

Introdução aos “testes de QI”

Por Hindenburg Melão Jr.
<http://www.sigmasociety.com>

Um pouco de história:

Os testes de inteligência começaram a ser usados na China, no século V d.C., mas só se tornaram instrumentos clínicos e passaram a seguir procedimentos padronizados para aplicação e interpretação no século XX, mais precisamente em 1904, com os trabalhos pioneiros de Alfred Binet e Theodore Simon, e foram publicados pela primeira vez em 1905. As escolas da França precisavam de um método objetivo, eficiente e rápido que possibilitasse diagnosticar deficiências mentais, e para isso Binet e Simon criaram o primeiro teste de inteligência do Ocidente. Originalmente mediam-se níveis mentais (Binet não gostava do termo “idade mental”) por meio da relação entre a idade cronológica e o desempenho em testes destinados a outras faixas etárias. Assim se uma criança com idade de 10 anos tivesse mesmo desempenho que a média das crianças de 8 anos, significaria que ela tinha um nível mental de 8 anos. Com isso se tornou possível medir o atraso ou o adiantamento mental das crianças, possibilitando oferecer a cada uma delas uma educação mais personalizada, adequada a seu próprio ritmo de aprendizado.

Em 1912, William Stern sugeriu o termo “QI” (quociente de inteligência) para representar o nível mental, e propôs que o QI fosse determinado pela divisão da idade mental pela idade cronológica. Assim a criança com idade mental de 10 anos e nível mental de 8 anos teria QI 0,8, porque $8/10 = 0,8$. Em 1916, Lewis Terman propôs multiplicar o QI por 100, a fim de eliminar a parte decimal, e esta fórmula foi mantida durante décadas e continua sendo usada por muitos até hoje: $QI = 100 \times IM/IC$, em que IM = idade mental e IC = idade cronológica. Por esta fórmula, uma criança de 10 anos com idade mental de 8 teria QI 80, indicando um atraso sensível em comparação à média das crianças de mesma idade, enquanto uma criança com 10 anos e idade mental de 12 teria QI 120, indicando um adiantamento sensível em comparação à média das crianças de sua idade.

Terman também propôs classificações para os diferentes níveis mentais em função dos QIs obtidos em testes. Estas classificações foram adotadas durante várias décadas e algumas clínicas continuam a usá-las até hoje:

Classificação proposta por Lewis Terman (~1916)

QI acima de 140:	Genialidade
120-140:	Inteligência muito superior
110-120:	Inteligência superior
90-110:	Inteligência normal (ou média)
80-90:	Embotamento
70-80:	Limítrofe
50-70:	Cretino
20-50:	Imbecil
QI abaixo de 20:	Idiota

De acordo com Stern, uma criança com 5 anos e idade mental de 10 anos teria QI 200, e outra criança com 7 anos e idade mental de 14 também teria QI 200. Para ele, tanto a criança de 5 aos quanto a de 7 anos, quando se tornassem adultas, provavelmente teriam QI 200. Essa

idéia também foi abraçada por Terman e todos os expoentes mundiais da Psicometria naquela época, no entanto, quando essas crianças se tornam adultas, aquela que tinha QI 200 aos 7 anos tende a se tornar um adulto com QI em torno de 160, enquanto aquela que tinha QI 200 aos 5 anos tende a se tornar um adulto com QI 145. Isso é mau para a teoria, porque a invariabilidade do QI é uma propriedade muito importante. Essa distorção foi extensivamente confirmada por um estudo desenvolvido pelo próprio Terman, ao longo de mais de três décadas, com um grupo de 1.528 crianças com QIs acima de 130. Nesse estudo, ele encontrou crianças com QIs acima de 200 num nível de abundância 10.000 vezes acima do que seria esperado numa população adulta, e a incidência era tanto maior se tanto mais jovens fossem as crianças, demonstrando que havia algo errado com o método adotado para determinar o QI. Por isso o conceito de QI como representação da divisão da idade mental pela cronológica foi substituído pelo conceito de QI em função da raridade. Essa mudança foi proposta por David Wechsler. Ele também sugeriu que o antigo conceito de QI passasse a ser chamado *ratio-IQ*, enquanto o novo conceito se chamaria *deviation-IQ*. Algumas referências on-line:

<http://sweb.uky.edu/~jcscov0/ratioiq.htm>

<http://www.psychologicaltesting.com/iqtest.htm>

<http://www.psych.usyd.edu.au/difference5/scholars/wechsler.html>

<http://www.wilderdom.com/intelligence/IQCautionsInterpretingIQ.html>

Na década de 1940, o uso de *deviation-IQ* em lugar de *ratio-IQ* já estava generalizado e as escalas de QI passaram a ser construídas com base em níveis de raridade. Para isso, utiliza-se a inversa de uma distribuição normal cumulativa, e o QI é determinado pelo escore padronizado z que representa a porcentagem de pessoas com desempenho mais baixo que o daquela cujo QI se pretende determinar. Assim, uma pessoa situada no percentil 50 tem QI 100, pois ela está acima de 50% da população e abaixo de outros 50%, logo ela tem exatamente o QI médio, que é 100. Uma pessoa situada no percentil 84% tem QI 115 (numa escala com $\sigma=15$), porque, numa distribuição normal, 84% dos escores devem ficar abaixo de $+1\sigma$, e um QI de 115 está no nível $+1\sigma$ (1 desvio-padrão acima da média). Uma pessoa situada no percentil 97,7% tem QI 130, porque 97,7% dos escores devem ficar abaixo de $+2\sigma$, e um QI de 130 está no nível $+2\sigma$. Uma pessoa situada no percentil 99,87% tem QI 145, porque 99,87% dos escores devem ficar abaixo de $+3\sigma$, e um QI de 145 está no nível $+3\sigma$. Por esse método, o QI de um sujeito é determinado com base na quantidade de pessoas de um determinado grupo que alcançam escores maiores ou menores do que ele. Se o grupo for constituído por uma população não-seleta, atribui-se à média desse grupo o QI 100. Se uma criança de 6 anos obtém um escore no topo de 2% de um grupo não-seleto constituído exclusivamente por crianças de 6 anos, então ela está no percentil 98 entre as crianças de 10 anos e seu QI é calculado em 130. Quando ela tiver 10 anos, provavelmente continuará no topo 2% das crianças de 10 anos e seu QI continuará sendo 130, e quando se tornar adulta (mais de 20 anos), provavelmente continuará no topo 2% de adultos e seu QI continuará sendo 130. A vantagem desse sistema de normatização é que na maioria das vezes o QI não apresenta variações significativas com a idade, tal como acontecia pelo método antigo de dividir a idade mental pela cronológica. Dos 6 anos até a idade adulta, o QI permanece quase inalterado, e é isso que pretendia Stern, mas só foi satisfatoriamente alcançado por Wechsler.

Além desse aprimoramento no sistema de padronização, Wechsler também propôs mudanças na nomenclatura das classificações de Terman, usando termos que hoje são reconhecidos como mais apropriados:

Classificação proposta por David Wechsler (~1940)

QI acima de 127: Superdotação (algumas fontes citam o termo “gênio” para QI>150)
120-127: Inteligência superior
110-120: Inteligência acima da média

90-110:	Inteligência média
80-90:	Embotamento ligeiro
65-80:	Limítrofe
50-65:	Debilidade ligeira
35-50:	Debilidade moderada
20-35:	Debilidade severa
QI abaixo de 20:	Debilidade profunda

Pouco tempo depois dos trabalhos iniciais de Wechsler, surgiram os primeiros rudimentos de Teoria de Resposta ao Item, começando com Richardson, em 1936, recebendo importantes contribuições de Lawley, Lord e Birnbaum e se consolidando com Georg Rasch, que em seu livro *“Probabilistic Models For Some Intelligence And Attainment Test”*, publicado em 1960, praticamente instaurou uma aliança entre a Matemática e a Psicologia, dando à Psicometria um aspecto mais robusto e mais científico. Nos anos seguintes, os avanços foram muito rápidos e significativos, em grande parte graças à rapidez de processamento proporcionada pelo advento do computador e alcançando seu apogeu com a popularização dos computadores pessoais, na década de 1980.

