

Gleiber Couto

**Desenvolvimento de Escalas com Interpretação Referenciada nos Itens para a
Bateria de Provas de Raciocínio**

**Itatiba
2007**

Universidade São Francisco

Programa de Pós-Graduação *Stricto Sensu* em Psicologia



**Desenvolvimento de Escalas com Interpretação Referenciada nos Itens para a
Bateria de Provas de Raciocínio**

Tese apresentada à Banca Examinadora do Programa de Pós Graduação *Stricto-Sensu* em Psicologia da Universidade São Francisco como parte dos requisitos para obtenção do título de Doutor em Psicologia na área de Avaliação Psicológica.

Orientador: Prof. Dr. Ricardo Primi

Gleiber Couto

Itatiba

2007

157.93 C91d	<p>Couto, Gleiber. Desenvolvimento de escalas com interpretação referenciada nos itens para a bateria de provas de raciocínio / Gleiber Couto. -- Itatiba, 2007. 174 p.</p> <p>Tese (doutorado) – Programa de Pós-Graduação em <i>Stricto Sensu</i> em Psicologia da Universidade São Francisco. Orientação de: Ricardo Primi</p> <p>1. Avaliação psicológica. 2. Psicologia cognitiva. 3. Psicometria. 4. Inteligência. 5. Tri. I. Primi, Ricardo. II. Título.</p>
----------------	---

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

Programa de Pós-Graduação “*Stricto Sensu*” em Psicologia

Desenvolvimento de Escalas com Interpretação Referenciada
nos Itens para a Bateria de Provas de Raciocínio

Autor(a): Gleiber Couto

Orientador(a): Prof. Dr. Ricardo Primi

Este exemplar corresponde à redação final da Tese de Doutorado
defendida por Gleiber Couto e aprovada pela comissão examinadora.

Data: 18 / 12 / 2007

COMISSÃO EXAMINADORA

Prof. Dr. Ricardo Primi (Presidente)

Prof^a. Dr^a. Alessandra Gotuzo Seabra Capovilla

Prof. Dr. Carlos Henrique Sancineto da Silva Nunes

Prof. Dr. Elizeu Coutinho de Macedo

Prof^a. Dr^a. Solange Muglia Wechsler

Itatiba

2007

Agradecimentos

Gostaria de agradecer em primeiro lugar aos meus pais que sempre me ensinaram a valorizar e buscar o conhecimento acima de qualquer outra aspiração; estiveram sempre ao meu lado dispostos a sacrificar os próprios sonhos para que eu pudesse realizar os meus. Sem vocês nada disso teria acontecido.

A todos os membros da minha família que torceram por mim, e de alguma forma colaboraram com este trabalho. Especialmente ao tio Silvio e tia Tina que mesmo à distância me acompanharam e torceram pelas minhas realizações e durante toda a minha vida sempre pude contar com sua ajuda.

Agradeço a todos os professores do programa de pós-graduação em Psicologia da Universidade São Francisco pelas valiosas lições sobre avaliação psicológica. Especialmente ao meu orientador Prof. Dr. Ricardo Primi que me recebeu com bastante disposição e aceitou a difícil tarefa de me orientar. Com o tempo se tornou também um bom amigo.

Aos amigos Sanyo Drummond Pires, Francisco Marini Junior, Renato Camargos Viana e Helton Rocha Campos, que são meus companheiros na trajetória profissional e sempre fonte de inspiração. Também agradeço aos amigos de infância, e aos amigos que fiz ao longo da vida, e todos aqueles que sentiram minha falta durante as nossas reuniões, mas sempre brindaram meu nome.

Aos colegas do LabAPE, Monalisa, Fabiano, Bia, Tatiana, Tchê & Maiana, Rodolfo, entre outros, com quem dividi muitas horas de aprendizado sobre psicometria e metodologia científica além de muitos momentos de trabalho bem humorado.

Resumo

Couto, G. (2007) *Desenvolvimento de Escalas com Interpretação Referenciada nos Itens para a Bateria de Provas de Raciocínio*. Tese de Doutorado, Programa de Pós-Graduação em Psicologia, Universidade São Francisco, Itatiba, SP.

Na teoria clássica dos testes uma alternativa frequentemente usada como medida das dimensões psicológicas é a referência às normas. Esse procedimento consiste em combinar um conjunto específico de tarefas de conteúdo psicológico definido, as quais são chamadas de itens, e comparar as respostas de um grupo determinado de pessoas nesses itens. A posição do sujeito no construto psicológico é inferida por meio de uma comparação sujeito-grupo. As medidas construídas com referência nesse modelo apresentam algumas desvantagens. Uma delas é que os resultados da medida, dificilmente permitem uma compreensão das capacidades do sujeito em termos de processos cognitivos. A Teoria de Resposta ao Item é um modelo que possibilita melhora nas escalas de medida. Isso porque ao contrário do tradicionalmente usado em psicometria clássica, permite que às respostas dos sujeitos sejam comparadas com a complexidade dos itens usados na sua avaliação. O presente trabalho teve por objetivo desenvolver escalas com interpretações referenciadas nos itens, para as provas de Raciocínio Abstrato (RA) e Raciocínio Numérico (RN) da Bateria de Provas de Raciocínio (BPR-5), usando os procedimentos metodológicos desenvolvidos pelo TRI, e de análise cognitiva dos itens. O resultado final foi a apresentação de uma escala descritiva dos tipos de operações que estão presentes em cada item em uma faixa de dificuldade. Suposições plausíveis podem ser feitas relacionando tais operações às atividades cognitivas por traz das respostas corretas em cada item. Portanto, conseqüentemente, suposições sobre as habilidades que cada sujeito, que acerta itens em uma faixa de dificuldade, é capaz de manifestar. Espera-se que os resultados desse trabalho colaborem com o desenvolvimento do método aplicado e, por conseqüência, estimulem a construção de instrumentos baseados nesses princípios, ao mesmo tempo em que desencorajem a construção artesanal e intuitiva de testes psicológicos.

Palavras-chave: psicometria; tri; inteligência; avaliação psicológica; psicologia cognitiva.

Abstract

Couto, G. (2007) *Development of rating scales with referenced interpretation of items for test batteries of reasoning*. Doctor's theses for post graduation in psychology, Universidade São Francisco, Itatiba, SP, Brazil

In classical test theory, measures of psychological dimensions used are often relative to the normal. In this procedure a set of items, which are tasks specific for the defined psychological content, is used and answers to these items, from members of a selected group, are compared. The position of a subject in this psychological construct is determined by comparing the subject to the group. Measures built, based on this model have certain disadvantages. One is that the result from such measure hardly ever allows the understanding of the subject's capacities, in the sense of cognitive processes. The Item Response Theory is a model that allows for better measuring scales. That is because this model, contrary to the classical model used in classical psychometrics, allows analysis of the subject's responses to be compared with the complexity of the items used in the evaluation. Goal of the present study, was to develop scales with referenced interpretation for Abstract Reasoning (RA) tests and Numeric Reasoning (NR) tests of the Bateria de Provas de Raciocínio (BPR-5), using methodological procedures developed by TRI, and cognitive analyses of the items. The end-result was a descriptive scale of the kind of operations that are present in each item of a of difficulty level. Relationships between the operations and the cognitive activities behind correct answers for each item can be determined. Thus statements can be made about the abilities of each subject that responds correctly to an item of a certain difficulty level. It is expected that the results of this work contribute to the development of the applied method and as a result promote the construction of instruments based on these principles and demotivate the construction of artisan laborious and intuitive psychological tests.

Key-Words: psycometry; tri; intelligency; psychological assessment; cognitive psychological.

Sumário

Índice de Tabelas	16
Índice de Figuras	17
Introdução	18
Apresentação	18
Capítulo I – Medida em Psicologia	21
Considerações sobre a Medida em Psicologia.....	21
Atribuição de Significado aos Escores de instrumentos de medida	30
Avanços da Psicometria e da Psicologia Cognitiva na Construção de Escalas.....	39
Capítulo II - Inteligência	49
Desenvolvimento das concepções sobre a Inteligência.....	49
Contribuições das Ciências Cognitivas para compreensão da Inteligência.....	71
Contribuições da Psicologia Cognitiva na compreensão do Raciocínio.	83
Capítulo III – A Bateria de Provas de Raciocínio	92
Prova de Raciocínio Abstrato	103
Prova de Raciocínio Numérico.....	116
Método.....	125
Fonte	125
Material.....	125
Procedimento	127
Resultados e Discussão.....	129
Prova de Raciocínio Abstrato (RA).....	129
Prova de Raciocínio Numérico (RN)	140
Conclusão das Escalas	151
Considerações Finais	163
Referências	168
Anexo I	177
Anexo II.....	Erro! Indicador não definido.

Índice de Tabelas

Tabela 1. Equalização das três formas do RA.....	131
Tabela 2. Equalização das três formas do RA (continuação).....	132
Tabela 3. Correlações entre os fatores de complexidade e a Dificuldade dos Itens.....	134
Tabela 4. Correlações entre os fatores de complexidade e a Dificuldade dos Itens.....	134
Tabela 5. Lista de Aglomeração.....	136
Tabela 6. Sumário do Modelo de Regressão.....	138
Tabela 7. Coeficientes da Regressão Logística.....	139
Tabela 8. Equalização das três formas do RN.....	142
Tabela 9. Equalização das três formas do RN (continuação).....	143
Tabela 10. Correlações entre complexidade e Dificuldade dos Itens da prova RN.....	145
Tabela 11. Correlações entre complexidade e Dificuldade dos Itens da prova RN.....	146
Tabela 12. Lista de Aglomeração.....	147
Tabela 13. Sumário do Modelo de Regressão para prova RN.....	149
Tabela 14. Coeficientes da Regressão Logística para a prova RN.....	150

Índice de Figuras

Figura 1. Curva Característica do Item.....	35
Figura 2. Teoria CHC das Habilidades Cognitivas.	56
Figura 3. Identidade. Adaptado de Jacobs e Vandeventer (1972).	104
Figura 4. Forma. Adaptado de Jacobs e Vandeventer (1972).	104
Figura 5. Sombreamento. Adaptado de Jacobs e Vandeventer (1972).	104
Figura 6. Tamanho. Adaptado de Jacobs e Vandeventer (1972).	105
Figura 7. Movimento no plano. Adaptado de Jacobs e Vandeventer (1972).	105
Figura 8. Séries. Adaptado de Jacobs e Vandeventer (1972).	105
Figura 9. Adição de Elementos. Adaptado de Jacobs e Vandeventer (1972).	105
Figura 10. Adição usando combinação. Adaptado de Jacobs e Vandeventer (1972).	105
Figura 11. Figuras de um conjunto. Adaptado de Jacobs e Vandeventer (1972).	106
Figura 12. Adição única. Adaptado de Jacobs e Vandeventer (1972).	106
Figura 13. Reflexão. Adaptado de Jacobs e Vandeventer (1972).	106
Figura 14. Inversão. Adaptado de Jacobs e Vandeventer (1972).	107
Figura 15. Matrizes com simetria universal. Adaptado de Ward e Fitzpatrick (1973).	108
Figura 16. Matrizes com Identidade de Elementos. Adaptado de Ward e Fitzpatrick (1973).	109
Figura 17. Matrizes com regra de elementos de conjunto. Adaptado de Ward e Fitzpatrick (1973).	109
Figura 18. Matrizes não simétricas. Adaptado de Ward e Fitzpatrick (1973).	110
Figura 19. Matrizes mostrando operações sobre o conteúdo das células. Adaptado de Ward e Fitzpatrick (1973).	111
Figura 20. Exemplo de item da prova RA da BPR-5.	126
Figura 21. Exemplo de item da prova RN da BPR-5.	126
Figura 22. Dendogramas de Cluster Hierárquico.	137
Figura 23. Gráfico de Dispersão da função de Regressão para prova RA.	140
Figura 24. Dendogramas de Cluster Hierárquico.	148
Figura 25. Gráfico de Dispersão da função de Regressão para prova RN.	151
Figura 26. Gráfico de Escalonamento dos Itens da prova RA pela dificuldade.	157
Figura 27. Gráfico de Escalonamento dos Itens da prova RA pela dificuldade.	162

Introdução

Apresentação

A psicometria, entendida aqui como um conjunto de procedimentos para a construção de instrumentos, passou por várias fases de desenvolvimento (Almeida, 1988). O primeiro, e mais estudado modelo, é chamado de Teoria Clássica dos Testes psicológicos (TCT), ou psicometria clássica. O paradigma de mensuração proposto nesse modelo pressupõe que as diferenças individuais de comportamento observadas representam a magnitude dos elementos mentais. Dessa forma, é possível medir tais elementos por meio da organização de tarefas que supostamente são representações comportamentais dos seus vários níveis de magnitude (Cronbach, 1996; Pasquali, 2000).

Nesse modelo, uma alternativa frequentemente usada como medida dos construtos psicológicos é a referência às normas criadas para grupos. Esse procedimento consiste em combinar um conjunto específico de tarefas de conteúdo psicológico definido, as quais são chamadas de itens, e comparar as respostas de um grupo determinado de pessoas nesses itens. Desta forma a medida é baseada em uma comparação entre o sujeito, objeto da medida, e o grupo, conhecido como grupo normativo ou amostra de padronização. A posição do sujeito no construto psicológico é inferida por meio de uma comparação sujeito-grupo. (Primi, 2000; Pasquali, 1996).

As medidas construídas com referência nesse modelo apresentam algumas desvantagens. Uma delas é que os resultados da medida dificilmente permitem uma compreensão das capacidades do sujeito em termos de processos cognitivos. Por exemplo, uma pessoa que obtém como resultado um $QI = 90$, num teste de inteligência qualquer. Esse resultado não apresenta uma descrição em termos de comportamentos de sua

capacidade de realização. O QI apenas ilustra o desempenho da pessoa num conjunto de tarefas em relação ao grupo normativo.

Segundo Muniz (1994) em 1968 o trabalho de Lord e Novick intitulado “*Statistical Theories of Mental Tests Scores*”, foi definido como marco do desenvolvimento da psicometria moderna com o modelo de análise de dados, que atualmente é conhecido por Teoria de Resposta ao Item (TRI). Ao contrário da psicometria clássica, nesse modelo, compara-se às respostas dos sujeitos com os parâmetros dos itens usados na avaliação, tornando os resultados independentes de amostras da população. Essa característica representa uma possibilidade para se avançar no nível das escalas de medida, do ordinal ao intervalar (Primi, 2000, 2002; Primi & Almeida, 2001).

O presente trabalho teve por objetivo desenvolver escalas de avaliação com interpretações referenciadas nos itens, para as provas de Raciocínio Abstrato (RA) e Raciocínio Numérico (RN) da Bateria de Provas de Raciocínio (BPR-5). Foram usados os procedimentos metodológicos desenvolvidos pela TRI, e análise cognitiva dos itens conforme à proposta por Embretson (1983, 1985, 1994 & 1999).

A apresentação dos fundamentos teóricos foi dividida em três partes. No Capítulo I são apresentadas algumas considerações sobre as concepções de medida em geral e como elas são utilizadas em psicologia. Discutiui-se sobre as conseqüências da medida para o processo de atribuição de significados aos escores dos Testes Psicológicos usando os métodos tradicionais. Os principais avanços da psicometria, com uma introdução à TRI, que permitem a integração com as descobertas da psicologia cognitiva nessa área. Por fim, discutiram-se as implicações desses avanços metodológicos para o procedimento de construção de testes, notadamente ao que se refere à validade de construto. No Capítulo II são apresentadas as principais abordagens de estudo sobre inteligência, como a análise fatorial, a explanação de algumas pesquisas relevantes sobre as bases neurais das atividades

cognitivas frequentemente associadas à inteligência, e os estudos da psicologia cognitiva sobre a natureza do raciocínio analítico-indutivo usando a análise componencial. O Capítulo III apresenta uma introdução a Bateria de Provas de Raciocínio (BPR - 5) com uma breve revisão das pesquisas realizadas até o momento, e também uma reflexão teórica sobre os estudos do raciocínio abstrato e raciocínio numérico. São apresentados modelos de análise sobre a composição dos itens usados na avaliação dos dois tipos de raciocínio. Em seguida é apresentada a parte empírica na qual é descrito o método usado na pesquisa, a fonte dos dados analisados, uma breve descrição das provas que foram analisadas, e os procedimentos adotados. O próximo passo foi apresentar os resultados encontrados para cada prova acompanhados de uma discussão que foi seguida pela descrição das interpretações propostas para as escalas que foram o objetivo do estudo. O capítulo final traz algumas considerações gerais sobre o trabalho, levando em conta os aspectos teóricos e metodológicos utilizados.

Capítulo I – Medida em Psicologia

Considerações sobre a Medida em Psicologia

A psicometria é a disciplina que trata das diversas formas de medida dos atributos da psicologia, contribuindo para construção e verificação de hipóteses científicas acerca dos fenômenos estudados por essa área. Disciplina que pode ser considerada equivalente, por exemplo, ao que representa a econometria em economia, a sociometria em sociologia. Porém, não é dessa forma que ela é frequentemente reconhecida, estando associada de forma quase que exclusiva aos testes e escalas psicométricas, embora não esteja incorreto, esse é seu sentido mais estrito (Pasquali, 1997).

A construção de medidas em psicologia passou por várias fases de desenvolvimento, começando pela psicometria clássica até a psicometria moderna (Almeida, 1988). A Teoria Clássica dos testes psicológicos, ou psicometria clássica, propõe um paradigma de mensuração dos construtos psicológicos, por meio de um conjunto de pressupostos teóricos. Entre os postulados do modelo está a pressuposição de que as diferenças individuais de comportamento estão relacionadas à quantidade e a estrutura de organização dos elementos mentais. Torna-se possível, desta maneira, medir as quantidades desses elementos, organizando tarefas que seriam as representações comportamentais dos seus vários níveis de magnitude e organização (Cronbach, 1996; Pasquali, 2000).

Sobre essa posição, Thurstone (1937) argumenta que a psicologia deve ser uma ciência quantitativa e racional dedicada a descobertas significativas, relevantes e consistentes dos elementos psicológicos que determinam o comportamento. Portanto, a formalização e matematização do enfoque psicométrico são métodos que estão a serviço da psicologia, e não constituem um fim em si mesmo para a psicometria.

Muniz (1994), por sua vez, discute que a preocupação com a medida dos fenômenos psicológicos é inerente a qualquer campo de estudo rigoroso sobre o comportamento humano, e o que constituiria a especificidade da psicometria seria sua especialização e ênfase nas propriedades métricas exigidas das medidas, não importando seu campo de aplicação ou os instrumentos utilizados. Assim, os conteúdos psicológicos sobre os quais trabalham os psicometristas são os mais amplos e variados possíveis.

Desta forma, a medida em psicologia ou psicometria se insere dentro da teoria da medida em geral, ou seja, a preocupação com a medida nesse contexto não é diferente da preocupação em outras áreas do conhecimento. A discussão epistemológica sobre a utilização do símbolo matemático, no caso o número, para representar os construtos psicológicos, passa pela demonstração de que são preservadas tanto as características do número, quanto dos fenômenos aos quais se querem representar por meio deste (Muniz, 1994).

Segundo Muniz (1994, 1996) e Pasquali (1997, 2003) a preocupação com a preservação dos axiomas do sistema numérico na medida dos construtos psicológicos é, em grande parte, sustentada pela natureza desses construtos. Nesse caso, a construção de medidas implica alguns problemas, dentre os quais três devem ser destacados, a representação, a unicidade e o erro.

A representação diz respeito à necessidade de se justificar o uso do número para designar os fenômenos da psicologia. O problema está na razão pela qual se deve acrescentar ao uso de procedimentos empíricos, como a observação, um resultado numérico. Nesse caso, o número possibilita o incremento de precisão nos resultados da observação. Desta maneira, o problema se constitui na demonstração de que tanto o número quanto o fenômeno descrito por meio dele preservam suas características estruturais.

A unicidade diz respeito ao fato de que a utilização do número é a única, ou a melhor forma de representação das propriedades dos fenômenos do mundo natural. No caso da psicologia é sabido que não se trata da única, portanto, é preciso demonstrar que a utilização do número acrescenta vantagem. Ainda assim, quando essa representação pode ser demonstrada como melhor, cabe ao pesquisador esclarecer que as diferenças de qualidade encontradas nas medidas podem estar relacionadas as características dos construtos que foram medidos. As diferenças de qualidade da medida realizada indicam o nível de medida que é expresso nas escalas. Quando todos os axiomas do sistema numérico podem ser mantidos para a descrição do fenômeno, isso implica no tipo mais sofisticado de medida, a escala de razão. Por exemplo, o peso é uma forma de medida na qual os valores podem ser manipulados usando vários tipos de operações matemáticas, ao contrário o QI é uma medida limitada pelo tipo de transformação matemática que se pode fazer de seus valores.

O erro de medida se refere ao fato que todo e qualquer procedimento empírico está sujeito a erros. Estes erros podem se originar de várias fontes, em primeiro lugar, podem ser atribuídos a características dos instrumentos usados na aferição. Por exemplo, a observação de um dado fenômeno pode ser adulterada pela qualidade do instrumento usado, neste caso o controle se dá pela calibração do instrumento. Em segundo lugar, os erros podem acontecer pelas diferenças individuais no julgamento da ocorrência de um dado fenômeno, neste caso o controle pode ser feito pelo treinamento dos examinadores. Em terceiro lugar, os erros podem ser aleatórios, cujas causas estão fora do alcance dos procedimentos empíricos, neste caso, é feito controle estatístico, procedendo a estimação da magnitude e a apresentação de um indicador do valor provável de erro contido na medida.

Uma vez justificada a utilização da medida no campo da psicologia, pela resposta às três questões descritas, a discussão se dirige para a sua operacionalização. A aplicação da

medida conduz a questões sobre a qualidade das informações e das interpretações que podem ser feitas dos valores expressos pelos seus resultados. Esses resultados descrevem certas características do fenômeno medido, e o grau de detalhamento que a medida é capaz de apresentar pode estar diretamente relacionado ao número de axiomas do sistema numérico que foram preservados nessa transição. Em outras palavras, o nível de medida da escala (Pasquali, 1997, 2003).

Segundo Muñiz (1994) dos vários axiomas presentes no sistema numérico, três conjuntos são importantes no caso da medida, aqueles que descrevem as propriedades numéricas de identidade, ordem e aditividade. O primeiro conjunto diz respeito a propriedade da *Identidade* que define que um número é idêntico a si mesmo e somente a si mesmo, apresenta três axiomas que expressam a relação de *igual a* ($=$),

reflexividade, os números são idênticos ou são diferentes $a = a$ e $a \neq b$;

simetria, se $a = b$ então $b = a$;

transitividade, duas coisas iguais a uma terceira são iguais entre si. Se $a = b$ e $b = c$ então $a = c$.

O segundo conjunto diz respeito a propriedade da *Ordem* que se baseia na desigualdade dos números, ou seja, todo número é diferente do outro. Essa desigualdade se caracteriza em termos da magnitude, isto é, um número é diferente do outro porque o primeiro é maior ou menor que o segundo. Desta forma eles podem ser colocados em uma seqüência invariável ao longo de uma escala linear conhecida tradicionalmente como escala monotônica crescente. Esse conjunto apresenta três axiomas que expressam a relação de *diferente de* (\neq) ou *maior que* ($>$),

assimetria, a ordem dos termos não pode ser invertida, se $a > b$ então $b < a$ e $a \neq b$;

transitividade, se um número é maior que um segundo que por sua vez é maior que um terceiro, então o primeiro número também é maior que o terceiro, se $a > b$ e $b > c$, então $a > c$;

conectividade, ou $a > b$ ou $b > a$.

O terceiro conjunto diz respeito a propriedade de *Aditividade* que descreve que, a exceção do zero, os números podem ser somados, ou seja, podem ser concatenados de modo que a soma de dois números produz um terceiro número diferente dos anteriores. As quatro operações fundamentais podem ser aplicadas aos números. Apresenta dois axiomas principais,

comutatividade, a ordem dos termos não altera o resultado da adição, $a + b = b + a$,

associatividade, a ordem de associação ou de combinação dos termos não afeta os resultados, $(a + b) + c$ é igual a, $a + (b + c)$.

Uma medida que preserva todos os axiomas do número é mais sofisticada, podendo aplicar todas as operações fundamentais da matemática aos seus resultados. As escalas de medida podem variar entre *nominal*, *ordinal*, *intervalar* e de *razão*. A escala *nominal* é aquela na qual apenas o axioma de identidade é preservado, neste caso, a única condição necessária é que o mesmo símbolo não pode ser duplicado para designar objetos diferentes, e também que símbolos diferentes não sejam usados para designar o mesmo objeto. Trata-se do nível mais elementar de escala, na qual ainda não existe medida propriamente dita, pois os números não são atribuídos a características dos fenômenos, mas o próprio fenômeno é identificado pelo rótulo numérico, sendo, portanto classificação ou contagem (Pasquali, 1997, 2003). Um exemplo cotidiano de escala nominal é o número de uma pessoa no *Cadastro de Pessoas Físicas* (CPF), nesse caso os números expressos pelo CPF

não apresentam qualquer qualidade relacionada a pessoa que o possui, trata-se apenas de identificação.

A escala *ordinal* é considerada o primeiro nível de medida e pode ser estabelecida quando os axiomas de identidade e ordem são preservados. Neste caso, os números, além de designar características diferentes dos fenômenos, lhes impõem uma ordem. Por exemplo, quando uma professora pede aos seus alunos que se organizem em fila por ordem crescente de tamanho. Não existe uma diferença padronizada de tamanho entre o primeiro e o segundo lugares nessa fila com os demais, ou seja, as diferenças de tamanho entre o primeiro e o segundo pode ser de dez centímetros ao passo que a diferença de tamanho entre a sexta e sétima posição na fila pode ser de dois centímetros. As únicas inferências possíveis são que o aluno que ocupa a primeira posição na fila é menor que seus colegas, e o aluno que ocupa a última posição na fila é maior que seus colegas, também que o segundo colocado é menor que o terceiro e assim por diante (Pasquali, 1997, 2003).

A escala *Intervalar* é o segundo nível de medida e é verificada quando os axiomas de identidade, ordem e parte do axioma de aditividade são preservados. Neste caso, além da ordenação dos fenômenos pode-se dizer que são mantidas iguais as distâncias entre cada ponto da escala. Por exemplo, a diferença de entre seis e sete é a mesma que a diferença entre 26 e 27, ou seja, um ponto, não importando em que parte da escala está localizada, pois os intervalos são sempre iguais (Pasquali, 1997, 2003).

A escala de *Razão* é sem dúvida a escala de medida mais sofisticada, pois preserva os axiomas de identidade, ordem e aditividade. Neste caso, além das distâncias iguais entre os pontos da escala, a origem é conhecida, ou seja, o zero absoluto. Pode-se notar que quanto mais axiomas do número a medida preserva, maior será o seu nível, isto é, mais ela se aproxima da escala numérica ou métrica. Portanto ela suporta maior quantidade de transformações matemáticas (Pasquali, 1997, 2003). Retomando o exemplo de comparação

entre o peso e o QI. Dada uma situação, na qual, João pesa 100 Kg e Maria pesa 50 Kg, então se pode concluir que João tem o dobro do peso de Maria; ou então, Paulo pesa 70 Kg e Pedro 73 Kg, enquanto Marcos pesa 90Kg e Julio pesa 93Kg, se pode concluir que a diferença de peso entre Paulo e Pedro é a mesma que existe entre Marcos e Julio. No caso, a medida da inteligência expressa em uma escala de QI não permite as mesmas conclusões, mesmo que os coeficientes apresentem as mesmas diferenças numéricas. Isso porque, as duas escalas se encontram em níveis diferentes de medida, o peso é expresso numa escala de razão que apresenta um ponto zero e intervalos iguais em todos os pontos da escala, enquanto o QI é expresso em uma escala Ordinal (Pasquali, 1997, 2003).

Todavia, o principal fator para definição do nível de medida não são as características do número, pois este, como objeto de estudo da matemática, possui todas essas características, ordem natural, origem e distâncias iguais entre si. Para a matemática, todas as escalas são de razão. Mas para a medida aplicada à psicologia, que trata da utilização dos números para descrever características dos construtos psicológicos, são as características de cada fenômeno a ser medido, tais como, se elas permitem uma ordem natural ou o conhecimento do ponto zero e a igualdade de distâncias entre suas mudanças de magnitude é que definem o nível de medida (Arias, 1996).

Em psicologia, existe o debate sobre qual o nível de medida alcançado pelos diversos construtos. Na psicometria, em muitos casos, se assume que os valores concedidos às respostas dos sujeitos num teste formam uma escala numérica de intervalo, ou seja, com propriedades de ordem e distâncias iguais. Porém, a maior parte das vezes é de conhecimento do pesquisador que elas não possuem essas características, e nem mesmo as teorias subjacentes são capazes de suportar tais caracterizações (Muniz, 1994; Arias, 1996). Segundo Lord e Novick (1968) existe a necessidade de se avaliar estas questões do ponto de vista pragmático, portanto, do ponto de vista da utilidade dos resultados da escala. Desta

forma, se as pontuações totais no teste proporcionam mais informação para uma classificação adequada dos sujeitos, e também a predição de determinadas condições, quando ele é construído como se fosse uma escala de intervalo, então se justifica o uso de procedimentos estatísticos característicos das escalas de intervalo. Também, essas pontuações sempre são examinadas quanto a sua fidedignidade e validade, portanto, é pouco provável que cumpram com as exigências requeridas nestes aspectos se não possuírem a propriedade de ordem e alguma aproximação com os intervalos de unidades iguais.

Em outras ocasiões, para se tratar as pontuações dos testes psicológicos em uma escala de intervalos se utiliza o argumento da relação entre os valores da escala e a distribuição normal. Portanto, pode-se dizer que no concernente a psicometria, tem se optado por uma aproximação da medida centrada no sujeito. Ou seja, o interesse fundamental está em situar o sujeito em algum ponto no continuum da dimensão psicológica, que geralmente não é observável, e na maior parte dos casos ilustrar as diferenças individuais entre os sujeitos (Arias, 1996).

A partir da compreensão de que a medida apresenta vários níveis, e de como cada um deles é alcançado, resta saber como providenciar o tipo adequado de medida. Existem várias formas de medida, ou seja, várias maneiras de se atribuir números às propriedades dos objetos. Novamente a forma da medida tem relação com as características dos objetos que se pretende medir, e existem várias taxionomias. Para este trabalho será adotada a divisão em três formas diferentes de mensuração, *medida fundamental*, *medida derivada* e *medida por Lei ou teoria* (Pasquali, 1997, 2003).

A medida fundamental é possível quando os objetos ou seus atributos possuem uma unidade base natural, ou seja, eles podem ser divididos em partes menores que mantêm as mesmas características originais. Por exemplo, o comprimento pode ter como base de

medida o metro ou o centímetro e em nenhuma dessas decomposições a medida perde qualquer de suas características elementares. Além de uma unidade base natural, os objetos devem possuir uma representação extensiva, que são dimensões que permitem a associação ou ligação entre elas para formar um outro objeto de mesma natureza, porém maior. Essas características permitem a construção de um instrumento que possui as mesmas dimensões que se pretende medir nos objetos, por exemplo, usar uma unidade de distância para medir o comprimento. Se a distância entre dois pontos quaisquer possui 90 centímetros pode-se dizer que ela possui 90 partes menores do metro que é formado por 100 dessas partes. Essas características fazem como que esse tipo de medida seja conhecido como medida direta (Pasquali, 1997, 2003).

A medida derivada é uma outra forma de medir objetos ou dimensões que não possuam as características exigidas para medida fundamental. Nem todos os objetos possuem unidade base natural ou representação extensiva, mas quando suas dimensões puderem ser associadas a objetos que possuam medida fundamental, eles são passíveis de medida derivada. As relações supostas precisam ser demonstradas empiricamente, assim quando um ou mais atributos de um objeto são afetados diretamente por dois ou mais componentes de um outro objeto, que possua medida fundamental, então se pode estabelecer uma medida derivada por meio de uma função de potência entre os componentes, com os quais o atributo se relaciona. Por exemplo, a medida da densidade de um objeto é derivada da relação que ela possui com outras dimensões que possuem medida fundamental. A massa pode ser medida em quilos e o volume é medido pelo cubo do comprimento (m^3), então densidade é medida pela relação entre massa e volume, ou seja,

$$Densidade = \frac{massa}{volume}.$$

A última forma de medida, chamada de medida por lei ou teoria, é sem dúvida a mais freqüente em psicologia. Os atributos da realidade que são objeto de estudo desta ciência em sua grande maioria não possuem representações extensivas ou unidade base de medida, portanto não são passíveis de medida direta ou fundamental. Nem mesmo são resultantes de relações com componentes de objetos que possuem representação extensiva, portanto não permitem medida derivada. A maioria dos fenômenos da psicologia é medida a partir das relações teóricas cientificamente estabelecidas sobre eles. Um construto psicológico específico pode ser medido indiretamente pela freqüência com que ele se manifesta em termos comportamentais em uma população, e também pela comparação dessa freqüência entre populações diferentes. Uma outra forma de construir medidas dos construtos é pela identificação de relações entre a apresentação do construto e a ocorrência de X 's ou Y 's variáveis observáveis. O mais importante nesse caso é que o instrumento esteja calibrado para medir a variável com a qual o construto esteja relacionado (Pasquali, 1997, 2003). Porém nem sempre as relações entre um construto e variáveis observáveis são claras. Algumas descrições teóricas propõem aspectos por demais subjetivos e que se relacionam apenas fracamente com determinados comportamentos, o que dificulta a demonstração empírica do construto. Dessa maneira a medida que se faz desse construto não pode alcançar o nível de qualidade desejado, pois está condicionada a características subjetivas do construto. Uma vez discutidos os tipos de medida, passar-se-á à discussão sobre a interpretação dos resultados da medida.

Atribuição de Significado aos Escores de instrumentos de medida

A partir do entendimento do nível de medida possível em psicologia, e também à forma como essa medida pode ser realizada, se impõe a questão de como isso afeta a

atribuição de significado aos resultados dos instrumentos de medida. A experiência mostrou que o uso dos resultados brutos é pobre no sentido de informar sobre as diferenças individuais, sobre a magnitude das habilidades, ou das características de personalidade (Cronbach, 1996).

É difícil atribuir significado aos resultados brutos, uma vez que as interpretações mantêm-se dependentes de certas características da prova que permanecem desconhecidas. Por exemplo, João acerta dez questões de uma prova de matemática contendo vinte questões, e acerta sete questões numa prova de língua Portuguesa contendo igualmente vinte questões. Para um avaliador inexperiente ou apressado, esses resultados podem levar a uma interpretação de que João é mais habilidoso em matemática do que em português, ou ainda, que João é mais habilidoso em matemática do que, seu colega Pedro que acertou apenas nove questões na mesma prova. A resposta às questões do tipo, depende de se tornarem conhecidos alguns parâmetros psicométricos das provas, por exemplo, a sua dificuldade (Cronbach, 1996). Tradicionalmente são usados três tipos de procedimentos para se atribuir significado aos resultados brutos que são denominados como interpretações com referência à norma, ao conteúdo e ao critério.

As interpretações com referência na norma são as mais comuns, e se caracterizam pela estimação do desempenho de uma dada população ao responderem aos itens de teste. Os resultados desse grupo são usados como parâmetro de comparação para se verificar qual a posição relativa de um determinado sujeito avaliado (Cronbach, 1996). Assim, os instrumentos de medida são normatizados, combinando um conjunto específico de tarefas de conteúdo psicológico definido, os itens, e comparando as respostas de um grupo determinado de pessoas nesses itens. O modelo psicométrico estabelece a avaliação das grandezas psicológicas com referência às normas criadas para cada grupo, determinando

uma comparação sujeito-grupo. Esses grupos são conhecidos como grupo normativo, ou amostra de normatização (Primi, 2000; Pasquali, 1996).

As interpretações com referência ao conteúdo são caracterizadas pela análise do tipo de problema presente nas provas. Nessa análise é verificada a abrangência das questões com relação às informações que se pretende obter sobre os sujeitos. Por exemplo, se o objetivo da prova é produzir a informação sobre o quanto um determinado indivíduo conhece sobre matemática, ela deve abranger os principais campos do conhecimento sobre matemática. Por outro lado, se a intenção é saber o quanto o indivíduo conhece sobre equações diferenciais, os problemas que constam na prova devem ser apenas sobre esse tema. As interpretações referenciadas no conteúdo consistem, basicamente, na descrição adequada dos problemas presentes nas provas em relação a um universo mais amplo, do qual aquele conteúdo foi extraído, e essa informação é utilizada como base para descrever operacionalmente o significado do resultado alcançado por um indivíduo.

O terceiro procedimento usado para se atribuir significado aos escores é o de se estabelecer relações entre eles e variáveis externas, as quais podem ser então preditas a partir da nota no teste. Essa predição do desempenho do sujeito nas variáveis externas será tão melhor conforme a magnitude das relações encontradas entre elas e as notas no teste. O estabelecimento de relações com variáveis externas pode possibilitar uma predição do desempenho do sujeito na variável, mas também uma descrição sobre sua condição atual no que diz respeito aos elementos dessa variável (Cronbach, 1996; AERA, APA & NCME, 1999).

Os métodos de atribuição de significado, presentes no modelo clássico de construção de instrumentos, trazem consigo algumas desvantagens, já que dificilmente permitem uma compreensão em detalhes da capacidade do sujeito. Resultados apresentados com referência nas normas, por exemplo, apesar de informarem a posição que uma pessoa

ocupa quando comparada com um grupo de indivíduos com características semelhantes, não possibilitam a inferência de que habilidades essa pessoa domina e o grau de sofisticação com que é capaz de realizar determinadas atividades.

Outra dificuldade encontrada na interpretação de resultados de testes calculados por esse modelo, é que eles não são comparáveis a outros instrumentos de mensuração. Acertar 50% dos itens em um teste para avaliação de conhecimento em matemática, por exemplo, não é equivalente a um resultado semelhante quando obtido em outro teste, mesmo que ambos avaliem exatamente os mesmos conteúdos. A comparabilidade de resultados de testes, quando calculados pelo modelo clássico, só é obtida em situações muito específicas, quando o nível de dificuldade de seus itens e dos testes como um todo, for bastante semelhante, bem como as suas medidas de dispersão (Embretson & Reise, 2000).

Os métodos de atribuição de significado aos escores dos testes podem ser usados como indicador do grau de desenvolvimento da ciência psicológica em relação ao seu objeto de estudo. Apesar de os métodos tradicionais, citados anteriormente, terem sua utilidade demonstrada, os problemas encontrados apontam para uma questão que tem recebido pouca atenção dos pesquisadores em geral, a saber, a métrica da medida em psicologia.

Blanton e Jaccard (2006) discutem o que consideram a arbitrariedade da métrica utilizada pelas medidas em psicologia. Os autores definem uma métrica como arbitrária quando não é conhecido o ponto, sobre uma dimensão psicológica latente, onde um determinado escore localiza um indivíduo, ou como a mudança de uma unidade no escore do teste reflete a mudança de magnitude da dimensão psicológica latente. Essa definição aponta a existência de uma função que relaciona o “escore real” de um indivíduo com o construto latente de interesse, e seus escores observados com a métrica da medida realizada. Quando a métrica é arbitrária, a função que descreve essa relação e os valores

dos parâmetros são desconhecidos. Segundo os autores, a métrica frequentemente utilizada na pesquisa e na prática da psicologia pode ser considerada arbitrária. Dessa forma, torna-se necessário ao usuário das medidas, muita cautela para se fazer inferências considerando a magnitude das mudanças em uma dimensão psicológica simplesmente baseado nas mudanças observadas nos escores de teste.

