

# Os princípios de organização perceptiva e a atividade inteligente: um estudo sobre testes de inteligência

Ricardo Primi

USP - IP\*

Eliana M. S. Rosado

PUCAMP\*\*

Hipotetizou-se que o processo cognitivo de eduzir relações segue a tendência econômica do princípio mínimo definido pela Teoria da Gestalt. Buscou-se explicar a dificuldade dos itens de testes de inteligência, em função de um tipo particular de organização onde relações irrelevantes são eduzidas em função de um arranjo perceptivo específico. Nestas condições a carga de informação se eleva, aumentando a probabilidade de erros. Sendo esta hipótese verdadeira, conseguiu-se reverter este arranjos, diminuir-se-ia a dificuldade dos itens. Assim, foram escolhidos 9 itens de dificuldade elevada dos testes D-48 e D-70 e compostos dois cadernos: um contendo os itens originais e o outro contendo os mesmos itens arranjados segundo os princípios de organização perceptiva da Gestalt, de forma a não produzir arranjos perceptivos irrelevantes. Estes cadernos foram aplicados em dois grupos emparelhados quanto ao sexo e nível de inteligência (medidos através do G 36). Cada grupo contendo 14 sujeitos respondeu a um dos cadernos. Observou-se que os escores dos cadernos alterados aumentaram significativamente, principalmente em sujeitos com menos habilidade.

**Palavras-chave:** testes de inteligência, análise de itens, princípio mínimo, teoria da Gestalt, D-48, D-70, raciocínio indutivo, princípios de organização perceptiva.

## Abstract

It was ipotetized that the inductive reasoning is related with economic principle of Gestalt's theory defined as minimum principle. It is purpose of this work to explain the difficulty of intelligence test itens with function of particular type of organization where irrelevant relations are infered. In this condition the load of information elevates causing the increase of errors. If this hypothesis is true, if we could reverse this particular organizations the item difficulty will decrease. It was selected 9 itens of high level o difficulty of the D-48 and D70 and created two experimental tests, one having the original itens and the other having the same itens but arranged with Gestalt principles of organization with the purpose of avoid irrelevant perceptual organizations. This tasks was aplied in two groups equiparated in sex an intelligence (measured by G36 test). Each group formed with 14 subjects answered to the one of experimental tests. It was observed that the scores of the modified tests increase significantly, mainly for the subjects of low intelligence level.

**Key words:** intelligence tests, item analisys, minimum principle, Gestalt theory, D-48, D-70, inductive reasoning, perceptual organizations principle.

## Aspectos ligados à mensuração da inteligência

Um ponto importante que deve ser considerado na construção de testes em psicologia, notadamente no que tange a avaliações da inteligência, é o

\* Bolsista do CNPq pelo projeto de doutoramento em Psicologia Escolar.

\*\* Professora titular do curso de Pós-Graduação em Psicologia Escolar.

Endereço para correspondência: Depto de Pós-Graduação em Psicologia da PUCAMP, rua Waldemar César da Silveira, 105, CEP 13045-270, Campinas, SP.

nível de mensuração. De um modo geral, o cientista procura "traduzir" o fenômeno estudado, através da atribuição de números às observações, buscando construir um sistema de representação matemático simbólico da realidade que está estudando. Este sistema (ou modelo assim construído) possibilita ao pesquisador trabalhar com relações entre números, sabendo que seu sistema corresponde a alterações análogas na realidade de onde este modelo foi extraído. Contudo, o cientista deve ter uma compreensão clara deste processo de mensuração e da variável

que está tentando medir, para poder decidir qual sistema numérico (incluídas aí as operações matemáticas admissíveis) deve atribuir à sua variável, portanto qual nível de mensuração é possível atingir. Segundo Lord & Novick (1974, pag 17) a mensuração:

“...é um procedimento de atribuir números (escores, medidas) a propriedades específicas de unidades experimentais, de forma a caracterizar e preservar as relações no domínio comportamental ... Nós chamamos qualquer operação que preserve a distinção feita por uma atribuição numérica uma operação *admissível*”.

É preciso ficar claro que a *variável* inteligência (considerada aqui como um constructo geral, deixando-se de lado o problema da dimensionalidade deste constructo) em sua natureza é contínua e quantitativa (intervalar). Contudo, isto não significa que os *escores* obtidos através de observações feitas por instrumentos de medida sejam da mesma natureza. Por exemplo, a medida tradicional de QI é muitas vezes mal interpretada como se fosse uma variável intervalar, quando na verdade este nível é assumido somente para fins estatísticos. Como cita Ferguson (1966, pag 15):

“No trabalho psicológico muitas variáveis são de fato ordinais, embora para propósitos estatísticos elas sejam, justificadamente, comumente tratadas como se fossem variáveis intervalares. Por exemplo, escores em testes de QI, aptidão, testes de aptidão acadêmica, testes de personalidade e outros parecidos são de fato variáveis ordinais, embora estas sejam comumente tratadas com se fossem variáveis intervalares ou de razão”.

Como na inteligência não se tem clareza quanto à unidade de mensuração, não se pode medir em escalas intervalares. Em ciências exatas como a Física, a medida de distância tem como uma de suas unidades de medida o metro. Portanto, uma distância de 8m representa o dobro da distância de 4m. Isto é possível porque nesta escala, os intervalos (de um em um metro) representam distâncias iguais. Na Psicologia, em contrapartida, não se pode dizer de uma

pessoa que acertou 8 itens de um teste de inteligência contendo 10 itens, que ela tenha o dobro da inteligência de uma pessoa que tenha acertado 4 itens. Isto porque alguns itens podem exigir muito mais da capacidade cognitiva de um indivíduo do que outros, e desse modo o “intervalo medido” varia de item para item.