Em 1946, foi fundada a Mensa, primeira sociedade para pessoas de elevado QI, que só admitia membros com escore acima de 98% da população mundial, o que corresponde a um QI 133 ou acima. Nas décadas seguintes, foram criadas outras sociedades similares, com critérios seletivos progressivamente mais difíceis de serem atendidos, até chegar em 99,9%, que é o nível de corte de International Society for Philosophical Inquiry, fundada em 1974 e que só aceita membros com QI acima de 150. Antes de chegarem nesse nível, já haviam deparado com o problema de os testes existentes na época não terem conteúdo adequado para medir corretamente QIs acima de 130. A Mensa estava no limiar do que os testes eram capazes de medir, enquanto sociedades como Intertel e International Society for Philosophical Inquiry extrapolavam os limites de validade dos testes, de modo que não havia como garantir que alguém com QI 155 era mais inteligente do que alguém com QI 135, já que ambas estavam fora dos limites do que os testes podiam aferir com alguma segurança.

Até 1970, o uso de TRI constituiu um dos mais importantes avanços no tratamento estatístico dos testes, mas esta ferramenta não oferecia nenhuma contribuição para o conteúdo semântico dos testes, que continuavam sendo constituídos por perguntas muito básicas, muito fáceis, inadequadas para medir desempenhos em altos níveis. Para resolver esse problema, Kevin Langdon publicou, em 1973, os primeiros testes de inteligência apropriados para medir QIs acima de 130, podendo chegar a medir corretamente até cerca de 160 ou 170. Se estes testes existissem na época de Terman, os resultados do estudo que ele realizou com as 1.528 crianças talentosas provavelmente teriam sido muito mais interessantes, já que as crianças teriam sido selecionadas com base em critérios ligados ao pensamento profundo, complexo e criativo, em vez de apenas considerar a rapidez para solucionar problemas simples e de nível elementar. O grupo talvez não tivesse 1.528 crianças e certamente não seriam as mesmas, mas sim crianças com potencial de produção intelectual muito maior. Muitas crianças geniais passaram despercebidas pelos testes de velocidade usados na época, enquanto outras sem nenhum brilho particularmente notável acabaram sendo escolhidas para o estudo apenas por serem rápidas para lidar com questões primárias. Será muito interessante se algum pesquisador moderno repetir o estudo de Terman, porém usando instrumentos de seleção mais adequados.

Os testes de Kevin Langdon representam um separador de águas na história do conteúdo dos itens, sendo suas contribuições tão ou mais importantes do que os aprimoramentos de Binet. Por analogia, o avanço alcançado por Binet em comparação aos testes de velocidade dos

reflexos de Galton é semelhante ao avanço alcançado por Langdon em comparação aos testes de pensamento básico de Terman, Binet, Simon, Otis etc.

A partir do momento que passaram a existir testes adequados para medir desempenhos em níveis estratosféricos, também foram criadas sociedades de alto QI mais seletas. Em 1982, foram fundadas por Ronald Hoeflin Prometheus Society e Mega Society. Prometheus destina-se a pessoas com escores acima de 99,997% da população mundial (QI acima de 164) e Mega Society destina-se a pessoas com escores acima de 99,9999% da população mundial (QI acima de 176). Mega Society foi registrada no Guinness Book de 1990, por ser a sociedade de alto QI mais exclusiva do mundo, com corte teórico de 1 aprovado entre 1.000.000 de potenciais candidatos. Em 1985, Hoeflin criou seus próprios testes para admissão em suas sociedades: Mega, Titan e Ultra, contudo tanto os testes de Langdon quanto os de Hoeflin não eram apropriados para medir desempenhos em níveis tão altos, e o próprio Langdon escreveu artigos contestando a acurácia nos escores mais altos desses testes. Posteriormente, Bob Seitz e Hindenburg Melão Jr. também revisaram os cálculos dos tetos reais destes testes, cada um usando um método diferente, mas os três concordaram que os tetos do Mega Test e Titan Test ficam em torno de 170, não em 193, como pretendiam nas normas publicadas. Estima-se os cortes reais de admissão em Mega Society e Prometheus sejam respectivamente cerca de 1 em 20.000 e 1 em 10.000, com algumas estimativas mais otimistas de 1 em 100.000, mas ninguém que tenha investigado com profundidade o assunto considera seriamente um corte de 1 em 1.000.000.

Outras sociedades pretensamente mais seletas continuaram a ser criadas, como Giga Society, Sigma VI, Sigma V, Pars e Olympiq. Atualmente as mais seletas do mundo são:

Sigma VI

Corte teórico 1 em 1.000.000.000 (QI acima de 196)
Provável corte real: 1 em 1.000.000

Giga

Corte teórico 1 em 1.000.000.000 (QI acima de 196)
Provável corte real: 1 em 50.000 ou 1 em 100.000

Sigma V

Corte teórico 1 em 3.500.000 (QI acima de 180)
Provável corte real: 1 em 50.000

Pars

Corte teórico 1 em 3.500.000 (QI acima de 180)
Provável corte real para membros aprovados em hard tests: 1 em 50.000
Provável corte real para membros aprovados em testes clínicos: 1 em 500

Olympiq

Corte teórico 1 em 3.500.000 (QI acima de 180)
Provável corte real para membros aprovados em hard tests: 1 em 50.000
Provável corte real para membros aprovados em testes clínicos: 1 em 500

Além destes grupos há cerca de 100 outras sociedades cujos cortes de QI variam entre 126 e 176, contudo a maioria não adota testes adequados para avaliação de QIs acima de 170, de modo que só as sociedades com cortes até cerca de 160 é que podem de fato reivindicar legitimidade para seus cortes teóricos. Na opinião de Alexandre Prata Maluf, por exemplo, membro em 3 das sociedades mais exclusivas do mundo (acima de 180), o teto dos testes

aceitos para admissão em Giga Society é cerca de 160. Para o Mega e o Titan, ele estima que o teto seja próximo a 170. No caso do Sigma Test, estima que o teto pode chegar em 180.

Em meados de 2000, foi publicado o Sigma Test, o primeiro teste com questões baseadas na resolução de problemas complexos do mundo real, problemas pretensamente capazes de medir habilidades como as de Da Vinci, Pascal, Galileu e Arquimedes. São questões envolvendo criatividade, engenhosidade e pensamento profundo em diferentes níveis, e este teste tem sido amplamente elogiado pelos principais expoentes das comunidades de alto QI, cujos depoimentos estão listados nesta página:

http://www.sigmasociety.com/sigma_teste/sigma_opiniao.asp

É possível que este seja o primeiro instrumento psicométrico capaz de medir corretamente a capacidade de produção intelectual em níveis de QI acima de 170, podendo chegar até 180 ou 190. Desde então o Sigma Test foi traduzido para 13 outros idiomas e aplicado em centenas de pessoas com escores muito altos nos testes existentes até então (média em torno de 150), inclusive membros de Giga Society (QI acima de 196), Pars (QI > 180), Olympiq (QI > 180), Pi Society (QI > 176) etc. Em outubro de 2003, após ter sido aplicado em mais de 100 pessoas, foi publicada a terceira norma do Sigma Test, introduzindo diversos novos conceitos e novas ferramentas psicométricas. Embora uma amostra de 100 pessoas possa ser pequena para testes coletivos ou mesmo para testes clínicos individuais, é bastante razoável para testes discursivos e sem limite de tempo, porque praticamente não há o fator sorte/azar interferindo na precisão da norma, de modo que se consegue uma qualidade comparável a uma amostra de 5.000 pessoas ou mais. Além disso, algumas das inovações introduzidas neste procedimento de normatização possibilitam alcançar acurácia maior do que qualquer método usado até então. Além dos aprimoramentos no que diz respeito ao tratamento estatístico dos dados, também houve vários aprimoramentos conceituais, entre os quais a idéia de "QI de potencial" é a mais importante, por possibilitar a obtenção de informações que de outro modo não seriam possíveis. Pelos conceitos e métodos anteriores, não havia como estabelecer proporções entre QIs. Por exemplo: uma pessoa tem QI 100 e outra tem QI 50; quantas vezes a pessoa com QI 100 é mais inteligente do que a outra com QI 50? Não se pode responder a esta pergunta simplesmente dividindo um QI pelo outro. Além disso, para que o cálculo seja possível e faça algum sentido, é necessário que os procedimentos para obtenção destes escores tenham sido apropriados. Nenhum teste anterior ao Sigma Test provê subsídios suficientes para responder a esta pergunta. No caso do Sigma Test, é o primeiro cujos escores são obtidos de tal maneira que se torna possível responder a esta pergunta de modo preciso e cientificamente bem fundamentado. A descrição detalhada do método pode ser encontrada nos documentos relativos às normas de 2003, 2004 e 2005 do Sigma Test, disponíveis on-line em http://www.sigmasociety.com/sigma_teste/sigma_sigma_teste.asp.

