALIAÇÃO DA INTELIGÊNCIA	30
Matrizes Progressivas Avançadas de Raven	36
Teste de Raciocínio Inferencial (RIn)	43
OBJETIVOS.....	49
Objetivo Geral	49
Objetivos Específicos	49
MÉTODO.....	50
RESULTADOS	54
DISCUSSÕES E CONSIDERAÇÕES FINAIS.....	64
REFERÊNCIAS	72
ANEXOS.....	83

LISTA DE TABELAS

Tabela 1. Pontuação dos participantes nos testes Raven e RIn	57
Tabela 2. Correlação de <i>Pearson</i> entre os acertos nas pontuações dos testes	57
Tabela 3. Correlação de <i>Pearson</i> entre os acertos nas pontuações dos testes, por sexo	58
Tabela 4. Correlação parcial entre os acertos nas pontuações dos testes controlando a variável sexo	59
Tabela 5. Teste <i>t Student</i> das pontuações dos testes por meio da variável sexo	59
Tabela 6. Correlação de <i>Pearson</i> entre a idade e os acertos nas pontuações dos testes	63
Tabela 7. Correlação parcial entre os acertos nas pontuações dos testes controlando a variável idade.....	63
Tabela 8. Análise de variância (ANOVA) entre a pontuação dos testes em relação à variável idade.....	64

LISTA DE FIGURAS

Figura 1. Frequência da pontuação total do teste Raven	55
Figura 2. Frequência da pontuação total do teste RIn	56
Figura 3. Média de pontuação dos participantes no teste Raven por idade	61
Figura 4. Média de pontuação dos participantes no teste RIn por idade	62

APRESENTAÇÃO

A área de avaliação psicológica tem-se desenvolvido ao longo da história da psicologia quanto ao dimensionamento de seus estudos e pesquisas, bem como as conclusões científicas aplicadas aos mais diferentes campos de atuação, seja na área da educação, no contexto organizacional, em áreas clínicas, na neuropsicologia, entre outros contextos. A criação, a padronização e a validação de testes são uma necessidade acadêmica e de prática social, e essa carência de instrumentos é uma questão apresentada por Sisto, Codenotti, Costa e Nascimento (1979), ao verificarem a ausência de materiais brasileiros, propiciando a utilização de instrumentos estrangeiros sem validade para um contexto e amostra específicos, dentro da realidade de nosso país. Essa prática de utilização de instrumentos estrangeiros nos trouxe problemas ao longo dos anos, principalmente no que diz respeito à maneira como as avaliações têm sido realizadas pelos profissionais, também enfatizada por estudiosos em anos mais recentes (Noronha, Freitas, Baldo, Barbini & Almeida, 2004).

Para Alves (2006), observa-se, na literatura internacional, uma preocupação na realização de estudos que verifiquem as implicações relativas à revisão dos testes de inteligência, como pode ser observado em artigos que se dedicam a esse tema no periódico *Psychological Assessment* (Nelson, 2000; Silverstein & Nelson, 2000). Em relação ao Brasil, durante muitos anos, os psicólogos tiveram de trabalhar com testes sem normas ou estudos nacionais, sendo verificado que muitos não atendem às exigências mínimas no que diz respeito às informações necessárias para escolha e utilização do instrumental mais adequado para as diferentes situações (Alchieri & Cruz, 2003; Cruz, Alchieri & Sarda Jr., 2002; Noronha, 1999; Noronha & Vendramini, 2003).

Como uma tentativa de contribuir de forma relevante para a área de avaliação psicológica, principalmente para os aspectos relacionados com o estudo da inteligência, este projeto de pesquisa se propõe a buscar evidência de validade para o teste Matrizes Progressivas Avançadas de Raven (Raven, Raven & Court, 1998) com o Teste de Raciocínio Inferencial (RIn) (Sisto, 2006). Objetiva-se também averiguar o construto inteligência de modo pragmático, por meio de instrumentos psicológicos, e verificar a relação da inteligência em função do gênero e da faixa etária dos sujeitos.

No que diz respeito aos testes utilizados nesta pesquisa, as Matrizes Progressivas Avançadas de Raven são utilizadas internacionalmente como medida de avaliação da inteligência relacionada com a seleção de pessoal de nível superior e como taxa de previsão dos progressos educacional e profissional, entre outros. Verificam-se estudos de validade e precisão para diversos países, tais como Alemanha, Austrália, Estados Unidos da América, Inglaterra e Nova Zelândia (Raven & cols., 1998); porém, verifica-se a ausência desses estudos para amostras brasileiras, constando na listagem dos testes avaliados pelo Sistema de Avaliação de Testes Psicológicos (Satepsi) (Conselho Federal de Psicologia, 2007) com parecer desfavorável.

O teste de Raciocínio Inferencial (RIn) (Sisto, 2006) avalia a inteligência, especificamente o fator g, por meio do raciocínio inferencial. No que se refere aos parâmetros psicométricos, apresenta, no manual, estudos de validade e precisão para amostras provenientes de dois estados brasileiros (São Paulo e Minas Gerais). Quanto ao caráter das evidências de validade do teste, verificam-se evidências de validade relativas à estrutura interna e de critério.

O primeiro capítulo deste trabalho visa a apresentar uma breve contextualização da área de avaliação psicológica e uma explicação das facetas dessa área, tais como as de investigação psicológica, testes psicológicos, instrumentos psicométricos e outras.

Em seguida, destacam-se considerações em relação à avaliação psicológica no contexto brasileiro, a importância de instrumentos validados e padronizados para essa população. Também se reforçam as exigências do Conselho Federal de Psicologia quanto à pesquisa e à utilização de instrumentos no Brasil.

Posteriormente, trata-se especificamente da definição do construto inteligência, em um histórico do estudo da inteligência ao longo do desenvolvimento da psicologia. A seguir, são abordados os instrumentos utilizados para a avaliação da inteligência, enfatizando-se os dois instrumentos empregados nesta pesquisa e seus aspectos psicométricos.

Depois da fundamentação teórica, são apresentados os objetivos do trabalho, aos que se segue o método, considerando-se os participantes, a descrição dos instrumentos, bem como o procedimento para a coleta de dados. A próxima seção versa o plano de análise de dados, os resultados e as considerações finais. Por fim, são apresentadas as referências e os anexos.

AVALIAÇÃO PSICOLÓGICA

A avaliação de fenômenos e processos psicológicos e a investigação de características dos comportamentos humanos estão respaldadas nos pressupostos derivados da medida, que se apresenta como um correlato de identificação e caracterização de determinado fenômeno psicológico. O instrumento, ou teste, é uma forma objetivamente definida para a mensuração de certas características do comportamento humano, o que exige um processo cuidadoso de elaboração. Em pesquisa, a maneira como essa atividade é conduzida está intimamente vinculada à especificidade dos instrumentos utilizados (Muñiz, 1996).

Em relação à investigação psicológica, assume-se que os testes psicológicos possam representar, pela medida, determinada ação que equivale a um comportamento e, assim, indiretamente, mensurar esse aspecto do comportamento. Na avaliação psicológica, os testes são instrumentos que informam sobre a organização dos comportamentos desencadeados pelos estímulos apresentados nos testes, sejam por figuras, sons, formas espaciais, entre outros, ou por estímulos não-estruturados. Dessa forma, avaliam e quantificam comportamentos observáveis, por meio de técnicas e metodologias específicas, formuladas cientificamente em construtos teóricos, que embasam a análise dos resultados (Aiken, 1996).

Pasquali (1996), entre outros autores, expõe que os testes são instrumentos de medida validados e padronizados, resultado da medida de fenômenos e processos psicológicos, ou seja, são operações empíricas pelas quais a psicologia estuda diferentes manifestações da conduta humana. A construção de instrumentos psicológicos tem por suposição a de que a melhor maneira de observar um fenômeno psicológico é por meio da medida.

Vale destacar, de forma que se busque um conceito sobre teste psicológico, que a palavra “teste” tem origem inglesa e significa “prova”. É um termo que deriva do latim *testis* e, internacionalmente, refere-se a uma modalidade de medição, conhecida hoje nos diferentes meios técnicos e científicos (Tyler, 1973). No que tange à sua finalidade, o teste visa a medir as diferenças existentes quanto a determinada característica, seja entre diversos sujeitos, ou a respeito do comportamento de um indivíduo em ocasiões diferentes, relacionadas com o próprio sujeito consigo mesmo, ou na relação dele com o meio.

Para Tyler (1973), o instrumento psicométrico mais típico é o teste. Contudo, ele destaca que não é a única ferramenta. Trata-se de uma situação estimuladora que é padronizada por meio dos itens do teste e do ambiente de aplicação, relacionado com o que o indivíduo responde. Entende-se por situação experimental tudo que compõe o teste e sua aplicação, considerando-se o material empregado, as instruções, o local da aplicação, os itens a serem respondidos, entre outros aspectos. Essas condições precisam ser padronizadas para que se evitem variações nas condições da administração e para que se possa mensurar aquilo a que se propõe, sem viés.

Sabendo-se que os testes psicológicos visam a fornecer dados confiáveis e informações sobre as características psicológicas dos indivíduos, destaca-se sua funcionalidade de sentido de servir como subsídio para a tomada de decisões e para a intervenção de vários profissionais (Cunha, 1974). Comumente, os testes são empregados como instrumentos integrantes de um processo mais amplo — o qual denominamos de avaliação psicológica — e devem ser utilizados em conjunto com outras técnicas, entre elas a entrevista psicológica e a observação do comportamento.

Vale destacar que, para que os testes psicológicos alcancem seu intuito, é necessária a existência de uma uniformidade em sua metodologia, desde as questões relacionadas com a padronização até as relacionadas com o desenvolvimento de

critérios para a interpretação dos resultados obtidos, o qual denominamos de normatização. Verificam-se três grandes grupos de testes psicológicos, a saber: os testes de inteligência, os testes psicomotores e os testes de personalidade. Esses instrumentos são frequentemente utilizados no processo psicodiagnóstico e são selecionados de acordo com a necessidade do caso investigado e das hipóteses levantadas nessa avaliação (Pasquali, 1997).

Os instrumentos relacionados com a avaliação da inteligência visam a auxiliar na análise do funcionamento intelectual de crianças, adolescentes, adultos e idosos e apresentam-se como recursos diagnósticos para a identificação e a medida de diferentes habilidades cognitivas. Os testes psicomotores avaliam a maturidade perceptivo-motora infantil, apresentando indicadores de distúrbios neurológicos. Os instrumentos de avaliação da personalidade investigam os aspectos afetivo-emocionais da personalidade e fornecem dados qualitativos do modo de funcionamento e da estrutura emocional do indivíduo. Assim, os testes são fontes de informação, que podem ser utilizadas para comparações entre pessoas ou para comparar a mesma pessoa sob diversas circunstâncias e entre grupos. Estão presentes na coleta de dados de pesquisas dos mais diversos âmbitos, tais como diferenças individuais, organização e traços psicológicos e aspectos relacionados com diferenças comportamentais, culturais, entre outros (Anastasi & Urbina, 2000).

Segundo Alchieri e Cruz (2003), a investigação do comportamento por meio dos testes se dá com a aplicação de diferentes técnicas, relacionadas com o tipo de avaliação que se pretende investigar. Cada teste apresenta sua finalidade específica, que visa a controlar e a excluir quaisquer variáveis que venham a interferir na investigação dos resultados, para que os dados encontrados sejam precisos e representem as reais condições do avaliado. A psicometria é uma das especialidades psicológicas que busca

aperfeiçoar as qualidades dos testes, e sua origem está relacionada com disciplinas tais como a estatística e as psicologias experimental, metodológica e computacional.

Para os autores, quatro condições são necessárias para que um instrumento, ou teste psicológico, tenha qualidade e possibilidade de uso seguro: elaboração e análise de itens e estudos de validade, precisão e padronização. No processo denominado de elaboração e análise dos itens, são elaboradas e avaliadas as questões individuais (itens) do teste, de forma que se busque uma representação de todas as possíveis variações que o atributo medido possa admitir, correndo-se o risco de não considerar a variedade das respostas e, com isso, deixar de medir características do comportamento que se pretende investigar.

Posteriormente, verifica-se a validade do teste, de modo que se garanta que realmente possa medir aquilo a que se propõe, expressa na equivalência de sua medida com outras formas similares de avaliação e na possibilidade de essa mensuração ter um caráter preditivo na avaliação do comportamento. Objetiva-se comparar o instrumento com outras formas semelhantes de ver aquele comportamento e as respostas de grupos de respondentes com outros testes relacionados. Para Cronbach (1996), a validação é um processo de coleta de informações que dêem suporte para as inferências as quais serão feitas com base nas pontuações do teste. Segundo Anastasi e Urbina (2000), a validade informa-nos sobre as possibilidades de interpretação dos dados encontrados em um teste e quanto o suporte teórico está representando seus construtos. Para que isso seja firmado, é necessário o acúmulo de evidências que se complementem, no sentido de fornecer uma compreensão mais ampla dos escores dos testes.

Em uma etapa posterior, deve-se identificar se a medida efetuada pelo instrumento pode ser considerada como consistente e não-passível de sofrer modificações alheias à manifestação do comportamento. A precisão, ou fidedignidade,

objetiva verificar a consistência das respostas dadas pelos respondentes. É uma maneira de calibrar o instrumento, ajustando-o especificamente à forma da avaliação. Com o auxílio da estatística e de procedimentos metodológicos, procura-se obter uma forma de estabilidade das respostas (Muñiz, 1996; Pasquali, 2001).

Em relação à padronização, Alchieri e Cruz (2003) destacam que o objetivo desse procedimento é o estabelecimento de normas para a utilização deste. Essas normas se referem à apresentação do instrumento, tais como o número de páginas, o tipo de papel, a disposição dos itens, a forma como deve ser aplicado, a idade em que é indicado seu uso, a escolaridade específica, a indicação de ser aplicado em grupo ou individualmente e outras normas.

Ainda em relação ao grupo de normas, estas se referem à classificação dos resultados obtidos para os sujeitos analisados, verificando percentil, escores padronizados ou classificação dos comportamentos, dentro das especificações distintas e indicadas para tal classificação (idade, escolaridade, gênero, entre outras). Destaca-se, ainda, que a padronização se apresenta como um caráter temporário, devendo ser atualizada em um período de tempo, com o objetivo de condizer com a realidade em que esse instrumento está inserido (Anastasi & Urbina, 2000).

Para que os testes de inteligência sejam utilizados de forma correta, bem como qualquer outro teste psicológico, faz-se necessário o estabelecimento de normas apropriadas para a população em que os testes serão utilizados. Anastasi e Urbina (1977) enfatizam que as normas de um teste são relacionadas apenas com a população normativa da qual foram derivadas, não podendo ser tratadas como absolutas, universais ou permanentes. A representatividade da amostra deve considerar características relevantes sobre a variável que está sendo avaliada. Assim, a avaliação de determinada característica psicológica não deve empregar tabelas normativas provenientes de outros

países, em função das diferenças sociais e culturais existentes. Em muitos casos, as normas devem ser específicas para cada subgrupo da população, pois pode haver diferenças consideráveis entre as pessoas nos resultados do teste. Os subgrupos podem estar relacionados com a região geográfica, o gênero, a escolaridade, a idade, entre outros aspectos (Alves, 1998).

Portanto, tendo em vista a necessidade de que as avaliações sejam mais confiáveis e precisas em todos os contextos de atuação, a realização de estudos sobre instrumentos psicológicos é indispensável para o desenvolvimento da área de avaliação psicológica.

AVALIAÇÃO PSICOLÓGICA NO CONTEXTO BRASILEIRO

Os testes psicológicos têm sido entendidos como instrumentos auxiliares na coleta de dados, que, com as demais informações organizadas pelo psicólogo, auxiliam a compreensão do problema estudado, de forma que se facilite a tomada de decisões. No Brasil, por muito tempo, o teste psicológico foi rejeitado na prática do profissional, uma vez que não oferecia características psicométricas necessárias que garantissem sua confiabilidade (Noronha & Vendramini, 2003).

Alchieri e Cruz (2003) relatam que a área de avaliação psicológica, no que tange à discussão quanto aos instrumentos psicológicos, se tem centrado na elaboração de indicadores, critérios e testes cada vez mais específicos e dirigidos às necessidades específicas, amparando-se na qualidade e na elaboração de instrumentos nacionais de avaliação psicológica que, por muitos anos, foram criticados em função do inexpressivo número de estudos científicos e do uso indiscriminado de muitos testes sem a realização de um único estudo que comprovasse suas características psicométricas.

Novos campos de atuação, atualização do conhecimento existente sobre avaliação psicológica, tendo em vista as diferentes necessidades sociais e científicas brasileiras, têm ocorrido a partir da validação de técnicas psicológicas. Intensificação da pesquisa sobre a atualização na área e ampliação da literatura brasileira sobre medidas psicológicas têm possibilitado um aperfeiçoamento qualitativo na intervenção profissional dos psicólogos. Destaca-se que há muito a ser feito para o desenvolvimento das medidas psicológicas no Brasil; entretanto, os esforços dos pesquisadores e especialistas em avaliação psicológica têm sido dirigidos para o aumento do intercâmbio de experiências e pela organização de trabalhos conjuntos. Observa-se uma necessidade de ajustar a conduta profissional por meio da organização e da socialização da produção de conhecimento sobre os processos de avaliação e medida de fenômenos psicológicos (Cruz, 2002b).

Noronha e Vendramini (2003) afirmam que os testes psicológicos são instrumentos que auxiliam na prática do psicólogo e podem fornecer importantes informações para a elaboração de um diagnóstico, quando do processo de avaliação. Para que os testes sejam úteis e eficientes, eles devem passar por estudos que comprovem suas qualidades psicométricas e garantir reconhecimento e credibilidade por parte da comunidade científica e de leigos.

Segundo o Conselho Federal de Psicologia (CFP, 2003), os testes são instrumentos exclusivos do psicólogo, e a psicologia 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. Em 2003, o CFP publicou uma resolução na qual regulamenta a elaboração, a comercialização e o uso dos instrumentos psicológicos.

De acordo com a Resolução n.º 02/2003, os testes devem ter requisitos mínimos para que sejam reconhecidos como tais, e será considerado falta ética o uso de instrumentos que se encontrem fora dos padrões exigidos pelo CFP. Destaque deve ser dado à determinação de que o manual do teste oriente o profissional na confecção de documentos. Acredita-se que tal obrigatoriedade estimule a pesquisa em torno dos instrumentos e desestimule o uso de materiais precários, com qualidade duvidosa, que porventura possam não garantir bons serviços para a sociedade em geral.

