



**FUNCIONAMENTO DIFERENCIAL DO ITEM: ESTUDO DE VALIDADE PELA ESTRUTURA  
INTERNA DE UMA MEDIDA DE INTELIGÊNCIA**

**DIFFERENTIAL ITEM FUNCTIONING: VALIDITY STUDY BY THE INTERNAL STRUCTURE  
OF AN INTELLIGENCE MEASURE**

**FUNCIONAMIENTO DIFERENCIAL DE LOS ÍTEMS: ESTUDIO DE VALIDEZ POR LA  
ESTRUCTURA INTERNA DE UNA MEDIDA DE INTELIGÊNCIA**

*Ivan Santana Rabelo<sup>1</sup>*

*Andrea da Silva Mazzariolli<sup>2</sup>*

*José Raimundo Evangelista da Costa<sup>3</sup>*

*Raphaella Maria Ferreira Alves da Silva De Martini Barbosa<sup>4</sup>*

*Nelimar Ribeiro de Castro<sup>5</sup>*

**Resumo:** Este artigo objetiva apresentar estudos psicométricos com uma medida cognitiva não verbal, que tem como objetivo avaliar o raciocínio geral em adultos, a partir de uma medida de inteligência geral não-verbal. A amostra que compôs o estudo normativo foi de 1.230 pessoas, sendo 57,2% de mulheres. As idades foram de 15 a 64 anos, e escolaridades variaram do Ensino Fundamental ao Superior, coletados nos estados de Minas Gerais, São Paulo, Bahia, Rio de Janeiro e Santa Catarina, entre os anos de 2013 a 2018. Conforme resultados, observou-se que os itens demonstraram bom poder de discriminação entre pessoas com alto e baixo desempenho. A distribuição dos itens em termos de dificuldade foi bastante equilibrada. Em análises de variância (ANOVA) e de provas *post hoc* de Tukey realizadas, constatou-se uma diminuição das pontuações com o avanço da idade, assim como diferenças de resultados em relação a escolaridade dos respondentes. Para estudos do Funcionamento Diferencial dos Itens (DIF) observou-se diferenças em razão do sexo. Foram observados também pouco DIF para o critério escolarização. Os coeficientes de precisão por diferentes métodos aplicados apresentaram bons níveis de confiabilidade no desempenho em tarefas relacionadas ao

---

<sup>1</sup> Universidade de São Paulo (USP). E-mail: [professorivanrabelo@gmail.com](mailto:professorivanrabelo@gmail.com) ORCID <https://orcid.org/0000-0003-1344-4113>

<sup>2</sup> Universidade de São Paulo (USP). E-mail: [andrea.mazariolli@usp.br](mailto:andrea.mazariolli@usp.br) ORCID <https://orcid.org/0000-0003-1126-7221>

<sup>3</sup> Universidade Paulista (UNIP). E-mail: [raievan@yahoo.com.br](mailto:raievan@yahoo.com.br) ORCID <https://orcid.org/0000-0002-5119-4752>

<sup>4</sup> Universidade Paulista (UNIP). E-mail: [raphaellademartini@gmail.com](mailto:raphaellademartini@gmail.com) ORCID <https://orcid.org/0000-0003-2638-2089>

<sup>5</sup> Universidade Federal de Juiz de Fora (UFJF). E-mail: [nelimar.de.castro@gmail.com](mailto:nelimar.de.castro@gmail.com) ORCID <https://orcid.org/0000-0002-1548-5632>



raciocínio inferencial, em deduzir relações entre objetos apoiado por um conjunto de suposições e informações, pensamentos, crenças e conhecimentos.

**Palavras-chave:** Inteligência, Autoavaliação, Psicometria, Confiabilidade e validade, Testes de inteligência.

**Abstract:** This article aims to present psychometric studies with a non-verbal cognitive measure, which aims to assess general reasoning in adults, based on a measure of non-verbal general intelligence. The sample that comprised the normative study was 1,230 people, being 57.2% women. The ages ranged from 15 to 64 years, and schooling varied from Elementary to Higher Education, collected in the states of Minas Gerais, São Paulo, Bahia, Rio de Janeiro and Santa Catarina, between the years 2013 to 2018. According to results, we observed It was noted that the items showed good power of discrimination between people with high and low performance. The distribution of items in terms of difficulty was quite balanced. In analysis of variance (ANOVA) and post hoc Tukey tests performed, a decrease in scores was found with advancing age, as well as differences in results in relation to respondents' education. For studies on the Differential Functioning of Items (DIF), DIF was observed due to sex. Little DIF was also observed for the schooling criterion. The precision coefficients by different applied methods showed good levels of reliability in performance in tasks related to inferential reasoning, in deducing relationships between objects supported by a set of assumptions and information, thoughts, beliefs and knowledge.