Outro importante avanço no que diz respeito ao conteúdo dos testes ocorreu com a publicação do Sigma Test VI, seguramente o mais difícil que existe e o único que pode determinar corretamente QIs acima de 190 e talvez acima de 200. Porém a norma deste teste ainda é muito imprecisa, baseada em amostra muito pequena, devido à extrema dificuldade dos itens, de modo que mais de 99,9% das pessoas não conseguem acertar nenhuma das questões, e assim acaba sendo difícil reunir dados amostrais para tratamento estatístico do teste (pois escores 0 não têm muita utilidade para o tratamento estatístico).

Além dos avanços conceituais, técnicos e estruturais resultantes da criação do Sigma Test e sua norma, bem como os avanços conceituais do Sigma Test VI, também foi proposta uma nova classificação de níveis mentais com base em critérios objetivos. Até então, os estratos eram arbitrariamente definidos de 10 em 10 pontos, algumas vezes de 15 em 15, sem que houvesse mais do que um palpite na determinação do tamanho desses intervalos. O intervalo

de inteligência “normal”, por exemplo, era definido como entre 90 e 110, ou entre 85 e 115, sem que houvesse nenhuma razão especial para a escolha deste intervalo. Em contraste a isso, a classificação proposta a seguir é muito mais criteriosa:

Classificação adotada em Sigma Society com base em Análise Fatorial Hierárquica (2005)

QI acima de 200:	Imensurável
185-200:	Gênio Universal
176-185:	Gênio altamente criativo (escassez de dados amostrais)
164-176:	Gênio criativo
151-164:	Gênio
142-151:	Muito talentoso
135-142:	Superdotado ou talentoso (portador de necessidades especiais / altas habilidades)
127-135:	Inteligência muito acima da média
116-127:	Inteligência acima da média
82-116:	Inteligência média ou normal (escassez de dados amostrais)
QI abaixo de 82:	(Ausência de dados amostrais para estabelecer uma classificação)

Agora temos a primeira classificação objetivamente determinada com base no grau de similaridade entre os sujeitos de cada grupo, de modo que pares de pessoas com QIs situados no mesmo estrato são mais semelhantes entre si (em comportamento e pensamento) do que se comparadas a pessoas de outros estratos, ou seja: duas pessoas, uma com QI 145 e outra com QI 150 estão (pela classificação indicada acima) no mesmo grupo, então elas são mais semelhantes entre si do que duas pessoas com QIs 140 e 145, que pertencem a grupos diferentes, ou duas pessoas com 150 e 155, que também pertencem a grupos diferentes. Embora nos três casos a diferença seja de 5 pontos e implique mesma proporção de potencial, os componentes cognitivos presentes e manifestos em diferentes níveis são determinantes no processo de agrupamento. Essa estratificação objetiva é possível mediante o uso de Análise Fatorial (nesse caso, foi usada Análise Fatorial Hierárquica com ligação Ward e distâncias de Minkowski). Análise Fatorial é uma das ferramentas estatísticas usadas na classificação e taxonomia de plantas e animais, por exemplo, e possibilita agrupar animais em espécies, gêneros, filós, famílias etc., deixando os animais mais semelhantes no mesmo grupo, em vez de usar critérios arbitrários como por exemplo: “até 15kg é gato, acima de 16kg é cachorro”. Portanto esta nova classificação representa mais um importante avanço para a Psicometria, tornando as classificações mais científicas, em vez de meros palpites, como estava sendo feito desde os tempos de Terman. Claro que pelo fato de amostra atual do Sigma Test ser pequena, esta classificação ainda precisa ser aprimorada, conforme a amostra for ampliada.

Em http://www.sigmasociety.com/artigos/afh_st.pdf pode-se encontrar uma Análise Fatorial dos itens do Sigma Test. Isso é diferente da Análise Fatorial dos sujeitos examinados com o Sigma Test. A classificação acima é baseada na Análise Fatorial dos sujeitos. No caso do artigo citado, os itens é que são estratificados por similaridade, e é muito recomendável ler o artigo para perceber melhor a vantagem de adotar a estratificação proposta acima, em lugar das classificações antigas. Um estudo mais completo, incluindo também com classificações para QIs abaixo de 82, será publicada em breve com base numa combinação de escores no Sigma Test e no TIG-NV.

O que se espera no futuro é que os testes sejam culturalmente mais inclusivos que os atuais, além de atingirem tetos reais mais elevados e tenham normas mais acuradas. Atualmente existem testes que são aplicados em gorilas e chimpanzés, e há experimentos que permitem estimar os QIs de golfinhos e baleias, além de estimativas de QIs de cães, porcos, gatos e outros animais. Mas ainda são instrumentos fortemente influenciados pela cultura humana e o que se pretende no futuro é que seja possível medir a inteligência de animais com rigor semelhante ao que se consegue medir a inteligência humana. Também existem, desde a

década de 1960, programas de computador capazes de resolver testes de inteligência baseados em séries numéricas, séries de figuras e outras formas de reconhecimento de padrões. Contudo ainda não existe nenhum programa para ler e interpretar enunciados, de forma realmente inteligente. Existem programas de computador com QI em torno de 205 para o jogo de Xadrez e acima de 220 para o jogo de Damas, além de programas QI acima de 300 para séries numéricas, programas com QI acima de 140 para séries de figuras e demonstração de teoremas, e num futuro próximo devem surgir não apenas programas capazes de alcançar elevados escores em testes mais variados como também devem ser criados testes voltados especialmente para avaliar a inteligência de máquinas. Esperamos que num futuro próximo existam testes capazes de medir inteligência de humanos, outros animais de nosso planeta, outros animais de outros planetas, máquinas inteligentes e até mesmo a inteligência de plantas e outras estruturas que apresentem comportamento inteligente. Esperamos também que além dos testes baseados em questionários, seja possível efetuar medidas de forma mais direta, por meio de interfaces eletrônicas conectando o cérebro a computadores, e medindo propriedades neurológicas, em vez de medir desempenho para responder a perguntas. Além disso, esperamos que surjam algumas mudanças imprevistas e revolucionárias, cujos conhecimentos atuais não nos permitem antever como serão, mas que terão grande impacto.

Porém o mais importante não é como serão os testes no futuro, mas sim a importância que terão para a humanidade, porque à medida que se tornam mais eficientes, mais confiáveis, mais precisos, ganham mais credibilidade para serem usados como instrumentos de seleção profissional, recrutamento militar, avaliação acadêmica, diagnósticos neuropsicológicos etc. Os testes de QI já são amplamente usados para todas estas finalidades, sobretudo pelas empresas mais bem sucedidas e dirigidas por grandes visionários, como Microsoft e IBM, e tudo leva a crer que os testes se disseminarão cada vez mais, tão logo mais pessoas se dêem conta de sua grande eficiência em comparação às avaliações subjetivas e altamente duvidosas que algumas empresas, instituições de ensino e clínicas continuam a usar. Atualmente ainda existem muitos preconceitos impostos por educadores e psicólogos leigos em Psicometria, que por desconhecerem as ferramentas lógicas e matemáticas que fundamentam os testes, preferem não usar estes testes a ter que aprender o que é necessário para compreendê-los. Mas esta mentalidade medieval não pode sobreviver por muito tempo, porque os fatos falam mais alto e se impõe, já que as empresas que usam testes como instrumentos de seleção se sobressaem em comparação àquelas que usam técnicas obsoletas e totalmente inócuas de dinâmicas de grupo, testes de personalidade e outros sistemas subjetivos de avaliação, e quanto melhores são os testes adotados pela empresa, tanto mais garantias de que os funcionários selecionados são de fato os mais competentes. O mesmo que se aplica no caso das empresas também se aplica no caso de instituições de ensino e qualquer que seja a finalidade que envolva avaliação cognitiva: o uso de testes objetivos é muito superior ao uso de sistemas idiossincráticos de avaliação. E dia virá em que os governantes das cidades, estados, países, continentes, planetas e galáxias serão decididos pela aplicação de testes, em vez de eleições, além de haver contratos em que os candidatos assumam o compromisso por escrito de cumprir o que tiver sido prometido em seus projetos. Isso resultará em menos desperdício de recursos com campanhas, já que não será necessário seduzir eleitores. Resultará na seleção de candidatos mais capacitados, já que os vencedores serão aqueles que demonstrarem mais competência e apresentarem melhores projetos, em vez de serem eleitos os que aparentam ser o que não são, e o fato de haver contratos exigindo o cumprimento dos programas propostos, sob pena de prisão, forçará os eleitos a manterem a palavra que deram durante a etapa em que eram candidatos. Isso já havia sido proposto por Platão, há quase 2.500 anos, e assim que alguém com o grau de visão de Platão ou Aristóteles ou similar chegar ao poder, as mudanças necessárias serão implementadas, e “será colocada ordem na casa”.