Entre os requisitos exigidos para que o instrumento possa ser utilizado pelos psicólogos, segundo o Anexo 1 da Resolução CFP n.º 002/2003, que versa sobre os critérios de avaliação da qualidade de testes psicológicos, exige-se um conjunto de propriedades básicas que os instrumentos psicológicos devem possuir, de acordo com os parâmetros internacionalmente definidos, para que sejam reconhecidos pelas comunidades científica e profissional. O formulário que o autor/pesquisador do instrumento deverá enviar para análise do conselho está dividido em três partes: a) descrição geral do teste, incluindo descrições do nome do teste, autor e editora (originais e no Brasil); classificação ampla do construto; detalhamento do construto/variável ou variáveis que o teste pretende avaliar; e outras informações; b) requisitos técnicos, nos quais devem contar evidências de validade, precisão/fidedignidade, sistema de correção e interpretação dos resultados (normatização), entre outros dados; e, por fim, c) consideração e análise dos requisitos mínimos, formada por uma tabela (formulário), em que estão sistematizados os requisitos mínimos em relação aos artigos 4º e 5º da presente resolução, para facilitar a elaboração do parecer. Para julgar o atendimento dos requisitos mínimos no que se refere à precisão e à validade, devem-se considerar os estudos brasileiros realizados a

partir da década de 1980. No que se refere à normatização, devem-se considerar os estudos brasileiros feitos a partir da década de 1990.

Em suma, nas primeiras duas seções, é analisada uma série de propriedades dos instrumentos, enquanto, na terceira e última parte, são apresentados os indicadores mínimos (forma do manual, precisão, validade e normatização) que deverão ser considerados na elaboração do parecer final, informando-se se o instrumento atende ou não aos requisitos mínimos. No caso de o instrumento não atender às condições mínimas, em seu parecer final serão explicitadas quais as condições que não foram atendidas e apresentadas sugestões, com vista à sua melhoria. Caso seja observada alguma limitação no instrumento, mesmo que ele atenda aos requisitos mínimos, serão indicadas também sugestões de melhoria. Tais sugestões serão encaminhadas ao responsável técnico pelo teste.

A elaboração e a produção de materiais científicos têm, como uma de suas premissas, levar o conhecimento às pessoas, de forma que possibilite o exercício da investigação e a interação de idéias e soluções para os problemas humanos. A avaliação da pesquisa científica é baseada no avanço da produção de conhecimento, propiciado pelo fruto do trabalho dos pesquisadores, que, com seus precursores, constroem e reconstroem esse conhecimento, submetendo-o à crítica (Cruz, 2002b).

O autor complementa que toda produção de conhecimento só tem função de existir se está à disposição da comunidade. A publicação de resultados de pesquisas tem, portanto, entre seus objetivos estimular a reflexão e a crítica, divulgar as descobertas científicas e alcançar o reconhecimento dentro da comunidade científica e fora dela sobre a relevância dos produtos da ciência. Nesse sentido, procura-se incentivar o ensino da avaliação psicológica e de disciplinas correlatas na graduação e na pós-graduação, tendo em vista a dificuldade de alunos e professores em compreender

a necessidade efetiva de medidas de avaliação de fenômenos e processos psicológicos, no que se refere à gênese de métodos de diagnósticos e de intervenção.

EVOLUÇÃO DO ESTUDO DA INTELIGÊNCIA

A inteligência tem despertado o interesse de pesquisadores desde o início da psicologia como ciência. No que se refere à avaliação da inteligência, seu começo e desenvolvimento deram-se ao mesmo tempo em que a psicologia se estabelecia como área de conhecimento (Anastasi & Urbina, 2000).

Investigações teóricas e empíricas relacionadas com o conceito de inteligência têm ocupado posição de destaque na história da psicologia, no entanto, repleta de discordâncias sobre a maneira de descrever, medir e investigar os fenômenos de interesse, principalmente no que se refere à utilização dos resultados provenientes dos testes que se propõem a medir o nível de inteligência dos indivíduos. De acordo com vários autores, muitas dessas discordâncias provêm da falta de consenso sobre a definição e a conceituação de inteligência (Almeida, 1996; Howard, 1993; Sternberg & Kaufman, 1998).

Segundo Sternberg (1982), em 1921, quando os editores da *Revista de Psicologia Educacional* perguntaram a 14 psicólogos famosos o que é inteligência, as respostas variavam, mas, geralmente, abrangiam dois temas, a saber: a inteligência envolve a capacidade para aprender a partir da experiência e a capacidade para adaptar-se ao meio ambiente. Em 1986, foi realizada a mesma pergunta a 24 psicólogos cognitivos, pesquisadores de inteligência. Eles também enfatizaram, no conceito de inteligência, a importância da aprendizagem por meio da experiência e adaptação ao

ambiente e ampliaram a definição para salientar a importância da metacognição, ou seja, a compreensão e o controle, pelas pessoas, de seus próprios processos de pensamento (Sternberg & Detterman, 1986).

Historicamente, Francis Galton (1822-1911) acreditava que a inteligência era uma função das capacidades psicofísicas e, por muitos anos, manteve um laboratório muito bem equipado, onde se media uma ampla variedade de habilidades psicofísicas e de sensibilidades, tais como discriminação de peso, sensibilidade à altura do som e vários outros testes de força física. Contudo, nenhuma associação pôde ser detectada, e a abordagem psicofísica para avaliar a inteligência logo enfraqueceu, caindo no esquecimento, embora reaparecesse muitos anos mais tarde (Tosi, 2007).

No que se refere às concepções teóricas de inteligência, não se pode deixar de fazer referência a Charles Spearman, que, em 1904, propôs a Teoria Bifatorial, ou Teoria Eclética dos Fatores, elaborada com base na análise fatorial, para tentar explicar um fator geral da inteligência. Sua teoria postula que toda atividade intelectual está expressa em um fator geral *g*, comum a toda atividade mental e possuidor de uma natureza fisiológica, e um fator específico *s*, associado a tarefas individualizadas. Apesar do nome de Teoria Bifatorial, em 1927, Spearman distinguiu três fatores da inteligência: o fator *g*, o fator *e* e um terceiro, chamado fatores de grupo, os quais englobariam aspectos da conduta humana que possuíssem comunalidade.

Em sua teoria da inteligência, um conjunto de leis que governam o conhecimento, chamadas noegenéticas, foi defendido por Spearman (1927). Ao estudar a natureza do fator *g* em termos de funcionamento cognitivo, definiu três processos básicos, a saber: apreensão da própria experiência, educação de relações e educação de correlatos.

A lei da apreensão da própria experiência diz que “uma pessoa tem uma maior ou menor capacidade de observação de sua própria mente” (Spearman, 1927, p. 154), significando uma capacidade ligada à percepção, à rapidez e à acuidade com que as pessoas percebem os estímulos, bem como aos processos de autopercepção da atividade consciente. Por sua vez, a lei de educação de relações afirma que “quando uma pessoa tem em mente duas ou mais idéias quaisquer (pensadas ou percebidas) tem também uma maior ou menor capacidade de incorporar em seu pensamento qualquer classe de relação essencial (derivadas da mesma natureza) entre elas” (Spearman, 1927, p. 155). Em outras palavras, referiu-se à capacidade maior ou menor de estabelecer relações entre duas ou mais idéias, sejam elas originadas da percepção, sejam originadas de representações mnêmicas (por exemplo, pensar ou ver um elefante e uma garça e relacioná-los em uma categoria: animais).

A educação de correlatos diz que “quando uma pessoa tem em sua mente uma idéia qualquer junto com uma relação também tem, ao mesmo tempo, uma maior ou menor capacidade de incorporar em seu pensamento uma idéia correlata” (Spearman, 1927, p. 156). Esta última se refere à capacidade maior ou menor que as pessoas demonstram de criar novas idéias a partir de uma idéia e uma relação (por exemplo: a partir de uma categoria, como a dos animais, evocar idéias particulares, como jacarés, gorilas, quatis, entre outros). Esses processos estariam presentes em todas as atividades mentais, não importando o conteúdo (matemático, verbal, artístico e demais).

Ainda se tratando das leis noenéticas, segundo Sisto (2006), o termo noenético envolve dois conceitos, quais sejam, o noético e o genético. Noético significa todo conhecimento, seja percepção, seja pensamento, baseado no próprio dado ou informação. Genético refere-se a todo conhecimento, na medida em que dá origem a

um conteúdo (fica excluída, assim, a mera reprodução). Para Spearman, essas duas qualidades são concomitantes — todo processo noético é genético, e vice-versa.

Spearman (1927) estudou as leis quantitativas sob três aspectos maiores, que permitem outras subdivisões. Um desses grupos está relacionado com a influência da idade. Controlando a variável idade, os dados encontrados sobre o fator *g* mostraram que o desenvolvimento segue uma linha constante, isto é, a criança tende a manter constante seu posto na classificação, apesar do passar dos anos e de seu conseqüente desenvolvimento. Até aproximadamente os 16 anos, há um aumento do fator *g*, tendo a possibilidade de parar o desenvolvimento um pouco antes. A princípio, infere-se que o crescimento é mais rápido; depois se estabiliza, mais ou menos entre os 15 e os 16 anos, mantendo-se semelhante até o princípio da senilidade, quando começa um rápido declínio.

Em relação à hereditariedade, Spearman ainda destacou a influência da nacionalidade e a influência da educação. Seus resultados não foram conclusivos em razão das dificuldades experimentais em controlar determinadas variáveis, mas apresentou algumas conclusões, entre elas a de que a educação influi mais nas habilidades específicas que diretamente no fator *g*, produzindo maior diversificação quanto às diferenças individuais.

Em relação ao gênero, os diferentes sexos comportam-se de maneira muito semelhante com referência ao fator *g*, sendo as diferenças encontradas consideradas desprezíveis. Em se tratando de habilidades mais específicas, em alguns aspectos houve uma superioridade feminina, enquanto, em outros, uma superioridade masculina.

Para Spearman (1927), o fato de o teste de QI (quociente intelectual) medir o fator *g* define-o como um instrumento para avaliar a habilidade para manipular a informação que subjaz as capacidades, tais como o raciocínio, a aprendizagem e a

resolução de problema. Nesse sentido, dentro dos argumentos que suportam a teoria de um fator geral da inteligência, que governa o nível de inteligência de um indivíduo, há estudos realizados conhecidos como correlações múltiplas positivas, ou seja, há uma alta correlação entre QI e tarefas cognitivas simples. Spearman mostrou, a partir de vários testes, cobrindo diferentes áreas de habilidades cognitivas, que havia uma correlação positiva entre esses testes para um mesmo indivíduo. Em outras palavras, se um indivíduo tinha bom desempenho em um teste de habilidade verbal, também teria um bom desempenho em outra habilidade cognitiva (por exemplo: habilidade matemática). Esse conjunto de correlações positivas entre testes de diferentes habilidades foi denominado de fator de inteligência geral, ou fator *g*.

Segundo Angelini e cols. (1999), Spearman nunca imaginou que o fator *g* englobasse todas as capacidades requeridas para um comportamento inteligente, ainda menos que os termos “fator *g*” ou “inteligência” poderiam ser usados como sinônimos de capacidade. A detecção de qualquer problema requer percepção contextual. Um indivíduo sempre inicia a procura de uma impressão global da informação, um esquema, que lhe possibilite ter em mente várias informações de uma vez. A educação não se limita a materiais pictóricos. De fato, Spearman notou, por exemplo, que nossa recordação imediata, depois da leitura de um trecho, é uma consciência concentrada de sua essência, mais do que cada uma de suas partes. Se o material for especialmente interessante, podemos desenvolver melhor um sentido de suas implicações, acima e além de seu conteúdo real.

Ainda sobre o estudo da inteligência, Thurstone (1938), entre outros autores, assumiu uma posição contrária à de Spearman e, por meio da análise fatorial, propôs sete fatores, entendidos como aptidões mentais diferenciadas. Para ele, a inteligência não se restringia a um único fator, mas a sete fatores, que identificou como capacidades

mentais primárias, a saber: compreensão verbal (medida por meio de testes de vocabulário), fluência verbal (medida pela quantidade de palavras que se iniciam com determinada letra, faladas durante certo tempo), raciocínio indutivo (avaliado por testes de analogias e séries numéricas), visualização espacial (avaliada em testes que exigem rotação mental de figuras de objetos no espaço), número (avaliado por meio de testes de cálculo e problemas matemáticos simples), memória (avaliada por meio de testes de evocação de palavras ou de imagens) e rapidez perceptiva (avaliada por meio do reconhecimento de pequenas diferenças entre as figuras).

Por sua vez, Cattell (1971) assumiu a existência de dois fatores gerais, os quais denominou de Inteligência Fluida (Gf) e Inteligência Cristalizada (Gc). A fluida pode ser entendida como a capacidade para raciocinar em situações novas ou inesperadas, enquanto a cristalizada, como representante da profundidade e quantidade de conhecimentos e experiências adquiridas (Schelini, 2000).

É importante destacar também Horn (1991), que, posteriormente a essas aferições, acrescentou à concepção de Inteligência Fluida e Inteligência Cristalizada outras seis capacidades cognitivas, sendo elas processamento visual, memória a curto prazo, armazenamento, recuperação a longo prazo e velocidade de processamento, rapidez para a decisão correta e processamento auditivo. Por meio de vários estudos de meta-análise, Carroll (1993) diferenciou as capacidades em três diferentes camadas ou estratos, a saber: camadas específicas, camadas amplas ou geral e camada relativa a uma única camada geral. Estudos mais recentes (Flanagan, McGrew & Ortiz, 2000) usam a integração entre os modelos de Cattell e Horn (teoria das Inteligências Fluida e Cristalizada) e o modelo de Carroll (teoria dos três estratos), denominada de Teoria Psicométrica das Habilidades Cognitivas de Cattell-Horn-Carroll (CHC).

O conceito de inteligência também se mistura a outros conceitos na literatura psicológica, particularmente ao raciocínio. Os dois conceitos têm sido concebidos como sinônimos em várias teorias psicológicas. Sternberg (1982) inicia um de seus trabalhos afirmando que esses dois conceitos são tão próximos que se torna difícil diferenciá-los, sendo esta uma posição semelhante à apresentada por Spearman (1927).

Segundo Sisto (1975), historicamente o raciocínio é um dos constructos mais estudados na psicologia e, até hoje, um dos mais importantes. É um processo comumente usado para avaliar uma informação com vista a inferir conclusões, conclusões essas que não pertencem explicitamente aos dados. Em outros termos, o raciocínio é um processo mental que resulta em uma conclusão com base em percepções, pensamentos ou afirmações. Os processos envolvidos nos diferentes tipos de raciocínio geralmente fazem parte do constructo inteligência.

Para Almeida (1988), o raciocínio consiste na aptidão intelectual por excelência, que integra os processos cognitivos mais elaborados do ponto de vista intelectual; porém, é apenas um dos componentes da inteligência. O autor completa essa definição tratando a inteligência como sendo todos os processos cognitivos, desde a recepção da informação à elaboração da resposta, passando pela codificação, pela memorização, pela aprendizagem, pela evocação e pelo relacionamento da informação. Desse modo, o raciocínio se situaria, preferencialmente, nos componentes de tratamento de informação.

Reportando-se às teorias cognitivistas da inteligência ou teorias do processamento da informação, modelos consolidados envolvem a análise dos tempos despendidos pelos sujeitos na realização das tarefas. Métodos componenciais foram usados para decompor as analogias e os silogismos, entre outras tarefas clássicas na avaliação da inteligência, nas várias funções cognitivas usadas pelos sujeitos em sua resolução (Sternberg, 1977; Hunt, 1978). Por meio dessa segmentação, procurou-se

estimar quais os componentes relevantes na resolução das tarefas e seu peso efetivo, avaliando-o por meio do tempo e do erro que traziam ao desempenho. A manipulação do formato das tarefas e do volume de informação disponível e a processar permitia, por meio de um método aditivo-subtrativo, inferir tais componentes e sua respectiva importância (Sternberg, 1977).

Segundo Ribeiro e Almeida (2005), adjacentes a esse modelo encontram-se estudos com os tempos de reação em tarefas cognitivas relativamente simples, como, por exemplo, apertar um botão em face de um som grave ou informar se duas letras são idênticas ou diferentes do ponto de vista semântico (vogal e consoante) ou físico (minúsculas e maiúsculas). Buscando-se os correlatos da mente ou da cognição, pretende-se verificar em que medida sujeitos mais e menos capazes do ponto de vista intelectual (por norma, recorrendo-se a testes de QI ou testes de fator g) se diferenciam nos tempos gastos e, se possível, interpretar tais diferenças em termos da capacidade e das estratégias usadas. A velocidade de processamento da informação aparece como indicador da capacidade intelectual dos indivíduos desde os primórdios da psicologia científica, sobretudo com os trabalhos laboratoriais em torno das tarefas simples e complexas de reação (Hick, 1952; Hyman, 1953).

A simples presença ou ausência de um estímulo a assinalar com uma única resposta tipifica as “tarefas simples”, e a existência de graus, formatos ou tipos diversos de estímulos exigindo respostas diferenciadas por parte do sujeito torna as “tarefas complexas” (Jensen & Munro, 1979). Essa complexidade, traduzida na inspeção do estímulo, em sua codificação, na diferenciação e na associação a determinado padrão de resposta, requererá mais tempo para seu processamento e resposta.

Estudos entre tempo de reação e níveis de inteligência têm provocado muita controvérsia. É difícil explicar tais correlações e suas implicações para a definição de

inteligência. Uma leitura estritamente neurológica parece pouco adequada, dada a presença de variáveis associadas ao treino, às estratégias e à metacognição dos sujeitos (Boris, MacLeod & Forrin, 1993; Lindley, Wilson, Smith & Bathurst, 1995).

Não obstante, tem-se verificado que a relação entre os níveis de correlação obtidos e a complexidade das tarefas não é linear. Se as correlações aumentam à medida que elevamos a complexidade das tarefas, é certo também que, a partir de determinado nível de complexidade, os coeficientes de correlação deixam de aumentar (Jensen, 1993). Esse dado parece sugerir que as correlações são mais elevadas quando a complexidade cognitiva inerente às tarefas se aproxima do limiar de capacidade de memória de trabalho dos sujeitos, isto é, quando ficam justamente aquém de provocar uma sobrecarga da informação a processar na memória de trabalho.

Os conceitos de dificuldade e complexidade aparecem interligados, podendo ser considerados como os dois lados de uma moeda, que se inter-influenciam mutuamente. A dificuldade dos itens corresponderá à probabilidade de serem respondidos corretamente. Um item com elevada porcentagem de respostas corretas é considerado um item fácil. A complexidade torna-se mais difícil de operacionalizar. A maioria dos estudos nessa área aborda a complexidade da tarefa pelo somatório de operações cognitivas fundamentais a efetuar (Sternberg, 1977). Se, por critério de complexidade da tarefa, se adotar a quantidade de componentes cognitivos elementares envolvidos, torna-se difícil uma descrição exaustiva dos componentes cognitivos que entram na realização de qualquer teste de inteligência.