**Keywords:** Intelligence, Self-Assessment, Psychometrics, Reliability and validity Intelligence tests.

**Resumen:** Este artículo tiene como objetivo presentar estudios psicométricos con una medida cognitiva no verbal, cuyo objetivo es evaluar el razonamiento general en adultos, basado en una medida de inteligencia general no verbal. La muestra que comprendió el estudio normativo fue de 1.230 personas, un 57,2% de mujeres. Las edades oscilaban entre los 15 y los 64 años, y la escolaridad variaba desde la educación primaria hasta la superior, recogida en los estados de Minas Gerais, São Paulo, Bahía, Río de Janeiro y Santa Catarina, entre los años 2013 a 2018. Según los resultados, Se observó que los ítems mostraron un buen poder de discriminación entre personas con alto y bajo rendimiento. La distribución de los ítems en términos de dificultad fue bastante equilibrada. En el análisis de varianza (ANOVA) y las pruebas post hoc de Tukey realizadas, se encontró una disminución en los puntajes con la edad avanzada, así como diferencias en los resultados en relación con la educación de los encuestados. Para los estudios sobre el funcionamiento diferencial de los ítems (DIF), se observó DIF debido al sexo. También se observó poca DIF para el criterio de escolarización. Los coeficientes de precisión por diferentes métodos aplicados mostraron buenos niveles de confiabilidad en el desempeño en tareas relacionadas con el razonamiento inferencial, al deducir las relaciones entre objetos respaldados por un conjunto de supuestos e información, pensamientos, creencias y conocimiento.

**Palabras clave:** Inteligencia, Autoevaluación, Psicometría, Fiabilidad y validez, Pruebas de inteligencia.

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).



De acordo com Sisto (2006), 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.

Almeida (1988) explica que o raciocínio consiste na aptidão intelectual por excelência, que integra os processos cognitivos mais elaborados do ponto de vista intelectual, mas que, porém, este 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 (Rabelo, 2008).

De acordo com Ribeiro e Almeida (2005), associados aos estudos com inteligência encontram-se pesquisas 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.

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, conforme apontado por Rabelo e Mazariolli (2019), 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*.

No que tange a influência da aprendizagem cultural, há 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).

Na pesquisa realizada por Deary, Whalley, Lemmon, Crawford e Starr (2000) com um grupo de pessoas com 80 anos foi verificado que as pessoas 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 *apud* Sisto, 2006).

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).

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* (1983) e a KAIT — *Kaufman Adolescent and Adult Intelligence Test* (1993), conforme apontado em Kaufman e Kaufman (1993).

Em pesquisas brasileiras realizadas com adultos (Alves, Colosio & Ruivo, 1992), foram constatadas diferenças significativas estatisticamente relacionadas a variável Sexo. 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. Não obstante, por se tratar de um estudo realizado com amostragem de candidatos em processo seletivo, não é possível garantir que seja representativa da população, principalmente porque a maioria dos cargos para o sexo feminino relacionavam-se a atividades de baixo prestígio social, o que, de certa forma, poderia explicar os resultados obtidos pelo sexo feminino, segundo os autores do estudo. Na pesquisa de padronização do D-70, também foram verificadas diferenças entre os sexos da amostra, tanto no grupo de escolares do Ensino Médio como no de adultos com escolaridade de Ensino



Médio e para adultos do Ensino Superior, o que levou à construção de Tabelas normativas para cada sexo (Alves, 1998).

Ao 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 (Rabelo, 2008).



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. 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, Raven e Court (1991).

Diante da importância da avaliação psicológica, como um processo integrado, que abrange várias etapas, uma delas podendo envolver a utilização de instrumentos psicológicos, Pasquali afirma que os testes psicológicos devem assegurar as características psicométricas básicas para sua utilização, permitindo qualificar uma avaliação como cientificamente respaldada (Pasquali, 2010).

Foi com esse objetivo de contribuir para a ampliação da diversidade e atualização de instrumentos para avaliar o referido construto, que o Teste Matrizes de Inteligência Geral (MIG) foi elaborado. Por se propor a ser um instrumento de avaliação psicológica, objetiva satisfazer as exigências científicas, sobretudo regulamentadas pela Psicologia, na medida em que pretende oferecer uma medida de inteligência consistente e válida empiricamente, expressa em um valor único e global, no modelo de teste de fator g, possibilitando contribuir para análises do raciocínio, utilizando-se de tarefas não verbais, mas detalhadamente elaboradas para possibilitar a análise de determinadas funções envolvidas no processamento de operações mentais, determinantes no processamento de cada tarefa.

Portanto, neste artigo serão descritos estudos psicométricos do Teste Matrizes de Inteligência Geral (MIG) com 28 itens. O teste objetiva mensurar, por meio do raciocínio inferencial, a inteligência fluida ou fator g. No instrumento observam-se séries de itens organizados em ordem de dificuldade, apresentando 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 sequência da figura matriz.

## **Método**

### **Participantes**

Os estudos para a realização de normas, evidências de validade e precisão do Teste Matrizes de Inteligência Geral (MIG). A amostra que compôs o estudo normativo foi de 1.230 pessoas, sendo 527 homens (42,8%) e 703 mulheres (57,2%). As idades variaram entre 15 e 64 anos, com média de 25,33 anos e desvio-padrão de 9,5 anos, porém 10 pessoas não deram essa informação.

No que se refere à escolaridade, esta variou de Ensino Fundamental incompleto a Ensino Superior completo, com maior concentração no Ensino Superior (62,6%) seguido do Ensino Médio (32,4%). Por outro lado, a menor concentração foi de Ensino Fundamental com 53 participantes (4,3%). A respeito da localização geográfica da amostra, os participantes eram provenientes de diversos estados do Brasil, com maior concentração de participantes foi do estado de Minas Gerais (47,3%), seguido do estado de São Paulo (33,4%), os demais estados com amostras coletadas foram Bahia (10,8%), Rio de Janeiro (5%) e Santa Catarina (3,4%).