O problema da arbitrariedade da métrica das medidas em psicologia aponta diretamente para o nível de medida que se consegue estabelecer dos fenômenos em questão. Embretson (2006) aponta que uma solução possível para esse problema pode ser encontrada na utilização dos procedimentos psicométricos desenvolvidos na Teoria de Resposta ao Item (TRI). Uma vez que existem diferenças significativas na forma de estimação dos parâmetros dos itens e do teste de modo a representar uma nova relação possível com o construto psicológico.

A TRI se caracteriza como um modelo matemático, inicialmente desenvolvido no bojo de uma preocupação com a equalização de provas. Uma das aplicações da TRI na avaliação de sujeitos possibilita avançar no nível das escalas de medida do ordinal ao intervalar. Ao contrário da psicometria clássica, nesse modelo, as respostas dos sujeitos formam uma função entre sua habilidade e a dificuldade dos itens usados na avaliação, tornando os resultados independentes de amostras da população. Possibilita, desta forma, uma análise mais detalhada da preservação das propriedades axiomáticas na relação entre um construto psicológico e o número usado para representá-lo (Primi, 2000, 2002; Primi & Almeida, 2001).

O estudo da relação entre a probabilidade de acerto e a dificuldade do item produz uma equação monotônica crescente que descreve a posição do sujeito no traço latente, ou seja, sua habilidade. A Figura 2 apresenta um exemplo da Curva Característica de Item (CCI), na qual se observa que sujeitos com maiores magnitudes do traço latente apresentam

probabilidades também maiores de acertar o item, ao passo que pessoas com menor magnitude do traço apresentam menor probabilidade de acertar o item (Primi & Almeida, 2001).

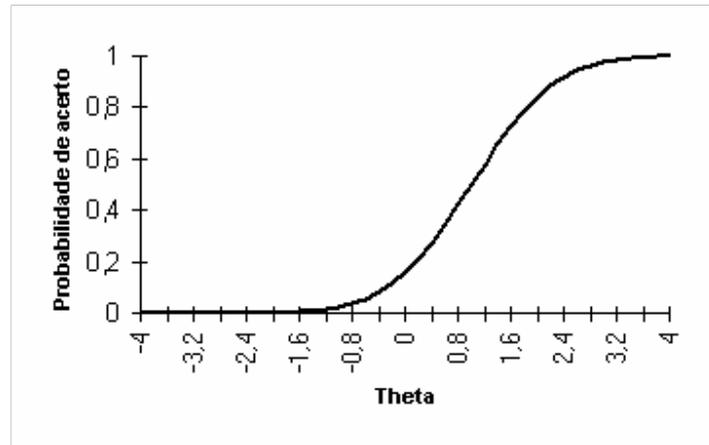


Figura 1. Curva Característica do Item.

Existem vários modelos matemáticos que podem ser usados para representar essa relação, porém os mais comuns são os de 1, 2 e 3 parâmetros, que controlam os seguintes parâmetros dos itens: a , que se denomina índice de discriminação do item, seu valor é dado pela inclinação da CCI em relação ao eixo das abscissas, o valor do ângulo formado por essa inclinação é proporcional ao valor de a , que será tanto maior quanto maior for o ângulo. b que se denomina índice de dificuldade do item, seu valor é dado pelo ponto, na reta, perpendicular ao eixo das abscissas, e representa o valor de (θ) quando a probabilidade de o sujeito acertar ao item é de 50%, quando $c = 0$. A característica mais importante desse parâmetro é que ele se encontra na mesma escala que a variável latente (θ) . E o último parâmetro c , que caracteriza a probabilidade de respostas certas dadas ao acaso, seu valor é dado por $P(\theta) \neq 0$ quando $\theta = -\infty$, ou seja, quando a habilidade do sujeito tende a menor representação possível de (θ) , e ainda assim a probabilidade de

acerto se mantém diferente de zero. Isso significa que o sujeito tem uma habilidade menor que a exigida para a execução adequada do item e, mesmo assim, ele acerta o item. (Hambleton, 1990; Muñiz, 1990; Baker, 2001).

O primeiro modelo é o logístico de um parâmetro, conhecido como modelo de Rasch em referência ao nome de seu idealizador. Trata-se do modelo mais difundido devido a sua parcimônia de medida e da simplicidade de sua lógica. Esse modelo é representado por uma função logística de um parâmetro, a qual considera que as respostas de um sujeito a um conjunto de itens dependem apenas de sua habilidade e da dificuldade dos respectivos itens (Baker, 2001). Sua expressão matemática é,

$$P_i(\theta) = \frac{e^{D(\theta-b_i)}}{1 + e^{D(\theta-b_i)}}, \text{ onde,}$$

$P_i(\theta)$, probabilidade de se acertar ao item i dado um determinado valor de (θ) ,

θ , valor do traço latente ou variável que se estiver medindo,

b_i , índice de dificuldade do item i ,

e , base dos logaritmos neperianos que vale (2,72),

D , constante de aproximação aos valores da curva normal acumulada (1,7).

O modelo é bastante simples e conforme a organização de suas variáveis interpreta-se que conhecendo a dificuldade do item e a habilidade do sujeito é possível prever qual é a probabilidade desse sujeito acertar o item. O valor de b é dado pelo valor de θ no qual a probabilidade de acertar o item é de 50%.

O modelo logístico de dois parâmetros, por sua vez, mantém todas as características do modelo de Rasch, porém, acrescenta a estimação de outro parâmetro do item que é o seu poder de discriminação. Pode-se assumir que o conceito de discriminação do item é o

mesmo usado na Teoria Clássica, ou seja, a capacidade do item diferenciar pessoas com diferentes níveis no construto avaliado. É importante notar que além de estar relacionado, na Teoria de Resposta ao Item, às curvas de informação do teste, a forma de estimativa desse parâmetro na TRI é diferente em relação a Teoria Clássica. Sua expressão matemática é,

$$P_i(\theta) = \frac{e^{Da_i(\theta-b_i)}}{1 + e^{Da_i(\theta-b_i)}}, \text{ onde,}$$

As variáveis são as mesmas descritas na equação anterior com a diferença que se acrescentou o valor de a , que representa o poder discriminativo do item. Uma relação importante entre os parâmetros θ e a é que a discriminação varia em função da habilidade do sujeito, ou seja, os itens discriminam melhor quando a habilidade coincidir com o ponto de inflexão da curva (Muñiz, 1990).

O modelo logístico de três parâmetros mantém todas as características dos modelos de um e dois parâmetros, porém acrescenta a possibilidade de se avaliar a probabilidade de o sujeito acertar o item ao acaso, ou seja, sem que tenha habilidade suficiente. Sua expressão matemática é,

$$P_i(\theta) = c_i + (1 - c_i) \frac{e^{Da_i(\theta-b_i)}}{1 + e^{Da_i(\theta-b_i)}}, \text{ onde,}$$

As variáveis dessa equação são as mesmas da equação anterior com o acréscimo da variável c que representa a probabilidade de acerto ao acaso. O valor de c coincide com o valor de $P_i(\theta)$ para um valor de $\theta = -\infty$. Uma consideração importante, quando se trata

dos modelos de um ou dois parâmetros é que para $\theta = b, P(\theta) = 0.50$ no modelo de três parâmetros se temos $\theta = b, P(\theta) = \frac{1+c}{2}$.

Segundo Muñiz (1990), alguns autores propõem um modelo logístico de quatro parâmetros que visa controlar circunstâncias aleatórias, que podem estar relacionadas com falhas do construtor no momento da elaboração dos itens, isso faz com que um sujeito com grande competência falhe na resolução do item. Sua expressão matemática é,

$$P_i(\theta) = c_i + (Y_i - c_i) \frac{e^{Da_i(\theta-b_i)}}{1 + e^{Da_i(\theta-b_i)}}, \text{ onde,}$$

Todas as variáveis são as mesmas do modelo de três parâmetros, sendo Y uma tentativa de representar a circunstância de falha nos itens que adquire valores pouco menores que um. Até o momento existem poucas pesquisas sobre esse modelo que demonstrem vantagens dele em relação ao modelo de três parâmetros.

Os três modelos da TRI apresentam uma relação direta entre as habilidades dos sujeitos e a probabilidade de acerto dos itens. Após o processo de estimação, os parâmetros dos itens se tornam conhecidos, e ao submeter uma pessoa ao processo de avaliação,

“...aplica-se um conjunto de itens previamente calibrados, isto é, com parâmetros conhecidos ... A atribuição do score ao sujeito é feita comparando-se o perfil de acertos de uma pessoa com as dificuldades dos problemas respondidos para atribuição de um valor numérico que indica a habilidade do sujeito” (Primi, 2004. p 55).

Dessa forma tem-se uma comparação sujeito-item produzindo uma escala referenciada nos itens e independente de amostras da população. O principal avanço do

modelo psicométrico da Teoria de Resposta ao Item, e sua conseqüente contribuição para o parâmetro de validade das interpretações dos escores de testes psicológicos, é a possibilidade de estimar a habilidade das pessoas e a dificuldade dos itens em uma escala comum (Woodcock, 1999).

Avanços da Psicometria e da Psicologia Cognitiva na Construção de Escalas

O caminho trilhado pela psicometria até o modelo da Teoria de Resposta ao Item proporcionou as condições metodológicas necessárias para uma associação com as descobertas sobre os processos mentais oriundos da psicologia cognitiva. Embretson (1983) descreve como a investigação sobre a validade de construto, originalmente proposta por Cronbach e Meehl's (1955), poderia ser decomposta em duas etapas distintas. A primeira, ela chamou de *Representação do Construto*, que visa a investigação sobre quais são os processos cognitivos básicos envolvidos na realização das tarefas presentes nos itens de testes psicológicos; e a segunda, denominada de *Rede Nomotética*, trata dos procedimentos conhecidos em psicometria como referência ao critério, ou mais recentemente, relação com variáveis externas, portanto visa a investigação das relações da medida com as variáveis externas que permitem explicar os resultados encontrados.

Como pode ser notado, a psicometria clássica tratava da validade com maior ênfase e sofisticação somente ao que Embretson (1994) chamou rede nomotética. Uma das formas proeminentes de se atribuir evidências de validade a um teste era por meio de correlações encontradas com medidas de critério. Como explica a autora:

“A conceituação através de representação de construto e medida nomotética permite à teoria cognitiva assumir o papel central no desenvolvimento dos testes. A conceituação clássica de validade de construto

não permite um claro desenvolvimento de itens e testes porque a representação de construto (significado do teste) vem da medida nomotética (significância do teste) portanto, o teste deve já existir. As correlações entre critérios externos e os escores do teste fornecem pouca informação sobre como redesenhar os itens ou os procedimentos de testagem. Na conceituação apresentada, no entanto, o significado do teste provém dos processos que estão envolvidos na resolução do item. Consequentemente a significância do teste segue do significado do teste porque os construtos latentes envolvidos na performance determinam as correlações com outras medidas” (p.109).

A primeira etapa está mais claramente associada com as contribuições da psicologia cognitiva e em especial com a teoria do processamento da informação. As relações entre a psicometria e a psicologia cognitiva se materializaram em dois procedimentos denominados, *Linear Logistic Latent Trait Model* (LLTM) e *Multicomponent Latent Trait Models* (MLMT). Eles associam, da psicometria, especialmente da TRI, a possibilidade de se estimar a habilidade do sujeito vinculada às características dos itens, e também da habilidade estimada estar na mesma escala que a dificuldade do item; e da psicologia cognitiva, a explicação da dificuldade dos itens vinculada às variações na estrutura interna do item. O uso desse tipo de procedimento de investigação é uma faceta da *representação do construto*, e proporcionou uma outra forma de se atribuir significado aos escores de testes, chamada de interpretação com referência nos itens.

O LLTM foi desenvolvido inicialmente nos trabalhos de Fischer (1973). O autor trabalhava com avaliação de processos educacionais, especialmente da possibilidade de se criar instruções ideais para o aprendizado, ou seja, qual a melhor forma de se ensinar determinados conteúdos. Ele identificou que a dificuldade dos itens pode ser explicada pelas características presentes dentro de cada um deles, ou seja, está implícita em cada

tarefa. Propôs a decomposição da dificuldade dos itens nas variáveis descritas na equação a seguir, como forma de estimar valores de dificuldade de cada um dos componentes.

$$b_i = \sum_m \eta_m q_{im} + a ,$$

Onde,

b_i = Parâmetro de dificuldade do item i ,

η_m = Dificuldade do fator de complexidade m ,

q_{im} = Representação da magnitude que o item i exige em relação ao fator de complexidade m ,

a = Constante de normalização,

m = Número de fatores de complexidade exigido para o item i .

Substituindo o b na fórmula do modelo de Rasch pela equação proposta por Fischer para a decomposição da dificuldade, temos a seguinte equação,

$$P_i(\theta) = \frac{e^{D(\theta - (\sum_m \eta_m q_{im} + a))}}{1 + e^{D(\theta - (\sum_m \eta_m q_{im} + a))}} ,$$

que apresenta as variáveis do modelo de Rasch explicadas anteriormente, porém com o parâmetro de dificuldade decomposto nas variáveis que representam os elementos estruturais de cada item. A equação resultante representa a possibilidade de estimação quantitativa da influência dos processos cognitivos, descritos pela psicologia cognitiva, presentes num determinado item de teste.

O MLTM foi desenvolvido por Embretson (1983), também com a intenção de operacionalizar a integração entre os dois campos, cognitivo e psicométrico, e também apresenta uma proposta de decomposição da dificuldade da tarefa, mas difere do método de

Fischer porque acrescenta uma proposta de decomposição da habilidade dos sujeitos nos componentes cognitivos associados a resolução do item. Dessa forma, o valor de θ , que é uma representação da habilidade do sujeito, é decomposto no conjunto de processos básicos que compõe a habilidade que é exigida para a resolução adequada da tarefa, permitindo uma quantificação da magnitude de cada processo. A equação a seguir descreve as variáveis tratadas pelo MLTM.

$$P_{(x_{ij}=1/\theta_j, b_i)} = \prod_k P_{(x_{ijk}=1/\theta_{jk}, b_{ik})} = \prod_k \frac{e^{(\theta_{jk}-b_{ik})}}{1 + e^{(\theta_{jk}-b_{ik})}},$$

Onde,

$P_{(x_{ij}=1/\theta_j, b_i)}$ = Probabilidade que o sujeito j acerte o Item i ,

$P_{(x_{ijk}=1/\theta_{jk}, b_{ik})}$ = Probabilidade que o sujeito j acerte a tarefa associada ao componente específico k envolvido na resolução do item i ,

θ_{jk} = Habilidade do sujeito j no componente k ,

b_{ik} = Dificuldade do item i no componente k ,

K = Subscrito usado para identificar os componentes.

Trata-se de uma das aplicações mais afortunadas da TRI, a possibilidade de agregar aos índices de dificuldade dos itens, informações mais precisas ou particularizadas referentes ao significado do desempenho. Essa aplicação encontrada nos trabalhos de Embretson (1983, 1985, 1994) e também pode ser vista no conjunto de procedimentos utilizados para realizar o escalonamento comportamental proposto por Carroll (1993b). O escalonamento comportamental agrega em sua proposta a possibilidade de se atribuir descrições comportamentais aos resultados de testes, em termos de realização ou conhecimento dos sujeitos avaliados. Assim, esses procedimentos elucidam aspectos da

validade, pois oferecem descrições de como os traços latentes se manifestam por meio de comportamentos característicos ou habilidades específicas.

O processo de construção de escalas cujo significado está referenciado nos componentes que influenciam a dificuldade dos itens possui algumas etapas. Em primeiro lugar é feita a estimação da habilidade dos sujeitos (θ), do parâmetro de dificuldade (b) dos itens e, conseqüentemente, da probabilidade que tem cada sujeito de acertar itens de dificuldade diferente. Esses dados permitem que seja construída a curva característica do sujeito na qual está agregada um conjunto de informação sobre em quais faixas de dificuldade seu desempenho é melhor ou pior. O próximo passo é decompor as tarefas nos seus fatores de complexidade (Carroll, 1993b).

O método de investigação dos construtos que estão por trás da realização de tarefas, ou *representação do construto*, tem por objetivo estabelecer procedimentos capazes de identificar os mecanismos cognitivos teóricos que podem explicar o desempenho dos sujeitos nas tarefas. A aplicação desses procedimentos parte do pressuposto que existe uma relativa dependência entre o desempenho numa tarefa e processos cognitivos, estratégias cognitivas e conhecimento armazenado. Dessa forma, a pesquisa de representação do construto consiste em decompor as tarefas ou itens nos menores componentes possíveis. Portanto, assume-se que os componentes e as transformações executadas por eles na dinâmica de solução de cada item de teste, são os elementos essenciais para descrição da dimensão psicológica em termos de processos mentais (Embretson, 1983).

A utilização desse método para a investigação da representação do construto depende da satisfação de quatro critérios gerais para demonstrar sua adequação, estes foram propostos por Embretson (1983). O primeiro critério diz que as performances das pessoas devem estar relacionadas com as características dos itens do teste. Diferentemente da análise psicométrica tradicional, na qual se avaliam as diferenças entre os sujeitos pela resposta ao item “íntacto”

numa relação com o grupo, essa proposta implica numa análise de decomposição dos itens, da diferença de dificuldade entre eles, e na conseqüente identificação de características que possam explicar tais diferenças e, em última instância, o desempenho de um único sujeito por vez. Presupõe-se que a resposta correta é encontrada pelo sujeito por meio da utilização de processos cognitivos diferentes para solução de cada característica encontrada nos itens. Como o objetivo dos testes psicológicos, também nessa perspectiva, é a medida das diferenças individuais e não das diferenças nas características dos itens, a identificação de tais diferenças são apenas um meio de se avaliar um ou mais construtos, e de demonstrar que as diferenças individuais nos respectivos construtos é que definem o desempenho no teste.

O segundo critério propõe que os métodos de representação do construto devem proporcionar a comparação entre hipóteses teóricas alternativas sobre os construtos que são exigidos para a realização dos itens de teste. Esse critério está relacionado com a proposta de Cronbach e Meehl's (1955) de que a demonstração de validade de construto é equivalente a construir uma teoria, e então demonstrá-la empiricamente. Uma forma efetiva de comparar hipóteses teóricas, segundo esses autores, é operacionalizá-las em modelos quantitativos para explicar o fenômeno, e então comparar o ajuste do modelo.

O terceiro critério exige que o método providencie a quantificação dos processos cognitivos, também chamados de construtos teóricos, por meio da estimação dos parâmetros psicométricos dos elementos nos quais foram decompostos os itens. Isto é, se os itens de um teste foram decompostos com sucesso em muitos componentes aos quais se possa relacionar com processos cognitivos, é necessário que se faça à estimação dos mesmos parâmetros que tradicionalmente se calcula para os itens, para cada um dos componentes que estão envolvidos na sua solução. Trata-se de uma forma de demonstrar a materialização dos processos cognitivos básicos, a partir da medida de sua representação pelos componentes.

É bastante conhecido em psicometria que as propriedades dos testes dependem das características estatísticas dos itens, e essas propriedades permitem avanços na interpretação de resultados, o que representa um incremento na validade. Por exemplo, a testagem adaptativa, na qual se seleciona os itens que melhor se ajustam à medida dos construtos que estão envolvidos, por meio de uma comparação entre os parâmetros de cada item com as características específicas dos sujeitos, expressas durante a realização dos itens que são apresentados. A satisfação desse critério permite que os testes possam ser comparados pelos parâmetros de dificuldade dos itens, tais parâmetros representam a magnitude das propriedades cognitivas teóricas que estão embutidas nos itens.

O último critério diz respeito ao fato de que os métodos de representação do construto devem providenciar medidas particulares, ou seja, medidas das habilidades individuais sobre os processos cognitivos. Isto é, devem possuir parâmetros individuais de rendimento que meçam diferenças individuais sobre os construtos. Ao passo que, ao se definirem quais são os processos cognitivos necessários à realização de cada item de teste e se estimar os valores dos parâmetros dos itens para cada construto, pode-se deduzir o valor das habilidades do sujeito, também em cada construto, necessárias à realização adequada do item.

A partir da conceituação dos critérios que a pesquisa de investigação dos construtos precisa cumprir, Embretson (1994) sistematizou um delineamento de construção de testes que colocou a psicologia cognitiva como ponto central. Esse delineamento, que no trabalho de 1994, foi chamado de *Cognitive Design System to Test Development* (CDSTD) estabelece que inicialmente seja feito o estudo da representação do construto, com ajuda da psicologia cognitiva, e só depois os estudos de validade de medida nomotética. Este método recomenda passos que têm por objetivo otimizar as propriedades cognitivas e psicométricas.

Em primeiro lugar, devem-se especificar os objetivos gerais da medida, definindo-se claramente os traços para os quais se deseja construir um instrumento. Em seguida se estabelecem as características gerais e específicas das tarefas que serão utilizadas como itens do instrumento. Essas características se referem aos atributos da tarefa potencialmente manipuláveis e que possivelmente estejam associados à dificuldade. Em tarefas cognitivas elas são chamadas de *fatores de complexidade*.

O terceiro passo trata do desenvolvimento de um modelo cognitivo para a tarefa. Esta fase é considerada uma das mais importantes dentro dessa perspectiva de construção de testes. Historicamente esse aspecto não recebeu a atenção necessária na construção da maioria dos testes de inteligência. Uma hipótese possível é de que não havia um desenvolvimento adequado da teoria psicológica sobre a maioria dos construtos a serem avaliados. Com o desenvolvimento da psicologia cognitiva e especialmente o modelo de processamento de informação esse cenário apresenta propensão a mudanças (Pasquali, 1997). São aspectos importantes dessa fase, a revisão bibliográfica dos estudos experimentais sobre a tarefa candidata a item de teste e a seleção ou criação de um modelo cognitivo para a tarefa.

O passo seguinte consiste na avaliação empírica do modelo adotado, ou seja, uma avaliação do modelo cognitivo quanto ao seu potencial psicométrico. Desta forma, é possível ter uma visão mais clara do efeito que os fatores de complexidade poderiam ter na dificuldade do item, para orientar o seu delineamento de construção.

O quinto passo estabelece a especificação dos itens orientada pelas variáveis cognitivas. Depois de definidos os fatores de complexidade, deve-se especificar as combinações possíveis entre os fatores para a construção dos itens e a quantidade de itens por combinação. Em seguida, passa-se a criação dos itens, baseando-se nas especificações definidas. Nesta fase pode-se ter o auxílio da informática como a inteligência artificial.

O sétimo e oitavo passos dizem respeito à avaliação das propriedades psicométricas dos itens. Envolve a aplicação dos itens criados a uma amostra representativa da população para se estimar suas propriedades cognitivas e psicométricas. Com relação às propriedades cognitivas, o objetivo é testar o modelo proposto, por exemplo, a partir da predição do tempo de reação, baseado na teoria que especifica os processos envolvidos na resolução dos itens. Com relação às propriedades psicométricas, devem ser utilizados métodos da psicometria clássica e da TRI, ou especialmente métodos multicomponenciais como os discutidos anteriormente, com o objetivo de prever a dificuldade dos itens por meio dos fatores de complexidade. O emprego desses procedimentos possibilita a atenção ao segundo critério da pesquisa de representação do construto que se refere ao teste de hipótese teórica propostas sobre o sistema cognitivo. Em seguida, deve-se finalmente estimar os parâmetros para os itens, incluindo os fatores de complexidade, para a calibração final dos parâmetros.

O nono passo envolve a criação final de um banco de itens que irão compor o teste. Caso o modelo tenha conseguido prever a dificuldade dos itens por meio dos fatores de complexidade, ou seja, sua informação estrutural, o resultado final é um banco de itens indexados por suas características estruturais. Este banco de itens pode ser usado em procedimentos de aplicação como avaliação adaptativa, ou avaliação de componentes específicos. Isto é possível porque, estão associadas a cada item, informações sobre a dificuldade dos processos cognitivos envolvidos em sua resolução via fatores de complexidade. E só então, como último passo, a verificação da *rede nomotética*, que como visto, busca verificar as relações entre os escores componenciais estimados para os sujeitos, e medidas externas de critério.

Embretson, (1994) apontou como principais contribuições do sistema CDSTD para a construção de testes que primeiro, o conteúdo do teste passa a ser formulado por princípios explícitos, segundo, o significado dos escores são associados aos processos cognitivos e

não somente a critérios externos, terceiro, os parâmetros dos itens estão vinculados aos indicadores de complexidade cognitiva, e por último, as habilidades dos sujeitos são medidas indiretas dos processos que são embutidos na dificuldade das tarefas.

Segundo a autora, estas características permitem ao desenvolvimento de itens de testes ser mais sistemático e científico e menos intuitivo e artístico. Os testes podem ter suas propriedades psicométricas aperfeiçoadas e passam a trazer informações diagnósticas mais específicas, vinculadas aos processos cognitivos embutidos nos itens. As escalas apresentam interpretações referenciadas nos itens e não simplesmente escores quantitativos.

O trabalho de Embretsom é uma contribuição importante para a psicologia, e especialmente para esse trabalho, pois foi o primeiro a sistematizar explicitamente um método para construção de testes que relaciona os estudos da psicologia cognitiva com os da psicometria. Por isso é o método mais abrangente disponível atualmente e foi empregado aqui na análise dos subtestes da Bateria de Provas de Raciocínio com o objetivo de desenvolver uma escala com interpretações referenciadas nos itens. Em seguida serão abordadas algumas concepções teóricas sobre a inteligência, especialmente aquelas ligadas a avaliação do raciocínio.

Capítulo II - Inteligência

Desenvolvimento das concepções sobre a Inteligência

Os estudos sobre a inteligência delimitam um dos campos de investigação mais importantes da psicologia desde seu surgimento enquanto uma disciplina científica até os dias atuais. O conceito passou por diversas transformações teóricas ao longo dos anos e aglutinou significativo material de pesquisa empírica que sustenta suas descrições atuais. Apesar de inúmeros trabalhos sobre o assunto terem sido realizados e publicados, a questão sobre a natureza da inteligência ainda não é consensual (Almeida, 1988).

Historicamente, no modelo psicométrico, o estudo da inteligência passou pelo esforço de se realizar sua medida, e suas abordagens passaram por várias fases de desenvolvimento. A partir de estudos com análise fatorial, Spearman (1927) concebeu a “teoria dos dois fatores” sobre a inteligência. Os pressupostos dessa teoria sugerem que por trás de toda atividade intelectual estaria a influência de um fator geral (*g*) e fatores específicos (*s*).

O fator *g* representa uma quantidade fixa de “Energia Mental”, ou seja, essa energia varia livremente de indivíduo para indivíduo, mas permanece inalterada para o mesmo indivíduo ao longo da atividade cognitiva necessária para a realização de uma tarefa a ele correlacionada. Apesar de se encontrar a influência do fator *g* em qualquer atividade cognitiva não significa que essa influência seja a mesma, pelo contrário, ela é relativa, apresentando influência maior para a realização de determinado tipo de tarefa cognitiva e menor para outras. Porém, quanto maior o fator *g* em um indivíduo melhor será seu desempenho em qualquer atividade, obedecendo à relação de influência do fator. Por exemplo, uma situação com dois indivíduos na qual o primeiro apresenta um *g*, de $2x$, e o

segundo tem um g de x ; a capacidade de realização de tarefas será maior no primeiro indivíduo do que no segundo, na mesma proporção da influência do fator.

O fator específico ou s , não só varia de indivíduo para indivíduo como varia também no mesmo indivíduo relacionado com cada tipo de atividade cognitiva que determinada tarefa exige. Segundo Spearman (1927) existe uma condição que eventualmente se torna obstáculo para a demonstração empírica de que a atividade cognitiva possa ser descrita por meio de dois fatores. Essa condição é que os fatores específicos somente são independentes entre si quando as atividades que estão subjacentes à realização das tarefas sejam também diferentes. Quando as atividades são semelhantes ocorre uma sobreposição dos fatores específicos. Essa sobreposição recebeu o nome de fatores de grupo. Por exemplo, quando uma tarefa consiste em marcar todas as letras d em um texto e outra tarefa em marcar todas as letras p , provavelmente, ao se analisar essas tarefas numa matriz de correlações fatoriais, serão encontrados não apenas o g em comum, mas também um outro fator que consiste na superposição ampla dos fatores específicos.

O autor definiu o g em termos cognitivos como responsável por três processos básicos. Segue uma descrição, apresentada por Almeida (1988), de cada processo tal como são entendidos atualmente. (1) Apreensão da experiência, capacidade relacionada à percepção, ou seja, à rapidez e acuidade com que as pessoas são capazes de discriminar os estímulos e a consciência que têm desse processo; (2) educação das relações, que seria a capacidade de estabelecer relações entre duas ou mais idéias, sejam elas percebidas no ambiente ou evocadas da memória por uma outra estimulação; (3) educação de correlatos, que seria a capacidade de derivar novas idéias ou relações a partir de uma idéia ou uma relação previamente estabelecida. Os indivíduos apresentam diferentes níveis de inteligência geral à medida que diferem em sua capacidade para cada um desses componentes.

A idéia de um fator geral de inteligência não demorou a ser questionada. Thurstone (1934) também por meio de estudos fatoriais discute que o cerne da inteligência não se resume a um fator, mas em nove fatores que ele chamou de habilidades mentais primárias. São elas: compreensão verbal; fluência verbal; raciocínio indutivo; raciocínio numérico; raciocínio aritmético; raciocínio dedutivo; visualização espacial; memória e rapidez perceptiva. Os fatores de raciocínio aritmético e dedutivo são menos evidentes e podem ser agrupados sob um único fator de raciocínio. As habilidades são chamadas de primárias ou básicas porque contribuem na realização de um conjunto variável de tarefas. O que diferencia cada uma dessas habilidades, e contrasta com a existência de um fator geral, é o fato de que cada uma delas aparece associada a um determinado tipo de tarefas e são praticamente ausentes em outras.

Ao longo dos anos, um conjunto de autores propuseram modelos conciliatórios que agregavam os aspectos divergentes das teorias de Spearman e Thurstone. McGrew (2007) apresenta um histórico do desenvolvimento das teorias sobre a inteligência, cuja perspectiva aponta duas tradições de pesquisa distintas. Por um lado, a “Tradição Inglesa” utiliza proeminentemente técnicas de análise fatorial que primeiro extraem, em uma matriz de correlações, o fator geral para em seguida extrair os fatores de grupo de menor extensão. Pesquisadores como Burt (1941, 1949) e Vernon (1950, 1961) sugeriram modelos de estrutura hierárquica para a inteligência. Burt postulou a divisão das aptidões em quatro níveis. O mais básico composto pelas sensações e reações, o segundo nível composto pelas percepções e coordenação motora, o terceiro pela memória e formação de hábitos, e o quarto formado pela compreensão e raciocínio. Vernon também postulou uma concepção de inteligência de quatro níveis. No primeiro nível se encontra o *g* de Spearman, e após a extração de *g* os resultados de testes tendem a se agrupar em dois grandes fatores que pertencem ao segundo nível e foram chamados, verbal-educativo e perceptivo-mecânico. Se

a análise for suficientemente detalhada, ou seja, se um conjunto amplo de testes fosse incluído, esses dois fatores podem dividir - se em vários outros que pertencem a um terceiro nível, e são semelhantes aos fatores postulados por Thurstone. Já no quarto nível estão os vários fatores específicos possíveis, relacionados às tarefas como leitura, ortografia, dentre outras.

À medida que se deslocam para o topo da hierarquia, os fatores vão diminuindo e se tornam extensivos às grandes áreas da realização intelectual humana. Ao contrário, quando se deslocam para a parte inferior da hierarquia, os fatores se tornam mais numerosos e relacionados a realização de tarefas específicas ligadas a atividade cotidiana das pessoas. No quarto nível a demonstração empírica se torna mais difícil de ser verificada em amostras populacionais diferentes, em virtude da instabilidade na matriz de correlações. Uma explicação possível é que os fatores começam a se confundir com várias atividades rotineiras que não são freqüentemente medidas por instrumentos, e sim pela observação direta em ambiente natural.

Por outro lado, a “Tradição Americana”, que primordialmente usava métodos de Análise Multifatorial, encontrou de sobremaneira fatores de segunda ordem. Dessa forma, pesquisadores como Thurstone, Cattell e posteriormente Horn propuseram teorias sobre a inteligência que prescindiam de um fator geral. Pode-se concluir, portanto, que a principal divergência entre os dois grupos de pesquisadores esteve baseada na existência ou não do fator geral, que de acordo com o método de análise fatorial utilizado poderia ser mais facilmente identificado.

Segundo McGrew (2007) a teoria de Cattell conhecida por teoria da Inteligência Fluída e Cristalizada apresentou uma visão dicotômica da inteligência, compreendendo dois grandes fatores. Primeiro, a inteligência fluída que se relaciona ao raciocínio que o sujeito apresenta frente a situações novas, e em seguida a inteligência cristalizada que é o resultado

do conhecimento e da aculturação acumulados. Aproximadamente em meados da década de 80, Horn, ex-aluno de Cattell, baseado nos dados das pesquisas disponíveis, identificou a presença de no mínimo sete habilidades amplas além da inteligência fluida e cristalizada. A apresentação teórica proposta por Horn ficou conhecida como teoria Horn-Cattell Gf-Gc, ou teoria da Inteligência Fluida e Cristalizada de Horn e Cattell. O trabalho de Horn teve raiz nos trabalhos de Thurstone e apresentava uma estrutura multidimensional hierárquica em dois níveis com fatores gerais e específicos, agregando também o papel das habilidades na aprendizagem. No nível das habilidades gerais estariam a inteligência fluída, inteligência cristalizada, processamento visual, processamento auditivo, memória de curto prazo, conhecimento quantitativo, velocidade de processamento e velocidade de decisão. No nível das habilidades específicas estariam as tarefas apresentadas aos sujeitos.

As habilidades, nesse modelo, foram organizadas em um continuum e a inteligência fluída localiza-se no extremo adaptativo, estando relacionada à capacidade de apreender uma configuração não familiar e reorganizá-la para satisfazer alguma exigência. Trata-se da atividade mental elementar associada ao raciocínio. A inteligência cristalizada, por sua vez, localiza-se no extremo dos conhecimentos adquiridos pelo indivíduo, corresponde às habilidades específicas já praticadas pelo indivíduo, fortemente associada a aprendizagem (McGrew, 1997).

Horn apresentou duas linhas de evidências contrárias à posição de *g* como um processo unitário. Primeiro, o resultado de medidas precisas, de vários processos mentais, usando baterias de testes bem selecionadas não se ajustam ao modelo do fator *g*. Mais precisamente *g* seria um fator de comunalidade, pois o fator geral extraído em um conjunto de dados gerado por uma bateria específica não é o mesmo que outro fator geral extraído de outro conjunto de dados gerado por uma bateria diferente. Desta forma, é possível, ao analisar os dados apresentados nos vários estudos que propõem a existência do fator *g*,

encontrar pelo menos oito g 's diferentes em cada estudo. Em outras palavras, o problema da teoria sobre a existência de um fator g está no fato de que os fatores encontrados não são os mesmos em um estudo e outro. Nesse caso, o g representa diferentes misturas de medidas e não uma inteligência geral. Segundo, as relações que os supostos indicadores de inteligência geral têm com variáveis de desenvolvimento, funções neurológicas, educação, realização e estrutura genética são muito variadas. Ou seja, tais indicadores, apresentam relações diferentes com os resultados de diferentes critérios na área de realização acadêmica, curvas de desenvolvimento, dentre outros. As diferentes relações encontradas indicam uma validade de construto para diferentes fatores de habilidades amplas e não um simples princípio unitário.

Seguindo com a discussão sobre a existência de g , Carroll (1993a) propõe que os resultados de uma série de análises fatoriais de segunda-ordem em medidas de habilidades amplas podem assegurar a existência de um fator geral. O autor argumenta, com relação ao posicionamento de Horn, sobre a multiplicidade de g 's que podem ser encontrados dependendo do conjunto de dados a serem analisados, que a explicação está no princípio fundamental sobre o qual está baseada a análise fatorial. A natureza de um fator simples descoberto nas intercorrelações entre as variáveis em uma matriz de correlações, não necessariamente relata características especiais das variáveis envolvidas, mas pelo contrário, relata sobre características latentes que são comuns àquelas variáveis.

A partir dos resultados de uma meta análise, Carroll (1993a) propôs uma forma integradora de compreender a relação entre os vários fatores mentais, a Teoria dos Três Estratos da Inteligência. Em seu estudo, o autor fez um levantamento dos últimos sessenta anos de pesquisas realizadas pela psicometria sobre a inteligência, e selecionou 1500 artigos nos quais obteve 460 conjuntos de dados. Esses dados foram então analisados e os resultados sugeriram uma estrutura fatorial da inteligência que se organiza em camadas que

estão ligadas hierarquicamente. O terceiro estrato, considerado o mais básico, se relacionava ao fator geral proposto por Spearman, foi descrito como um complexo de processos cognitivos superiores que são comuns a todas as atividades intelectuais. O segundo estrato, intermediário, composto de habilidades ou capacidades inferidas da inteligência, incluindo a inteligência fluída, cristalizada, processos de aprendizagem e memória, percepção visual, percepção auditiva, habilidade de recuperação, velocidade cognitiva e velocidade de processamento e decisão. E o primeiro estrato, considerado mais específico, relacionado com as características das tarefas usadas na avaliação.

Mais recentemente a teoria dos três estratos foi integrada ao modelo da inteligência fluída e cristalizada (Gf-Gc), passando a ser conhecida como teoria de Cattell-Horn-Carroll (CHC) das habilidades cognitivas (Carroll, 1997). A partir da integração entre as duas principais teorias sobre a inteligência a discussão sobre a existência de *g* deu lugar a discussão sobre a natureza de *g*. A principal diferença entre as duas teorias refere-se ao fator geral que no modelo dos três estratos é obtido por uma análise fatorial de segunda ordem. Concluiu-se que o fator geral é um tipo de processamento mental subjacente a todas as atividades cognitivas descritas no segundo estrato. Ao passo que no modelo de Cattell e Horn o fator geral foi subdividido em dois, a Inteligência Fluída e Cristalizada. Outras diferenças encontradas nos dois modelos diziam respeito à quantidade de fatores presentes no segundo estrato. No caso, o modelo de Carroll apresentava apenas sete fatores que são bastante semelhantes aos dez apresentados no modelo de Cattell e Horn. No modelo integrado CHC (Figura 1), pode-se notar a separação entre conhecimento quantitativo que fazia parte da inteligência fluída, o mesmo aconteceu com a habilidade de leitura e escrita que estava associada a inteligência cristalizada (McGrew & Flanagan, 1998). Com relação ao primeiro estrato, havia apenas 69 habilidades específicas no modelo dos três estratos, outras seis habilidades identificadas por Carroll, desempenho matemático, desempenho em

geografia, informações de ciência geral, informações sobre a cultura, informações gerais, conhecimento de usos da linguagem, foram consideradas como aspectos relacionados ao conhecimento e desempenho acadêmico e não como características da inteligência. Na versão do modelo CHC foram incluídas as seis habilidades ampliando o primeiro estrato para 75 habilidades, fazendo dessa teoria a descrição mais completa da atualidade sobre a inteligência (McGrew & Flanagan, 1998).

