

DM
RIBE/R1

Instituto Superior de Psicologia Aplicada

Mestrado em Comportamento Organizacional

Dissertação de Mestrado

Os tempos de latência nas repostas aos itens de testes informatizados:


Contributos para a compreensão do processamento cognitivo

Rui Manuel Bártolo Ribeiro, nº 155

Orientador: Prof. Doutor Leandro Almeida
Universidade do Minho

1999/2001

C

 I SPA | Instituto Superior de Psicologia Aplicada

Centro de
Documentação

Registo: 15597
Data: 01.03.05

Tel: 21 001 17 00 • libspa@ispa.pt

*“O homem só se descobre
quando se confronta com um obstáculo”*

Antoine De Saint Exupéry

Dedico este trabalho à minha filha Joana.

Agradecimentos

Por imperativo da responsabilidade do texto final irei agradecer em nome pessoal e começaria pelo orientador, Prof. Doutor Leandro de Almeida, que apesar da sua permanente ocupação em projectos vários, esteve sempre disponível desde o momento em que foi convidado. Com uma orientação plena, não coarctando a liberdade de expressão, pelo que será justo reconhecer que as falhas aqui encontradas apenas poderão ser atribuídas ao autor.

Agradeço à Sónia Silva pelo prestimosa ajuda na recolha bibliográfica e na discussão do plano metodológico.

Um agradecimento forte ao Miguel Pina e Cunha, colega de Mestrado, amigo e companheiro do trajecto académico, pelo seu incentivo para a conclusão do Mestrado, com um preciosa ajuda na recta final, bem reveladora da sua competência técnica, mas sobretudo dos valores que norteiam a sua conduta como profissional e como Homem.

À Mané um agradecimento muito profundo, pelo incentivo constante, pela disponibilidade para acaloradas e proficuas discussões e pelos comentários à versão final, efectuados por quem sabe mas com um formato característico de quem se é amigo.

Do ISPA agradeço ao Director e amigo Prof. Doutor Frederico Pereira. Aos colegas e amigos um agradecimento muito especial pelas palavras de incentivo e disponibilidade para a ajuda. Agradeço (por ordem alfabética) a: Ana Cristina Martins, Ângela Vila-Real, Carlos Marques Cristina Quelhas, Emília Marques, Fernando Silva, Jorge Gomes, Luís Andrade, Olga Romão,

Olivia Ribeiro, Pedro Almeida, Teresa d'Oliveira, Teresa Garcia-Marques e Victor Cláudio. Deste grupo de colegas realço a ajuda na versão final da Teresa d'Oliveira e o incentivo constante do Victor.

Gostaria de agradecer a elementos da Força Aérea, nomeadamente, ao Coronel Médico Gabriel Olim e ao Tenente-Coronel Piloto Carlos Macário pelas facilidades concedidas para efeitos de investigação. À própria Força Aérea, na pessoa do seu Comandante de Pessoal, General Piloto Aviador Rui Fidalgo pela autorização concedida para a elaboração de uma parte significativa do trabalho. À Rafaela Coelho pelos sucessivos processamentos das alterações efectuadas ao texto. Aos meus camaradas, mas também amigos, Paulo Marques, José Valente e António Surrador pelo apoio.

Agradeço de forma sentida à Gi, não só pela revisão do texto final, mas sobretudo, pela inteligência em compreender e gerir de forma exímia as emoções, proporcionando-me as condições necessárias para a conclusão deste trabalho. Mais uma vez, obrigado.

À minha filha agradeço a compreensão pelo tempo roubado à nossa relação.

Não podendo ser exaustivo e considerando o longo período de elaboração deste trabalho, sinto que não a agradeço, de forma reconhecida, a um conjunto grande de pessoas que contribuíram directa ou indirectamente para a realização deste trabalho. A todos vós... OBRIGADO.

Índice

Resumo	11
Introdução	14

Parte Teórica

1. As aptidões cognitivas e o processamento de informação	20
1.1. Aptidões cognitivas: Diferentes terminologias e diferentes modelos	20
1.2. Memória a curto-prazo e memória de trabalho	35
1.3. Velocidade mental	41
1.3.1. Trade-off entre correcção e rapidez	44
1.3.2. Complexidade da tarefa	50
1.3.3. Limitação do tempo de aplicação	53
1.4. Os contributos da psicologia cognitiva na avaliação das aptidões cognitivas	58
2. Avaliação psicológica informatizada	61
2.1. Diferentes tecnologias na avaliação psicológica	63
2.2. Vantagens da informatização para a avaliação psicológica	66
2.2.1. Construção e/ou escolha do teste	67
2.2.1.1. Testes adaptativos	68
2.2.1.2. Testes novos, novas dimensões	70
2.2.2. Aplicação	73

2.2.3. Correção e padronização	73
2.2.4. Interpretação dos resultados e redacção do relatório	74
2.2.5. Tomada de decisão	75
2.3. Testes de papel-e-lápis versus testes informatizados	76
2.4. Potencialidades e limites dos sistemas informatizados	81
2.5. Comentários finais	83
3. Avaliação psicológica em contexto da selecção de pessoal	85
3.1. Os princípios da selecção	87
3.2. O paradigma da selecção de pessoal	89
3.3. A selecção de pessoal como processo	91
3.3.1. Análise de funções	93
3.3.1.1. Especificações da função	94
3.3.2. Escolha dos métodos de selecção	96
3.3.3. Validade relativa ao critério	97
3.3.4. Restrições aos coeficientes de validade	99
3.3.5. Métodos de selecção e validade preditiva	100
3.4. Validade <i>versus</i> utilidade	107

Parte Empírica

Introdução aos estudos empíricos	114
4. Estudo 1 - Estratégia nos testes de escolha múltipla	116
4.1. Instrumento	116
4.2. Amostra	120
4.3. Apresentação e discussão de resultados	121
4.3.1. Qualidades métricas	121
4.3.2. Análise das médias e desvios-padrão dos tempos de latência aos itens ..	122

4.3.3. Factores não cognitivos	126
4.3.4. Identificação dos factores avaliados	131
4.4. Sumário dos resultados	134
5. Estudo 2 - Desempenho nas Matrizes Progressivas de Raven	136
5.1. Amostra	136
5.2. Instrumento	138
5.3. Análise e discussão de resultados	138
5.3.1. Distribuição dos resultados	139
5.3.2. Consistência interna	141
5.4. Velocidade mental e inteligência psicométrica	145
5.5. Comentários finais	151
Considerações finais	154
Referências bibliográficas	162

Índice de Figuras

Figura 1.1	Curva hipotética Exactidão-Tempo de Latência. O sujeito B é mais lento e mais exacto que o sujeito A (retirado de Lohman 1994b, p. 159)	45
Figura 1.2	Parâmetros da curva hipotética Exactidão-Tempo de Latência, segundo W. A. Wickelgren (1997)	47
Figura 3.1	Paradigma de selecção baseado em Smith e Robertson (1989)	93
Figura 3.2	Representação gráfica de dois hipotéticos coeficientes de correlação	98
Figura 4.1	Exemplo de um item do teste “Padrões”	117
Figura 4.2	Médias dos tempos de resposta aos itens	122
Figura 4.3	Desvios-padrão nos tempos de resposta aos itens	123
Figura 4.4	Relação entre o tempo médio de resposta e o número de respostas correctas nos itens 1 e 12	126
Figura 4.5	Médias dos tempos de latência e o número de respostas incorrectas para cada um dos itens	129
Figura 4.6	Representação espacial das variáveis tempo e erros do teste de padrões e dos testes de inteligência geral e rapidez perceptiva	133
Figura 5.1	Distribuição dos sujeitos por idades	137
Figura 5.2	Distribuição dos sujeitos por grupos de habilitações literárias	137
Figura 5.3	Gráfico de barras com as frequências dos resultados finais em termos de respostas correctas	140
Figura 5.4	Histograma com os tempos gastos nos itens respondidos correctamente	141
Figura 5.5	Gráfico de dispersão dos itens pelo número total de respostas correctas	142
Figura 5.6	Gráfico de dispersão dos itens pelo número total de respostas correctas e pelas médias dos tempos de latência	145

Índice de Quadros

Quadro 1.1	Processo do tempo de reacção de escolha segundo o modelo de D. P. Keating e B. L. Bobbitt (1978)	34
Quadro 3.1	Amplitude dos coeficientes de validade para um conjunto de métodos de selecção segundo I. Robertson e M. Smith (1989)	101
Quadro 3.2	Comparação entre dois métodos com diferentes validades num estudo de utilidade	111
Quadro 4.1	Identificação das variáveis do teste “Padrões”	120
Quadro 4.2	Médias e desvios-padrão dos tempos de resposta aos itens respondidos correcta e incorrectamente e correspondente análise da significância da diferença de médias	124
Quadro 4.3	Médias e desvios-padrão dos tempos de resposta por item e as respostas pretendidas	128
Quadro 4.4	Teste de diferença de médias entre os itens em que a figura-alvo não aparece nas figuras complexas e os itens em que a figura-alvo aparece nas duas figuras complexas	130
Quadro 4.5	Correlações momento-produto, níveis de significância e número de sujeitos entre dois testes psicométricos e as notas globais do “Padrões”	131
Quadro 4.6	Matriz factorial após rotação ortogonal	133
Quadro 5.1	Estatística descritiva para as variáveis respostas correctas (RC) e tempo gasto em segundos (TEMP) para a totalidade do teste	139
Quadro 5.2	Hipotéticos alphas para as Matrizes Progressivas de Raven na forma Standard se os itens respectivos forem eliminados	142
Quadro 5.3	Distribuição dos sujeitos pelos três níveis de capacidade mental geral	146
Quadro 5.4	Médias e desvios-padrão dos tempos de latência, e número de respostas correctas para cada um dos itens por nível de capacidade mental geral e respectivas análises de variância	147

Resumo

A velocidade de processamento de informação está fortemente relacionada com o constructo inteligência geral (Jensen, 1982; Vernon, Nador, & Kantor, 1985b), assim como, a velocidade mental, tal como a capacidade da memória de trabalho poderão ser considerados medidas indirectas da inteligência geral. Se bem que a capacidade da memória de trabalho possa ser operacionalizada pela velocidade mental, a verdade é de que não existe racional teórico que sustente a redução da capacidade da memória de trabalho à velocidade mental (Stankov & Roberts, 1997).

Por outro lado, o rendimento numa tarefa cognitiva, irá ser, não só uma função da capacidade de armazenamento/processamento da memória de trabalho como também da complexidade da mesma percebida pelo sujeito (Marshalek, Lohman, & Snow, 1983; Stankov, 2000). A percepção da estrutura da tarefa e sua complexidade permite a emergência de factores não-cognitivos limitadores da expressão plena da capacidade cognitiva do sujeito (Stankov, 2000; Wilhelm & Schulze, 2002).

Nos testes informatizados, a velocidade mental no processamento de informação relativa a provas cognitivas pode ser operacionalizada pelos tempos de latência, que corresponde ao tempo gasto pelo sujeito, desde que o item lhe é apresentado até expressar a resposta.

Foram realizados dois estudos centrados nos tempos de latência a itens de testes de realização cognitiva fortemente saturados de factor *g* (um do tipo das figuras embebidas e outro de

inteligência geral) aplicados a sujeitos em situação real de selecção: 731 candidatos ao curso de pilotagem da Força Aérea Portuguesa no primeiro estudo e a 300 candidatos a condutores de veículos pesados de transporte de passageiros.

Os tempos de latência médios são diferentes quando os sujeitos acertam ou erram os itens. Aos itens respondidos correctamente corresponderam tempos de latência mais rápidos e mais homogéneos do que aos incorrectos. Verificou-se a emergência de factores não-cognitivos que limitaram a eficiência dos sujeitos em termos de velocidade mental, tendo-se verificado diferenças significativas na comparação entre as diferenças alternativas de respostas. Os candidatos foram significativamente mais rápidos quando a figura-alvo se encontrava nas duas figuras complexas, do que quando não se encontrava em nenhuma. Numa prova de escolha múltipla a incerteza da resposta pode potenciar a complexidade do item proporcionando a emergência de factores não cognitivos que limitam a eficiência da memória de trabalho, baixando a velocidade de resposta. O número de respostas correctas decresce, aumentando o tempo de latência, à medida que aumenta o grau de complexidade Schweizer (1996). Este aspecto é agravado pelo conhecimento do tempo restante para terminar.

No segundo estudo, foram utilizadas as Matrizes Progressivas de Raven na sua forma *Standard*, mas de aplicação em suporte informático. Os resultados, à semelhança do estudo anterior, apontam para a confirmação da hipótese de que à medida que os itens vão apresentando maior índice de complexidade, maior é o tempo gasto na sua resolução.