### **Instrumento**

O teste Matrizes de Inteligência Geral (MIG (Rabelo et al., 2021) é um instrumento que objetiva mensurar desempenhos característicos dos testes de inteligência não verbais, a qual permite identificar os tipos de raciocínios e os processamentos envolvidos na sua execução, entre eles, identidade vertical e horizontal e com espelhamentos, analogia com soma e com subtração de elementos, rotação, quantificação, subtração, deslocamento-

posição, sequenciação, orientação espacial, relações gestálticas e relações complexas; além das classificações habituais do potencial intelectual. O teste é composto por figuras, incluindo formas geométricas, desenhos e outras figuras variadas, que apresentam itens em ordem crescente de dificuldade, a partir das quais o sujeito deve escolher a resposta correta e registrar na folha apropriada.

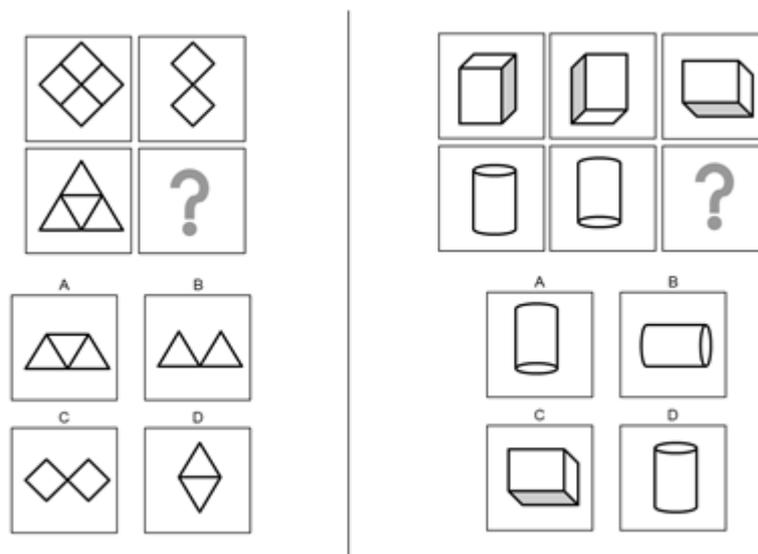


Figura 1. Exemplos de itens do teste

Fonte: Adaptado de Rabelo et al. (2021)

Entre os materiais que compõe o teste está presente um caderno de aplicação com 28 itens, nos quais há uma matriz, e o respondente deverá identificar em 4 opções de resposta aquela que mais adequadamente completa a matriz. O instrumento tem tempo de aplicação determinado de 6 minutos. Deve ser aplicado de forma individual ou coletiva. É um instrumento que objetiva avaliar a capacidade de educação de correlatos, definida por Spearman como um dos elementos do fator *g* ou inteligência geral.

O teste já apresenta outros estudos publicados em artigos e capítulos de livros (Rabelo, Anunciação, Cruz, Castro, 2020; Rabelo, Cruz, Castro, 2020; Rabelo & Mazariolli,



2019), sendo a presente pesquisa, mais um dos estudos de evidência de validade para a medida.

## **Procedimentos**

Este trabalho faz parte de uma pesquisa maior, atualmente ainda em andamento, que ocorre para o desenvolvimento de novos instrumentos de medida de funções psicológicas para utilização em contextos de avaliação neuropsicológica e psicológica. A pesquisa foi aprovada pelo Comitê de Ética em Pesquisa (CEP) da Universidade UNIFEBE, sob CAAE: 15956419.3.0000.5636 e número do parecer: 3.605.922.

Os instrumentos coletados nesta pesquisa foram aplicados por uma psicóloga com experiência em avaliação psicológica e utilização de testes, de forma coletiva, nas salas de aulas, durante o horário letivo. Responderam aos instrumentos apenas aqueles mediante assinatura do Termo de Consentimento Livre e Esclarecido.

Os estudos psicométricos com o teste foram realizados analisando a relação com a variável sexo, idade, escolaridade, sendo apresentadas análises de variância em relação às variáveis da amostra. Por fim, são apresentadas as estimativas de precisão, relacionadas à consistência interna, método das metades, Rasch e Teste-reteste.

## **Resultados**

Quanto aos resultados, a média da amostra ficou em 14,8 pontos na pontuação do teste MIG, praticamente no ponto central da pontuação possível, que é 14. A mediana ficou em 15 e a pontuação mais frequente foi de 14, em 9,3% da amostra. Nota-se que nenhum participante do grupo amostral atingiu o escore máximo no teste que é 28.

Com o objetivo de analisar as características dos itens, foram realizados três estudos: análise da dificuldade dos itens segundo pressupostos da Teoria Clássica dos Testes (TCT), análise da discriminação dos itens por meio do Índice D e pela diferença de média. A análise



de dos itens pela TRI foi descrita em tópico específico. A avaliação da dificuldade dos itens pela TCT pode ser obtida pela frequência de erros/acertos de cada item, assim, quanto maior a porcentagem de erros, mais difícil o item se apresenta, sendo o inverso verdadeiro, ou seja: quanto maior a porcentagem de acertos, mais fácil o item (Pasquali, 2010).