Leia mais sobre a história dos testes em <http://www.sigmasociety.com/artigos/historia.pdf>

Mitos e fatos sobre inteligência e testes de QI:

Pessoas inteligentes são fisicamente mais frágeis.

Mito. Pessoas com QI elevado geralmente são mais fortes, mais saudáveis e mais ágeis do que média da população.

Pessoas com alto QI sempre são bons alunos.

Mito. As tarefas escolares são estimulantes para QIs até 125 ou 130. Acima deste nível, as tarefas se tornam desprovidas de interesse e isso pode causar queda no rendimento escolar. Estudos realizados por Leta Stollingsworth, desde 1970, mostram um alarmante índice de evasão escolar entre crianças e jovens com QI acima de 160 e principalmente acima de 180.

A inteligência é um traço genético transmitido dos pais aos filhos.

Parcialmente certo. A inteligência fluida (pensamento lógico, criatividade, aptidões cognitivas diversas) é predominantemente hereditária, enquanto a inteligência cristalizada (cultura, fluência lingüística etc.) é mais fortemente influenciada pelo meio em que a pessoa vive, pelos estímulos que recebe etc.

Escores em testes de QI são fiéis representações da capacidade intelectual das pessoas.

Mito. Se os testes forem bem construídos, bem padronizados e bem aplicados, o escore pode ser razoavelmente fidedigno. Mesmo os bons testes estão sujeitos a flutuações relacionadas a oscilações do humor das pessoas examinadas. O assunto é complexo e extenso. O fato é que os testes, embora não sejam exatos (incerteza em torno de 5 pontos), são os melhores instrumentos que existem para medir o desempenho intelectual até cerca de 130. Para QIs acima de 130 até 170 existem os hard tests. Para QIs acima de 170 é muito difícil avaliar corretamente o desempenho e geralmente recorre-se a estimativas baseadas na produção intelectual. Usar um teste típico de clínica para medir o QI de Einstein, por exemplo, seria como tentar medir o talento literário de Shakespeare com base na velocidade com que ele soletra palavras, isto é, seria tentar medir um conjunto complexo de habilidades requintadas por meio de um sistema primário de habilidades simplórias. O fato de Shakespeare ser um dos grandes gênios da Literatura não significa que ele seria capaz de soletrar mais rapidamente do que uma criança mediana, assim como o fato de Einstein ser um dos maiores gênios da Física não significa que ele poderia resolver, num prazo reduzido, alguns problemas simples de Aritmética elementar.

Os escores em testes não variam depois dos 16 anos.

Mito. Testes como o WAIS e o Raven foram aplicados em milhões de pessoas de vários países e de diferentes faixas etárias, e ficou constatado que o QI aumenta rapidamente até cerca de 16 anos, continua aumentando mais lentamente até 25 anos, estabiliza até 40 anos e então começa a decair lentamente até cerca de 70 anos. Não há dados suficientes para dizer com segurança, mas parece que aos 70 anos se estabiliza novamente (não há queda substancial dos 70 até os 89 anos).

Testes de QI são bons indicativos da capacidade intelectual.

Fato para QIs até 130. Nos casos de QIs mais elevados, os hard tests continuam avaliando corretamente até cerca de 160 ou 170. Acima do nível de 1 em 1.000.000, existem poucos testes que conseguem medir corretamente. O Sigma Test e o Power Test são dois exemplos que conseguem medir até um pouco acima de 170 ou mesmo acima de 180.

Existem diferenças étnicas, sexuais, regionais de QI.

Fato. O QI médio do brasileiro é cerca de 87. O QI médio mundial é 100. Não existem diferenças entre o QI global de homens e mulheres, mas em pontos específicos as diferenças

são marcantes. Por exemplo: mulheres se sobressaem em coordenação motora fina, em linguagem, em encenação e em sensibilidade para perceber sentimentos e intenções. Os homens se sobressaem em atividades mecânicas, em Lógica e Matemática e motricidade robusta. Entre grupos étnicos também há diferenças que são geralmente distorcidas e esse é um tema complexo que não poderia ser abordado em poucos minutos. O fato é que os desempenhos de judeus são, em média, melhores do que os de outros povos.

Um alto QI é garantia de sucesso profissional e acadêmico.

Mito. O QI é um dos ingredientes necessários e um dos mais importantes, porém existem outros fatores dos quais também não se pode prescindir, como a disciplina, motivação, habilidade para se relacionar, perseverança e até mesmo a sorte.

Toda criança precoce será um gênio quando adulto.

Mito. A precocidade no sentido de atingir em idade mais tenra o grau de desenvolvimento observado em crianças mais velhas não implica necessariamente a presença de habilidades especiais que não estão presentes nas outras crianças de mesma idade. Um gênio não é alguém que aprende matemática ou música mais rápido. É alguém que produz matemática nova e compõe novas e harmoniosas melodias. Por isso aprender mais cedo ou mais rápido nem sempre é indicativo de genialidade. Aliás, na grande maioria das vezes as crianças precoces não se tornam adultos geniais. Um estudo realizado por Terman, ao longo de várias décadas, com 1528 crianças com QIs entre 135 e 200+, revelou que praticamente nenhuma delas se tornou um gênio, na acepção mais rigorosa da palavra.

Todo adulto gênio foi precoce quando era criança.

Mito. Einstein, por exemplo, não foi uma criança precoce. A genialidade está mais intimamente relacionada à habilidades especiais de criar do que à precocidade na manifestação de habilidades triviais. Mas muitas vezes os gênios de fato foram crianças precoces. Gauss, Pascal, Capablanca, Beethoven e a maioria dos gênios de fato foram crianças muito precoces. É importante perceber a diferença entre esta questão e a anterior, pois a maioria das crianças precoces não se tornam gênios, mas a maioria dos gênios foram crianças precoces. Isso acontece porque a precocidade é um dos mais importantes ingredientes para a genialidade, e na ausência deste ingrediente, raramente se forma um gênio, mas com apenas este ingrediente também raramente se forma um gênio.

Algumas informações básicas e curiosidades sobre o tema:

Existem diversas escalas de inteligência, entre as quais as principais são as de Wechsler, Stanford-Binet e Cattell (além de SAT, GRE, T etc.). Um escore 180 na escala Cattell, por exemplo, representa menor desempenho do que 160 pela escala Wechsler. Por isso é importante conhecer não apenas o escore, mas também a escala em que aquele escore foi obtido. Para converter escores de uma escala A para outra escala B, basta subtrair o valor médio da escala A, dividir pelo desvio-padrão da escala A, multiplicar pelo desvio-padrão da escala B e somar o valor médio da escala B. Por exemplo: João teve escore 148 num teste com média 100 e desvio-padrão 24 e Pedro teve escore 136 num teste com média 100 e desvio-padrão 15. Qual deles teve melhor desempenho? Para responder a isso, tanto faz converter o escore de João para a escala do teste usado em Pedro quanto converter o escore de Pedro para a escala do teste usado em João. Vamos converter o escore de João para a escala do teste usado em Pedro: $148 - 100 = 48$. Agora fazemos $48/24 = 2$. Depois fazemos $2 \times 15 = 30$. E, por fim, fazemos $30 + 100 = 130$. Então o escore 136 de Pedro no teste que foi aplicado em Pedro é intrinsecamente mais alto do que o escore 148 de João no teste que foi aplicado em João, pois o escore 148 de João no teste que lhe foi aplicado equivale 130 no teste que foi aplicado em Pedro. Para quem está habituado a converter escalas de temperatura (Celsius, Kelvin, Fahrenheit, Réaumur e Rankine), o procedimento é idêntico.

Existem centenas de testes de inteligências, alguns enfatizando o desempenho lingüístico, outros enfatizando o desempenho em Aritmética, outros enfatizando o desempenho em pensamento abstrato, outros enfatizando a cultura etc. O Raven Progressive Matrices, por exemplo, mede diretamente apenas um conjunto pequeno de faculdades cognitivas relacionadas ao pensamento abstrato, mas este pequeno conjunto de traços cognitivos está fartamente saturado do fator “g”, que é o fator comum a um vasto conjunto de habilidades. A vantagem deste tipo de teste é conseguir avaliar a “inteligência geral” com um conjunto relativamente pequeno de questões e não ser muito afetado por fatores culturais. Por outro, o WAIS mede várias habilidades diferentes em seus 14 subtestes, e isso possibilita fazer um diagnóstico estratificado sobre as habilidades de uma pessoa para diferentes atividades, enquanto o Raven mede apenas o conjunto de todas as habilidades somadas.