A decorrência dessa reflexão exige do pesquisador a prudência em assumir o nível ordinal para esta medida, já que neste nível de mensuração não há exigências da delimitação exata e homogênea do intervalo entre uma posição e outra da escala. Em suma, pode-se dizer que uma pessoa que acerta 8 itens tem “mais inteligência” do que uma pessoa que acerte 4 itens, mas não é possível dizer *o quanto* ela é mais inteligente. Em termos práticos, pode-se interpretar a medida de QI como uma indicação da posição de uma pessoa em relação às outras. O conhecimento do QI de um indivíduo, como por exemplo situado acima de 132 - em um teste que tenha desvio padrão de 15 - indica que esta pessoa ocupa uma posição dentre os 2.14% primeiros sujeitos mais inteligentes do grupo de referência (norma).

Embora a limitação quanto ao nível de mensuração seja definida, estatisticamente os dados são tratados como se estes representassem medidas intervalares. Sob este aspecto Lord & Novik (1974) sustentam que este procedimento, “incorreto” em sua natureza, se justifica quando a utilidade das escalas baseadas nele seja evidente (a utilidade vista como o poder de predição de algum critério relevante). Portanto, embora do ponto de vista teórico não seja prudente considerar uma escala intervalar quando na verdade ela é ordinal, do ponto de vista pragmático este procedimento pode ser adotado.

Como conclusão das discussões acima, pode-se dizer que uma questão fundamental para aprimoramento da medida da inteligência é o conhecimento da “quantidade” de inteligência que cada item está medindo. Em nossa opinião, uma maneira de se aproximar desta medida é o *estudo de aspectos qualitativos dos itens*, entendendo-se com clareza a quais processos mentais aqueles itens fazem exigência. Conhecendo quais são os processos cognitivos exigidos pelos diferentes itens analisados, pode-se aproximar de uma compreensão mais precisa sobre qual é o peso (ou seja, qual intervalo que

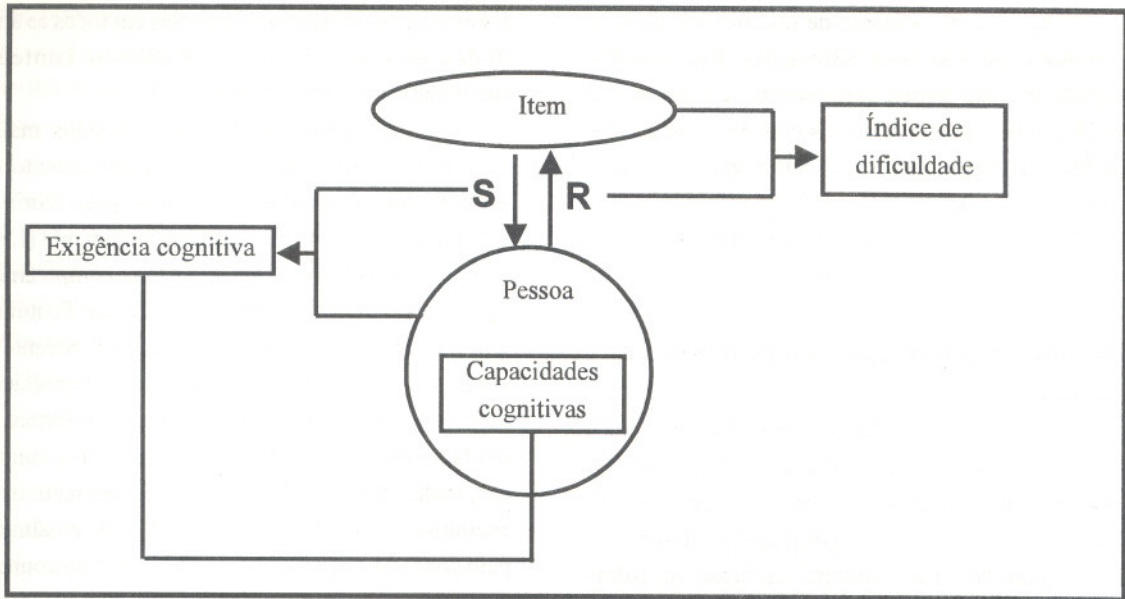


Figura 1

cada item cobre) na caracterização da inteligência. Neste aspecto, a relação com o processo de construção dos testes de inteligência é direta.

Um fase importante na construção de testes de inteligência é a fase de análise dos itens. É comum serem tomadas algumas propriedades estatísticas dos mesmos, dentre elas, o *índice de dificuldade*, que reflete a proporção de acertos no item analisado (Almeida, 1988, 1993). Este índice pode ser um indicativo do grau de exigência cognitiva de um item particular. No entanto, o resultado em si não qualifica quais são estas exigências. O presente trabalho procurou analisar experimentalmente a dificuldade de itens, buscando relacionar a exigência cognitiva com a dificuldade dos itens. Observou-se aqui a lógica descrita a seguir.

A dificuldade do item é determinada pela sua exigência em termos de processos cognitivos: quanto mais exigente em termos de processos superiores de raciocínio (capacidades cognitivas), mais difícil é o item. Esta afirmação pode ser representada na Figura 1, acima. Um item particular é um estímulo que faz exigências a determinadas capacidades cognitivas, de uma determinada pessoa. A resposta correta ao item ocorre se a pessoa possui determinadas capacidades que a possibilitem processar aquelas informações de maneira adequada.

Portanto, de um modo geral, o índice de dificuldade  $ID_i$  de um item particular  $i$ , aplicado a uma

amostra particular de  $N$  sujeitos, está condicionado pela sua exigência mínima em termos de processos cognitivos, portanto uma posição mínima  $\theta_i$  na escala de habilidade, para que sua resolução seja conseguida. Sendo assim, a dificuldade do item está condicionada à amostra particular estudada, ou seja, ele será diretamente proporcional ao número de sujeitos que tenham habilidade igual ou acima de  $\theta_i$ . Portanto o índice de dificuldade não é invariante quanto à amostra de onde ele foi extraído. Quanto a este aspecto, o leitor é remetido aos estudos da Teoria de Resposta ao Item que vêm se aprofundando nestes e noutros aspectos da análise dos itens (Baker, 1992; Muniz, 1990; Hamblenton & Colbs, 1978).