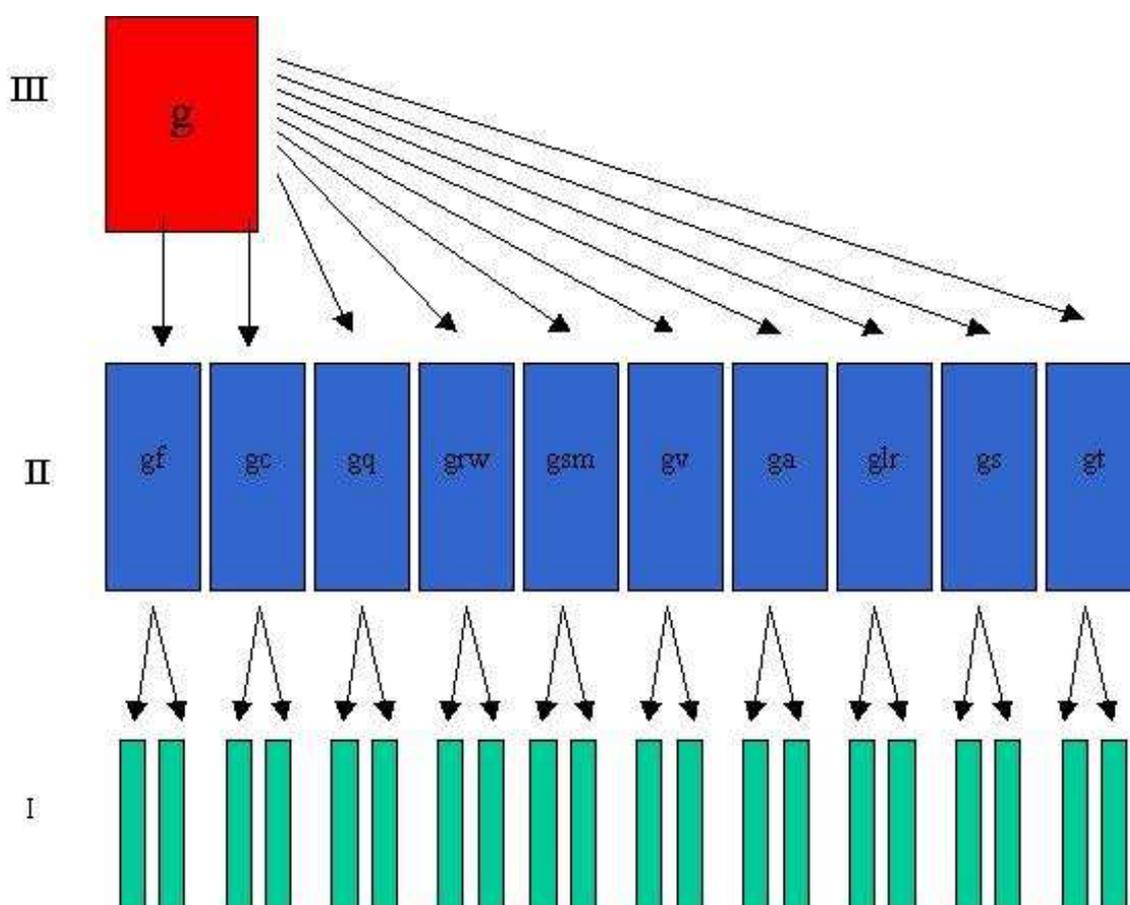


Figura 2. Teoria CHC das Habilidades Cognitivas.

Uma revisão das pesquisas, usando análise fatorial, realizadas no período de 1993 e 2003 sugeriu um conjunto de extensões para o segundo estrato na taxionomia do modelo CHC. Segundo Stankov (2000) diferentes pesquisas apontam que as extensões estruturais

podem ser classificadas em *internas* e *externas*. As primeiras descrevem a natureza dos fatores amplos bem estabelecidos, ao passo que a segunda sugerem novos domínios de habilidades amplas que têm sido apenas parcialmente investigados. Em seguida serão apresentadas as definições e uma breve descrição dos estudos, sobre as dez habilidades amplas, e as extensões estruturais *internas* e *externas* que compõe o segundo estrato do modelo CHC, conforme apresentada por Carroll (1993a) e McGerw (2007).

O fator *Inteligência Fluída (Gf)*, também denominado não verbal ou geral, refere-se às operações mentais que o organismo deve realizar quando enfrenta uma tarefa relativamente nova que não pode ser realizada automaticamente. Essas operações envolvem relacionar idéias, induzir conceitos abstratos, compreender implicações, solucionar problemas utilizando-se principalmente de raciocínio indutivo e dedutivo e processamento integrado de representações visuais e auditivas. Trata-se do fator mais relacionado ao fator *g* apresentado no terceiro estrato.

O fator *Inteligência Cristalizada (Gc)*, também conhecido como raciocínio verbal, refere-se à extensão e profundidade dos conhecimentos desenvolvidos pela cultura dominante. Trata-se de uma habilidade de raciocínio baseada na aplicação efetiva de conhecimentos aprendidos previamente. Arquiva principalmente conhecimento verbal ou baseado na linguagem, resultante do investimento de outras habilidades durante as experiências escolares e cotidianas. Envolve habilidades associadas à modalidade auditivo-simbólica de representação e informações. Uma das formas frequentemente usadas para avaliar *Gc* é pedir aos indivíduos que definam ou solucionem problemas que exigem o uso e a compreensão de palavras que variam de freqüência em uma cultura. Vocabulário, informação geral e conceitos associados com ocupações especializadas ou profissões, e domínios de conhecimento específico são geralmente evitadas na avaliação e pesquisas sobre *Gc*.

McGrew (2007) aponta que altas performances em um ou mais domínios de conhecimento específicos são geralmente associadas ao construto de sabedoria ou *expertise*. O conhecimento psicométrico produzido até o momento encontra dificuldades em estabelecer claramente os limites entre os construtos de Gc e um outro fator indicado no segundo extrato *Conhecimento Geral de Domínio Específico (Gkn)*. Enquanto Gc é caracterizado pelo domínio e apropriação que o indivíduo apresenta em relação aos aspectos próprios da cultura em que está inserido, Gkn é caracterizado como a amplitude e profundidade de conhecimentos adquiridos sobre um domínio especial que tipicamente não representam experiências universais gerais de um indivíduo com a cultura. Por exemplo, conhecimento do Inglês como uma segunda língua, extensão do conhecimento em geografia; conhecimento mecânico, conhecimento sobre a função, terminologia e operação de máquinas e equipamentos; informação sobre ciência geral, física, biologia, engenharia, etc.

Um exemplo dos estudos que tentaram encontrar os limites entre as duas habilidades é o de Rolfhus e Ackermam (1999) que investigaram as relações entre medidas tradicionais de Gc, um conjunto de vinte provas de conhecimento especializado (por exemplo, história, literatura e artes, astronomia, biologia, economia, eletrônica, direito, negócios, etc.) e medidas tradicionais de habilidades numéricas e espaciais, que foram aplicadas em uma amostra de universitários. O primeiro passo foi realizar uma análise fatorial de segunda ordem das vinte provas, que apontou como resultado quatro fatores específicos de conhecimento (humanidades, ciências, civismo e mecânica) ligados a um fator geral de conhecimento (Gkn). Foi encontrado que os fatores Gc e Gkn se correlacionaram em um nível moderado, indicando que Gc, tal como é avaliado, é um construto relacionado, porém independente de Gkn. Os autores apontaram ainda, que seja passível de questionamento se a separação dos fatores amplos de habilidade (Gc, Gkn),

como a apresentada nos seus resultados, pode ser encontrada também nos primeiros níveis de desenvolvimento, por exemplo, com amostras de crianças. Isso porque existe uma ampla fonte de variância educacional comum, devido a uma homogeneização cultural dos processos de escolarização. Embora pareça viável uma distinção entre os dois fatores a partir do ensino médio.

McGrew (2007) aponta que a inclusão do fator amplo Gkn em uma teoria da inteligência é importante, e que esse fator pode ser entendido como uma especialização de Gc que ocorre ao longo da vida adulta. Em uma sociedade altamente especializada, o conhecimento não é uma entidade unitária, especialmente em altos níveis de funcionamento e amadurecimento. As habilidades de Gkn são o resultado de experiências com atividades de domínio específico, tipicamente dependentes de prática e treinamento regular, freqüente e sistemático depois de no mínimo uma década.

Horn e Masunaga (2000) estudaram o construto de *expertise*, que na perspectiva da teoria CHC se relaciona com Gkn, e hipotetizaram que o raciocínio envolvido no exercício de *expertise* é largamente baseado em conhecimento e dedução, ao contrário do raciocínio característico de Gf, que é baseada em indução. Os autores concluíram que a desempenho superior dos *experts* é caracterizado por uma memória de longo prazo. Ou seja, dentro de um domínio de conhecimento específico, essa memória abastece o *expert*, em uma situação imediata, com muito mais informação do que é possível por um sistema de retenção de curto prazo. Esse sistema de longo prazo aperfeiçoa uma forma de raciocínio dedutivo que utiliza um armazenamento complexo de informação capaz de efetivamente antecipar, prever, avaliar, verificar, analisar e monitorar a resolução de um problema dentro de um domínio de conhecimento. Essas habilidades caracterizam a maturidade de expressão da inteligência.

O fator de *Recuperação da Memória de Longo Prazo (Glr)* pode ser caracterizado pela apreensão e recuperação de longo prazo, fluência e flexibilidade. Refere-se à habilidade para armazenar informações e a fluência de recuperação por meio de associações. Está relacionada não somente a extensão da informação armazenada, mas à facilidade de sua recuperação.

O fator de *Conhecimento Quantitativo (Gq)* foi definido como o montante de conhecimento de procedimentos armazenados pelo organismo que envolve a habilidade de compreender conceitos e relações quantitativas e manipular símbolos numéricos. O conhecimento quantitativo se diferencia da inteligência fluída no fato de que o primeiro consiste basicamente em conhecimentos e procedimentos matemáticos (operações mentais) e a última envolve a habilidade de raciocinar indutiva e dedutivamente e não implica em conhecimento.

O fator de *Habilidade de Leitura e Escrita (Grw)* refere-se ao total de conhecimentos armazenados que incluem habilidades básicas de compreensão da linguagem escrita e a expressão do pensamento por meio da escrita. Envolve habilidades básicas como decodificação e complexas como a compreensão de um texto e ainda conhecimento de uso da língua e velocidade de leitura.

O fator *Memória de Curto Prazo (Gsm)* refere-se à capacidade de aprender ou apreender, reter e manter as informações na consciência por um momento para poder recuperá-las e usá-las de imediato durante um período curto. Pode ser também definida como a habilidade de simultaneamente manter e processar as informações (memória de trabalho). A definição apresentada encampa dois processos diferentes de memória que representam uma discussão entre os pesquisadores da inteligência. A distinção entre a Memória de Curto Prazo (MCP) e Memória de Trabalho (MT) como construtos separados. Por um lado, Engle, Tuholski, Laughlin e Conway (1999) testaram se MCP e MT são

construtos distintos. Participaram do estudo 133 estudantes universitários que responderam a três tarefas simples de armazenamento de curto prazo e três tarefas que requeriam processamento complexo e armazenagem. Foram usadas técnicas de análise fatorial confirmatória para testar se os dados se ajustavam melhor a um modelo de um ou dois fatores. Os resultados indicaram que o modelo usando dois fatores apresentava melhor ajuste aos dados, mas que os fatores latentes apresentavam uma correlação de 0,68. Vários outros estudos apresentaram resultados semelhantes, como por exemplo, Kail e Hall (2001), Conway, Cowan, Bunting, Therriault e Minkoff's (2002), que conjuntamente sugerem a interpretação de que MCP e MT são construtos separados, mas correlacionados. Por outro lado, a Memória de Trabalho foi conceitualizada por Baddeley (1976) como um construto que consiste em pelo menos três componentes, a saber, memória viso-espacial, processamento fonológico e executivo central. Portanto os resultados dos estudos que sugerem que os dois construtos são separados, mas relacionados, podem estar indicando na verdade uma relação entre aspectos do mesmo construto. Uma explicação possível pode ser encontrada na organização das tarefas usadas em cada um dos três estudos citados. Em cada um deles são utilizados sempre problemas simples de memória versus problemas complexos. Por exemplo, McGrew e Woodcock (2001) realizaram um estudo utilizando métodos de análise fatorial confirmatória e exploratória nas provas do teste Woodcock e Johnson III (WJ III) e os resultados apontam que quando um conjunto mais amplo de medidas do CHC é utilizado, as medidas de MT não apontam para um construto separado, pelo contrário, indicam uma mistura complexa de habilidades compatível com a descrição teórica do construto. Dessa forma os autores concluem que mais estudos são necessários para decompor e medir os vários processos por trás da MT.

O fator de *Processamento Viso - espacial (Gv)* definido como a habilidade de gerar, perceber, analisar, sintetizar, guardar, lembrar, manipular, transformar e raciocinar com

representações ou padrões visuais. Inclui a habilidade de imaginar as características dos objetos após aplicação de transformações diversas como rotações, reflexões e está relacionada à modalidade visual de representação de informações. Essa habilidade tem apresentado uma posição dicotomizada nos estudos sobre inteligência. Por um lado, ela tem sido relegada a um *status* secundário devido em parte a alguns aspectos, primeiro, além de um nível mínimo de proficiência essa habilidade não apresenta consistência na previsão de sucesso nas atividades escolares ou no trabalho. Segundo, as relações entre Gv e um critério externo são geralmente diminuídas quando outro preditor mais poderoso, no caso Gf ou Gc, são incluídos no estudo. E em terceiro lugar, as medidas de Gv usadas nos estudos podem ser pobres para representar o funcionamento viso-espacial. Por outro lado, essa habilidade continua presente nos estudos sobre inteligência devido a sua corrente associação com realizações consideradas de grande prestígio em profissões exigentes como engenharia, arquitetura, física, e também em atividades mais prosaicas como carpintaria, mecânica e ocupações industriais (McGrew, 2007).

As características estruturas dessa habilidade foram pouco estudadas até o momento, por exemplo, Juhel (1991) estudou a estrutura de Gv usando métodos de análise fatorial exploratória com uma amostra de estudantes de graduação. Seus resultados apontaram Gv como uma função cognitiva complexa derivada de tarefas que envolvem Visualização (Vz), Relações Espaciais (RE) e Memória Visual (MV). As tarefas de visualização foram apontadas como envolvendo um processamento cognitivo mais complexo, com altas cargas em Gv quando é aumentada a complexidade das manipulações e transformações. Já as tarefas que envolvem RE e MV aparecem em um nível menos complexo, caracterizando habilidades estreitas (compatíveis com o terceiro estrato da teoria CHC) ou mais próximas da periferia, considerando o modelo hierárquico. Posteriormente Miyake, Friedman, Rettinger, Shah e Hegarty's (2001) hipotetizaram e confirmaram que as

habilidades Vz, RE e MV se diferenciam em função da relativa exigência sobre o sistema de memória de trabalho, especialmente memória de trabalho visoespacial e componentes do funcionamento executivo. Além disso, as três habilidades requerem um grau substancial de armazenamento visoespacial, e a manutenção de representações visoespaciais envolvidas no desempenho sobre testes de habilidade espacial pode ser fortemente ligada ao funcionamento executivo e atenção controlada.

O fator de *Processamento Auditivo (Ga)*, está associado a modalidade auditiva de representação das informações que afetam o desenvolvimento, embora não tenham efeito sobre a compreensão. Depende do aparelho auditivo e refere-se ao grau com que o organismo pode exercer controle cognitivo sobre a percepção dos estímulos auditivos, compreender e sintetizar padrões de estímulos auditivos incluindo a percepção de padrões sonoros apresentados em condições inadequadas ou distorcidos, apreender ordem e ritmo entre sons, compreensão de grupos de sons apresentados simultaneamente. Segundo McGrew (2007), da mesma forma que o processamento visoespacial, historicamente as habilidades auditivas têm recebido pouca atenção na literatura psicométrica. Afortunadamente o interesse em pesquisas sobre o domínio Ga foi incrementado por vários aspectos, três entre eles foram destacados. Primeiro, o avanço tecnológico ofereceu equipamentos que facilitaram a condução de pesquisas sobre esse domínio; segundo, o aumento do interesse na psicofísica da percepção auditiva; e terceiro, uma explosão de pesquisas que focaram as relações entre as habilidades de Ga de processamento fonológico e consciência fonológica com o desenvolvimento precoce de leitura e distúrbios de leitura.

McGrew (2007) relata que inicialmente um pequeno número de estudos usando métodos exploratórios e confirmatórios sugeriam que a consciência fonológica (CF) podia ser dividida em duas habilidades menores de análise e síntese. Porém, estudos posteriores começaram a indicar que a CF é uma habilidade unidimensional. Por exemplo, Wagner,

Torgesen e Rashotte (1994) realizaram um estudo longitudinal com 244 crianças. Os pesquisadores usaram o método de análise fatorial confirmatória e especificaram previamente suas hipóteses sobre a divisão da consciência fonológica em dois fatores, análise e síntese. Os resultados apontaram altas correlações entre os dois fatores gerando um problema de alta multi-colinearidade quando ambos eram incluídos no modelo de predição. Os autores concluíram que os dois fatores deviam representar o mesmo construto e modificaram seu modelo para incluir um fator simples de consciência fonológica. Em outro estudo Anvari, Trainor, Woodside e Levy (2002) exploraram as relações entre consciência fonológica, percepção musical e leitura precoce em uma amostra de 100 crianças de quatro e cinco anos de idade. Os resultados apontaram que as medidas de consciência fonológica revelaram um único fator nas duas faixas de idade, porém a análise das tarefas de percepção musical indicou um fator único (fator musical) para o grupo de quatro anos e dois fatores (percepção de frequência e percepção de ritmo) para o grupo de cinco anos. Os resultados indicam uma independência entre percepção musical e desempenho da consciência fonológica.

A consideração da “velocidade mental” como um indicador de brilhantismo ou inteligência pessoal tem ocupado uma posição central nos estudos sobre as habilidades cognitivas humanas. Algumas perspectivas são mais frequentes no estudo da velocidade do processo mental, em primeiro lugar desponta a perspectiva mais associada com a aplicação das baterias de inteligência e com a velocidade no desempenho de tarefas com pouca dificuldade ou tarefas treinadas exaustivamente. A velocidade do processamento cognitivo, entendida como um fator amplo, é denominada, *Velocidade de processamento (Gs)* e diz respeito à habilidade de realizar uma tarefa cognitiva de forma rápida e automática, mantendo a concentração e a atenção focada mesmo quando sob pressão. Está envolvida com a rapidez de realocação dos limitados recursos cognitivos dos organismos para realizar

diversas tarefas (McGrew, 2007). Os estudos de Ackerman e Kanfer (1993) e Ackerman e Cianciolo (2000) demonstraram que a velocidade perceptual tradicional está no alto de uma hierarquia que inclui pelo menos quatro fatores de velocidade perceptual incluindo, (1) velocidade de reorganizar padrões visuais simples, (2) examinar, comparar e visualizar o estímulo, (3) realizar tarefas que exigem significativa demanda de memória imediata, (4) realizar tarefas de reconhecimento que impõem demandas cognitivas adicionais como visualização espacial, estimativa, intercalar e levantar a extensão da memória. Esses quatro fatores podem ser localizados no nível de habilidade do III estrato da teoria CHC.

Em segundo lugar está a perspectiva de velocidade mais associada com o modelo experimental que emprega medidas cronometradas de tempo de reação e decisão. Chamada de *Velocidade de Decisão (Gt)* é a habilidade de responder de forma eficiente e correta a tarefas cognitivas mais complexas que envolvem problemas de compreensão, raciocínio, e solução de problemas (rapidez de reação). É tipicamente medida considerando o tempo que um indivíduo necessita para apresentar uma resposta (correta ou não) para uma variedade de tarefas. Segundo McGrew (2007), uma revisão da literatura sobre as medidas de velocidade de processamento de informação básica, como medidas de Tempo de Reação (TR) e Tempo de Inspeção (TI), tem apontado relações entre elas e g em uma amplitude de 0,30 à 0,50. De um modo geral quando medidas de velocidade tradicionais em psicometria são fatoradas junto com medidas de tempo de reação e inspeção dois fatores separados emergem, um de TR e outro interpretado como tempo de movimento. Os resultados apontam para uma relação próxima entre as medidas de velocidade e a capacidade de inteligência geral (g). Apesar disso, não há consenso sobre qual dos dois processos (TR ou TI) são medidos, e qual a implicação dessas medidas sobre a teoria da inteligência e a prática aplicada.

Em adição às duas primeiras categorias de velocidade, pode ser acrescentada uma terceira, chamada de *Velocidade Psicomotora Geral (Gps)* definida como a habilidade para rapidamente e fluentemente executar movimentos motores com o corpo, independentemente de controle cognitivo. Essa categoria pode ser caracterizada pela habilidade para copiar palavras ou sentenças repetidamente ou escrever palavras ou sentenças tão depressa quanto possível. E também pela habilidade de rapidamente executar sucessivos movimentos com a articulação muscular, não sendo considerado importante a precisão desses movimentos.

Até então, a taxionomia do modelo CHC mantinha uma lacuna significativa a respeito de alguns domínios sensoriais humanos, como habilidades táteis, cinestésicas e olfativas. Habitualmente se reconhece a importância de compreender como tais processos são usados por indivíduos com déficits visuais para se conduzir, ou por *experts* em atividades de *gourmets, sommelier*, etc. Dessa maneira é comum que nas principais baterias de avaliação neuropsicológica estejam presentes medidas de habilidades sensoriais (Lezak, 1995). Apesar disso, as evidências estruturais para esses domínios têm sido escassas (Carroll, 1993).

Porém, desde a proposta do modelo CHC, um grande número de estudos tem analisado bancos de dados contendo variáveis sensoriais em conjunto com variáveis cognitivas. No geral, tais estudos têm sugerido que essas habilidades sensoriais podem ampliar a taxionomia do modelo CHC (Stankov, 2000; Bowman, Markham & Roberts, 2002). Duas habilidades amplas têm sido sugeridas para o domínio tátil e cinestésico. A *Habilidade Tátil (Gh)* foi definida como aquela que depende dos receptores sensoriais do sistema tátil e de *inputs* sobre o aparato tátil. Trata-se da habilidade de realizar discriminações finas de pressão sobre a superfície da pele. E também, *Habilidade Cinestésica (Gk)* que depende dos receptores sensoriais que detectam a posição corporal,

peso ou movimento dos músculos, tendões e juntas. Habilidade para detectar ou ter resposta ou movimento do corpo ou partes do corpo, incluindo movimento dos membros superiores e a habilidade para reorganizar uma rota previamente explorada sem *input*.

Stankov (2000) isolou diferentes habilidades no nível mais específico (primeiro estrato) envolvendo processamento refinado tátil e cinestésico. As habilidades táteis são processamentos que dependem de fina discriminação de textura e sensibilidade de pressão sobre a pele, enquanto que habilidades cinestésicas envolvem a consciência dos movimentos dos membros superiores e a habilidade do sujeito de reconhecer visualmente um caminho que foi percorrido de olhos vendados. Tais habilidades compõem um fator de performance tátil representado pelo desempenho nessas tarefas. Porém o autor considera que é difícil desmembrar a influência de outros fatores do segundo estrato, especialmente raciocínio viso-espacial e inteligência fluida, sobre o desempenho nessas tarefas, porque o processo de resolução parece envolver habilidade de visualização.

Uma terceira habilidade ligada aos movimentos do corpo tem sido sugerida e foi definida como *Habilidade Psicomotora (Gp)* caracterizada pela precisão para desempenhar movimentos motores com o corpo (movimentos dos dedos, mãos e pernas) com coordenação ou força. E também pela coordenação de muitos membros, destreza manual, habilidade para executar precisamente movimentos coordenados com a mão ou a mão e um objeto. Além disso, habilidade de executar precisamente e fluentemente uma seqüência de movimentos coordenados para um determinado fim.

No que diz respeito ao processamento olfativo, Carroll (1993a) relata o estudo de Jones (1957) que investigou a estrutura do domínio olfativo a partir da análise da sensibilidade de discriminação de cheiros. Os resultados sugeriram uma possível estrutura hierárquica composta pela discriminação de quatro odores específicos no primeiro nível e dois fatores de ordem superior que não foram nomeados. Em outro estudo, Danthiir, Pallier,

Roberts e Stankov (2001) examinaram o lugar do processamento olfativo dentro da esfera das habilidades cognitivas humanas. Nesse estudo foram utilizadas dezessete tarefas, cinco das quais foram apresentadas como supostas medidas de cognição olfativa. Os resultados mostraram quatro fatores dos quais três também apresentaram relações com Raciocínio viso-espacial e Inteligência fluida, semelhante aos resultados apresentados pelas tarefas de modalidade tátil-cinestésica. Porém um dos fatores pôde ser predominantemente definido por tarefas de memória olfativa e se mostrou independente de outras habilidades cognitivas presentes no segundo estrato e também de simples acuidade olfativa. O autor concluiu que esse fator pode ser considerado uma parte da taxionomia do modelo CHC das habilidades humanas e foi definido como a *Habilidade Olfativa (Go)* que depende dos receptores sensoriais do sistema olfativo principal, e envolve a habilidade para memória de odores e também a sensibilidade para diferenciar odores.

O advento da teoria CHC sobre inteligência representou uma amenização das divergências históricas nas posições em relação a existência de *g*. Os esforços de pesquisa gradualmente voltaram-se para tentativa de se compreender o processo mental por trás de *g*. Nessa medida a integração de aproximadamente um século de pesquisas psicométricas com teorias contemporâneas de processamento de informação resultaram em um importante progresso na compreensão sobre a inteligência em geral e a capacidade de raciocínio em particular. Apesar de alguns modelos de processamento de informação terem sido hipotetizados e pesquisados, em geral, quatro tipos de processos podem ser considerados consensuais na literatura para apresentar relação com habilidades cognitivas de ordem superior. São eles Conhecimento Procedural e Declarativo (CP, CD), Velocidade de Processamento (Gs) e Memória de Trabalho (MT) (Kyllonen, 1996).

McGrew e Woodcock (2001) usando procedimentos característicos da pesquisa em psicometria, especialmente as técnicas de Modelagem de Equações Estruturais (MEE),

tiveram a intenção de verificar as relações entre as medidas de habilidades cognitivas consideradas como eficiência de processamento de informação (MT, Gs, e MCP) e habilidades complexas (*g*). Para tanto os autores selecionaram testes construídos para medir os fatores do segundo estrato da teoria CHC. O *g* foi operacionalmente definido como um fator latente de segunda ordem extraído de cinco medidas desses fatores (Gf, Gc, Glr, Ga e Gv). Os testes foram aplicados em uma amostra dividida em cinco grupos de idade diferenciada (6-8; 9-13; 14-19; 20-39; 40-90+), contendo aproximadamente entre 1000 e 2200 sujeitos cada. Os resultados apontaram em todos os grupos da amostra que MT apresentou um amplo efeito causal sobre a performance em *g* e que esse efeito tende a decrescer com o aumento da idade (0,93; 0,90; 0,82; 0,83; 0,73). Também, consistentemente com os resultados de outras pesquisas, por exemplo, Engle, Tuholski, Laughlin e Conway (1999), Gs pôde ser considerado um precursor direto da MT, embora esse efeito só tenha sido encontrado de forma significativa a partir dos 14 anos. Os resultados observados nos respectivos grupos [14-19 (0,07); 20-39 (0,09) e 40-90+ (0,22)] sugerem que Gs apresenta uma pequena influência sobre MT na adolescência que vai crescendo até atingir seu ápice depois dos 40 anos.

Segundo McGrew (2003) desde os anos 90 o conceito de MT tem sido usado como ponto central em pesquisas que tentam explicar as diferenças individuais em habilidades cognitivas de alto nível, como compreensão da linguagem, leitura e matemática, dentre outras. Os resultados das pesquisas têm apontado que diferenças individuais em capacidade de MT podem ser responsáveis pelas diferenças individuais encontradas em habilidade geral (*g*). Por exemplo, Kyllonen e Cristal (1990) e Kyllonen (1996) encontraram correlações entre os fatores latentes de medidas de MT e Gf de no mínimo de 0,80 com média de 0,90. Uma série de pesquisas usando técnicas de MEE tem encontrado que a MT demonstra um efeito significativo sobre todas as variáveis dependentes, por exemplo, Fry e

Hale (2000) encontraram correlações significativas de 0,51 entre Gs e MT e de 0,38 entre MT e Gf; Engle e cols. (1999) encontraram correlações de 0,59 entre MT e Gf; Conway, Cowan, Bunting, Theriault e Minkoff (2002) encontraram valores de 0,60 entre MT e Gf, de 0,40 entre Gs e Gsm e 0,18 entre Gsm e Gf; também Sub, Oberauer, Wittmann, Wilhelm e Schulze (2002) encontraram 0,65 entre MT e *g* e 0,79; 0,71; 0,90 respectivamente entre *g* e Gs, Glr, Gf. O autor considera que, em primeiro lugar, esse conjunto de resultados aponta que MT pode ser considerado um fator causal significativo no desempenho de tarefas nas quais as habilidades cognitivas complexas são exigidas, tipicamente as habilidades do segundo estrato da teoria CHC; segundo, que em pelo menos nos três primeiros estudos os resultados têm apontado Gs como um precursor direto para a MT; e terceiro, é necessário que se realizem mais estudos que incluam a matriz completa de fatores do segundo estrato do CHC.

A explicação para a constante força da relação entre MT e as habilidades cognitivas complexas aponta para a exigência de diferentes recursos cognitivos que estão por trás da performance em tarefas de MT. Por exemplo, capacidade de armazenamento, eficiência de processamento, executivo central, processos de domínio específico e atenção controlada. Vários pesquisadores têm hipotetizado que a razão pela qual MT é fortemente associada aos construtos cognitivos complexos (*g* e Gf) é que muitas informações precisam ser ativamente mantidas na memória de trabalho, especialmente quando transformações ativas dessas informações são exigidas por uma tarefa específica. Mesmo quando o processo de transformação é efetivo isso precisa ser realizado dentro dos limites do sistema de memória de trabalho (McGrew, 2003). Outras áreas do conhecimento realizaram pesquisas que exploraram as relações entre a memória de trabalho e outros processos cognitivos utilizando métodos de pesquisa diferentes dos que são geralmente utilizados na pesquisa

psicométrica, especialmente modelos experimentais associados a técnicas de neuroimagens. Esses procedimentos são típicos da pesquisa em neurociências e serão abordados a seguir.

Contribuições das Ciências Cognitivas para compreensão da Inteligência

Segundo Imbert (1998) o campo das Ciências Cognitivas pode ser definido como uma área de pesquisa que se caracteriza pela associação estreita entre as Ciências do Cérebro com a Psicologia, Lingüística, Antropologia, Filosofia e Informática. Apresenta como objetivo principal o estudo da inteligência, desde sua estrutura formal até o seu substrato biológico. Preocupa-se inicialmente com a concepção dos modelos teóricos percorrendo as principais expressões da atividade inteligente sejam elas, psicológicas, antropológicas ou lingüísticas.

A relação entre as Ciências do Cérebro e a Psicologia pode ser ilustrada pela analogia *hardware – software* vs cérebro – mente. Assim, a implementação dos fenômenos cognitivos, em especial da inteligência, depende dos mecanismos cerebrais de forma análoga em que o funcionamento de um sistema computacional depende das características de funcionamento de seus circuitos. O tratamento das informações recolhidas do meio ambiente via sistema sensorial possibilitam, num primeiro nível, estruturar o conhecimento útil. O planejamento entendido como uma função cognitiva que visa decidir sobre o melhor desenvolvimento de uma ação para se atingir um objetivo, depende da combinação de conhecimentos úteis para representar de forma adaptada o ambiente.

Uma das contribuições das Neurociências para as Ciências Cognitivas em geral, e para a Psicologia em particular, pode ser descrita pela função disciplinadora. A partir do conhecimento sobre o funcionamento do cérebro pode-se estabelecer os limites biológicos das teorias sobre os processos cognitivos de modo a preservar um isomorfismo biológico.

As teorias psicológicas ganham força na medida em que conseguem descrever precisamente a maneira como as operações mentais são executadas, na solução de problemas cotidianos, e tais descrições se correlacionam com os conhecimentos sobre as operações que o cérebro realiza diante dos mesmos problemas. Ao passo que, a contribuição das teorias psicológicas é de promover a integração do conhecimento atual sobre o funcionamento do sistema nervoso em um conjunto capaz de explicar detalhadamente o funcionamento mental (Imbert, 1998).

Este ponto de vista está intimamente relacionado a uma visão de que a mente, objeto de estudo da Psicologia, é um produto do funcionamento cerebral. A distinção entre o que representa a atividade neural, que é objeto de interesse da Neurologia, e a atividade neural que é de interesse da Psicologia pode ser visualizada pela analogia da água em ebulição. A água ao ser aquecida só muda de estado físico, passando do líquido ao gasoso, quando a adição de temperatura alcança o limite de aproximadamente 100 graus centígrados. Da mesma forma, uma atividade cerebral passa a ser compreendida como uma atividade mental, quando as ações somadas dos neurônios individuais passam a compor um sistema complexo de redes de neurônios ou assembléia de neurônios (Del Nero, 1997).

Inúmeras pesquisas foram realizadas com o intuito de procurar compreender as relações entre o funcionamento mental em geral e a atividade cerebral subjacente. Um exemplo recente pode ser encontrado no trabalho de Jung-Beeman e cols. (2004), que pesquisaram a atividade neural em pessoas resolvendo problemas verbais por *insight* com o objetivo de descobrir se diferentes processos neurais e cognitivos estão envolvidos na solução de problemas ou se as estratégias de solução (com ou sem *insight*) são diferentes apenas na sua expressão subjetiva. Os sujeitos responderam a um conjunto de problemas verbais e depois a análise do padrão de performance de respostas certas e o hemisfério cerebral envolvido indicavam se eles tinham resolvido com ou sem *insight*. Foram

utilizados dois procedimentos de medida enquanto os sujeitos respondiam, a Imagem por Ressonância Magnética Funcional (IRMF) e o Eletroencefalograma (EEG). Os resultados mostraram dois correlatos neurais objetivos ligados ao *insight*. A IRFM mostrou um incremento na atividade do giro temporal superior anterior no hemisfério direito para respostas de *insight* quando comparadas a respostas de não *insight*. Já o EEG mostrou um inesperado disparo de alta frequência de ondas gama na atividade neural da mesma área iniciando aos 0,3 seg. para soluções com *insight*. A área temporal anterior direita é freqüentemente associada com a criação de cruzamento de conexões para informações que apresentam relações distantes durante a compreensão. Tais resultados sugerem que embora toda solução de problemas envolva uma ampla rede cortical compartilhada, os flashes inesperados de *insight* acontecem quando distintos processos neurais e cognitivos se empenham na solução de modo que permitem a visualização das conexões.

No que concerne especificamente à inteligência, Lezak (1995) relata que, segundo o modelo da neuropsicologia clássica, os primeiros conceitos sobre inteligência, definidos pela psicométrica tradicional insipiente, apresentaram uma aplicação limitada para a avaliação neuropsicológica. Essa proposição estava baseada em um conhecimento que até então era consensual na neuropsicologia. Os conceitos sobre inteligência eram genéricos e permitiam que toda a atividade cognitiva pudesse ser encerrada neles. As primeiras concepções sobre inteligência foram tão válidas para a psicologia quanto a expressão “A força do solo” é para o agricultor. Apesar de plausível, é algo bastante arcaico. O funcionamento cognitivo como um atributo continuum, observado nos indivíduos saudáveis de cérebro intacto, reflete a harmonia de várias funções trabalhando de forma integrada.

O conceito de fator *g* de Spearman tratava a inteligência como uma variável unitária, ou seja, a “energia mental” que podia ser comparada à “força física”. O conhecimento sobre a natureza da inteligência era ainda incipiente e uma característica

marcante é a de que a inteligência apresentava um crescimento ao longo do curso normal de desenvolvimento e um declínio relacionado à dano no tecido nervoso. Tal característica chamou a atenção para a existência de uma relação entre os eventos do sistema nervoso e o desempenho nas tarefas comumente usadas para medir inteligência.

Porém, a relação entre os danos no sistema nervoso e a capacidade de realização de determinadas tarefas não era tão evidente, e os escores nos primeiros testes para avaliar a inteligência não conduziam a uma predição relacionada, por exemplo, com o tamanho de uma lesão cerebral. As pontuações apresentadas na forma de *Quociente Intelectual*, usado pelas primeiras baterias para medir a inteligência, representavam a atividade de uma multiplicidade de funções cognitivas avaliadas pelos itens desses testes. Dessa forma, cada teste apresentava um conjunto de itens que podiam ser separados em várias classes específicas. A resolução correta dos itens de cada classe exigia habilidade em funções diferenciadas, que na prática da avaliação neuropsicológica precisavam ser dissociadas (Lezak, 1995).

Nos anos que se sucederam, as concepções sobre inteligência, especialmente derivadas das pesquisas psicométricas, incorporaram as informações oriundas da pesquisa sobre dissociação de função em neuropsicologia. As habilidades que compunham o escore de um teste de fator *g* foram gradualmente separadas em capacidades específicas que passaram a compor o modelo conceitual sobre inteligência em psicologia, por exemplo a teoria CHC. Desta maneira, a visão da neuropsicologia de que o conceito de fator *g* apresentava uma aplicação limitada deu lugar a um projeto de pesquisas integrando metodologias que passam a permitir uma melhor compreensão dos fenômenos cognitivos em geral e da inteligência humana em particular (Deary, 2000, 2001).

Segundo Deary (2000, 2001) os componentes cognitivos ou elementos mentais, que estão por trás das diferenças de inteligência, podem se tornar conhecidos a partir de no

mínimo três tipos ou fontes de procedimentos de investigação. Primeiro, a pesquisa psicométrica, especialmente o artifício de decomposição dos itens em partes menores e elementos específicos que possam ser associados a atividades cognitivas, especialmente as unidades de processamento. As propostas teóricas da psicologia cognitiva capazes de explicar a associação entre os elementos específicos dos itens e as unidades de processamento cognitivo permitiram a inferência aos fatores de complexidade. Um procedimento, em psicometria, que objetiva realizar essa operação é o *Linear Logistic Traits Models* (LLTM) (Fischer, 1973) foi detalhado anteriormente. O segundo conjunto de procedimentos é oriundo da psicologia cognitiva experimental, mais especificamente as pesquisas sobre tempo de reação, que envolvem a manipulação da relação estímulo-resposta pela adição ou subtração de componentes ou elementos de processamento mentais. A associação entre o tempo de reação e o desempenho em testes de inteligência não é recente, e tem sido encontrada desde as primeiras décadas do século vinte. Tais resultados sugerem que pessoas com alta inteligência tendem a apresentar tempos de reação menores e menos variáveis. O terceiro conjunto de procedimentos diz respeito ao processamento de informação sensorial individual, comuns em pesquisa psicofísica. Por exemplo, o processamento de informação visual ou auditiva. Alguns resultados indicam uma associação entre o tempo de inspeção e escores de testes de inteligência da ordem de $r = 0,40$.

A evolução da concepção da inteligência para um construto multifacetado em associação com a concepção do cérebro, como um sistema massivamente interconectado, no qual regiões especializadas contribuem com o desempenho em várias tarefas cognitivas, esclareceu algumas dúvidas sobre a natureza do fator *g*. Especialmente a questão da baixa correlação entre o desempenho nas baterias padronizadas para medir inteligência e o dano cerebral, ou a atividade de um local específico do cérebro.

Pesquisas recentes, adotando combinação de metodologias, encontraram fortes associações entre funções cognitivas e estruturas cerebrais específicas. Por exemplo, Duncan, Burgess e Emslie (1995) pesquisaram as relações entre as funções do lobo frontal (funções executivas) e a inteligência geral e especificamente inteligência fluida, em pacientes com lesões no córtex pré-frontal. Particular interesse foi dado aos casos nos quais os pacientes lesados manifestavam prejuízos cognitivos e, no entanto, apresentavam escores de QI no *Wechsler Adult Intelligence Scale* (WAIS) preservados. Comumente casos desta natureza, na prática clínica da neuropsicologia, foram usados para desafiar a existência de uma base neural para a inteligência geral (Fator *g*).