Devido a estas diferenças, uma pessoa pode pontuar muito bem num determinado teste (lingüístico, por exemplo) e não tão bem em outro, ou até muito mal em outro (séries numéricas, por exemplo). Contudo existe um fator geral, o “g” de Spearman, que é uma medida da capacidade intelectual geral, para praticamente todas as áreas intelectuais. O fator “g” é o construto primário da inteligência, usado tanto para Matemática quanto para Comunicação, para atividades mecânicas etc. Uma pessoa com alto escore “g” pode se sobressair em várias áreas diferentes. Além do fator “g” existem muitos fatores secundários que possibilitam bom desempenho em áreas específicas. Assim uma pessoa com saturação no fator secundário para Matemática, mas não para Literatura, poderá se sobressair apenas em Matemática (e outras áreas ligadas à Matemática).

O termo QE, introduzido por Goleman, para representar a “Inteligência Emocional”, tem sido interpretado incorretamente por muitos. O próprio Goleman, em seus artigos, demonstra compreender claramente a importância do QI e criou o QE para ser um complemento (não um substituto) aos testes de QI. Porém muitas pessoas, talvez pelo fato de não terem bons escores em testes de QI, distorcem este fato e tentam divulgar a informação de que o QI deixou de ser uma medida adequada para prognosticar desempenho profissional e acadêmico, e agora o QE é que tem esta função. Isso está errado e contradiz o que é professado pelo próprio Goleman. O fato é que o QI combinado ao QE são dois importantes indicadores de desempenho, sendo o QI mais fidedigno e mais preditivo, por ser objetivo e não depender tanto da capacidade do examinador na interpretação idiossincrática das respostas.

Dúvidas frequentes:

1 – Eu fiz um teste on-line (ou numa revista) e meu QI foi 134. O que isso significa?

Resp.: não é recomendável que faça testes on-line, a menos que estes testes apresentem informações precisas sobre a padronização e outros dados que possam dar algum respaldo científico a estes instrumentos. Quase todos os testes on-line não são padronizados, e os poucos que passaram por algum processo de padronização geralmente apresentam várias falhas neste processo. Por estes motivos é preferível que vá a uma clínica e seja avaliado por um profissional capacitado. Mesmo assim você não estará a salvo de receber diagnósticos e orientações incorretos, porque no Brasil mais de 99% dos psicólogos não conhecem o mínimo necessário para aplicar e interpretar testes. Esse problema preocupa os dirigentes do Conselho Federal de Psicologia, que há alguns anos estão empreendendo esforços no sentido de tirar de circulação os testes de baixa qualidade usados indiscriminadamente em empresas, colégios e até mesmo em clínicas. Também estão planejando cursos e certificações para credenciar os psicólogos e outros profissionais verdadeiramente qualificados para a correta aplicação de testes. Em outras palavras: seu escore 134 pode significar coisa, assim como pode não significar nada.

Se após ser examinado por um profissional capacitado, você obtiver um escore acima de 120 e principalmente se pontuar acima de 130, é recomendável que faça também um hard test, no qual poderá ter uma avaliação mais rigorosa de sua capacidade em níveis mais altos, já que os testes de clínica não são capazes de medir corretamente desempenhos acima de 130 ou 135. Alguns testes clínicos disponíveis no Brasil que recomendamos são estes: TIG-NV, TNVRA, G-38, G-36, WAIS-III, WISC-III (para crianças), BPR-5, DAT, Raven. Alguns hard tests que recomendamos são estes: Sigma Test (disponível em 14 idiomas, inclusive português), Sigma Test VI (8 idiomas, inclusive português, porém não disponível no momento), Teste de Desempenho Mental (disponível em português e inglês), Hepta Test (disponível em português). Hard tests em outros idiomas: Sigma Associations Test, Sigma Analogy Test, Power Test, Eureka Test, G-Test, Archimedes Test, Titan Test, Mega Test, Logima Strictica, .

2 – Eu fiz um teste vocacional (ou de personalidade, ou de aptidão ou de qualquer coisa ligada a sexo, relacionamento, perfil profissional etc.) on-line (ou numa revista) e o resultado foi O que isso significa?

Resp.: repete-se o comentário anterior. Qualquer classificação ou diagnóstico feito nestas condições tem todas as chances de ser incorreto. Os testes de personalidade (e similares) apresentam ainda mais erros, e são erros mais graves, do que os testes de inteligência, porque além de as avaliações serem muito subjetivas, ainda por cima as classificações e nomeações costumam ser totalmente esotéricas, baseadas unicamente no que o autor do teste acredita que sejam as características que ele gostaria de medir, sendo que muitas vezes as questões são totalmente inadequadas e medem características diferentes daquelas que o autor pretendia. Esta crítica não aplica apenas aos testes on-line e publicados em revistas femininas, mas é extensiva também à esmagadora maioria dos testes usados em clínicas e universidades. São raríssimos os testes de personalidade (e assemelhados) que apresentam o mínimo de fundamentação científica.

3 – Eu fiz um teste vocacional (ou de personalidade, ou de aptidão ou de qualquer coisa ligada a sexo, relacionamento, perfil profissional etc.) on-line (ou numa revista) e o resultado foi O que isso significa?

Resp.: repete-se o comentário anterior. Qualquer classificação ou diagnóstico feito nestas condições tem todas as chances de ser incorreto.

Porém, assim como qualquer outros instrumento, os testes tanto podem ser usados para propósitos benéficos e científicos como podem ser usados para

4 – Onde posso ser examinado(a) com um bom teste e por um bom profissional?

Resp.: Para avaliar QIs abaixo de 120, a melhor opção é ir a uma boa clínica psicológica, porém mais de 99% dos psicólogos brasileiros não estão familiarizados com os conceitos elementares necessários para aplicar e interpretar os escores de testes, por isso é importante pesquisar as referências sobre o examinador no âmbito específico de construção e aplicação de testes. Alguns psicólogos gozam de boa reputação como psicanalistas ou behavioristas, mas não conhecem o básico sobre testes. Como a aplicação de testes não é tarefa muito simples, se isso for feito por um profissional que não seja especializado no assunto e não conheça as limitações dos instrumentos em uso, corre-se o risco de receber um diagnóstico equivocado, além de avaliação e orientação inadequadas. Este é um problema particularmente grave se levarmos em consideração que quase todos os testes disponíveis para uso clínico apresentam dezenas de falhas.

Para avaliar QIs acima de 120 e abaixo de 180, os hard tests são mais apropriados. Os hard tests geralmente apresentam conteúdo mais requintado do que os testes usados em clínicas, porém os testes de clínica geralmente passaram por processos mais rigorosos de padronização, porque muitos autores de hard tests são talentosos para elaborar questões originais e interessantes, mas não possuem os conhecimentos necessários para realizar uma padronização formal. O Sigma Test tem a virtude de combinar o melhor conteúdo semântico com o melhor sistema de padronização que existe (único que possibilita gerar escores em escala de proporção). O Sigma Test VI possui conteúdo semântico no mesmo nível do Sigma Test (e mais difícil), porém a padronização se baseia numa amostra muito pequena e fortemente influenciada pela ego-seleção.

Para obter informações sobre testes usados em clínicas, recomendamos que entre em contato com a Casa do Psicólogo ou com o Conselho Federal de Psicologia. Nestas instituições poderá receber a orientação necessária sobre onde encontrar clínicas e psicólogos em sua cidade.

Para fazer hard tests, nós disponibilizamos uma lista com os instrumentos que são aceitos para admissão em Sigma e em outras associações similares. Visite nossas seções “Sigma Test”, “Critérios para admissão” e “Links”.

5 – O que são sociedades de alto QI e o que é necessário para se associar a elas?