Entendemos que uma forma experimental de se estudar qualitativamente um item, ou seja, identificar quais as capacidades cognitivas de processamento que este item exige (em outras palavras, definir psicologicamente o traço  $\theta$  que este item mede) se torna possível se:

- (a) formularmos alguns conceitos sobre a natureza do processamento exigido pelo item (portanto, afirmações sobre sua exigência cognitiva)
- (b) manipularmos experimentalmente estas variáveis em sua construção, prevendo quais itens irão exigir mais (ou menos) destas capacidades.

Analisando o índice de dificuldade, pode-se corroborar ou não estas afirmações. Este trabalho baseou-se nesta lógica experimental e buscou manipular a dificuldade dos itens com base em formulações qualitativas sobre a natureza do processamento exigido. Buscou-se formular estas afirmações apoiando-se em alguns conceitos da teoria da Gestalt, como se verá a seguir.

### Organização perceptiva e processo de raciocínio

Estudos experimentais tem sido realizados com o propósito de compreender a natureza dos processos de raciocínio presentes em testes de QI. Estes estudos tem sido conduzidos dentro da Abordagem do Processamento Humano da Informação, como refere Sternberg (1992a, 1992b). Geralmente, os processos mentais avaliados pelos testes de fator g (Raven, G36, INV, D-48, D70) envolvem três funções psicológicas, como descreveu Spearman (apud Almeida, 1988):

a) *Apreensão das experiências*: é uma capacidade ligada à *percepção*. Diz respeito à rapidez e à acuidade com que as pessoas percebem os *estímulos*, bem como aos processos de *auto-percepção* da atividade consciente.

b) *Educação de relações*: é uma capacidade maior ou menor de estabelecer *relações* entre duas ou mais idéias, sejam elas originadas da percepção ou de representações mnêmicas (como na figura 2, descobrir as relações entre as duas figuras geométricas na parte superior do desenho, ou entre as duas palavras)

c) *Educação de correlatos*: é uma capacidade maior ou menor que as pessoas demonstram de *criar novas idéias a partir de uma idéia e uma relação*.

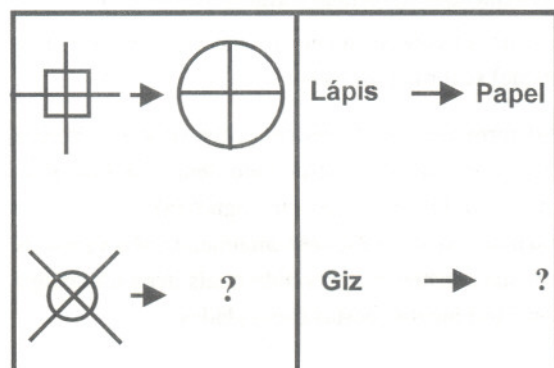


Figura 2

Estes processos estariam presentes em todas as atividades mentais, não importando o conteúdo (matemático, verbal, artístico etc.).

Basicamente, há processos ligados mais à percepção e outros ligados ao relacionamento das informações. Os estudos iniciados pela teoria da Gestalt, e que orientam pesquisas atualmente (Leeuwenberg & Boselie, 1988), trazem compreensões sobre os processos ligados à percepção. Postula-se que na percepção atua uma tendência inerente ao sistema nervoso, que faz com que interpretações perceptivas dos estímulos sensoriais (as formas ou unidades percebidas) sejam as mais simples, simétricas, bem formadas, dadas as condições naturais da estimulação. Esta tendência é conhecida atualmente pelo conceito de *princípio mínimo*. Os bem conhecidos princípios de organização perceptiva - semelhança, proximidade, boa continuidade, fechamento (Köhler, 1945) são especificações desta tendência à economia. Na figura 3, tende-se a ver dois círculos e uma seta sobre eles e não dois círculos irregulares. Isto porque esta última configuração é mais complexa do que as iniciais, que representam interpretações fechadas e bem formadas. Este princípio tem sido testado experimentalmente em várias áreas, nos estudos de percepção (ibidem).

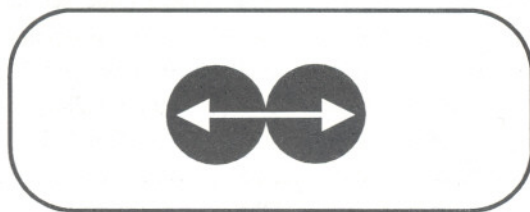


Figura 3

Os princípios dinâmicos da percepção podem ser estendidos aos processos de raciocínio envolvidos em tarefas presentes em testes de inteligência não verbal. Nelas há um movimento dinâmico onde a estimulação é organizada, resultando formas finais simples, simétricas e bem formadas. Neste caso, as figuras são organizadas em “todos” logicamente coerentes que são, portanto, configurações finais mais simples do que um aglomerado de formas não relacionadas. Nestes itens o sujeito deve formar uma espécie de conceito, que explique as relações entre as formas (educação de relações ou raciocínio indutivo)

e lhe possibilite produzir novas formas coerentes com o todo, que geralmente ele deve apresentar como resposta (edução de correlatos ou raciocínio dedutivo).

Nos estudos de Simon (1972) e Simon & Kotovsky (1963) é estabelecida uma relação entre *economia de informação* e *dificuldade* em processos superiores de raciocínio. Estes autores estudaram a formação de conceitos de padrões seqüenciais presentes em testes de QI, como é exemplificado pela seqüência abaixo:

**atbataatbat** \_\_

Nestas atividades, os sujeitos devem perceber as relações que governam a sucessão de símbolos - assim como a edução de relações discutida acima - e posteriormente extrapolar esta seqüência, apontando a próxima letra da série - assim como a edução de correlatos.

Para realizar esta atividade, os autores entendem que o sujeito, inicialmente, procura encontrar uma periodicidade entre as letras, buscando relações que se repetem em intervalos regulares, ou senão relações interrompidas em intervalos regulares. Abaixo é apresentado o exemplo citado acima já dividido em períodos:

**atb ata atb at** \_\_

Feito isto, ele procura os detalhes destes padrões, buscando novamente relações dentro de um período, ou entre símbolos em posições correspondentes de períodos sucessivos. No exemplo acima, cada período é composto por três letras. Observa-se que as duas primeiras posições de cada período sempre são ocupadas pelas letras "a" e "t". Já a terceira posição de cada período é ocupada pelas letras "b" e "a" intercaladamente.