Quando se dividiu a amostra em dois grupos com diferentes níveis de rendimento intelectual global, verificou-se que os sujeitos com maior capacidade mental geral tendem a responder mais rapidamente que os sujeitos com menor índice de capacidade mental geral. Apesar de as diferenças dos tempos de latência entre estes dois grupos tenderem a aumentar à medida que aumenta o grau de dificuldade dos itens, este padrão é significativamente influenciado pela limitação temporal da aplicação do teste. O sujeito mais inteligente tem vantagens em realizar

um teste de aptidão mental com complexidade crescente e tempo limite de aplicação. Nos itens iniciais (teoricamente menos complexos) acede mais rapidamente à informação necessária e responde mais rapidamente às perguntas com a excepção dos itens demasiadamente simples. Nas questões mais complexas, os mais inteligentes, ao serem mais rápidos a processarem informação, têm a possibilidade de executá-las sem que a capacidade (limitada) da memória de trabalho entre em sobrecarga.

No segundo estudo, também se verificou a influência de factores de ordem não-cognitiva na resolução de tarefas cognitivas. Se considerarmos as cinco séries de doze itens que compõem a prova, os tempos de latência gastos pelos sujeitos nos itens iniciais, intermédios e finais de cada série são diferentes entre si. Quando comparamos os sujeitos em termos do seu nível de capacidade mental geral, verifica-se que as diferenças entre os grupos são apenas estatisticamente significativas para os itens intermédios de cada série. Os resultados convergem com os do estudo de Stankov (2000), em que o aumento de complexidade cognitiva está associado à maior probabilidade de emergência de factores não cognitivos na realização de tarefas intelectuais reduzindo a eficiência da memória de trabalho e baixando consequentemente a velocidade mental. O rendimento do sujeito é, assim, afectado não só pelas suas limitações em termos de processamento de informação, como também por aspectos de ordem não-cognitiva. As características não-cognitivas, nomeadamente as de personalidade poderão interferir na estratégia cognitiva individual de controlo do esforço e tempo.

Dos dois estudos realizados, parece ser claro que o acesso ao tempo de latência das respostas, conseguido através da informatização dos testes, permite-nos conhecer melhor, não só as diferenças individuais ao nível do atributo cognitivo que se está a avaliar, como também das implicações da própria estrutura da prova na estratégia de resolução dos itens.

Introdução

O paradigma de selecção apresentado por Smith e Robertson (1989) e assumido por Smith, Farr e Schuler (1993), parece ser o que reúne maior consenso entre os responsáveis pela selecção de pessoal. De acordo com este modelo, o processo de selecção engloba a análise de funções e correspondente identificação das exigências, a atracção dos candidatos, a escolha e aplicação dos métodos de selecção, a escolha dos candidatos e a validação do processo de modo a retroalimentá-lo qualitativamente.

Os testes psicológicos são um dos métodos mais utilizados (Bártolo Ribeiro, 1995) para avaliar os atributos psicológicos necessários ao melhor desempenho da função para que está a concorrer. Fornecem um prognóstico do desempenho futuro dos candidatos tão rigoroso quanto as suas qualidades psicométricas o permitem. Neste âmbito, a validade relativa ao critério tem um papel de grande destaque.

Se bem que a maioria dos testes utilizados em contexto de selecção de pessoal para a avaliação das aptidões perceptivo-cognitivas, psicomotoras e características da personalidade, tem sido aplicada no formato papel-e-lápis, é cada vez mais frequente desenvolverem-se processos de selecção em que as provas são aplicadas informaticamente. A utilização de testes informatizados, quer criados a partir de provas de papel-e-lápis, quer desenvolvidos especificamente para serem aplicados em computador, não pode ser considerada, por si só, uma mais-valia do processo de selecção. O ganho com a utilização de testes no contexto da selecção de pessoal deriva sobretudo das qualidades métricas que esses testes possuem e não

propriamente da sofisticação da plataforma utilizada. No entanto, os testes aplicados com base em computadores poderão contribuir para um melhor conhecimento do funcionamento psicológico do sujeito, nomeadamente no que diz respeito à esfera cognitiva.

A maioria dos estudos sobre as aptidões com vista à sistematização da sua categorização têm por base os resultados obtidos em testes aplicados no formato papel-e-lápis (Carroll, 1993), em que os investigadores trabalharam essencialmente com *scores* resultantes do número de respostas correctas. A construção de provas informáticas tem a vantagem de aceder a informação não disponível nos testes de papel-e-lápis, como seja o tempo de latência nas respostas aos itens o que proporciona um estudo de maior profundidade da evolução dos sujeitos ao longo da prova. Uma outra grande vantagem da informatização consiste na construção de novas provas sem equivalente no tradicional formato papel-e-lápis, como os testes com movimento, a utilização de sistemas multimedia e a execução de tarefas simultâneas (*time-sharing*), permitindo avaliar atributos psicológicos não investigados anteriormente (Drasgow, Olson, Keenan, Moberg & Mead, 1993; Pellegrino & Hunt, 1989). A avaliação de novos constructos por intermédio de testes computadorizados permite maior conhecimento das capacidades mentais, mas também, cria a necessidade de reformulação de muito do que se escreveu sobre as aptidões cognitivas. A evolução tecnológica dos computadores (sobretudo em termos do *display*) e a criatividade dos construtores de testes irão contribuir fortemente para uma análise mais rica das capacidades cognitivas. Fernández-Ballesteros (1995) reforça ainda mais esta posição, ao afirmar que uma das vias de desenvolvimento da avaliação psicológica será efectuada através da avaliação assistida por computador e pela realidade virtual, como resultante da intercomunicação entre as novas tecnologias e as ciências cognitivas.

A grande motivação que orientou este nosso trabalho para o estudo de provas informatizadas foi a de explorar as mais-valias da utilização de sistemas informáticos no âmbito da avaliação

psicológica no contexto específico da selecção de pessoal. Procurou-se, por intermédio de dois estudos, analisar o processamento de informação do sujeito a partir dos tempos de latência das respostas aos itens. O nosso objectivo centrou-se mais na análise do comportamento dos sujeitos na execução de tarefas cognitivas a partir dos tempos de resposta às questões do que propriamente no estudo em profundidade dos conteúdos cognitivos avaliados pelos testes.

No nosso primeiro capítulo pretende-se dar uma visão dos diferentes conceitos relativos às aptidões cognitivas. Se por um lado a perspectiva psicométrica tem dominado o estudo dos atributos psicológicos, o acesso aos tempos de latência das respostas tem trazido à discussão a perspectiva cognitivista na análise das tarefas intelectuais, permitindo conhecer as aptidões cognitivas e o processamento de informação pela perspectiva da velocidade mental.

No segundo capítulo apresenta-se uma panorâmica dos diferentes sistemas informáticos possíveis de serem utilizados em diferentes etapas da avaliação psicológica, evidenciando os ganhos e perdas, comparativamente com a versão clássica de papel-e-lápis.

No terceiro e último capítulo da parte teórica tivemos por objectivo dar um enquadramento mais específico à avaliação psicológica, abordando os princípios base da selecção de pessoal, onde o uso dos testes é determinante para o conhecimento dos recursos pessoais dos candidatos com vista à tomada de decisão relativamente à sua admissão.

Na componente prática deste trabalho procurou-se abordar a avaliação psicológica na intersecção de três aspectos: conteúdo avaliado, forma e contexto de aplicação. O conteúdo esteve essencialmente orientado para a avaliação das aptidões cognitivas, excluindo, deste modo, as aptidões motoras, as características do âmbito da personalidade e dos interesses. As provas foram aplicadas de forma informatizada e num contexto real de selecção de pessoal.

Desenvolvemos dois estudos com objectivos diferenciados mas com um denominador

comum, que é a análise dos tempos de latência às respostas, como meio de conhecermos melhor as diferentes variáveis intervenientes no funcionamento cognitivo na realização de provas psicométricas. No primeiro estudo utilizamos um teste cognitivo fortemente saturado ao nível perceptivo, do tipo das figuras embebidas, em que o sujeito tem de identificar se uma figura-alvo se encontra ou não nas figuras complexas apresentadas como figuras distractoras (“*perceptual closure*”). Pretendeu-se saber se a eficiência do sujeito, ao nível da velocidade mental, está apenas ancorada à complexidade/dificuldade da tarefa, ou se podem ser considerados outros factores intervenientes no processamento de informação.

O segundo estudo tem como base o teste das Matrizes Progressivas de Raven na versão *Standard*, de aplicação informatizada. Tomando em consideração variáveis como a cotação atribuída ao item respondido pelo sujeito, o tempo gasto nas respostas correctas, o resultado final em termos do número de respostas correctas e as características específicas dos itens, vai ser analisado como os sujeitos se diferenciam em termos de evolução ao longo da prova. Neste estudo também se vão comparar os sujeitos com diferentes níveis de capacidade intelectual, com o objectivo de testar a hipótese de que os sujeitos com um nível intelectual mais elevado, em termos de inteligência psicométrica, são mais rápidos a processar informação (tempos de latência mais baixos) quando o nível de complexidade da tarefa é elevado, comparativamente com sujeitos com um nível intelectual mais baixo.

Os testes utilizados nos dois estudos não foram desenhados de propósito para esta investigação. Fazem parte da própria bateria utilizada nos respectivos processos de selecção, assim como a sua aplicação foi efectuada em contexto real de selecção de pessoal, pelo que o procedimento não possui um carácter experimental.

Reforçamos mais uma vez que a preocupação central do nosso trabalho está focalizada no estudo dos tempos de latência às respostas a itens de testes informatizados, de modo a questionar a importância desta dimensão na utilização de estratégias de resolução de tarefas

cognitivas, com vista à determinação do incremento da capacidade preditiva do sucesso posterior no desempenho profissional de funções exigentes ao nível do processamento rápido de informação (pilotos, condutores, etc.).

Parte Teórica

1. As Aptidões Cognitivas e o Processamento de Informação

Iremos começar este capítulo por apresentar diferentes termos relacionados com o funcionamento cognitivo, tanto na língua inglesa como na portuguesa. Passaremos, de seguida, a abordar a memória e a velocidade mental como dois constructos fortemente implicados no processamento de informação, mas também como medidas indirectas da avaliação da inteligência. Terminaremos com a descrição do contributo da psicologia cognitiva para um melhor conhecimento do funcionamento mental.

1.1. Aptidões cognitivas: Diferentes terminologias e diferentes modelos

De acordo com Cronbach (1984), os conceitos de *intelligence* e *capacity* são usados na cultura britânica como indicador do potencial do sujeito (descrito pela nota no teste) como garante do nível que o sujeito atingiria se lhe fosse dado todo o suporte educativo. Ainda segundo o mesmo autor, o conceito de *ability* está directamente relacionado com o rendimento (*performance*) quando lhe é pedido que dê o seu máximo. O resultado da prestação do sujeito numa determinada prova, quando este dá o seu melhor, será a sua *ability*. O rendimento será, assim, o equivalente a um estado e não propriamente a um traço, pelo que a *ability* está relacionada com o rendimento quando são proporcionadas adequadas condições e é solicitado o máximo empenho.

Por outro lado, os testes avaliadores de *aptitude* visam predizer o sucesso numa função ou num curso. São testes específicos a essa função ou curso, como sejam, a título de exemplo, a *aptitude* para a matemática ou para a música (Cronbach, 1984). Desta forma, um teste que meça *aptitude* para a engenharia pode incluir secções que avaliem a designada *general mental ability*, raciocínio mecânico e espacial e a proficiência em matemática.

O termo *aptitude* é frequentemente usado como uma capacidade, conveniência ou mesmo inteligência. No âmbito da psicologia é geralmente referido como um conjunto de *abilities* que surgem como importantes para uma cultura particular (Kline, 2000). Mas a diferenciação entre testes avaliadores de *aptitudes* e testes avaliadores de *abilities* não é clara. Kline (2000, p. 230) afirma, inclusive “... *the verbal aptitude test is nothing more than a test of verbal ability factor.*”. Este autor assume uma posição radical relativamente ao uso do conceito de *aptitude* como conceito psicométrico ou científico, afirmando que deve ser abandonado enquanto tal, já que envolve dois outros conceitos bem diferentes e que são: *Abilities* e *attainments*. Para este autor, *attainment* é visto como a posse de *skills* que foram adquiridos através da experiência ou treino deliberado, enquanto que *capacity* ou *ability* referem-se ao potencial para adquirir tais *skills*. Se considerarmos como tradução de *attainment* a realização, podemos considerar que os testes de realização são tangenciais aos testes psicométricos.

Os testes de *aptitude* distinguem-se dos de *achievement* mais pela perspectiva do que pelo conteúdo. Enquanto os últimos visam essencialmente avaliar o desenvolvimento e o conhecimento até ao momento, os primeiros têm um carácter de previsão.