A capacidade de eduzir relações está relacionada com a capacidade de perceber variáveis que não são óbvias em si mesmas, para que as relações possam ser reconhecidas. Isso é o que Spearman (1927) quis dizer quando escreveu que a percepção de uma variável tendia a evocar instantaneamente o conhecimento de uma relação, e

vice-versa. Sem uma percepção do todo, não se pode ver nada, e muito menos analisar. As associações simbólicas que sustentam a percepção são mais importantes que os dados sensoriais brutos. Não é, prioritariamente, o discernimento das semelhanças e diferenças o foco, mas a capacidade de eduzir constructos que tornem possível discerni-las (Angelini & cols., 1999).

Os mesmos autores complementam que, para analisar, devemos ser capazes de perceber mais do que o todo e devemos ter hipóteses, possivelmente não-verbalizadas, sobre o que é importante prestar atenção no todo. Análise não significa cortar em pedaços ao acaso. Significa investigar relações potenciais sugeridas pela nossa compreensão do todo. É necessário ter também representações simbólicas de linhas, círculos, quadrados e triângulos. A capacidade para perceber todas essas coisas é baseada na experiência e na aprendizagem cultural.

Para Contrera e Hattori (2003), dependendo do foco de estudo, a imagem é um termo comumente utilizado para designar representações gráficas ou verbais de algo que existe ou poderia existir. Em outras palavras, é a representação de algo por semelhança. Dentre as vantagens do uso de imagens, destaca-se o fato de ser universal, pois vence a barreira da linguagem, podendo, por meio de um entendimento imediato, ser compreendido por pessoas de língua e cultura diversas. Outra vantagem, comumente apontada, é a de que a imagem permite uma leitura em menos tempo que o requerido pelo texto escrito.

Em um artigo publicado por Alves (1998), é tratada a necessidade de considerar e de controlar algumas variáveis que têm demonstrado uma influência significativa nos resultados dos testes de inteligência. Dentre elas, podem-se destacar o nível socioeconômico, a idade, o sexo e o grau de escolaridade. Enquanto a variável nível socioeconômico se mostrou relevante no Teste de Goodenough, nas Matrizes

Progressivas Coloridas de Raven e na Escala de Maturidade Mental Colúmbia, a variável sexo apresentou diferenças significantes entre homens e mulheres no Teste de Goodenough, no Teste R-1 e no D-70, mas não nos testes Raven e Colúmbia. Em relação à idade cronológica, sabe-se que esta sempre deve ser considerada na avaliação da inteligência de crianças, pois os resultados nos testes devem aumentar à medida que a criança cresce. Ocorre o mesmo em relação à escolaridade.

Em uma pesquisa com pré-escolares utilizando os testes Goodenough e Goodenough-Harris, foram encontradas diferenças significativas quanto ao sexo, no qual as meninas apresentaram os resultados mais altos nessas duas escalas (Alves, 1979; 1981).

Por outro lado, Alves e Duarte (1993), na pesquisa de padronização da Escala de Maturidade Colúmbia, não verificaram diferenças por gênero nos respondentes avaliados. Inferem que, por se tratar de um teste não-verbal que avalia a capacidade geral de raciocínio, é possível que o resultado no teste não seja influenciado pelo sexo do sujeito. Isso corrobora com a pesquisa de padronização do Teste das Matrizes Coloridas de Raven, em que não foram constatadas diferenças entre os sexos, com exceção das crianças de 11 anos e 11 anos e meio, que apresentaram variação (Alves, Angelini, Custódio & Duarte, 1988).

Melo (1998) obteve resultados análogos ao comparar meninos e meninas de 10 e de 13 anos na Escala Verbal Brasileira do WISC adaptada por Lemgruber e Paine (1981) e na Escala Verbal do WISC traduzida por Poppovic (Wechsler, 1964). Não foram verificadas diferenças entre os sexos em todos os subtestes verbais, exceto para o subteste Compreensão na escala traduzida para a idade de 10 anos e para a amostra total. Nesse caso, as meninas obtiveram resultados mais altos, o que permite inferir

maior maturidade social das meninas aos 10 anos, mas que não aparece mais aos 13 anos.

Em pesquisas brasileiras realizadas com adultos (Alves, Colosio & Ruivo, 1992), foram constatadas diferenças estatisticamente significativas entre os sexos. Em uma pesquisa que utilizou o teste R-1, aplicado em adultos de três níveis de escolaridade, os sujeitos do sexo masculino obtiveram resultados mais altos. Contudo, sendo o estudo realizado com uma amostra de candidatos em um processo seletivo para uma vaga de emprego, não é possível garantir que seja representativa da população, principalmente porque a maioria dos cargos para o sexo feminino era de atividades de baixo prestígio social, o que, de certa forma, poderia explicar os resultados obtidos pelo sexo feminino. Na pesquisa de padronização do D-70, também foram encontradas diferenças entre os sexos, tanto no grupo de escolares do 2º grau como no de adultos com escolaridade de 2º grau e para adultos do 3º grau, o que levou à construção de tabelas normativas para cada sexo (Alves, 1998).

Em um artigo publicado por Weiss e cols. (2003), as diferenças de gênero no funcionamento neuropsicológico de pacientes com desordens psiquiátricas fornecem resultados contrários, segundo o autor, atribuídos à complexidade de variáveis que influenciam em diferenças sexuais cognitivas. No estudo, tentou-se avaliar a magnitude das diferenças de gênero em funções verbais e espaciais visuais e uma correlação dos resultados com uma auto-avaliação dessas capacidades em homens e mulheres. Foram avaliados 97 estudantes de ensino médio, sendo 51 mulheres e 46 homens, todos examinados em uma bateria neuropsicológica, com foco nas habilidades verbais e espaciais visuais.

Uma diferença significativa entre homens e mulheres foi encontrada em três de quatro medidas verbais e uma das três tarefas espaciais visuais. Os resultados

mostraram que as mulheres apresentaram um nível mais alto que homens nos testes verbais, enquanto os homens tendem a superar as mulheres em tarefas espaciais visuais. Ressalta-se, no entanto, que essas diferenças foram consideradas pequenas. Adicionalmente, foi aplicada uma escala de auto-avaliação, na qual os homens classificaram suas capacidades espaciais significativamente superiores às das mulheres, enquanto as mulheres não classificaram suas capacidades verbais superiores às dos homens. Por fim, concluiu-se que, ao medir os QIs verbal e não-verbal, não se encontrou nenhuma diferença de gênero em relação ao QI não-verbal.

Santos e cols. (2002), em um artigo denominado “Epidemiologia do desenvolvimento cognitivo de escolares em Jequié, Bahia, Brasil”, relataram que a média de escore geral no teste Raven, ajustada por idade, foi maior para os meninos ($21,6 \pm 10,7$ pontos) que para as meninas ($18,0 \pm 8,9$ pontos), diferença essa considerada estatisticamente significativa ($p < 0,001$). Quanto à diferença no desempenho cognitivo entre os sexos, observou-se que a média de escores obtidos nos testes Raven foi melhor entre os meninos, enquanto a diferença de desempenho no rendimento escolar favoreceu as meninas.

Malta, Goulart e Costa (1998) observaram, em estudo realizado com alunos de primeira série em Belo Horizonte, que o desempenho escolar das meninas foi superior ao dos meninos. Nesse sentido, verificou-se maior probabilidade de repetência escolar entre os meninos. Por outro lado, os meninos demonstraram melhor desempenho nos testes psicológicos, o que permite indagar acerca de questões relativas a gênero e a aspectos mensurados pelos testes empregados.

Em um artigo publicado por Colom e Garciápz (2002) sobre diferenças sexuais relacionadas com a Inteligência Fluida (Gf) entre pessoas de escolaridade superior, verificou-se evidência empírica contraditória acerca do problema. Segundo os autores,

alguns pesquisadores afirmam que há uma diferença sexual, confiando na adição das diferenças sexuais padronizadas em testes cognitivos que medem capacidades verbais e espaciais, e de raciocínio. Outros teóricos asseguram que não há uma diferença sexual na inteligência geral. No estudo, o conflito foi verificado por meio de medidas de Inteligência Fluida (Gf). Três medidas de Gf foram usadas no presente estudo, sendo *Cattell's Culture-Fair intelligence test (Scale 3)*, *Inductive Reasoning Tests taken from the Primary Mental Abilities Battery (PMA-R)* e *Advanced Progressive Matrices (APM)*.

Participaram da pesquisa 4.072 sujeitos com escolaridade superior, sendo 1.772 mulheres e 2.300 homens. Os resultados indicaram que não há nenhuma diferença entre sexo significativa no *Cattell's Culture-Fair intelligence test* ($t = 1.285$, ns). As mulheres superaram os homens no *PMA-R* ($t = 5.616$, $p < 0,001$), e os homens superaram as mulheres no *APM* ($t = 3.575$, $p < 0,001$). Esses dados contradizem o que alguns pesquisadores propõem (Allik & cols., 1999; Lynn, 1999): que há vantagem masculina em testes de inteligência, pois, para isso ser considerado “verdadeiro”, os homens deveriam ter superado sistematicamente as mulheres nos três testes investigados, dado esse que não foi verificado na pesquisa. Por isso, considerando que não há nenhuma diferença sistemática que favorece qualquer sexo nas medidas de Gf, conclui-se que a diferença sexual na Inteligência Fluida não existe.

Estudos realizados por Jensen (1998), Colom e cols. (2000) e Colom, Garciálpz e cols. (2001) afirmam que não há nenhuma diferença entre sexo na inteligência geral definida como g. O que é verificado na literatura, segundo os autores, é que as mulheres superam os homens em algumas tarefas verbais, enquanto os homens superam as mulheres em algumas tarefas espaciais (Halpern, 1997; Neisser & cols., 1996). Os testes

de Inteligência Fluida são baseados em figuras abstratas e material não-verbal para eliminar tais diferenças.

Além da variável sexo, este estudo se propõe a investigar a influência da idade nos resultados obtidos nos testes de inteligência. Cunha (1974, p. 101), por meio de um levantamento da literatura afirma que “a curva do crescimento intelectual do homem tem aceleração positiva na infância e na pré-adolescência; esta aceleração decresce, sensivelmente na adolescência, tornando-se nula na idade adulta. Da idade madura para a velhice, a curva apresentaria um declínio significativo”. Na pesquisa de padronização do teste *Wechsler Adult Intelligence Scale* (WAIS), foi observado que o resultado máximo ocorria entre 20 e 30 anos, com um leve declínio até os 60 anos, acentuando-se a partir dessa idade (Anastasi, 1977).

Resultados semelhantes foram apresentados por Raven, Court e Raven (1985), considerando que a capacidade intelectual declina a partir dos 25 anos de idade. Matarazzo (1976) ressalta a função do papel da educação, considerando que, em um estudo transversal, os indivíduos mais velhos viveram em um meio educacional com experiências diferentes dos mais jovens e, em média, tiveram menos anos de escolaridade. Segundo o autor, esses dados levam à impressão de um declínio com a idade, que é apenas aparente, visto que as pesquisas longitudinais demonstraram que a inteligência do adulto continua a aumentar discretamente até os 50 anos e, então, começa a mostrar um pequeno declínio acima dessa idade. A partir desses resultados, Matarazzo lança a problemática da necessidade de atualizações constantes de normas dos testes, visando a corrigir possíveis distorções.

Raven, Raven e Court (1991) relatam que, ao se compararem estudos de épocas diferentes, em relação a pessoas nascidas nos mesmos anos, mas que tivessem diferentes idades nos anos em que os estudos foram realizados, estes apresentavam

resultados semelhantes, independentemente do ano em que foram testados. Isso permite inferir que não há um declínio com a idade, mas que o resultado está relacionado com o ano de nascimento, o que faz com que pessoas nascidas há mais tempo tenham resultados menores no teste de Raven, mostrando a influência das variáveis culturais e nutricionais a fim de aumentar os resultados médios no teste com o passar dos anos.

Para Alves (1998), o declínio no número de pontos em teste de inteligência com o aumento da idade também foi verificado na pesquisa com o teste R-1 (Alves, Colosio & Ruivo, 1992). Contudo, ao se realizar um levantamento da escolaridade em relação à idade, observou-se que há maior frequência de pessoas com escolaridade até o ensino básico entre os 30 e 45 anos, com ensino fundamental entre os 20 e 35 anos e com ensino médio entre os 20 e 30 anos. Isso demonstra que pessoas de mais idade apresentaram um grau de escolaridade menor, o que parece contribuir para tal resultado, corroborando com as afirmações de Matarazzo (1976) e de Raven e cols. (1991).

Nesse sentido, ainda na pesquisa realizada por Alves (1998), observou-se um aumento no número médio de pontos com a idade no teste Goodenough e sua revisão Goodenough-Harris, na Escala de Maturidade Mental Colúmbia e nas Matrizes Progressivas Coloridas de Raven. Dessa forma, os testes devem apresentar normas para cada idade, ou até mesmo para cada meio ano, em função de diferenças encontradas. Contudo, em adultos, a idade é uma variável mais complexa, porque se espera uma estagnação da inteligência em função da idade. Em tempos anteriores, autores discutiram se a inteligência parava de aumentar por volta dos 16 ou 18 anos, limites esses usados na Escala Stanford-Binet de 1937 e 1960 (Anastasi, 1977). Esses questionamentos se davam porque os resultados médios obtidos nos testes paravam de aumentar significativamente nessa idade.

Outro aspecto a ser destacado está relacionado à influência ambiental na determinação da inteligência. Flynn (2003) refere-se ao ganho intelectual, refletido no aumento das pontuações em testes de inteligência durante o século 20, identificado em vários países, principalmente naqueles com alto desenvolvimento socioeconômico. A taxa de crescimento intelectual gira em torno de três pontos de quociente intelectual (QI) por década, estando o ganho maior relacionado mais com os testes de Inteligência Fluida que com os de Inteligência Cristalizada.

Flynn (2006) propôs explicações para o aumento intelectual, agrupadas em duas amplas abordagens, sendo uma a hipótese nutricional e a outra hipótese da estimulação cognitiva. A primeira está relacionada com a influência dos melhores padrões nutricionais e cuidados médicos da população, enquanto a outra se refere à presença e à disponibilidade de desafios cognitivos apresentados pelos meios de comunicação, tais como televisão, jornais, computador, internet, entre outros, e, acima de tudo, à demanda cognitiva de trabalhos cada vez mais complexos.

Ao lado disso, no início da segunda metade do século passado, observa-se uma crescente migração populacional do campo para a cidade, pois é na cidade que se encontram meios de comunicação, transporte, educação, cuidados médicos e sanitários muito mais sofisticados. O aumento da urbanização e, com ela, a modernização implicam, portanto, intervenções ambientais que melhoram as condições de vida da população (Flynn, 2003). Essas melhorias se refletem nas menores taxas de mortalidade infantil, na maior longevidade da população e no maior controle de doenças endêmicas nas zonas urbanas. Dentre as melhorias psicológicas, a variável especialmente a ser destacada é o aumento intelectual da população ao longo do tempo, chamado de Efeito *Flynn*.

Santos e cols. (2002), observando dados de uma amostra brasileira de escolares entre os 8 e 12 anos, submetidos a uma avaliação de competência cognitiva pelo teste Raven, verificaram uma média de acertos de $27,7 \pm 5,6$ pontos para aqueles que freqüentavam escola particular, enquanto o grupo da escola pública atingiu somente $17,0 \pm 4,4$ pontos (Frey & Pinelli Jr., 1991), achado coerente com resultados aqui obtidos em uma amostra proveniente de escolas públicas do interior da Bahia.

Em um estudo realizado por Candeias e Almeida (2000) na Universidade do Minho, verificou-se que as diferenças individuais permitem considerar diferentes processos cognitivos e formas de inteligência. A justificativa para as diferenças individuais no desempenho intelectual está ligada à maior relação da própria inteligência com o meio cultural, à aprendizagem e aos contextos de vida de cada pessoa. Os resultados mais recentes na investigação no domínio da cognição sugerem ser necessário complementar os testes de QI e de fator *g* com formas diferentes de avaliação da inteligência e das funções cognitivas.

INSTRUMENTOS DE AVALIAÇÃO DA INTELIGÊNCIA

Ao longo da História, verificam-se registros de instrumentos e métodos variados para medir os tamanhos dos objetos, desde as coisas mais simples até o tamanho de galáxias, e todas as coisas presentes nelas. Nada escapa ao olhar científico por meio da medida. Criam-se métodos para mensurar a velocidade da luz e do som, a pressão atmosférica, a quantidade de moléculas presente nas coisas, características genéticas das pessoas, as probabilidades de ganhar em jogos de sorte, entre outras tantas. Registram-se as medidas de muitas coisas, e a criação de métodos de mensuração apropriados para medir a inteligência também se faz presente na História.

Pasquali (1997), reportando-se a Dubois, em 1970, afirma que os testes já eram usados na China três mil anos antes de Cristo. Anastasi (1977), reportando-se a Bowman, diz que os testes chineses foram usados durante cerca de dois mil anos, sem especificar a época, e também comenta rapidamente que foram usados exames na Grécia Antiga e na Europa Medieval. Mas, em todos esses casos, tratava-se de testes de cultura, com baixo teor de Inteligência Fluida, de modo que os testes chineses do século 5 d.C. provavelmente foram os primeiros testes de inteligência.

Desde o século 5, os chineses já aplicavam exames com o objetivo de medir a capacidade intelectual e, assim, dividir a população em castas, com funções e responsabilidades de acordo com as habilidades de cada um. Esses testes consistiam simplesmente em avaliar a capacidade mnemônica, a capacidade para interpretar os textos clássicos e a habilidade para escrever poemas. Esses provavelmente foram os primeiros testes cognitivos de que se tem registro (Pasquali, 2003).

No Ocidente, os trabalhos pioneiros foram realizados pelo médico Esquirol (1838), que associou os diferentes níveis de retardo mental aos diferentes níveis de fluência verbal. A idéia de Esquirol é correta, mas seria importante fazer duas distinções relacionadas com o gênero. Para a maioria dos traços cognitivos, inclusive a linguagem, os níveis de habilidade entre as mulheres se distribuem com uma dispersão claramente menor que a dispersão observada entre os homens. Em outras palavras, para a maioria das habilidades cognitivas e personalógicas, o desvio-padrão para homens é maior que para mulheres. Esse fato é amplamente documentado em cerca de 30 milhões de pessoas examinadas pelo SAT ao longo de várias décadas, além de 12 mil pessoas examinadas pelo DAT. As mulheres apresentam, em média, fluência verbal superior à dos homens. De modo geral, as mulheres saem-se melhor que os homens em praticamente todas as atividades que exigem rapidez para tarefas simples, enquanto os homens sobressaem em

atividades que exigem profundidade para lidar com problemas complexos. Esquirol deu um passo importante no sentido de conceber meios de estimar níveis mentais, pois ele tinha razão sobre haver correlação entre deficiências na linguagem e no atraso mental; porém, trata-se de uma correlação fraca (Anastasi, 1977).