Foram verificadas as porcentagens de acerto/erro para cada item do Teste MIG, sendo verificado ao analisar os percentuais de erro dos itens que nove deles (itens 1, 2, 3, 4, 6, 7, 10, 11 e 13) obtiveram erros inferiores a 25%, ou seja, mais de 75% das pessoas os acertaram; 6 itens (8, 9, 12, 14, 15 e 17) ficaram entre 25 e 50% de erro, 6 entre 51 e 75% (itens 5, 18, 19, 20, 21 e 22) e 8 (itens 16, 23, 24, 25, 26, 27 e 28) apresentou mais de 75% de erros. Nota-se, deste modo, que a distribuição dos itens em termos de dificuldade foi bastante equilibrada.

A partir destas análises, verificou-se que cinco itens (17, 18, 19, 20 e 21) obtiveram índices de discriminação acima de 70, enquanto seis ficaram abaixo de 20 (itens 1, 3, 16, 26, 27 e 28), enquanto o restante ficou entre 20 e 70%. Esses resultados indicaram que os itens têm bom poder de discriminação entre pessoas com alto e baixo desempenho, no geral, exceto os itens, pelos seis itens, com índice D menor do que 205. Contudo, estes itens são justamente os itens nos extremos da escala de dificuldade.

O Índice D permite, de forma clara, avaliar a discriminação dos itens e compará-las identificando os mais ou menos discriminativos. Contudo, segundo Pasquali (2010), o uso do Teste *t* de Student permite uma medida mais exata desta discriminação, pois pode vir a indicar se de fato o grupo superior apresentou média de acerto significativamente diferente do grupo inferior. Assim, nos resultados do Teste *t* de Student para cada item com objetivo de verificar a diferença no desempenho entre os grupos com alto e baixo desempenho, os resultados indicaram que em todos os itens o grupo de desempenho superior apresentou resultado significativamente melhor que o grupo inferior, indicando que os itens discriminaram adequadamente pessoas com alto e baixo desempenho no Teste Matrizes de Inteligência Geral (MIG).

Em seguida foram realizados estudos de validade. A validade de um teste é determinada por meio de estudos estatísticos e indica o grau em que o teste mede,



efetivamente, o que pretende medir (Anunciação, Portugal, Rabelo, Cruz, Landeira-Fernandez, 2020). Portanto, foi realizada uma análise fatorial pelo método dos Componentes Principais, rotação Promax. O KMO de 0,895 e o Teste de Esfericidade de Bartlett de 12089,270 ( $df= 378$ ;  $p<0,001$ ) sugeriram a possibilidade de extração de fatores. Os *eigenvalues* indicaram a possibilidade de extração de sete fatores com variância explicada de 58,49%.

As análises fatoriais subsequentes indicaram como uma solução possível um modelo bifatorial explicando 35,21% da variância. O fator 1 englobou os itens de 16 a 28, e o fator 2 os de 1 a 15, de 17 a 22. As cargas fatoriais variaram de 0,78 a 0,12, sendo que apenas os itens 16 e 27 apresentaram valores menores do que 0,30.

Embora os dois fatores tenham sido extraídos com índices adequados, houve uma tendência à saturação de um fator sem a rotação, sugerindo possibilidades de organização em um único fator. Ao se analisar o conteúdo e estrutura dos itens de cada fator, pode-se observar que os itens de 1 a 15, que compuseram o fator 2, implicam em itens que obtiveram índices de dificuldade menores do que os itens de 16 a 28 que compuseram o fator 2, além disso, nenhuma diferença significativa entre os itens destes fatores justificam a existência de uma segunda dimensão. Deste modo, atribui-se esta segunda dimensão à interferência da dificuldade dos itens.

Posteriormente foram realizados estudos de validade que relacionam o teste MIG a variáveis externas. Algumas das variáveis que serviram de critério foram de natureza sociodemográfica, como idade, sexo e nível educacional. Assim sendo, inicialmente foram avaliadas possíveis diferenças entre homens e mulheres, por meio do teste *t* de Student. Porém, na amostra estudada, não foi observada diferença estatística significativa entre os desempenhos de homens e mulheres nos exercícios propostos pelo teste, apesar da média maior para o grupo de mulheres.

Em relação à variável Idade, investigou-se a correlação, por meio do coeficiente de Pearson, entre a idade e o desempenho no teste MIG. A correlação obtida foi moderada quase baixa, negativa e significativa estatisticamente ( $r=-0,33$ ;  $p<0,001$ ), indicando decréscimo no desempenho no teste MIG em razão do aumento da idade. Quando efetuadas correlações

controlando o efeito das variáveis Sexo e Nível Educacional, isoladas e em conjunto, observou-se pequeno, quase inexpressivo, decréscimo no índice de correlação pela influência da variável Nível Educacional, embora se mantendo a mesma classificação de magnitude moderada a baixa, negativa e significativa.

Para se definir tais faixas etárias, foi realizada uma análise da frequência das idades, seguido do levantamento do gráfico para inspeção visual dos resultados da pontuação no teste em relação a variável idade. Consta-se uma diminuição das pontuações com o avanço da idade no teste, contudo, observam-se picos de pontuação e maior variabilidade nas idades mais avançadas.