Utilizámos propositadamente quatro termos que na língua inglesa estão fortemente associados à avaliação intelectual: *capacity*, *ability*, *aptitude* e *achievement*. A separação entre estes conceitos não é, de modo algum, clara e pacífica. Numa primeira abordagem, poderemos considerar que *capacity* corresponde a algo adquirido geneticamente e que está direccionado

para uma área específica e que vai influenciar a realização nos factores que lhe estão subjacentes. O conceito de *ability* deriva directamente da *capacity* geral mais lata, derivando em várias *abilities* específicas, cujo valor resulta da nota obtida numa prova específica para a avaliar. Seria, numa linguagem simbólica, a parte visível da *capacity* do sujeito. O conceito de *aptitude* estaria relacionado com o sucesso futuro numa área profissional específica. Essa área profissional exigiria que o sujeito possuísse várias *aptitudes*. Por último, o conceito de *achievement* estaria fortemente associado a conhecimentos adquiridos.

Mas esta tentativa de separação de conceitos não encontra eco em todos os autores. Howe (1996) no seu capítulo “*Human abilities: Their nature and measurement*” evidencia que um dos problemas para compreender as *human abilities* está relacionado com razões conceptuais. Para este autor existe mais do que uma maneira de utilizarmos a palavra *ability*; ela pode ser usada para fazer afirmações relativamente ao que a pessoa é capaz de executar (perspectiva descritiva), mas também pode ser utilizada para nos referirmos ao que a pessoa possui e que lhe possibilita executar determinada tarefa (perspectiva explicativa).

Cooper (1999), numa tentativa de definição dos conceitos, efectua uma retrospectiva histórica sobre o que significam os termos *intelligence*, *mental ability* e *aptitude*. Refere que Sternberg e Salter (1982) ao se verem confrontados com a dificuldade em decidir entre impor uma definição única de inteligência aos diferentes autores da obra *Handbook of Human Intelligence* (Sternberg, 1982), ou encontrar uma definição comum a todos os autores que contribuíram para o livro, optaram por considerar a inteligência como um comportamento adaptado. Numa frase, poder-se-ia dizer que, ser inteligente, é ter recursos e saber utilizá-los de forma optimizada com vista a um comportamento adaptado a um objectivo/contexto.

Para Cooper (1999), *abilities* em sentido lato, são todos os comportamentos passíveis de serem avaliados. Como exemplos, temos digitalizar textos, conhecimento das linhas

ferroviárias de um país, corrida, leitura de um mapa, cozinhar, desenhar pontes, ajudar pessoas deprimidas, resolver anagramas, etc. Como aspecto constitutivo determinante está a possibilidade de avaliar a eficácia de cada uma dessas actividades, quer através da monitorização do comportamento (contar o número de palavras digitadas por minuto, por exemplo), quer pedindo às pessoas que avaliem o comportamento (como, por exemplo, pedir às pessoas sujeitas a um tratamento psicoterapêutico que avaliem a eficácia do mesmo).

Numa tentativa de passarmos a utilizar os termos em português, verificamos que, de acordo com o dicionário ilustrado de Inglês-Português da Michaelis, os termos até agora abordados têm a seguinte tradução: para *capacity*, capacidade, faculdade, competência, aptidão; para *ability*, capacidade, habilidade, competência, agilidade, destreza, talento, aptidão, faculdade; para *aptitude*, aptidão, capacidade, habilidade, inteligência, perspicácia; e para *achievement*, façanha, realização.

A confusão sobre o que os termos ingleses significam e a sua tradução para a língua portuguesa parece evidente. O único termo que, na tradução, surge com significado próprio e independente é *achievement*, que podemos doravante considerar como realização e que traduz o nível alcançado pelo sujeito na aprendizagem de uma área específica do conhecimento. Em relação aos outros conceitos, encontrámos capacidade e aptidão como os mais evocados, mas também se constata a referência a faculdade, competência, habilidade e talento. Num esforço de efectuarmos uma correspondência objectiva de significados entre os termos portugueses e os anglo-saxónicos, procurou-se em dicionários especializados o respectivo significado.

De acordo com o léxico Inglês-Português de Doron e Parot (2001) os termos ingleses que temos vindo a referir têm o seguinte significado na nossa língua: *ability* - destreza; *achievement* - não tem correspondente; *aptitude* - aptidão; e *capacity* - capacidade. Ainda com

base na obra de Doron e Parot, a aptidão é definida como o “... êxito de qualquer actividade, seja ela uma tarefa isolada, uma conduta complexa, uma aprendizagem ou ainda o exercício de uma profissão, exige que o sujeito possua as capacidades e as motivações adequadas. A capacidade é, em si mesma, condicionada pela presença de uma aptidão que pode ser desenvolvida através da experiência e da formação. Deste modo, qualquer um pode ter uma boa aptidão verbal; com a experiência e treino, falar em público pode tornar-se excelente. Com efeito, o termo aptidão assume acepções diferentes e mesmo opostas. C. Spearman utilizou-a para designar a inteligência, enquanto a mesma palavra, acompanhada por um qualificativo, é empregue a propósito de aptidões específicas, não intelectuais (sociais ou motoras, por exemplo), podendo mesmo designar as aquisições necessárias para o exercício de uma actividade muito especializada, como, por exemplo, escrever à máquina ou conduzir um veículo. Por último, como relembra M. Reuchlin, é necessário distinguir aptidão, capacidade e factores, na medida em que os factores derivam de uma análise factorial que nada presume sobre a sua origem inata ou adquirida.” (p. 79-80, redigido por C. Lévy-Leboyer, in Doron & Parot, 2001).

O termo capacidade é descrito por C. Prévost na obra de Doron e Parot (2001) do seguinte modo “Possibilidade de êxito e de competência num domínio prático ou especulativo. Ela pode ser verificada directamente num contexto escolar ou profissional e indirectamente por testes diversos (orientação escolar ou profissional). A análise factorial (C. Spearman e L. Thurstone) através dos seus modelos matemáticos (unitários ou hierárquicos) esforça-se por medir aptidões primárias, substractos das capacidades...” (Doron & Parot, 2001, p.124).

Por sua vez, o termo destreza resultante da tradução do termo *ability* é descrito por M. Durand, na obra de Doron e Parot (2001), como “... a precisão dos ajustamentos motores no decorrer de tarefas de pontaria, de antecipação-coincidência ou de perseguição, nas quais

existem grandes diferenças individuais. Estas características não correspondem a estádios de tratamento da informação nem a aptidões especiais. Os progressos registados na compreensão dos processos de controlo motor levam a adoptar uma atitude crítica face a este termo, devido sobretudo à sua polissemia e à inexistência de factos experimentais susceptíveis de a apoiar.” (Doron & Parot, 2001, p. 232). Este enfoque na componente motora está completamente desfasado do seu correspondente inglês *ability*, que na maioria das vezes é utilizado para melhor caracterizar as *mental abilities*.

Com base no Dicionário de Língua Portuguesa Contemporânea da Academia das Ciências de Lisboa e da Editora Verbo (2001), o termo aptidão é definido como a capacidade natural ou adquirida de alguém para determinada actividade. Também pode ser traduzido como a preparação física ou intelectual, ou o conjunto de capacidades para realizar uma tarefa. O termo capacidade é descrito como aptidão ou habilidade. Por outro lado, o talento é visto como uma aptidão digna de nota, natural ou intelectual, para alguma coisa, para uma actividade.

Considerando a grande confusão entre os termos capacidade e aptidão, vamos considerar, ao longo deste trabalho, a aptidão como tradução de *ability*, a qual tem como principal significado a quantidade de capacidade num atributo específico capaz de se expressar na realização de uma tarefa particular. Por capacidade (tradução de *capacity*), entendemos que será todo o volume do atributo em causa que o sujeito possui, e que se identificaria com a respectiva aptidão se a tarefa proposta ao sujeito fosse uma representação perfeita desse atributo e tivesse sido desempenhada nas condições mais favoráveis.

Carroll (1993) apresenta de forma relativamente clara, do ponto de vista operacional, a diferença entre *ability* e *aptitude*. Considerando que as aptidões cognitivas (*cognitive abilities*) são, no mínimo, relativamente estáveis e relativamente resistentes às tentativas de

mudança através da educação ou do treino e, ao mesmo tempo, capazes de predição do sucesso futuro, são também vistas como *aptitudes*. Assim, o termo *aptitude*, refere-se, em sentido restrito, a uma *cognitive ability*, possivelmente preditora de um conjunto de aspectos relacionados com o sucesso futuro na aprendizagem. Desta forma, *ability* será um termo utilizado para descrever atributos, enquanto *aptitude* terá associada uma componente de predição. Como é referido por Carroll (1993, p.17), uma *ability* será uma *aptitude* se ajudar a prever o grau de aprendizagem, com base numa predição do nível anterior de aprendizagem.

Se as aptidões correspondem à parte observável das capacidades que os sujeitos possuem relativamente a um determinado atributo, então quando elas são avaliadas com vista a se estabelecer uma previsão de sucesso no processo de aprendizagem e/ou o futuro desempenho, estaremos a referir-nos a um outro conceito que será, em língua inglesa, *aptitude* e que eventualmente poderá corresponder ao termo português de potencial.

Por sua vez, os termos desempenho, *performance*, rendimento ou realização será sempre o reflexo objectivo e mensurável do que foi aprendido/treinado.

Numa tentativa de melhor clarificação apresentamos um exemplo que conjuga os diferentes conceitos até agora abordados e que, de algum modo, está relacionado com o contexto aeronáutico. O desempenho dos pilotos militares em cenários de treino combate ar-ar, está dependente, entre outros aspectos, da aptidão espacial. Com o nascimento recebemos uma determinação de ordem genética condicionante do desenvolvimento das nossas capacidades, incluindo a espacial. Caso sejam preenchidas favoravelmente todas as condições possíveis, relativamente à estimulação, será atingida a cota máxima possível, que será diferente de sujeito para sujeito. A capacidade possível de cada sujeito irá determinar as limitações do processo de aprendizagem de tarefas onde esse atributo seja solicitado. Através de provas constituídas por itens específicos, é possível solicitar ao sujeito determinados

comportamentos que ajudem a ter uma noção da sua capacidade. O que o sujeito demonstra na execução dessas provas, é o que aqui vamos considerar como traço manifesto ou estado, e que corresponderá à aptidão do sujeito naquele atributo específico. Quanto mais válidos e mais representativos forem os itens que constituam o teste, e mais favoráveis as condições de realização da prova, maior a probabilidade de aproximação entre a aptidão demonstrada e a correspondente capacidade do sujeito. Uma nota elevada num teste de orientação espacial não só nos dá uma estimativa da capacidade real do sujeito naquele atributo específico, como pode constituir um preditor do sucesso no processo de aprendizagem da leitura e interpretação dos instrumentos do avião. A nota obtida no teste é, assim, um indicador do potencial desse avaliado para aprender um conjunto de matérias práticas (leitura de mapas; execução de manobras acrobáticas a baixa altitude; etc.) construídas com base nesse atributo e eventualmente noutros. A aprendizagem que o sujeito vai ter da função de pilotagem, e o conseqüente treino de determinadas manobras, irá posteriormente ser avaliada num teste de conhecimentos ou de rendimento. A associação estatística entre o resultado no teste de aptidão e o resultado da prova de conhecimentos e/ou de avaliação de desempenho, traduz-se na capacidade preditiva daquele teste de aptidão. Se a validade preditiva desse teste for elevada, a nota no teste de aptidão constituirá um indicador do potencial do sujeito para aquisição desses conhecimentos específicos. O potencial permite-nos estimar a percentagem de sucesso prevista num programa específico de treino de uma outra tarefa igualmente saturada de aptidão espacial.

Mas se a caracterização destes quatro conceitos aparenta existirem fronteiras suficientemente delimitadas, o significado dos termos carece de melhor clarificação. Esta confusão também aparece na língua espanhola. Marañón (1997), ao definir o conceito de *capacidad humana*, associa o termo aptidão *aptitud* ao de capacidade *capacidad* dando-lhe o mesmo significado. Com base numa entrevista efectuada a Kyllonen, Marañón (1997) afirma que o conceito de

“aptidão ou capacidade cognitiva” traduz a aptidão para operar na esfera cognitiva. Não distingue aptidão de capacidade, mas sim entre serem básicas ou serem efectivas/aplicadas. A primeira, aptidão/capacidade básica, corresponde ao aspecto mais estável, que é sobretudo herdado, e é, também, quem determina até onde se pode chegar no processo de aprendizagem através da realização de uma específica tarefa cognitiva de certo nível de complexidade. A aptidão/capacidade básica não é passível de ser avaliada directamente, pelo que o que é observado quando se realiza uma determinada tarefa cognitiva, é a aptidão/capacidade efectiva.

Desta forma, a aptidão/capacidade efectiva resulta não somente da aptidão/capacidade básica, mas também do conhecimento conceptual, procedimental e estratégico de realização da tarefa que a pessoa dispõe (Marañón, 1997). De acordo com a terminologia que adoptámos neste nosso trabalho, à aptidão/capacidade básica a que Marañón se refere, corresponde o termo capacidade, enquanto que a aptidão/capacidade efectiva corresponderá ao termo aptidão.