De acordo com a autora, o pai da psicologia experimental foi Francis Galton, e o pai dos testes psicométricos foi Alfred Binet. Galton, a partir de 1884, projetou os primeiros testes que visavam a medir a inteligência; contudo, os resultados não chegaram a ser satisfatórios. Ele acreditava que a inteligência poderia ser o resultado de um conjunto de características simples, como diâmetro da cabeça, velocidade dos reflexos, acuidade visual etc. Entretanto, as correlações encontradas não eram tão boas como ele esperava. A correlação entre QI e tamanho da cabeça, por exemplo, é cerca de 0,4, enquanto a correlação entre QI e velocidade dos reflexos para tarefas elementares é cerca de 0,3; portanto, ambas muito baixas para que possam legitimar sua hipótese.

Os primeiros testes de inteligência propriamente ditos, que deram origem aos modernos testes de QI, foram desenvolvidos por Alfred Binet, no início do século 20, e começaram a ser utilizados em 1904. Em princípio, Binet pretendia criar um instrumento que possibilitasse o diagnóstico objetivo de deficiências mentais, além de medir a gravidade da deficiência. O objetivo educativo de identificar as crianças mal-sucedidas na escola por razões intelectuais leva-o à construção, em 1905, da Escala de Inteligência Binet-Simon, assumida como um marco histórico na avaliação da inteligência (Almeida & Buela-Casal, 1997).

Em sua versão inicial, a Escala de Inteligência Binet-Simon era constituída por 30 itens, distribuídos por ordem crescente de dificuldade e dirigida a crianças dos 3 aos 12 anos (Almeida, 1988; Richardson, 1991). Os itens foram agrupados de acordo com a proporção de acertos das crianças por faixas etárias. A partir desse procedimento, dá-se

início à noção de “idade mental” (desenvolvimento mental) que, associada à noção de “idade cronológica”, permite o cálculo do QI (QI de razão) ou a percepção de um desenvolvimento intelectual normal, superior ou inferior por referência à idade do sujeito.

Segundo Anastasi (1977), reportando-se a T. H. Wolf, Binet não utilizava o termo “idade mental” e preferia usar a expressão “nível mental”. Isso demonstra a refinada percepção de Binet, porque, objetivamente, o termo “nível mental” teria evitado diversas confusões que ocorrem em virtude do uso inadequado da expressão “idade mental”. Pela comparação entre idade mental (medida pelo teste) e idade cronológica (ou biológica), era possível saber se uma criança tinha desenvolvimento inferior ou superior e era possível, ainda, saber quão avançada ou atrasada ela estava em comparação com outras crianças da mesma idade. Binet e seu colega nesse trabalho, Theodor Simon, compreenderam claramente as limitações do método que usavam e não ousaram ir além desse ponto, sem antes desenvolver uma metodologia mais apropriada e ter uma compreensão melhor do assunto. No entanto, em 1906, na Universidade de Stanford, Lewis Madison Terman publicou uma versão aprimorada dos testes de Binet, que prontamente foi reconhecida como a melhor bateria de testes de inteligência da época, e, em 1916, por sugestão de William Stern em 1912, foi introduzido o conceito de “quociente de inteligência”, representando a idade mental multiplicada por 100 e dividida pela idade cronológica. Foi assim que surgiu o termo “QI”.

Outro dado relevante no estudo da medida da inteligência está relacionado com a aprendizagem cultural, havendo discussões entre os psicometristas quanto ao fato de o fator *g* medir algo mais profundo, em razão de que o QI se correlaciona com o desempenho em testes não-verbais, mostrando-se menos carregado culturalmente. Em

paralelo, pesquisas com gêmeos univitelinos criados em famílias e ambientes muito diferentes acusam um componente genético bastante forte (Sisto, 2006).

O autor cita ainda um estudo de Deary, Whalley, Lemmon, Crawford e Starr (2000) com um grupo de pessoas com 80 anos, dos quais obteve registro do QI da infância. O estudo mostrou que as que tiveram maiores pontuações foram as mais bem-sucedidas em termos de saúde, longevidade e no cotidiano da vida. Há, porém, outros pesquisadores que argumentam que o fator g é relevante somente quando habilidades acadêmicas estão envolvidas e não seriam tão úteis a muitas atividades do mundo real. Nesse sentido, tem-se encontrado uma forte relação entre a inteligência e a exposição à instrução formal, como havia mostrado Spearman (Cahan & Cohen, 1989; Green, Hoffman, Morse, Hayes & Morgan, 1964; entre outros).

McGrew e Flanagan (1998) relatam que o modelo hierárquico de Três Estratos de Cattell-Horn-Carroll (CHC) foi utilizado como embasamento teórico para a criação de muitos testes e consiste em uma visão multidimensional, composta de dez fatores ligados à área do funcionamento cognitivo: a Inteligência Fluida (Gf), referente às operações mentais de raciocínio que o indivíduo realiza diante de novas situações e que dependem de conhecimentos adquiridos. Essas operações envolvem relacionar idéias, solucionar problemas, induzir conceitos abstratos, entre outros aspectos. A Inteligência Cristalizada (Gc) refere-se à profundidade dos conhecimentos adquiridos e à aplicação dos conhecimentos aprendidos previamente.

Em continuidade a esse modelo, temos habilidades tais como o Conhecimento Quantitativo (Gq), que se refere ao reconhecimento e a entendimentos da informação real sobre os objetos; Leitura e Escrita (Grw), que diz respeito ao conhecimento adquirido em habilidades básicas exigidas no entendimento de textos e vocábulos escritos; Memória de curto prazo imediata (Gsm), relativa à capacidade de manutenção

de informações na consciência por um espaço de tempo curto para poder readequá-las; Processamento Visual (Gv), que está ligada à capacidade de gerar, perceber, analisar, armazenar, lembrar, manipular, transformar e raciocinar com representações ou padrões visuais; Processamento Auditivo (Ga), que concerne à capacidade ligada à percepção, à análise e à síntese de padrões sonoros, incluindo a percepção de estruturas musicais complexas; Capacidade de Armazenamento e Recuperação de Memória de Longo Prazo (Glr), que é referente à extensão e fluência de informação ou conceitos adquiridos de memória de longo prazo; Velocidade Cognitiva Geral (Gs), que inclui a habilidade de realizar uma tarefa cognitiva de forma rápida e automática, mantendo a atenção focalizada e concentração; Velocidade de Processamento/Rapidez de Decisão (Gt), que reúne a velocidade de reação ligada à capacidade de responder as tarefas cognitivas mais complexas, como, por exemplo, a solução de problemas associados à rapidez em reagir ou tomar decisões (Almeida, 2002).

Essa forma de avaliação da inteligência por fatores tornou-se bastante conhecida na psicologia. Esse modelo possibilitou a construção de diferentes escalas, como a WISC — *Wechsler Intelligence Scale for Children* (1949), a WAIS — *Wechsler Adult Intelligence Scale* (1955), a WPPSI — *Wechsler Preschool and Primary Scale of Intelligence* (1967), a K-ABC — *Kaufman Assessment Battery for Children* (Kaufman & Kaufman, 1983) e a KAIT — *Kaufman Adolescent and Adult Intelligence Test* (Kaufman & Kaufman, 1993).

Para Pasquali (1997), tanto os antigos testes chineses do século 5 d.C. quanto os testes modernos medem o desempenho da pessoa na realização de tarefas específicas. Acredita-se que os escores atribuídos aos desempenhos nesses testes se correlacionam positivamente com o que se denomina como inteligência. Os testes não medem a inteligência propriamente dita, mas medem o desempenho da inteligência enquanto ela

se manifesta na resolução de problemas. Se o teste for estruturado de modo que exija que a inteligência tenha uma participação ampla e profunda no processo de discriminação, a ponto de imprimir de forma balanceada suas virtudes globais no escore final, então, o teste será representativo da capacidade intelectual global. Caso contrário, será baixo o coeficiente de correlação entre os escores obtidos no teste e o verdadeiro nível intelectual das pessoas testadas.

Matrizes Progressivas Avançadas de Raven

A primeira escala desenvolvida por Raven ficou conhecida no Brasil como Matrizes Progressivas de Raven Escala Geral (MPR). Contudo, além dela, o autor desenvolveu mais duas escalas: a Matrizes Progressivas Coloridas (MPC), para ser empregada com crianças pequenas, pessoas idosas e deficientes mentais, e a Matrizes Progressivas Avançadas (MPA), desenvolvida para testar pessoas com capacidade intelectual superior à média, sendo mais usada para pessoas com escolaridade universitária, escala essa da qual se deriva esta pesquisa (Raven & cols., 1998).

Segundo Raven (1962), as Séries A, B, C, D e E das Matrizes Progressivas, Escala Geral foram preparadas para abranger a mais ampla área da capacidade intelectual e ser empregadas com pessoas de qualquer idade, independentemente de sua instrução, nacionalidade ou condições físicas. Dada tão extensa aplicabilidade, esperava-se que apresentasse certas limitações. Previstas para serem empregadas tanto com crianças como com adultos, as duas primeiras séries da escala, bem como os problemas iniciais da terceira e quarta, não ofereciam aos adultos muito mais que uma oportunidade para se exercitarem no método de trabalho. Uma breve série de exercícios, preparada para exemplificar o método, teria sido da mesma utilidade.

A escala visava também a fornecer uma estimativa válida da capacidade de uma pessoa para pensar com clareza quando se lhe permitia trabalhar continuamente em seu próprio ritmo, desde o começo até o fim, sem interrupção. Dado que todas as cinco séries de itens que formam a escala começam com problemas fáceis e terminam com outros mais difíceis, não se esperavam resultados muito válidos se a escala se utilizasse como teste de eficiência intelectual (rapidez), limitando o tempo de execução, o que impediria a muitas pessoas atingir as últimas questões dentro do tempo concedido. Quando o teste foi construído, já se reconheciam essas limitações. Considerou-se possível, partindo das informações obtidas com a escala básica, construir formas derivadas para atender especificamente a requisitos particulares, capazes de indicar, em poucos minutos, se uma pessoa se situa intelectualmente abaixo, dentro ou acima da média. Procurava-se também um teste válido de eficiência intelectual que pudesse ser aplicado a pessoas de inteligência superior à média e que discriminasse facilmente, mesmo entre indivíduos de capacidade superior.

Segundo o autor, as Matrizes avaliam a capacidade de eduzir relações por meio de variáveis nas quais as relações não são óbvias em si. Para que as variáveis sejam reconhecidas, a relação deve ser passível de ser discriminada pelo sujeito; em suma, a pessoa deverá ser capaz de apreender em meio a algo confuso ou extrair uma possibilidade de resolução do problema. Isso é o que Spearman (1927) procurou aferir ao afirmar que a percepção de uma variável tendia a evocar instantaneamente um conhecimento de uma relação, e vice-versa.

Segundo Pasquali, Wechsler e Bensusan (2002), Raven identificou o fator g essencialmente como uma capacidade de fazer comparações, raciocinar por analogia e desenvolver um método lógico de pensar independentemente de informação previamente adquirida. Essa capacidade foi definida por ele como “a apreensão da

própria experiência, a educação de relações e educação de correlatos”, ou seja, fazer observações e extrair princípios. Destaca-se que essa habilidade se desenvolve de forma linear, adjacente à maturação do organismo, assim sendo com a idade cronológica. Esse desenvolvimento segue etapas numa evolução que poderíamos chamar a teoria do desenvolvimento cognitivo de Raven.

Esse comportamento educativo exige um processo mais complexo que simplesmente reprodutivo, relacionando-se com solucionar o problema ou satisfazer-se com uma solução específica coerente com todas as evidências disponíveis (Raven, 1962).

Diante de uma tarefa, deve-se primeiramente identificar o que está sendo observado, passar a analisar o que está ocorrendo e, valendo-se disso, poder identificar o espaço a ser preenchido nas Matrizes. Se não adotar uma postura analítica, o respondente escolherá uma reprodução do todo para preencher o espaço, e não a peça que completa o padrão corretamente. O que está sendo avaliado não é o discernimento entre semelhanças e diferenças, mas, sim, a capacidade de eduzir soluções que tornem possível discerni-las. Portanto, espera-se que o indivíduo tenha, possivelmente de modo não-verbalizado, alguma hipótese quanto àquilo a que deve prestar atenção dentro do todo (Angelini & cols., 1999).

Para melhor captar a originalidade genética desse fator geral e, assim, evitar influências de elementos culturais e de treinamento, Raven optou por itens de caráter gráfico em vez de verbal para seu teste. Ressalta-se que o próprio Spearman considerou o teste de Raven a melhor prova não-verbal para avaliar o fator *g* ou a educação. Opinião essa compartilhada na Inglaterra na época, inclusive por Philip Vernon, da Universidade de Calgary, Alberta (Savage, 1968).

Segundo Paín (1995), as Matrizes de Raven apresentam-se como distintas de outros instrumentos de avaliação do constructo inteligência. Enquanto outros testes se utilizam de provas e jogos buscando abranger uma série de comportamentos muito diferentes a partir de estimulações, situações e instruções variadas, como nos testes Binet, WISC, Colúmbia, Bender, entre outros, as Matrizes de Raven apresentam ao sujeito uma tarefa única organizada sob a forma de uma tabela de dupla entrada, pela qual recebe o nome de matrizes progressivas. Matriz, porque cada dado se relaciona com qualquer dos outros em uma interseção, e progressiva, porque sua formulação deriva da aplicação de critérios cada vez mais complexos. Além disso, como se trata de uma prova de escolha múltipla, a análise não se esgotará na matriz, mas será possível interpretar a escolha de certas alternativas que representam uma abordagem específica da tarefa.

Tal autora complementa que Raven (1962) construiu essa prova para destacar a presença de um fator geral, único e comum, que Spearman (1927) havia conseguido determinar por meio da conversão fatorial das correlações entre provas em saturações. Como Spearman adotou uma atitude realista, atribuiu à presença matemática do fator g uma entidade descrita como energia mental, capaz de promover atos legitimados pelas chamadas leis neogenéticas. Raven construiu os itens da prova baseando-se em tais leis, já que a resolução destes supõe a educação.

Enquanto as Matrizes Progressivas Escala Geral não discriminam significativamente as pessoas de inteligência muito elevada, as Matrizes Progressivas Escala Avançada se propõem a realizar tal distinção. Essa segunda escala foi construída para ser um teste válido de eficiência intelectual que pudesse ser aplicado em pessoas de inteligência superior à média e que discriminasse facilmente, mesmo entre indivíduos de capacidade superior. Sua forma original data de 1943, quando foram usadas nas Juntas

de Seleção de Oficiais do Gabinete de Guerra. Em 1947, preparou-se uma revisão para uso geral como teste não-verbal de inteligência, em que, durante a execução, uma pessoa pudesse demonstrar sua capacidade para estabelecer comparações entre figuras e desenvolver um método lógico de raciocínio. Tem a vantagem, sobre os testes verbais, de que a clareza dos processos do pensamento é avaliada com independência das aquisições culturais (Raven & cols., 1998).

As Matrizes Progressivas Avançadas de Raven constam apenas de duas séries: a Série I, com 12 itens; e a Série II, com 36. Os problemas da Série I vão crescendo rapidamente de dificuldade, dos muito fáceis aos muito difíceis. Cumprem dessa forma dois objetivos: 1) servir como teste de triagem, pois somente os que atingirem certo número de acertos seriam indicados para submeter-se à Série II; 2) oferecer um treinamento, rápido, econômico e eficiente, aos que não estão familiarizados com esse tipo de solução de problemas. Procura-se, assim, anular a vantagem dos que conhecem testes similares, colocando todos eles, dentro do possível, em igualdade de circunstâncias. Na Série II, os 36 problemas estão ordenados por ordem de dificuldade, que aumenta mais paulatinamente, chegando, porém, a atingir níveis muito altos. Permite, assim, discriminar eficientemente entre pessoas dotadas de grande capacidade de raciocínio.

Segundo os autores, enquanto o teste não havia sido aplicado a uma ampla gama de adultos de inteligência superior, não era possível determinar a ordenação mais adequada dos problemas na escala. Essa ordenação se tornou possível, a partir do trabalho experimental, edição de 1947, empreendido por G. A. Foulds, com a análise de itens, realizada por A. R. Forbes e H. G. Bevans, dos protocolos do teste, fornecidos pelas instituições que utilizaram a edição de 1947 para orientação ou seleção profissional. Na edição de 1962, foram excluídos da Série II doze problemas que nada

contribuíam à distribuição dos escores de adultos de inteligência superior à média. Os problemas restantes foram ordenados de acordo com a frequência com que eram resolvidos, à medida que o escore total da Série revisada variava de 0 a 36 pontos.

Em relação ao tempo e nível intelectual, segundo Campos (2002), a edição de 1962 das Matrizes Progressivas Avançadas de Raven, como um teste que se dispõe a medir a capacidade intelectual, está disposta de forma que pode ser aplicada sob duas modalidades com o mesmo valor: em tempo livre, para avaliar a capacidade total de observação e clareza de pensamento de uma pessoa, ou com tempo limitado, para a eficiência de seu trabalho intelectual. Na primeira tentativa, nenhum participante, nem mesmo de alta capacidade intelectual, resolve todos os problemas em menos de 40 minutos, enquanto o interesse e a atenção de uma pessoa se mantêm perfeitamente por mais de uma hora, sem fadiga.

Para avaliar a capacidade total de observação e a clareza de pensamento de uma pessoa, pode-se mostrar a Série I e explicar a mecânica do teste. Pode-se, então, apresentar a Série II e permitir que trabalhe em seu ritmo próprio, do o começo até o fim, sem interrupção. Quando se trata de avaliar o rendimento intelectual, pode-se dar a Série I como um breve teste de prática, seguido da Série II como teste de rapidez. O tempo concedido para a Série II pode variar; um período de aproximadamente 40 minutos proporciona a distribuição mais satisfatória de escores. A Série II permite explorar todas as operações de análise e integração envolvidas nos mais elevados processos de pensamento e discrimina claramente, mesmo entre pessoas de inteligência superior (Raven, 1962).

Em relação aos parâmetros psicométricos sob os quais o teste foi submetido, verifica-se a análise de itens, desenvolvida a partir da aplicação do instrumento em 2.256 sujeitos, com a finalidade de estudar o comportamento de cada item, selecionar os

itens a serem eliminados da escala e as figuras errôneas como resposta de cada problema apresentado. A verificação da precisão foi realizada por meio do teste-reteste, aplicado a uma amostra composta de 444 sujeitos, sendo apresentados coeficientes 0,76 e 0,86 entre os adolescentes e de 0,91 para os adultos, demonstrando inferioridade do primeiro grupo em relação ao segundo.

Em se tratando da validade do instrumento, os resultados obtidos se referem a uma comparação entre estudantes de escolas técnicas e universitários, além de estudos comparativos quanto ao tempo de execução e entre as diferentes escalas das Matrizes Progressivas. Tanto os estudos de validade quanto os de precisão foram realizados em diversos países, sendo eles Inglaterra, Alemanha, Estados Unidos da América, Austrália e Nova Zelândia (Raven & cols., 1998).