A hipótese aventada pelos autores, para explicar os casos, foi de que os resultados podiam estar relacionados ao tipo de medida que estava sendo usada. Nesse caso, testes padronizados cuja resolução adequada de seus itens depende, em grande parte, de conhecimento prévio (*Gc*). Para testar essa hipótese, na pesquisa foram examinados, A) dois pacientes que apresentavam lesão frontal e escores de QI no WAIS considerados superiores e preservados; B) um terceiro paciente que apresentava lesões mais extensas, lobo frontal e outras áreas, mas que apresentava um perfil cognitivo semelhante; C) um grupo misto de cinco pacientes portadores de lesões focais em áreas posteriores; D) também foi incluído na análise um grupo controle composto por sujeitos normais com perfil sócio demográfico e cognitivo semelhante ao dos pacientes do grupo clínico. Os sujeitos foram avaliados, e tiveram seus escores de QI estimados, por meio do WAIS, que se trata de uma medida relacionada a *Gc* e do *Cattell's Culture Fair* (CCF), mais relacionada a *Gf*. Os resultados encontrados mostraram que os pacientes com lesão frontal não apresentaram diferença no QI do WAIS em relação aos respectivos controles, mas apresentaram diferença significativa no QI do CCF. Em relação aos pacientes com lesões não frontais, esses não apresentaram diferenças significativas nos QI's estimados nos dois

testes. Os autores consideraram que os resultados foram favoráveis às hipóteses iniciais e sugerem que o *g* pode ser um reflexo da atividade ampla de funções frontais.

Em outra pesquisa Duncan, e cols (2000), tiveram como objetivo demonstrar a existência de *g* e suas relações com a atividade do lobo frontal. Se por um lado, a hipótese do *g*, proposta por Spearman, é que por trás de acentuada habilidade em diversas modalidades cognitivas existe um processo mental mais expressivo. Por outro lado, uma hipótese proposta por Thomson, é que a execução adequada de qualquer tarefa exige a contribuição de um conjunto de funções de processamento de informação. Portanto, o padrão de correlações encontrado pode ser explicado pelo fato de que qualquer conjunto de tarefas usadas para gerar as matrizes de correlações em questão compartilhava parte de seus componentes. Os dois autores derivaram suas hipóteses da análise de matrizes de correlações construídas a partir de dados oriundos de respostas a um conjunto de testes envolvendo várias funções cognitivas. Os resultados mostravam um padrão de correlações positivas. Quanto melhor o desempenho dos sujeitos em um conjunto de tarefas específicas, melhor também era seu desempenho geral.

Para testar as hipóteses acima, os autores compararam os resultados da Tomografia por Emissão de Pósitrons (PET – SCAN) de sujeitos que realizaram tarefas de conteúdo espacial, verbal e percepto-motor divididas em dois níveis de associação com o *g*. Foram adotadas correlações acima de 0,55 como altas associações e correlações abaixo de 0,41 como baixas associações. Os resultados apontaram que, em contraste com a visão proposta por Thomson, de que *g* reflete uma larga amostra de funções cognitivas, as tarefas altamente associadas a *g* não mostraram o recrutamento difuso de muitas regiões cerebrais. Em vez disso, elas estão associadas com o recrutamento seletivo do córtex frontal lateral em um ou nos dois hemisférios. Apesar das tarefas serem de conteúdo muito diferente, e contrastadas em alta e baixa associação com *g*, o recrutamento do córtex frontal lateral foi

similar em cada caso. Esses resultados sugerem a plausibilidade do *g*, e que ele deriva de um importante e específico sistema de controle de diversas formas de comportamento, localizado no córtex frontal.

Também Geake e Hansen (2005), estudaram os correlatos neurais da inteligência por meio do mapeamento da atividade cerebral, usando procedimento de Imagens por Ressonância Magnética Funcional (IRMf) enquanto os sujeitos respondiam a tarefas envolvendo analogias. Participaram da pesquisa 12 sujeitos adultos destros, sendo 4 homens, com idades entre 18 e 54 anos, estudantes de pós-graduação com inteligência acima da média segundo o “National Adult Reading Test” (NART). Sendo que os escores do NART estão teoricamente associados à inteligência cristalizada (*Gc*). Os pesquisadores conjecturaram que o desempenho em analogias sustenta toda atividade intelectual, incluindo *Gc* como categoria emparelhada. Em seu procedimento eles utilizaram 55 itens de analogias formadas por seqüências de letras com quatro alternativas de resposta. Cada alternativa recebia uma nota que variava de 1 à 5 pontos, dependendo das transformações sofridas pelos elementos em cada alternativa de solução do problema, por exemplo, cópia de letras, inserção de letras, mudar letras de lugar, etc. Os sujeitos foram solicitados a responder aos problemas enquanto era feito o mapeamento da atividade cerebral por meio do IRMf. Os resultados encontrados apontaram uma rede de trabalho formada por áreas frontais do cérebro ativadas na resolução das analogias, a saber, região pré-frontal lateral superior e lateral inferior. Tais descobertas sugeriram a interpretação de que a inteligência geral esta associada a tarefas de analogia, primeiro, porque existe uma sobreposição, entre as regiões de ativação, associadas com a resolução de tarefas como analogias usando seqüências de letras e usando analogias simples. Segundo, pelo grau de similaridades entre as regiões de ativação associadas com a resolução de problemas semelhantes a essas

analogias, e os itens altamente correlacionados com fator g. Também porque a ativação do giro frontal medial esquerdo aparece correlacionada com os escores do NART.

Particularmente no que diz respeito à capacidade de raciocínio, as pesquisas psicométricas vêm apontando correlações entre g e Gf com capacidade de Memória de Trabalho (MT), especialmente com o subcomponente de Atenção Executiva (AE) (McGrew, 2003). Da mesma forma, as pesquisas em Neurociências vêm sugerindo a ligação entre as funções cognitivas envolvidas na MT e a atividade do cortex pré-frontal. Esses resultados sugerem que Gf é um construto que deve ser investigado a partir da combinação de abordagens de pesquisas, porque as diferenças individuais de desempenho nas tarefas de raciocínio, podem ser relacionadas a uma variedade de demandas de processamento presentes nestas tarefas, ligadas ao controle de atenção e capacidade de memória (Kane, 2005).

Alguns exemplos de como as pesquisas em neurociências vêm abordando as relações entre os construtos g, Gf, MT, AE são apresentados a seguir. Segundo Kane (2005) as contribuições da memória para a habilidade intelectual em geral são conhecidas há bastante tempo, por exemplo, Baddeley e Hitch (1974) discutem a idéia de que a memória imediata é útil para a cognição complexa na medida em que permite que estímulos críticos, idéias ou respostas produzidas se mantenham acessíveis por tempo suficiente para influenciar no desempenho. Os mesmos autores (Baddeley & Hitch, 1974; Baddeley, 1986) propuseram um modelo de Memória de Trabalho (MT) no qual o domínio específico de memória interage com recursos de atenção ou estruturas de controle (executivo central). Os componentes de atenção e memória desse sistema trabalham juntos melhorando operações de processamento e armazenagem. Uma medida da capacidade de Memória de Trabalho foi proposta por Daneman e Carpenter (1980), na qual os participantes liam em voz alta ou escutavam uma série de duas até seis sentenças e depois deveriam relembrar a palavra final

de cada sentença. Dessa forma, foi embutido, dentro de uma tarefa tradicional de *span* de memória de curto prazo, um componente secundário de leitura que acrescenta uma demanda de atenção executiva além dos processos de armazenagem.

As tarefas de Memória de Trabalho (MT) geralmente estão mais associadas às diferenças individuais de habilidades cognitivas do que as tarefas construídas para medir a capacidade da Memória de Curto Prazo (MCP). Uma explicação para isso é que nas tarefas de MT encontra-se embutido um componente de habilidade de controle da atenção que exige a manutenção da informação relevante acessível, mesmo diante de interferência e estímulos distratores. O sucesso nesse tipo de tarefa exige que os sujeitos mantenham em mente os elementos alvo, ou seja, a informação relevante, apesar da interferência proativa presente em outros elementos dos itens que são irrelevantes para a resposta. A manutenção desses estímulos em mente é desafiada intermitentemente pela mudança no foco da atenção entre os estímulos presentes na memória e os processos secundários das tarefas. Os processos de controle executivo ajudam a manter acessíveis os elementos relevantes em face à interferência, provocada pelos elementos distratores, e as mudanças no foco de atenção. Desta maneira, surge a hipótese de que a capacidade de atenção executiva é responsável pela ampla utilidade preditiva das tarefas de memória de trabalho em relação aos processos cognitivos complexos (Kane, 2005).

Argumentos empiricamente sustentados sobre essa hipótese, de que as demandas de atenção e não as de armazenagem presentes nas tarefas de memória de trabalho são as responsáveis pelo seu amplo poder preditivo de processos cognitivos complexos, foram apresentados por Engle, Tuholski, Laughlin e Conway (1999). Os autores investigaram a relação entre tarefas de MT, MCP e tarefas clássicas para medida de inteligência fluida, como as Matrizes progressivas de Raven e o *Cattell's Culture Fair Tests* (CCT). Os resultados da análise fatorial confirmatória indicaram que os fatores derivados das tarefas

de MT e de MCP são separados, porém correlacionados, permitindo a interpretação de que as tarefas de MT exigem em alguma medida a capacidade de atenção e não apenas simples armazenagem. Ao lado disso, a hipótese crítica da atenção foi testada por meio da Modelagem de Equações Estruturais. Depois de removida a variância comum para as variáveis latentes, MT e MCP, que representa a parte de exigências de armazenagem, a variância residual do fator de MT, que representa uma exigência única de atenção executiva mostrou correlações com Gf. Considerando as duas análises, os resultados mostraram que a variância residual de MT explica aproximadamente de 25% a 35% da variância em Gf, e a variância comum com MCP explica 9% ou menos da variância de Gf. Assim, o fator crítico na relação entre memória de trabalho e inteligência fluida parece ser a atenção executiva.

Um outro conjunto de evidências de ligação entre a capacidade de memória de trabalho e atenção executiva pode ser encontrado em estudos como o conduzido por Kane e Engle (2000). Os autores avaliaram um conjunto de sujeitos, divididos em grupos de alto e baixo desempenho em tarefas de MT, com uma tarefa de atenção do tipo Brown-Peterson, que é composta por três listas de dez palavras cada, todas pertencentes ao mesmo campo semântico. Segundo os autores, a interferência proativa é comumente encontrada em tarefas semelhantes, como a lembrança do conjunto de itens anterior, interferindo na lembrança dos itens de cada lista apresentada subsequentemente. Nesse estudo, alguns participantes tiveram sua atenção dividida tendo de lembrar continuamente uma seqüência apontada. A lógica é de que a resistência à interferência proativa requer controle da atenção, neste caso a vulnerabilidade à interferência deve estar associada a capacidade de MT e demandas de atenção dividida.

Os resultados mostraram que entre os indivíduos que realizaram a tarefa na condição simples, sem atenção dividida, os grupos de alta e baixa habilidade em MT não diferem na lembrança da lista 1, porém os indivíduos do grupo de baixa habilidade se

mostraram mais vulneráveis a interferência proativa na lista 3, apresentando uma diferença de desempenho em torno de 20%. Ao lado disso, entre os indivíduos que realizaram a tarefa na condição complexa, com atenção dividida, o grupo de alta habilidade em MT tornou-se tão suscetível à interferência proativa quanto o grupo de baixa habilidade em MT, enquanto o grupo de baixa habilidade manteve sua suscetibilidade inalterada. Esses resultados sugerem que a capacidade de atenção executiva está associada ao alto desempenho em tarefas de MT, pois o grupo de alta habilidade em MT apresentou diminuição no desempenho em tarefas complexas quando comparados com tarefas simples. Também, o fato de que o desempenho do grupo de baixa habilidade e MT não teve seu desempenho prejudicado nas tarefas complexas quando comparado com o desempenho em tarefas simples, sugere que esse grupo não utiliza processos de atenção executiva para combater a interferência proativa.

As pesquisas apontadas aqui, especialmente aquelas usando procedimentos de neuroimagem, têm servido para esclarecer quais estruturas cerebrais são importantes para compreender a inteligência geral. Os seus resultados representam as principais contribuições da área de Neurociência Cognitiva para esse tema. Porém, os resultados mais importantes podem ser descritos como a combinação de procedimentos de diversas áreas para compreensão da inteligência enquanto um fenômeno humano. Os trabalhos futuros devem manter a tradição, inaugurada recentemente nessa área, de integrar os conhecimentos teóricos e metodológicos da Psicologia Cognitiva, Comportamental e Psicometria com a Neurociência, pois são promissores os avanços na compreensão não somente de como a mente funciona, mas também como e por que esse funcionamento se distingue em cada indivíduo.

Contribuições da Psicologia Cognitiva na compreensão do Raciocínio.

Segundo Sternberg (1983) a psicometria se dedica à tentativa de encontrar os fatores subjacentes às habilidades humanas, tem por objetivo prever resultados e, em função do método utilizado, menor capacidade de elucidação dos conceitos teóricos importantes para a compreensão dos construtos. Portanto, de se entender as razões que permitem a predição dos resultados. A compreensão sobre os processos subjacentes a cada fator de habilidade descrito foi limitada pela metodologia empregada pelo modelo psicométrico tradicional. A aproximação do modelo psicométrico com a psicologia cognitiva permitiu avanços na compreensão da inteligência, bem como na possibilidade de estudo empírico. Os resultados dessa aproximação pôde gradualmente pôr termo à argumentação teórica de que o modelo psicométrico se preocupa apenas com o resultado final e não com o processo cognitivo que leva às respostas.

A psicologia cognitiva pode ser apontada como uma das disciplinas que compõe o projeto científico das Ciências Cognitivas. Um de seus objetivos é sistematizar os processos mentais envolvidos na cognição em geral tais como, sensação, percepção, memória, entre outros. Sobre a inteligência, ela visa identificar e descrever os processos cognitivos que estão envolvidos na resolução de problemas. Utiliza o método experimental de análise e apresenta uma preocupação de se compreender o próprio funcionamento cognitivo, pretendendo descrever a natureza da Inteligência. Um dos seus preceitos é que o conhecimento das capacidades em termos de processos e a sua operacionalização é a base para avaliações adequadas que vão permitir intervenções individualizadas.

A psicologia cognitiva aqui relatada adota o modelo do processamento de informação, e também se baseia na analogia cérebro – mente vs *hardware* – *software*. Nesse caso a analogia é interpretada fazendo uma comparação entre o sistema de

processamento de informação humano (cérebro – mente) ao sistema de processamento de informação computacional (hardware – software). Uma dedução formal é que a mente humana obedece a uma série de leis e regras de funcionamento tal como ocorre com os softwares. Portanto, ao serem descobertos os detalhes de tal funcionamento, podem ser criadas estratégias de intervenção mais adequadas aos problemas comumente encontrados. Os princípios adotados pela psicologia cognitiva servem de base para o estudo dos processos cognitivos que então implicados em cada uma das habilidades propostas pelos fatores amplos descritos pela teoria CHC.

Um modelo bastante difundido na perspectiva da psicologia cognitiva é a teoria triárquica da inteligência proposta por Sternberg (1985). Segundo o autor, essa proposta teórica trata principalmente dos indivíduos e suas relações com seu mundo interno e externo. Dessa maneira, procura explicar de forma integrada as relações existentes entre a inteligência e os mecanismos mentais subjacentes ao comportamento inteligente, como também as formas como os indivíduos usam esses mecanismos no cotidiano, de modo a promover o ajustamento ao meio.

Portanto, para uma descrição da inteligência que privilegie todas essas facetas, a teoria triárquica é subdividida em três sub teorias. A primeira, denominada de sub teoria contextual, é composta por três processos, a adaptação, a modelação e a seleção. A *adaptação* é o processo a partir do qual o sujeito tenta se adequar ao meio em que vive buscando um melhor ajustamento. Considerando que as exigências para adaptação diferem de um meio para outro, este aspecto reflete os efeitos da cultura sobre a inteligência. A *modelação* consiste na tentativa do sujeito de reorganizar o meio conforme suas necessidades. Nem sempre ocorre quando o sujeito não consegue adaptar-se ao meio, e pode ocorrer antes de uma tentativa frustrada de adaptação. A *seleção*, por sua vez, consiste em uma busca por ambientes mais adequados (Sternberg, 1985).

A segunda é a sub teoria de dupla faceta ou experiencial, que é composta por dois processos. Primeiro a *capacidade de processar informações novas*, pode envolver a compreensão de tarefas novas, e a execução de ações dirigidas pela compreensão que se faz dessas tarefas para a solução adequada da mesma. Segundo, a *capacidade de automatizar o processamento da informação*, algumas tarefas complexas parecem ser possíveis apenas porque uma parte das operações mentais requeridas pode ser realizada com a diminuição do investimento de recursos da atenção controlada. Por conseguinte, a dificuldade em automatizar os processos envolvidos na realização de determinadas tarefas pode prejudicar o desempenho final. Os dois aspectos dessa sub teoria estão correlacionados, pois a capacidade de administrar os recursos da atenção requeridos por parte das operações mentais, libera recursos para serem usados para analisar situações novas.

A terceira é denominada de sub teoria componencial e visa descrever os processos mentais implicados na atividade inteligente a partir da tentativa de isolamento de suas unidades básicas. Essa unidade básica foi denominada de *componente de processamento* ou simplesmente *componente* e possui três importantes propriedades que a princípios são independentes entre si. São elas, duração, dificuldade - que sinaliza a possibilidade de ser executada erroneamente - e a probabilidade de execução (Sternberg, 1983).

Os componentes de processamento são divididos em três categorias e apresentam três tipos de funções diferentes. A primeira categoria é a dos *metacomponentes*, caracterizados como processos executivos de ordem superior utilizados para planejar as atividades a serem realizadas, monitorá-las enquanto estão sendo realizadas, e avaliar o resultado final. Incluem o reconhecimento da existência de um problema, a decisão sobre a natureza do problema, e a seleção de um conjunto de processos hierarquicamente inferiores necessários à resolução do problema, a seleção de uma estratégia para combinar os componentes inferiores, e de uma representação mental adequada à operacionalização dos

processos para a resolução do problema. Resumindo, os metacomponentes administram a atribuição de recursos mentais, a monitoração do processo de resolução e a avaliação dos resultados.

A segunda categoria é a dos *componentes de performance*, considerados processos de ordem inferior responsáveis pela execução das várias estratégias de performance das tarefas selecionadas pelos metacomponentes. Um exemplo de componente de performance é a *inferência*, responsável pela identificação de uma figura particular pelas similaridades ou diferenças entre um conjunto de objetos. Muitos dos componentes de performance são específicos para determinado tipo de tarefa (Sternberg, 1983).

A terceira categoria é a dos *componentes de aquisição de conhecimento*, conforme a expressão do nome, esses componentes estão envolvidos nos processos de aprendizagem de informação nova e sua estocagem na memória. São de três tipos diferentes, relevantes para a aquisição de conhecimento processual e declarativo, supostamente, aplicados em todos os domínios de conhecimento. A *codificação seletiva*, relacionada à capacidade de separar ou selecionar a informação relevante da que é irrelevante; a *combinação seletiva*, relacionada ao agrupamento ordenado e plausível de toda a informação relevante codificada previamente de modo a maximizar a coerência interna; e a *comparação seletiva*, envolve o estabelecimento de relacionamentos entre as informações recentemente adquiridas com a informação existente estocada na memória, de modo a maximizar o encadeamento de uma estrutura de conhecimento formada recentemente com uma estrutura formada previamente (Sternberg, 1983).

Os componentes apresentam ainda, níveis de generalidades entre si. Podem ser *gerais*, quando são utilizados em todas as tarefas de um dado universo; de *classe*, quando são utilizados em um sub conjunto de tarefas de um universo; e *específicos*, quando são usados em uma única tarefa dentro de um universo. Embora os componentes possam ser

isolados experimentalmente, na prática eles funcionam de forma integrada e de difícil identificação. Os metacomponentes ativam os componentes de performance e os componentes de aquisição de conhecimento que por sua vez fornecem o *feedback* para os metacomponentes.

Os princípios da teoria triárquica da inteligência foram usados para tentar descrever quais processos mentais estavam implicados na resolução adequada de determinados tipos de tarefas. Um conjunto de tarefas fortemente associadas ao raciocínio é a solução de analogias. Segundo Mulholland, Pellegrino e Glaser (1980), dois estudos nos quais podem ser encontradas as bases teóricas e empíricas para o exame dos componentes de desempenho em tarefas de solucionar analogias de conteúdo figural ou geométrico, são de Evans (1968) e Sternberg (1977). A sub teoria componencial propõe que alguns processos cognitivos comuns são aplicados na resolução das tarefas de raciocínio analógico com o formato, A está para B assim como C está para D. São eles, *codificação* dos termos individuais da analogia; *inferência* da relação entre o primeiro e o segundo termo; *mapeamento* da relação entre o primeiro e o terceiro termo; *aplicação* dos resultados dos processos de mapeamento e inferência para gerar um quarto termo que é usado para avaliar as alternativas de resposta apresentadas. Ainda há um componente opcional de *justificativa* que é usado para selecionar entre alternativas quando nenhuma parece ser a resposta ideal; e o processo de *responder* indicando a escolha de uma alternativa de resposta.

A proposta teórica advoga que os termos de uma analogia são codificados como lista de categorias. Comparações entre as características dos termos A e B, produzem uma lista de mudanças nas características que ocorrem de A para B. Então é mapeada uma comparação de características entre os termos A e C, considerando as mudanças mapeadas de A para B, de modo que aplicadas ao termo C possam ser generalizadas para um termo D ideal. Assim as alternativas são analisadas até que se encontre um termo D no qual esteja

embutido um conjunto de mudanças semelhantes as que ocorreram de A para B, que melhor represente as mudanças apropriadas do termo C.

Conforme a descrição de Mulholland, Pellegrino e Glaser (1980), foram testados muitos modelos representando diferentes concepções sobre a natureza do processamento de codificação de características e comparação de componentes. Por exemplo, o processamento pode ser *serial*, que dizer, uma operação é realizada de cada vez enquanto isso outros processos ficam inativos; ou *paralelo*, quer dizer, vários processos ocorrem ao mesmo tempo. E também, se são *exaustivos*, ou seja, depois de iniciados não podem ser interrompidos até que todas as operações sejam executadas; ou *auto-determinados*, quer dizer, que nem todas as operações precisam ser executadas. Os resultados sugeriram que a codificação e a inferência de mudanças de características nos termos A para B é *serial* e *exaustivo*, ao passo que o mapeamento e a aplicação das características respectivas aos termos C e D são *seriais* e *auto-determinadas*.

O ponto fraco da teoria componencial é a idéia de que o processo de codificação é exaustivo e que a ativação e representação de uma característica associada com um dado termo da analogia são independentes do contexto. Durante o processo de codificação de uma nova palavra ou figura em um problema de analogia, um elemento presente nessa operação, inicialmente ou subseqüentemente, deve proporcionar o contexto para um adequado nível de representação do item na memória. Os autores consideram que o cerne da pesquisa e das concepções teóricas de Sternberg estavam voltados para a comparação de processos e não sobre as diferenças de conteúdo usados em diferentes analogias. Portanto, a natureza do processo de codificação de características não foi detalhado, tão pouco, foram feitas descrições da codificação de características presentes em diferentes tipos de analogias. A noção de mudanças de características foi operacionalmente definida em termos da classificação subjetiva da distância entre os termos da analogia. Uma grande

distância entre dois termos era interpretada como refletindo a ocorrência de um grande número de mudanças de características entre eles (Mulholland, Pellegrino & Glaser, 1980).

Por outro lado, no trabalho de Evans (1968), podem ser encontradas especificações das concepções de processamento e representação necessárias para solução de analogias com figuras geométricas. Ele construiu um programa de computador que era capaz de resolver um conjunto de analogia com figuras geométricas. Os algoritmos básicos de programação utilizados foram usados para desenvolver e operacionalizar uma teoria que permitiu uma descrição detalhada da representação dos termos individuais de uma analogia e as regras ligadas a esses termos. O programa apresentava três importantes fases de processamento, a) *descrição da transformação das figuras*, cuja operação principal é a decomposição das partes que compõe cada termo de uma analogia em sub-partes ou elementos; b) *combinação de regras*, determinação das transformações que relacionam os elementos nos pares de termos A – B; c) *diferença de medições*. Primeiro, ocorre a decomposição dos termos A e B e a comparação das figuras para se determinar os elementos que são comuns a ambas. Após a identificação dos elementos que constituem os termos A e B, esses elementos são emparelhados, e então é derivado um conjunto de regras de transformações lógicas e espaciais relacionadas a eles. Em seguida, é feita uma comparação entre a informação apresentada nos termos A e B, com a informação apresentada em cada par de termos formados por C, e as respectivas alternativas de D.

No que diz respeito à organização e ordenação que representam um nível geral de estratégias de solução para problemas em forma de analogias, podem ser apontadas diferenças entre a teoria componencial e a teoria derivada do trabalho de Evans. Por exemplo, a primeira apresenta um procedimento de solução baseado em “inferência – aplicação – teste” ao passo que a segunda apresenta um procedimento de solução baseado em “inferência – inferência – comparação”. Por outro lado, em um nível mais detalhado dos

componentes individuais de processamento as duas teorias apresentam similaridades que podem ser reunidas em três grandes grupos. Para a teoria componencial os processos podem ser agrupados em (1) descoberta do atributo “codificação”, (2) comparação de atributos “inferência, mapeamento e aplicação”, e um processo opcional (3) “justificativa”; que correspondem aproximadamente aos seguintes agrupamentos da teoria de Evans (1) Padrões de comparação e decomposição, (2) geração de regras e combinação, e o também opcional (3) discriminação de regras. Porém, o mais importante é que provavelmente as similaridades são uma função, tanto das características do processamento de informação, priorizado pela teoria componencial, como das características internas dos problemas, privilegiada na teoria de Evans, que necessariamente estão ligadas durante a solução de problemas de analogia. Desta maneira, os aspectos específicos do conteúdo dos itens são os elementos individuais usados para construir os termos de uma analogia ao passo que as transformações lógicas e espaciais aplicadas a esses elementos são usadas para construir as regras que explicam as mudanças nos termos da analogia (Mulholland, Pellegrino & Glaser, 1980).

Partindo desse ponto de vista sobre o processo cognitivo implicado na resolução de problemas de analogia, Mulholland, Pellegrino e Glaser (1980) buscaram relacionar a performance nas tarefas com os modelos de processamento de componentes de informação nos itens, especificar e avaliar uma teoria sobre o desempenho em analogias e identificar fontes potenciais de diferenças individuais na habilidade de solução de analogias. Para isso foram construídos dois conjuntos de analogias experimentais usando figuras geométricas aplicando os princípios teóricos descritos acima. O primeiro conjunto com 230 itens foi constituído por analogias verdadeiras compostas respectivamente por um, dois e três elementos que sofria, nenhuma, uma, duas e três transformações respectivamente. O segundo conjunto, também com 230 itens, foi formado por analogias falsas divididas em

um-terço de itens repetindo um ou mais elementos com o termo D da analogia incorreto e dois-terços contendo nos termos C ou D uma ou mais transformações corretas relacionadas com a primeira parte da analogia (termos A e B). Esse conjunto de analogias experimentais foi aplicado juntamente com o subteste de analogias geométricas do Teste de Habilidades Cognitivas (em inglês: *Cognitive Abilities Test*) de Thorndike e Hagen (1974) a uma amostra composta por 28 estudantes universitários igualmente distribuídos em relação ao sexo e com idade entre 18 e 28 anos. Foi solicitado dos sujeitos que marcassem as analogias como falsas ou verdadeiras e registrado o tempo de resposta.

Os resultados mostraram que o aumento no número de elementos aumenta sistematicamente o tempo de resposta. O tempo de processamento é constante dentro de cada condição de transformação, porém varia em função do número de transformações, geralmente com incremento por elemento e por número de transformações. Esses dados são coerentes com as propostas teóricas de que os elementos que compõem os termos das analogias são processados um a um e de que existe aditividade sobre os elementos e as transformações sofridas por eles. As conclusões gerais indicam que o desempenho em tarefas de analogia é designado sistematicamente pelos elementos e transformações aplicadas nas estruturas dos itens.

Adiante serão tratados os principais aspectos da Bateria de Provas de Raciocínio com o objetivo de desmembrar os itens de cada prova em uma cadeia de processos cognitivos básicos, os componentes, descrevendo mais detalhadamente o raciocínio indutivo. Levando em conta que o formato da maioria dos subtestes - analogias nas provas RA e RV e séries nas provas RN e RE - correspondem a problemas de raciocínio indutivo, isto é, problemas nos quais se deve analisar as informações para descobrir as regularidades inferindo as regras gerais e depois aplicá-las na produção de idéias análogas.

Capítulo III – A Bateria de Provas de Raciocínio

Dentre os instrumentos disponíveis no mercado brasileiro para avaliação da inteligência encontra-se a bateria de provas de raciocínio (BPR-5) que começou a ser desenvolvida por Almeida e Primi em 1995 com a aplicação da bateria de provas de raciocínio diferencial (BPRD) na população Brasileira. O BPRD é amplamente utilizado em Portugal desde a década de 80, o que permitiu reunir considerável quantidade de informação sobre sua precisão e validade. Esse primeiro estudo com o BPRD produziu resultados que levaram os autores a propor uma forma reduzida da bateria (Almeida & Primi, 1998).

A essa forma reduzida do BPRD deu-se o nome de BPR-5, que é atualmente composto de cinco provas, cada uma possui conteúdo destinado a avaliação de diferentes processos de raciocínio, a saber, prova de raciocínio abstrato (RA), raciocínio verbal (RV), raciocínio espacial (RE), raciocínio numérico (RN), e raciocínio mecânico (RM). Cada uma delas pode ser analisada segundo os conceitos propostos pelo modelo CHC, desenvolvido por Carroll em 1993 e 1997, atualmente usado para entendimento das capacidades cognitivas.

A prova RA é a mais ligada a *gf*, é composta de itens que envolvem analogias entre figuras geométricas nas quais existe uma relação entre os dois primeiros elementos, a relação deve ser descoberta e aplicada ao terceiro elemento para que se descubra qual alternativa apresenta as transformações necessárias para completar a série. A prova RV é ligada a *gf* e também a *gc*, é composta por analogias como em RA, porém envolvendo palavras. A tarefa do sujeito é descobrir qual a relação entre o primeiro par de palavras para que se descubra qual será a quarta palavra que completa a série.

A prova RE está relacionada aos fatores gf e gv , e é composta por vários cubos tridimensionais em movimentos que podem ser constantes ou alternados. O objetivo proposto para o sujeito é descobrir o movimento pela análise das faces e então indicar qual a posição do cubo que completa a série. A prova RN está relacionada aos fatores gf e gq , e é composta por uma série de números relacionados por uma regra aritmética que deve ser descoberta para que o sujeito possa indicar corretamente quais os dois últimos números da série. A prova RM está relacionada a gf e ao conhecimento mecânico prático pouco associado a conhecimento acadêmico. É composto por gravuras que apresentam problemas práticos de conteúdo físico-mecânico para os quais o sujeito deve descobrir qual alternativa apresenta a melhor resposta ao problema apresentado pela figura (Almeida & Primi, 1998).

De um modo geral pode-se dizer que o BPR-5 é um teste de avaliação da Inteligência Fluida gf que se trata do fator mais fortemente ligado ao *fator g*. Mas a resolução adequada de seus itens depende, ainda que em menor proporção, da influência de outros fatores. Estudos detalhados de precisão e validade realizados para sua padronização em concomitância a estudos realizados em Portugal com mesmo fim, oferecem possibilidade de comparação entre as duas amostras normativas (Almeida & Primi, 1998).

Desde a sua publicação no Brasil, em 1998, muitas pesquisas foram realizadas nos mais variados contextos de utilização com fins de se verificar as possibilidades interpretativas dos seus escores. Primi, Guntert e Alchieri (2000) estudaram a relação entre quatro provas da bateria (RA, RE, RN e RV), o teste R-1 e o Teste de Raciocínio Não Verbal (TNRV) com indicadores de capacidade intelectual oferecidas pelo teste projetivo Zulliger. Foram criados dois grupos de desempenho extremo, G1 formado por pessoas com percentil abaixo de 25 ($P < 25$) nas provas de raciocínio e G2 por pessoas com $P > 75$, nas seis medidas cognitivas, em seguida foram comparadas as diferenças entre os dois grupos nos indicadores do Zulliger. Entre os principais resultados obtidos destacam-se, por um lado,

correlações positivas entre as medidas de inteligência e as seguintes variáveis do teste de Zulliger, número de respostas de forma bem vista (F+), respostas associadas à forma e a cor (FC), respostas com dimensionalidade (FK, KF e F), e também resposta de conteúdo animal, humano e botânico, além de movimentos de figuras humanas inteiras. Por outro lado, foram encontradas correlações negativas com as variáveis, forma distorcida (F-), e respostas de conteúdo artístico, de anatomia e antropológicos. Tais resultados permitem concluir que existe uma maior frequência das variáveis do Zulliger habitualmente associadas ao bom desenvolvimento cognitivo entre os indivíduos com maior desempenho nas provas de inteligência utilizadas. Entretanto as associações foram baixas, o que limita a aplicação desses indicadores como medida de inteligência, sugerindo também que eles medem aspectos diferentes do funcionamento cognitivo.

Primi, Biguetti, Munhoz, Noronha, Polydoro, Di Nucci e Pellegrini (2002), buscaram por associações entre as habilidades cognitivas, medidas pelo BPR-5; interesse, medido pelo Levantamento de Interesses Profissionais (LIP); e traços de personalidade, medidos pelo Questionário de Personalidade 16 PF. Os resultados mostraram associações significativas entre os raciocínios: abstrato (RA), mecânico (RM) e espacial (RE) com interesses por ciências físicas e cálculo. Outra associação significativa foi entre os subtestes RA, RM e RE com o Fator III do 16 PF.

Almeida e Primi (2004) exploraram os perfis de capacidades cognitivas da BPR-5, utilizando os escores da amostra de padronização da Forma B da bateria, no Brasil (n=1.144) e em Portugal (n=1.595). A partir das 243 possibilidades de configuração de perfis encontradas, foram estabelecidas oito categorias de classificação, a saber: 1) *Perfis médios*, quando os escores em quatro ou cinco provas se encontravam na faixa média de capacidade; 2) *Perfis Superiores*, quando os escores nas cinco provas estivessem acima da média ou quatro estivessem acima da média e uma fosse média; 3) *Perfis Inferiores*,

quando os escores nas cinco provas estivessem abaixo da média ou quando quatro estivessem abaixo e uma fosse média; 4) *Perfis médio inferiores*, quando os escores estivessem abaixo da média em três provas e médios nas outras duas; 5) *Perfis médio superiores*, quando os escores estivessem acima da média em três provas e médios nas outras duas; 6) *Perfis com discrepâncias*, quando fosse encontrado escore pelo menos em uma prova inferior à média e uma superior à média; 7) *Perfis médios negativos*, quando em três provas os escores estivessem na média e nas outras duas fossem inferior à média; 8) *Perfis médios positivos*, quando os escores em três provas estivessem na média e nas outras duas fossem superior à média.

Os resultados apontaram que nos dois países, a categoria perfis com discrepância foi a mais freqüente, porém, puderam-se notar algumas diferenças significativas nas categorias de perfis entre os países ($\chi^2 [7] = 26,8$; $p < 0,001$) e uma diferença na distribuição dos estudantes nas oito categorias (Brasil $\chi^2 [35] = 126,7$; $p < 0,001$ e Portugal $\chi^2 [35] = 60,8$; $p < 0,01$). Pôde ser observada também uma diminuição do número de perfis médios e um aumento dos perfis com discrepância nos alunos mais velhos, o que sugere uma especialização das capacidades cognitivas. Quanto à análise dos alunos com perfis de discrepância, foram encontradas semelhanças entre os dois países, sugerindo que existe uma estabilidade na estrutura das capacidades avaliadas pelo BPR-5. Foram identificadas também a existência de competências distintas em função das variáveis utilizadas como critério externo, por exemplo, gênero, idade, reprovação e área de estudo (Almeida & Primi, 2004).

Baumgartl (2004) buscou evidências de validade de critério para o BPR-5, BFM-1 e PMK usando como critério externo a associação com medidas relacionadas ao desempenho no trabalho. Participaram do estudo 79 funcionários de uma empresa de energia elétrica. Os resultados apontaram correlações significativas entre o BPR-5 com o desempenho quando

os funcionários apresentam menor tempo de experiência. Esse resultado faz sentido uma vez que a bateria é um instrumento que avalia a inteligência fluida que se refere à capacidade do indivíduo em lidar com situações novas e pouco estruturadas. Desta forma a autora concluiu que os resultados obtidos apontaram que o BPR-5 é um instrumento adequado ao processo de avaliação de eletricitários.

Baumgartl e Nascimento (2004) realizaram uma outra pesquisa com 57 eletricitários com idade entre 19 e 45 anos, com o objetivo de estabelecer normas para o BPR-5 e de investigar a relação entre variáveis demográficas e ocupacionais com o desempenho na forma B da bateria. As normas foram estabelecidas especificamente para os funcionários do setor elétrico de uma companhia energética, por ocasião da reimplantação do processo de avaliação psicológica. Os resultados apontaram que, em relação à função exercida, o desempenho variou entre 90 e 110 na escala de escore padrão normalizado (EPN), ou seja, dentro da faixa de desempenho considerada média. No que diz respeito à comparação das médias de desempenho nos subtestes e a função exercida, não ocorreu diferença estatisticamente significativa entre os grupos avaliados. Os coeficientes de correlação indicaram que o subteste RN apresenta uma correlação pequena, mas significativa, com a função exercida. Os coeficientes de correlação entre os resultados nos subtestes e a escolaridade foram positivos e significativos, com exceção do subteste RA. E entre os subtestes e a idade foram negativos e significativos, sugerindo uma queda no desempenho dos sujeitos com mais idade.

As autoras encontraram diferenças significativas ($t=-2,93$; $p=0,025$) entre as médias, com base no EPN Total, usando as normas do manual e as do estudo em questão. A diferença foi encontrada nos escores dos subtestes RA ($t=10,843$; $p=0,000$), RM ($t=-7,786$; $p=0,000$) e RE ($t=5,804$; $p=0,000$). As autoras observaram que existe a necessidade da criação de normas para adultos, assim como para grupos específicos, considerando a

realidade da avaliação a ser realizada. Ou seja, os resultados revelaram a importância do estabelecimento de normas para a empresa, visto que a amostra organizacional apresenta características peculiares em relação à amostra original de estudo da BPR-5.

Costa (2004) buscou elucidar as relações entre variáveis do teste das Pirâmides de Pfister e o BPR-5. Participaram da pesquisa 200 estudantes do ensino médio, dos períodos diurno e noturno de escolas da rede estadual e privada de um município do Estado de Minas Gerais. Ao ser comparado o desempenho no BPR-5 entre as séries testadas, houve uma tendência entre os alunos da 1ª série a apresentarem melhores resultados do que os alunos da 3ª série. Segundo a autora, estes dados indicam a ocorrência de um declínio nos padrões de desenvolvimento do desempenho em testes saturados de inteligência fluida (*Gf*), que atinge o seu auge por volta dos 18 anos de idade.