Este processo é denominado **gerador de padrões**, sugerindo que, ao final dele, o sujeito terá, em sua memória, a curto prazo um conjunto de símbolos que representam o padrão analisado. Este conjunto é denominado **descrição de padrões**. Abaixo é exemplificada uma forma de representação do exemplo citado:

**[a, t, (b, a)]**

Esta descrição representa os períodos. Sendo assim, as letras "a" e "t", depois da chave, significam que a primeira e a segunda letras de cada período são um "a" e um "t". As letras "b,a" dentro dos parênteses significam que no primeiro período a terceira letra é "b" e no segundo "a".

Portanto, é concebido que os sujeitos descubram os padrões, construam representações destes padrões, armazenem na memória a curto prazo e, finalmente, derivem dos padrões listas de processamento que os habilitem a operar transformações nas informações simbólicas, tornando possível a extrapolação das seqüências.

Nota-se, portanto, que as *descrições de padrões* são conceitos, ou maneiras econômicas de se representar, particularmente, no experimento discutido, as seqüências de letras, mas, de um modo geral, diversos padrões seqüenciais. Os autores chegaram a analisar a dificuldade dos itens, como função direta do número de informações ou símbolos requeridos para formar as *descrições de padrões* para estes itens. Contando o número de símbolos de uma dada descrição de padrão, tem-se um valor da carga de informação requerida para representar uma seqüência. (este índice é denominado "*I*"). Portanto, os itens com uma carga alta de informação a ser armazenada nas *descrições de padrões* são mais complexos do que itens com baixa carga de informação. Uma das explicações para isto é que nos itens com um valor alto de "*I*" há uma exigência maior das capacidades de memória a curto prazo.

Outro ponto sugerido, mas não discutido sistematicamente, se refere às relações espúrias que poderiam ser feitas levando a respostas erradas. É o caso do exemplo apresentado abaixo:

**urtustuttu** \_\_

Que dividido em períodos resulta em:

**urt ust utt u** \_\_

Neste exemplo, em cada período de três letras a primeira e última posição são ocupadas pelas letras "u" e "t". A segunda posição de cada período é sucessivamente ocupada por uma progressão do alfabeto começando, no primeiro período, pela letra "r".

Embora a descrição de padrões para esta seqüência seja relativamente simples em termos de número de símbolos requeridos para sua representação, há algumas *relações por semelhança* que podem levar a organizações incorretas dos períodos da seqüência, por exemplo com as últimas quatro letras: *u t t u*. Quanto a este aspecto, parece-nos, portanto, que os conceitos da Gestalt sobre os princípios de organização perceptiva podem aprimorar a compreensão da dificuldade dos itens.

Retornando à percepção visual, tem-se que há diferentes níveis de organização atuando em uma tarefa complexa como a presente, em testes de QI. Um nível mais básico está ligado à organização perceptiva. Em nível superior está a organização destas formas, em relacionamentos mais sutis e abstratos. As interações entre estes níveis pode ser tomada como uma explicação da dificuldade dos itens. Isto porque organizações preferidas por serem simples em nível de organização perceptiva podem dificultar os processos de relacionamento em nível superior, ou vice versa. Convencionou-se desde já que o primeiro tipo de interferência será denominado: *interferência positiva* e o segundo, *interferência negativa*. Portanto, quando existem interferências negativas, pode-se dizer que há uma sobrecarga de informação irrelevante, já que as relações encontradas em nível perceptivo não coincidem com as relações que devem ser descobertas em nível superior mais abstrato. Sabe-se que no desenvolvimento da capacidade de raciocínio ocorre um gradativo desprendimento dos níveis concretos de atividade mental, atingindo, portanto, um ponto onde se chega a uma autonomia dos processos de raciocínio sobre algumas áreas da percepção (Vygotsky, 1989)

É de se salientar que os avanços atuais da Psicologia Cognitiva parecem gradualmente recuperar conceitos e leis desenvolvidas pela Gestalt, mas que por falta de uma metodologia de pesquisa mais apurada acabaram sendo esquecidos durante certo tempo. Anderson (1990) cita os estudos sobre organização perceptiva como sendo parte integrante dos processos básicos de percepção e atenção. Além dos estudos citados nos EUA, pesquisadores de laboratórios europeus também seguem esta diretriz. Por exemplo, estudos sobre a audição organizadora, feito

por Demany e McAdams (1992) buscaram aprofundar os conhecimentos sobre percepção auditiva, tomando como referência os trabalhos de Wertheimer (1923) sobre as leis das formas.

Em suma, pode ser dito que a dificuldade de itens de teste de inteligência está ligada a exigências quanto à capacidade de seleção de informações relevantes em níveis perceptivos e com a capacidade de trabalhar várias informações ao mesmo tempo na memória a curto prazo. Neste trabalho procuramos desenvolver o primeiro ponto ligado ao nível perceptivo. Basicamente a hipótese geral foi que a dificuldade de itens de raciocínio não verbal ligados ao fator g estaria ligada a interferências negativas entre os processos ligados à percepção e ao relacionamento de informação. Isto ocorreria nestes casos, pois o sujeito deve ter uma maior capacidade de desprendimento dos níveis concretos de raciocínio para que consiga reestruturar as organizações concretas (processo de seleção de informações relevantes), buscando relações mais equilibradas em níveis abstratos. Portanto, nestes casos, o sujeito estaria mais propenso ao erro. Reversamente, se conseguíssemos reverter esta interferência de negativa para positiva, conseguiríamos facilitar o item. Isto foi o que se buscou neste trabalho.

## Método

### Sujeitos

Participaram desta pesquisa 28 sujeitos, sendo 8 do sexo masculino e 20 do sexo feminino, abrangendo as idades de 14 a 21 anos. Estes sujeitos faziam parte de uma classe de 1º ano do segundo grau de uma escola estadual no interior do Estado de São Paulo.