Com base nestas informações, e em distintas tentativas de agrupamentos por meio de análises de variância (ANOVA) e de provas *post hoc* de Tukey realizadas, chegou-se a uma divisão do grupo conforme apresentado na tabela 1 de Análise de Variância (ANOVA) e a prova *post hoc* de Tukey.

Tabela 1.

ANOVA e Prova de Tukey para as Faixas Etárias.

ANOVA			Subconjuntos para Alfa = 0,05		
<i>gl</i>	<i>F</i>	<i>p</i>	1	2	3
3	55,173	0,000	acima de 41 anos	10,4	
			de 29 a 40 anos		12,72
			de 18 a 28 anos		15,6
			Até 17 anos		16,4
			<i>p</i>	1,000	1,000
					0,538

Para realizar as análises por nível de educacional, investigou-se a associação entre as variáveis Escolaridade e a pontuação no teste MIG, por meio da correlação de Pearson. O índice de correlação encontrado foi considerado baixo, positivo e significativo estatisticamente, sugerindo que o aumento do Nível Educacional correspondeu ao aumento do desempenho no

teste MIG, porém com magnitude pouco expressiva na amostra investigada. O controle das variáveis sexo e idade indicaram pequeno decréscimo no índice de correlação, apenas em se tratando da variável idade.

As diferenças entre o desempenho do teste entre distintos grupos de escolaridade podem ser melhor observadas a seguir. Na tabela 2 são demonstradas a ANOVA e a prova *post hoc* de Tukey em relação às variáveis.

Tabela 2.

ANOVA e Prova de Tukey para Escolaridade.

ANOVA			Subconjuntos para Alfa = 0,05		
<i>gl</i>	<i>F</i>	<i>p</i>		1	2
2	52,294	0,000	Ensino Fundamental	7,94	
			Ensino Médio		14,81
			Ensino Superior		15,26
			<i>P</i>	1,000	0,751

Para os estudos com a TRI, inicialmente foram descritos os parâmetros de ajuste ao modelo Rasch, utilizados para as análises, seguidos da apresentação dos índices de dificuldades dos itens comparativamente aos índices de habilidade das pessoas. Posteriormente, foram descritos os resultados das análises de funcionamento diferencial dos itens (DIF) considerando algumas variáveis relevantes para o teste MIG, sendo elas sexo, faixa etária, nível educacional, estado de residência e variáveis relativas ao contexto do trânsito.

Deste modo, podem ser observados os parâmetros de ajuste ao modelo Rasch, destacando os *infits* e *oufits* de itens e pessoas. Em ambos os casos, espera-se que os valores obtidos estejam entre 0,30 e 1,50, sendo considerados itens ou pessoas em desajuste aqueles cujos *infits* e *oufits* estiverem fora deste intervalo. As análises com o teste MIG indicaram que nenhum dos itens apresentou *infits* fora deste intervalo, indicando bom ajuste,

mas 12 (42,86%) itens apresentaram *outfits* acima de 1,50. No caso das pessoas, nenhum *infits* esteve abaixo de 0,30, mas 175 (14,23%) esteve acima do intervalo adequado, enquanto para os *outfits* 208 (16,91%) estiveram abaixo de 0,30, e 268 (21,79%) estiveram acima de 1,50. Além disso, as médias dos *infits* e *outfits* estiveram próximas ao padrão esperado, que é 1,00. Esses resultados na tabela 3 sugeriram um bom ajuste dos itens ao modelo utilizado.

Tabela 3.

Parâmetros de ajuste ao modelo Rasch.

Parâmetros	Itens			Pessoas		
	<i>Infit</i>	<i>Outfit</i>	Erro	<i>Infit</i>	<i>Outfit</i>	Erro
Média	0,97	1,26	0,58	0,97	1,26	0,58
DP	0,50	1,77	0,07	0,50	1,77	0,09
Máximo	2,95	9,90	1,16	2,95	3,85	1,67
Mínimo	0,36	0,09	0,52	0,36	0,18	0,52
< 0,30	0	0		0	208(16,9%)	
> 2,0	0	12(42,9%)		175(14,2%)	268(21,8%)	

Quanto aos índices de dificuldade dos itens e de habilidade das pessoas, nota-se que a média dos itens foi zero (DP= 2,47), variando entre -3,69 a 5,48. Os índices de dificuldade dos itens, com o erro padrão e os *infits* e *outfits* de cada item podem ser encontrados na tabela 4.

Tabela 4.

Índice de dificuldade dos itens, erro padrão, *infit* e *outfit*.