No que respeita à relação entre a variável nível formal do teste de Pfister e o desempenho na prova de raciocínio, os resultados indicaram que os indivíduos que executaram pelo menos uma pirâmide em estrutura apresentaram elevada pontuação no BPR-5. Pôde ser observado também, que os indivíduos que não executaram tapetes apresentaram um desempenho melhor no BPR-5 do que os que fizeram pelo menos um tapete. A autora verificou que tanto no grupo com alto desempenho quanto no grupo com baixo desempenho, há uma ocorrência alta de porcentagem de tapetes. Neste caso, considerando a interpretação para as pirâmides em estruturas e para as pirâmides em tapetes, foram comparadas as estruturas com as respostas FC. Os resultados apontaram que a pessoa que faz estrutura consegue aplicar o controle intelectual às excitações afetivas, adequando-se no seu campo psicológico, na persistência ou não do afeto, dentro de uma modulação própria, de mais agradável a menos agradável. Por outro lado, os indivíduos que fizeram tapete apresentaram maior propensão a sofrer impacto emocional por carecer de uma ação analítica sobre a afetividade representada pelas 360 etiquetas coloridas do Pfister.

Segundo a autora, esta interpretação é possível, uma vez que foram encontradas no grupo com alto desempenho no BPR-5 pessoas que executaram pirâmides em tapetes (18%). Ou seja, se a perturbação emocional for muito grande, mesmo em pessoas com um nível intelectual melhor pode-se observar a ocorrência de tapetes. Portanto, tais resultados se apóiam na correlação existente entre a estrutura e o alto desempenho e entre tapete e baixo desempenho.

Em outro estudo usando o BPR-5 no contexto organizacional, Filizatti (2004) comparou os resultados nos subtestes e o raciocínio geral, com os resultados da avaliação de desempenho no trabalho realizada pelo supervisor, pelo colega de trabalho e a média de avaliação de ambos. Os resultados indicaram que existe uma correlação significativa entre a avaliação de desempenho no trabalho e os subtestes da BPR-5. Os coeficientes foram, com o subteste RN ($r=0,260$; $p=0,004$), e com o RE ($r=0,364$; $p<0,001$). Quanto à avaliação realizada pelo colega de trabalho, foi encontrada apenas no subteste RA uma correlação significativa ($r=0,281$; $p=0,007$), e correlação marginal nos subtestes RE ($r=0,163$; $p=0,071$) e RN ($r=0,162$; $p=0,073$). Também foi encontrada uma correlação positiva e significativa entre o resultado geral e a avaliação do supervisor ($r=0,321$; $p<0,001$) e do colega ($r=0,192$; $p<0,034$).

Também na área organizacional, Mansão (2005) realizou um estudo correlacional entre o teste Self-Directed Search Career Explorer (SDS) e o BPR-5. Os dados indicaram que as pessoas com interesses profissionais representados no SDS pela tipologia Realista (R), geralmente tendem a apresentar escores mais altos em Raciocínio Mecânico (RM). Ou seja, as características que descrevem o perfil realista, como por exemplo, a preferência por profissões ligadas a áreas de ciências e áreas exatas, podem ser associadas às habilidades mecânicas avaliadas pela prova RM da BPR-5.

Couto (2005), em seu trabalho, realizou a estimação dos parâmetros dos itens das provas da BPR-5 utilizando o modelo de Rasch, dois e três parâmetros da Teoria de Resposta ao Item (TRI). Em seguida investigou as relações entre os parâmetros estimados por modelos diferentes e entre populações diferentes. Foram usados neste estudo os dados da amostra de padronização para a população brasileira do teste e dados de estudos realizados com a população portuguesa. Os resultados apontaram à invariância dos parâmetros em relação ao modelo usado na estimação e em relação à população. Os dados da BPR-5 revelaram se ajustar melhor ao modelo de 3-parâmetros, mas houve também ajuste ao modelo de 2-parâmetros e ao modelo de Rasch. O estudo sugeriu que devido às vantagens alcançadas recomenda-se utilizar os valores dos parâmetros estimados pelo modelo de Rasch.

Campos (2005) estudou a relação entre os conteúdos das provas que compõe o BPR-5 e a dificuldade dos itens, estimada pelo modelo de Rasch, apresentadas no estudo de Couto (2005). O conteúdo dos itens foi classificado em diferentes categorias conforme a prova avaliada. Para a prova RV, as categorias, classe gramatical e tipo de relação entre as palavras; Para a prova RN, tamanho da seqüência numérica, operações matemáticas utilizadas, número de seqüências lógicas utilizadas, complexidade das seqüências lógicas; Prova RA, tipos de regras de modificação e número de tipos de regras de modificação; Prova RE, número de eixos de rotação, sentidos de rotação, faces visíveis dos cubos e presença de estímulos visuais nas bordas das faces dos cubos; Prova RM, tipo de conhecimento em Física. Os coeficientes das varáveis de cada categoria foram correlacionadas com o índice de dificuldade dos itens.

Segundo o autor, os resultados das análises sugerem que o conteúdo das provas da BPR-5 estão ligadas aos fatores do segundo estrato, G_f , G_c e G_v , conforme descritas no modelo CHC. Pôde-se observar ainda que o conteúdo das provas está relacionado às

habilidades específicas, descritas no primeiro extrato, como raciocínio seqüencial geral, raciocínio indutivo, raciocínio quantitativo, conhecimento lexical, visualização, relações espaciais e manipulação de imagens. A análise de relação entre o conteúdo e a dificuldade do item resultou em coeficientes significativos para parte das variáveis pertencentes às categorias classe gramatical, tipo de relação entre as palavras, seqüência numérica, operações matemáticas utilizadas, complexidade das seqüências lógicas, tipos de regras de modificação, número de tipos de regras de modificação, número de eixos de rotação, presença de estímulos visuais nas bordas das faces dos cubos e tipo de conhecimento em Física. De acordo com o autor, a análise do conteúdo dos itens possibilitaram maior detalhamento das habilidades medidas pelo instrumento, assim como a identificação de prováveis explicações para as dificuldades dos itens.

Adotando procedimentos diferentes dos observados nos estudos anteriormente citados, Ferreira, Almeida e Guisande (2006) realizaram uma pesquisa procurando apresentar todo um conjunto de investigações que relacionam a memória de trabalho e as aptidões humanas, tendo em vista o modelo CHC. Os resultados encontrados sugerem que a dimensão memória de trabalho, claramente associada ao estrato I do modelo teórico CHC, surge representada em várias baterias da inteligência, inclusive na Bateria de Provas de Raciocínio.

Cunha, Nunes, Primi, Cruz e Almeida (2006) analisaram as habilidades cognitivas de 24 músicos estudantes e profissionais já graduados pela Escola de Música da Universidade Federal da Bahia usando o BPR-5. Os resultados apontaram que dos raciocínios avaliados pela BPR-5, o raciocínio verbal, avaliado pela Prova RV, foi o único que apresentou diferença estatisticamente significativa entre o grupo avaliado e a amostra normativa.

Barbosa (2006) em seu estudo buscou evidências de validade para o BPR-5 utilizando como critério uma amostra de idosos. Com objetivo de verificar a validade foi adotado procedimento de correlação entre os resultados na BPR-5 e a idade, o gênero, a escolaridade e o tempo gasto para a execução das tarefas. Dos 43 participantes submetidos à BPR-5, 19 responderam a Forma A, sendo que 78,9% têm de 4 a 7 anos de estudo; e dos 24 que responderam a Forma B, 82,4% têm de 11 a 14 anos de estudo. Foi realizado um registro do tempo de aplicação dos subtestes, considerando o desempenho no tempo previsto pelo manual (número de itens respondido no tempo padronizado) e o tempo de execução de cada tarefa pelo grupo de idosos. No estudo com os resultados do tempo necessário para a execução das tarefas, os dados apontaram que os idosos utilizam aproximadamente 50% a mais que o tempo limite nos estudos com alunos do ensino médio e fundamental. Em relação a comparação das médias dos grupos etários, foi verificado uma diminuição do desempenho entre os grupos com 60-64 e 65-69 anos, em todos os subtestes; essa diminuição das médias ocorreu também entre a faixa de 70-74 e 75-79 anos, exceto para RA e RV, que aumentam.

Considerando as duas médias nos subtestes e nos escores totais, com ou sem limite de tempo, foi realizado um estudo de correlação e o Teste *t*. Os dados revelaram uma diferença significativa na relação tempo X escore, indicando que o tempo é uma variável que interfere na obtenção da estimativa das habilidades cognitivas a partir da BPR-5, no caso dos idosos (Barbosa, 2006). Outro estudo realizado procurando verificar o desempenho dos idosos (escores obtidos com o tempo livre e a idade) demonstrou que ocorre um declínio das médias em todos os subtestes e mais acentuado no RE, RM e RN. A correlação dos escores sem tempo limite e a idade foi no RV de ($r=-0,227$; $p>0,05$) e no RA de ($r=-0,123$; $p>0,05$); e com tempo limite foi de ($r=-0,373$; $p=0,05$) no RV e de ($r=-0,323$; $p=0,05$) no RA. Por sua vez, a correlação entre as médias nos subtestes com e sem tempo

limite e o gênero, os valores demonstraram uma diferença significativa apenas nos resultados da prova RM com tempo limite ($t [35] = 4,677; p=0,000$), com tempo limite livre ($t [41] = 5,037; p=0,000$); do EG 5 com tempo limite ($t [35] = 2,245; p=0,031$), com tempo livre ($t [41] = 2,163; p=0,036$).

Na análise de variância entre escolaridade e escores nas provas da BPR-5 com tempo limite, Barbosa (2006) observou que na prova RV e RN não há diferenças significativas entre os escores dos idosos do ensino médio e superior; na RM, RE e RN, os idosos não apresentaram diferenças significativas entre os escores no desempenho de ensino fundamental e médio.

A partir dos estudos citados, verifica-se que há uma grande diversidade de pesquisas utilizando a Bateria de Provas de Raciocínio. Além disso, as pesquisas indicam que a BPR tem se mostrado um instrumento eficiente para a avaliação de estudantes, avaliação de músicos, avaliação de candidatos no contexto organizacional e para a avaliação de idosos. No entanto o instrumento apresenta a limitação imposta pelo modelo da psicometria que lhe serviu de esteio, no caso a psicometria clássica. O presente estudo tem por objetivo desenvolver escalas de avaliação referenciadas nos itens, com base nos resultados encontrados pela análise psicométrica realizada por Couto (2005) e de análise cognitiva semelhante à proposta por Embretson (1983, 1985, 1994 & 1999).

Prova de Raciocínio Abstrato

O raciocínio analógico é um algoritmo típico de expressão do pensamento humano e sua investigação é importante para a psicologia na medida em que providencia conhecimento sobre o funcionamento de uma das formas de expressão da inteligência em particular. A frequência com que podem ser encontradas tarefas de analogia presente em testes de inteligência sugere que o desempenho em tarefas desse tipo, materializa as diferenças individuais no funcionamento de processos cognitivos básicos. Esse conjunto de testes revela aspectos significantes do comportamento que são relacionados ao processamento de informação verbal e não verbal de grande utilidade para discriminar entre indivíduos que possuem maior ou menor probabilidade de obter sucesso acadêmico ou em atividades relacionadas.

A validade preditiva atribuída aos instrumentos pode ser apontada como um incremento de qualidade fornecido pela psicometria, e frequentemente leva a questões sobre as diferenças individuais em termos de cognição que acompanham as diferenças de desempenho nos testes. Porém, ainda é bastante incipiente na psicologia o conhecimento sobre a natureza dos processos que estão na base do desempenho nesse tipo de tarefa. Melhor dizendo, sabe-se pouco sobre como diferenças individuais de processamento cognitivo estão ligadas a determinadas características da tarefa de modo a influenciar as diferenças no desempenho (Mulholland, Pellegrino & Glaser, 1980).

Estudos que tiveram como objetivo central tratar desse problema não são recentes apesar de se apresentarem com uma frequência modesta. No que diz respeito especificamente a expressão do raciocínio analógico em tarefas organizadas em matrizes usando figuras abstratas podem ser citados, por exemplo, o estudo de Jacobs e Vandeventer (1972) no qual foram analisados qualitativamente o conteúdo dos itens de 201 testes de

inteligência, buscando descobrir quais eram as regras inferenciais utilizadas na sua construção. A idéia geral é que tais regras poderiam explicar a dificuldade dos itens a partir das relações com determinadas funções cognitivas. Como resultado os autores apresentaram uma taxonomia que incluía doze classes de regras, que são descritas a seguir.

Identidade: consiste na repetição de um elemento sem mudanças nas células em linha de uma matriz, mas permite mudanças nas células das colunas da matriz, por exemplo,

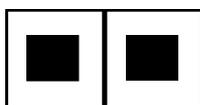


Figura 3. Identidade. Adaptado de Jacobs e Vandeventer (1972).

Forma: Os elementos apresentados nas células em linha de uma matriz sofrem mudança completa da forma,



Figura 4. Forma. Adaptado de Jacobs e Vandeventer (1972).

Sombreamento: se caracteriza pela alteração completa ou parcial do sombreamento interno do elemento nas células da matriz,



Figura 5. Sombreamento. Adaptado de Jacobs e Vandeventer (1972).

Tamanho: se caracteriza pelo aumento ou diminuição de um elemento apresentado nas células da matriz,

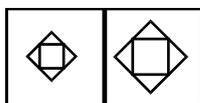


Figura 6. Tamanho. Adaptado de Jacobs e Vandeventer (1972).

Movimento no plano: se caracteriza pela movimentação do elemento em graus regulares no plano apresentado nas células da matriz,

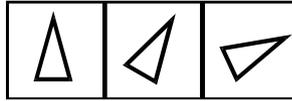


Figura 7. Movimento no plano. Adaptado de Jacobs e Vandeventer (1972).

Séries: consiste na apresentação dos elementos nas células de ma matriz em séries progressivas ou não, com números ou aspectos geométricos,

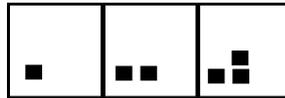


Figura 8. Séries. Adaptado de Jacobs e Vandeventer (1972).

Adição de elemento: se caracteriza pela adição ou remoção progressiva de um elemento ao longo das células de uma matriz,



Figura 9. Adição de Elementos. Adaptado de Jacobs e Vandeventer (1972).

Adição usando combinação: se caracteriza pelo elemento de uma coluna ser adicionado, justaposto, superposto ou subtraído de um elemento de outra coluna para produzir um terceiro elemento na célula seguinte da matriz,

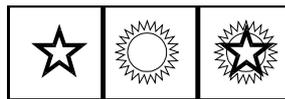


Figura 10. Adição usando combinação. Adaptado de Jacobs e Vandeventer (1972).

Elementos de um conjunto: se caracteriza por três elementos ou três valores de um atributo categórico serem distribuídos nas células ao longo de uma linha ou coluna na matriz, ou seja, os elementos aparecem três vezes em uma matriz 3X3,

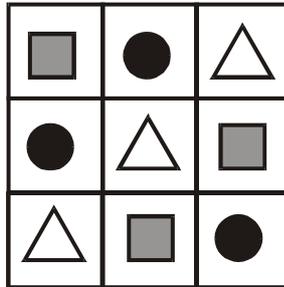


Figura 11. Figuras de um conjunto. Adaptado de Jacobs e Vandeventer (1972).

Adição única: consiste em dois elementos serem sobrepostos ao longo das células da matriz e os elementos comuns serem tratados de modo diferencial, ou seja, adicionados enquanto elementos comuns são removidos,

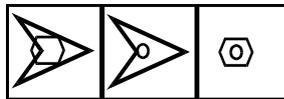


Figura 12. Adição única. Adaptado de Jacobs e Vandeventer (1972).

Reflexão: se caracteriza por um elemento na célula seguinte aparecer como o reflexo em espelho do primeiro elemento,

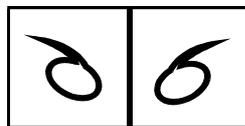


Figura 13. Reflexão. Adaptado de Jacobs e Vandeventer (1972).

Inversão: consiste no intercâmbio ou troca de alguma transformação sofrida pelo elemento, ou propriedade, como forma, sombreamento e/ou posição,

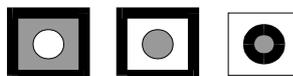


Figura 14. Inversão. Adaptado de Jacobs e Vandeventer (1972).

Em outro estudo, Ward e Fitzpatrick (1973) também examinaram as características de itens de raciocínio figural apresentados na forma de matrizes. O mais conhecido teste que apresenta itens com esse formato é o teste das Matrizes Progressivas de Raven (Raven, 1956) e apesar do volume de literatura acumulada, ainda podem ser encontradas entre os pesquisadores dúvidas e discordâncias sobre o que exatamente ele mede.

Ward e Fitzpatrick (1973) relatam que o seu interesse em detalhar as características dos elementos presentes nas matrizes cresceu depois da construção de uma versão “piloto” para um teste de raciocínio abstrato. Nesse caso foram utilizadas no lugar dos tradicionais modelos de múltipla escolha, um formato de questões abertas na qual o sujeito completa a lacuna em branco. O teste piloto era composto de 120 itens divididos igualmente em duas formas paralelas e foi aplicado a uma amostra estratificada de 300 estudantes com idade variando entre sete e 15 anos. A análise de itens mostrou que as características psicométricas dos itens nas duas formas foram notavelmente similares, e o resultado mais destacado foi a alta correlação entre as formas (0,97). Pôde ser observado também que a dificuldade dos itens tendia a aumentar ou diminuir em função do número de atributos das células em uma matriz e da complexidade das operações requeridas para a solução. Daí a possibilidade de que os itens pudessem ser distinguidos efetivamente em termos de seu conteúdo e do processo cognitivo subjacente.

O exame das matrizes sugeriu que na descrição de seus elementos três características principais podem ser consideradas, a) *características gerais da matriz*; b) *número e arranjo dos atributos das células*; c) *as operações realizadas nos atributos das células*. No que diz respeito a primeira e segunda características a análise dos itens sugere que basicamente matrizes construídas em forma de quadrado podem ser descritas como

simétricas e não simétricas. As simétricas não mudam o nível ou o contorno dos elementos de uma célula para outra, a tarefa do sujeito é completar a figura de acordo com as indicações das outras células. As não simétricas apresentam um formato desigual ou irregular entre as células que pode se manifestar de muitas maneiras.

As matrizes simétricas podem ser divididas de acordo com as características dos elementos em: *simetria pura*, nos quais está envolvido o fechamento de uma figura ou completar uma gestalt, que pode aparecer na figura como um todo ou nos eixos horizontal, vertical e diagonal da matriz. A Figura 15 apresenta exemplos desses diferentes tipos de matrizes, Em A um exemplo de simetria universal, em B simetria horizontal, C simetria vertical e D simetria diagonal.

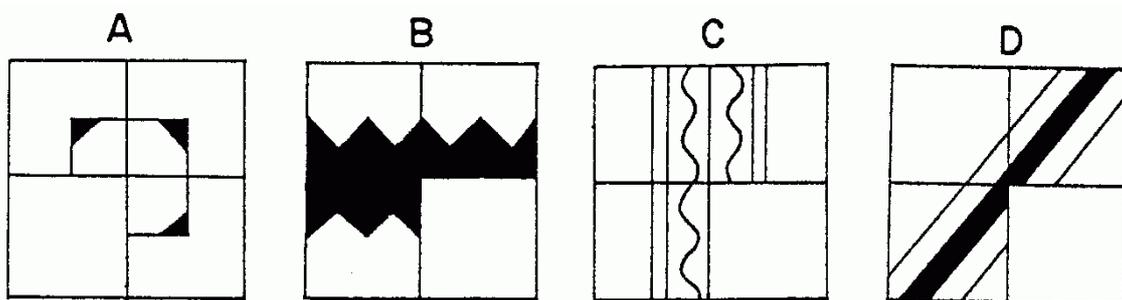


Figura 15. Matrizes com simetria universal. Adaptado de Ward e Fitzpatrick (1973).

Identidade dos elementos, característica que diz respeito a simetria dos elementos, eles podem ser completamente simétricos, por exemplo, nas matrizes onde itens idênticos são apresentados nas colunas ou linhas, ou parcialmente simétricos, como em matrizes mais complexas cujos itens combinam padrões de identidade presente nas linhas e colunas, nestes casos o incremento da dificuldade fica a cargo da adição de um conjunto de formas ou atributos. A figura 16 apresenta exemplos de matrizes com essas características, em A

identidade simples, em B identidade por colunas, em C identidade por linhas, em D identidade combinando linhas e colunas.

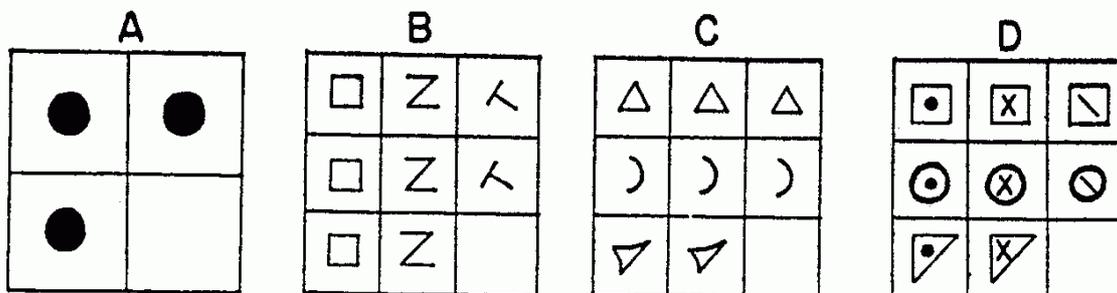


Figura 16. Matrizes com Identidade de Elementos. Adaptado de Ward e Fitzpatrick (1973).

Elementos de conjunto (latin square items), possuem simetria de ordem na qual o conteúdo das células é variado e balanceado sistematicamente. Podem aparecer de forma simples, ou alternando a localização da célula em branco. A dificuldade pode ser incrementada combinando dois ou três efeitos de conjunto. A Figura 17 apresenta exemplos das diferentes combinações possíveis de conjunto, em A combinação simples, em B, combinação simples variando a localização da célula em branco, em C combinação de dois conjuntos de elementos, em D a combinação de conjunto com identidade.

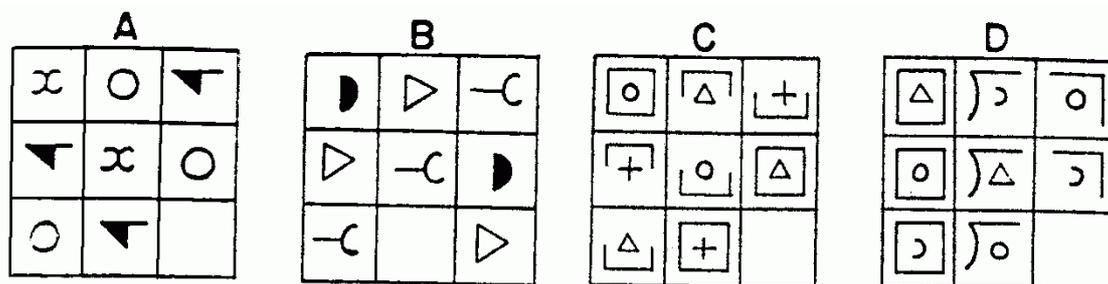


Figura 17. Matrizes com regra de elementos de conjunto. Adaptado de Ward e Fitzpatrick (1973).

Ao lado disso, as matrizes não simétricas podem ser produzidas combinando o conteúdo das células por somas de vários tipos. Por exemplo, executando uma soma simples do conteúdo das células por linhas ou por colunas ou por linhas e colunas ao mesmo tempo. Nesses casos os elementos possuem uma afinidade com a característica de *identidade dos elementos* nas matrizes simétricas descrito anteriormente. Uma outra forma de acrescentar dificuldade aos itens pode ser a manipulação do conteúdo das células exigindo um reconhecimento de elementos comuns ou de intercessão ao longo das linhas, ou ainda a manipulação de elementos que exigem o conceito de união. A figura 18 apresenta exemplos de matrizes não simétricas, em A soma de elementos por colunas ou linhas, em B soma de elementos por linhas, em C soma de elementos por coluna, em D interseção de elementos, em E união de conjuntos.

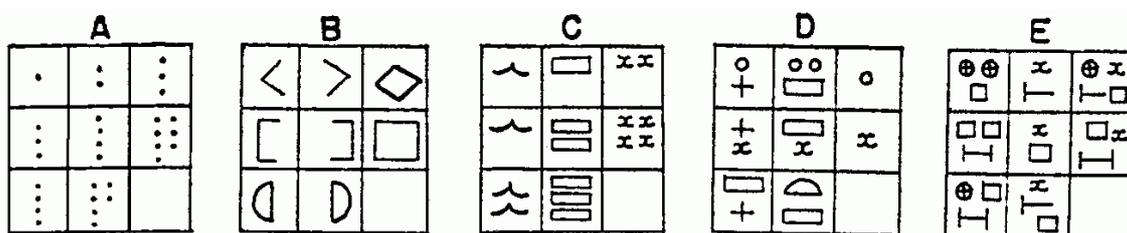


Figura 18. Matrizes não simétricas. Adaptado de Ward e Fitzpatrick (1973).

No caso da terceira característica, *operações realizadas nos elementos das células*, diferentes tipos podem ser utilizadas para escrever uma extensa quantidade de itens. Por exemplo, seguindo uma ordem crescente de dificuldade, pode-se inicialmente alterar o tamanho dos elementos por linhas, por colunas ou ambos ao mesmo tempo (Figura 19 A e B). Acrescentar crescimento seriado dos elementos presentes nas células (Figura 19 C). Formar itens inserindo as regras de adição ou multiplicação dos elementos das células para formar a resposta correta (Figura 19 D e E). Outro conjunto de regras que pode ser aplicado

é o da reordenação, a rotação e reflexão dos elementos presentes em cada célula (Figura 19 F, G e H). E o último conjunto proposto pelos autores diz respeito a inclusão de movimento e de sombreamento sobre os elementos das células (Figura 19 I, J e K).

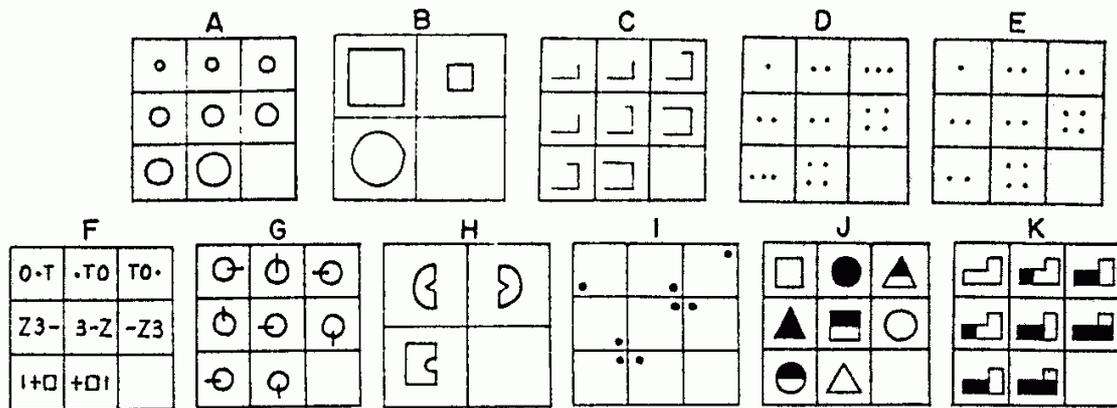


Figura 19. Matrizes mostrando operações sobre o conteúdo das células. Adaptado de Ward e Fitzpatrick (1973).

O resultado do exercício de examinar as matrizes pode ser útil, principalmente ilustrando as potencialidades que o cuidado com o conteúdo na construção de itens pode ter para futuras utilizações do teste. Apesar de os autores objetivamente terem descrito sua taxionomia como rude, o processo pode ser de interesse para os pesquisadores envolvidos na construção de medidas cognitivas (Ward & Fitzpatrick, 1973).

Hornke e Habon (1986) acrescentam que a construção de itens considerando os componentes de dificuldade apresenta como importante contribuição a aproximação entre os modelos psicométricos e a teoria psicológica. Este aspecto pode ser essencial para futuras aplicações e a consolidação das interpretações para o teste. Em seu trabalho os autores objetivaram a construção de um banco de itens que pudesse ser usado para realizar testagem adaptativa. Assim, além da necessidade de se estimar os parâmetros psicométricos dos itens, buscaram relacionar esses parâmetros psicométricos a demandas de processamento cognitivo, ou seja, compreender o desempenho a partir das funções

cognitivas envolvidas. Para se testar a natureza componencial da solução dos itens foi aplicado o *Linear Logistic Test Model* (LLTM; Fischer, 1973) que permite a decomposição da dificuldade dos itens em elementos que podem ser associados as demandas cognitivas.

Visando esse fim, os autores primeiro formularam as seguintes regras, (1) oito maneiras de compor as partes dos itens foram definidas: Identidade, adição, subtração, interseção, adição única, seriação, variação de gestalts abertas e fechadas. (2) as relações entre os elementos podem ocorrer por linhas, por colunas ou ambas. (3) ambas as partes dos itens podem ser separadas, integradas ou embutidas. As regras para construção dos itens levaram em consideração as informações apresentadas na literatura (Ward & Fitzpatrick, 1973; Jacobs & Vandeventer, 1972). Em seguida escreveram um banco completo de itens (num total de 616 itens) por meio da combinação de todas as operações cognitivas formuladas que foram distribuídos em 35 conjuntos de testes contendo 24 itens cada, sendo que haviam no mínimo seis itens ancora por conjunto. Os itens foram aplicados a 7400 sujeitos provenientes das forças armadas alemãs, e por fim avaliaram os parâmetros de dificuldade dos itens à luz das operações hipotetizadas.

Os resultados indicaram que ao se analisar os índices de desajuste estimados entre os parâmetros psicométricos pelo menos 21,1% dos itens apresentaram valores críticos. Entre os problemas encontrados nos itens podem ser destacadas falhas na construção que permitiam a integração de componentes e a decomposição das gestalts, com relação a construção de regras foram encontrados problemas na composição da subtração que nesses casos incluíram duas regras. A estimação dos índices de contribuição de cada componente para a dificuldade do item, usando o LLTM, apontaram em primeiro lugar com respeito as operações sobre as partes dos itens de Separação, Integração e Encaixe (partes embutidas), que os itens envolvendo separação foram os mais fáceis seguidos dos que envolviam integração, porém com um grande número de elementos das duas operações com

dificuldades sobrepostas, ou seja, diferentes elementos contribuem da mesma maneira para a dificuldade. Com relação as oito formas de se compor as partes dos itens, as operações de identidade foram as mais fáceis e as de interseção as mais difíceis, mas também com grande parte dos elementos apresentando sobreposição. Com relação ao sentido em que as transformações ocorreram, linha, coluna ou ambos, as transformações por colunas foram as mais difíceis e a combinação de linhas e colunas as mais fáceis, porém apresentando grande sobreposição.

A análise das sobreposições indicou que elas variaram em uma extensão de mais de duas unidades com tendência de se concentrar em torno de 0,75 numa escala de theta que variou de -2 a 2. Os autores apontaram que parece ser duvidoso que isso possa ser explicado pelo número de itens em cada conjunto, porque depois das eliminações cada conjunto apresentava entre 27 e no máximo 60 itens. Naturalmente restringindo o número de vezes que uma operação aparece num conjunto produzindo combinação com outras operações. Também a taxonomia para os tipos de regra pode ter apresentado falhas não solucionadas como a categorização de transformações semelhantes em categorias distintas, por exemplo, interseção e adição única. Uma outra possibilidade pode ser a distribuição da amostra no que se refere a quantidade de sujeitos que responde a um número pequeno de itens em comum.

De qualquer modo, devido as restrições do LLTM em analisar uma grande quantidade de dados, foram efetuadas análises de regressão considerando o valor da dificuldade dos itens (b) como variável dependente e as 15 operações como variáveis independentes. Os resultados apontaram que 40% da variância da dificuldade foi explicada por uma função composta por 14 das 15 operações cognitivas hipotetizadas. Em seguida os resultados da regressão foram correlacionados com os resultados estimados pelo LLTM e apresentaram correlações em média de $r = 0,99$ com o mínimo de 0,76 e o máximo de 0,99.

Esses resultados permitem dizer que de um modo geral as metas iniciais foram cumpridas e que parte razoável da variância da dificuldade dos itens foi adequadamente explicada pelas operações hipotetizadas. Uma boa parte dos problemas com os itens puderam ser identificados ao longo das análises e apesar da técnica usada não ser perfeita, por ser um modelo linear de decomposição da dificuldade, uma outra técnica mais facilmente aplicável a construção de itens ainda não foi vista (Hornke & Habon, 1986).

Outro estudo que partindo dos mesmos pressupostos teóricos buscou identificar e testar os componentes de dificuldade dos itens dos testes Matrizes Coloridas de Raven (MCR) e Matrizes Progressivas de Raven (MPR) foi realizado por Green e Kluever (1991). Para tanto, os dois testes foram aplicados a uma amostra de 269 crianças com idades entre dois anos e nove meses até 11 anos e oito meses. Os coeficientes de dificuldade foram estimados pelo modelo logístico de um parâmetro da TRI usando o software BICAL, os itens desajustados foram eliminados da análise.

Em seguida foram analisados os itens do MCR na tentativa de identificar potenciais componentes de dificuldade que foram identificados nos trabalhos anteriormente citados de Ward e Fitzpatrick, (1973) e Mulholland, Pellegrino e Glaser (1980) dentre outros. Os potenciais componentes de dificuldade incluídos na análise foram (a) orientação nos eixos verticais e horizontais vs orientação nos eixos diagonais, (b) simetria vs assimetria, (c) progressão vs não progressão, (d) número de dimensões de variação, (e) linhas retas vs linhas curvas, (f) número de linha ou sólidos, (g) densidade do desenho, (h) cores vs preto e branco. As primeiras variáveis foram marcadas como dicotômicas (0-1) com exceção do número de dimensões que variou de (0-3). Outras variáveis analisadas foram (i) número de opções de resposta, (j) opções contando com progressão, rotação e reflexão de elementos, (k) número de direções das opções (horizontal, vertical e etc.), (l) número de elementos do

desenho, (m) reversão entre primeiro e segundo plano. As características citadas foram avaliadas por de forma independente pelos autores,

Os dados gerados para MCR foram analisados usando os coeficientes de dificuldade como variável dependente e as características dos itens como variáveis independentes em uma equação de regressão usando o método de entrada forçada e o método stepwise. Os resultados apontaram que, em primeiro lugar, considerando o método de entrada forçada apresentou um valor de R de 0,90 e, em segundo lugar, considerando o método stepwise apontou uma equação com as variáveis, número de opções (0,17), reflexão (0,43), número de dimensões (0,75). Opções de direções de opção (0,29) contribuindo significativamente ($p < 0,05$) com um R de 0,88. O passo seguinte foi testar a equação proposta pelo modelo stepwise usando os dados gerados pelo MPR. Os resultados apresentaram uma equação com a seguinte contribuição, número de opções (0,31), reflexão (0,18), número de dimensões (0,39). Opções de direções de opção (0,07) contribuindo significativamente ($p < 0,05$) com um R de 0,69.

Os autores consideraram que os resultados encontrados foram congruentes com as explicações sobre os modelos de processamento de informação de analogias geométricas descritos na literatura. Especialmente o número de elementos das figuras e o número de dimensões do problema. Porém, as diferenças de dificuldade produzidas por diferentes transformações como rotação e reflexão precisam ser melhor entendidas.

Os trabalhos acima descritos representam um avanço no entendimento sobre os componentes que são processados cognitivamente presentes nos itens que compõe a medida do raciocínio abstrato. No entanto, é necessário que se integre esse tipo de procedimento de pesquisa com as novas descobertas sobre o funcionamento cognitivo alcançado pela pesquisa em ciências cognitivas. Desta maneira pode ser possível entender que tipo de processo cognitivo é representado pelos elementos que compõe o item, qual a contribuição

de cada elemento para a dificuldade desse item, e também as possíveis ramificações de atividade neural envolvidas na resolução adequada desse item.

Prova de Raciocínio Numérico

Simon e Kotovsky (1963) e Simon (1979) apresentam uma explanação teórica sobre como sujeitos humanos processam padrões de seqüência de símbolos. Dentre os objetivos do seu trabalho estavam a exploração de como são armazenados e lembrados os padrões seriais; e como são adquiridos os conceitos ou regras induzidos de um exemplo. Um aspecto importante tratado no primeiro trabalho é o da aquisição de conceitos. Os conceitos podem ser definidos como o conjunto de regras que descrevem as características dos objetos, e/ou a relação entre essas características. Geralmente os conceitos podem ser usados para separar uma subclasse dentro de algumas classes de objetos, ou mesmo identificar se um objeto em particular pertence ou não a uma classe ou subclasse. Uma evidência empírica comumente usada para demonstrar que um sujeito aprendeu ou compreendeu um conceito, é sua habilidade de organizar objetos em um conjunto segundo as regras descritas pelo conceito, e separá-los de outros objetos que não pertencem ao mesmo conjunto. Por exemplo, pode-se dizer que um sujeito compreende o conceito de vermelho quando ele é capaz de separar os objetos vermelhos em um conjunto com vários objetos coloridos.

A esse respeito os autores destacaram que a habilidade para identificar objetos que exemplificam um conceito não está necessariamente relacionada a habilidade de produzir exemplos desse mesmo conceito. A falta de relação apontada entre as duas habilidades pode ser vista na discrepância observada entre o vocabulário de leitura e o usado na fala de um indivíduo. Essa discrepância pode ser entendida como a diferença entre a habilidade de

compreender e a de produzir palavras em um idioma, e está ligada a diferentes aspectos do sistema de memória, como, a memória de reconhecimento e de evocação (Simon & Kotovsky, 1963).

Essas duas habilidades podem ser medidas usando várias classes de conceitos apresentados proeminentemente na forma de *Padrões Seriais*. Assim, pode-se medir a compreensão que um sujeito tem de um conceito apresentando a ele uma seqüência de caracteres e perguntando se aquela seqüência segue as regras descritas pelo conceito ou não. Porém, o que é mais freqüente, na medida de aquisição de conceitos por padrões seriais, é apresentar uma seqüência incompleta e solicitar que o sujeito complete essa seqüência. Por exemplo, é apresentada ao sujeito uma seqüência *abababab...*, cuja regra é a simples alternância de caracteres *a* e *b*, e lhe é solicitado continuar a seqüência. Dessa forma, se ele consegue continuar adequadamente essa seqüência, pode-se dizer que o sujeito deduziu a regra e reconheceu o padrão envolvido na série (Simon & Kotovsky, 1963).

Um outro aspecto considerado pelos autores diz respeito a compreensão do processo de geração de seqüências e da influência da “bagagem” dos sujeitos para a complexidade das tarefas de *Padrões Seriais*. O objetivo foi entender o mecanismo cognitivo que permite, a um sujeito qualquer, induzir o padrão de descrição em seguimentos de séries de letras, mante-lo na memória e continuar adequadamente a seqüência em questão. Portanto, em seu trabalho, os autores evitaram a construção de um mecanismo cognitivo complexo e específico para essa tarefa, em vez disso adotaram um sistema hipotético composto pelos mesmos mecanismos elementares que têm sido usados pela psicologia cognitiva para descrever os processos gerais de solução de problemas, e de percepção e memória. Desta forma, foi apresentada uma descrição dos componentes cognitivos básicos que fazem parte das etapas de resolução das tarefas. Para tanto, foram analisadas as respostas de 79 sujeitos

à um conjunto de 15 problemas de completar séries de letras. Os sujeitos foram divididos em dois grupos, o primeiro composto de 12 sujeitos com curso superior e o segundo composto por 67 sujeitos estudantes do final do segundo grau.