Contudo, verifica-se tanto no manual original (Raven & cols., 1998) quanto no manual em português (Campos, 2002) a ausência de estudos de evidência de validade e precisão para amostras brasileiras. Esse instrumento consta na listagem dos testes avaliados pelo Sistema de Avaliação de Testes Psicológicos (Satepsi) (Conselho Federal de Psicologia, 2007) com parecer desfavorável, ressaltando-se as seguintes deficiências: ausência de informações sobre precisão no Brasil; ausência de informações sobre validação no Brasil; ausência de informações sobre normas brasileiras; e ausência de definição de construto. Todos esses itens faltam no manual brasileiro.

Verificam-se, no entanto, estudos brasileiros com o instrumento, entre os quais se encontra um artigo publicado por Primi, Flores-Mendonza e Castilho (1998) relacionado com um estudo exploratório das propriedades psicométricas do teste Matrizes Progressivas Avançadas de Raven. Nesse estudo, foram investigadas as propriedades psicométricas de uma versão experimental informatizada do teste, no qual participaram 360 universitários que responderam ao instrumento de forma coletiva, sem

limite de tempo. Para a Série II, a dificuldade dos itens se correlacionou em 0,96 com a versão original do teste. Foi observada, também, a precisão do instrumento por meio da fórmula de Kuder-Richardson (0,83) e coeficiente de validade comparando a pontuação do teste com o desempenho acadêmico dos participantes, apresentando uma correlação de 0,43. Tais resultados permitiram concluir que a versão informatizada foi equivalente à original no que se refere às propriedades psicométricas.

Teste de Raciocínio Inferencial (RIn)

O Teste de Raciocínio Inferencial RIn de Sisto (2006) avalia, por meio do raciocínio inferencial, a inteligência ou fator g. Formado por quatro séries de itens organizados em ordem de dificuldade, apresenta, em todo início de série, exercícios que possibilitam o entendimento do raciocínio esperado. A partir do raciocínio inferencial, o sujeito deduz relações entre objetos, apoiado em um conjunto de suposições e informações que vêm à mente, anteriormente observado por meio de estímulos perceptuais, memória, crenças e conhecimentos, e de raciocínios dedutivos, e utiliza-se disso para escolher, no teste, entre as figuras que se apresentam na parte de baixo da folha, as quais completam a seqüência da figura acima.

A respeito do processo de raciocínio, a forma de agir e o conhecimento que se tem do mundo influenciam os dados que são selecionados, as interpretações que se fazem e as conclusões que são tiradas. Por isso, as pessoas inferem diferentes conclusões, valendo-se de diversos dados que, tratados como conclusões óbvias, permitem que não se sinta necessidade de justificar como se chegou a tais conclusões. A inferência é feita sempre que uma pessoa vai além da evidência dada pela informação. É o ato ou processo de tirar uma conclusão baseada unicamente no que se sabe, em uma informação que a pessoa possui (Sisto, 2006).

Como exemplo, segundo o mesmo autor, se você descobre que está chovendo ao olhar pela janela, é possível inferir que o céu está nublado. Em uma perspectiva lógica as inferências básicas são: a dedução, inferência do universal ao particular; e a indução, que recorre o caminho inverso. O raciocínio dedutivo satisfaz um padrão muito elevado da exatidão e traz consigo a afirmação de que a verdade de suas premissas garante a verdade de sua conclusão, o que chamamos de inferência dedutiva. Todavia, quando a verdade das premissas possibilita afirmar apenas que a conclusão é provável, trata-se de inferência indutiva. Nesse sentido, o padrão da exatidão do raciocínio indutivo é muito mais flexível, pois as premissas fornecem bases limitadas para aceitar a conclusão. Os argumentos dedutivos afirmam garantir as conclusões, enquanto os argumentos indutivos apenas as recomendam.

O teste RIn contou com uma amostra de 1.102 sujeitos, provenientes de cidades do estado de São Paulo (Assis, Bauru, Bragança Paulista, Jundiaí, Itatiba e São Paulo) e Minas Gerais (Pouso Alegre e Uberlândia). A idade dos sujeitos investigados apresentou média de 18,99 anos e desvio-padrão de 6,06, sendo a idade mínima de 11 anos e a idade máxima de 73 anos.

Em relação aos aspectos psicométricos do teste, foi realizado um estudo de validade de constructo por análise fatorial de todos os itens do instrumento, no qual se verificou que os fatores encontrados não se referiam a um constructo psicológico, mas às dificuldades dos itens. Considerando-se que as séries estão em ordem de dificuldade, itens fáceis representados pelo alto índice de acerto, aumentando em grau de dificuldade, optou-se por considerar cada série um subtteste e realizar uma análise fatorial com esses dados.

A medida de adequação da amostra por meio do método de Kaiser-Meyer-Olkin, apresentou resultado $KMO = 0,69$, indicação essa de que o grau de variância comum

entre as quatro variáveis poderia ser considerado medíocre, possibilitando a explicação de uma quantidade não muito grande de variância. Ao lado disso, usando o teste de esfericidade de Bartlett, com resultado de $\chi^2 = 796,893$, $df = 6$ e $p = 0,00$, pôde-se concluir que a matriz é apropriada para a extração de fatores. Por conta de tais resultados, destaca-se a necessidade de cautela na interpretação dos dados.

Para a verificação da unidimensionalidade do teste, o que permite inferir a presença de um fator geral, realizou-se a análise fatorial pelo método de extração por fatoração de eixos principais (PAF), com rotação *Oblimin*. Contudo, apenas um fator alcançou um *eigenvalue* igual ou superior a 1, elucidando 52,41% da variância. Tal indicativo de evidência de validade está relacionado também com sua estrutura interna.

O estudo da dimensionalidade do teste assume grande importância quando é considerada em relação à evidência de validade de um teste, posto que evidência de validade de constructo informa se uma interpretação teórica pode ser dada às pontuações do teste (APA, 1985). A unidimensionalidade é uma característica de testes que medem um único constructo psicológico, medindo apenas uma única habilidade ou característica do examinando.

Ainda de acordo com o manual do teste RIn (Sisto, 2006), entre as evidências de validade do instrumento se encontram estudos de análise dos itens por meio do modelo Rasch (1960), método esse empregado para escalar os itens em termos unidimensionais. O resultado da análise do ajuste forneceu informações sobre a dificuldade dos itens, o erro-padrão e os valores de *infit* e *outfit*. A média da medida Rasch foi 0,00, desvio-padrão de 1,28 e erro-padrão médio de 0,08. Pelos índices do *infit*, todos os itens apresentaram ajuste ao modelo, enquadrados nos intervalos 0,70 a 1,30, normalmente usado. Em relação ao *outfit*, dois itens não se ajustaram a esse critério, o que, no entanto, não é preocupante em razão do fato de seus valores estarem muito próximos

(1,31). O resultado dessa análise permitiu uma representação dos itens em uma escala do mais difícil para o mais fácil.

No que tange à evidência de validade relativa à estrutura interna, foi realizado uma análise por meio do funcionamento diferencial do item (DIF). Estudos com essa característica estão relacionados às implicações éticas e sociais, relacionadas à utilização de testes, no que se refere à possibilidade de avaliar sistematicamente as capacidades de certos grupos em função de sua cultura, etnia, gênero ou quaisquer outras características diferenciadoras, justificando a importância de estudo para neutralizar esse tipo de efeito.

Introduzido por Holland e Thayer, em 1988, citado por Sisto (2006), o termo funcionamento diferencial dos itens (*Differential Item Functioning*) foi proposto em substituição ao termo viés. A função desse tipo de análise é determinar os itens com probabilidades de acertos diferentes para subgrupos, nos quais os indivíduos possuem o mesmo nível na variável medida.

Estudos sobre o viés são importantes no contexto da validade de constructo dos itens, pois representam o grau em que um item ou conjunto de itens mede um traço ou constructo (Sisto, 2006). Dos 40 itens estudados no teste, apenas dois deles indicaram diferenciar gênero, A1 e B6; porém, como um favorece o sexo masculino e outro o feminino, não se realizaram mudanças no teste, pelo equilíbrio de tal interferência.

Entre as evidências de validade de critério, uma delas está relacionada com o desenvolvimento dos sujeitos investigados. Foram feitas análises para estudar as relações entre o RIn e as idades e gênero dos participantes. Essas análises, além de fornecerem mais informações sobre o comportamento do instrumento, estarão relacionadas com a elaboração das normas, no sentido de verificar se todas as idades necessitam de normas específicas ou se algumas idades se agrupam, porque suas diferenças podem ser atribuídas ao acaso.

Foram realizadas também análises em relação ao gênero dos sujeitos. Contudo, as correlações encontradas foram muito baixas e apenas duas delas significativas. Para o sexo masculino, os dados demonstram que, conforme o aumento da idade, se aumentam também as pontuações no teste, enquanto, para as mulheres, conforme aumenta a idade, a pontuação diminui.

Verificou-se também o indicativo de que há um aumento das pontuações até determinada idade e depois inicia uma diminuição. Em razão disso, os participantes foram divididos em dois grupos, sendo um deles formado por sujeitos de até 22 anos e outro com idades superiores a essa. Foram calculados novamente os coeficientes de correlação em razão dos grupos formados. Constatou-se por esses coeficientes um aumento em relação à correlação geral (-0,01). Verificou-se também que até os 22 anos a correlação é positiva, o que possibilita inferir que conforme o aumento da idade aumenta-se também as pontuações obtidas no RIn. Já para o grupo com idade acima de 22 anos, a correlação foi negativa, o que permite interpretar que, conforme aumenta a idade, há uma diminuição da pontuação no RIn.

Sisto (2006) destaca que essa tendência condiz com as pesquisas tradicionais sobre inteligência, que apontam um crescimento até um dado momento da vida do indivíduo e depois uma diminuição ordenada. Como essa diminuição já aparece entre os 20 e 30 anos, é possível inferir que se refere mais a uma inteligência fluida ou fator g do que a uma inteligência cristalizada, cuja diminuição é bem mais tardia.

Em relação aos estudos de precisão, segundo o manual do teste, o coeficiente encontrado demonstrou estabilidade dos resultados obtidos, o que indica que a medida da avaliação não está relacionada com erros casuais. Essa estabilidade pode se referir aos itens estarem de forma semelhante ao longo do instrumento ou se referir a possibilidade de que em uma segunda aplicação o mesmo resultado da primeira seria

obtido. Sendo assim, confirma-se que o teste apresenta uma precisão ou fidedignidade razoável para um diagnóstico do que se pretende medir.

A fidedignidade do teste foi estudada em relação às faixas etárias e sexo por meio do coeficiente alfa de Cronbach e as duas metades de Spearman-Brown. Calculou-se, também, a precisão do instrumento em sua totalidade pelo modelo Rasch. Foram encontrados os coeficientes considerados bastante satisfatórios, variando de 0,71 a 0,85 pelo coeficiente de consistência interna (alfa de Cronbach) e de 0,73 a 0,93 pelo método das duas metades (Spearman-Brown). Esses coeficientes verificados por esses dois métodos demonstram uma freqüência constante ao longo das faixas etárias.

No que tange ao instrumento e aos participantes que responderam ao teste, os indicadores alfa, Spearman-Brown e Rasch apresentaram resultados também bastante satisfatórios. Pela análise dos coeficientes alfa, no caso de que algum item fosse retirado do instrumento, observou-se que não haveria aumento da precisão.

OBJETIVOS

Objetivo Geral

Buscar evidências de validade para o teste Matrizes Progressivas Avançadas de Raven por meio do teste Raciocínio Inferencial (RIn).

Objetivos Específicos

Correlacionar o teste Matrizes Progressivas Avançadas de Raven com o teste Raciocínio Inferencial – RIn;

Verificar a relação entre os escores obtidos em função do gênero dos sujeitos analisados;

Averiguar as diferenças entre os escores relacionando-os com a faixa etária dos sujeitos.

MÉTODO

Participantes

A aplicação dos instrumentos para coleta dos dados foi realizada em 291 estudantes de ensino superior de duas universidades privadas no estado de São Paulo, sendo 232 (79,7%) de uma universidade na cidade de São Paulo e 59 (20,3%) de uma universidade de São José dos Campos selecionados em amostragem por conveniência. Dos participantes investigados, 87 (29,9%) são do sexo masculino e 204 (70,1%) do sexo feminino. As idades variaram entre os 17 e 63 anos, com média de 28 anos e desvio-padrão (DP) de 9,42.

Em relação aos cursos, verificam-se 107 (36,8%) alunos de Administração de Empresas, 52 (17,9%) alunos de Gestão de Recursos Humanos e 132 (45,4%) alunos de Psicologia. No que tange aos instrumentos aplicados, todos os participantes analisados neste estudo responderam ao teste Raven e teste RIn.

Instrumentos

Para a coleta de dados foram utilizados o Teste de Matrizes Progressivas Avançadas de Raven e o Teste de Raciocínio Inferencial (RIn). O material é constituído de caderno de aplicação e folha de respostas, de ambos os testes, e lápis para anotação.

a) Matrizes Progressivas Avançadas de Raven

O Teste das Matrizes Progressivas Avançadas de Raven, padronizado e publicado em 1938, foi desenvolvido por Raven para abranger pessoas com capacidade intelectual superior à média, sendo mais usado para pessoas com escolaridade universitária.

O nome original do instrumento é *Advanced Progressive Matrices (APM)*, de autoria de J. C. Raven, com tradução e adaptação brasileira de Francisco Campos, produzido pelo CEPA e manual mais recente com data de 2002. A aplicação pode ser individual ou coletiva, e o tempo máximo de realização é 60 minutos. O material consiste em manual, caderno de aplicação, folha de resposta e crivo de apuração. O instrumento destina-se à avaliação de adolescentes (a partir de 12 anos) e adultos.

A aplicação das Matrizes Progressivas Avançadas de Raven é realizada por meio de dois cadernos de resposta, denominados Séries I e II. A Série I é composta de 12 itens, e a Série II, de 36 itens. O nível de dificuldade aumenta a cada item em ambas as séries. A Série II possui um nível de dificuldade superior à Série I. Entre as funções da primeira série, esta serve como exemplo para a resolução da série seguinte e, principalmente, para indicar quando a segunda série não deve ser aplicada nos respondentes que apresentarem dificuldades em responder a Série I e, conseqüentemente, não conseguirão completar o restante do teste.

Em cada item é apresentado um quadro com uma figura na qual falta um pedaço. Abaixo desse quadro, são apresentadas oito opções com possíveis complementos para esse pedaço. O participante deve assinalar o pedaço que melhor completar a figura do quadro, anotando a opção na folha de resposta.

A correção do teste é feita atribuindo-se um ponto à resposta correta e zero à resposta considerada errada. Se o sujeito realizar a marcação de mais de uma resposta, ela é considerada incorreta, atribuindo-se pontuação zero, conforme explicado anteriormente, a menos que o sujeito tenha corrigido, como instruído na folha de resposta, riscando a marcação rejeitada e assinalando a resposta correta ao lado. Em seguida, soma-se a pontuação obtida e transformam-se os dados em percentil observando as tabelas de padronização.

Neste estudo, foi utilizado apenas o segundo caderno, denominado de Série II. O objetivo foi diminuir o desgaste ao responder o teste, por causa da utilização de dois instrumentos que medem inteligência geral por meio de estímulos não-verbais. Outro aspecto está relacionado com a escolaridade dos respondentes, da qual, por se tratarem estes de sujeitos cursando o ensino superior, se esperava que não houvesse necessidade de testar a capacidade de resolução do teste, acreditando-se que eles tivessem um nível de inteligência mínimo necessário para responder as questões da segunda série.

b) Raciocínio Inferencial (RIn)

Desenvolvido por Sisto (2006), o teste avalia a inteligência ou fator *g*, por meio do raciocínio inferencial. Composto de quatro séries organizadas em ordem crescente de dificuldade apresenta, no início de cada série, exercícios que podem possibilitar o entendimento do raciocínio e a resolução da tarefa.

Para cada exercício é oferecido, na parte superior da página, um quadro faltando uma parte. Logo abaixo desse quadro, são apresentadas possibilidades de respostas para complementar a parte faltante. Cabe ao analisando informar qual das opções completa melhor a figura em questão.

O RIn é um instrumento destinado a pessoas dos 10 a 70 anos, tendo um tempo máximo de 25 minutos de aplicação, tempo limite para a resolução dos exercícios. O teste pode ser aplicado tanto individual quanto coletivamente, em grupos de 20 a 30 pessoas. É conveniente aplicar a prova quando as pessoas não estiverem cansadas e em uma única sessão.

A instrução do exercício deve ser lida com o respondente e somente em seguida pedir que resolvam as questões. Deve-se solicitar a todos que marquem a alternativa na folha de resposta, verificando se marcaram corretamente as respostas, como também

perguntar se todos marcaram e se alguém tem dúvida. Durante a aplicação, deve-se solucionar dúvidas somente se necessário, restringindo-se às instruções específicas do teste. Também, depois de iniciada a prova, deve-se evitar interromper as pessoas ou distrair sua atenção.

A correção é feita atribuindo-se um ponto à resposta correta e zero ponto à resposta considerada errada. Duas medidas são apresentadas para a interpretação dos resultados. Uma delas se refere à pontuação bruta, ou seja, a soma de pontos obtidos pelos acertos no teste e a outra é uma medida derivada do modelo Rasch.

Procedimento

A pesquisa foi submetida à aprovação do Comitê de Ética em Pesquisa da Universidade São Francisco. As instituições assinaram um termo de adesão à pesquisa. A aplicação foi realizada coletivamente nas salas de aula, pelo próprio pesquisador, em horário combinado com a instituição sempre precedida pela leitura e assinatura do Termo de Consentimento Livre e Esclarecido. A aplicação dos instrumentos teve uma duração em torno de uma hora e vinte e cinco minutos.

Outro fator de que se procurou tratar no decorrer da aplicação da pesquisa foi a fadiga. Assim, os testes foram aplicados sob três formas: alternando a sua ordem, aplicando-se primeiro o RIn e, em seguida, o Raven, e vice-versa, e aplicando-se um em um dia e o outro com intervalo de sete dias.

RESULTADOS

As informações a seguir referem-se aos resultados obtidos a partir da pontuação total do teste Raven e do teste RIn, e também dos totais das séries do segundo. Inicialmente realizaram-se análises da precisão dos instrumentos, por meio do alfa de Cronbach. O teste Matrizes Progressivas Avançadas de Raven apresentou um coeficiente de 0,87 e o Teste de Raciocínio Inferencial (RIn) um alfa de 0,80. Como os instrumentos apresentaram índices satisfatórios, isto é, iguais ou superiores a 0,80, pôde-se inferir que os testes forneceram um bom nível de precisão.

Em relação à pontuação total no teste Raven, que pode variar de 0 a 36 acertos, verificou-se que os participantes apresentaram média de 14,94 pontos, com DP de 6,52, 15 pontos de mediana e 21 de moda. A menor pontuação encontrada foi de 1 acerto variando até 33 a maior pontuação. A figura 1 oferece a representação da pontuação total do teste Raven de acordo com a frequência dos acertos.