Resp.: existem atualmente dezenas de sociedades de alto QI, com diferentes objetivos, diferentes critérios para admissão e diferentes perfis de associados. Algumas estabelecem níveis de corte de 1 em 20 (a cada 20 candidatos, 1 é aprovado), outras estabelecem níveis teóricos de corte em milhões ou até bilhões, o que é geralmente fictício. Sigma VI, por exemplo, tem corte teórico de 1 em 1 bilhão (entre 6,4 bilhões de pessoas que existem atualmente, apenas 6 seriam aprovadas), mas na prática o corte é cerca de 1 em 1 milhão ou algo assim. Giga Society também tem corte teórico de 1 em 1 bilhão, mas na prática o corte é cerca de 1 em 100 mil. A legitimidade do corte depende muito dos testes aceitos para admissão. Se os testes forem normatizados adequadamente e tiverem questões com nível apropriado de dificuldade, o corte acaba sendo respeitado, caso contrário, as distorções podem ser grandes. Há muitas sociedades que aceitam testes de clínica para cortes acima de 1 em 100, o que constitui um erro grave, porque o limite de dificuldade dos itens dos testes de clínica não ultrapassa o nível de 1 em 100 e qualquer medida acima disso acaba representando a

velocidade para resolver problemas elementares, e isso não reflete necessariamente a habilidade para resolver problemas complexos.

De modo geral, as comunidades de alto QI atuam muito abaixo do que poderiam, no sentido de canalizar o potencial de seus membros para prestar serviços à sociedade. Em parte isso se deve à falta de interesse do governo, no sentido de explorar o potencial latente destes núcleos de intelectuais, e em parte a culpa é dos próprios membros e dirigentes, por não se mobilizarem para a execução de projetos concretos. Uma parceria entre o estado e estes grupos ou entre empresas e estes grupos traria muitos benefícios à população, e provavelmente, num futuro próximo, vários empreendimentos deste gênero entrarão em vigor.

6 – Existe alguma relação entre QI e formação acadêmica?

Resp.: até certo ponto, sim, pois as habilidades necessárias para ter bons escores em testes são semelhantes àquelas necessárias ao bom desempenho acadêmico. As disparidades começam a surgir quando as exigências da vida acadêmica ou dos testes se tornam muito diferentes. Por exemplo: ganhadores de prêmio Nobel ou de medalha Fields apresentam um nível muito acima do que é medido pelos testes clínicos. Por outro lado, os hard tests exigem habilidades muito mais requintadas do que aquelas exigidas de um Ph.D. típico. Como resultado, para pessoas com QI entre 70 e 140 há forte correlação entre QI e graduação acadêmica. Para QIs acima de 150 e principalmente acima de 160, ocorre uma destoaância entre o baixo nível das exigências da vida acadêmica em comparação ao alto nível de produção intelectual destes sujeitos, e como resultado observa-se uma índice muito alto de abandono dos estudos entre crianças e jovens com QI acima de 160. Estudos sobre grupos de crianças com QI acima de 160 até 180 podem ser encontrados no site do instituto de Leta Stetter Hollingworth: <http://www.hollingworth.org>. De modo geral, os QIs médios dos diferentes níveis acadêmicos são aproximadamente estes:

Graduação	QI
Ensino fundamental	104
Ensino médio	108
Ensino superior	111
Pós-graduação	115
Mestrado	120
Doutorado	125
Pós doutorado	130
Professor Emérito	140
Prêmios nacionais	150
Prêmio Nobel	160
Medalha Fields	170

Estes escores variam em diferentes países, diferentes épocas e diferentes instituições. Também há diferenças de QI entre profissões e um estudo detalhado sobre o tema pode ser encontrado em <http://www.ssc.wisc.edu/cde/cdewp/98-07.pdf>.

7 – É verdade que a humanidade está se tornando mais inteligente com o passar do tempo?

Resp.: O fato é que o QI médio de populações não-seletas atuais é mais alto do que o QI médio de populações equivalentes de 80 anos atrás. Este efeito “inflacionário” foi descoberto e publicado em 1984, no *Psychological Bulletin*, por James R. Flynn, baseado em pesquisas realizadas entre 1932 e 1978. O efeito já havia sido previsto décadas antes, por vários pesquisadores, porém não se sabia se o QI médio estava aumentando ou diminuindo. A

maioria acreditava que a queda seria mais provável, já que as populações nos países mais desenvolvidos estão crescendo mais lentamente do que nos países mais pobres, além disso, pessoas mais instruídas e com melhor alimentação geram menos descendentes do que as menos instruídas, com menos acesso à cultura e algumas vezes com deficiências nutricionais. Nesse contexto, seria esperado que, em média, a parte menos favorecida da população se tornasse cada vez mais numerosa e assim causasse uma queda no QI médio mundial. Mas o que se tem observado é que, embora essa hipótese seja correta, o fator que tem mais peso é que a qualidade de vida de todos está, em média, melhor do que antes. Isso é mais fácil de perceber se compararmos o padrão de vida de um cidadão urbano médio de nossos dias ao padrão de vida de um nobre europeu medieval. O nobre medieval perdia os dentes precocemente, vivia com menos higiene, não tinha nenhum recurso tecnológico similar aos atuais, a Medicina não era muito diferente dos curandeiros tribais modernos, não havia água tratada, nem eletrodomésticos etc. Por isso a expectativa de vida na Idade Média era cerca de 30 anos, enquanto hoje é cerca de 70 anos, e mesmo nos países mais pobres do mundo, como Etiópia e Camboja, vive-se hoje, em média, até os 45 anos. Portanto, embora as populações mais pobres cresçam mais rapidamente, a tecnologia avança suficientemente rápido para proporcionar a estas populações melhores recursos do que os que estavam disponíveis aos mais abastados de alguns anos atrás. Os telefones celulares, por exemplo, hoje são usados por todos, mas há 20 anos não existiam e há 10 anos eram artigo de luxo. Além desse fator de caráter fisiológico (melhor alimentação, mais higiene, melhor medicina etc.), há diversos fatores culturais envolvidos. Um deles é que os computadores contribuem para disseminar rapidamente a cultura, os videogames são excelentes exercícios mentais, Antes da invenção da imprensa, os livros eram copiados manualmente e tinham preço altíssimo. Depois caíram vertiginosamente de preço, mas continuavam relativamente caros devido aos preços da matéria prima, que com o passar do tempo foi caindo de preço. Hoje os livros digitais tem custo baixíssimo. No caso dos testes, quando foram criados não havia questões típicas de testes em parte alguma, exceto nos próprios testes. Algum tempo depois, começaram a ser comercializados livros e revistas com questões semelhantes às dos testes de Cattell e Binet, e estes livros foram se tornando cada vez mais numerosos e mais populares, de modo que mais pessoas passaram a conhecer questões típicas de testes antes de serem examinadas, assim quando eram testadas pela primeira vez, as questões já lhe eram mais familiares do que às pessoas que viveram alguns anos antes. Agora, com testes gratuitos disponíveis on-line, muito mais pessoas têm acesso a questões típicas de testes e se preparam (intencionalmente ou não) melhor do que as que pessoas que eram testadas há 50 anos. Por todos estes motivos, além de vários outros, o QI da população mundial está subindo cerca de 3 pontos por década, variando de acordo com o país, a idade, o nível de QI etc. Por exemplo: Se o QI médio na China era 103 e agora é 109 não significa que o topo 1% da China com QI médio 140 agora tenha subido para 146, pois nos níveis mais altos o crescimento pode ter sido muito mais lento ou muito mais rápido, e o estudo de Flynn menciona apenas variações na média (seria preciso fazer estudos mais completos). Além disso, em alguns países o QI subiu mais de 30 pontos em 50 anos. Em outros países subiu apenas 5 pontos no mesmo período.

Por todos estes motivos, embora saibamos que o QI médio da população, medido por testes concebidos no início do século XX, aumentou cerca de 30 pontos nos últimos 100 anos, não podemos dizer que este aumento no QI indique um aumento na inteligência propriamente dita. O mais provável é que parte do efeito seja uma real representação da evolução da inteligência, mas a maior parte do efeito deve ter sido causada pela simples popularização das questões que integram os testes mais comuns, como Cattell, Raven, Wechsler e Binet. Mais informações sobre o efeito Flynn podem ser encontradas aqui:

<http://www.indiana.edu/~intell/flynneffect.shtml>

8 – Os testes de QI realmente medem alguma coisa?