Para Carroll (1992), uma aptidão é vista como um traço latente ou uma característica dos indivíduos que se expressa em rendimentos diferenciados numa categoria de tarefas que variam nas suas exigências. Para caracterização das aptidões é importante especificar bem quais são as tarefas que vão ser utilizadas para as medir. A definição de aptidão está muito dependente do tipo de tarefa escolhida para a avaliar.

A definição da aptidão cognitiva envolve, naturalmente, a definição do conceito capacidade. Ao longo deste trabalho o conceito aptidão está directamente relacionado com a palavra inglesa *ability*, que para Carroll (1993, pp.8) "*... refers to the possible variations over individuals in the liminal levels of task difficulty (or in derived measurements based on such liminal levels) at which, on any given occasion in which all conditions appear favorable, individuals perform successfully on a defined class of tasks*". Esta definição de aptidão está

intimamente relacionada com a tarefa e com a marcação dos seus limiares. A avaliação do atributo a que a aptidão faz referência, está dependente da amplitude entre os limiares considerados para a tarefa. Poderíamos afirmar que o resultado do sujeito na realização de uma tarefa está directamente relacionado com os seus níveis de dificuldade, e não tão somente com a aptidão do sujeito para a sua execução. As tarefas variam em termos de dificuldade ou seja, em termos de probabilidade que os sujeitos têm de serem capazes de as realizar correctamente.

Falar de medição levanta sempre a problemática da estabilidade dos resultados ao longo do tempo. Pretende-se sempre saber se o resultado do sujeito na tarefa é possível de se voltar a verificar, caso todas as condições envolventes sejam mantidas. Por parte do sujeito será de esperar que a aptidão se mantenha ao longo do tempo, excepto em situações limite, como seja, a título de exemplo, sofrer um acidente.

O estudo das aptidões cognitivas tem sido desenvolvido sobretudo na perspectiva psicométrica, ou seja, de acordo com as regras da teoria factorial. O procedimento da análise factorial, como apoio à estruturação das aptidões cognitivas, continua a ser utilizado, não obstante as críticas que são feitas pelos cognitivistas. Carroll (1993) reanalisou, por intermédio da análise factorial exploratória, 460 conjuntos de dados retirados da literatura. Para ele o modelo hierárquico de três estratos será o melhor processo de estruturar as aptidões. No terceiro estrato teríamos um único factor (*g*); um conjunto de aptidões amplas no segundo; e um conjunto alargado de aptidões específicas no primeiro estrato. Investigações anteriores (Gustafsson, 1984, 1989; e Unheim & Gustafsson, 1987, citados por Carroll, 1992) já apontavam para esse modelo, à excepção de considerarem o factor *g* do terceiro estrato como equivalente a um factor de aptidão fluida do segundo estrato. Se, de acordo com Carroll (1992), o factor *g* não explica mais do que 50% da variância de uma determinada medida, será

de esperar que os tais factores mais específicos (que se situam no primeiro estrato) possam contribuir significativamente para a predição de um critério externo.

O atributo exigido para a correcta execução de uma tarefa cognitiva está directamente relacionado com a aptidão cognitiva a avaliar. Entenda-se por tarefa cognitiva aquela que exige um correcto ou apropriado processamento de informação mental para uma realização com sucesso (Carroll, 1993). Partindo desta conceptualização do que é tarefa cognitiva, será fácil hipotetizar que aptidão ou aptidões cognitivas estão directamente implicadas.

Para uma melhor compreensão dos conteúdos das aptidões cognitivas, é fundamental conhecer os processos cognitivos e os tipos de processamento de informação mental. Se entendermos o processo como uma acção ou um conjunto de acções nas quais algo é efectivado com vista a atingir um resultado, o processo cognitivo utilizará conteúdos mentais nas suas operações. Dentro destes conteúdos mentais é frequente fazer-se referência a representações ou codificações (*encodings*) tanto de estímulos externos como de imagens, conhecimentos, regras e material semelhante da memória a curto e a longo prazo. As respostas resultantes do processamento cognitivo podem ser mais explícitas ou mais encobertas (Carroll, 1993).

Os trabalhos desenvolvidos do ponto de vista psicométrico que envolvem conceitos da psicologia cognitiva podem ajudar a melhor categorizar as aptidões cognitivas. Um exemplo da investigação psicométrica que emprega conceitos de psicologia cognitiva é o trabalho de Kyllonen e Christal (1990) que envolve o estudo do papel da memória de trabalho. A memória de trabalho correlaciona-se de forma acentuada com a velocidade de processamento em testes de raciocínio. Carroll (1992) defende, no entanto, que o formato convencional (papel-e-lápis) em que as aptidões foram medidas, não permite, por intermédio da análise factorial, uma diferenciação clara entre a dificuldade e a velocidade de processamento de

informação. Acredita-se que a velocidade, variável capaz de ser avaliada por intermédio dos testes informatizados, pode também contribuir para um melhor conhecimento de como as aptidões estão estruturadas e interligadas em termos de funcionamento.

Sendo as tarefas cognitivas complexas, elas tendem a ser analisadas de forma parcelar, em processos, estádios ou componentes. Sternberg (1977, citado por Carroll, 1993), propõe a análise de determinadas tarefas cognitivas (como seja o exemplo de analogias) a partir de componentes, que a seguir se descrevem. Esta transcrição do processo cognitivo inclui já a componente comparação, acrescentada mais tarde por Sternberg e Gardner (1983):

- a. Codificação (*Encoding*): Processo de tradução de cada estímulo numa representação interna a partir da qual podem ser realizadas outras operações mentais.
- b. Inferência (*Inference*): Processo de descoberta de uma regra que permita relacionar os dois primeiros termos da analogia e correspondente armazenamento na memória de trabalho.
- c. Correspondência (*Mapping*): Descoberta de uma regra de um nível mais elevado, que permita a relação entre o 1º e o 3º termos da relação, armazenando o resultado na memória de trabalho.
- d. Aplicação (*Application*): Corresponde ao emprego da relação inferida do terceiro termo em cada alternativa de resposta.
- e. Comparação (*Comparing*): Processo de comparações sucessivas das várias alternativas de resposta em relação à resposta idealizada para o problema.
- f. Justificação (*Justification*): Componente processada ocasionalmente e que se

baseia na escolha de uma alternativa quando nenhuma das opções dadas parece totalmente correcta.

- g. Resposta (*Response*): Forma de o sujeito expressar a resposta que considera adequada.

A análise componencial de uma tarefa possibilita conhecer, com maior pormenor, o processo de resolução de um problema. É uma análise mais rica do que a extrapolação da capacidade do sujeito a partir do somatório de respostas correctas a um conjunto de questões.

Para validação desta teoria, Sternberg (1977) concebeu modelos experimentais que permitem estimar ou observar não só a correcção/precisão das respostas como também o tempo gasto no processamento. Carroll (1980b), na análise aos experimentos laboratoriais de Sternberg, refere que os resultados alcançados, apesar de utilizarem reduzidas amostras, podem ser generalizados, de alguma forma, para medidas de aptidão psicométrica, como as da componente rapidez de codificação e de resposta. Por sua vez, Almeida (1988) chama a atenção para as dificuldades que se colocam ao nível da análise componencial do acto inteligente "...algumas questões subsistem, por exemplo, se todas as componentes são referidas independentemente do grau de complexidade e de dificuldade das tarefas, ou do treino das capacidades ou do nível de desenvolvimento cognitivo dos indivíduos" (p.169). Uma investigação neste sentido aproxima-nos mais da metodologia experimental, por oposição à correlacional, onde a maioria dos estudos psicométricos têm sido desenvolvidos.

O estudo das capacidades/aptidões cognitivas, quer na perspectiva estrutural quer processual, passa forçosamente por uma boa caracterização da tarefa cognitiva. Neste sentido, Carroll (1993) propõe o modelo designado por Tarefa Cognitiva Elementar (*Elementary Cognitive Task*), numa tentativa de clarificar a noção de tarefa cognitiva. Para este autor, uma Tarefa

Cognitiva Elementar (TCE), corresponde a um largo leque de tarefas em que os resultados podem ser caracterizados a partir de um número relativamente reduzido de processos ou operações mentais. O sucesso dos resultados estará forçosamente dependente das instruções dadas ou dos planos ou conjunto de planos adaptados pelos sujeitos. Esta perspectiva sobre o acto cognitivo apresentada por Carroll é mais abrangente do que somente a análise do processo em si mesmo, pois engloba não só as operações processadas pelo indivíduo, como também um conjunto de operações relativas ao experimentador, como as instruções e os exercícios de treino, entre outros.

Como forma de representação da análise da TCE, Carroll (1980a) apresenta um esquema de desenvolvimento de uma tarefa apresentada ao sujeito e a correspondente decomposição das suas reacções cognitivas ao longo do tempo, e que designou por *Dual Time Representation* (cf. Quadro 1.1). Este esquema foi utilizado por Keating e Bobbitt (1978, citado por Carroll, 1993) para ilustrar a avaliação do tempo de reacção de escolha. Nesta tarefa a operação cognitiva em jogo consiste na percepção da cor de uma luz que é apresentada ao sujeito, e correspondente resposta por intermédio da pressão do botão associado aquela cor. O Tempo de Reacção está directamente relacionado com a tradução dos códigos associados às cores.

Da análise componencial de uma tarefa cognitiva extremamente simples e idêntica à descrita no Quadro 1.1, verificamos três aspectos principais. O primeiro diz respeito a que nem todas as componentes referidas por Sternberg (1977) e Sternberg e Gardner (1983) para análise da realização de uma tarefa cognitiva como a analogia, estão presentes noutra tipo de tarefas. Parece, assim, que a utilização das componentes descritas é activada em função da complexidade da tarefa cognitiva. O segundo aspecto prende-se com a representatividade das instruções para a realização da tarefa. Parece-nos que a tarefa descrita (tempo de reacção de escolha) está intimamente relacionada com a utilização da memória. Nesta tarefa o sujeito

somente tem de codificar o estímulo descrito (verde ou encarnado) e associar-lhe um código de resposta. O sujeito, neste caso, codifica duas relações. Uma para o estímulo luz verde, e outra para o estímulo luz encarnada. Quanto maior for o número de estímulos a apresentar, maior será o número de relações que o sujeito necessita de memorizar.

Quadro 1.1. Processo do tempo de reacção de escolha segundo o modelo de D. P. Keating e B. L. Bobbitt (1978).

Instruções:	Guarda em memória relações indicadas por código de luzes
1. Espera pelo estímulo	Reavivar na memória de trabalho os códigos que podem ser utilizados
2. O estímulo luz é apresentado	Codificação: Traduz o tipo de estímulo luz; Comparação: Compara os vários códigos do estímulo luz com os da relação anteriormente armazenada na memória; Justificação: Selecciona a primeira opção que é idêntica ao código da luz identificando a resposta a dar
3. Pressiona o botão de acordo	Resposta: Activa o movimento para dar a resposta pressionando o botão.

O número de relações que o indivíduo poderá conter em memória, está limitado, no essencial, pela sua capacidade mnésica e pelo tempo de que dispõe para o trabalhar. O tempo que é dado ao sujeito para incorporar cognitivamente as instruções de uma tarefa assume elevada importância para o sucesso na realização da tarefa cognitiva, sendo fundamental que o conteúdo das instruções seja idêntico para todos os sujeitos. O tempo necessário ao candidato para a memorização das instruções pode, também, constituir informação pertinente para o avaliador.

1.2. Memória a curto-prazo e memória de trabalho

A memória de trabalho consiste numa componente de armazenamento e de atenção, tendo como principal função a manutenção de representações mnésicas para fazer face a processamentos concorrentes, distrações e/ou mudanças de atenção (Baddeley & Hitch, 1974; Conway, Cowan, Bunting, Theriault & Minkoff, 2002). A memória de trabalho distingue-se da memória a curto-prazo na medida em que esta última tem uma função mais simples e que é a de funcionar como *buffer* de armazenamento em que a sua capacidade é determinada pela competência e estratégias desenvolvidas (Conway *et al.*, 2002). Nesta perspectiva a memória de trabalho tem uma função mais complexa associada ao processamento e à atenção.

Considerando a pertinência de ambas as memórias no modelo de processamento de informação (Almeida, 1988) e o modo como influenciam a execução de tarefas cognitivas, passaremos a apresentar alguns estudos que relacionam estas duas memórias com a inteligência e as aptidões cognitivas.

Drasgow e colaboradores (1993) distinguem duas tarefas que permitem a avaliação da memória a curto-prazo: Uma em que letras são apresentadas sequencialmente, após o que é apresentado ao sujeito uma letra que estava contida na série e que o sujeito terá de identificar o lugar que ocupava na série. Na outra, designada por tarefa de memória simultânea, é basicamente exigido o mesmo aos sujeitos, só que o conjunto a memorizar é apresentado simultaneamente. Estas tarefas apresentam elevado índice de fidelidade e quando cruzadas com a *performance* e o treino, nalgumas actividades profissionais, apresentam bons índices de validade preditiva (Carretta, 1987a, 1987b; Drasgow *et al.*, 1993).