À conclusão desse estudo os autores propuseram um modelo hipotético do processamento necessário à solução dos itens de completar séries, e também os componentes básicos nos quais se podem decompor essas tarefas. O primeiro componente é a *identificação da relação* e consiste no exame da série para encontrar como um elemento se liga ao outro. Nesse caso três relações são possíveis; a identidade, que consiste na repetição da letra; o movimento para frente, ou para a próxima letra no alfabeto; e o movimento para trás, ou a próxima letra do alfabeto no sentido reverso. O segundo componente é *identificação da periodicidade*, consiste na descoberta do comprimento do período de uma série que completa um ciclo do padrão. Por exemplo, na série *abababab* o comprimento do período é de um, porque a mesma relação se aplica a cada elemento consecutivamente. Já na série *aabbaabbaabb* o comprimento do período é de dois porque a relação se aplica a cada par de elementos da série. O terceiro componente é *complementação do padrão de descrição* depois do comprimento de período da série ter sido encontrado o sujeito precisa identificar as relações que controlam todas as posições de cada elemento dentro de um período em ordem para definir completamente a regra para a seqüência. O quarto e último componente é a *extrapolação* que consiste em identificar a posição no período ocupada pela célula em branco que completa a seqüência, isolar a parte da regra que controla essa posição e aplicar essa regra para gerar a resposta que completa a seqüência (Simon & Kotovsky, 1963; Simon, 1979).

Seguindo os mesmos pressupostos teóricos, porém aplicados a tarefas de indução de regras em problemas de séries numéricas, Holzman, Pellegrino e Glaser (1982, 1983) examinaram a influência de dimensões de processamento e conteúdo de conhecimento no

desempenho de itens individuais que representam tarefas de séries numéricas usadas nos principais testes de inteligência. Especificamente foi examinada a contribuição de diferentes características dos itens para a dificuldade, ou seja, os fatores de complexidade. Atenção especial foi dada a investigação de como a idade e diferenças de habilidade interagem com as dimensões de processamento e conteúdo de conhecimento para determinar a dificuldade dos itens. Participaram da pesquisa 54 pessoas divididas em três grupos. O primeiro grupo, chamado de grupo de adultos, composto por 18 estudantes de graduação; o segundo grupo, chamado de grupo de crianças de QI alto, composto por crianças da quarta e quinta séries da escola elementar, com QI acima de 130 na escala de inteligência do Stanford-Binet, e o terceiro grupo chamado de grupo de QI médio, composto por crianças da mesma série escolar, porém com QI abaixo de 103 estimado pela mesma escala.

No primeiro estudo Holzman, Pellegrino e Glaser (1982) definiram a variável dimensão de processamento pela análise do conteúdo das analogias numéricas em dois tipos, *Número de Operações* e *Ambigüidade de Operações*. O número de operações poderia variar entre 0 (nenhuma operação aritmética), 1 (uma operação aritmética simples), ou 2 (combinação de operações aritméticas). A ambigüidade de operações poderia variar entre 0 (quando as relações não eram ambíguas) e 1 (quando o problema permite mais que uma inferência sobre as relações aritméticas aplicadas no primeiro par, por exemplo, $11 : 22 :: 7 : 18 :: 28 : _$. No qual podem ser aplicadas as operações de $\times 2$ ou $+11$). Já a variável conteúdo de conhecimento foi dividida qualitativamente entre os tipos de relações aritméticas, em contraste com a adição, aplicadas aos itens (Adição vs Subtração, Adição vs Multiplicação, Adição vs Divisão), a magnitude do valor relacional, e também a ambigüidade da operação.

Foram realizadas análises de variância para estudar os efeitos de contraste das variáveis idade e habilidade entre os grupos, e análises de regressão para encontrar a influência dos fatores de complexidade na previsão da dificuldade dos itens. Os resultados da análise de variância apontaram, com esperado pelos autores, que as tarefas de analogias de séries numéricas discriminaram satisfatoriamente [$F(1,51) = 53,07$; $p < 0,001$] entre os grupos de idade (adultos vs crianças) e também [$F(1,51) = 5,55$; $p < 0,05$] entre os grupos de habilidade (Alto QI vs Médio QI). Os resultados das análises de regressão realizadas para as variáveis componentes das dimensões de processamento apontaram um modelo com variância explicada de 0,73; 0,78 e 0,72 entre os três grupos, de adultos, crianças de QI alto e crianças de QI médio respectivamente. A variável *Número de Operações* é a que mais contribui com a previsão do modelo apresentando valores de β de -0,52; -0,72 e -0,84 respectivamente entre os grupos de adultos, crianças de alto e médio QI. O aumento progressivo da contribuição dessa variável ao longo dos grupos contrasta com a diminuição da contribuição da variável *Ambigüidade de Operações* que apresentou valores de β de -0,44; -0,21 e -0,01 entre os grupos. A explicação hipotetizada pelos autores seria a de que crianças com QI médio usam menos operações de multiplicação e divisão para inferir as relações aplicadas aos itens, que são comuns nos itens com operações ambíguas, desta forma evitando o efeito de ambigüidade como um fator de complexidade.

Os resultados das análises de regressão entre os componentes da variável conteúdo de conhecimento produziu um modelo de previsão com variância explicada de 0,69 e 0,59 respectivamente para os grupos de QI médio e alto. Para o grupo de QI médio apenas as variáveis magnitude, divisão e multiplicação apresentaram coeficientes significativos β - 0,69; -0,34 e 0,51 respectivamente. O valor negativo do β da operação de divisão significa que essa operação é mais difícil em contraste com a adição. O grupo de QI alto apenas a

variável magnitude apresentou coeficiente significativo β de -0,72. Os autores concluíram que ambos, as dimensões de processamento e conteúdo de conhecimento influenciam o desempenho dos sujeitos em todos os três grupos, embora diferentes componentes de solução apresentem variação nos níveis relativos de dificuldade e na interação com os grupos de sujeitos examinados. A natureza dessa interação é de especial interesse para ampliar a compreensão do desenvolvimento intelectual e das diferenças individuais sobre a atividade acadêmica.

No segundo estudo Holzman, Pellegrino e Glaser (1983) novas variáveis foram adotadas para a análise com o objetivo de compreender melhor a indução de regras examinando as dimensões de processamento e o conteúdo de conhecimento que apresentaram relação com o desempenho nas tarefas. As variáveis construídas para essa análise foram: *Demandas sobre a Memória de Trabalho*, caracterizada pelo número de elementos presente em cada série; *Comprimento do Período*, caracterizado pelo tamanho dos períodos em cada série e *Complementação do Padrão de Descrição*, caracterizadas conforme descrição de Simon e Kotovsky (1963) apresentada acima; *Complexidade das Relações*, caracterizada por relações hierárquicas envolvendo os elementos do item, por exemplo, na série 36 : 34 : 30 : 22 : ..., a regra de solução envolve duas operações de forma hierárquica, primeiro a subtração, em seguida, a multiplicação do valor subtraído no elemento anterior por dois. Ou seja, a aplicação de mais que uma operação aritmética para explicar a mudança dos elementos no período foi classificada como complexa; *Categoria de Operações Aritméticas*, caracterizada pela relação com a Identidade dos elementos na série, quando não havia operação aritmética aplicada à série e sim a simples repetição dos seus elementos o valor atribuído a essa categoria foi de zero, quando alguma operação foi identificada em contraste com a Identidade o valor atribuído foi de 1 (está variável é dividida em quatro, Adição vs I, Subtração vs I, Multiplicação vs I e Divisão vs I).

Comprimento da Série, caracterizado pelo tamanho de elementos que caracterizam cada série; *Conflito de Direção*, quando mais que uma regra pode ser deduzida e não há clareza sobre qual regra controla as mudanças dentro da série.

Os resultados da análise de regressão usando como variável dependente a dificuldade dos itens e variáveis independentes o conjunto de variáveis descrito acima, apontou para os três grupos (Adultos, Crianças com QI Alto e QI Médio) respectivamente, valores de R^2 de 70,4; 76,4 e 77,6 todos com $p < 0,001$. Porém, as variáveis que contribuem significativamente com a explicação da variância nos três grupos foram *Demandas sobre a Memória de Trabalho* apresentando valores de β de -0,89; -0,89 e -0,67 respectivamente, e as quatro subcategorias de *Categoria de Operações Aritméticas*, apresentando o valor mais alto de β de -0,43 para o grupo de QI Médio e o valor mais baixo de β de -0,14 para o grupo de adultos. As categorias de operações aritméticas de Adição e Subtração apresentaram valores positivos, sendo os mais altos de 0,27 e 0,32 para o grupo de adultos. Esses valores sugerem que essas categorias são mais fáceis que as demais. A categoria de *Complexidade das operações* apresentou valor de $\beta = -0,17$ significativo apenas para o grupo de crianças com QI Médio. Quando a variável *Demandas sobre a Memória de Trabalho* é retirada da análise os valores de R^2 apresentam queda sensível (59,9; 65,9; 71,4; nos três grupos respectivamente) indicando que a variância associada a memória de trabalho é maior que a associada as outras variáveis presentes na análise.

Uma nova análise de regressão foi realizada introduzindo duas novas variáveis, a saber, *Descoberta da Periodicidade*, caracterizada pela utilização das mudanças nos elementos adjacentes ao período para determinação da regra de mudança que controla a série. Por exemplo, na série 12 : 12 : 13 : 13 : 14 : 14 ..., o elemento adjacente ao período imediatamente anterior sofre uma mudança que sinaliza a regra que controla as mudanças

na série, por outro lado na série 45 : 36 : 44 : 36 : 43 : 36 : ..., o elemento adjacente não apresenta mudança que forneça indicações sobre a regra que controla a série. Quando é possível induzir a regra baseada nos elementos adjacentes essa variável recebe o valor de zero e ao contrário recebe o valor de um. E a *Magnitude da Operação* é caracterizada pela quantidade de operações que são aplicadas aos períodos que compõe a série e permitem a resposta adequada.

Os resultados apresentaram valores de R^2 de 74,7; 84,9 e 82,5 todos com $p < 0,001$ respectivamente para os grupos de Adultos, Crianças com QI Alto e QI Médio. As variáveis que apresentaram valores significativos para os três grupos foram *Demandas sobre a Memória de Trabalho* com o valor mais alto de $\beta = -0,60$ para o grupo de adultos e a *Descoberta da Periodicidade* com o valor mais alto $-0,36$ para o grupo de crianças de QI Alto. Esses resultados levaram os autores a considerar que as demandas sobre a memória de trabalho presentes em cada item estão fortemente relacionadas ao desempenho, embora ambos os grupos tenha sido influenciados por essa variável seu efeito é maior em crianças do que em adultos. Outra característica considerada importante foi a complexidade das regras de transformação dos elementos dos itens, que pode ser observada nas variáveis complexidade de operações de também nas categorias de operações que sempre apresentaram contribuições altas e significativas na previsão da dificuldade dos itens.

De um modo geral os resultados encontrados pelos estudos de Simon e Kotovsky (1963) e também de Holzman, Pellegrino e Glaser (1982, 1983) apresentaram um modelo considerado adequado para explorar os fatores de complexidade presente em itens de analogia envolvendo problemas de séries numéricas. Dessa forma, com o objetivo de produzir uma escala na qual a variação da dificuldade possa ser entendida pelas características dos itens, as análises propostas no presente estudo vão adotar procedimentos

semelhantes, e também as variáveis que serão usadas na análise cognitiva da prova de raciocínio numérico serão baseadas nas variáveis propostas pelos autores.

Método

Fonte

Grupo 1: uma amostra composta por 1722 protocolos sendo que 603 sujeitos responderam a forma A do teste, distribuídos em 30,4% dos alunos na sexta série (183), 29,3% na sétima série (175) e 40,3% na oitava série (245). As idades variavam entre 11 e 18 anos; 48,9% dos sujeitos eram do sexo masculino (295) e 51,1% do sexo feminino (308). Já 1119 sujeitos responderam a forma B do teste, sendo 25,2% na primeira série do ensino médio (282), 52,1% na segunda série (583) e 22,7% na terceira série (254). As idades variavam entre 14 e 21 anos, 42,7% dos sujeitos eram do sexo masculino (478) e 57,3% do sexo feminino (641). Os protocolos compõem o banco de dados do LabAPE para o BPR-5.

Grupo 2: uma amostra composta por 150 estudantes de escolas públicas e 20 de instituições particulares de ensino, localizadas no interior do Estado de São Paulo, que cursam de 1ª à 5ª séries do Ensino Fundamental, do período matutino e vespertino, com idade entre 06 a 11 anos, de ambos os sexos, escolhidos por conveniência. Só participaram desse estudo aqueles sujeitos cujos pais ou responsáveis assinaram o termo de consentimento

Material

Prova de Raciocínio Abstrato (RA) e Raciocínio Numérico (RN) da BPR-5: a Bateria de Provas de Raciocínio – 5 é um instrumento de medida das capacidades cognitivas, especificamente da inteligência fluida (Gf), que foi desenvolvido por Primi & Almeida e publicado em 1998. É organizada em duas formas A e B, com cinco subtestes cada, destinados à avaliação de alunos de 6º a 8º série do ensino fundamental, forma A, e alunos da primeira à terceira série do ensino médio, forma B, respectivamente. Os subtestes

são: Prova de Raciocínio Abstrato (RA), Prova de Raciocínio Verbal (RV), Prova de Raciocínio Mecânico (RM), Prova de Raciocínio Espacial (RE), Prova de Raciocínio Numérico (RN). Neste trabalho foram analisadas as provas RA e RN. A prova RA é composta por 25 itens envolvendo analogia com figuras geométricas nos quais, o sujeito deve completar o termo D da analogia a partir da observação das características do termo C e aplicação de regras deduzidas dos termos A e B (exemplo, Figura 20), e o tempo limite para resolução dos problemas é de 12 minutos.

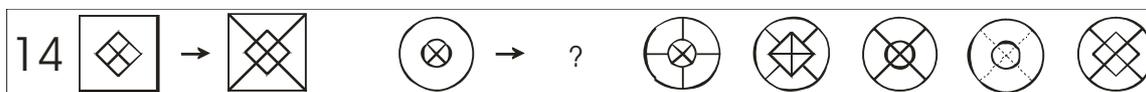


Figura 20. Exemplo de item da prova RA da BPR-5.

A Prova RN é composta por 20 itens nos quais existem séries de números lineares ou alternados, a tarefa do sujeito é descobrir as regras aritméticas que regem a progressão das séries de modo que possa descobrir quais números completam a série (exemplo, Figura 21), o tempo limite é de 18 minutos;

A. 3 6 9 12 15 ? ?

Figura 21. Exemplo de item da prova RN da BPR-5.

Para ambas as provas são feitas as contagens dos acertos dos sujeitos em cada item e depois o resultado bruto são transformados em Escore Padrão Normalizado e Percentis.

Prova de Raciocínio Abstrato (RA) e Raciocínio Numérico (RN) da BPR-5i: a Bateria de Provas de Raciocínio Infantil foi desenvolvida por Almeida, Primi e Cruz (2007) com objetivo de avaliar a realização cognitiva dos estudantes de 1ª a 5ª séries do ensino fundamental, no que respeita à apreensão de relações entre elementos (raciocínio indutivo) e aplicação das relações inferidas a novas situações (raciocínio dedutivo). Esta bateria é

formada por quatro provas distintas, sendo três delas com itens semelhantes aos da BPR-5, que são, prova RA, prova RV e prova RN e uma exclusiva chamada de prova de raciocínio pratico RP. Todas as provas têm em comum a avaliação do raciocínio como uma habilidade específica relativa à Inteligência fluida. Para este trabalho foram analisadas as provas RA e RN.

SPSS, Winsteps: Trata-se de softwares de pacotes estatísticos necessários à realização da estimação dos parâmetros psicométricos, o primeiro foi usado para realizar a Análise de Regressão Múltipla, Análise de Grupamentos (Cluster) e as Correlações, o segundo, foi usado para estimar os parâmetros de dificuldade dos itens usando o modelo de Rasch da TRI.

Procedimento

O primeiro passo para realização do trabalho foi a estimação do parâmetro de dificuldade (b) dos itens usando o modelo de Rasch. Couto (2005) realizou uma equalização entre as formas A e B das cinco provas que compõe o BPR-5. As respostas dos sujeitos foram organizadas em um banco de dados comum, no caso, a prova de Raciocínio Abstrato (RA) dispunha de 19 itens comuns, e a prova de Raciocínio Numérico de 12 itens comuns que foram usados como ancoras e os parâmetros dos itens foram estimados por meio do modelo de Rasch. O resultado apontou um conjunto de 31 itens para a prova RA e 28 itens para a prova RN com o parâmetro b descrito em uma escala comum, com índices de ajuste considerados adequados. Os resultados encontrados no trabalho de Couto (2005) foram usados como referência para estimar os parâmetros psicométricos das duas provas no BPR-5 infantil (Cruz, 2007). Os valores de b de 10 itens que são comuns as três formas da BPR-5 (A, B e infantil) nas duas provas RA e RN, foram fixados para estimação dos parâmetros dos itens da forma infantil da BPR-5i.

Em seguida foram realizadas análises de conteúdo com objetivo de decompor os itens em fatores cognitivos de complexidade. Para tanto, na prova RA, foi realizada uma adaptação a partir dos conjuntos de categorias de classificação sugeridos por Jacobs e Vandeventer (1972) e Hornke e Habon (1986) que derivou em três conjuntos, denominados A, B e C, de categorias de regras de construção de itens considerados como fatores de complexidade. Para a prova RN foram adaptadas as categorias sugeridas por Simon e Kotovsky (1963), Simon (1972) e Holzman, Pellegrino e Glaser (1982, 1983).

O próximo passo foi a estimação das relações entre cada categoria e de cada uma delas com o valor de b . Foi realizada correlação ponto bisserial de Pearson entre o b e as variáveis do Conjunto A e B dos fatores de complexidade da prova RA, entre o b e os fatores de complexidade do Conjunto A da prova RN, e correlação de Pearson entre o b e as categorias do Conjunto C da prova RA e entre o b e os fatores do Conjunto B da prova RN. Em seguida as variáveis do conjunto A e B da prova RA e do conjunto A da prova RN foram submetidas a uma análise hierárquica de Cluster, usando o método de *Ward* com medidas de associação entre as variáveis binárias estimadas pelo padrão da diferença de distância entre elas. O objetivo da análise foi de identificar uma estrutura de dados formada pelo agrupamento de observações por similaridade. E por fim, com o objetivo de estimar como os fatores cognitivos descritos pelas categorias propostas (fatores de complexidade) prevêm o índice de dificuldade dos itens, foram realizadas Análises de Regressão Múltipla pelo método *Stepwise* usando o índice de dificuldade (b) dos itens, como variável dependente e os fatores de complexidade, separados por provas, como variáveis independentes. A partir dos resultados da análise cognitiva e psicométrica proposta foram desenvolvidas escalas com interpretação referenciada nos itens para as provas RA e RN da BPR-5.

Resultados e Discussão

Prova de Raciocínio Abstrato (RA)

O processo de equalização das três formas de apresentação da prova RA da BPR-5 apresentou como resultado um conjunto de 51 itens para avaliação do raciocínio abstrato, distribuídos em uma escala comum de dificuldade. Nas Tabelas 1 e 2 a primeira, segunda e terceira colunas apresentam os itens ordenados pela forma como eles aparecem na forma A, B e infantil da bateria e o gabarito de cada forma, respectivamente. A quarta coluna e as seguintes, respectivamente, apresentam o valor de b , os parâmetros de ajustamento e a correlação item-total. Os resultados da estimação mostram que o parâmetro b variou entre -3,75 e 2,42 ($M = -0,67$; $DP = 1,56$), indicando que, de um modo geral os itens podem ser considerados de dificuldade moderada em relação a amostra de respondentes, e a dispersão em torno da média aponta para uma frequência também razoável de valores altos de dificuldade, sendo o valor máximo referente a aproximadamente dois desvios padrão.

No caso das medidas de ajuste os parâmetros *infit* e *outfit* indicam situações incoerentes, nas quais, sujeitos com habilidade maior que a exigência de dificuldade de um item erram esse item, ou a situação contrária, na qual sujeitos com habilidade menor que a dificuldade de um item acertam o item. A literatura aponta que valores máximos para que essas medidas sejam consideradas adequadas variam entre 1,3 e 1,5 (Linacre & Wright, 1994a e 1994b; Wright & Stone, 2004). O *infit* descreve as situações de discrepância nas quais os valores de habilidade se encontram próximos aos valores de dificuldade dos itens. Os resultados mostraram valores variando entre 0,59 e 1,45 ($M = 0,97$; $DP = 0,17$), sugerindo que a média de *infit* dos itens é adequada, o valor da dispersão indica pouca variação entre os 51 itens, e que valores próximos ao máximo encontrado são pouco

frequentes na amostra examinada. Já o *outfit* descreve as situações extremas nas quais os valores de habilidade e dificuldade são distantes. Os resultados variaram entre 0,29 e 2,98 (M=1,03; DP=0,48) sugerindo que em média os itens são ajustados, porém o valor da dispersão indica uma variabilidade maior do que a encontrada nos *infits* e uma frequência razoável de valores mais próximos do máximo encontrado. De um modo geral pôde-se observar que por um lado os itens apresentaram ajuste adequado e dificuldade bem distribuída. Por outro lado, alguns itens necessitam de revisão que permita formular hipóteses explicativas sobre os desajustes. Os itens desajustados podem ser encontrados em destaque na Tabela 2 e em sua maioria pertencem ao conjunto da BPR-5i ainda em construção. No caso da correlação item-total os valores variaram entre 0,18 e 0,75 (M= 0,49; DP= 0,12), em média os valores são adequados e a dispersão indica uma boa homogeneidade entre os itens, esses resultados sugerem que os itens apresentam boa concordância entre si para avaliar o raciocínio abstrato, nesse caso apenas quatro itens apresentaram valores menores que 0,30; esses itens estão em destaque nas Tabelas 1 e 2.

Tabela 1. Equalização das três formas do RA

Ordem	BPR-5 A - B - i	Gabarito A - B - i	b	Infit	Outfit	Correlação Item-total
1	RA01_x_12	d_x_d	-1,79	0,86	0,89	0,55
2	RA02_01_x	b	-2,57	0,82	0,56	0,46
3	RA03_03_19	e_e_d	-1,42	0,73	0,49	0,59
4	RA04_x_x	a	-1,54	0,82	0,69	0,59
5	RA05_04_14	d_d_d	-1,18	0,71	0,53	0,62
6	RA06_05_15	c_a_c	-0,9	0,85	0,81	0,55
7	RA07_09_22	a_a_a	-0,52	0,93	0,83	0,53
8	RA08_06_x	d_d	-0,85	0,89	0,74	0,54
9	RA09_07_24	c_c_b	0,34	0,93	0,89	0,54
10	RA10_x_27	c_x_c	0,1	1,15	1,22	0,43
11	RA11_x_17	e_x_e	-0,36	0,96	0,89	0,55
12	RA12_x_x	c	-0,04	0,93	0,94	0,56
13	RA13_x_x	b	-0,45	1,11	1,27	0,45
14	RA14_10_x	c_c	-1,3	0,87	0,77	0,52
15	RA15_08_x	c_c	-0,08	1,06	1,12	0,46
16	RA16_12_21	d_d_d	0,29	1,01	1,02	0,49
17	RA17_13_x	a_a	-0,56	1,07	1,09	0,44
18	RA18_16_x	d_d	-0,07	1,10	1,13	0,43
19	RA19_14_23	b_b_b	0,64	0,98	1,00	0,50
20	RA20_15_x	a_a	1,37	0,99	1,04	0,48
21	RA21_18_x	a_a	1,75	1,23	1,62	0,29
22	RA22_19_x	e_e	0,86	1,12	1,14	0,41
23	RA23_24_x	e_d	0,87	1,06	1,19	0,43
24	RA24_22_x	a_a	1,82	1,04	1,52	0,40
25	RA25_23_x	e_e	2,42	1,03	1,52	0,36
26	RAx_02_x	x_c	-1,72	0,82	0,55	0,45

Tabela 2. Equalização das três formas do RA (continuação).

Ordem	BPR-5 A - B - i	Gabarito A - B - i	b	Infit	Outfit	Correlação Item-total
27	RAx_11_x	c ou e	-0,74	1,16	1,40	0,26
28	RAx_17_x	x_c	1,23	1,07	1,14	0,38
29	RAx_20_x	x_c	0,95	0,94	0,94	0,48
30	RAx_21_x	x_b	1,57	0,98	1,00	0,44
31	RAx_25_x	x_e	1,87	1,06	1,15	0,36
32	RAx_x_01	x_x_a	-1,87	1,01	0,94	0,55
33	RAx_x_02	x_x_e	-3,52	1,08	1,20	0,39
34	RAx_x_03	x_x_a	-2,62	0,78	0,76	0,61
35	RAx_x_04	x_x_e	-3,52	0,82	0,65	0,52
36	RAx_x_05	x_x_e	-3,75	0,71	0,29	0,57
37	RAx_x_06	x_x_a	-3,25	0,90	0,79	0,51
38	RAx_x_07	x_x_e	-3,25	0,85	0,57	0,55
39	RAx_x_08	x_x_b	-2,45	0,73	0,48	0,68
40	RAx_x_09	x_x_c	-2,79	0,83	0,55	0,61
41	RAx_x_10	x_x_b	-1,98	0,88	0,79	0,61
42	RAx_x_11	x_x_c	-2,27	0,59	0,40	0,75
43	RAx_x_13	x_x_e	-1,64	0,81	0,70	0,66
44	RAx_x_16	x_x_b	-0,77	1,03	1,08	0,53
45	RAx_x_18	x_x_c	-0,68	0,80	0,73	0,65
46	RAx_x_20	x_x_a	-0,58	0,82	0,80	0,63
47	RAx_x_25	x_x_c	-0,69	1,45	1,99	0,27
48	RAx_x_26	x_x_a	-0,25	1,30	1,88	0,32
49	RAx_x_28	x_x_d	0,45	1,13	2,02	0,35
50	RAx_x_29	x_x_d	0,56	1,37	2,98	0,18
51	RAx_x_30	x_x_b	0,79	1,15	1,59	0,35

Ao se proceder a análise do conteúdo dos 51 itens da prova RA, das três formas de apresentação da BPR-5 (Forma A, B e Infantil), foram identificados três conjuntos de categorias de regras de construção denominadas fatores de complexidade. Conjunto (A), formado por doze elementos que designam as partes dos itens: *Identidade, Forma, Sombreamento, Tamanho, Movimento no plano, Séries, Adição ou subtração de Elementos, Adição ou subtração de termos por combinação, Elemento do conjunto, Adição Única, Reflexão, Inversão*. Conjunto (B), formado pela relação entre os elementos presentes num item, cujas partes podem ser claramente separadas ou independentes uma da outra, *Separação de Componentes*; as características dos elementos presentes no item podem estar

integradas, por exemplo, o elemento muda de forma e textura ao longo das células da matriz, *Integração de Componentes*; e os elementos dentro de um item podem estar embutidos, ou seja, os elementos das células são combinados para formar a figura na próxima célula a partir de partes diferentes de uma gestalt, *Componentes Embutidos*. Conjunto (C), formado pela quantidade de elementos presente num item, *Número de Elementos*; a quantidade de transformações que ocorrem com os elementos dos itens, *Número de Transformações*; e quantos tipos diferentes de transformações ocorrem com os elementos dos itens, *Tipos de Transformações*.

À partir da classificação dos itens segundo as 18 categorias mencionadas, pôde-se notar que os itens não apresentavam as características descritas pelas categorias de *Identidade*, *Elementos do Conjunto* e *Adição Única*. Por essa razão essas três categorias foram descartadas. A tabela com o resultado da análise cognitiva que descreveu a decomposição dos itens nas partes descritas pelo conjunto de categorias acima pode ser encontrada no Anexo I. As 15 categorias restantes foram então analisadas quanto às suas relações entre si e com o parâmetro *b*. A tabela 3 mostra o valor das correlações entre as categorias do Conjunto C, como pode ser observado a maior relação foi entre o *Número de Transformações* e o *Tipo de Transformações*, o que era previsto uma vez que a segunda variável é derivada da primeira.

De um modo geral as variáveis desse conjunto estão bastante relacionadas entre si e também com o valor do parâmetro *b*, explicando, na pior das possibilidades, pelo menos 0,29% da variância da dificuldade. Uma explicação plausível é de que ao considerar um item quanto ao seu conjunto de regras de construção, quanto maior o número de elementos presentes, maior a tendência de aumentar o número de transformações sofridas pelos elementos e consequentemente aumentar os tipos de transformações possíveis e a maior

freqüência dessas regras aplicadas a um item aumenta consideravelmente a dificuldade desse item.

Tabela 3. Correlações entre os fatores de complexidade e a Dificuldade dos Itens.

	<i>b</i>	Número de Elementos	Número de Transformações
Número de Elementos	0,54**	1	-
Número de Transformações	0,55**	0,55**	1
Tipos de Transformações	0,57**	0,55**	0,95**

**p= 0,001; N= 51

A Tabela 4 mostra as variáveis do Conjunto A, B e o parâmetro *b*, que apresentaram correlações significativas. Pôde-se notar que apenas cinco variáveis de um total de 16 submetidas à análise apresentaram correlações significativas, dessas apenas as regras de *Movimento no Plano* e *Separação de Componentes* apresentaram relação positiva com o aumento da dificuldade explicando 17% e 15% da variância respectivamente. É interessante notar que a regra de *Série* apresentou correlação negativa e significativa com a dificuldade, permitindo a especulação de que quanto maior a freqüência desse tipo de regra presente em um item, mais fácil ele se torna.

Tabela 4. Correlações entre os fatores de complexidade e a Dificuldade dos Itens.

	Movimento no Plano	Separação de Componentes	Sombreado	Adição Elementos	Séries	Inversão	Componentes Embutidos
Separação de Componentes	0,15	1	-	-	-	-	-
Sombreado	-0,16	-0,15	1	-	-	-	-
Adição Elementos	0,05	0,15	-0,38**	1	-	-	-
Séries	-0,04	-0,10	-0,22	-0,12	1	-	-
Inversão	-0,04	-0,10	-0,07	-0,27	-0,08	1	-
Componentes Embutidos	0,08	-0,07	0,26	0,02	-0,06	0,32*	1
<i>b</i>	0,42**	0,39**	-0,04	0,18	-0,28*	0,23	0,23

**p=0,001; *p= 0,005; N= 51

Para explorar melhor a existência de uma possível estrutura de relações entre os aspectos cognitivos presentes nas regras de construção de itens dos Conjuntos A e B foi realizada uma análise de Agrupamentos. Os resultados apresentados na Tabela 5 mostram os coeficientes de associação formados pelo padrão de diferença de distância entre as variáveis e a ordem em que cada variável é selecionada para ser adicionada ao grupo. O valor do coeficiente de similaridade indica a proximidade das variáveis, quanto menor o valor desse coeficiente maior a probabilidade de formarem um agrupamento. A primeira combinação foi feita entre *Sombreamento* e *Componentes Embutidos*, essa associação foi notada anteriormente quando foram examinadas as correlações, porém naquele caso não eram significativas (Tabela 4). Em seguida são feitas as associações entre *Reflexão* e *Separação de Componentes* e também entre *Inversão* e *Integração de Componentes*. Esses três pares de associação constituem os núcleos dos três agrupamentos que podem ser apontados como a solução mais adequada para a análise. Seguindo com a adição de variáveis a regra de *Tamanho* é acrescentada ao segundo par e a regra de *Série* seguida de *Forma* são acrescentadas ao terceiro par respectivamente. A partir dessas associações o acréscimo de variáveis nos agrupamentos passa a exigir aproximadamente o dobro do valor do coeficiente de similaridade, nesse caso, reduzindo a homogeneidade dos grupos.

Tabela 5. Lista de Aglomeração.

Passo	Combinação Cluster		Coeficientes	Primeiro Estágio Cluster		Prox. Passo
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	2	11	0,000	0	0	9
2	7	9	0,001	0	0	5
3	8	10	0,002	0	0	4
4	5	8	0,005	0	3	6
5	3	7	0,009	0	2	7
6	1	5	0,013	0	4	7
7	1	3	0,021	6	5	8
8	1	4	0,036	7	0	9
9	1	2	0,055	8	1	10
10	1	6	0,080	9	0	0

A decisão sobre o número ideal de agrupamentos, é considerada importante na análise de Cluster, por exemplo, com a decisão de formar três agrupamentos as regras de *Movimento no Plano* e *Adição de Elementos*, foram consideradas como variáveis atípicas, ou seja, não apresentam características suficientemente comuns capazes de formar um agrupamento nitidamente visível no espaço representado no Dendograma. Essa decisão foi sustentada com referência nos valores do coeficiente de associação, que depois do sexto passo (Tabela 5) começa a apresentar valores próximos ao dobro dos encontrados nos passos anteriores. A Figura 22 apresenta dois dendogramas, sendo que o primeiro ilustra os resultados descritos na Tabela 5 no qual se pode observar a distância necessária para se considerar uma solução de dois agrupamentos incluindo a variável *Movimento no Plano* no segundo agrupamento. O segundo dendograma representa a associação entre as variáveis depois da exclusão daquelas consideradas atípicas e a realização de uma nova estimação, da mesma forma a solução em três agrupamentos continua sendo a ideal tendo em vista a distancia entre as variáveis em uma solução com dois agrupamentos.

primeira sozinha prevê aproximadamente 32% da variância de b , enquanto as outras quatro juntas prevêem 22%.

Tabela 6. Sumário do Modelo de Regressão.

Passos	R	R ²	EP	Estatísticas de Mudança			
				Mudança de R ²	Mudança de F	gl	p
1	0,565	0,319	1,29997	0,319	23,387	50	0,000
2	0,631	0,399	1,23365	0,080	6,520	49	0,014
3	0,676	0,457	1,18496	0,058	5,110	48	0,028
4	0,707	0,500	1,14837	0,044	4,107	47	0,048
5	0,735	0,541	1,11278	0,040	4,055	46	0,050

A Tabela 7 apresenta a ordem de entrada de cada variável na função de Regressão com os respectivos valores de Beta que podem ser interpretados como o valor da contribuição de cada variável para o modelo de predição da variável dependente. É interessante notar que se por um lado, as variáveis *Tipos de Transformação*, *Número de Elementos* e *Componentes Embutidos* contribuem positivamente com a previsão da variância de b , ou seja, quanto maior a frequência dessas variáveis mais difíceis os itens se tornam, por outro lado, as variáveis *Série* e *Sombreamento* apresentam uma contribuição negativa, essa relação pode ser interpretada como uma indicação de quanto maior a frequência desses dois tipos de regras está presente em um item, mais fácil esse item tende a ser.

Tabela 7. Coeficientes da Regressão Logística.

Passos		Coeficientes Não Padronizados		Coeficientes Padronizados	t	p
		B	EP	β		
1	Constante	-2,293	0,382		-6,010	0,000
	Tipos de Transformações	0,929	0,192	0,565	4,836	0,000
2	Constante	-2,167	0,365		-5,930	0,000
	Tipos de Transformações	0,929	0,182	0,565	5,096	0,000
	Série	-1,639	0,642	-0,283	-2,553	0,014
3	Constante	-2,648	0,411		-6,451	0,000
	Tipos de Transformações	0,667	0,210	0,405	3,171	0,003
	Série	-1,427	0,624	-0,246	-2,287	0,027
	Número de Elementos	0,486	0,215	0,291	2,260	0,028
4	Constante	-2,466	0,408		-6,045	0,000
	Tipos de Transformações	0,757	0,209	0,460	3,630	0,001
	Série	-1,720	0,622	-0,297	-2,767	0,008
	Número de Elementos	0,454	0,209	0,272	2,173	0,035
	Sombreamento	-0,704	0,347	-0,220	-2,027	0,048
5	Constante	-2,436	0,396		-6,158	0,000
	Tipos de Transformações	0,701	0,204	0,426	3,438	0,001
	Série	-1,698	0,602	-0,293	-2,819	0,007
	Número de Elementos	0,485	0,203	0,290	2,388	0,021
	Sombreamento	-0,862	0,346	-0,269	-2,494	0,016
	Componentes Embutidos	1,691	0,840	0,211	2,014	0,050

A Figura 23 apresenta o gráfico de dispersão entre os valores Z preditos pela função de Regressão e os valores da variável dependente b . Foi traçada a linha de ajuste ao modelo de previsão com referencia ao valor de R^2 . A distribuição dos pontos no eixo de X representa as variações de dificuldade dos itens ao passo que a distribuição no eixo de Y

representa as variações no valor de Z calculado pela função. Quanto mais próximos os pontos estiverem da reta melhor a predição do modelo, e maior seria o valor de R^2 , como pode ser observado o modelo apresenta uma boa predição.

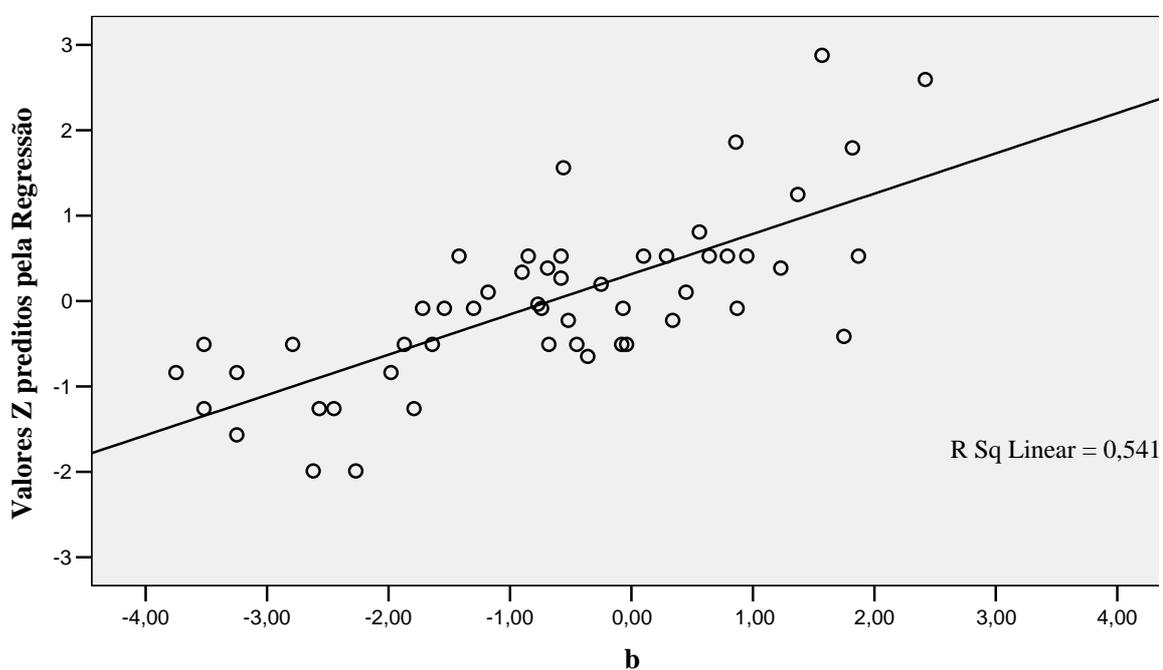


Figura 23. Gráfico de Dispersão da função de Regressão para prova RA.

Prova de Raciocínio Numérico (RN)

No caso da prova RN, o processo de equalização das três formas da BPR-5 apresentou como resultado um conjunto de 48 itens para avaliação do raciocínio numérico, distribuídos em uma escala comum de dificuldade. As Tabelas 8 e 9 apresentam os itens

ordenados como aparecem na forma A, B e infantil da bateria, respectivamente. Os resultados da estimação mostram que o parâmetro b variou entre -8,83 e 2,67 ($M = -1,74$; $DP = 3,08$), indicando que, de um modo geral os itens podem ser considerados fáceis, a dispersão em torno da média aponta para uma heterogeneidade dos valores de dificuldade, ou seja, alguns itens são muito fáceis e outros muito difíceis, nesse caso o valor mínimo chama atenção, pois está a aproximadamente três desvios padrão abaixo da média.