Resp.: Conforme já foi dito, existem diferentes tipos de testes, destinados a medir diferentes características. Todos medem alguma coisa. O que precisa ser questionado é se eles medem fato o que propõem medir (ou se medem algo diferente do que é proposto) e o quão bem eles medem. Quando medimos a altura de uma pessoa, por exemplo, sabemos com precisão o significado de “altura”, dispomos de instrumentos precisos para medir comprimentos e a medição é direta. Quando tentamos medir a inteligência, surgem várias dificuldades. A primeira é que há diferentes definições para o termo “inteligência”. A segunda é que os instrumentos usados não medem a totalidade da inteligência, mas apenas algumas partes que julgamos serem representativas da totalidade. A terceira é que a medição não é feita diretamente, mas com base numa performance ou num conjunto de tarefas ou realizações. Isso torna a medida da inteligência menos precisa do que a medida da altura, além disso, corre-se o risco de que aquilo que se está medindo não seja propriamente a inteligência. Apesar disso, existem procedimentos estatísticos que possibilitam superar todas estas dificuldades e obter escores que sejam representativos da inteligência. Quando se fazem pesquisas pré-eleitorais, por exemplo, entrevistam-se alguns milhares de pessoas de diversas classes sociais e culturais, de diferentes cidades, e com isso se consegue projetar como serão as eleições com toda a população, com quase 100 milhões de eleitores, com margem de erro em torno de 1% a 5%, ou seja, com mais de 95% de exatidão. O mesmo acontece ao tentar medir a inteligência total com base nas respostas a um pequeno conjunto selecionado de questões ou na execução de determinadas tarefas. O resultado acaba sendo muito bom, com erro menor do que 5%. Portanto, desde que o conteúdo dos itens for adequado às características que se pretende medir, desde que eles sejam suficientemente abrangentes e diversificados de modo a cobrir toda a amplitude das tais características, desde que os procedimentos estatísticos na normatização sejam meticulosos e rigorosos, então o que o teste mede é confiável.

Uma crítica muito freqüente e muito injusta feita aos testes de QI é que eles não medem talento musical, nem vocação para artes cênicas, nem habilidade para futebol etc. Tal crítica equivale a dizer que um alicate é uma ferramenta ruim porque ele não serve para cortar papel, nem para abrir latas, nem para costurar. De fato ele não serve para nada disso porque ele não foi construído com estes propósitos. Ele é bom, útil e eficiente quando usado corretamente, visando o objetivo para o qual foi criado e por pessoas capacitadas para usá-lo. Do mesmo modo, quando uma pessoa capacitada para usar determinado teste, conhecedora das aplicações e das limitações deste teste, usa este teste de maneira correta, o resultado será útil e confiável. O problema é que na grande maioria dos casos, os autores dos testes falham ao tentar determinar o que seus testes são capazes de medir. Muitas vezes um teste que teoricamente deveria medir “raciocínio abstrato” (que envolve música, matemática, pintura, linguagem etc.) na verdade mede apenas um conjunto muito restrito e primário de pensamento abstrato, ligado à percepção de mudanças de padrões visuais em séries de matrizes. Isto evidentemente não é “raciocínio abstrato”. Isso é “capacidade para reconhecer e reproduzir padrões que se repetem em séries de figuras”. Outro exemplo são os testes que dizem medir “raciocínio espacial”, quando na verdade medem apenas uma fração muito pequena da vastíssima variedade de possíveis formas de pensamento espacial. Esta falha sistemática observada na nomenclatura de praticamente TODOS os testes que existem, inclusive no WAIS, Stanford-Binet etc., acaba alimentando as críticas dos que são contra o uso de testes. De fato, estas críticas são procedentes, mas em vez de simplesmente dizer: “estes testes não prestam porque não medem aquilo que propõem medir”, uma postura mais idônea e mais louvável seria dizer: “estes testes não medem o que propõem medir, mas medem tais e tais características, e estas características correlatam positivamente com estas outras etc.” É o que acontece com o Raven, por exemplo, que mede um pequeno conjunto de habilidades primárias saturadas de “g” que correlatam com um grande conjunto de habilidades mais complexas também saturadas de “g”. É como medir o pé de uma pessoa e com base no tamanho do pé determinar também o tamanho da mão, a altura, o peso etc., com base nas proporções entre pé, mão, altura e peso

da média da população. Desse modo, como há forte correlação entre tamanho do pé, da mão e altura, determinando um destes pode-se calcular com boa precisão os tamanhos de outras partes. No caso do peso, a correlação não é tão boa com altura, tamanho do pé e da mão, portanto acaba sendo menos preciso determinar o peso por este método. O mesmo acontece no caso de testes. Um escore no Raven pode servir bem para estimar o desempenho em algumas áreas de exatas, mas não em linguagem. Vejamos a seguir o que medem alguns testes:

O que mede o Sigma Test?

É um teste baseado em problemas da vida real, que exigem criatividade, engenhosidade, planejamento, organização, pensamento analítico, pensamento lógico profundo e complexo, pensamento lógico simples e superficial, sensibilidade para perceber sutilezas. Esquemáticamente podemos dizer que o Sigma Test mede as seguintes capacidades:

- Capacidade para compreender o pensamento de outras pessoas em questões profundas.
- Capacidade para compreender o pensamento de outras pessoas em questões superficiais.
- Capacidade para compreender a Natureza em questões profundas.
- Capacidade para compreender a Natureza em questões superficiais.
- Capacidade para compreender a Lógica em questões profundas.
- Capacidade para compreender a Lógica em questões superficiais.
- Capacidade para interpretar textos simples e complexos.
- Capacidade para perceber sutilezas evidentes em questões lógicas e situações da vida real.
- Capacidade para perceber sutilezas recônditas em questões lógicas e situações da vida real.
- Capacidade para usar rudimentos de Matemática elementar para resolver problemas simples do cotidiano em diferentes níveis de dificuldade.

As questões do Sigma Teste que medem a capacidade para compreender o pensamento de outras pessoas são como a 35 e 36, mas existem outras capacidades para compreender o pensamento de outras pessoas, em termos emotivos e sociológicos, que não são medidas pelo Sigma Test. Tanto na questão 35 quanto na 36, mede-se a capacidade para compreender o pensamento de outras pessoas no âmbito cognitivo.

O que mede o Sigma Test VI?

É uma versão mais difícil do Sigma Test. É baseado em problemas da vida real que exigem criatividade, engenhosidade, pensamento lógico profundo e complexo, sensibilidade para perceber sutilezas e capacidade para descobrir e inventar soluções engenhosas, elaboradas e eficientes para problemas de nível olímpico. Esquemáticamente podemos dizer que o Sigma Test VI mede as seguintes capacidades:

- Capacidade para compreender o pensamento de outras entidades conscientes.
- Capacidade para compreender a Natureza e os limites do que é possível.
- Capacidade para compreender a Lógica em questões de nível olímpico.
- Capacidade para interpretar a mente e a natureza.
- Capacidade para formular soluções pertinentes que envolvam estruturas complexas.
- Capacidade de usar rudimentos de Matemática para resolver problemas de nível olímpico.

O que medem o Stanford-Binet e o WAIS-verbal?

São testes baseados em perguntas acadêmicas triviais, exigem atenção para evitar descuidos, pensamento lógico simples, superficial e rápido. Esquemáticamente podemos dizer que medem as seguintes capacidades:

- Capacidade para ler e escrever.
- Capacidade para conhecer os significados de algumas palavras e reconhecer sinônimos.
- Capacidade para ter adquirido algumas informações variadas.
- Capacidade para ler e interpretar textos.
- Capacidade para aplicar fórmulas matemáticas simples na resolução de problemas elementares.
- Capacidade para perceber diferenças evidentes entre coisas semelhantes.
- Capacidade para perceber semelhanças evidentes entre coisas diferentes.
- O WAIS mede também a memória em curto prazo num nível muito básico.

O que medem o Cattell e o RAPM?

São testes baseados em séries de figuras, exigem atenção para evitar descuidos, pensamento lógico simples, superficial e rápido. Esquemáticamente podemos dizer que medem as seguintes capacidades:

- Capacidade para perceber padrões de mudanças em desenhos simples.
- Capacidade para perceber tendências, com base nas mudanças nos desenhos simples.
- Capacidade para deduzir mudanças futuras, com base nas tendências observadas.
- Capacidade para estabelecer analogias simples entre figuras.