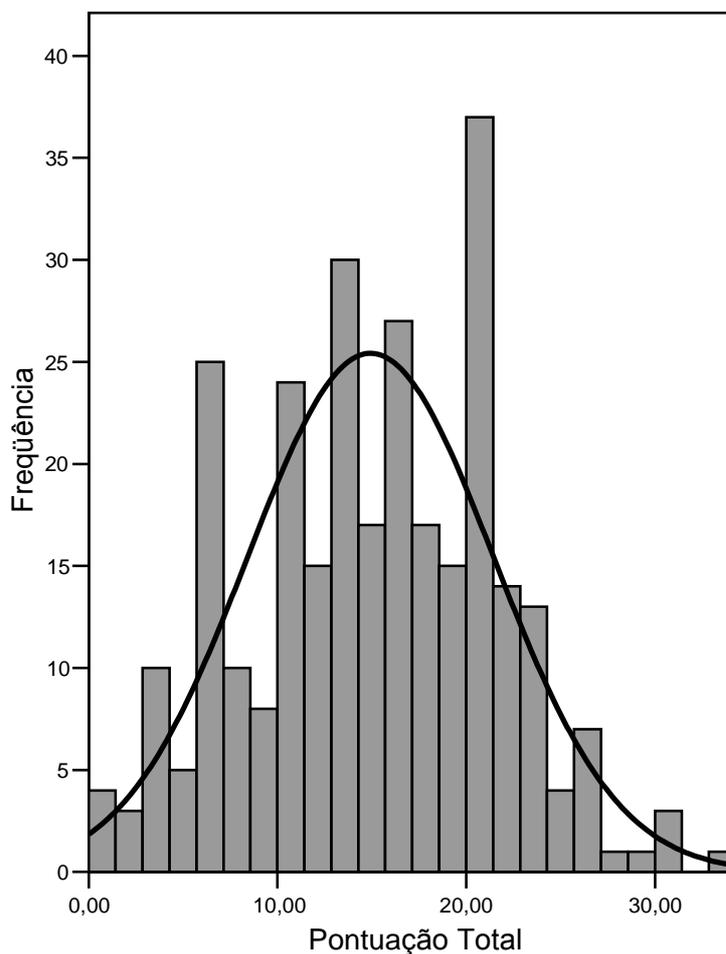


Figura 1. Frequência da pontuação total do teste Raven

Observa-se que a maior parte dos respondentes (55,6%) apresentaram uma pontuação entre 10 e 20 acertos, 22,3% obtiveram uma pontuação entre 1 e 9 pontos e uma terceira parte dos participantes apresentaram acertos entre 21 e 33 pontos (21,8%). Esses dados possibilitam inferir que o teste apresentou uma boa distribuição na frequência de pontuação na amostra investigada. Contudo, verifica-se que em relação às pontuações mais altas, apenas 2,9% apresentaram pontuações acima de 26 acertos e nenhum respondente apresentou mais que 33 acertos.

No que se refere ao total de pontos do teste RIn, observou-se uma média de 23,03 acertos, com DP de 5,79, 24 pontos de mediana e moda. Apesar de a pontuação

total variar de 0 a 40 acertos, no grupo estudado, o mínimo foi de 4 pontos e o máximo de 35. A figura 2 apresenta essa distribuição em razão da pontuação total do teste RIn.

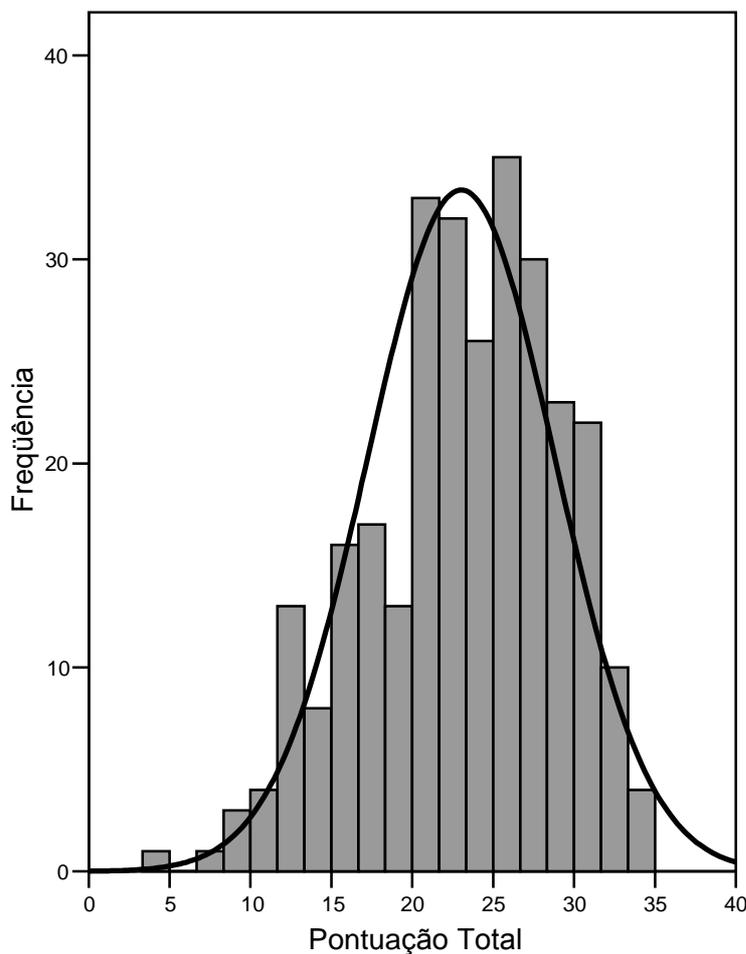


Figura 2. Frequência da pontuação total do teste RIn

No que tange a frequência da pontuação observada na figura 2, verifica-se a ausência de pontuações muito baixas, próximas ao zero, sendo que os primeiros 10 pontos da escala comportaram apenas 1,7% de participantes, 24,4% acertaram entre 11 e 19 pontos. A grande parte dos respondentes (67%) apresentaram de 20 a 30 pontos. Em virtude disso, verifica-se uma tendência dos participantes investigados a obterem uma pontuação acima da média. Vale ressaltar que apenas 6,7% obtiveram entre 31 e 35 acertos, e que não foi atingida a pontuação máxima do teste.

Os valores de média, DP, mediana, moda, mínimo e máximo de cada uma das séries do teste RIn também foram investigados. Esses dados podem ser observados na tabela 1.

Tabela 1. Pontuação dos participantes nos testes Raven e RIn

	<i>RIn</i> <i>Série A</i>	<i>RIn</i> <i>Série B</i>	<i>RIn</i> <i>Série C</i>	<i>RIn</i> <i>Série D</i>	<i>Total RIn</i>	<i>Total Raven</i>
Média	9,33	4,71	2,63	6,36	23,03	14,94
Mediana	10,00	5,00	2,00	7,00	24,00	15,00
Moda	11,00	5,00	2,00	7,00	24,00	21,00
Desvio-padrão	1,94	1,90	1,63	2,79	5,79	6,52
Mínimo	1,00	1,00	0,00	0,00	4,00	1,00
Máximo	12,00	8,00	6,00	13,00	35,00	33,00
Número de itens	12	8	6	14	40	36

Para a realização da análise das associações entre os instrumentos em questão, realizou-se correlações de *Pearson* entre os acertos nas pontuações dos testes. Com base na tabela 2, observou-se correlações com magnitudes significativas entre a pontuação total do teste Raven e as séries e o total do teste RIn.

Verificaram-se correlações moderadas positivas entre a pontuação total do Raven e as três primeiras séries do teste RIn respectivamente ($r = 0,52$; $r = 0,47$; $r = 0,49$) e uma correlação baixa positiva ($r = 0,30$) com a Série D. Em relação as pontuações totais dos testes, observa-se uma correlação alta positiva ($r = 0,61$).

Tabela 2. Correlação de *Pearson* entre os acertos nas pontuações dos testes

		<i>Total Raven</i>
RIn Série A	<i>R</i>	0,52
RIn Série B	<i>R</i>	0,47
RIn Série C	<i>R</i>	0,49
RIn Série D	<i>R</i>	0,30
Total RIn	<i>R</i>	0,61

Procurou-se também, neste estudo, verificar as diferenças nas pontuações dos testes em relação ao gênero dos participantes. Para tal, realizaram-se as mesmas correlações de *Pearson* entre os testes para o gênero masculino e feminino conforme pode ser observado na tabela 3.

Tabela 3. Correlação de *Pearson* entre os acertos nas pontuações dos testes, por sexo.

<i>Sexo</i>		<i>r</i>	<i>Total Raven</i>
Masculino (N=87)	RIn Série A	<i>r</i>	0,50
	RIn Série B	<i>r</i>	0,41
	RIn Série C	<i>r</i>	0,46
	RIn Série D	<i>r</i>	0,45
	Total RIn	<i>r</i>	0,63
Feminino (N=204)	RIn Série A	<i>r</i>	0,53
	RIn Série B	<i>r</i>	0,49
	RIn Série C	<i>r</i>	0,50
	RIn Série D	<i>r</i>	0,24
	Total RIn	<i>r</i>	0,60

Observou-se, por meio da tabela 3, que os resultados mantiveram correlações similares aos dados obtidos na tabela 2, com exceção da correlação entre a pontuação total do teste Raven e a Série D do teste RIn. Vale destacar a discrepância observada nesta última, pois, os homens apresentaram uma correlação considerada moderada positiva ($r=0,45$), enquanto a mulheres obtiveram uma correlação muito baixa, considerada estatisticamente fraca ($r= 0,24$).

Ainda no que se refere ao gênero dos participantes e os resultados encontrados nos testes, com o objetivo de verificar a influência da variável sexo sobre as correlações entre os instrumentos, realizou-se também uma correlação parcial entre os totais do Raven e RIn e suas séries. Contudo, conforme observado na tabela 4, as mesmas correlações foram evidenciadas.

Tabela 4. Correlação parcial entre os acertos nas pontuações dos testes controlando a variável sexo

		<i>Total Raven</i>
RIn Série A	<i>r</i>	0,52
RIn Série B	<i>r</i>	0,47
RIn Série C	<i>r</i>	0,49
RIn Série D	<i>r</i>	0,30
Total RIn	<i>r</i>	0,61

Procurando evidenciar as possíveis variações entre os gêneros dos participantes, realizou-se um teste *t* de Student das pontuações encontradas nos testes. Por meio da análise de variância verifica-se que o sexo masculino apresentou uma média de 23,7 e DP 5,86 no teste Raven e média de 16,37 e DP 6,61 no teste RIn, médias essas superiores às das mulheres que obtiveram no Raven uma média de 22,74 e DP 5,76 e no RIn uma média de 14,32 e DP 6,41. Os resultados obtidos podem ser apreciados na tabela 5.

Tabela 5. Teste *t Student* das pontuações dos testes por meio da variável sexo

	<i>Sexo</i>	<i>M</i>	<i>DP</i>	<i>t</i>	<i>p</i>
Total RIn	Masculino	23,7	5,86	1,30	0,196
	Feminino	22,74	5,76		
Total Raven	Masculino	16,37	6,61	2,469	0,014
	Feminino	14,32	6,41		

Entre os dados obtidos observa-se, contudo, que apenas a diferença de média na pontuação total do teste Raven pode ser considerada estatisticamente significativa. Verifica-se uma diferença de médias entre homens e mulheres de 2,04 pontos e um $t = 2,47$. Segundo Sisto (2006), por meio da prova *t* de *Student*, calcula-se se a diferença pode ser atribuída ao acaso ou não, o que pode ser observado, verificando-se um valor

de t maior do que 2, considerado assim, significativo. Contudo, em relação ao RIn, apesar de se observar uma diferença de médias entre homens e mulheres, estes valores não são considerados estatisticamente significativos. Tais informações corroboram com as pesquisas realizadas por Sisto (2006), descritas no manual do instrumento.

No que tange aos estudos relacionados à variável idade, verifica-se por meio das figuras 3 e 4 a pontuação média dos participantes por idade nos testes, independentemente de sexo. Esse gráfico sugere uma leve diminuição das pontuações com o avanço da idade no teste Raven; contudo, observam-se picos de pontuação nas idades mais avançadas. Ressalta-se que esse leve declínio não ficou evidenciado em relação ao teste RIn, mas este apresentou também muitos picos de pontuação nas idades mais avançadas.

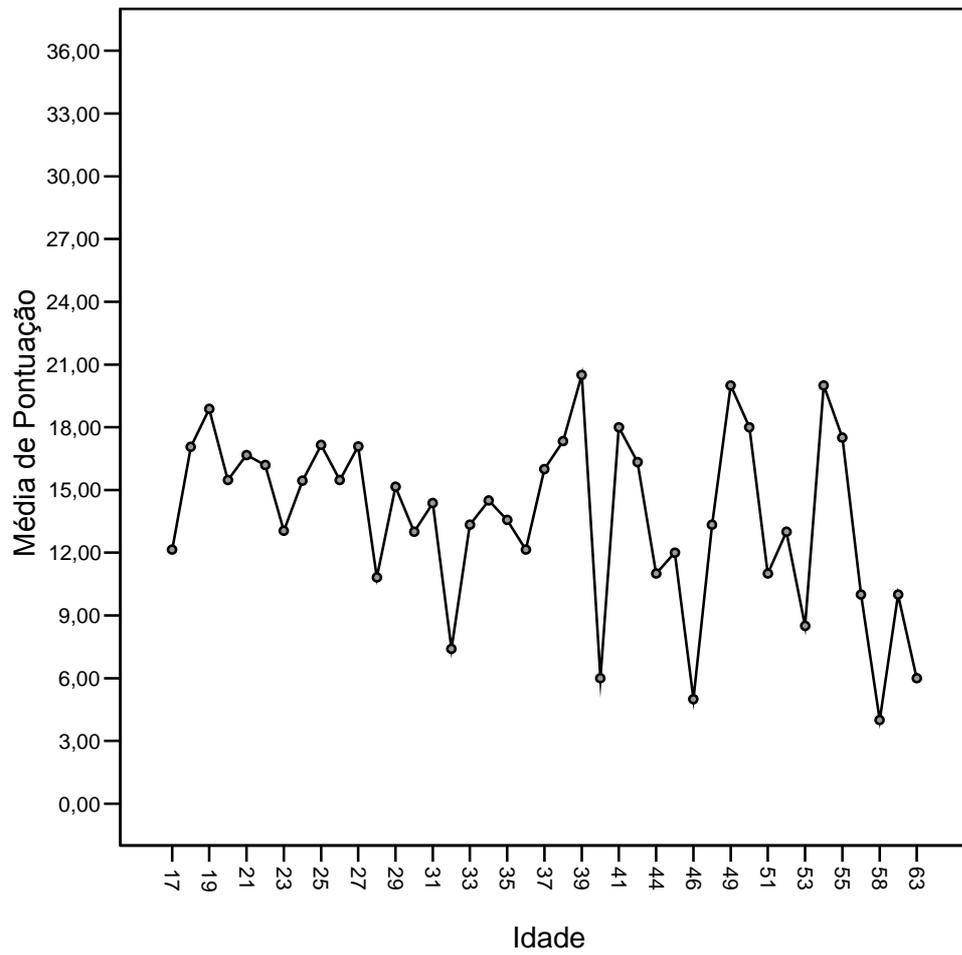


Figura 3. Média de pontuação dos participantes no teste Raven por idade

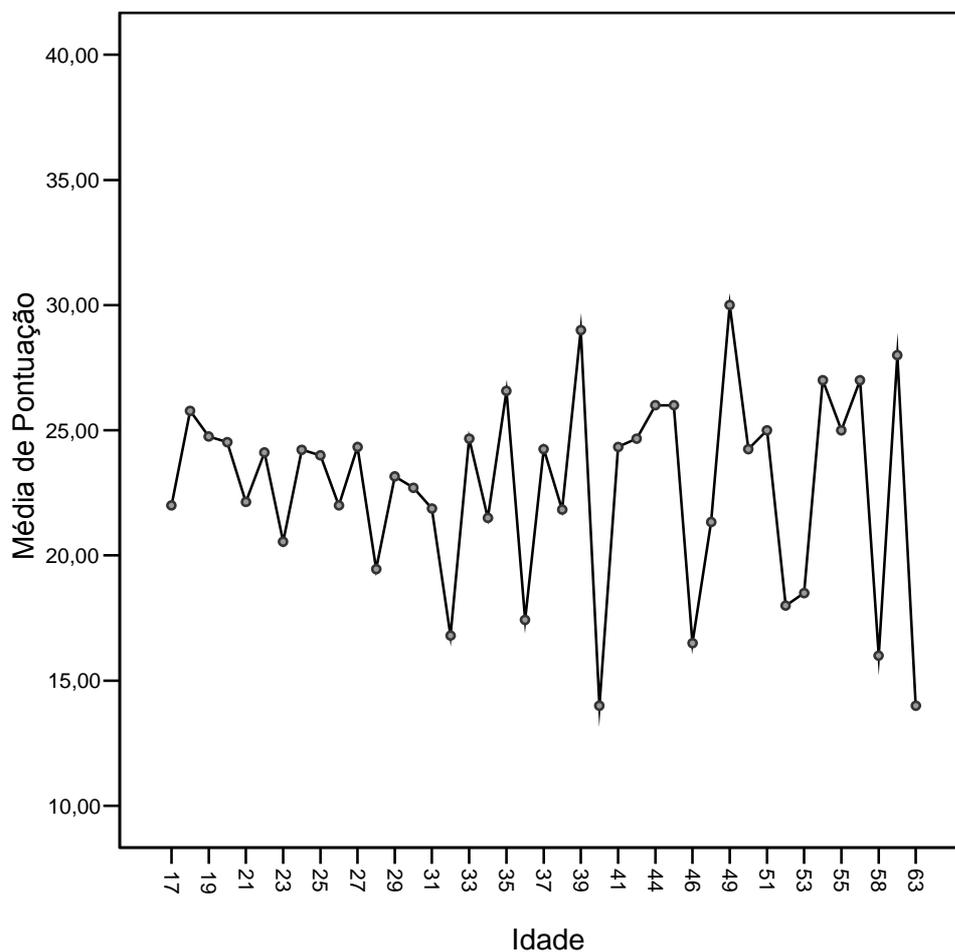


Figura 4. Média de pontuação dos participantes no teste RIn por idade

Visando à investigação da influência da faixa etária dos respondentes sobre as pontuações nos testes, realizou-se uma correlação de *Pearson* entre a variável idade e os acertos nas pontuações dos testes. Conforme pode ser observado na tabela 6, não se evidenciou qualquer correlação estatisticamente significativa entre as pontuações e as idades dos participantes.