Tratando das medidas de ajuste *infit* e *outfit*, os resultados, para o *infit*, mostraram valores variando entre 0,72 e 1,44 ($M = 1,00$; $DP = 0,15$), o valor máximo e da média de *infit* dos itens podem ser considerados adequados. O valor da dispersão indica pouca variação entre os 48 itens, ou seja, que valores próximos ao máximo encontrado são pouco frequentes na amostra, sugerindo que os itens estão ajustados. Já para o *outfit* os resultados variaram entre 0,24 e 9,90 ($M = 1,56$; $DP = 1,86$) sugerindo que em média os itens apresentam pequena violação, porém o valor da dispersão indica uma variabilidade maior do que a encontrada nos *infits*, ou seja, uma frequência razoável de valores mais altos que podem ser considerados desajustados. Nesse caso, pode-se dizer que por um lado a maioria os itens apresentaram ajuste adequado e dificuldade bem distribuída. Por outro lado, os itens que apresentaram desajuste, em sua maioria com *outfit* acima do recomendado, apresentam valores de dificuldade muito baixos. Esse fato permite formular a hipótese de falta de atenção do examinado ao responder ao item para explicar os desajustes. Os itens desajustados podem ser encontrados em destaque na Tabela 8 e 9, e da mesma forma que ocorreu na prova RA a maioria pertence ao conjunto da BPR-5i. Examinando os valores da correlação item-total os valores variaram entre 0,30 e 0,75 ($M = 0,57$; $DP = 0,12$), em média os valores podem ser considerados muito bons e a dispersão indica uma boa homogeneidade entre os itens, esses resultados sugerem que os itens apresentam boa

concordância entre si para avaliar o raciocínio abstrato, nesse caso nenhum item apresentou valor menor que 0,30.

Tabela 8. Equalização das três formas do RN.

Ordem	BPR-5 A – B - i	Gabarito A - B - i	b	Infit	Outfit	Correlação Item-total
1	RN01_01_16	18 21	-4,34	1,29	2,87	0,35
2	RN02_x_22	17 17	-0,90	0,92	0,87	0,70
3	RN03_02_17	51 56	-3,43	1,14	2,63	0,43
4	RN04_x_21	28 26	-0,89	1,08	1,16	0,65
5	RN05_04_27	6 3	-1,62	1,11	1,22	0,53
6	RN06_x_20	9 12	-1,45	0,76	0,67	0,75
7	RN07_05_x	20 25	-0,23	1,01	1,00	0,58
8	RN08_x_x	2 30	-1,36	0,82	0,89	0,73
9	RN09_13_x	48 19	-0,21	0,83	0,76	0,65
10	RN10_06_23	10 3	-0,32	0,99	0,91	0,59
11	RN11_09_x	75 28	0,06	0,92	0,85	0,61
12	RN12_x_x	81 243	1,52	1,18	1,86	0,44
13	RN13_10_26	44 45	0,39	1,03	1,00	0,56
14	RN14_11_x	19 27	1,08	0,96	0,89	0,57
15	RN15_12_x	26 16	0,35	0,84	0,84	0,64
16	RN16_x_x	172 36	1,02	0,93	0,84	0,61
17	RN17_14_x	12 1	1,09	1,00	0,93	0,56
18	RN18_x_x	50 51	1,44	0,87	0,80	0,60
19	RN19_18_x	11 12	1,56	1,07	1,04	0,51
20	RN20_x_x	74 138	2,67	0,88	0,55	0,45
21	RNx_03_19	11 3	-2,78	0,92	1,06	0,43
22	RNx_07_25	5 64	-1,89	1,04	1,56	0,42
23	RNx_08_x	23 29	0,31	1,13	1,16	0,50
24	RNx_15_x	7 39	1,13	0,87	0,76	0,62

Tabela 9. Equalização das três formas do RN (continuação).

Ordem	BPR-5 A – B	Gabarito A - B - i	b	Infit	Outfit	Correlação Item-total
25	RNx_16_x	37 42	1,31	0,98	0,92	0,57
26	RNx_17_x	105 110	0,85	1,14	1,13	0,49
27	RNx_19_x	12 15	2,25	0,91	0,78	0,58
28	RNx_20_x	120 720	2,40	0,93	0,95	0,54
29	RNx_x_01	6 7	-8,83	1,44	9,90	0,40
30	RNx_x_02	1 1	-7,05	1,20	0,93	0,56
31	RNx_x_03	2 3	-6,25	1,27	1,03	0,60
32	RNx_x_04	0 4	-5,50	1,22	1,51	0,62
33	RNx_x_05	2 3	-4,41	1,35	3,30	0,59
34	RNx_x_06	4 3	-5,77	1,06	0,59	0,67
35	RNx_x_07	5 1	-6,20	1,02	2,46	0,64
36	RNx_x_08	5 10	-5,25	0,81	2,44	0,70
37	RNx_x_09	6 5	-5,21	0,89	1,40	0,69
38	RNx_x_10	10 14	-5,17	0,72	0,87	0,73
39	RNx_x_11	24 0	-5,21	1,07	0,93	0,67
40	RNx_x_12	20 4	-4,94	0,93	2,04	0,68
41	RNx_x_13	14 13	-4,94	0,92	1,12	0,70
42	RNx_x_14	30 1	-4,69	0,86	1,07	0,71
43	RNx_x_15	35 0	-4,98	0,85	0,71	0,72
44	RNx_x_18	7 5	-2,73	0,97	1,12	0,65
45	RNx_x_24	22 31	-0,93	0,87	0,90	0,58
46	RNx_x_28	38 76	1,98	0,96	0,24	0,30
47	RNx_x_29	6 4	1,10	1,08	1,48	0,32
48	RNx_x_30	23 19	1,60	0,96	9,90	0,30

Dando prosseguimento às análises necessárias ao cumprimento dos objetivos desse trabalho, examinou-se o conteúdo do conjunto de 48 itens da prova RN, das três formas de apresentação da BPR-5 (A, B e Infantil). Como resultado foram identificados dois conjuntos de categorias de regras de construção denominadas fatores de complexidade. O primeiro conjunto chamado de A foi formado pelas seguintes variáveis: *Tipo de Operações*, que diz respeito à quantidade de categorias de operações matemáticas que foram aplicadas aos elementos da série. *Ambigüidade de Operações*, essa variável diz respeito à possibilidade de se aplicar mais de uma regra para resolver corretamente o item, quando

isso era possível o item recebia nota um, do contrário nota zero. *Categorias de Operações* informa quais os tipos de operação foram aplicadas nos elementos da série, foi subdividida em *Identidade*, quando um elemento se repete ao longo da série sem sofrer nenhuma operação matemática. *Adição*, quando essa é a regra aplicada aos elementos da seqüência para se induzir a resposta correta, o item recebe nota um nessa categoria. *Subtração*, quando essa é a regra aplicada aos elementos da seqüência para se induzir a resposta correta, o item recebe nota um nessa categoria. *Multiplificação*, quando essa é a regra aplicada aos elementos da série para se induzir a resposta correta, o item recebe nota um nessa categoria. *Divisão*, quando essa é a regra aplicada aos elementos da série para se induzir a resposta correta, o item recebe nota um nessa categoria. *Complexidade*, quando é possível identificar relações hierárquicas entre as operações matemáticas aplicadas nos elementos ao longo da série o item recebe nota um nessa categoria.

O segundo conjunto de regras chamado de B é composto pelas seguintes variáveis: *Número de Operações*, variável que representa a quantidade de vezes que operações diferentes ou envolvendo valores diferentes foram aplicadas aos elementos ao longo da série. *Memória de Trabalho*, informa a quantidade de informação deve ser manipulada ou ficar disponível na memória de trabalho para a solução adequada do item. *Padrão de Descrição*, diz respeito a quantidade de regras que estão aplicadas aos elementos da série são necessárias para se chegar a resposta correta, lembrando que nem sempre é necessário descobrir todas as regras para se chegar a resposta correta que completa a série. *Número de Períodos* trata da quantidade de séries que podem ser encontradas pela combinação de elementos dentro de uma seqüência. *Comprimento da Seqüência* refere-se ao número de elementos que compõe a seqüência apresentada no item. Todos os 48 itens foram analisados e pontuados segundo esse conjunto de categorias de regras de construção, a

tabela com a pontuação dos itens em cada categoria pode ser encontrada no Anexo II desse trabalho.

O próximo passo foi verificar qual a relação entre as categorias e a dificuldade dos itens. A Tabela 10 mostra os resultados da correlação ponto-bisserial entre as categorias do Conjunto A e o parâmetro de dificuldade (*b*) dos itens. Como pode ser observado existe pequena frequência de associação entre os fatores de complexidade, e dentre eles as variáveis que estão mais associadas com a dificuldade são *Complexidade* com 44% de variância explicada, *Identidade* com 32% de variância explicada, sendo que a correlação negativa indica que quanto maior a presença dessa variável menor a dificuldade do item e *Multiplicação* com aproximadamente 12% de variância explicada. Outra associação interessante é entre as *Categorias de Operação*, no caso a *Multiplicação* e *Divisão* apresentam uma associação moderada e inversa com *Subtração* e *Adição*, respectivamente. Essa associação pode sugerir direções opostas de influência entre esses variáveis e a dificuldade aos itens.

Tabela 10. Correlações entre complexidade e Dificuldade dos Itens da prova RN.

	Tipos de Operações	Ambigüidade de Operações	Identidade	Adição	Subtração	Multiplicação	Divisão	Complexidade
Ambigüidade de Operações	-0,05	1	-	-	-	-	-	-
Identidade	0,36*	-0,27	1	-	-	-	-	-
Adição	0,40**	0,12	-0,09	1	-	-	-	-
Subtração	0,21	0,11	-0,34*	-0,25	1	-	-	-
Multiplicação	0,13	-0,16	-0,17	-0,10	-0,31*	1	-	-
Divisão	-0,03	0,24	-0,15	-0,31*	0,06	-0,09	1	-
Complexidade	-0,40	0,15	-0,29**	0,07	0,06	0,10	0,08	1
<i>b</i>	0,08	0,24	-0,57**	0,13	-0,20	0,36*	0,10	0,67**

**p=0,001; *p= 0,005; N= 48

A Tabela 11 apresenta o resultado das correlações de Pearson entre as variáveis do Conjunto B e o *b* dos itens. Como pôde ser observado, as seis variáveis apresentaram níveis de associação de moderado a alto, porém o valor de nenhuma associação justifica a retirada de qualquer variável entre os fatores de complexidade da análise de previsão da dificuldade pelo excesso de variância compartilhada. A maior variância explicada de um fator de complexidade em relação a outro é entre a variável *Memória de Trabalho e Comprimento da Seqüência* de aproximadamente 40%. Esse resultado é esperado, uma vez que quanto maior a seqüência aumentam-se as possibilidades de aumentar também o número de elementos que necessitam ser manipuladas mentalmente pelo sujeito para se obter a resposta correta.

Tabela 11. Correlações entre complexidade e Dificuldade dos Itens da prova RN.

	Número de Operações	Memória de Trabalho	Padrão de Descrição	Número de Períodos	Comprimento da Seqüência
Memória de Trabalho	0,60**	1	-	-	-
Padrão de Descrição	0,50**	0,47**	1	-	-
Número de Períodos	0,18	0,34*	0,38**	1	-
Comprimento da Seqüência	0,49**	0,64**	0,23	0,41**	1
<i>b</i>	0,64**	0,74**	0,41**	0,08	0,36*

**p=0,001; *p= 0,005; N= 48

Em seguida foi realizada uma análise de Agrupamentos com o objetivo de explorar melhor a existência de uma possível estrutura de relações entre os fatores de complexidade descritos no Conjunto A da prova RN. Os resultados apresentados na Tabela 12 mostram os coeficientes de associação formados pelo padrão de diferença entre as variáveis que foi o critério adotado para estabelecer a ordem em que cada variável é selecionada para ser adicionada ao grupo. A primeira combinação foi feita entre *Tipos de Operações* e

Complexidade, essa associação foi notada anteriormente quando foram examinadas as correlações, porém naquele caso não eram significativas (Tabela 10), e entre *Ambigüidade de Operações* e *Adição*. Em seguida, com pequeno aumento no coeficiente de similaridade, são feitas as associações entre *Multiplicação* e *Divisão*. Esses três pares de associação constituem os núcleos dos três agrupamentos que podem ser apontados como a solução mais adequada para a análise. O próximo passo é a adição da variável *Identidade* ao terceiro par. A partir dessas associações o acréscimo de variáveis nos agrupamentos passa a exigir mais que o dobro do valor do coeficiente de similaridade, nesse caso, reduzindo a homogeneidade dos grupos.

Tabela 12. Lista de Aglomeração.

Passo	Combinação Cluster		Coeficientes	Primeiro Estágio Cluster		Prox. Passo
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	1	8	0,000	0	0	5
2	2	4	0,000	0	0	5
3	6	7	0,002	0	0	4
4	3	6	0,008	0	3	6
5	1	2	0,017	1	2	7
6	3	5	0,032	4	0	7
7	1	3	0,073	5	6	0

A Figura 24 apresenta dois dendogramas, no primeiro são ilustrados os resultados apresentados na Tabela 12 é se pode observar a distância necessária para se considerar uma solução de dois agrupamentos que, nesse caso, incluem a junção do primeiro e segundo agrupamentos e a adição da variável *Subtração* no terceiro agrupamento. O segundo dendograma representa a associação entre as variáveis depois da exclusão da *Subtração* considerada atípica e a realização de uma nova estimação. Da mesma forma a solução em

três agrupamentos continua sendo a ideal tendo em vista a perda de homogeneidade dos grupos em uma solução com dois agrupamentos. A perda de homogeneidade indica uma mudança no perfil de combinações de características de complexidade, ou seja, uma diminuição das correlações entre essas características.

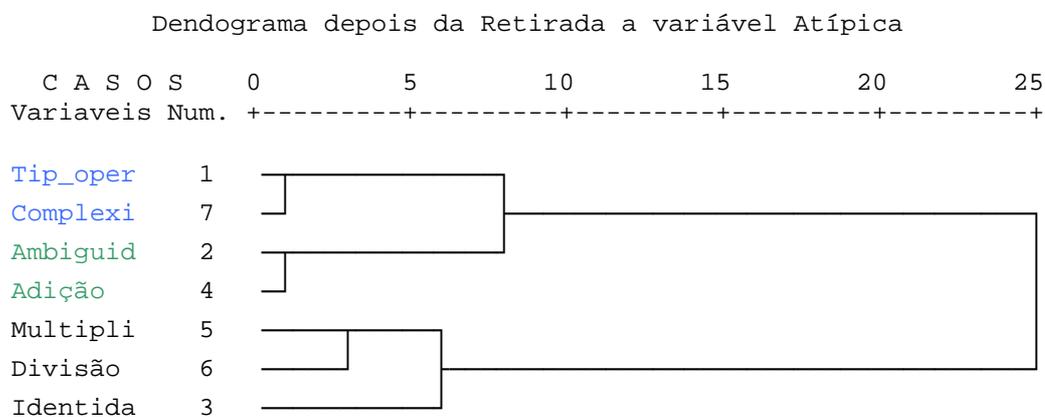
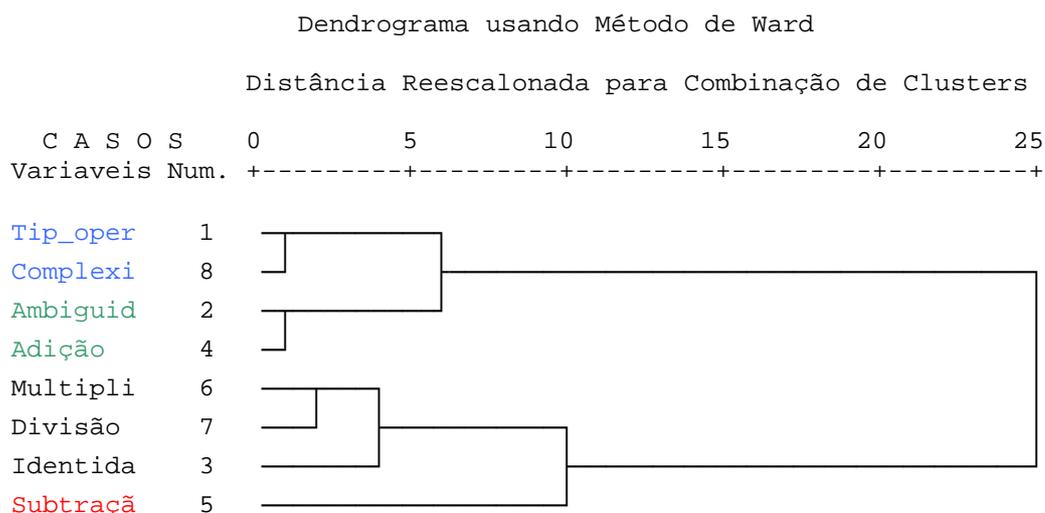


Figura 24. Dendogramas de Cluster Hierárquico.

Foi realizada uma análise de regressão múltipla para se verificar como os fatores de complexidade prevêm as variações do parâmetro de dificuldade dos itens. A Tabela 6 apresenta o resultado dessa análise que apontou uma função composta por quatro variáveis que são capazes de explicar aproximadamente 75% da variância da dificuldade dos itens. A

seqüência de passos mostra que a variável que mais prevê a dificuldade é *Memória de Trabalho* e a que menos prevê é *Multiplicação*, sendo que a primeira sozinha prevê aproximadamente 54% da variância de *b* enquanto as outras três o juntas prevêem 20%.

Tabela 13. Sumário do Modelo de Regressão para prova RN.

Passos	R	R ²	EP	Estatísticas de Mudança			
				Mudança de R ²	Mudança de F	gl	p
1	0,736	0,542	2,10798	0,542	54,389	46	0,000
2	0,801	0,641	1,88689	0,099	12,412	45	0,001
3	0,835	0,698	1,75012	0,057	8,308	44	0,006
4	0,864	0,747	1,62021	0,049	8,339	43	0,006

Na Tabela 14 são apresentadas a ordem de entrada de cada variável na função de regressão com os respectivos valores de Beta. O valor desse coeficiente é interpretado como a contribuição de cada variável para o modelo de predição da variável dependente, que nesse caso é a dificuldade dos itens. É importante ressaltar que por um lado, as variáveis *Memória de Trabalho* e *Número de Operações* contribuem positivamente com a previsão da variância de *b*, ou seja, quanto mais freqüente esse tipo de regra é presente nos itens mais difíceis eles se tornam, por outro lado, as variáveis de *Categorias de Operação de Identidade* e *Multiplicação* apresentam uma contribuição negativa, essa relação pode ser interpretada como uma indicação de quanto maior a freqüência desses dois tipos de regras está presente em um item, mais fácil esse item tende a ser.

Tabela 14. Coeficientes da Regressão Logística para a prova RN.

Passos		Coeficientes Não Padronizados		Coeficientes Padronizados		
		B	EP	β	t	p
1	Constante	-7,118	0,791		-9,004	0,000
	Memória de Trabalho	3,354	0,455	0,736	7,375	0,000
2	Constante	-5,457	0,850		-6,418	0,000
	Memória de Trabalho	2,775	0,439	0,609	6,321	0,000
	Identidade	-2,195	0,623	-0,339	-3,523	0,001
3	Constante	-5,449	0,789		-6,909	0,000
	Memória de Trabalho	1,966	0,495	0,431	3,976	0,000
	Identidade	-2,147	0,578	-0,332	-3,714	0,001
	Número de Operações	0,437	0,151	0,299	2,882	0,006
4	Constante	-5,615	0,732		-7,667	0,000
	Memória de Trabalho	1,831	0,460	0,402	3,979	0,000
	Identidade	-1,962	0,539	-0,303	-3,639	0,001
	Número de Operações	0,449	0,140	0,308	3,200	0,003
	Multiplificação	1,952	0,676	-0,226	2,888	0,006

A Figura 25 apresenta o gráfico de dispersão entre os valores Z preditos pela função de Regressão e os valores da variável dependente b . Foi traçada a linha de ajuste ao modelo de previsão com referencia ao valor de R^2 . A distribuição dos pontos no eixo de X representa as variações de dificuldade dos itens ao passo que a distribuição no eixo de Y representa as variações no valor de Z calculado pela função. Quanto mais próximos os

pontos estiverem da reta melhor a predição do modelo, e maior seria o valor de R^2 , como pode ser observado o modelo apresenta uma predição considerada ótima.

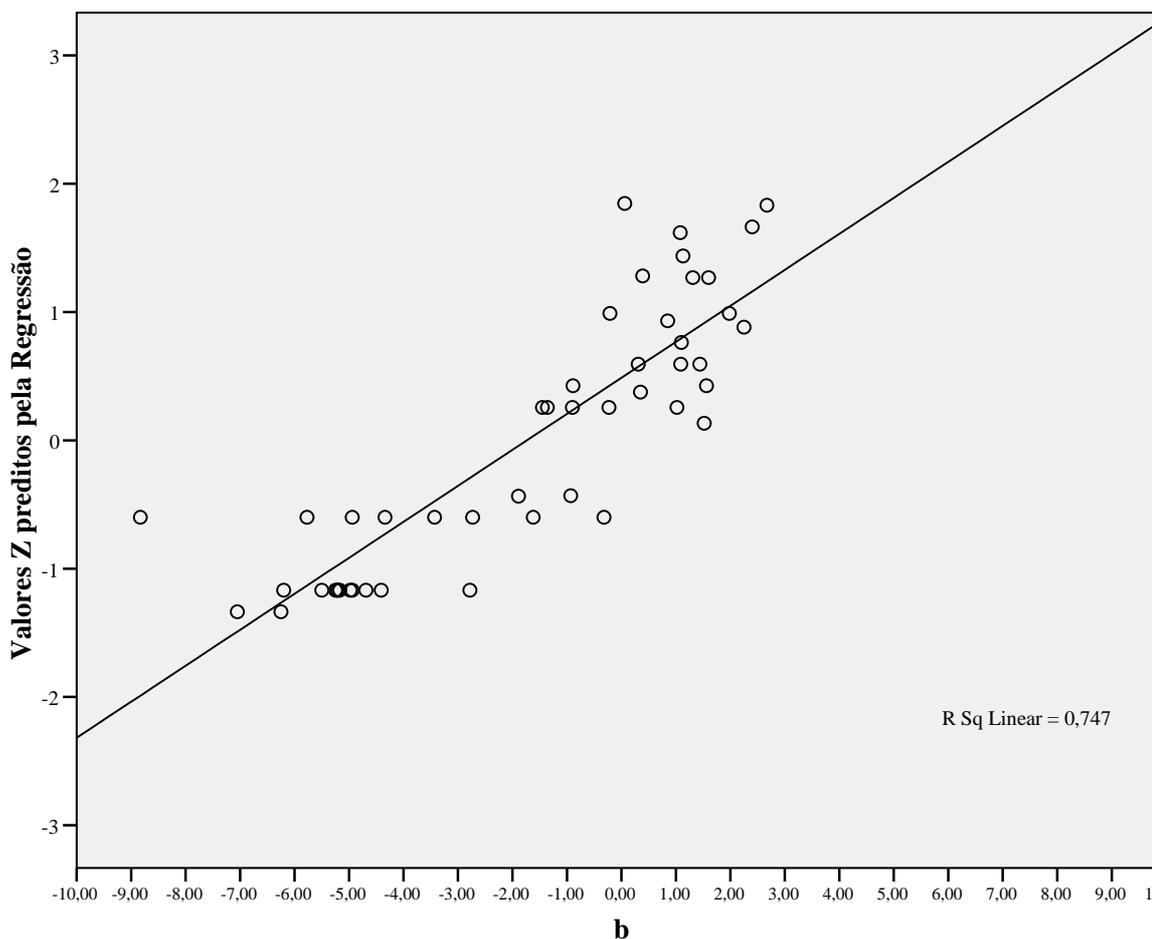


Figura 25. Gráfico de Dispersão da função de Regressão para prova RN.

Conclusão das Escalas

A análise dos resultados de associação entre os fatores de complexidade desenvolvidos nesse estudo, e sua conseqüente associação com a previsibilidade da dificuldade dos itens, ofereceu subsídios para tecer hipóteses sobre as características do escalonamento dos itens. Como os fatores de complexidade desenvolvidos pela análise cognitiva foram testados em sua capacidade de previsão da dificuldade dos itens, a presença

de um fator ou conjunto de fatores fornece uma evidência empírica para a explicação de porque os itens variam em termos de dificuldade.

Em primeiro lugar, a partir dos resultados da análise de regressão, puderam ser enumerados os fatores que melhor prevêem a dificuldade, tanto aqueles associados ao seu aumento quanto os associados a sua diminuição. Por exemplo, a contribuição do fator de complexidade *Série* para explicar a dificuldade dos itens da prova RA. Um aspecto a ser ressaltado sobre o resultado dessa análise é que as variáveis que não entraram na função de previsão também apresentam contribuição para a dificuldade dos itens, a sua exclusão da função se deu em razão da colinearidade entre as variáveis. Em segundo lugar, a colinearidade entre as variáveis pôde ser observada pelo perfil de combinação apresentado pelos fatores de complexidade, oferecidos pela análise de aglomerados. Pode-se dizer, de maneira geral, que os aglomerados foram formados pela presença ou ausência conjunta de cada fator em um item, produzindo uma medida de similaridade que cada variável do conjunto compartilha. Por exemplo, no aglomerado formado pelos fatores *Inversão*, *Componentes Integrados*, *Série* e *Forma*, da prova RA, apenas a forma entrou na função de regressão.

Os resultados dessas análises oferecem informações sobre as relações entre os fatores de complexidade e o parâmetro de dificuldade dos itens (b), permitindo estabelecer um tipo de escalonamento no qual a explicação da dificuldade é feita a partir das características de cada fator de complexidade. É possível também fazer especulações sobre o conjunto de manipulações mentais que foram realizadas pelos sujeitos durante a realização das tarefas. Nesse caso, como processos cognitivos são atributos dos sujeitos e não dos itens, deve-se ressaltar que tais manipulações são inferidas das características das tarefas, como se em cada item estivessem cristalizados o conjunto de operações mentais que foram aplicadas pelo autor na sua construção.

No caso da prova RA, os resultados apontam que os fatores que melhor prevêem a dificuldade são os *Tipos de Transformações*, *Número de Elementos*, *Componentes Embutidos*, *Série* e *Sombreamento*. Os três primeiros, por um lado, apresentaram uma relação positiva com a previsão da dificuldade, isso quer dizer que quanto maior o número de elementos que passam por tipos diferentes de transformações presente em um item maior sua dificuldade. Os dois últimos, por outro lado, apresentaram uma relação negativa com a dificuldade, sugerindo que quando um item apresenta elementos que passam por essas transformações mais fácil ele se torna.

As variações nos valores da dificuldade dos itens podem ser entendidas em termos da quantidade de informações a ser processada na memória e a algumas características dessas informações que estão envolvidas na realização adequada da tarefa. Por exemplo, na prova RA, 54% da variância da dificuldade pode ser explicada por uma função que envolve cinco variáveis, sendo que as duas que mais contribuem com a previsão são ligadas a quantidade de informações a ser processada *Tipos de Transformações* 43% e *Numero de Elementos* 29% enquanto que as características da informação a ser processada contribui com 29% com a previsão do modelo. Uma explicação geral da dificuldade incide sobre a quantidade de transformações diferentes que são aplicadas aos itens e na quantidade de elementos que os compõe, porém um item com um grande número de elementos e transformações pode se tornar mais fácil que outro com menos elementos dependendo do tipo de transformações que lhe foram aplicadas, no caso, *Série* e *Sombreamento* são consideradas transformações que se relacionam com a “facilidade” do item.

A Figura 26 mostra o escalamento dos itens da prova RA segundo o parâmetro b . A esquerda estão posicionados os sujeitos conforme as habilidades distribuídas em uma faixa de Theta, à direita estão posicionados os itens conforme o valor de sua dificuldade

distribuída em uma faixa de b . Nota-se que tanto Theta quanto b são apresentados em uma escala comum, sendo essa característica importante ao se acrescentar o resultado da análise dos fatores de complexidade para explicação da dificuldade. A letra M que aparece na barra central indica a média da medida, as letras S e T indicam respectivamente a posição do primeiro e segundo desvios padronizados em relação à média. A escala varia de -4 à 4 em valores de Theta e a média de habilidade dos sujeitos é aproximadamente um ponto superior a média dificuldade dos itens caracterizando que o teste é relativamente fácil para essa população. Em seguida são discutidas as interpretações propostas para cada faixa de dificuldade (b).

Faixa de b de -4 à -3: Nesta faixa de dificuldade são encontrados os itens que apresentam em suas regras de construção, de modo geral, apenas um elemento que passa por apenas uma transformação na seqüência da analogia ($A \rightarrow B$ e $C \rightarrow ?$). As transformações variam comumente entre os tipos de fatores de complexidade de *Sombreamento* e *Tamanho*. Sujeitos com essa faixa de habilidade na prova RA apresenta capacidade de resolver problemas que envolvam mudanças simples relacionadas a alterações de tonalidade de sombreamento dos elementos e alterações no seu tamanho relativo.

Faixa de b de -3 à -2: Nesta faixa de dificuldade os itens apresentam em suas regras de construção, comumente, um ou dois elementos que passam por uma transformação que varia entre os fatores de complexidade de *Sombreamento*, *Tamanho* e *Séries*. Sujeitos com essa faixa de habilidade apresentam capacidade de resolver problemas que envolvam mudanças simples relacionadas a alterações no sombreamento dos elementos, no seu tamanho relativo, e a repetição de elementos em escala progressiva.

Faixa de b de -2 à -1: Nesta faixa de dificuldade os itens apresentam, de modo geral, entre um, dois e três elementos que sofrem uma ou duas transformações, na maioria

das vezes de apenas um tipo. Nesse caso, as transformações variam além dos tipos de complexidade de *Sombreamento*, *Séries*, *Tamanho*, aparecem também *Inversão*, *Reflexão* e *Forma*. Sujeitos com essa faixa de habilidade apresentam além da capacidade de resolver problemas que envolvam mudanças simples nos seus elementos caracterizadas por alterações no sombreamento, no seu tamanho relativo, e a repetição de elementos em escala progressiva, são capazes também de identificar a inversão de regras entre os elementos, o espelhamento do item, e alterações sistemáticas na forma.

Faixa de *b* de -1 à 0: Nesta faixa de dificuldade os itens apresentam entre dois e três elementos que passam por duas transformações, de tipos diferentes. As transformações simples de *Tamanho* são substituídas pela regra de *Inversão*, e mantêm os tipos de *Sombreamento*, *Séries*, *Reflexão* e *Forma*. Sujeitos com essa faixa de habilidade apresentam a capacidade de resolver problemas que envolvam a combinação de mudanças nos seus elementos, caracterizadas pelas alterações citadas na faixa de dificuldade anterior, porém aplicada a um número maior de elementos.

Faixa de *b* de 0 à 1: Nesta faixa de dificuldade os itens apresentam dois e três elementos que sofrem duas ou três transformações diferentes combinadas. Os fatores de *Série* não são mais observados e as transformações combinam além dos fatores de complexidade citados anteriormente a *Adição de Elementos*, *Separação de Componentes*, *Integração de Componentes e Movimento no Plano*. Sujeitos com essa faixa de habilidade apresentam a capacidade de resolver problemas que envolvam a combinação de mudanças nos seus elementos caracterizadas por alterações combinando, por exemplo, a tonalidade do sombreamento e tamanho relativo; identificação da inversão de regras entre os elementos, o espelhamento do item e alterações sistemáticas na forma; integração de partes de elementos do item ou a separação dos elementos em unidades independentes umas das outras de modo a não formar um conjunto unificado.

Faixa de *b* de 1 à 2: Nesta faixa de dificuldade os itens apresentam três elementos que sofrem três transformações diferentes combinadas. As transformações podem combinar todos os fatores de complexidade citados anteriormente. Sujeitos com essa faixa de habilidade apresentam a capacidade de resolver problemas que envolvam combinações complexas de mudanças nos seus elementos especialmente a separação dos elementos em unidades independentes umas das outras de modo a não formar um conjunto unificado e o movimento dos elementos no plano alternando mudanças de posição, tamanho e sombreado.

Faixa de *b* de 2 à 3: Nesta faixa de dificuldade os itens apresentam três ou mais elementos que sofrem três transformações diferentes combinadas. As transformações podem combinar todos os fatores de complexidade citados anteriormente. Sujeitos com essa faixa de habilidade apresentam a capacidade de resolver problemas envolvendo todas as características citadas acima, nos dois sentidos da analogia ($A \leftrightarrow B \quad C \leftrightarrow ?$). Isto é, o sujeito é capaz de identificar as mudanças progressivas nos componentes do item quando ele muda da posição A para B na analogia, e aplicar a lógica da mudança ao termo C, para encontrar o termo D que melhor representa as mudanças identificadas, mas também, é capaz de inverter a lógica quando o termo C apresenta o componente já incorporando as mudanças inferidas de A para B, e a tarefa passa então a ser de identificar qual alternativa representa melhor o componente antes de sofrer as transformações.

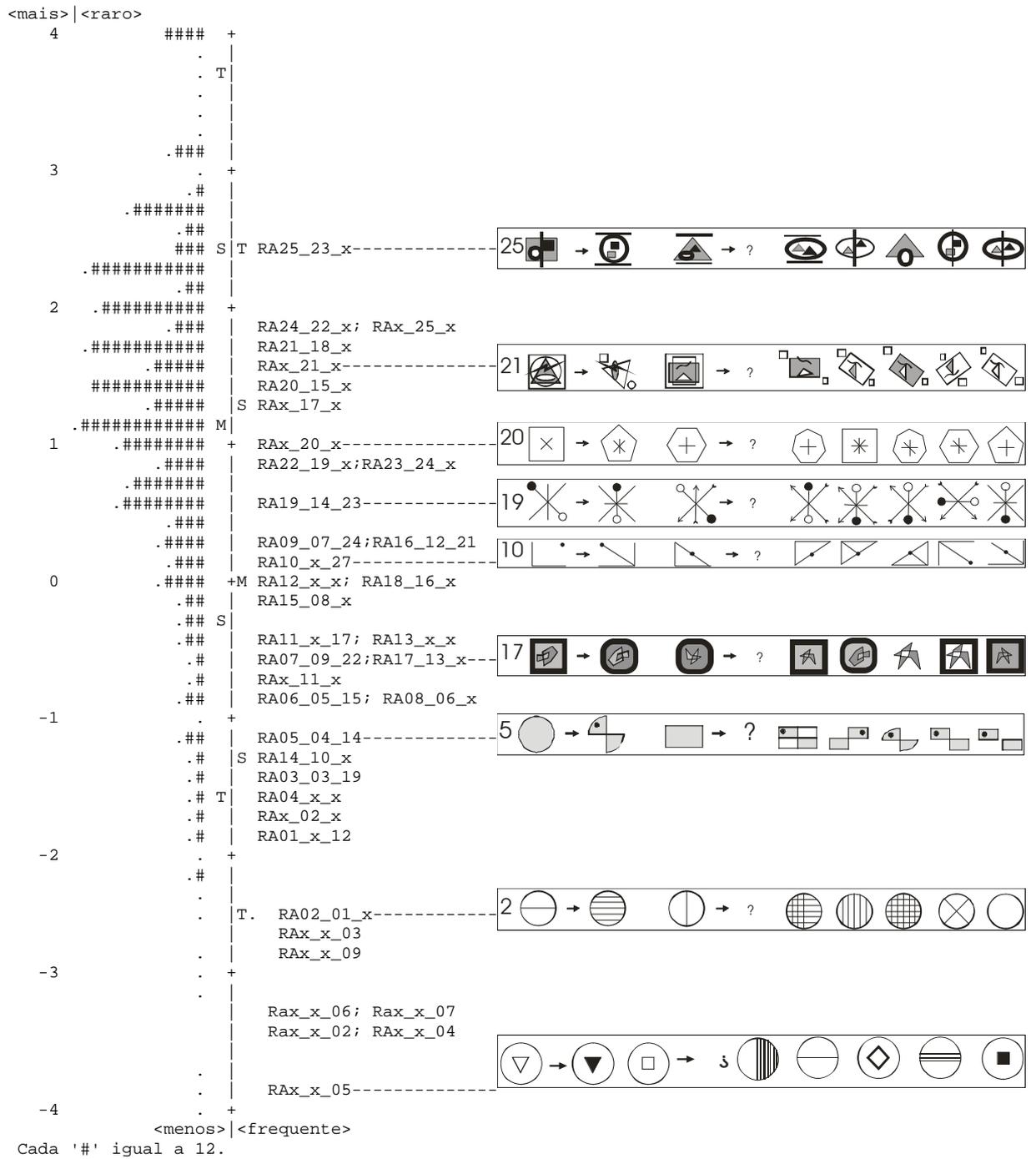


Figura 26. Gráfico de Escalonamento dos Itens da prova RA pela dificuldade.

No caso da prova RN, os resultados apontam que os fatores que melhor prevêm a dificuldade são *Memória de Trabalho, Número de Operações, Identidade, Multiplicação*.

Os dois primeiros apresentaram uma relação positiva com a previsão da dificuldade, isso quer dizer que quanto maior o número de informações que necessitam estar disponíveis ou de operações a serem realizadas, presente em um item, maior sua dificuldade. Os dois últimos apresentaram uma relação negativa com a dificuldade, sugerindo que quando um item apresenta operações simples de identidade ou de multiplicação mais fácil ele se torna.

Da mesma forma, as variações nos valores da dificuldade dos itens dessa prova podem ser entendidas em termos da quantidade de informações a ser processada na memória, e a algumas características dessas informações que estão envolvidas na realização adequada da tarefa. Por exemplo, aproximadamente 75% da variância da dificuldade pode ser explicada por uma função que envolve quatro variáveis, sendo que as duas que mais contribuem com a previsão são ligadas a quantidade de informações a ser processada *Memória de Trabalho* 40% e a quantidade de diferentes tipos de operações *Numero de Operações* 31%, enquanto que dois tipos de operação a ser realizada contribui com 29% com a previsão do modelo. Uma explicação geral da dificuldade dos itens dessa prova é que quanto maior a quantidade de informações que um sujeito deve manter disponível, e a quantidade de operações que devem ser realizadas para encontrar a resposta adequada, maior a dificuldade. Porém, dependendo do tipo de transformações que são aplicadas, aos itens, nesse caso, *Identidade* e *Multiplicação* a quantidade de informação que necessita ser mantida ativa durante a realização da tarefa diminui sensivelmente, tornando esses itens mais fáceis, ou seja, essas transformações estão relacionadas com a diminuição da dificuldade dos itens.

A Figura 27 apresenta a representação gráfica da escala dos itens da prova RN, do lado esquerdo estão posicionados os sujeitos conforme as habilidades na faixa de Theta, do lado direito estão posicionados os itens pela dificuldade na faixa de b . A letra M que aparece na barra central indica a média da medida, a letra S e a letra T indicam

respectivamente a posição do primeiro e do segundo desvio padronizado em relação a média. A escala possui uma amplitude de nove pontos em valores de Theta, e a média de habilidade dos sujeitos é pouco superior a média dificuldade dos itens, sugerindo que a dificuldade da prova é compatível com a habilidade da amostra.

Faixa de b de -5 à -4: Nesta faixa de dificuldade são encontrados os itens que passam por apenas uma transformação em apenas uma série da seqüência. Essa transformação varia comumente entre os tipos de *Identidade*, *Adição* ou *subtração*. Sujeitos com essa faixa de habilidade na prova RN apresentam capacidade de resolver problemas que envolvam operações matemáticas simples de adição ou subtração de uma ou mais unidades em uma série, sendo que a segunda série intercalada na seqüência geralmente apresenta o fator de complexidade identidade.