Conforme podemos perceber, o conjunto do que o Sigma Test mede é mais abrangente e mais profundo do que o conjunto do que é medido pelos testes tradicionais. Além disso, as perguntas são muito mais semelhantes a situações da vida real, por isso medem muito melhor a habilidade para lidar com problemas reais. Soma-se a isso que o WAIS apresenta algumas dezenas de inexatidões no gabarito, nos enunciados, na contagem de pontos e na padronização, o Raven tem 7 itens problemáticos e pelo menos 1 item ambíguo, além de graves falhas psicométricas. O Sigma Test também não está isento de falhas, mas em comparação à maioria dos hard tests é um dos mais bem revisados, e em comparação aos supervised tests é incomparável melhor revisado. O Sigma Test VI, no entanto, contém uma falha na questão 3, que foi apontada pelo Campeão Mundial da Olimpíada da Matemática de 1980, Ph.D. por Princeton, pesquisador no IMPA e UMPA e coordenador nacional da OBM Nicolau Corção Saldanha. Enquanto a questão não for substituída, o teste permanecerá indisponível. O Sigma Test também teve questões revisadas: questão 31 (por Rodrigo de Almeida Rodrigues) e questão 32 (por Peter David Bentley, post doctoral por Oxford, obteve terceiro maior score no Sigma Test). A questão 29 também foi aprimorada por Peter David Bentley e Albert Frank (Albert é professor de Lógica em Bruxelas e referee internacional para várias revistas especializadas em Lógica e Matemática).

No caso do WAIS, embora apresente uma extensa lista com centenas de revisores, basta uma rápida revisão de alguns minutos para apontar vários erros que continuam presentes. O mesmo problema afeta o DAT, BPR-5, Raven e outros.

Convém chamar a atenção para o fato de que, embora praticamente nenhum teste seja perfeito, nada no mundo é perfeito, e estando as inexatidões compreendidas dentro de limites aceitáveis, isso não chega a invalidar estes instrumentos. O WAIS, por exemplo, apesar das falhas, é o

melhor teste individual disponível para uso clínico. Não seria recomendável retirá-lo de circulação, já que sua utilidade, mesmo que com escores ligeiramente distorcidos, é muito importante para numerosas finalidades, inclusive diagnósticos de várias patologias mentais, orientações vocacionais, diagnóstico de superdotação, de talentos específicos etc. A solução não é eliminar estes testes, mas revisá-los e aprimorá-los para que se tornem instrumentos mais fidedignos e mais acurados.

Sigma Society VI

Sigma Society VI foi inaugurada em 2002 e possivelmente constitui o grupo intelectual mais exclusivo do mundo, com cut-off teórico em 99,9999999% e provável cut-off real em 99,99985%, que representa um QI acima de 244 pela escala Cattell ou acima de 196 pela escala Stanford-Binet. O recorde anterior cabia à Mega Society, com corte teórico em 99,9999% e corte real próximo a 99,999%. Mega Society está registrada no Guinness Book de 1990. Atualmente outra sociedade também reivindica o recorde de ser a mais exclusiva do mundo. Trata-se de Giga Society, com corte teórico igual ao de Sigma VI, mas com provável corte real semelhante ao de Mega Society. Para ser admitido em Sigma VI é necessário obter escore acima de 196 no Sigma Test ou no Sigma Test VI.

Quem são os membros de Sigma VI?

ID 001: Hindenburg Melão Jr. (do Brasil)

Três recordes mundiais em atividades intelectuais, um dos quais registrado no Guinness Book, ed. 1998, p. 100-111 (mate anunciado mais longo em simultâneas de Xadrez às cegas, mate anunciado mais longo em Xadrez Epistolar e fundador da sociedade de alto QI mais exclusiva do mundo).

Autor do primeiro método do mundo para construir testes em escalas de proporção.

Autor de novidade teórica Top-10 mundial, eleita pelo júri do Sahovski Informator.

Autor de projeto de máquina da invisibilidade, construída por Susumu Tachi e usado na NASA e no MIT.

Autor de aprimoramento no método usado pela NASA e pela ESA para cálculo de paralaxes estelares.

Autor de correção na fórmula recomendada pela OMS para cálculo de IMC.

Autor de método para cálculo de fatoriais de números não-inteiros sem o uso de Cálculo.

Autor de outras inovações e aprimoramentos em Matemática, Física, Astronomia, Psicometria, Econometria, Educação e outros campos.

Autor de inovações e aprimoramentos em fórmulas e estratégias para aplicações financeiras.

Autor do Sigma Test, disponíveis em 14 idiomas e reconhecido para admissão em dezenas de associações culturais de vários países.

Membro honorário em associações culturais de vários países, inclusive Pars Society, para pessoas com QI acima de 220 (Cattell).

Atualmente aguardando análise de um possível recorde brasileiro de maior ganho na Bolsa de Ações em um único dia (50,4%).

28º colocado no simulador de investimentos da Folhainvest em dezembro, acima de 99,97% dos 102.000 participantes. Classificado acima de 99,92% em julho, acima de 99,4% em junho, acima de 97% em maio e acima de 91% em agosto.

ID 002: Petri Widsten (da Finlândia, atualmente residindo na Austrália)

Melhor performance gauged IQ e non-gauged IQ no Sigma Test.

Distinguido com prêmio de melhor tese de doutorado da Finlândia em Engenharia Química no biênio 2003-2004.

Ph.D. Summa Cum Laude em Química pela Universidade de Tecnologia de Helsinque (uma das mais bem reputadas do mundo).

Vencedor do concurso de Lógica X-Test de ISI Society.

Vice-campeão no Mensa International Contest 2001.

Membro em Pi Society, da Grécia, para pessoas com QI acima de 212 pela escala Cattell.

Autor de importantes inovações em Química, especialmente em fibras de madeira.

Fala fluentemente 9 idiomas (inclusive português).

ID 003: Peter David Bentley (do Reino Unido, atualmente residindo em Hong Kong)

Terceira melhor performance gauged IQ no Sigma Test e segunda melhor performance non-gauged IQ.

Co-fundador de Ludomind, associação anglo-belga para pessoas com QI acima de 150 e autoras de puzzles difíceis, bonitos e originais.

Post Doctoral pela Universidade de Oxford, da Inglaterra (universidade mais bem reputada do mundo, ao lado de Harvard e Moscou).

D. Phil (= Ph.D.) pela Universidade de Oxford.

Diretor da INFCON.

ID 004: Rauno Lindström (da Finlândia)

Segunda melhor performance gauged IQ no Sigma Test e terceira melhor performance non-gauged IQ.

Post Doctoral em Física pela Universidade de McMaster, do Canadá.

Ph.D. em Física pela Universidades de Turku, da Finlândia.

Ph.D. em Física pela Universidades de Joensuu, da Finlândia.

Orientador de estudantes de Ph.D. em vários países (inclusive do Brasil).

Ganhador de 2 medalhas de ouro e 5 medalhas de prata em competições esportivas (Tiro ao alvo, Dardo e Natação).

Pesquisador sobre natureza dos círculos em campos de milho (mesmos círculos do filme "Sinais", protagonizado por Mel Gibson).

ID 005: Kristian Heide (da Noruega)

Melhor performance no Sigma Test VI.

Recordista mundial em Brain Master Challenge Series (talvez a pessoa com pensamento mais rápido do mundo).

M.Sc. em Astrofísica.

Detentor de dois recordes nacionais (Noruega) envolvendo motricidade, equilíbrio e potência física, um dos quais registrado no Guinness Book (lançamento de ovo à distância e maior tempo sobre uma cadeira equilibrada em duas pernas).

Classificação dos níveis de inteligência

Classificação proposta por Lewis Terman (~1916)

QI acima de 140:	Genialidade
120-140:	Inteligência muito superior
110-120:	Inteligência superior
90-110:	Inteligência normal (ou média)
80-90:	Embotamento
70-80:	Limítrofe
50-70:	Cretino
20-50:	Imbecil
QI abaixo de 20:	Idiota

Classificação proposta por David Wechsler (~1940)

QI acima de 127:	Superdotação (algumas fontes citam o termo "gênio > 150")
120-127:	Inteligência superior
110-120:	Inteligência acima da média
90-110:	Inteligência média
80-90:	Embotamento ligeiro
65-80:	Limítrofe
50-65:	Debilidade ligeira
35-50:	Debilidade moderada
20-35:	Debilidade severa
QI abaixo de 20:	Debilidade profunda

Classificação adotada em Sigma Society com base em Análise Fatorial Hierárquica (2005)

QI acima de 200:	Imensurável
185-200:	Gênio Universal
176-185:	Gênio altamente criativo (escassez de dados amostrais)
164-176:	Gênio criativo
151-164:	Gênio
142-151:	Muito talentoso
135-142:	Superdotado ou talentoso (portador de necessidades especiais)
127-135:	Inteligência muito acima da média
116-127:	Inteligência acima da média
82-116:	Inteligência média ou normal (escassez de dados amostrais)
QI abaixo de 82:	(Ausência de dados amostrais para estabelecer uma classificação)