Tabela 6. Correlação de *Pearson* entre a idade e os acertos nas pontuações dos testes

		<i>Idade</i>
RIn Série A	<i>r</i>	-0,09
RIn Série B	<i>r</i>	-0,02
RIn Série C	<i>r</i>	-0,02
RIn Série D	<i>r</i>	-0,06
Total RIn	<i>r</i>	-0,07
Total Raven	<i>r</i>	-0,15

Realizou-se também uma correlação parcial entre as pontuações totais dos testes e as séries do RIn, a fim de confirmar a influência da variável idade, sobre os resultados. Contudo, verifica-se, por meio da tabela 7, que as correlações mantêm suas características já avaliadas na tabela 2, sugerindo que as pontuações não sofrem influência da variável independente em questão.

Tabela 7. Correlação parcial entre os acertos nas pontuações dos testes controlando a variável idade

		<i>Total Raven</i>
RIn Série A	<i>R</i>	0,52
RIn Série B	<i>R</i>	0,48
RIn Série C	<i>R</i>	0,49
RIn Série D	<i>R</i>	0,30
Total RIn	<i>R</i>	0,61

Por fim, com vista à exploração de diferenças relacionadas à variável idade, recorreu-se à análise de variância (ANOVA), que analisa as diferentes fontes de variação que podem ocorrer em um conjunto de valores. A análise de variância pode ser usada para testar a hipótese de que várias médias são iguais e a significância das possíveis diferenças (Dancey & Reidy, 2006).

Os resultados encontrados estão descritos na tabela 8. Contudo, ressalta-se que não se evidenciou nenhuma diferença estatisticamente relevante.

Tabela 8. Análise de variância (ANOVA) entre a pontuação dos testes em relação à variável idade.

	Média dos quadrados	<i>F</i>	<i>p</i>
Total RIn	44,374	1,39	0,070
Total Raven	55,965	1,39	0,071

No que se refere às médias de acertos entre as idades, pôde-se observar que, mesmo havendo diferenças nas pontuações para a faixa etária, tais diferenças não apresentam valores consideráveis estatisticamente. Tais ausências de significâncias estatísticas ficam evidenciadas em ambos os testes.

DISCUSSÕES E CONSIDERAÇÕES FINAIS

Diante da importância da avaliação psicológica, como um processo integrado, que abrange várias etapas, uma delas podendo envolver a utilização de instrumentos psicológicos (Anastasi & Urbina, 2000; Pasquali, 2001), o presente estudo contribuiu para o aprofundamento de pesquisas sobre a inteligência em participantes do ensino superior. Pasquali (1997) afirma que os testes psicológicos devem assegurar as características psicométricas básicas para sua utilização, permitindo qualificar uma avaliação como cientificamente respaldada. No processo de escolha do instrumento, é possível encontrar testes que preencham apenas parcialmente os requisitos necessários, exigindo uma análise que busca investigar até que ponto o teste avalia aquele aspecto que se pretende medir. Entre as análises do instrumental a ser usado, deve-se atentar para os estudos sobre a validade, a precisão e a padronização do instrumento.

Como objetivo principal deste estudo, buscou-se verificar evidências de validade para o instrumento Matrizes Progressivas Avançadas de Raven. Assim, envolve um componente operacional, referente à sua avaliação sistemática, geralmente comparando-se com um critério externo — neste caso, realizando-se a comparação entre os dois instrumentos por meio de um estudo correlacional com o teste RIn. Sendo este segundo um teste aprovado como apto para a utilização com a população brasileira, com diferentes estudos estatísticos evidenciando sua possibilidade de utilização apropriada, segundo normas do CFP, e tendo-se verificado uma correlação alta positiva entre os totais dos dois testes, infere-se que se trata de construtos relacionados. Mais especificamente, esse tipo de validade do instrumento pode ser definida como a capacidade de realmente medir aquilo a que se propõe (Anastasi & Urbina, 1977).

Uma vez verificada a validade dos instrumentos, corroborando com os estudos preconizados nas propostas dos autores de cada teste, outro ponto expressivo é o grau com que o instrumento avalia sem variar na medida (Anastasi, 1977). Denomina-se esses estudos de precisão ou fidedignidade, entendida como a estabilidade que o teste tem em manter seu resultado depois de determinado tempo ou mesmo de sua consistência/homogeneidade, sendo este segundo a forma de verificação neste trabalho.

Observou-se, então, por meio da análise de confiabilidade, que os instrumentos utilizados apresentaram índices satisfatórios pelo alfa de Cronbach. Esses dados nos permitem inferir que os testes possuem um bom nível de precisão, podendo ser considerados instrumentos confiáveis para a medida do respectivo construto neste grupo analisado.

Em relação às médias de acertos nos totais dos instrumentos, observou-se no teste Raven uma concentração nas pontuações entre 10 e 20 acertos, o que corresponderia a uma curva com boa distribuição na frequência de pontuação na

amostra investigada. Contudo, ainda no mesmo teste, verificaram-se poucos acertos nas pontuações mais altas.

No teste RIn, verificou-se a ausência de pontuações muito baixas, próximas ao zero. As pontuações começam a ter maior representatividade a partir de 10 pontos. A grande parte dos respondentes apresentou de 20 a 30 pontos, sugerindo uma tendência dos participantes investigados a obter uma pontuação acima da média. Destacou-se também que apenas uma minoria obteve entre 31 e 35 acertos e que não foi atingida a pontuação máxima do teste.

Levando-se em consideração que o teste Raven foi construído para avaliar indivíduos superiores à média, neste estudo, em alunos de nível superior, e sendo o RIn um instrumento padronizado com estudantes desde o ensino fundamental ao superior, esperava-se que a média de pontuação no teste Raven fosse mais baixa do que no testes RIn, representando, assim, essa escolaridade mais alta. As análises anteriormente destacadas corroboram com a hipótese levantada.

Em relação ao decréscimo na média de acertos observado entre as séries do teste RIn, isso é esperado, considerando-se que, de acordo com a construção do instrumento, há uma ordem de dificuldade dos itens que aumenta a cada série. Esses dados corroboram com os estudos realizados com o teste (Sisto, 2006). Vale destacar que, apesar de a média de acertos na última série ter sido mais alta, em relação às duas séries anteriores, isso se dá em razão do número de itens maior na última série. Contudo, proporcionalmente, tal decréscimo fica evidenciado neste estudo, como já informado.

Diante do pressuposto de que há uma relação entre o teste Matrizes Progressivas Avançadas de Raven e o Teste de Raciocínio Inferencial (RIn), verificou-se, entre as pontuações totais dos testes, uma correlação alta positiva. Observaram-se também correlações moderadas positivas entre a pontuação total do Raven e as três primeiras

séries do teste RIn e uma correlação positiva baixa com a Série D. Inferiu-se, por meio desses dados, que os testes avaliam construtos relacionados, neste caso, ambos os instrumentos avaliam a inteligência por meio de um fator geral, determinado de fator *g* de acordo com a teoria de Spearman. Essa relação entre inteligência e raciocínio é tratada por diferentes autores (Almeida, 1988; Sisto, 1975; Stenberg, 1982). Sternberg (1982) trata esses dois conceitos tão próximos tornando complexa sua diferenciação, sendo esta uma posição semelhante à apresentada por Spearman (1927). Sisto (1975) afirma que os processos envolvidos nos tipos de raciocínio comumente fazem parte do constructo inteligência. Para Almeida (1988), o raciocínio nada mais é do que uma aptidão intelectual, que integra os processos cognitivos mais elaborados do ponto de vista intelectual; porém, é apenas um dos componentes da inteligência.

Com o objetivo de verificar se houve diferenças entre os participantes do sexo masculino e feminino, realizaram-se as mesmas correlações de *Pearson* entre os testes; contudo, diferenciando-se grupos de homens e mulheres. Os resultados mantiveram correlações similares aos dados obtidos antes da divisão de gêneros, com exceção da correlação entre a pontuação total do teste Raven e a Série D do teste RIn, na qual os homens apresentaram uma correlação moderada positiva, enquanto as mulheres apresentaram uma correlação fraca. Uma das justificativas que poderiam relacionar-se a esta discrepância foi a possível ausência de controle do viés relacionado ao gênero da amostra investigada na padronização do teste Raven, visto que, nos dados de padronização do teste RIn, não se verificam diferenças de gênero entre as pontuações, sabendo-se que foi realizado um controle por meio de uma análise de funcionamento diferencial dos itens (DIF), objetivando tal influência.

O estudo sobre viés dos itens está relacionado à validade de constructo, ou seja, o grau em que um item ou o conjunto dos itens medem um traço ou constructo. Sendo

esse estudo realizado no teste Rin, possibilitou informar que os itens do instrumento mantiveram o mesmo comportamento estatístico, quando comparados em grupos de sujeitos pertencentes ao sexo masculino e feminino da mesma população, comprovando, assim, a ausência de funcionamento diferencial dos itens (Sisto, 2006).

Com o mesmo objetivo, o de verificar a influência da variável sexo sobre as correlações entre os instrumentos, realizou-se também uma correlação parcial entre os totais do Raven e RIn e suas séries. Contudo, as mesmas correlações foram evidenciadas. Esses dados sugerem que a variável sexo não influencia significativamente os dados encontrados nas correlações entre os dois testes, conforme observado nos estudos realizados por Alves e Duarte (1993), Jensen (1998), Colom e cols. (2000), Colom, Garcíapz e cols. (2001), Meio (1998), entre outros. Os resultados corroboram também com o manual do teste RIn, em Sisto (2006), que, entre as evidências de validade de critério, apresentaram análises para estudar as relações entre o RIn e as idades e gênero dos participantes. Contudo, as correlações encontradas foram muito baixas.

Procurando ainda evidenciar as possíveis variações entre os gêneros dos participantes, realizou-se um teste *t* de Student das pontuações encontradas nos testes. Por meio dos estudos, verificou-se que os homens apresentaram uma média superior a das mulheres. Entre os dados obtidos, observa-se, contudo, que apenas a diferença de média na pontuação total do teste Raven pode ser considerada estatisticamente significativa. Essa diferença de média encontrada corrobora com as pesquisas realizadas por Allik e cols. (1999), Alves (1998), Lynn (1999), Santos e cols. (2002), entre outros, que demonstram, em suas pesquisas, diferenças de pontuação entre homens e mulheres em instrumentos de avaliação da inteligência. No teste RIn, apesar de este apresentar uma diferença de médias entre homens e mulheres, tais valores não foram considerados

estatisticamente significativos, confirmando, por meio de mais essa análise, a ausência de diferenças entre sexos, conforme descrito no manual do instrumento (Sisto, 2006).

No que tange aos estudos relacionados à variável idade, verificou-se, por meio dos gráficos, uma leve diminuição das pontuações com o avanço da idade no teste Raven; contudo, observam-se picos de pontuação nas idades mais avançadas. Esses achados corroboram com a pesquisa de padronização do teste WAIS, no qual foi observado um leve declínio a partir dos 30 até os 60 anos (Anastasi, 1977). Ressalta-se que esse leve declínio ficou ainda menos evidenciado em relação ao teste RIn. Este apresentou também muitos picos de pontuação nas idades mais avançadas. Em ambos os testes, sugerem-se outras pesquisas com idades avançadas para observar se este declínio na inteligência se mostra mais acentuado com a senilidade, conforme pesquisas realizadas por outros autores (Cunha, 1974; Raven & cols., 1985).

Visando a aprofundar um pouco mais a investigação da influência da idade, realizou-se uma correlação de *Pearson* entre a variável idade e os acertos nas pontuações dos testes. Contudo, não se evidenciou qualquer correlação de magnitude representativa entre as pontuações e as idades dos participantes.

Realizou-se, em seguida, uma correlação parcial entre as pontuações totais dos testes e as séries do RIn, a fim de confirmar a influência da variável idade, sobre os resultados. Entretanto, as correlações não sofreram alterações, o que sugere a ausência de influência da variável em destaque nas pontuações dos instrumentos.

Por fim, com vista à exploração de diferenças relacionadas à faixa etária dos participantes, recorreu-se à análise de variância (ANOVA). Não obstante, ressalta-se que não se evidenciou nenhuma diferença estatisticamente relevante.

Assim sendo, confirma-se a ausência de diferenças estatisticamente significativas entre as pontuações e as idades dos participantes, permitindo, assim, a

compreensão de que não há nesse grupo uma diferença de desempenho nos testes relacionado com a idade do respondente.

Segundo Matarazzo (1976), dados de avaliação da inteligência em relação à idade levam à impressão de um declínio com a idade, que é apenas aparente, visto que as pesquisas longitudinais evidenciam que a inteligência continua a aumentar levemente até os 50 anos e, então, começa a mostrar um pequeno declínio acima dessa idade. Raven e cols. (1991) relatam que não há um declínio com a idade, mas que o ano de nascimento influencia no resultado, ou seja, pessoas nascidas há mais tempo têm resultados menores, mostrando a influência de outras variáveis no sentido de aumentar os resultados médios com o passar dos anos.

Os resultados observados nesta pesquisa com a variável idade corroboram com Anastasi (1977), que afirma que, no passado, autores discutiram se a inteligência parava de aumentar por volta dos 16 ou 18 anos, sendo isso observado nos limites usados pela Escala Stanford-Binet (1937; 1960). Contudo, em adultos, a idade é uma variável mais complexa, porque se espera uma estagnação da inteligência em função da idade (Alves, 1998).

Vale destacar a importância de análises de critério para os instrumentos, pois, além de fornecerem mais informações sobre o comportamento dos instrumentos, propiciam a elaboração de normas específicas para cada gênero, idade ou outras variáveis, além de informar quando as possíveis diferenças encontradas podem ser atribuídas, ou não, ao acaso (Sisto, 2006).

Em suma, os estudos apresentados e os resultados obtidos neste trabalho associam-se às preocupações da comunidade científica e do Conselho Federal de Psicologia (CFP), visto que este se propõe a suplantam a precariedade dos instrumentos psicológicos utilizados no Brasil e a restringir as práticas errôneas, o que torna evidente

a necessidade de construção e de aperfeiçoamento dos instrumentos e práticas utilizadas na avaliação psicológica. Sem embargo, Sisto e cols. (1979) já destacavam a problematização das avaliações errôneas, realizadas com testes sem padronização brasileira ou sem análises psicométricas adequadas, assunto esse tratado mais recentemente por outros pesquisadores (Alchieri & Cruz, 2003; Cruz, Alchieri & Sarda Jr., 2002; Alves, 2006; Noronha, Freitas, Baldo, Barbini & Almeida, 2004; Noronha & Vendramini, 2003).

Propõe-se que novos estudos sejam realizados com as Matrizes Progressivas Avançadas de Raven, relacionados à DIF, constructo, unidimensionalidade, entre outros, visto que o teste RIn foi analisado sob tais aspectos, o que pôde contribuir para maior avaliação do instrumento, além de comprovar uma única dimensão. Sugere-se também que novas pesquisas sejam realizadas com os respectivos testes e com os demais instrumentos de avaliação psicológica, especificamente do constructo inteligência.

REFERÊNCIAS

- Aiken, L R (1996) *Tests Psicológicos y Evaluación*. Mexico: Praticce Hall Hispano.
- Alchieri, J. C. & Cruz, R. M. (2003). *Avaliação Psicológica: conceito, métodos e instrumentos*. São Paulo: Casa do Psicólogo.
- Alchieri, J. C., Noronha, A. P. P. & Primi, R. (2003). *Guia de referência: testes psicológicos comercializados no Brasil*. São Paulo: Casa do Psicólogo.
- Allik, J., Must, O., & Lynn, R. (1999). Sex differences in general intelligence among high school graduates: some results from Estonia. *Personality and Individual Differences, 26*, 1137–1141.
- Almeida, L.S. (1988). Teorias da inteligência. Porto: Edições Jornal de Psicologia.
- Almeida, L.S. (1996). Considerações em torno da medida da inteligência. Em L. Pasquali (Org.), *Teoria e Métodos de Medida em Ciências do Comportamento* (pp. 199-223). Brasília: INEP.
- Almeida, L.S. (2002). As aptidões na definição e avaliação da inteligência: o concurso da análise fatorial.
- Almeida, L.S. e Buela-Casal, G. (1997). Evaluación de la inteligencia general. Em G. Buela- Casal e J.C. Sierra (Orgs.), *Manual de evaluación psicológica: Fundamentos, técnicas y aplicaciones*. Madrid: Siglo XXI de España Editores.
- Alves, I. C. B. (1979). *O Teste Goodenough-Harris em uma população pré-escolar paulistana*. Dissertação de Mestrado - Instituto de Psicologia da Universidade de São Paulo, São Paulo.
- Alves, I. C. B. (1981). O Teste Goodenough-Harris em Pré-Escolares Paulistanos. *Boletim de Psicologia, XXXIII*, 80, 40-52.
- Alves, I. C. B. (1998). Variáveis significativas na avaliação da inteligência. *Psicol. esc.*

educ., 2 (2), 109-114.

Alves, I. C. B. e Duarte, J. L. M. (1993). Padronização Brasileira da Escala de Maturidade Mental Colúmbia. Em B. B. Burgemeister, L. H. Blum, e I. Lorge. *Escala de Maturidade Mental Colúmbia*. (3ª Ed.). Manual para Aplicação e Interpretação. São Paulo: Casa do Psicólogo.

Alves, I. C. B.; Angelini, A. L.; Custódio, E. M.; Duarte, W. F. (1988). O Teste das Matrizes Progressivas Coloridas de Raven: "Um estudo das variáveis idade, sexo e escolaridade", *Resumos da XVIII Reunião da Sociedade de Psicologia de Ribeirão Preto*, 18, 163.

Alves, I. C. B.; Colosio, R. e Ruivo, R. 1. (1992). O Teste R-I: Um estudo das variáveis idade, sexo e escolaridade. *Anais da 44ª Reunião Anual da Sociedade Brasileira para o Progresso da Ciência*, 44, 904.

Alves, I.C.B. (2006). *Novos Estudos Psicométricos do Teste D.70*. Avaliação Psicológica, 5(2), 251-253.

Anastasi, A. (1977). *Testes Psicológicos*. São Paulo: EPU.

Anastasi, A., & Urbina, S. (2000). *Testagem Psicológica*. Porto Alegre: Artes Médicas.

Angelini, A.L., Alves, I.C.B.; Custódio, E.M.; Duarte, W.F. & Duarte, J.L.M. (1999.) *Matrizes Progressivas Coloridas de Raven: Escala Especial*. Manual. São Paulo: CETEPP.

APA (1985). American Psychological Association, American Educational Research Association & National Council on Measurement in Education. *Standards for educational and psychological tests*. Washington, DC: APA.

Boris, D. A., MacLeod, C. M. & Forrin, B. (1993). *Eliminating the IQ-IT correlation by eliminating an experimental confound*. (pp. 475-500). *Intelligence*, 17.

Cahan, S.; & Cohen, N. (1989). Age versus schooling effects on intelligence

development. *Child Development*, 60, 1239-1249. Em F. F. Sisto (2006). *Teste de Raciocínio Inferencial (RIIn). Manual*. Vetor Editora Psicopedagógica Ltda.

Campos, F. (2002). *Matrizes Progressivas – Escala Avançada*. Rio de Janeiro: CEPA.