Item	Índice de Dificuldade	SE	<i>Infit</i>	<i>Outfit</i>
Item 27	5,48	0,23	1,00	0,77
Item 28	4,69	0,17	0,90	1,84
Item 26	3,62	0,12	0,92	0,89
Item 16	3,37	0,11	1,49	5,63

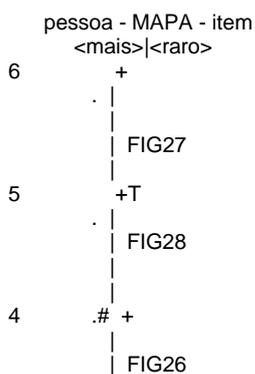


Item 25	2,95	0,10	0,81	1,03
Item 23	2,37	0,09	0,83	1,00
Item 24	2,14	0,09	0,76	0,59
Item 22	1,59	0,08	0,71	0,89
Item 21	1,04	0,07	0,71	0,66
Item 20	0,48	0,07	0,69	0,60
Item 19	0,41	0,07	0,68	0,64
Item 18	0,19	0,07	0,76	0,76
Item 05	0,14	0,07	1,57	2,28
Item 17	-0,14	0,07	0,81	0,78
Item 09	-0,33	0,07	1,29	1,64
Item 15	-0,49	0,07	1,07	1,39
Item 08	-1,20	0,07	1,24	1,69
Item 12	-1,31	0,08	0,96	1,12
Item 14	-1,45	0,08	0,89	0,92
Item 07	-1,66	0,08	1,07	1,51
Item 11	-1,92	0,08	0,93	1,76
Item 13	-2,19	0,09	0,76	0,59
Item 10	-2,32	0,09	1,01	2,22
Item 02	-2,61	0,10	1,11	1,81
Item 06	-2,65	0,10	1,07	1,68
Item 04	-3,17	0,11	1,02	2,19
Item 03	-3,34	0,12	1,06	1,49
Item 01	-3,69	0,13	1,05	3,32
Média	0,00	0,10	0,97	1,49
Desvio padrão	2,47	0,04	0,22	1,03



No caso das pessoas, os índices de habilidade variaram entre -5,01 e 5,81, com média de 0,90 (DP= 1,71). Assim, pode-se verificar que a média das habilidades das pessoas foi mais elevada do que dos índices de dificuldade dos itens, embora com pequena diferença. Além disso, ocorreram casos de pessoas com nível de habilidade maior e menor que a dificuldade dos itens do teste MIG.

Ao observar o mapa de distribuição de itens e pessoas pode-se verificar essa informação e perceber que houve uma boa distribuição dos itens e pessoas ao longo de toda a escala. No caso dos itens, houve uma distribuição uniforme, indicando itens com níveis de dificuldade em toda a escala, enquanto as pessoas se concentraram em valores mais medianos de habilidade, o que era o desejado, sugerindo que o teste não se mostrou nem fácil nem difícil para a maioria das pessoas, sugerindo boa discriminação.







atribuída ao acaso ou não. Um valor de  $t$  maior que 2,4 é considerado, segundo Draba (1977), um ponto bom para a significância estatística, quando se está analisando mais de 20 itens, como é o caso específico do teste MIG (total de 28 itens).

Assim, o estudo de DIF em razão do sexo apontou viés de resposta em apenas dois itens, os de número 2 e 3, sendo ambos favoreceram os homens. Já em razão do Nível Educacional foram observados cinco itens com DIF em oito interações de um total de 84, indicando pouco DIF para o critério escolarização. Os itens 3, 4 e 16 favoreceram as pessoas do Ensino Médio, em detrimento daquelas que apresentavam apenas o Ensino Fundamental; o item 16 favoreceu aos que apresentavam o Ensino Superior e Médio em detrimento dos que apresentavam Ensino Fundamental; o item 26 favorecer aos que apresentavam o Ensino Superior em detrimento daqueles com Ensino Médio e Fundamental, enquanto o item 17 favoreceu ao Ensino Fundamental, em detrimento do Ensino Médio e Superior. Este foi o único item que favoreceu níveis mais baixos de escolarização.

Por fim, para a realização dos estudos de fidedignidade, foram realizados alguns procedimentos, visando estimar a precisão das medidas a partir de diferentes fontes. Assim, serão apresentados por prova, os resultados obtidos a partir da verificação da precisão por consistência interna (Alfa de Cronbach e pelo método de Rasch), método das metades (*split half*) e teste-reteste. O estudo de teste-reteste foi realizado com 92 pessoas com idade entre 19 e 53 ( $M=26,47$ ;  $DP= 8,90$ ) anos, sendo que 32 (34,78%) estavam entre 19 e 22 anos. Do total, 63 (68,48%) eram mulheres e 29 (31,52%), homens. Todos cursavam o Ensino Superior.

Foram realizados procedimentos de verificação da consistência interna pelos métodos Alfa de Cronbach e Rasch, método das metades, considerando itens pares e ímpares, e teste-reteste. As análises foram realizadas com a amostra normativa, com 1.230 sujeitos. Os coeficientes de Precisão por Diferentes Métodos aplicados apresentaram, respectivamente, Alfa de Cronbach de 0,87; no Modelo de Rasch 0,86; no método das Metades 0,92; e no Teste-reteste 0,88.

Tais resultados evidenciam que os coeficientes ficaram entre bons e excelentes, variando entre 0,87 a 0,92. Isso indica que o teste faz a avaliação com rebaixados coeficientes

### **Discussão**

Pode-se verificar que, na amostra estudada, o desempenho no teste de inteligência está significativamente associado à idade, com uma tendência para pontuações menores à medida que as idades aumentam. Observa-se que até os 28 anos as faixas etárias não se diferenciam quanto ao desempenho, enquanto entre as demais, houve diferenças significativas.

A respeito da variável idade, segundo Matarazzo (1976), dados de avaliação da inteligência levam à impressão de um declínio com a idade, que é apenas aparente, visto que as pesquisas longitudinais evidenciam que a inteligência pode continuar a aumentar levemente até os 50 anos e, então, começar a mostrar um pequeno declínio acima dessa idade. Por sua vez 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.