### Materiais

Inicialmente foram escolhidos seis itens do teste D-48 (Pichot, 1961) e seis itens do teste D-70 (Editions du Centre de Psychologie Appliquée, 1985). Dos doze itens selecionados, nove apresentavam dificuldade elevada e três dificuldade baixa. Como os manuais dos testes não apresentam o índice de dificuldade dos itens, decidiu-se escolher os últimos itens apresentados nos cadernos, já que os testes são organizados em ordem crescente de dificuldade.

Foram selecionados dois exemplos de cada teste relacionados aos itens escolhidos. Construiu-se, portanto, um caderno composto por quatro exemplos e mais 12 itens (o qual chamamos de caderno controle).

Um segundo caderno foi construído (chamado caderno experimental), mantendo a mesma estrutura e os mesmos itens, alterando-se no entanto a organização perceptiva segundo os princípios da Gestalt descritos anteriormente. Buscou-se, neste caderno, tornar mais concreta a relação entre os dominós, como é exemplificado na figura 4.

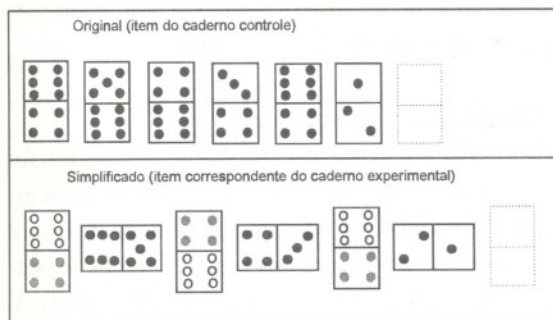


Figura 4

Tal estratégia traduz a tentativa de impor-se à organização perceptiva de cada item uma interferência positiva. O artifício empregado para tal intento,

neste item, contou com duas estratégias: (1) a decomposição das duas relações principais a serem inferidas entre os dominós, inicialmente através da mudança de posição dos mesmos, própria a cada relação. Assim, as peças que traziam a pontuação seis-quatro, e que aparecem na vertical no caderno controle, foram mantidas nesta posição, enquanto que as peças que variavam com pontuações cinco-seis, três-quatro, um-dois, foram colocadas em linha horizontal. Destacam-se aqui as relações: inversão dos mesmos pontos (seis-quatro) e decréscimo de pontuação (de seis a um) (2) Dentro da relação inversa, a pigmentação dos pontos foi manipulada, onde a metade com seis pontos aparece sempre em uma cor e a metade com quatro pontos apresenta outra cor.<sup>1</sup>

Assim como mostra a figura 4, o item original específico do primeiro caderno (caderno controle) apresentava as duas sequências alternadas. Na sua versão simplificada (pertencente ao caderno experimental), os dominós que fazem parte de cada sequência se tornam mais evidentes por agrupamento de semelhança. Por outro lado, a progressão decrescente fica mais evidente do que nos itens originais em função do princípio de organização de boa continuidade.

Na Tabela 1 são apresentados o número dos itens empregados que foram extraídos dos testes originais e sua ordem de apresentação, bem como os

Tabela 1 - Origem dos itens e princípios de organização utilizados na construção do caderno experimental

Nº do item nos testes originais	Teste	Ordem de apresentação no caderno de pesquisa	Semelhança de cor	Semelhança de posição	Proximidade	Boa Continuidade
4	D-48	1	X			
6	D-70	2	X			
19	D-48	3	X		X	
16	D-70	4	X			
5	D-48	5	X			
9	D-48	6	X		X	
20	D-70	7	X	X		X
35	D-70	8	X	X		X
32	D-48	9	X			X
44	D-48	10	X		X	
39	D-70	11	X			
41	D-70	12	X	X		

1. No caderno original foram usadas diferentes cores. Aqui, no entanto, estas cores são representadas por variações de cinza.

princípios usados na simplificação. Além destes dois cadernos, foi usado também o teste de raciocínio não verbal G-36 (Boccalandro, 1984), aplicado com objetivo de fixar um ponto de referência em termos de Fator g, para todos os sujeitos avaliados por meio da norma do respectivo teste.

### Procedimento

Depois de informar os sujeitos sobre os objetivos da pesquisa, e obtida a concordância dos mesmos em participar das atividades, iniciou-se a fase de coleta de dados. Esta foi feita em três encontros, durante aulas de Psicologia

Inicialmente, todos os sujeitos responderam o teste G-36. Em função dos resultados, os 28 sujeitos foram divididos em dois grupos, compostos por 14 sujeitos cada, emparelhados quanto ao nível de inteligência e ao sexo. Com esta divisão objetivou-se criar dois grupos semelhantes quanto às variáveis diferenciais analisadas. Cada grupo respondeu somente a um caderno de pesquisa (o grupo controle recebeu o caderno controle e o grupo experimental, o caderno experimental).

Tecnicamente, os escores de um grupo de sujeitos a um determinado teste estão condicionados a duas variáveis principais (não considerando aqui o erro de medida): (1) dificuldade dos itens do teste (2) grau de capacidade dos sujeitos que o responderam.

Sendo o objetivo principal deste trabalho compreender a dificuldade que os itens impõem ao processo de raciocínio em função de conceitos teóricos sobre a estrutura dos itens, objetivou-se, com a divisão citada acima, manter constante o nível de inteligência entre os dois grupos. Sendo assim,

havendo variância entre os escores dos cadernos controle e do caderno experimental, esta deveria ser atribuída primordialmente às alterações introduzidas nos itens. Em outras palavras, havendo diferenças entre os grupos, estas deveriam ser atribuídas à condição 1 (dificuldade dos itens) já que a condição 2 (diferenças de inteligência entre os dois grupos) foi controlada. Portanto, buscou-se com este delineamento isolar a variância ligada somente à dificuldade dos itens.

Resumindo, foi empregado um delineamento experimental onde a variável independente foi as alterações nos itens (com o objetivo de simplificá-los) e a variável dependente, o número de acertos. No grupo controle foram aplicados os itens normais e no grupo experimental, os itens alterados.