Qual é, então, a função da memória a curto-prazo no processamento da informação cognitiva?

Segundo Estes (1982), ela tem uma função dupla. É um local temporário de armazenamento de informação que foi recentemente recebida e está a ser processada para posterior envio para a memória a longo-prazo. A outra função corresponde à recepção de informação proveniente quer do meio exterior quer da memória a longo-prazo, com o objectivo de ser trabalhada numa tarefa que esteja em curso.

O processamento de informação toma em consideração não somente a estimulação que vem do exterior, como também a tem de relacionar com a informação armazenada anteriormente. Parece haver, assim, uma relação entre a capacidade de memória a curto-prazo e a rapidez/eficiência no processamento de informação. Para Estes (1982), os indivíduos com maior capacidade de memória a curto-prazo processam a informação de forma mais rápida e eficiente, enquanto que os indivíduos mais limitados ao nível da capacidade de memória a curto-prazo terão de desenvolver estratégias, por ensaios sucessivos, para conseguir resolver os mesmos problemas. Se considerarmos os tempos de latência como uma medida de velocidade no processamento de informação, será de esperar uma relação positiva entre os tempos gastos pelos sujeitos na realização dos itens correctos e a rapidez na recuperação da informação para a memória a curto-prazo. Segundo Estes, será de esperar que quanto maior a capacidade da memória a curto-prazo, mais rapidamente se processa a informação.

O estudo de Miller e Vernon (1992) confirma a importância de, pelo menos, duas componentes distintas para o estudo da inteligência: Memória a curto-prazo e velocidade de processamento de informação. Segundo a investigação dos autores, a memória a curto-prazo é uma das variáveis que mais se relaciona com a inteligência. O facto de a correlação entre a inteligência e o tempo de reacção poder baixar até valores não significativos quando os efeitos da memória a curto-prazo são controlados, pode implicar que esta se comporte como uma variável mediadora da relação entre a inteligência e o tempo de reacção. Miller e Vernon

(1992) verificaram que os resultados dos tempos de reacção, testes de inteligência e memória a curto-prazo se encontram significativamente correlacionados. Para estes autores, a memória a curto-prazo tem um efeito moderador na relação entre inteligência e tempo de reacção. Para Weiss (1986) a velocidade de processamento de informação é mediada pelo espaço de memória (*memory span*). De acordo com a interpretação de Miller e Vernon (1992) à proposta de Weiss (1986), a capacidade de memória a curto-prazo, tal como medida pelo *memory span*, constituirá uma limitação do raciocínio à resolução de problemas. Quanto maior for a capacidade de memória a curto-prazo mais rápidos e eficientes serão os sujeitos a armazenar e a recuperar informação.

Pellegrino e Glaser (1979) identificaram a existência de uma relação entre a inteligência e a memória. Para estes autores, indivíduos com resultados elevados em provas de aptidão verbal estão mais aptos a trabalhar a informação na memória a curto-prazo, assim como a aceder à informação contida na memória a longo-prazo.

Autores como Whimbley, Fischhof e Silikowitz (1969) encontraram, inclusive, uma correlação elevada, de 0,77 entre uma prova de *memory span* e o cálculo numérico, sugerindo que este tipo de memória é essencial para um armazenamento efectivo, recuperação e processamento de informação.

Vernon, Nador e Kantor (1985b) também encontraram associações entre variáveis de velocidade de processamento e diferentes sub-testes da prova de inteligência que foi utilizada (*Multidimensional Aptitude Battery* [MAB] de Jackson, 1983). Os sub-testes verbais correlacionam-se mais fortemente com as variáveis de recuperação da informação contida na memória a longo-prazo, que as relacionadas com a memória a curto-prazo. Por outro lado, o teste avaliador da capacidade de armazenamento da memória a curto-prazo correlacionou-se de forma fraca com os testes verbais e de forma elevada com os de rendimento,

particularmente com os espaciais. No teste de aritmética encontrou-se coeficientes com idêntica magnitude tanto para as medidas das tarefas de memória a longo-prazo como para as de curto-prazo, o que é coerente com as hipóteses iniciais, já que o teste de memória aritmética tanto faz recurso aos processos de memória a curto-prazo como de longo-prazo.

Apesar do risco que se corre em generalizar os resultados unicamente a partir de dois estudos, o trabalho de Necka (1992) chama a atenção para duas conclusões gerais que considera pertinentes: Primeiro, os dados sugerem que as diferenças individuais relativas ao funcionamento da memória de trabalho, podem ser consideradas peças elementares do processamento cognitivo da informação. Os dados recolhidos neste estudo de Necka (1992) confirmam a posição empírica, de que, de acordo com Baddeley (1986), a memória de trabalho é considerada como um mecanismo de execução central (*central executive mechanism*) para o processamento de informação. A segunda conclusão que Necka (1992) considera pertinente no seu estudo relaciona-se com o facto de não ter encontrado uma tão forte correlação entre memória de trabalho e a velocidade no processamento de informação que leve a atribuir importância a este factor para o estudo da teoria da inteligência. Esta conclusão, retirada exclusivamente a partir de um tipo de velocidade, que é a que diz respeito ao exame (*scan*) dos elementos retidos na memória de trabalho, deverá ser analisada com cuidado. Qualquer avaliação da velocidade de processamento contém sempre uma elevada especificidade, e, nos estudos de Necka, a velocidade foi sobretudo avaliada ao nível das tarefas relativas à memória de trabalho.

Na concepção de Necka (1992), a inteligência é um fenómeno multifacetado em que a memória de trabalho joga um papel importante. Embora outros autores (Kyllonen & Christal, 1990) sugiram que a percentagem de variância explicada da inteligência pela memória de trabalho seja na ordem dos 80%, o estudo de Necka apenas chega aos 20%, o que, para este

autor, significa que o nível intelectual geral do sujeito é explicado em cerca de 20% pela sua capacidade da memória de trabalho. Nesta perspectiva, fica ainda uma grande percentagem de variância por explicar, a qual será devida não só a outros factores ao nível dos mecanismos cognitivos, como também aos não cognitivos e que afectam o rendimento nos testes de aptidão, como a ansiedade, motivação, autoconfiança, entre outros (Almeida, 1994; Stankov, 2000).

O conhecimento da natureza da memória de trabalho está ainda revestido de algum mistério. Para Weiss (1989) o espaço de memória está relacionado com a energia do metabolismo cerebral. Indivíduos com um QI mais elevado têm uma energia do metabolismo cerebral mais eficiente e, conseqüentemente, têm maior espaço de memória. Haier, Siegel, Nuechterlein, Hazlett, Wu, Paek, Browning e Buchsbaum (1988) dão suporte a esta teoria de carácter mais fisiológico, ao verificar no seu estudo que o metabolismo da glucose cortical está correlacionado negativamente com a *performance* nas Matrizes Progressivas de Raven na forma avançada. Já em 1974, Maxwell, Fenwick, Fenton e Dollimore (citado por Haier, 1990) tinham lançado a hipótese da eficiência cerebral, que se pode traduzir, de uma forma simplista, na seguinte frase: Os indivíduos com melhores resultados nos testes de inteligência geral ou com um funcionamento cognitivo eficiente necessitam de activar um número menor de neurónios para processar a informação. Estudos no âmbito das neurociências, ao utilizarem a tomografia por emissão de positrões (*Positron Emission Tomography* - PET) em sujeitos voluntários durante a realização de um teste de inteligência, como por exemplo as Matrizes Progressivas de Raven na forma avançada, reforçam a hipótese da eficiência cerebral (Haier, 1990). A correlação inversa encontrada entre os resultados nas Matrizes Progressivas de Raven na forma avançada e o funcionamento cerebral localizado, significa que quanto mais inteligente for o sujeito, menor é a rede neuronal activada. Apesar de a ideia de conjugação de esforços entre a utilização da alta tecnologia das neurociências e dos estudos factoriais com

testes psicométricos poder contribuir para o estudo da inteligência, o custo é ainda bastante elevado e conseqüentemente limitativo da sua realização.

Para Kyllonen (1994), a arquitectura cognitiva humana consiste na memória a curto-prazo e memória de trabalho, as quais contêm a matéria para processamento num dado momento, e duas memórias a longo prazo. Uma contém o conhecimento declarativo e a outra o procedimental, onde são armazenados os resultados do processamento anterior. O projecto *Cognitive Abilities Measurement* (CAM) foi desenvolvido no cruzamento de duas aproximações possíveis à teoria das aptidões cognitivas (Kyllonen, 1994): A abordagem *top-down* que assenta essencialmente na teoria, sendo um trabalho de confirmação; e a abordagem mais exploratória, mais centrada nos dados, comumente designada por *bottom-up*. O enquadramento do projecto CAM foi inspirado nas duas perspectivas, já que elas são consideradas complementares. Na perspectiva *top-down*, chegou-se ao consenso sobre o que é o modelo de processamento de informação humana.

Um outro modelo da memória de trabalho, dando maior ênfase à sua natureza dinâmica, designado por Baddeley-Hitch (Baddeley & Hitch, 1974), considera a existência de três componentes essenciais da memória a curto-prazo: O centro executivo, o *loop articulatory* e o *visuo-spatial sketchpad*, estes dois últimos comandados pelo primeiro.

Süß, Oberauer, Wittmann, Wilhelm e Schulze (2002) utilizaram uma bateria de 17 tarefas avaliadoras da memória de trabalho, que aplicaram a 128 jovens adultos juntamente com um teste de inteligência geral, tendo verificado uma elevada correlação entre a capacidade geral da memória de trabalho e a inteligência geral. Para estes autores a capacidade da memória de trabalho é o melhor preditor da inteligência no estado actual da pesquisa sobre a cognição. No entanto, a forte associação entre estas duas variáveis, apesar de permitir um melhor conhecimento dos constructos das aptidões psicométricas não permite explicar tudo, até

porque a memória de trabalho, só por si, não é um constructo perfeitamente definido. É largamente aceite que a capacidade da memória de trabalho é um recurso limitador significativo de uma cognição complexa. Contudo, as funções da memória de trabalho que afectam algumas componentes do processo cognitivo numa determinada tarefa de raciocínio ainda não estão perfeitamente compreendidas. Poder-se-á questionar se é a capacidade de armazenamento, a capacidade de processamento ou a combinação de ambas, ou outras como a coordenação dos conteúdos mentais, que limitam a realização de determinadas fases de resolução de um problema cognitivo.

Apesar da tentativa de dissociação entre a memória a curto-prazo e a memória de trabalho, para um melhor conhecimento da relação de ambas com a inteligência e o processamento de informação nas tarefas cognitivas, é importante considerar que nenhuma tarefa se pode considerar como uma medida “pura” quer da capacidade da memória a curto-prazo, quer da capacidade da memória de trabalho, segundo Conway e colaboradores (2002).

1.3. Velocidade mental

O interesse pelo estudo da velocidade de processamento de informação, como forma de incrementar o conhecimento das diferenças individuais ao nível da inteligência é antigo (cf. Almeida, 1994), embora se tenha verificado um reavivar do mesmo em tempos recentes, patente na quantidade de artigos recentemente publicados nas revistas especializadas. No ano de 2001, a revista *Intelligence* publicou um número especial dedicado ao Tempo de Inspeção como forma de comemorar os vinte e cinco anos de pesquisa da relação entre o Tempo de Inspeção e a Inteligência.

A investigação da velocidade mental, no seu nível mais elementar (a partir do Tempo de

Inspecção), enquanto medida indirecta da inteligência, teve a sua origem com a investigação de Nettelbeck e Lally (1976). Estes autores encontraram uma correlação significativa entre o tempo de inspecção e o nível de inteligência avaliado psicometricamente. De acordo com Petrill, Luo, Thompson e Detterman (2001), esta associação é uma descoberta importante para um melhor conhecimento dos processos cognitivos e da associação possível dos processos biológicos como suporte às diferenças da inteligência humana.

Para além do Tempo de Inspecção, o Tempo de Reacção também é frequentemente referenciado como uma variável pertinente no estudo da velocidade mental. Enquanto o Tempo de Inspecção corresponde ao tempo que o sujeito demora a diferenciar estímulos semelhantes, o Tempo de Reacção corresponde ao tempo de resposta do sujeito a um estímulo. Ainda ao nível da velocidade mental é comum considerar-se o Tempo de Reacção de escolha e o Tempo de Latência às respostas. O Tempo de Reacção de Escolha é uma variável de velocidade em que o sujeito tem de efectuar uma opção, tal como no Tempo de Inspecção, mas em que os estímulos não são tão similares.

Apesar da quantidade de informação ser substancialmente reduzida nas tarefas de Tempos de Reacção, elas não são, segundo Almeida (1994), exclusivamente neurofisiológicas já que estão enriquecidas cognitivamente por envolverem a compreensão de instruções, motivação, familiaridade com o equipamento, acuidade sensorial e estratégias diversas.