Candeias, A. M. & Almeida, L. S. (2000). Contributos da inteligência social ao estudo da

inteligência e da sobredotação. *Sobredotação*, 1 (2), 129-146.

Carpenter, A.P.; Just, M.A. & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, 97, 404-431.

Carroll, J.B. (1993). *Human cognitive abilities*. New York: Cambridge University Press.

Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Boston: Houghton Mifflin. Em Schelini, P.W. (2000). *Bateria multidimensional de inteligência infantil: Proposta de instrumento*. Tese de Doutorado. PUC-Campinas. Campinas-SP.

Colom R. & Garcíapz O. (2002). Sex differences in fluid intelligence among high school graduates. *Personality and Individual Differences*, 32, 445–451.

Colom, R., Juan-Espinosa, M., Abad, F. J., & Garcíapz L. F. (2000). Negligible sex differences in general intelligence. *Intelligence*, 28(1), 57–68.

Colom, R., Quiroga, M. A. & Juan-Espinosa, M. (1999). Are cognitive sex differences disappearing? Evidence from Spanish populations. *Personality and Individual Differences*, 27, 1189–1195.

Conselho Federal de Psicologia – CFP (2007). *Sistema de Avaliação de Testes Psicológicos*. Disponível em <http://www.pol.org.br/satepsi> . Acessado em 03/05/2007.

Contrera, M. S. Hattori, O. T. (2003) *Publicidade e Cia*. São Paulo: Thompson.

Cronbach, J.C. (1970). *Essentials of psychological testing*. (3ª ed.). New York: Harper

& Row Publs.

Cronbach, L. J. (1996). *Fundamentos da testagem Psicológica*. (5ª ed.). Porto Alegre: Artes Médicas.

Cruz, R.M. (2002a). Medidas psicológicas em psicologia do trabalho e das organizações. Em R.M. Cruz, J.C. Alchieri & J.J. Sarda Júnior (Orgs.) *Avaliação e medidas psicológicas: produção do conhecimento e da intervenção profissional*. (pp.173-179) São Paulo: Casa do Psicólogo.

Cruz, R.M. (2002b) O processo de conhecer em avaliação psicológica. Em R.M. Cruz, J.C. Alchieri & J.J. Sarda Júnior (Orgs.) *Avaliação e medidas psicológicas: produção do conhecimento e da intervenção profissional*. (pp.15-26) São Paulo: Casa do Psicólogo.

Cruz, R.M., Alchieri, J.C. & Sarda Júnior, J.J. (Orgs.) (2002). *Avaliação e medidas psicológicas: produção do conhecimento e da intervenção profissional*. São Paulo: Casa do Psicólogo.

Cunha, S. E. (1974). A Psicometria da Inteligência e a dimensão idade. *Arquivos Brasileiros de Psicologia Aplicada*, 26, 3, 100-110.

Dancey, C. P., & Reidy, J. (2006). *Estatística sem matemática para a Psicologia*. Porto Alegre, RS: Artmed.

Deary, I.J.; Whalley, L.J.; Lemmon, H.; Crawford, J.R.; & Starr, J.M. (2000). The stability of individual differences in mental ability from childhood to old age: Follow-up of the 1932 Scottish Mental Survey. *Intelligence*, 28, 49-55. Em F. F. Sisto (2006). *Teste de Raciocínio Inferencial (RIIn)*. Manual. Vetor Editora Psicopedagógica Ltda.

Flanagan, D. P., McGrew, K. S., & Ortiz, S. O. (2000). *The Weschsler Intelligence Scales and CHC Theory: A contemporary approach to interpretation*. Boston: Allyn & Bacon.

Flynn, J. R. (2003). *Movies about intelligence: the limitations of g*. Current Directions in Psychology, 12, 95-99.

Flynn, J. R. (2006). O efeito Flynn: repensando a inteligência e aquilo que a afeta. Em C. Flores-Mendoza, R. Colom, A. Andrés-Pueyo, R.J. Haier & S.A. Petrill. *Introdução à Psicologia das diferenças individuais* (p.387-411). Porto Alegre: ArtMed.

Frey, P. D. & Pinelli Jr., B. (1991). Visual discrimination and visuomotor integration among two classes of Brazilian children. *Perceptual and Motor Skills*, 72, 847-850.

Green, R.L.; Hoffman, L.J.; Morse, R.J.; Hayes, M.E.; & Morgan, R.F. (1964). The educational status of children in a district without public schools. Bureau of Educational Research, College of Education, Michigan State University, Lansing, Michigan. Em F. F. Sisto (2006). *Teste de Raciocínio Inferencial (RIIn). Manual*. Vetor Editora Psicopedagógica Ltda.

Halpern, D. (1997). Sex differences in intelligence. *American Psychologist*, 52(10), 1091–1102.

Hick, W. (1952). On the rate of gain information. *Quarterly Journal of Experimental Psychology*, 4, 11-46.

Horn, J. L. (1991). Measurement of Intellectual Capabilities: a review of theory. Em K. S. McGrew, J. K. Werder, & R. W. Woodcock. *Woodcock-Johnson technical manual* (pp. 197-232). Chicago: Riverside.

Hoskins, S. L., Newstead, S. E., & Denni, I. (1997). *Degree performance as a function of age, gender, prior qualifications and discipline studied*. Assessment & Evaluation in Higher Education, 22 (3), 317-328.

Howard, R.W. (1993). *On what intelligence is*. British Journal of Psychology, 84, 27-37.

Hunt, E. (1978). *The mechanics of verbal ability*. (pp.109-130). Psychological Review,

85.

Hyman, R. (1953). *Stimulus information as a determinant of reaction time*. (pp.188-196). *Journal of Experimental Psychology*, 45.

Jensen, A. (1998). *The g factor*. London: Praeger.

Jensen, A. R. & Munro, E. (1979). *Reaction time, movement time and intelligence*. (pp. 121-126) *Intelligence*, 3.

Jensen, A. R. (1993). *Why is reaction time correlated with psychometric g?* (pp. 53-56). *Current Directions in Psychological Science*, 2.

Kaufman, A.S. & Kaufman, N.L. (1993). *Kaufman Adolescent and Adult Intelligence Test (KAIT)*. Circle Pines: American Guidance Service.

Lemgruber, V. e Paine, P. A. (1981). A Adaptação Brasileira da Escala Verbal do WISC. *Arquivos Brasileiros de Psicologia*, 33, 1-2,32-56.

Lindley, R. H., Wilson, S. M., Smith, W. R. & Bathurst, K. (1995). *Reaction time (RT) and IQ: Shape of the task complexity function*. (pp.339-345). *Personality and Individual Differences*, 18.

Lynn, R. (1994). Sex differences in intelligence and brain size: a paradox resolved. *Personality and Individual Differences*, 17(2), 257-271.

Malta, D. C.; Goulart, M. E. A. & Costa, M. F. F. L. (1998). Estado nutricional e variáveis sócio-econômicas na repetência escolar: Um estudo prospectivo em crianças da primeira série em Belo Horizonte, Brasil. *Cadernos de Saúde Pública*, 14, 157-164.

Matarazzo, J. (1976). *Wechsler. Medida e Avaliação da Inteligência do Adulto*. São Paulo:

Manole.

Melo, M. G. (1998). *Estudo Comparativo da Escala Verbal do WISC Adaptada por Paine e Lemgruber com a Escala Traduzida por Poppovic*. Dissertação de Mestrado -

Instituto de Psicologia da Universidade de São Paulo, São Paulo.

Muñiz, J (Org.). (1996) *Psicometria*. Madrid: Universitas.

Neisser, U., Boodoo, G., Bouchard, T.J., Boykin, A.W., Brody, N., Ceci, S.J., Halpern, D.F., Loehlin, J.C., Prtloff, R., Stemberg, R.J. & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, 2, 77-101.

Nelson, L. (2000). Introduction to the special section on methods and implications of revising assessment instruments. Em Alves, I.C.B. (2006). *Novos Estudos Psicométricos do Teste D.70*. (p. 251-253) *Avaliação Psicológica*, 5(2).

Noronha, A. P. P. (1999). *Avaliação Psicológica Segundo Psicólogos: usos e problemas com ênfase nos testes*. Tese de Doutorado não publicada. Pontifícia Universidade Católica de Campinas, Campinas.

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

Noronha, A. P. P. Vendramini, C. M. M. (2003) *Parâmetros psicométricos: estudo comparativo entre testes de inteligência e de personalidade*. *Psicol. Reflex. Crit.*, Porto Alegre, v. 16, n. 1, 2003. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-79722003000100018&lng=pt&nrm=iso>. Acesso em: 31/05/2007.

Noronha, A. P. P., Freitas, J. V., Baldo, C. R., Barbin, P. F., & Almeida, M. C. (2004). Conhecimento de estudantes a respeito de conceitos de avaliação psicológica. *Psicologia em Estudo*, 9(2), 263-269.

Noronha, A. P. P., Oliveira, K. L., & Beraldo, F. N. M. (2003). Instrumentos psicológicos mais conhecidos e utilizados por estudantes e profissionais de Psicologia. *Psicologia Escolar e Educacional*, 7(1), 47-56.

Pain, S. (1995). *Psicometria Genética*. (I.C.B.Alves, trad.). São Paulo: Casa do Psicólogo. (Original Publicado em 1985).

Pasquali, L. (1996) *Teoria e métodos de medida em ciências do Comportamento*. Unb/Finep: Brasília.

Pasquali, L. (1997) *Psicometria: teoria e aplicações*. Unb/Finep: Brasília.

Pasquali, L. (2001). *Técnicas de exame psicológico (TEP) – manual: Fundamentos das técnicas psicológicas*. (v. 1). São Paulo: Casa do Psicólogo, Conselho Federal de Psicologia.

Pasquali, L. (2003). *Teoria dos testes na Psicologia e na Educação*. Petrópolis: Vozes.

Pasquali, L., Wechsler, S. & Bensusan E. (2002). Matrizes Progressivas do Raven Infantil: Um Estudo de Validação para o Brasil. *Revista Avaliação Psicológica*, 2, 95-110.

Primi, R.; Flores-Mendoza, C. E. & Castilho, A. V. (1998). Estudo exploratório das propriedades psicométricas do teste Raven-escala avançada informatizado. *Bol. Psicol*, 48(109), 67-75.

Raven, J. C.; Court, 1. H. e Raven, J. (1985). *Manual for Raven's Progressive Matrices and*

Vocabulary Scales. Section I. General Overvie. London: H. K. Lewis & Co.

Raven, J., Raven J.C. & Court, J.H. (1998). *Manual for Raven's Progressive Matrices and Vocabulary Scales: Section 4 Advanced Progressive Matrices Sets I & II*. Oxford: Oxford Psychologists Press.

Raven, J.; Raven, J. C. e Court, J. H. (1991). *Manual for Raven's Progressive Matrices and*

Vocabulary Seales. Section I. General Overview. Oxford: Oxford Psychologists Press.

Raven, J.C. (1962). *Matrizes Progressivas - Escala Avançada - Série I e II*. Rio de

Janeiro: CEPA .

Ribeiro, R.B. & Almeida, L.S. (2005) *Tempos de reação e inteligência: a robustez dos dados face à fragilidade da sua interpretação*. (pp.95-103). *Aval. psicol.*, 4(2).

Richardson, J. T. E. (1994). *Mature students in higher education: I. A literatura survey on approaches to studying*. *Studies in Higher Education*, 19 (3), 309-325.

Santos, D. N. ; Borges, A. P. ; Pereira, P. S. ; Chalhub, A. A. ; Happé, F. ; Silva, R. R. ; Assis, A. M. ; Blanton, R. ; Parraga, I. ; Reis, M. ; Almeida Filho, N. & Barreto, M. L. (2002). Epidemiologia do desenvolvimento cognitivo de escolares em Jequié, Bahia, Brasil: procedimentos de avaliação e resultados gerais. *Cadernos de Saúde Pública*, Rio de Janeiro, 18(3), 723-733.

Savage, R.D. (1968). *Psychometric assessment of the individual child*. Middlesex, England: Penguin Books, Inc.

Schelini, P.W. (2000). *Bateria multidimensional de inteligência infantil: Proposta de instrumento*. Tese de Doutorado. PUC-Campinas. Campinas-SP.

Silva, M. J. M., & Santos, A. A. A. (2004). A avaliação da compreensão em leitura e o desempenho acadêmico de universitários. *Psicologia em Estudo*, 9 (3), 459-457.

Silverstein, M.L. & Nelson, L.D. (2000). Clinical Research implications of revising psychological tests. Em Alves, I.C.B. (2006). *Novos Estudos Psicométricos do Teste D.70*. (pp.251-253) *Avaliação Psicológica*, 5(2).

Simonite, V. (1997). *Academic achievement of mature students on a modular degree course*. *Journal of Further and Higher Education*, 21 (2), 241-249.

Sisto, F. F. (1975). *El factor 'g' y el pensamiento operatorio formal (grupo INRC)*. Tesis Doctoral. Madrid: Universidad Complutense de Madrid.

Sisto, F. F. (2006). *Teste de Raciocínio Inferencial (RIn). Manual*. Vetor Editora Psicopedagógica Ltda.

- Sisto, F. F., Codenotti, N., Costa, C. A. J., & Nascimento, T. C. N. (1979). *Testes Psicológicos no Brasil: que medem realmente?* Educação e Sociedade, 1, 152-165.
- Sisto, F. F., Sbardelini, E. T. B., & Primi, R. (2001). *Contextos e questões da avaliação psicológica*. São Paulo: Casa do Psicólogo.
- Spearman, C. (1904). "General Intelligence", objectively determined and measured. *American Journal of Psychology*, 15, 201-293.
- Spearman, C. (1927). *Las Habilidades del Hombre: su naturaleza y medición*. Buenos Aires: Paidós.
- Sternberg, R. J. & Detterman, D.K. (1986). *What is intelligence? Contemporary viewpoints on its nature and definition*. Norwood, NJ: Ablex
- Sternberg, R. J. (1977). *Intelligence, information processing and analogical reasoning: The componential analysis of human abilities*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Sternberg, R. J. (1982). Reasoning, problem-solving, and intelligence. Em R.J. Sternberg (Ed.) *Handbook of Human Intelligence*. Cambridge, M.A.: Cambridge Univ. Press.
- Sternberg, R.J., & Kaufman, J.C. (1998). *Human abilities*. *Annual Review of Psychology*, 49, 479-502.
- Thurstone, L. L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press. Em P.W. Schelini, (2000). *Bateria multidimensional de inteligência infantil: Proposta de instrumento*. Tese de Doutorado. PUC-Campinas. Campinas-SP.
- Tosi, S. M. V. D. (2007). *Teste de Inteligência Geral Não-Verbal*. São Paulo: Casa do Psicólogo.
- Trueman, M. & Hartley, J. (1996). *A comparison between the time-management skills and academic performance of mature and traditional-entry university students*.

University of Keele, 199-215.

Tyler, L.E. (1973). *Testes e Medidas*. (3a ed.). Rio de Janeiro: Zahar.

Weiss, E. M.; Kemmler, G.; Deisenhammer, E. A.; Fleischhacker, W. W. & Delazer, M. (2003). Sex differences in cognitive functions. *Personality and Individual Differences*, 35 863–875.

ANEXOS

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO (1ª via)

TÍTULO DA PESQUISA: Matrizes Progressivas Avançadas de Raven e Teste de Raciocínio Inferencial:
Evidências de Validade

Eu,, idade, RG, residente à abaixo assinado, dou meu consentimento livre e esclarecido para participar como voluntário do projeto de pesquisa supra-citado, sob a responsabilidade do pesquisador Ivan Sant'Ana Rabelo, aluno 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 dois testes de inteligência;
- 2- Durante o estudo realizarei duas tarefas que avaliam inteligência geral;
- 3- A participação nesta pesquisa não oferece risco a minha saúde físico e/ou mental, mas poderá me causar constrangimento;
- 4 - Obtive todas as informações necessárias para poder decidir conscientemente sobre a minha participação na referida pesquisa ou ensaio;
- 5 - Estou livre para interromper a qualquer momento minha participação na pesquisa sem qualquer prejuízo;
- 6 - Meus dados pessoais serão mantidos em sigilo e os resultados gerais obtidos através da pesquisa serão utilizados apenas para alcançar os objetivos do trabalho, expostos acima, incluída sua publicação na literatura científica especializada;
- 7 - Poderei contatar o Comitê de Ética em Pesquisa da Universidade São Francisco para apresentar recursos ou reclamações em relação à pesquisa ou ensaio clínico através do telefone: 11 - 4534-8023;
- 8 - Não receberei qualquer remuneração por minha participação nesta pesquisa;
- 9 - Poderei entrar em contato com o responsável pelo estudo, Ivan Sant'Ana Rabelo sempre que julgar necessário pelo telefone (11) 45348040.
- 10 - Este Termo de Consentimento é feito em duas vias, sendo que uma permanecerá em meu poder e outra com o pesquisador responsável.

São Paulo, de de 200....

Assinatura do Voluntário:

.....

Pesquisador Responsável pelo estudo: Ivan Sant'Ana Rabelo

assinatura:.....

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO (2ª via)

TÍTULO DA PESQUISA: Matrizes Progressivas Avançadas de Raven e Teste de Raciocínio Inferencial:
Evidências de Validade

Eu,, idade
....., RG, residente à
abaixo assinado, dou meu consentimento livre e esclarecido para participar como voluntário do projeto
de pesquisa supra-citado, sob a responsabilidade do pesquisador Ivan Sant'Ana Rabelo, aluno 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 dois testes de inteligência;
- 2- Durante o estudo realizarei duas tarefas que avaliam inteligência geral;
- 3- A participação nesta pesquisa não oferece risco a minha saúde físico e/ou mental, mas poderá me causar constrangimento;
- 4 - Obtive todas as informações necessárias para poder decidir conscientemente sobre a minha participação na referida pesquisa ou ensaio;
- 5 - Estou livre para interromper a qualquer momento minha participação na pesquisa sem qualquer prejuízo;
- 6 - Meus dados pessoais serão mantidos em sigilo e os resultados gerais obtidos através da pesquisa serão utilizados apenas para alcançar os objetivos do trabalho, expostos acima, incluída sua publicação na literatura científica especializada;
- 7 - Poderei contatar o Comitê de Ética em Pesquisa da Universidade São Francisco para apresentar recursos ou reclamações em relação à pesquisa ou ensaio clínico através do telefone: 11 - 4534-8023;
- 8 - Não receberei qualquer remuneração por minha participação nesta pesquisa;
- 9 - Poderei entrar em contato com o responsável pelo estudo, Ivan Sant'Ana Rabelo sempre que julgar necessário pelo telefone (11) 45348040.
- 10 - Este Termo de Consentimento é feito em duas vias, sendo que uma permanecerá em meu poder e outra com o pesquisador responsável.

São Paulo, de de 200....

Assinatura do Voluntário:

.....

Pesquisador Responsável pelo estudo: Ivan Sant'Ana Rabelo

assinatura:.....