Faixa de b de -4 à -3: Nesta faixa de dificuldade são encontrados os itens que passam por apenas uma transformação em cada série da seqüência. As transformações são comumente dos tipos de *Identidade*, *Adição* ou *subtração*. Sujeitos com essa faixa de habilidade apresentam capacidade de resolver problemas que envolvam operações matemáticas simples de adição ou subtração de uma ou mais unidades nas duas séries intercaladas na seqüência.

Faixa de b de -3 à -2: Nesta faixa de dificuldade são encontrados os itens que passam por apenas uma transformação em cada série da seqüência. As transformações são comumente dos tipos de *Identidade*, *Adição* e *subtração*, geralmente nesse caso a regra aplicada à seqüência representa uma mudança na tendência apresentada nas seqüências anteriores. Sujeitos com essa faixa de habilidade apresentam capacidade de resolver problemas que envolvam operações matemáticas simples de adição ou subtração de uma ou mais unidades em casas alternadas na seqüência, além de perceberem mudanças na tendência geral das operações aplicadas até então.

Faixa de b de -2 à -1: Nesta faixa de dificuldade são encontrados os itens que passam por apenas uma transformação em cada série da seqüência. Além das transformações de *Identidade*, *Adição* e *subtração* são acrescentadas *Multiplificação* e *Divisão*. Sujeitos com essa faixa de habilidade apresentam capacidade de resolver problemas que envolvam operações matemáticas simples de adição ou subtração, multiplicação ou divisão por pequenas unidades em séries alternadas na seqüência.

Faixa de b de -1 à 0: Nesta faixa de dificuldade são encontrados os itens que passam por transformações envolvendo mais que uma operação, ou uma operação envolvendo valores diferentes, aplicadas a cada série da seqüência. As transformações podem ser de *Identidade*, *Adição*, *subtração*, *Multiplificação* e *Divisão*. Sujeitos com essa faixa de habilidade apresentam capacidade de resolver problemas em uma seqüência que envolvam, em uma série operações matemáticas simples de adição ou subtração de unidades que aumentam progressivamente, alternada por outra série que usa operações de multiplicação ou divisão por pequenas unidades.

Faixa de b de 0 à 1: Nesta faixa de dificuldade são encontrados os itens que passam por transformações envolvendo duas operações diferentes ou com valores diferentes em cada série da seqüência. Transformações de *Identidade*, *Adição*, *subtração*, *Multiplificação* e *Divisão*. Sujeitos com essa faixa de habilidade apresentam capacidade de resolver problemas em uma seqüência que envolvem, em uma série, operações matemáticas de adição ou subtração de unidades que aumentam progressivamente alternada por outra série que usa operações de multiplicação ou divisão por pequenas unidades.

Faixa de b de 1 à 2: Nesta faixa de dificuldade são encontrados os itens que passam por transformações envolvendo duas ou mais operações diferentes ou com valores diferentes em cada série da seqüência. As transformações envolvem, além dos tipos citados, *Complexidade*, *Ambiguidade*, *subtração*, *Multiplificação* e *Divisão*. Sujeitos com essa faixa

de habilidade apresentam capacidade de resolver problemas em uma seqüência que envolvem, em uma série, operações matemáticas simples de adição ou subtração de unidades que aumentam ou diminuem progressivamente, alternada por outra série que usa operações de multiplicação ou divisão por pequenas unidades, ou que apresentam uma segunda série que progride de duas em duas casas.

Faixa de *b* de 2 à 3: Nesta faixa de dificuldade são encontrados os itens que passam por transformações envolvendo operações diferentes organizadas hierarquicamente ou a mesma operação com valores diferentes em cada série da seqüência. Transformações de *Identidade, Adição, subtração, Multiplicação e Divisão*. Sujeitos com essa faixa de habilidade apresentam capacidade de resolver problemas em uma seqüência que envolvam, em uma série, operações matemáticas de adição ou subtração de unidades que aumentam progressivamente alternada por outra série que usa operações de multiplicação ou divisão.

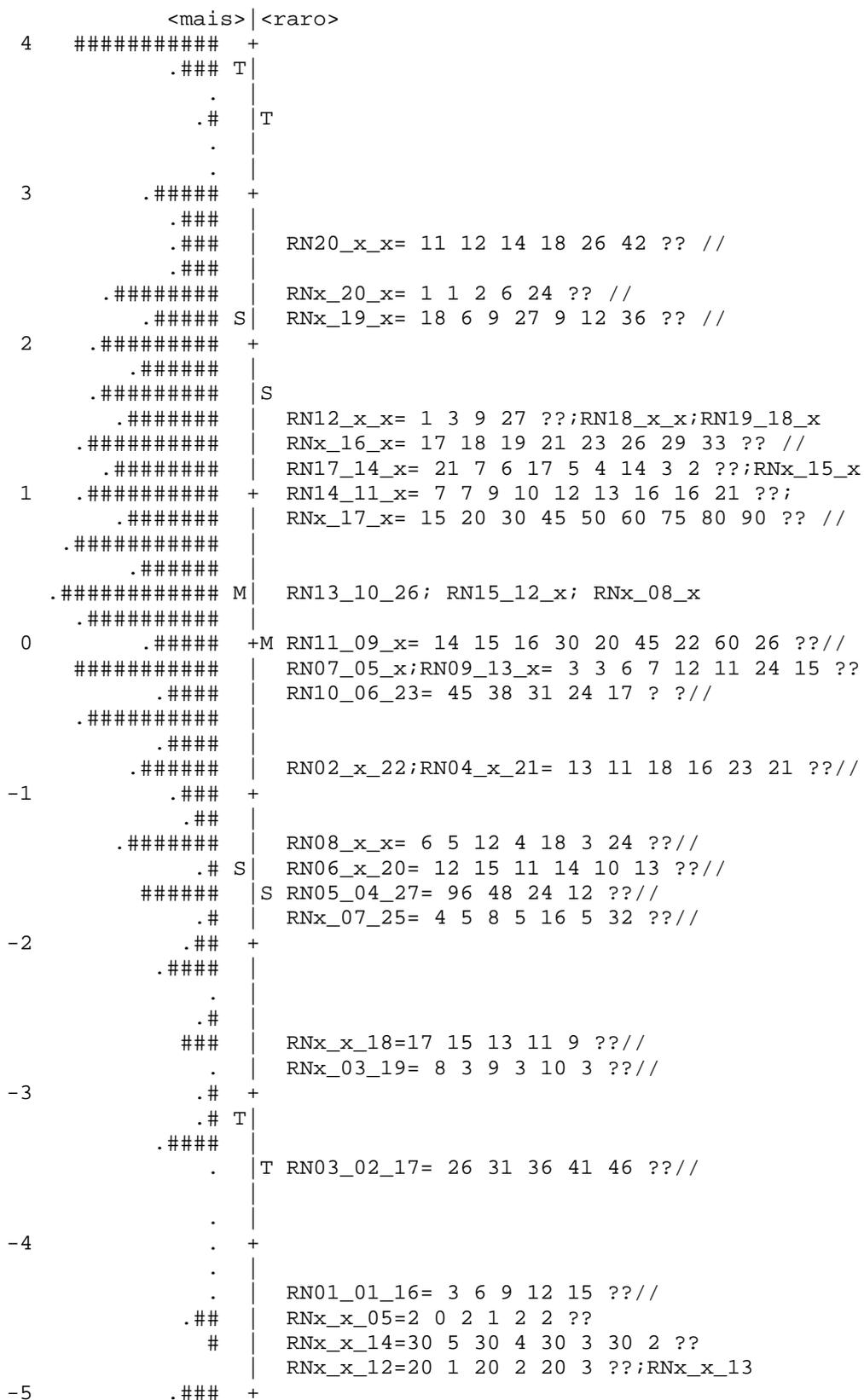


Figura 27. Gráfico de Escalonamento dos Itens da prova RA pela dificuldade.

Considerações Finais

As principais questões abordadas nesse trabalho sobre a construção de medidas dos fenômenos psicológicos, dizem respeito à ciência psicológica em geral, pois tratam da formulação e demonstração empírica dos seus construtos. Desde os primeiros anos de desenvolvimento da psicologia enquanto uma disciplina científica, questões sobre o que é a inteligência, quais são, e como estão organizados seus elementos foram tema de debate e, principalmente de motivação para desenvolvimento de novos estudos. Os elementos que sustentam esses apontamentos podem ser encontrados também nos trabalhos de Spearman (1927) e Thurstone (1934).

Thurstone (1937) apontou que o aspecto mais importante do debate sobre os elementos da mente diz respeito aos métodos utilizados na sua investigação, especialmente seu alcance e limitações. A psicometria se dedicou a construção das medidas desses fenômenos, e no curso do seu desenvolvimento contribuiu significativamente não apenas no campo das medidas, mas também no desenvolvimento das teorias sobre os construtos psicológicos que se propôs a medir. A operacionalização de um construto descrito teoricamente em termos dos comportamentos mais relevantes para designá-lo, produz um embate entre a descrição teórica e o dado empírico produzido. Muitas vezes os resultados obtidos pelo exame empírico desafiam as interpretações postuladas pela teoria sobre a dimensão psicológica em questão, apresentando como consequência a necessidade de uma revisão da teoria proposta. Esta relação entre o desenvolvimento da medida do construto psicológico, e o incremento da descrição teórica sobre esse construto está ligada ao conceito psicométrico de validade de construto, proposto por Cronbach e Meehl's (1955).

Os procedimentos adotados pela psicometria, durante sua primeira fase de desenvolvimento, foram, de sobremaneira, orientados para o que Embretson (1994) chamou

de rede nomotética, pois eram capazes de produzir informações sobre eventos externos como os quais essa medida se relacionava. O procedimento mais usado para demonstrar a validade do teste era relacionando seus resultados a medidas de critério externo, dessa maneira a validade de construto, que designa o significado do teste, era comprovada pela significância da relação com outras medidas. O desenvolvimento da psicometria para os modelos de resposta ao item, lançou as bases metodológicas que permitiram uma aproximação com desenvolvimentos da psicologia cognitiva, permitindo uma outra possibilidade de demonstrar o significado do teste, por meio da relação entre os escores e processos psicológicos que estão implicados na resposta aos itens.

O presente trabalho teve como objetivo a construção de escalas com interpretações referenciadas nos itens, portanto apresenta uma contribuição ao desenvolvimento da metodologia empregada na investigação dos processos de resposta sugeridos como fonte de evidência de validade para os instrumentos. O primeiro passo consistiu na análise do conteúdo dos itens de cada prova visando sua decomposição em partes menores que representam unidades de processamento necessárias a realização da tarefa proposta pelo item. Essa análise se baseou, no caso da prova RA, no conjunto de categorias proposto nos trabalhos de Jacobs e Vandeventer (1972) e Ward e Fitzpatrick (1973); e para a prova RN, nas regras propostas por Simon e Kotovsky (1963) e Simon (1979). O resultado dessa análise pôde ser considerado satisfatório, pois apresentou um conjunto de categorias capazes de decompor todos os itens de cada prova.

O segundo passo foi designar as categorias utilizadas para decompor os itens em fatores de complexidade, ou seja, investigar como essas categorias se relacionam umas com as outras, e qual a sua capacidade para explicar a dificuldade dos itens. Para a prova RA, os resultados encontrados vão ao encontro dos resultados encontrados por outros autores, por exemplo, nos resultados do trabalho de Hornke e Habon (1986) a estimação do índice de

contribuição de cada componente para a dificuldade do item, apontou que os itens envolvendo Separação foram os mais fáceis seguidos dos que envolviam Integração, sendo mais difíceis os itens que envolviam elementos embutidos. As operações de identidade foram as mais fáceis e as de interseção as mais difíceis, porém muitos fatores de complexidade apresentaram sobreposição, ou seja, contribuíram de forma semelhante para a dificuldade. Também no trabalho de Green e Kluever (1991) os resultados apontaram que as variáveis que melhor contribuía com a dificuldade dos itens foram respectivamente, número de dimensões, reflexão, número de opções. No caso do presente estudo, de modo semelhante, os resultados apontaram que itens que envolvem componentes embutidos, maior número de elementos e tipos de transformações são mais difíceis, os passos que itens envolvendo série e sombreamento são mais fáceis. Desta maneira, pôde-se considerar que os resultados encontrados foram congruentes com as explicações sobre os modelos de processamento de informação de analogias descritos na literatura, especialmente as exigências sobre a memória de trabalho presente nos fatores de complexidade de número de elementos das figuras, número de tipos de transformação e o número de dimensões do problema. Porém, as diferenças de contribuição para a dificuldade produzidas por diferentes transformações como tamanho e reflexão, série e sombreamento, componentes embutidos e separados, precisam ser melhor entendidas.

Considerando as análises da prova RN, da mesma forma, os resultados encontrados se alinham com aqueles descritos na literatura, por exemplo, no trabalho de Holzman, Pellegrino e Glaser (1982, 1983) os resultados das análises de regressão realizadas apontaram que a variável Memória de Trabalho é a que mais contribui com a previsão do modelo em todos os grupos, seguida por Número de Operações e Ambigüidade de Operações com contribuições significativas. Quando a variável Memória de Trabalho é retirada da análise os valores de R^2 apresentam queda sensível, indicando que a variância

associada a memória de trabalho é maior que a associada as outras variáveis presentes na análise. Entre as quatro subcategorias de Categoria de Operações Aritméticas, as variáveis, magnitude, divisão e multiplicação apresentaram coeficientes significativos, sendo que Adição e Subtração são mais fáceis que as demais. No caso do presente estudo os resultados vão à mesma direção, com as variáveis Memória de Trabalho e Número de Operações sendo as que mais contribuem com a dificuldade do item, e a identidade e multiplicação contribuindo para a facilitação do item. Dessa maneira, pôde-se concluir que estas variáveis são de grande importância para a determinação da dificuldade dos itens e influenciam o desempenho dos sujeitos, embora apresentem variação nos níveis relativos de dificuldade em cada estudo. Podemos considerar que a compreensão das características de cada fator é de especial interesse para ampliar a compreensão do desenvolvimento intelectual e das diferenças individuais sobre a atividade de raciocínio envolvendo padrões seriais.

O resultado final foi a apresentação de uma escala descritiva dos tipos de operações que estão presentes em cada item em uma faixa de dificuldade. Suposições plausíveis podem ser feitas relacionando tais operações às atividades cognitivas por trás das respostas corretas em cada item. Portanto, conseqüentemente, suposições sobre as habilidades que cada sujeito, que acerta itens em uma faixa de dificuldade específica, foi capaz de manifestar.

O aperfeiçoamento das descrições de cada faixa de habilidade, na escala, pode ser conseguido a partir da extrapolação dos elementos de desajustamento observados nesse estudo. Também, é possível acrescentar informações de fontes clássicas, como as interpretações referenciadas na norma, e em critérios externos relevantes, de modo a aumentar o poder explicativo das escalas tornando-as mais robustas. Esperamos que os resultados desse trabalho colaborem com uma melhor compreensão dos métodos aplicados

e, por consequência, estimulem a construção de instrumentos baseados nesses princípios, ao mesmo tempo em que desencorajem a construção artesanal e intuitiva de testes psicológicos.

Referências

- Almeida, L. S. (1988). *Teorias da Inteligência*. Porto, Edições Jornal de Psicologia.
- Almeida, L. S. & Primi, R. (1998). *Bateria de Provas de Raciocínio (BPR-5): Manual Técnico*. São Paulo: Casa do Psicólogo.
- American Educational Research Association, American Psychological Association & National Council on Measurement in Education (1999). *Standards of Educational and Psychological Testing*. Washington, DC : American Educational Research Association.
- Anastasi, A. & Urbina, S. (2000). *Testagem Psicológica*. Porto Alegre: Artmed.
- Arendasy, M. & Sommer, M. (2005). The Effect of Different Types of Perceptual Manipulations on the Dimensionality of Automatically Generated Figural Matrices. *Intelligence*, 33, 307-324.
- Arias, R. M. (1996). *Psicometria: teoria de los tests psicológicos y educativos*. Madrid: Editorial Sintesis.
- Baddeley, D. A. (1976) *The Psychology of Memory*. New York: Basic Books.
- Baker, F. B. (2001). *The Basics of Item Response Theory*. USA: ERIC Clearinghouse on Assessment and Evaluation.
- Barbosa, M. A. (2006). *Evidências de Validade da Bateria BPR-5 para Idosos*. Dissertação de Mestrado não publicada. Programa de pós-graduação *stricto-sensu* em Psicologia, Universidade São Francisco, Itatiba, São Paulo.
- Baumgartl, V. O. (2004). *Evidencia de Validade da BPR-5, BFM-1 e PMK em Eletricitários*. Dissertação de Mestrado não publicada. Programa de pós-graduação *stricto-sensu* em Psicologia, Universidade São Francisco, Itatiba, São Paulo.
- Baumgartl, V. O.; Nascimento, E. (2004). A Bateria de Provas de Raciocínio (BPR-5) aplicada ao contexto organizacional. *Psico-USF*, 9(1), 1-10.

Blanton, H.; Jaccard, J. (2006). Arbitrary Metrics in Psychology. *American Psychological Association*, 1(61), 27-41

Bowman, D. B.; Markham, P. M.; Roberts, R. D. (2002). Expanding the frontier of human cognitive abilities: so much more than (plain) g! *Learning and Individual Differences*, 13, 127-158.

Campos, H. R. (2005). *Análise de conteúdo e sua relação com a dificuldade dos itens da bateria de provas de raciocínio (BPR-5)*. Dissertação de Mestrado não publicada. Programa de pós-graduação *stricto-sensu* em Psicologia, Universidade São Francisco, Itatiba, São Paulo.

Carrol, J. B. (1993a). *Human Cognitive Abilities: A Survey of factor-Analytic Studies*. New York: Cambridge University Press.

Carrol, J. B. (1993b). Test Theory and the Behavioral Scaling of Test Performance. Em: N. Frederiksen, R. J. Mislevy, I. I. Bejar, (Eds). *Test Theory of a New Generation of Tests*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Carrol, J. B. (1997). The Three-Stratum Theory of Cognitive Abilities. Em: D. P. Flanagan, J. L. Gensheft, P. L. Harrison, (Eds). *Contemporary Intellectual Assessment: theories, tests and issues*. New York, London: The Guilford Press.

Costa, O. R. S. (2005). *Um estudo correlacional das pirâmides de Pfister e a BPR-5*. Dissertação de Mestrado não publicada. Programa de pós-graduação *stricto-sensu* em Psicologia, Universidade São Francisco, Itatiba, São Paulo.

Couto, G. (2005). *Análise Psicométrica da Bateria de Provas de Raciocínio. Utilizando a Teoria de Resposta ao Item*. Dissertação de Mestrado não publicada. Programa de pós-graduação *stricto-sensu* em Psicologia, Universidade São Francisco, Itatiba, São Paulo.

Conway, A. R. A.; Cowan, N.; Bunting, M. F.; Theriault, D. J. & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, 30(2),163-183.

Cronbach, L. J. (1996). *Fundamentos da Testagem Psicológica*. Porto Alegre: Artes Médicas.

Cruz, M. B. Z. (em preparação). *Estudo de Validade e Precisão da Bateria de Provas de Raciocínio Infantil – BPR-5i*. Dissertação de Mestrado. Programa de pós-graduação *stricto-sensu* em Psicologia, Universidade São Francisco, Itatiba, São Paulo.

Deary, I. J. (2000). *Looking Down on Human Intelligence: From Psychometrics to the Brain*. Oxford University Press.

Deary, I. J. (2001). Human Intelligence Differences: Towards a Combined Experimental – Differential Approach. *Trends in Cognitive Sciences*, 5(4), 164 – 170.

Del Nero, H. S. (1997). *O Sítio da Mente: Pensamento, Emoção e Vontade no Cérebro Humano*. São Paulo: Collegium Cognitio.

Danthiir, V.; Pallier, G.; Roberts, R. D.; Stankov, L. (2001). What the Nose Knows: Olfaction Within the Structure of Human Cognitive Abilities. *Intelligence*, 30, 337–361.

Duncan, J.; Burgess, P.; Emslie, H. (1995). Fluid Intelligence After Frontal Lobe Lesions. *Neuropsychologia*,33(3), 261-268.

Duncan, J.; Seitz, R. J.; Kolodny, J.; Bor, D.; Herzog, H.; Ahmed, A.; Newell, F. N.; Emslie, H. (2000). A Neural Basis for General Intelligence. *Science*, 289, 457-460.

Embretson, S. E. (1983). Construct Validity: Construct Representation Versus Nomothetic Span. *Psychological Bulletin*, 93(1), 179-197.

Embretson, S. E. (1985). Studying Intelligence Whith Test Theory Models. Em: D. K. Detterman (Ed.). *Current topics in human intelligence*. Norwood, NJ: Ablex.

Embretson, S. E. (1994). Applications of Cognitive Design Systems to Test Development. Em: C. R. Reynolds, (org.). *Cognitive assessment: a multidisciplinary perspective*. New York: Plenum Press.

Embretson, S. E. (1999). IRT Applications in Cognitive and Developmental Assessment. Em: S. E. Embretson, S. L. Hershberger (Edts.). *The New Rules of Measurement: what every psychologist and educator should know*. Mahwah, New Jersey, London: Lawrence Erlbaum Associates.

Embretson, S. E. (2006). The continued Search for Nonarbitrary Metrics in Psychology. *American Psychological Association*, 1(61), 50-55.

Engle, R. W.; Tuholski, S. W.; Laughlin, J. E.; Conway, A. R. A. (1999). Working Memory, Short-Term Memory, and General Fluid Intelligence: A Latent-Variable Approach. *Journal of Experimental Psychology General*, 128(3), 309-331.

Filizatti, R. (2004). *Estudo de validação dos testes 16PF e BPR5 no contexto organizacional*. Dissertação de Mestrado não publicada. Programa de pós-graduação *stricto-sensu* em Psicologia, Universidade São Francisco, Itatiba, São Paulo.

Fischer, G. H. (1973). The Linear Logistic Test Model as an Instrument in Educational Research. *Acta Psychologica*, 37, 359-374.

Fry, A. F., Hale, S. (2002). Relationships among processing speed, working memory, and fluid intelligence in children. *Biol Psychol.* 54(1-3),1-34.

Geake, J. G. & Hansen, P. C. (2005). Neural Correlates of Intelligence as Revealed by fMRI of Fluid Analogies. *NeuroImage*, 26, 555-564.

Green, K. E. & Kluever, R. C. (1991). Components of Item Difficulty of Raven's Matrices. *The Journal of General Psychology*, 119(2), 189-199.

Hambleton, R. K. (1990). Item Response Theory: Introduction and Bibliography. *Psicothema*, 2(1), 97-107.

Heitz, R. P.; Unsworth, N.; Engle, R. W. (2005). Working Memory Capacity, Attention Control, and Fluid Intelligence. Em: O. Wilhelm, R. W. Engle, (Edts). *Handbook of Understanding and Measuring Intelligence*. Thousand Oaks, London, New Delhi: Sage Publications, Inc.

Holzman, T. G.; Pellegrino, J. W.; Glaser, R. (1982). Cognitive Dimensions of Numerical Rule Induction. *Journal of Educational Psychology*, 3(74), 360-373.

Holzman, T. G.; Pellegrino, J. W.; Glaser, R. (1983). Cognitive Variables in Series Completion. *Journal of Educational Psychology*, 4(75), 603-618.

Hornke, L. F.; Habon, M. W. (1986). Rule-Based Item Bank Construction and Evaluation within the Linear Logistic Framework. *Applied Psychological Measurement*, 10(4), 369-380.

Hornke, L. F. (2001). Item-Generation Models for Higher Order Cognitive Function. Em: (...)

Imbert, M. (1998). Neurociências e Ciências Cognitivas. Em: D. Andler (Org). *Introdução às Ciências Cognitivas*. Rio Grande do Sul, São Leopoldo: Ed. Unisinos.

Jacobs, P. J.; Vandeventer, M. (1972). Evaluating the Teaching of Intelligence. *Educational and Psychological Measurement*, 32, 235-248.

Jung-Beeman, M.; Bowden, E. M.; Haberman, J.; Flymiare, J. L.; Arambel-Liu, S.; Greenblatt, R.; Reber, P. J.; Kounios, J. (2004). Neural Activity When People Solve Verbal Problems with Insight. *Plos Biology*, 2(4), 500-510.

Kane, M. J.; Engle, R. W. (2000). Working Memory Capacity, proactive interference, and divided attention: limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 26, 333-358.

Kane, M. J. (2003). The Intelligent Brain in Conflict. *Trends in Cognitive Science*, 7(9), 375-377.

Kane, M. J. (2005). Full Frontal Fluidity? Looking in on the neuroimaging of reasoning and intelligence. In: O. Wilhelm, R. W. Engle (Edts). *Handbook of Understanding and Measuring Intelligence*. Thousand Oaks, London, New Delhi: Sage Publications.

Kyllonen, P. C. & Christal, R. (1990). Reasoning ability is (little more than) working memory capacity?!. *Intelligence*, 14, 89-433.

Kyllonen, P. C. (1996) Is working memory capacity Spearman's? Em: I. Dennis & P. Tapsfield (Orgs.), *Human abilities: their nature and measurement* (pp. 49-75). Mahwah: Erlbaum.

Lezak, M. D. (1995). *Neuropsychological Assessment*. New York: Oxford University Press.

Lee, K. H.; Choi, Y. Y.; Gray, J. R.; Cho, S. H.; Chae, J.; Lee, S.; Kim, K. (In press) Neural Correlates of Superior Intelligence: Stronger Recruitment of Posterior Parietal Cortex. *NeuroImage*.

Linacre, J. M. & Wright, B. D. (1994a). Chi-square fit statistics. *Rasch Measurement Transactions*, 8(2), 350. Disponível: <http://www.rasch.org/rmt/rmt82a.htm>. Acessado: 31/01/2001.

Linacre, J. M. & Wright, B. D. (1994b). Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8(2), 370. Disponível: <http://www.rasch.org/rmt/rmt83.htm>. Accessed: Acessado: 31/01/2001.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent variable analysis. *Journal of Experimental Psychology: General*, 130, 621-640.

McGrew, K. S. (1997). Analysis of the Major Intelligence Batteries According to a Proposed Comprehensive Gf-Gc Framework. Em: D. P. Flanagan, J. L. Gensheft, P. L. Harrison (Eds). *Contemporary Intellectual Assessment Theories, Tests, and Issues*. New York, London: The Guilford Press.

McGrew, K. S. & Flanagan, D. P. (1998). *The Intelligence Test Desk Reference (ITDR): Gf-Gc Cross-battery Assessment*. Needham Heights: Allyn & Bacon.

Mulholland, T. M.; Pellegrino, J. W.; Glaser, R. (1980). Components of Geometric Analogy Solution. *Cognitive Psychology*, 12, 252-284.

Muñiz, J. (1990). *Teoría de Respuesta a los Ítems: un nuevo enfoque en la evolución psicológica y educativa*. Madri: Ediciones Pirámide, S. A.

Muñiz, J. (1994). *Teoría Clássica de los Tests*. Madrid: Ediciones Pirâmides, S. A.

Muñiz, J. (org) (1996). *Psicometria*. Madrid: Editorial Universitas, S. A.

Pasquali, L. (1996). *Teoria e Métodos de Medida em Ciência do Comportamento*. Brasília: INEP.

Pasquali, L. (1997). *Psicometria: teoria e Aplicações – A Teoria Clássica dos Testes Psicológicos*. Brasília: Editora UnB.

Pasquali, L. (2000). *Psicometria: Teoria dos Testes Psicológicos*. Brasília: LabPAM.

Pasquali, L. (org) (2001). *Técnica de Exame Psicológico – TEP*. São Paulo: Casa do Psicólogo.

Primi, R. (1996). *Construção de um instrumento para avaliação do raciocínio indutivo: aplicação da psicologia cognitiva e da teoria de resposta ao item*. Projeto de qualificação não publicado. Instituto de Psicologia, Programa de pós-graduação em Psicologia Escolar e do Desenvolvimento Humano, Universidade de São Paulo, São Paulo.

Primi, R. (1998). *Desenvolvimento de um instrumento informatizado para avaliação do raciocínio analítico*. Tese de Doutorado não publicada. Instituto de Psicologia, Programa de pós-graduação em Psicologia Escolar e do Desenvolvimento Humano, Universidade de São Paulo, São Paulo.

Primi, R. & Almeida, L. S. (2001). Teoria de Resposta ao Item. Em: E. M. Fernandes & L. S. Almeida (Org.). *Métodos e Técnicas de Avaliação: Contributo para a Prática e Investigação Psicológica*. (pp. 205-232). Braga, Portugal: Centro de estudos em Educação e Psicologia, Universidade do Minho.

Primi, R. (2002). Avanços na concepção psicométrica da inteligência. Em: F. C. Capovilla, (org.) *Neuropsicologia e Aprendizagem: uma abordagem multidisciplinar*. Sociedade Brasileira de Neuropsicologia - SBNp

Primi, R.; Guntert, A. E. A.; Alchieri, J. C. (2002). Estudo Correlacional Entre a BPR-5, R1, TNRV, com o Teste de Zulliger. *Psicologia: Reflexão e Crítica*, 1(33), 53-76.

Primi, R.; Bighetti, A. A.; Munhoz, A. H.; Noronha, A. P. P.; Polidoro, S. A. J.; Nucci, E. P. D.; Pellegrini, M. C. K. (2002). Personalidade, Interesses e Habilidades: Um Estudo Correlacional da BPR-5, LIP, 16PF. *Avaliação Psicológica*, 1, 61 – 72.

Primi, R. (2004). Avanços na Interpretação de Escalas com a Aplicação da Teoria de Resposta ao Item. *Avaliação Psicológica*, 1(3), 53-58.

Santos, A. A. A.; Primi, R.; Taxa, F.; Vendramini, C. M. M. (2002). O Teste de Cloze na Avaliação da Compreensão em Leitura. *Psicologia: Reflexão e Crítica*, 15(3), 594-560.

Simon, H. A.; Kotovsky, K. (1963). Human Acquisition of Concepts for Sequential Patterns. *Psychological Review*, 70(6), 534-546.

Simon, H. A. (1972). Complexity and the Representation of Patterned Sequences of Symbols. *Psychological Review*, 79(5), 369-382.

- Sternberg, R. J. (1983). Components of Human Intelligence. *Cognition*, *15*, 1-48
- Sternberg, R. J., & Powell, J. S. (1983). The development of intelligence. In P. H. Mussen (Org.), *Handbook of child psychology* (pp. 341-419). Nova York: John Wiley & Sons.
- Spearman, C. (1927). *Las Habilidades del Hombre Su Naturaleza y Medición*. Buenos Aires: Editorial Paidós.
- Stankov, L. (2000). Structural Extensions of a Hierarchical View on Human Cognitives Abilities. *Learning and Individuals Differences*, *12*, 35-51.
- Thurstone, L. L. (1934). The Vectors of Mind. *Psychological Review*, *41*, 1-32.
- Thurstone, L. L. (1937). Psychology as a quantitative rational science. *Science*, *85*, 227-232.
- Vigil-Colet, A.; Perez-Olle, J.; Fernandez, M. (1997). The Relationships of Basic Information processing Measures with Fluid and Crystallized Intelligence. *Personality and Individual Difference*, *23*(1), 55-65.
- Ward, J.; Fitzpatrick, T. F. (1973). Characteristics of Matrices Items. *Perceptual and Motor Skills*, *36*, 987-993.
- Woodcock, R. W. (1999). What can Rasch-Based Scores Convey about a Person's Test Performance? Em: S. E. Embretson, & S. L. Hershberger, (Eds). *The New Rules of Measurement: what every psychologist and educator should know*. Mahwah, N. J: Lawrence Erlbaum Associates.
- Wright, B. D.; Stone, M. H. (2004). *Making Measure*. Chicago: The Phanon Press.
- Zelazo, P. D.; Carter, A.; Reznick, J. S. (1997). Early Development of Executive Function: A Problem-Solving Framework. *Review of General Psychology*, *1*(2), 198-226.
- Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are

visuospatial working memory, executive functioning, and spatial abilities related? A latent variable analysis. *Journal of Experimental Psychology: General*, 130, 621-640.

Anexo I

Tabela 1. Categorias de Classificação dos itens em Fatores de Complexidade prova RA.

Ordem	N Ele	N Transf.	Tip Tranf	Id	Forma	Sombrea	Tamanho	Mov. plano	Séries	Ad Ele	Ad comb	Ele conj	Ad Única	Refle	Inver	Sep Comp	Inte Comp	Comp Embu
1	1	1	1	0	0	1	0	0	?	0	0	0	0	0	0	0	0	0
2	1	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
3	2	2	2	0	?1	0	0	0	0	1	0	0	0	0	0	0	0	0
4	2	1	1	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0
5	1	2	2	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
6	3	2	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
7	2	2	2	0	0	1	1	0	0	0	0	0	0	0	0	0	1?	0
8	2	2	2	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0
9	2	3	2	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0
10	2	2	2	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0
11	1	2	2	0	1	1	0	0	0	0	0	0	0	0	0	0	1?	0
12	1	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
13	1	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
14	2	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
15	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
16	2	2	2	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0
17	3	3	3	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0
18	2	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
19	2	3	2	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0
20	2	2	2	0	0	1	0	0	0	0	0	0	0	0	1	0	0	1?
21	3	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
22	2	3	3	0	0	1	0	1	0	1	0	0	0	0	0	0	0	1?
23	2	2	1	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0
24	5	2	2	0	0	0	0	1	0	1	0	0	0	0	0	1	0	0
25	4	5	4	0	1	0	1	0	0	1	0	0	0	0	0	1	0	0
26	2	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0

Tabela 1. Categorias de Classificação dos itens em Fatores de Complexidade prova RA (Continuação).

Ordem	N Ele	N Transf.	Tip Tranf	Id	Forma	Sombrea	Tamanho	Mov. plano	Séries	Ad Ele	Ad comb	Ele conj	Ad Única	Refle	Inver	Sep Comp	Inte Comp	Comp Embu
27	2	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
28	2	4	3	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0
29	2	2	2	0	1	0	0	0	0	1	0	0	0	0	0	1	0	0
30	5	7	5	0	0	1	1	1	0	0	0	0	0	1	0	1	0	0
31	2	2	2	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0
32	1	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
33	1	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
34	1	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
35	1	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
36	2	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
37	2	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
38	2	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
39	1	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
40	1	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
41	2	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
42	1	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
43	1	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
44	1	5	3	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0
45	1	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
46	2	4	4	0	1	0	0	1	1	1	0	0	0	0	0	0	0	0
47	2	4	3	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0
48	3	2	2	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0
49	1	2	2	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0
50	3	4	3	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0
51	2	2	2	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0

Anexo II

Tabela 1. Categorias de Classificação dos itens em Fatores de Complexidade da prova RN.

Ordem	N Eleme.	N Oper.	Tip Oper.	Ambigüidade Oper.	Iden.	Adição	Subtração	Multiplicação	Div.	Memó. de Trab	Padrão de Descr.	Num. Peri	Comp. Peri	Complex.	Comp. Série	Conflito de Direção
1	1	1	0	0	1	0	0	0	1	1	1	1	0	7	0	1
2	2	2	0	0	1	1	0	0	2	2	2	1	0	8	0	2
3	1	1	0	0	1	0	0	0	1	1	1	1	0	7	0	1
4	3	2	1	0	1	1	0	0	2	3	2	1	0	8	0	3
5	1	1	0	0	0	0	0	1	1	1	1	1	0	6	0	1
6	2	2	1	0	1	1	0	0	2	3	2	1	0	8	0	2
7	2	2	0	0	1	1	0	0	2	2	2	1	0	9	0	2
8	2	2	0	0	1	1	0	0	2	2	2	1	0	9	0	2
9	2	2	0	0	1	0	1	0	2	2	2	1	0	10	0	2
10	1	1	0	0	0	1	0	0	1	1	1	1	0	7	0	1
11	3	2	0	0	1	0	1	0	3	3	2	1	1	11	0	3
12	1	1	0	0	0	0	1	0	1	1	1	1	0	6	0	1
13	4	1	0	0	1	0	0	0	3	1	2	1	1	10	0	4
14	6	1	0	0	1	0	0	0	3	2	2	1	1	11	0	6
15	3	2	0	1	0	1	0	0	3	3	2	1	1	11	0	3
16	2	2	0	0	1	1	0	0	2	2	2	1	0	10	0	2
17	4	1	0	0	0	1	0	0	2	3	2	1	1	11	0	4
18	4	1	0	0	1	0	0	0	2	1	1	1	1	10	0	4
19	3	2	0	0	1	1	0	0	2	1	1	1	1	11	0	3
20	7	2	0	0	1	0	1	0	2	2	1	1	1	8	0	7
21	2	2	0	1	1	0	0	0	1	1	2	1	0	8	0	2
22	2	2	0	1	0	0	1	0	1	1	2	1	0	9	0	2
23	4	1	1	0	1	0	0	0	2	3	2	1	0	10	0	4
24	9	2	0	0	1	1	0	0	2	2	2	1	1	11	0	9

Tabela 1. Categorias de Classificação dos itens em Fatores de Complexidade da prova RN (Continuação).

Ordem	N Eleme.	N Oper.	Tip Oper.	Ambigüidade Oper.	Iden.	Adição	Subtração	Multiplicação	Div.	Memó. de Trab.	Padrão de Descr.	Num. Peri	Comp. Peri	Complex.	Comp. Série	Conflito de Direção	N Eleme.	N Oper.
25	8	1	1	0	1	0	0	0	2	3	2	1	1	10	0	8	1	1
26	6	1	1	0	1	0	0	0	2	3	2	1	1	11	0	6	1	1
27	6	2	0	1	1	0	0	0	3	3	3	1	1	9	0	6	2	0
28	6	1	0	0	0	0	1	0	2	6	1	1	1	7	0	6	1	0
29	1	1	0	0	1	0	0	0	1	1	1	1	0	8	0	1	1	0
30	1	1	0	1	0	0	0	0	1	1	1	1	0	8	0	1	1	0
31	1	1	0	1	0	0	0	0	1	2	2	1	0	8	0	1	1	0
32	2	2	0	1	1	0	0	0	1	2	2	1	0	8	0	2	2	0
33	2	2	0	1	1	0	0	0	1	2	2	1	0	8	0	2	2	0
34	1	1	0	0	0	1	0	0	1	1	1	1	0	8	0	1	1	0
35	2	1	0	1	0	0	0	0	1	1	2	1	0	8	0	2	1	0
36	2	2	0	1	1	0	0	0	1	2	2	1	0	8	0	2	2	0
37	2	2	0	1	1	0	0	0	1	2	2	1	0	10	0	2	2	0
38	2	2	0	1	1	0	0	0	1	2	2	1	0	8	0	2	2	0
39	2	2	0	1	1	0	0	0	1	2	2	1	0	8	0	2	2	0
40	2	2	0	1	1	0	0	0	1	2	2	1	0	8	0	2	2	0
41	1	1	0	0	0	1	0	0	1	1	1	1	0	8	0	1	1	0
42	2	2	0	1	0	1	0	0	1	2	2	1	0	8	0	2	2	0
43	2	2	0	1	1	0	0	0	1	2	2	1	0	10	0	2	2	0
44	1	1	0	0	0	1	0	0	1	1	1	1	0	7	0	1	1	0
45	2	1	0	0	1	0	0	0	1	2	2	1	0	8	0	2	1	0
46	2	2	0	0	1	0	1	0	2	2	2	1	0	8	0	2	2	0
47	5	2	1	0	0	1	0	1	2	3	2	1	1	8	0	5	2	1
48	8	2	0	0	1	1	0	0	2	2	1	1	1	9	0	8	2	0