Dentro dos vários investigadores que estudaram o Tempo de Reacção enquanto medida indirecta da inteligência, destaca-se Arthur R. Jensen. Para ele, o Tempo de Reacção é uma medida que nos permite um melhor conhecimento teórico da inteligência (Jensen, 1982b). Os estudos de Jensen e de outros investigadores centraram-se essencialmente na medição do tempo de resposta a estimulações visuais simples (Tempo de Reacção Simples) e de escolha (Tempo de Reacção de Escolha).

A maioria dos estudos confirmaram o paradigma de Hick em que se pretende medir o tempo de reacção do sujeito a responder a um estímulo simples ou complexo. Esta classificação de simples ou complexo é uma função da quantidade de informação mas também do número de respostas que o sujeito tem como alternativas. De acordo com o paradigma de Hick, os tempos gastos pelos sujeitos a responder a uma estimulação visual aumenta de forma logarítmica à medida que aumenta a quantidade de informação que o sujeito tem de processar para escolher a resposta a dar. Do ponto de vista da validade preditiva relativa ao rendimento no trabalho em determinadas funções, as medidas baseadas no Tempo de Reacção, assim como os testes avaliadores da rapidez em termos perceptivos e psicomotores, apresentam coeficientes elevados (McHenry, Hough, Toquam, Hanson & Ashworth, 1990).

Para Necka (1991a), o facto de o indivíduo ser inteligente não implica que seja rápido em termos globais, independentemente do tipo de tarefa e do nível de análise. A relação entre os tempos de reacção e a inteligência está dependente da complexidade da tarefa. Quanto mais complexas as tarefas a desenvolver, mais elevados são os coeficientes de correlação entre as duas variáveis (Vernon & Kantor, 1986).

De acordo com os estudos de Jensen (1982b) e de Necka (1991b, citado por Necka, 1992), os indivíduos inteligentes não são somente rápidos como também são mais estáveis nas suas reacções, ou seja, as variações intra-individuais dos seus tempos de reacção são relativamente baixas. Esta apreciação refere-se aos tempos de reacção avaliados na realização de tarefas psicomotoras elementares. No entanto, Vernon, Nador e Kantor (1985a), num estudo com a utilização de diferentes medidas cognitivas, também consideraram os desvios-padrão intra-individuais como outro parâmetro necessário à eficiência do sistema de processamento de informação. Eles correlacionaram-se, com a mesma ou maior magnitude que as médias dos tempos de reacção, com as medidas de inteligência, pelo que poderão ser considerados como

melhores estimativas. As médias e a variabilidade intra-individual dos tempos de reacção estão contudo fortemente correlacionados, pelo que se pode sugerir que ambas as medidas são necessárias a um sistema eficiente de processamento de informação. Aliás a existência de um factor geral de medidas de tempo de reacção, conseguido através da análise factorial, correlaciona-se de forma significativa com o factor geral intelectual, conseguido através de testes psicométricos como seja o caso das Matrizes Progressivas de Raven na forma Avançada (Vernon, 1989).

Os tempos que irão ser considerados nos nossos estudos dizem respeito ao Tempo de Latência, que difere do Tempo de Reacção de Escolha. Este último diz respeito, essencialmente, a estímulos visuais com fraca carga intelectual. Por sua vez, os tempos de latência compreendem o tempo desde a apresentação do estímulo, até ao registo da resposta, incluindo todo o tempo gasto na escolha da solução apropriada, de acordo com as instruções previamente assimiladas.

1.3.1. *Trade-off* entre correcção e rapidez

A investigação ao nível das diferenças individuais, no que diz respeito aos processos cognitivos, tem-se debruçado sobre a relação entre a correcção e a rapidez no desempenho de determinadas tarefas. A psicologia cognitiva está essencialmente interessada nos processos internos utilizados pelo sujeito para dar a resposta aos problemas, porque é aprofundando o conhecimento dos processos inerentes à realização cognitiva que melhor se poderá conhecer como os sujeitos se diferenciam entre si. Pretende-se conhecer as diferenças cognitivas entre os indivíduos, a partir das diferenças verificadas nas estratégias utilizadas.

Torna-se extremamente difícil o estudo das aptidões cognitivas segundo a perspectiva

cognitivista, considerando exclusivamente a variável correcção, pelo que os testes de aplicação convencional (papel-e-lápis) pouco podem ajudar neste campo. Com o recurso a sistemas informáticos e a consequente informatização de algumas tarefas cognitivas, é possível chegar-se um pouco mais longe. Pelo menos, essa expectativa já está criada. A execução dos testes em formato informático permite a medição do tempo de resposta a um dado item, o que lançou um novo ânimo no estudo das diferenças individuais, mas acrescentou uma dificuldade. A velocidade na resposta não está totalmente dissociada da correcção. Os aumentos em termos de rapidez são inversamente proporcionais à correcção. Quanto mais aumenta a velocidade mais diminui a correcção e vice-versa. Poderíamos dizer que existe uma curva teórica de associação entre estas duas variáveis (cf. Figura 1.1)

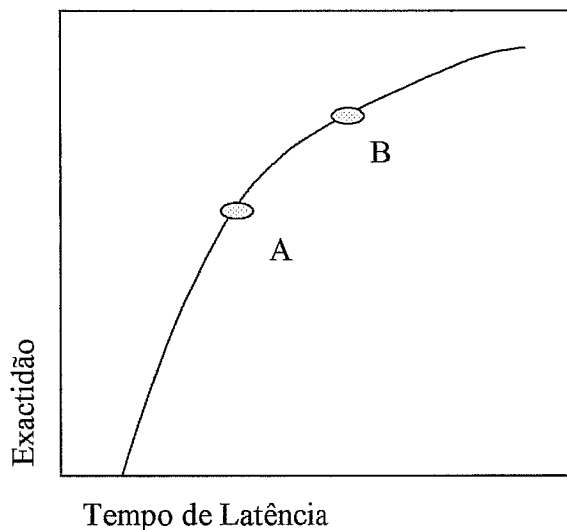


Figura 1.1. Curva hipotética da relação Exactidão-Tempo de Latência. O sujeito B é mais lento e mais exacto que o sujeito A (retirado de Lohman 1994b, p. 159).

Encontramos os *trade-offs* entre velocidade e correcção tanto no mesmo sujeito ao longo da

realização da mesma tarefa, como entre sujeitos. Contudo, o cansaço ou o ganho de confiança podem alterar a relação entre a velocidade e a correcção para o mesmo sujeito, partindo do princípio que o grau de dificuldade se mantém constante ao longo da prova. A diferença entre sujeitos é mais facilmente aceite porque pode ser justificada por características não cognitivas dos mesmos, nomeadamente, a impulsividade e a ponderação. Esta divergência leva a que exista bastante dificuldade na comparação intelectual dos sujeitos com base exclusivamente nos tempos de latência a uma determinada tarefa. Assim, não é possível deduzir se existe *trade-off* entre a correcção e a velocidade a partir, exclusivamente, da existência ou não de correlação entre estas duas variáveis como é claramente explicado por Lohman (1994b).

Um outro problema sobre esta temática deriva do facto de nos basearmos nos tempos de latência gastos nas respostas erradas, o que não nos permite fazer grandes extrapolações para o comportamento dos sujeitos nas respostas dadas correctamente.

Quando se tenta correlacionar as medidas de correcção e de rapidez com os factores que lhe estão subjacentes, verifica-se que as correlações tendem a diminuir à medida que aumenta o grau de dificuldade, quando se trata da rapidez. O inverso ocorre com a correcção das respostas. À medida que aumenta o grau de dificuldade da tarefa, as correlações entre o número de respostas correctas e o factor avaliado tende a ser maior (Lohman, 1989). Lohman conclui que os tempos de latência devam apenas ser considerados relativamente aos itens sem erro.

Se, por um lado, podemos concluir que a rapidez e a correcção são dois aspectos diferentes da *performance* do sujeito, não se sabe muito da natureza do que está subjacente ao controlo da compensação entre eles. Lohman (1994b) defende que as estatísticas multivariadas, nomeadamente a correlação canónica, podem ajudar a compreender melhor o fenómeno do *trade-off* entre a correcção e a rapidez. No entanto, e considerando as desvantagens destes

modelos, nomeadamente de as relações não serem lineares, podem ser tomadas em consideração outras formas de abordagem do *trade-off*.

Uma forma de compreensão do fenómeno de *trade-off* entre correcção e rapidez é através do cálculo de uma nota que possa incluir as duas dimensões, de acordo com uma fórmula que minimize a influência do *trade-off* entre velocidade e correcção (Dennis & Evans, 1990, 1991, citados por Lohman, 1994b). De acordo com Lohman (1994b) só faz sentido arranjar uma nota única a partir da velocidade e correcção se estas duas dimensões forem consideradas partes integrantes do domínio que está a ser avaliado. Se reflectem diferentes aspectos da aptidão, então não existem maneiras de as combinar.

Uma outra forma de análise será através da representação gráfica da curva. Wickelgren (1977, citado por Lohman, 1994b) apresentou uma fórmula que permitia traçar a curva do *trade-off* entre a correcção e a latência de resposta de cada sujeito (cf. Figura 1.2). Este modelo é uma função de três parâmetros: Assíntota (λ); intercepção (δ); e a curvatura (β).

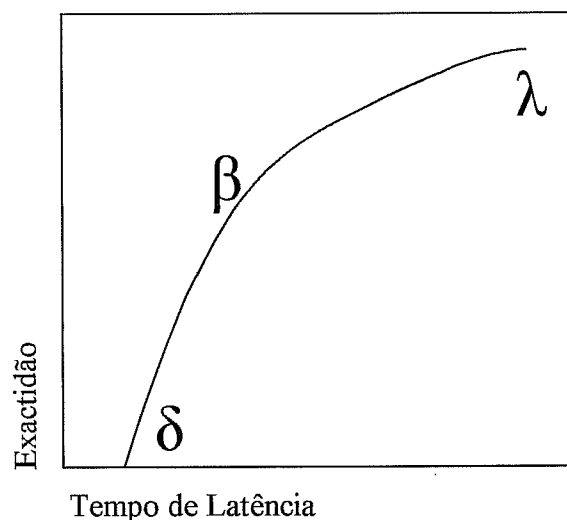


Figura 1.2. Parâmetros da curva hipotética Exactidão-Tempo de Latência, segundo W. A. Wickelgren (1977).

A assíntota (λ) indica-nos o nível de exactidão do sujeito, quando não existe limite de tempo. A intercepção (δ) indica o ponto a partir do qual a curva se forma sem que haja lugar a respostas correctas respondidas aleatoriamente. O terceiro parâmetro, corresponde à curvatura (β). Uma curvatura muito íngreme representa um rápido incremento em termos de exactidão para cada unidade adicional de tempo de processamento.

O modelo em cascata de McClelland (1979) traduz a maneira como a activação é acumulada em diferentes níveis num sistema hierarquicamente ordenado. Para ele os parâmetros assumem um significado diferente. O ponto chave neste sistema de processamento de informação é a activação. A intercepção da curva reflecte o tempo gasto pelos processos relativamente rápidos do sistema. A curvatura é dominada pelos processos mais lentos do sistema. O aumento de taxa de execução deste processo torna a curva mais escarpada; quando se baixa a taxa de execução, a curva fica mais plana. A assíntota, por sua vez, reflecte o nível final de activação alcançado. É um outro modelo de interpretação dos três parâmetros que são comumente referenciados nos estudos relativamente à velocidade-exactidão.

Para Lohman (1994b), consegue-se ter uma imagem muito mais realista do rendimento do sujeito se considerarmos os três parâmetros como pontos limite da curva: (1) o ponto de intercepção que corresponde à quantidade mínima de tempo requerida para processar a informação disponível com vista à formulação de uma resposta não aleatória; (2) a assíntota que corresponde ao outro extremo, ou seja, ao ponto onde a exactidão atingiu o limite; e, (3) a curvatura que será o ponto intermédio, que corresponderá ao ponto de transição entre os dois extremos definidos anteriormente.

A execução de uma tarefa cognitiva com tempo limite, onde o sujeito erra por falta de tempo e não por incapacidade própria, influencia os resultados obtidos nos estudos sobre esta temática, pelo que seria positivo que o sujeito fosse avaliado em distintas situações,

permitindo uma imagem mais real das suas capacidades.

Na perspectiva da análise do treino, a curva de rendimento entre a correcção e a rapidez, poderá dar informações muito úteis relativamente ao efeito da prática na aquisição de conhecimentos. Modificações ao nível da curvatura poderão traduzir a rapidez como esses conhecimentos poderão estar a ser assimilados. Lohman (1994b) atribui grande importância ao parâmetro curvatura por ser o indicador com maior utilidade para o conhecimento da *performance* do sujeito, embora também refira que necessita de melhor conhecimento da sua consistência interna, pois é para ele o elemento diferenciador entre indivíduos.