### Resultados e Discussão

Para as análises estatísticas dos resultados, foi utilizado o pacote informatizado "Statistics for Windows". Na tabela 2 são apresentadas as médias e desvios-padrão dos resultados nos cadernos de pesquisa nos grupos controle (GC) e experimental (GE). Dentro de cada grupo, os resultados são subdivididos em função dos escores obtidos no G-36: agrupados sob a designação "baixo", sujeitos com escores menores ou iguais à mediana; no grupo "alto", sujeitos com escores acima da mediana.

Para uma pontuação máxima de 12 pontos em ambos cadernos, nota-se que no grupo controle a média total de acertos foi 7,00 e no grupo experimental 8,46. Tanto no grupo controle quanto no experimental as médias de acertos no teste G-36 permaneceram constantes, aproximadamente 23,5

**Tabela 2** - Médias e desvios padrão dos resultados nos cadernos em função do resultado no teste G36 e dos grupos experimental e controle

Controle GC/ Experimental GE	Escore no teste G-36	Média nos cadernos de pesquisa	N	S	Média no G-36	N	S
GC	BAIXO	6,28	7	1,70	19,75	8	3,91
GC	ALTO	7,83	6	1,94	28,66	6	2,73
GC Total		7,00	13	1,91	23,57	14	5,66
GE	BAIXO	8,14	7	2,03	19,37	8	5,22
GE	ALTO	8,83	6	1,16	28,66	6	2,92
GE Total		8,46	13	1,66	23,35	14	6,39
Média Total		7,73	26	1,90	23,46	28	5,97

(com resultado máximo possível de 36), ou seja, há tanto sujeitos com capacidade alta quanto baixa nos dois grupos, como objetivou-se inicialmente, garantindo a equiparação entre os grupos.

Para avaliar a diferença de resultados entre os grupos foi utilizado o teste de Wilcoxon, que mostrou ser aquela significativa estatisticamente ( $T=7$  e  $Z=2,31$ ,  $p,05$ ). Assim, pode-se afirmar que há um aumento do número de acertos em função das manipulações experimentais, ou seja, uma variância sistemática do número de acertos em função do tratamento experimental, como inicialmente se objetivou.

O delineamento empregado pode ser interpretado de uma segunda forma. Pode-se supor que houve um delineamento fatorial para a criação dos grupos em função de duas variáveis independentes: resultados altos/baixos no teste G-36 e grupo controle/grupo experimental. Portanto, um delineamento 2x2. Sendo assim, seria de se esperar que o número de acertos estivesse relacionado aos efeitos principais e independentes das duas variáveis. Assim, poder-se-ia prever os melhores resultados nos cadernos experimentais, como segue abaixo (revelando, portanto, a aditividade hipotetizada):

Sujeitos do grupo experimental e que obtiveram escores altos no teste G-36; Sujeitos do grupo experimental e que obtiveram escores baixos no teste G-36; Sujeitos do grupo controle e que obtiveram escores altos no teste G-36; Sujeitos do grupo controle e que obtiveram escores baixos no teste G-36.

Para testar esta hipótese, foi utilizado um método de análise de tabelas cruzadas de múltipla entrada denominando "Log linear analysis". (Presente no pacote estatístico Statistics for Windows).

Inicialmente, foi construída uma variável dicotômica, baseada nos escores nos cadernos de pesquisa (tanto do caderno controle como do caderno experimental): os sujeitos com resultado *menor do que a mediana* foram classificados como pertencentes ao *grupo com baixo desempenho* e sujeitos com resultado *igual ou maior do que a mediana* como pertencentes ao *grupo com alto desempenho* (nota-se que a mediana foi extraída dos escores do grupo controle e experimental em conjunto). Assim, os

resultados podem ser estruturados em uma tabela 2(Desempenho nos cadernos de pesquisa) X 2(Grupo Controle ou Experimental) X 2(Grupo de desempenho no teste G-36), como se vê na tabela 3:

**Tabela 3** - Distribuição dos sujeitos segundo escores alto e baixo nos cadernos de pesquisa, em função do resultado no teste G36 e dos grupos experimental e controle

Escore nos cadernos	Grupo Controle	Grupo Experimental	Total
Escore baixo no G-36			
Baixo	5	2	7
Alto	2	5	7
Total	7	7	14
Escore alto no G-36			
Baixo	3	1	4
Alto	3	5	8
Total	6	6	12

Nota-se na tabela acima uma relação entre desempenho nos cadernos de pesquisa e pertença aos grupos experimental e controle, ou seja, a distribuição dos escores (categorizados como alto e baixo) se altera quando se consideram o grupo controle e experimental: no grupo experimental (terceira coluna da tabela), geralmente há mais pessoas com escores altos (5 para 2 e 5 para 1), enquanto que no grupo de controle há um número maior de pessoas com escores baixos para aqueles com mau desempenho no teste G-36 (5 para 2), e um número idêntico de pessoas com escores baixos e altos para aqueles com bom desempenho no teste G-36 (3 para 3).

Estes resultados se tornam mais evidentes por meio da análise do histograma bivariado (figura 5). Nele estão representados os resultados do escore dicotômico em função das variáveis escore no G36 e Grupo controle e experimental. Nota-se nele o tratamento experimental e o escore no teste. Como foi discutido acima, há geralmente mais pessoas com escores altos no grupo experimental (GE). No grupo controle, por outro lado, quando se considera os sujeitos com baixa habilidade (GC G36 baixo), nota-se uma grande concentração de escores baixos. Já no

grupo com alto desempenho (GC G36 alto), a distribuição de escores é homogênea.