Apesar de, no passado, autores discutiram se a inteligência parava de aumentar por volta dos 16 ou 18 anos, em adultos, que refere-se a amostra desta pesquisa, a idade é uma variável mais complexa, porque se espera uma estagnação da inteligência em função da idade (Alves, 1998). Sugerem-se outras pesquisas com idades avançadas para observar se este declínio nos resultados do teste se mostra mais acentuado com a senilidade, conforme ocorre em estudos com testes de inteligência, como pesquisas realizadas por outros autores (Cunha, 1974; Raven & cols., 1985; Alves, 1998; Tosi, 2008; Rabelo & cols., 2011; Rabelo, 2013).



Em relação à influência da variável escolaridade dos respondentes, os dados da tabela 2 informaram que os três níveis de escolaridade representados na amostra deste estudo, tiveram desempenhos significativamente diferentes, observando um aumento nas pontuações no teste com o aumento da escolaridade. Porém, a partir do agrupamento feito, observa-se que apenas os participantes com Ensino Fundamental diferenciaram-se estatisticamente de maneira significativa dos outros dois agrupamentos de escolaridade. Ainda assim, é possível afirmar que, na amostra estudada, há uma influência da escolaridade no desempenho das pontuações no teste.

Os estudos com modelo Rasch com o Teste Matrizes de Inteligência Geral (MIG) indicaram que nenhum dos itens apresentou *infits* fora deste intervalo, indicando bom ajuste, mas 12 (42,86%) itens apresentaram *outfits* acima de 1,50. No caso das pessoas, nenhum *infits* esteve abaixo de 0,30, mas 175(14,23%) esteve acima do intervalo adequado, enquanto para os *outfits* 208 (16,91%) estiveram abaixo de 0,30, e 268 (21,79%) estiveram acima de 1,50. Tais resultados sugeriram um bom ajuste dos itens ao modelo utilizado.

O estudo de DIF em razão do sexo apontou viés de resposta em apenas dois itens, favorecendo os homens. Em razão do nível educacional foram observados cinco itens com DIF em oito interações de um total de 84, indicando pouco DIF para o critério escolarização.

Portanto, na amostra estudada, não foi observada grandes diferenças estatísticas significativas entre os desempenhos de homens e mulheres nos exercícios propostos pelo teste, apesar da média maior para o grupo de mulheres. 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, Must, & Lynn, 1999; Lynn, 1994): 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 (1993, 1998), Colom, Juan-Espinosa, Abad e Garciápz (2000), Colom e Garciápz (2002), 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.

Ao final, os estudos de confiabilidade com o teste indicaram que o teste faz a mensuração com rebaixados coeficientes de erro de medida, no que diz respeito à consistência interna, assim como à estabilidade temporal.

### **Conclusão**

Em suma, destaca-se que os estudos psicométricos 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 contribuir para a qualidade 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 (Anunciação et al., 2020).

Entre os resultados, observou-se que os itens demonstraram bom poder de discriminação entre pessoas com alto e baixo desempenho, exceto seis dos itens. A distribuição dos itens em termos de dificuldade foi bastante equilibrada. Foi realizada uma análise fatorial pelo método dos Componentes Principais, rotação Promax. Tais análises fatoriais subsequentes indicaram como uma solução possível um modelo bifatorial explicando 35,21% da variância. Em análises de variância (ANOVA) e de provas post hoc de Tukey realizadas, assim como, estudos correlacionais diversos, constatou-se uma diminuição das pontuações com o avanço da idade, assim como diferenças de resultados em relação a escolaridade dos respondentes.

Especificamente nos estudos do Funcionamento Diferencial dos Itens (DIF) foram escolhidas as variáveis sexo e nível educacional. Observou-se DIF em razão do sexo com viés de resposta em apenas dois itens, sendo que ambos favoreceram os homens. Em razão do Nível Educacional foram observados cinco itens com DIF em oito interações de um total de 84, indicando pouco DIF para o critério escolarização. Os coeficientes de precisão por diferentes métodos aplicados apresentaram, respectivamente, Alfa de Cronbach de 0,87; no Modelo de Rasch 0,86; no método das Metades 0,92; e no Teste-reteste 0,88.

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 que vem sendo tratado por distintos pesquisadores (Alves, 2006; Hutz, 2009; Primi, Muniz & Nunes, 2009; Pasquali, 2010; Rabelo, 2008; Sisto, 2006).

Destaca-se, por fim, que a avaliação psicológica em qualquer contexto demandará do examinador tanto experiência clínica e experiência na utilização de testes quanto conhecimentos acerca do funcionamento da cognição normal e patológica. Esses conhecimentos servirão de parâmetro para comparar os resultados obtidos nas tarefas ou testes executados pelo respondente a fim de identificar sua capacidade cognitiva. Tal



identificação pode servir a diversos fins: avaliação de pessoas que desejam obter ou manter autorização para dirigir; investigação de fatores vinculados a fenômenos psicológicos ligados ao contexto do trabalho ou trânsito (fadiga, uso de álcool e drogas, comportamento seguro), assim como, avaliações e psicodiagnósticos no contexto clínico, pericial, entre outros.