A informação sobre o modo como os sujeitos resolvem uma tarefa, tem um valor meramente informativo quando as próprias tarefas admitem uma variedade de estratégias de solução que podem ser sistematicamente relacionadas com uma aptidão ou um outro constructo que pretendemos compreender. Por este facto, uma análise cognitiva mais minuciosa do desempenho poderá dar um contributo importante tanto na construção dos testes de aptidões, como na interpretação dos seus resultados. Conhecendo as possíveis estratégias de resolução dos itens dos testes de aptidão, podemos efectuar uma interpretação mais sustentada do rendimento do sujeito, que vai para além de uma análise quantitativa dos resultados obtidos. Também pode contribuir para a criação de novas provas com a inclusão de itens com diferentes níveis de complexidade cognitiva. Por sua vez, um melhor uso dos tempos de latência e dos erros, pode ser conseguido por intermédio do modelo de rapidez-exactidão, nomeadamente com a utilização dos modelos de curva. Estes modelos podem ser úteis, tanto para os problemas práticos da selecção e classificação, como na compreensão do que se está a avaliar, ou do que pensamos estar a avaliar (Lohman, 1994b).

1.3.2. Complexidade da tarefa

A maioria dos estudos sobre a relação entre a velocidade mental e os vários aspectos da inteligência incluem tarefas com informação reduzida, o que dá suporte às teorias de inteligência baseadas nos aspectos biológicos. Estas teorias dão ênfase às propriedades da transmissão de sinal do sistema nervoso em vez dos aspectos relacionados com o processamento cognitivo (Eysenck, 1986). No entanto, a complexidade cognitiva das tarefas utilizadas na avaliação da inteligência e das aptidões cognitivas por intermédio da velocidade de processamento de informação é uma variável que forçosamente tem de ser tomada em consideração.

Hunt (1980, citado por Larson, Merritt & Williams, 1988) afirma que as tarefas de processamento de informação simples não conseguem explicar mais do que 10% da variância de um teste de inteligência geral. Por outro lado, estudos como o de Cohn, Carlson e Jensen (1985) evidenciam forte associação entre o aumento de complexidade da tarefa e o rendimento intelectual dos sujeitos. Cohn e colaboradores (1985) encontraram um coeficiente de correlação de 0,94 entre a complexidade das tarefas de tempo de reacção (indicadas com latência média) e a magnitude das diferenças entre grupos em termos de rendimento. Uma vez que as tarefas mais complexas exigem maior manipulação mental do seu conteúdo, a eficiência na sua resolução poderá ser explicada pela capacidade da memória de trabalho. Os grupos com maior capacidade de memória de trabalho terão maior sucesso na execução das tarefas mais complexas.

Larson e colaboradores (1988), ao utilizarem três medidas de tempo de reacção com diferentes níveis de complexidade, verificaram que todas elas se correlacionam com os resultados obtidos nos testes psicométricos de inteligência. Analisando os resultados isoladamente, poderíamos afirmar que o estudo reforça o ponto de vista de Galton sobre a

inteligência, já que a tarefa mais simples do tempo de reacção envolvia um processamento mínimo de informação - o sujeito estava instruído para premir a barra de espaços do teclado de um computador IBM PC/XT quando um quadrado apresentado no ecrã ficasse iluminado. No entanto, nas restantes duas tarefas, em que o sujeito tinha de optar entre duas teclas, porque a tarefa proporcionava mais que dois tipos de estímulos e duas respostas possíveis, verificou-se que a correlação destes resultados com as provas de inteligência eram superiores à primeira. Quanto maior for a complexidade da tarefa apresentada ao sujeito, maior é a probabilidade de encontrarmos uma correlação elevada entre o tempo médio gasto na sua realização e o nível intelectual medido por um teste psicométrico de inteligência geral.

Se bem que a complexidade da tarefa no estudo de Larson e colaboradores (1988) tenha sido operacionalizada pelo aumento da probabilidade de aparecimento de múltiplos estímulos e possibilidades de respostas diferentes, torna-se necessário clarificar o que é a complexidade da tarefa em termos de exigências cognitivas. Para Marshalek, Lohman e Snow (1983), a compreensão da complexidade é um factor-chave para um melhor conhecimento da teoria da inteligência. As tarefas avaliadoras da velocidade de processamento têm a sua complexidade relacionada com as exigências em termos de codificação e a rápida manipulação dos dados na memória de trabalho. Esta ideia da pertinência da memória de trabalho para a inteligência é também defendida por outros autores, entre os quais Vernon (1986).

De acordo com Sanders (1983, citado por Larson *et al.*, 1988) poderíamos considerar duas grandes correntes para o estudo da complexidade da tarefa: O modelo das estruturas cognitivas e o dos recursos de energia. No modelo das estruturas cognitivas e tendo por base os trabalhos de S. Sternberg (1969), R. Sternberg (1980), e de Carroll (1976), referidos por Larson e colaboradores (1988), estão incluídos os conceitos de etapas de processamento, componentes e operações cognitivas elementares, ou seja, as unidades funcionais da cognição.

Desta forma, a complexidade da tarefa estaria hipoteticamente relacionada com a quantidade de operações cognitivas requeridas para a sua execução (Larson *et al.*, 1988). Ainda nesta perspectiva da complexidade da tarefa, g corresponderia ao universo de operações cognitivas. No entanto, de acordo com Larson e colaboradores (1988) e numa abordagem estrutural, o problema situar-se-ia não só na quantidade de operações cognitivas mas também no tipo de operações cognitivas requeridas.

O modelo dos recursos de energia apresenta uma ligação mais directa entre as exigências da tarefa e os recursos disponíveis para a enfrentar. Partindo do princípio que o espaço disponível para o processamento inerente à resolução do problema (que será, segundo Larson *et al.*, [1988] a memória a curto-prazo ou a memória de trabalho) tem uma capacidade limitada, o sujeito falha ao realizar o problema quando a complexidade deste (composto pelo número de elementos considerados e suas transformações) é superior à sua capacidade. Esta ideia está associada à sobrecarga de trabalho, em que o rendimento do sujeito baixa por a reserva de recursos estar exausta.

O modelo lançado por Vernon e colaboradores (1985b) é compatível com a hipótese de que a correlação entre inteligência e rapidez de processamento tende a aumentar à medida que aumenta a complexidade da tarefa em que se está a medir o tempo de reacção. Por outras palavras, poderíamos afirmar que quanto mais exigente for a tarefa em termos de quantidade de informação a armazenar e a processar, mais ela se aproxima dos itens que compõem os testes de inteligência. Poderíamos, assim, deduzir que, se os tempos de reacção fossem substituídos pelos tempos de latência nas respostas aos itens do teste de inteligência (instrumento de avaliação do nível intelectual dos sujeitos), seria de esperar correlações mais elevadas entre o resultado global e a média dos tempos de latência nos itens mais complexos do que nos menos complexos.

A complexidade de uma tarefa é, por vezes, confundida com a dificuldade da mesma. Stankov (2000) considera o conceito de complexidade cognitiva baseado em dois pontos base: (a) estar directamente relacionado com a carga que o teste tem de um determinado factor cognitivo relativamente abrangente (e.g., inteligência fluida); e (b) que é possível manipular as exigências da tarefa de modo a afectar a carga dos testes, logo avaliando empiricamente a eficácia de tais manipulações. Segundo este autor, é possível diferenciar a complexidade da dificuldade da tarefa. No contexto do seu estudo, a dificuldade refere-se à *performance* global no teste por um conjunto de pessoas e que se expressa por uma média aritmética. De acordo com esta perspectiva, é possível que testes de graus de complexidade diferentes sejam de igual grau de dificuldade. Por oposição, alguns testes psicológicos podem diferir em dificuldade sem que sejam diferentes ao nível da complexidade. Enquanto o conceito de dificuldade está mais associado à performance do grupo, o de complexidade está fortemente associado às características dos itens que o compõem. Na tentativa de melhor clarificar a relação entre estes dois conceitos, poderíamos referir que a dificuldade de um item resulta da forma como uma amostra de sujeitos lida com o seu grau de complexidade.

1.3.3. Limitação do tempo de aplicação

A determinação do tempo de aplicação de um teste de aptidão cognitiva corresponde a uma fase do desenvolvimento do mesmo, e tem por objectivo dotar o mesmo da qualidade métrica sensibilidade, ou seja, de permitir que ele consiga discriminar os sujeitos de acordo com o atributo psicológico que está a ser avaliado. A forma de cálculo do tempo de aplicação de uma prova está dependente do que está a ser avaliado e de como se pretende avaliar (Kline, 1986). Cronbach (1990) divide os testes de aptidão em dois grupos, consoante o tempo de aplicação: Os testes de rendimento e os de velocidade. Os testes de rendimento em que se pretende

avaliar uma aptidão e em que o resultado do sujeito deverá depender da capacidade que ele possui e não propriamente do tempo disponível. Um exemplo de testes de rendimento é as Matrizes Progressivas de Raven, aplicadas sem tempo limite, ou com um tempo que garanta que cerca de 75% da população consiga chegar ao fim do mesmo. Por outro lado, os testes de velocidade são geralmente formados por questões com baixo grau de complexidade onde se pretende que os sujeitos sejam diferenciados em função da velocidade. Estes últimos têm geralmente um tempo de aplicação inferior ou igual a 10 minutos, enquanto que os de rendimento são aplicados com um tempo superior a dez minutos.

A investigação relativa aos testes aplicados com tempo limite ou sem tempo limite em termos de exactidão e rapidez tem produzido resultados controversos. Nesta secção iremos apresentar alguns trabalhos que analisaram a relação da limitação do tempo de aplicação nos resultados obtidos pelos sujeitos ao nível da velocidade de processamento de informação, e da quantidade de respostas dadas correctamente.

Vernon e colaboradores (1985b) aplicaram um conjunto de testes avaliadores de tempos de reacção, que mediam a velocidade com que realizavam determinadas tarefas cognitivas elementares, e um grupo de testes de inteligência com tempo limite e sem tempo limite, a um conjunto de oitenta e um estudantes universitários. Verificaram, por intermédio da regressão múltipla, que a velocidade de processamento de informação tanto é boa preditora do sucesso nas provas de inteligência com tempo limite como em provas similares sem tempo limite. Os coeficientes foram aproximadamente equivalentes.

De acordo com os estudos de Kyllonen e Christal (1990) e Vernon e colaboradores (1985b), a velocidade de processamento traz duas vantagens para o sujeito mais inteligente ao realizar um teste de aptidão mental com dificuldade crescente. Nos itens iniciais (geralmente mais fáceis) o indivíduo acede mais rapidamente à informação necessária e responde mais

rapidamente às perguntas. Nas questões mais difíceis, os sujeitos mais inteligentes ao possuírem maior capacidade de memória de trabalho, são mais rápidos a processarem informação sem que a capacidade limitada entre em sobrecarga.

Se considerarmos a velocidade de processamento como uma medida indirecta de inteligência, será esperado que as pessoas com uma velocidade de processamento mais elevada (avaliada através dos tempos de latência) sejam capazes de responder rapidamente a um número maior de questões durante um teste de inteligência com tempo limite. A maior rapidez em termos de processamento também pode ser uma vantagem na realização de testes de rendimento sem tempo limite, uma vez que os sujeitos lidam melhor com as maiores exigências em termos de armazenamento e processamento dos itens mais complexos.

Esta análise progressiva da evolução dos sujeitos mais inteligentes ao longo de um teste de inteligência com tempo limite só faz sentido se o teste estiver bem construído, o que implica necessariamente uma correcta determinação do tempo de aplicação. Se o tempo limite for muito superior ao necessário para se realizar a globalidade do teste, está a introduzir-se uma nova variável. O sujeito ao verificar que lhe faltam poucos itens para terminar o teste e se sabe que ainda tem tempo, pode alterar a estratégia utilizada até então (dar as respostas o mais rapidamente possível) e passar a gastar mais tempo para se assegurar da resposta que vai dar. O inverso também é verdadeiro. Um tempo de aplicação significativamente inferior ao necessário pode alterar a estratégia de resolução dos exercícios, se souber o tempo disponível. Por exemplo, se um sujeito souber que lhe falta pouco tempo para terminar o teste, e que ainda lhe faltam alguns itens para responder e que as respostas erradas não descontam nas respondidas correctamente, pode alterar a estratégia, passando a arriscar mais e não gastando tanto tempo na confirmação da resposta.

O conhecimento antecipado do tempo de duração do teste, ou o tempo restante, são outras

condições a ter em consideração quando se estuda a limitação do tempo de aplicação na realização de provas mentais. Nalguns testes informatizados, o tempo restante é apresentado no ecrã. No entanto, no estudo de Vernon e colaboradores (1985b), o facto de considerarem, num dos testes, com tempo limite, apenas os primeiros cinco minutos de prova, poderá falsear o estudo já que os primeiros cinco minutos vão permitir responder aos itens mais fáceis, considerando que estão ordenados pelo grau de dificuldade. De acordo com o procedimento descrito neste estudo, os sujeitos davam as respostas nos primeiros cinco minutos com um lápis de uma determinada cor e completavam o restante teste com lápis de outra cor.