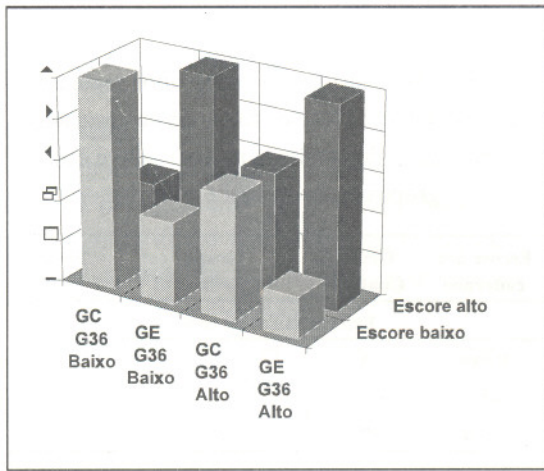


Figura 5 - Histograma bivariado do número de acertos

Um ponto interessante é que quando se considera sujeitos com pouca habilidade (G36-baixo) nota-se uma inversão dos escores. (5 para 2 e 2 para 5). No grupo de maior habilidade, esta inversão não é tão evidente, já que há poucos sujeitos com baixo escore. *Isto sugere que o tratamento obteve um efeito maior quando se considera sujeitos com baixa habilidade.* Portanto, não houve um efeito aditivo, como se supôs anteriormante.

Pode-se dizer que o tratamento experimental aproximou o escore de sujeitos com baixa habilidade dos escores de sujeitos com boa habilidade. Portanto, a simplificação pode ter aproximado os dois grupos em termos de desempenho, sendo que tanto pessoas com bom desempenho quanto aquelas com mau desempenho atingiram resultados altos no teste experimental.

A tabela 2 confirma este fato, já que que no grupo controle a média de acertos no teste experimental para sujeitos com alto e baixo desempenho no teste G-36 é diferenciada, ou seja, sujeitos com bom desempenho acertam mais (média=6,28 S=1.70) do que sujeitos com mau desempenho (média 7,83, S=1.94). Já no gupo experimental, esta diferença é menos evidente (8,14, S=2,03 versus 8,83, S=1,16).

Para analisar esta tabela foi utilizado o procedimento Log linear. Nesta análise foram testados vários modelos onde ora se supunha ser importante só os efeitos prinipais, ora interações de primeira ordem, segunda ordem etc. O teste então avalia quando as freqüências esperadas sob um dado modelo são significativamente diferentes das freqüências observadas. Se forem diferentes, o modelo é rejeitado. Procedendo esta análise, notouse que o modelo que melhor se adaptou aos dados aqui apresentados foi o que prevê o desempenho relacionado com a pertença ao grupo controle ou experimental, onde o Qui Quadrado de Máxima Verossimilhança = 0.86, gl=4 p0.92 (Maximum Likelihood Chi-Square). A hipótese inicial dizia que este estaria relacionado também com o desempenho no teste G-36, mas o efeito deste fator não participa significativamente na adaptação dos dados. O ponto principal mostrado por estes resultados é que uma parte do número de acertos esteve condicionada às manipulações experimentais, como se hipotizou no início deste trabalho.

## Conclusão

Geralmente, na elaboração de testes é pressuposto que a inteligência seja normalmente distribuída e busca-se construir instrumentos que revelem esta distribuição. Ou seja, elabora-se num teste uma grande parte de itens com dificuldade mediana, poucos itens com dificuldade elevada e poucos com dificuldade baixa. Estes são escolhidos em função da percentagem de acertos. Um item é considerado mais difícil do que outro quando menos pessoas conseguem resolvê-lo. Todavia, tal análise não desvenda sua dificuldade em termos do tipo de natureza do processamento mental exigido para sua resolução.

Além dessa análise via média de acertos, todos os itens são somados algébricamente para compor o escore ou resultado bruto. A comparação entre indivíduos é inevitável: uma pessoa que acerta mais itens é mais inteligente do que outra que acerta menos, sem contar limitações já conhecidas dos testes de inteligência, que geralmente devem restringir-se à exploração de poucos (por vezes um único) modos de raciocínio dentre outros possíveis do sujeito - por exemplo, o raciocínio analógico abstrato.

Uma análise da estrutura em si dos instrumentos deve ser aprofundada.

Contudo, mantendo um olhar mais apurado para os itens do teste de inteligência, pode-se logo notar os problemas com os quais nos defrontamos quando estamos tentando medir a inteligência. Como foi dito anteriormente, não se pode afirmar que o escore no teste seja uma variável contínua como o é o traço latente em análise. A suposição de que cada item esteja medindo um intervalo desta escala é um pressuposto estatístico que se faz necessário face à utilidade prática que um modelo baseado nestas pressuposições poderia trazer. Entretanto, em nível da realidade psicológica, pode-se dizer que os itens trazem informações sobre várias componentes da atividade cognitiva (ora sobre a apreensão, ora sobre a educação de relações, ora sobre a educação de correlatos), e em diferentes níveis, havendo portanto uma certa “impureza” de medida do traço em análise. Ou seja, há uma grande probabilidade que as componentes exigidas e a intensidade desta exigência variem de item para item. Como consequência, os “intervalos” que cada item mede podem estar sendo diferentes. Portanto, quando somamos o resultado nos itens para compor um escore bruto devemos ter consciência de que podemos estar “somando coisas diferentes”, sendo este procedimento somente justificável em termos pragmáticos, como foi discutido anteriormente.

O presente trabalho traz uma contribuição para a melhor compreensão da qualidade do “intervalo” que está se medindo, por ter separado a variância da dificuldade dos itens em função de um conceito sobre a natureza do raciocínio. Por outro lado, parece-nos importante o fato de termos identificado um “continuum”, na medida em que mudando a organização perceptiva mas mantendo a mesma tarefa a ser realizada, esta parece ser vista como mais simples ou mais complexa. Isto pode levar a uma melhor compreensão da intensidade com que um item mede determinada componente. Acreditamos que com isto seja possível iniciar uma linha de pesquisa que busque aprimorar o desenvolvimento dos testes de inteligência, baseados mais profundamente em concepções sobre o conceito de inteligên-

cia associadas intimamente a uma metodologia experimental de validação.