### Referências

- Allik, J., Must, O., & Lynn, R. (1999). Sex differences in general intelligence among high school graduates. *Personality and Individual Differences*, 26, 1137–1141.
- Almeida, L. S. (1988). *Teorias da inteligência*. Edições Jornal de Psicologia.
- Almeida, L. S., & Buela-Casal, G. (1997). Evaluación de la inteligencia general. In: G. Buela-Casal, & J. C. Sierra (Orgs.), *Manual de evaluación psicológica: Fundamentos, técnicas y aplicaciones*. Madrid: Siglo XXI de España Editores.
- Alves, I. C. B. (1998). Variáveis significativas na avaliação da inteligência. *Psicol. esc. educ.*, 2 (2), 109-114.
- Alves, I. C. B., & Duarte, J. L. M. (1993). Padronização Brasileira da Escala de Maturidade Mental Colúmbia. In: B. B. Burgemeister, L. H. Blum, & I. Lorge, *Escala de Maturidade Mental Colúmbia*. Manual para Aplicação e Interpretação. Casa do Psicólogo.
- Alves, I. C. B.; Colosio, R. & Ruivo, R. I. (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*. EPU.
- Anastasi, A., & Urbina, S. (2000). *Testagem Psicológica*. Artes Médicas.
- Anuniação, L., de Almeida Portugal, A. C., Rabelo, I., Cruz, R., & Landeira-Fernandez, J. (2020). Propriedades psicométricas de instrumento de Memória Visual de Curto Prazo (MEMORE). *Neuropsicologia Latinoamericana*, 12(2), 44-58.



Colom, R., & Garcia Lopez, 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., & Garciápz L. F. (2000). Negligible sex differences in general intelligence. *Intelligence*, 28(1), 57–68.

Cunha, J. A. (2003). *Psicodiagnóstico-V*. (5. ed.). Artmed.

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

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.

Draba, R. E. (1977). *The identification and interpretation of item Bias*. The University of Chicago.

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. In: C. Flores-Mendez et al. (Eds.). *Introdução à Psicologia das diferenças individuais* (pp. 387-411). ArtMed.

Hutz, C. S. (org.) (2009) *Avanços e Polêmicas em Avaliação Psicológica*. Casa do Psicólogo.

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

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

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

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

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

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



- Pasquali, L. (org.) (2010). *Instrumentação Psicológica: fundamentos e práticas*. Artmed.
- Primi, R., Muniz, M., & Nunes, C. H. S. S. (2009). Definições Contemporâneas de Validade dos Testes Psicológicos. Em C. S. Hutz (Org.), *Avanços e Polêmicas em Avaliação Psicológica* (pp. 243-265). Casa do Psicólogo.
- Rabelo, I.S. (2008). *Matrizes progressivas avançadas de Raven e teste de raciocínio inferencial: evidências de validade*. (Dissertação de Mestrado). Itatiba, USF.
- Rabelo, I. S. (2013). *Teste Memória Visual de Trânsito - MVT: manual*. Casa do Psicólogo.
- Rabelo, I. S., Portugal, A.C.A., Cruz, R. M., & Castro, N. R.. Landeira-Fernandez, J., Anunciação, L., (2021). *Matrizes de Inteligência Geral (MIG)* (1a ed.). São Paulo: Editora Nila Press.
- Rabelo, I. S., Anunciação, L., Cruz, R.M., & Castro, N.R. (2020). *Teste de Memória de Reconhecimento - MEMORE. Manual Técnico*. Nila Press.
- Rabelo, I. S., Cruz, R.M., & Castro, N.R. (2020). *Bateria Rotas de Atenção. Manual Técnico*. Nila Press.
- Rabelo, I. S., Mazariolli, A. S. (2019). Avaliação Cognitiva de Policiais Militares e Universitários em Medidas Padronizadas de Memória, Atenção e Inteligência. *Meta: Avaliação*, 11(32), 468-494.
- Rabelo, I. S., Pacanaro, S. V., Leme, I. F. S., Ambiel, R. A. M. & Alves, G. A. S. (2011). *Teste Não Verbal de Inteligência - BETA-III*. Casa do Psicólogo.
- Rabelo, I. S., Portugal, A. C. A., Cruz, R. M., Castro, N. R., Landeira-Fernandez, J., & Anunciação, L. (2021). *Matrizes de Inteligência Geral Não Verbal - MIG. Manual Técnico*. Nila Press.
- Raven, J. C., Court, J. H., & Raven, J. (1985). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Section I. General Overvie*. H. K. Lewis & Co.
- Raven, J.; Raven, J. C., & Court, J. H. (1991). *Manual for Raven's Progressive Matrices and Vocabulary Seales*. General Overview. Oxford Psychologists Press.
- Ribeiro, R. B., & Almeida, L. S. (2005) Tempos de reação e inteligência: a robustez dos dados face à fragilidade da sua interpretação. *Aval. psicol.*, 4(2), 95-103.



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

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

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

Sternberg, R. J. (1982). Reasoning, problem-solving, and intelligence. In: R. J. Sternberg (Ed.) *Handbook of Human Intelligence*. Cambridge Univ. Press.

Tosi, S. M. V. D. (2008). *Teste de Inteligência Geral Não-Verbal*. Casa do Psicólogo.