Todavia, mais e mais entende-se a atividade cognitiva como um complexo e particular emaranhado de componentes ativadas e articuladas pelo sujeito, visto de modo contextualizado. Assim, perceber o ambiente, identificar e selecionar os elementos de informação que serão efetivamente trabalhados durante a resolução de uma tarefa, são aspectos essenciais no estudo da atividade inteligente, bem como a natureza da tarefa proposta, a situação em que tal relação ocorre, a representação que o sujeito elabora da tarefa que lhe é proposta etc. (Richard, Bonnet & Guijhione, 1990). O aprofundamento dos conhecimentos sobre cognição enfocando qualquer uma dessas componentes rapidamente enumeradas é o caminho mais fecundo que a Psicologia está trilhando nesses últimos anos.

Atualmente uma grande mudança vem ocorrendo nos procedimentos de validação de testes, de um modo geral. Enquanto que, os métodos estatísticos clássicos estão centrados em escores baseados em soma de itens, atualmente a unidade de análise é o item. Esta metodologia vem se consolidando sob o nome de Teoria de Resposta ao Item, como já citamos anteriormente (*Item Response Theory*). Basicamente são estudadas equações que relacionam a habilidade cognitiva (traço latente inobservável diretamente, denominado theta) com a probabilidade de acerto no item. Há modelos que incluem a dificuldade do item e a capacidade discriminativa nestas equações que são chamadas curva característica do item (Baker, 1992). Um destes modelos, o qual a nosso ver pode trazer uma grande contribuição para a avaliação da inteligência, foi desenvolvido por Whitely (1980a, 1980b, 1980c). Neste modelo, as equações que descrevem as curvas características dos itens envolvem parâmetros de dificuldade para cada componente cognitiva. Uma perspectiva futura para o presente trabalho é estudar como as alterações introduzidas alteram o parâmetro de dificuldade da curva característica dos itens e bem como em que componente esta alteração atua.

Com relação à dimensão teórica, confirma-se, tal como no trabalho de outros pesquisadores, que a teoria da Gestalt pode trazer importantes con-

tribuições no que se refere à compreensão da organização dos processos cognitivos. Tal abordagem possibilitou transformar a informação oferecida ao sujeito, facilitando-lhe a resposta. Estes dados parecem promissores e convidam os pesquisadores interessados ao aprofundamento e à implementação de novos trabalhos na área.

## Referências

- Almeida, L. S. (1988a). *Teorias da inteligência*. Porto: Edições Jornal de Psicologia.
- Almeida, L. S. (1988b). *O raciocínio diferencial dos jovens*. Porto: Instituto Nacional de Investigação Científica.
- Almeida, L. S. (1993). *Relatório da disciplina de métodos de observação e investigação Psicológica-1º ano*. Braga: Universidade do Minho.
- Anderson J. R. (1990). *Cognitive psychology and its implications*. New York: W. H. Freeman.
- Baker, F. B. (1992). *Item response theory parameter estimation techniques*. New York: Marcel Dekker.
- Boccalandro, E. R. (1984). *G-36 Teste Não Verbal de Inteligência*. São Paulo: Vetor.
- Boselie, F. & Leeuwenberg, E. (1986). A test of the minimum principle requires a perceptual coding system. *Perception*, 15, 331-354.
- Demany, L. & McAdams, S. (1992). *L'audition organisatrice*. Courrier du CNRS, Dossiers Scientifiques Sciences Cognitives, N° 79, octobre.
- Editions du Centre de Psychologie Appliquée. (1985) *Teste D 70*. São Paulo: Casa do Psicólogo.
- Ferguson, G. A. (1966). *Statistical analysis in psychology and education*. New York: McGraw-Hill.
- Hambleton, H. K., Swaminatham, H., Cook, L. L., Eignor, D. R. Gifford, J. A. (1978). Developments in latent trait theory: models, technical issues, and applications. *Review of Educational Research*, 48, n° 4, 467-510.
- Hatfield, G. & Epstein, W. (1985). The status of the minimum principle in theoretical analysis of visual perception. *Psychological Bulletin*, 97, n° 2, 155-186.
- Leeuwenberg, E. & Boselie, F. (1988). Against the likelihood principle in visual Form perception. *Psychological Review*, 95, n° 4, 485-491.
- Lord, F. M. & Novick, N. R. (1974). *Statistical theories of mental test scores*. Reading: Addison -Wesley.
- Muniz, J. (1990). *Teoría de respuesta a los items: un nuevo enfoque en la evolución psicológica y educativa*. Madrid: Ediciones Pirámide.
- Palmer, S. E. (1992). Common region: A new principle of perceptual grouping. *Cognitive Psychology*, 24, 346-447.
- Pichot, P. (1961). *Teste Dominós D 48*. Rio de Janeiro: Cepa.
- Richard, J. F.; Bonnet, C. & Guijhione, R. (1990). *Traité de psychologie cognitive*, Tome 2. "Le traitement de information symbolique" Paris: Dunod.
- Simon, H. A. (1972). Complexity and the representation of patterned sequences of symbols. *Psychological Review*, 79, n° 5, 369-382.
- Simon, H. A. & Kotovsky, K. (1963). Human acquisition of concepts for sequential patterns. *Psychological Review*, 70, n° 6, 534-546.
- Sternberg, R. (1992a). O que é uma abordagem em processamento de informações às capacidades intelectuais humanas. In Sternberg, R. J., ed. *As capacidades intelectuais humanas: uma abordagem de processamento de informações*. Porto Alegre: Artes Médicas, Pág 13-16.
- Sternberg, R. (1992b). A capacidade intelectual geral. In Sternberg, R. J., ed. *As capacidades intelectuais humanas: uma abordagem de processamento de informações*. Porto Alegre: Artes Médicas. Pág 17-42.
- Vygotsky, L. S. (1989). *A formação social da mente*. São Paulo: Martins Fontes.
- Whitely, S. E. (1980a). Latent trait models in the study of intelligence. *Intelligence*, 4, 97-132.
- Whitely, S. E. (1980b). Multicomponent latent trait models for ability tests. *Psychometrika*, 45, 479-494.
- Whitely, S. E. (1980c). Modeling aptitude test validity from cognitive components. *Journal of Educational Psychology*, Vol 72, N° 6, 750-769.