Na parte prática deste trabalho foram analisadas as diferenças em termos de velocidade de processamento de informação quando se comparam sujeitos com diferentes capacidades cognitivas. Neste sentido, colocou-se a hipótese de que sujeitos com diferentes capacidades intelectuais gerais obtêm idênticos resultados em termos de tempos de latência quando os itens são respondidos por um número considerável de sujeitos – itens fáceis. Se os itens fáceis forem pouco exigentes em termos de informação a ser trabalhada (baixa complexidade), então não haverá diferenciação estatisticamente significativa entre os sujeitos com diferentes níveis de capacidade cognitiva. Desde que a capacidade de armazenamento/processamento de informação na memória de trabalho dos sujeitos mais inteligentes e dos menos inteligentes, seja suficiente para lidar com a informação exigida pelo item, ambos gastam idêntico tempo na sua resolução. A velocidade superior de processamento pelos sujeitos mais inteligentes não cria vantagem significativa quando se está na presença de itens fáceis. Se a hipótese por nós colocada estiver correcta, podemos deduzir que nos itens em que se verifica uma elevada taxa de respostas correctas, os tempos médios de resposta nestas questões podem dever-se a outros factores como, por exemplo, motivação para a tarefa.

Ainda dentro desta linha de pensamento, podemos concluir que os itens mais complexos, ou

seja, aqueles que têm uma maior carga de informação ou maior número de alternativas de resposta, são os que suscitam uma estratégia de resolução menos linear. Estes itens tenderão a provocar maior diferenciação entre os sujeitos inteligentes e os menos inteligentes, em termos do tempo de latência nas respostas. Os sujeitos mais inteligentes, ao possuírem maior capacidade de memória de trabalho conseguem ser mais rápidos a determinar a solução correcta, ficando com mais tempo disponível para realizar um maior número de itens, o que logicamente lhes aumenta a probabilidade de resolver mais itens correctamente, obtendo melhores resultados no final. Esta hipótese está de acordo com Vernon e colaboradores (1985b) que atribuem vantagem aos indivíduos com velocidades de processamento mais elevadas na realização de um teste não limitado no tempo. Neste tipo de provas, como os últimos itens do teste são geralmente os mais difíceis, somente os indivíduos com um eficiente sistema de processamento de informação conseguem dar respostas correctas.

Um outro parâmetro para definir a eficiência do sistema de processamento de informação é a consistência da rapidez de processamento, medida pela variabilidade dos tempos de reacção individuais. Os desvios-padrão geralmente correlacionam-se com as medidas de inteligência com uma magnitude tão elevada como a alcançada com as médias dos tempos de latência (Vernon *et al.*, 1985b). O facto de as médias dos tempos de reacção e os desvios-padrão intra-individuais estarem fortemente correlacionados entre si, pode significar que ambos são necessários à eficiência do sistema de processamento. Para os mesmos autores, os sujeitos que possuem uma elevada velocidade de processamento e menor variabilidade interna, são os mais aptos a resolverem os problemas mais complexos e mais rapidamente.

1.4. Contributos da psicologia cognitiva na avaliação das aptidões cognitivas

Lohman (1994b), ao tentar sintetizar o esforço dispendido durante quase duas décadas para aplicar os conhecimentos da psicologia cognitiva à avaliação psicológica, considera que o impacto não foi tão significativo como seria esperado. Para este autor, muita da pesquisa relativa ao processamento de informação nas aptidões é orientada por duas assunções base, das quais considera a primeira falsa e a segunda induz em erro.

De acordo com a primeira, somos capazes de decompor todas as diferenças individuais envolvidas na realização de uma tarefa ou de um factor de aptidão em componentes quantificáveis em termos de processos mentais. O que pressupõe que seria possível identificar todos os processos mentais envolvidos na realização de uma tarefa, o que iria permitir estimar as diferenças individuais relativamente a cada um deles.

A maior parte da investigação efectuada no âmbito das aptidões, de acordo com o paradigma do processamento de informação, procurou decompor as diferenças individuais em constructos mais abrangentes tais como a aptidão verbal, ou o raciocínio, em unidades baseadas no processamento. Ainda segundo Lohman (1994b), algumas dessas componentes não se decompõem, não podendo, assim, explicar as diferenças individuais no conjunto do rendimento em tais tarefas. Esta orientação pode ser útil, mas não explica as diferenças individuais na tarefa.

A segunda suposição é a de que os testes psicométricos proporcionam boas tarefas cognitivas, do mesmo modo que tarefas cognitivas produzem bons testes psicométricos. A perspectiva do processamento de informação sobre as diferenças individuais diz respeito à forma como o sujeito processa a informação. Só que, quando se tenta operacionalizar as tarefas cognitivas, elas são decompostas em elementos tão simples, que se perde a noção da estratégia utilizada

pelo sujeito para resolver a tarefa em termos globais. Há, assim, uma redução da qualidade das diferenças em termos da estratégia utilizada. Se se pretende, com base no paradigma do processamento de informação, descobrir informação interessante sobre as diferenças individuais no processamento cognitivo, então a análise do processamento de informação deverá ser aplicada a tarefas desenhadas explicitamente para a identificação das diferenças relativas a conhecimentos ou estratégias.

A grande promessa do estudo das aptidões, pela abordagem do processamento de informação, era a de que iria permitir uma forma de operacionalizar as noções vagamente delineadas nos processos mentais em modelos que pudessem dar a entender como eles se combinam de forma a permitir aos sujeitos produzir as respostas.

Um dos modelos centra-se nas notas das componentes, as quais podem ser estimadas de vários modos. O método mais simples consiste em subtrair à média do rendimento do sujeito num tipo particular de tarefa, a média desse mesmo sujeito num outro tipo de tarefa. Podem também serem utilizados esquemas mais complexos, baseados no tipo do estímulo apresentado e no tempo gasto na resposta dada pelo sujeito. No exemplo referenciado em Sternberg (1977), os sujeitos executam uma tarefa de rotação mental onde lhes são apresentados dois estímulos que diferem em orientação. Estes devem determinar como os dois estímulos podem convergir. Shepard e Metzler (1971) consideram que os sujeitos confrontados com este tipo de problemas têm imagens mentais do estímulo, efectuem a rotação delas numa distância requerida, comparam-nas e respondem. O modelo foi testado por estes autores através da regressão, em que analisaram o efeito da latência da resposta na separação angular entre os estímulos. O declive da recta da função dá-nos uma estimativa do ângulo em que o estímulo foi rodado, e a intercepção o tempo gasto na codificação, comparação e processo de resposta. A utilização do peso b (no modelo de regressão) pode

servir como medida da componente processo de rotação, para depois correlacionar a nota desta componente com outras variáveis (Sternberg, 1977).

Lohman (1994b) considera ainda uma terceira suposição no contexto de aproximação entre a psicologia cognitiva e a psicometria, que se refere à interdependência entre dois aspectos da *performance*: Os erros e o tempo de latência. Esta assunção significa que é possível explicar as diferenças individuais num dos aspectos medindo o outro. Para Lohman (1994a) a latência da resposta e o erro de resposta estão interligados de um modo não linear uma vez que representam diferentes aspectos do rendimento do sujeito, havendo razões teóricas e práticas para explicarem o modo como elas covariam para cada avaliado. Ainda segundo este autor, as notas combinadas em termos de erros e tempo constituem melhores indicadores de *performance* dos sujeitos que as mesmas consideradas isoladamente. Torna-se necessário equacionar modelos de conjugação destes dois elementos para que providenciem uma nota que seja um bom indicador das diferenças individuais. Os tempos de latência tendem a aumentar nos itens ou tarefas mais complexas, porque os testados têm maior dificuldade em manter a mesma taxa de erros que nos itens mais fáceis. As equações de regressão poderão fornecer estimativas das diferenças individuais. A intercepção, que corresponde à componente residual, poderá ser o centro da variância das diferenças individuais que é consistente entre ensaios.

2. Avaliação Psicológica Informatizada

A avaliação sistematizada de atributos psicológicos pode ser desenvolvida de acordo com vários métodos e técnicas, dos quais se salientam os testes psicológicos que ainda hoje, quando utilizados em contexto de selecção de pessoal, são designados por testes psicotécnicos. Apesar do uso de testes psicológicos ser uma prática corrente da avaliação nas várias áreas de aplicação da Psicologia, não existe uma definição suficientemente objectiva e abrangente de teste psicológico que seja aceite de modo pacífico por todos os práticos e investigadores.

Se considerarmos o modelo de classificação de Guilford (1982), comumente designado por “Estrutura da Inteligência” ou “Estrutura do Intelecto”, podemos distinguir três parâmetros de análise da realização de uma tarefa intelectual: (1) o conteúdo; (2) a operação mental solicitada; e (3) o resultado esperado. O modelo contempla cinco operações mentais, cinco conteúdos das tarefas e seis produtos possíveis do trabalho mental, permitindo um sistema classificatório de 150 aptidões diferentes, alargando, forçosamente, a amplitude da definição de teste psicológico. Cronbach (1990) apresenta uma definição de teste psicológico suficientemente abrangente de modo a contemplar um conjunto de procedimentos que considera importantes. Para ele, um teste é um procedimento sistemático de observação de comportamentos e de os descrever com a ajuda de escalas numéricas ou de categorias fixas. A definição de teste psicológico por nós defendida corresponde a todo o tipo de instrumento padronizado de medida que permita a avaliação de um atributo psicológico ou a predição de

um comportamento. Esta definição é convergente com a posição de Kaplan e Sacuzzo (1997) que consideram que um teste é um instrumento ou técnica de medida usado para quantificar um comportamento ou ajudar na compreensão e predição do comportamento em causa.

A definição de Kaplan e Sacuzzo (1997) é suficientemente abrangente para gerar consensos de várias áreas. Os instrumentos que iremos utilizar na parte prática estão de acordo com esta definição, pois não só permitem quantificar um determinado atributo, como ajudam a prever comportamentos associados a esse atributo.

Em 2001 comemoraram-se os vinte anos do lançamento da máquina que passou a ser referência no âmbito da informática: o IBM PC. Na altura (1981) tinha como características base: uma velocidade de 4,77 MHz, 16Kb de memória RAM e não possuía disco rígido. Vinte anos depois, é vulgar encontrar em qualquer grande superfície comercial computadores com processadores a uma velocidade superior a 1 Gigahertz, 126 Mb de memória RAM e discos rígidos a armazenarem 40Gb de informação. A evolução tem sido exponencial e tem arrastado atrás de si o desenvolvimento de *software* com elevadas capacidades multimédia.

A evolução extremamente acelerada das tecnologias de informação tem trazido elevadas vantagens para diferentes áreas funcionais nas organizações, incluindo a gestão dos Recursos Humanos, que à semelhança de outras áreas, tem vindo a utilizar de forma mais alargada os computadores. Um dos objectivos mais comuns da utilização dos sistemas informáticos no âmbito da Psicologia em contexto organizacional, tem sido a avaliação psicológica para fins de selecção de pessoal.

Os primeiros trabalhos de aplicação da informática à avaliação psicológica estavam essencialmente relacionados com a alteração da forma de administração da prova. Somente na década de 70 começaram a efectuar-se os primeiros estudos com o objectivo de avaliar as capacidades da máquina de forma a adaptar a sequência de apresentação dos itens em função

do desempenho do sujeito. Era o início da construção dos testes adaptativos.

Hoje em dia é possível utilizar os sistemas informáticos para a construção de situações em ambiente de realidade virtual, que permitem avaliar orientação espacial, aprendizagem, interação interpessoal, entre outras dimensões. As situações de realidade virtual facilitam o estudo da relação entre avaliador/avaliado, tal como a interação entre sujeitos.

Ao nível da metodologia estatística aplicada à avaliação psicológica, surge, com a avaliação psicológica computadorizada, a Teoria de Resposta ao Item (*Item Response Theory*) como forma de complementar as abordagens psicométricas das qualidades dos itens e consequentemente dos testes.

Por outro lado, as vantagens da informatização também têm sido questionadas. Se considerarmos as vantagens e desvantagens da informatização dos testes psicológicos, verificamos que a balança parece favorecer a avaliação computadorizada tanto ao nível organizacional como na pesquisa científica (Drasgow *et al.*, 1993).

A evolução tecnológica dos equipamentos informáticos tem estimulado, com efeito, o desenvolvimento de provas psicológicas mais sofisticadas e mais próximas da realidade. Neste capítulo iremos apresentar os diferentes sistemas tecnológicos utilizados na avaliação psicológica para, depois, apresentarmos as vantagens da sua informatização. Este capítulo irá terminar com uma descrição das potencialidades e limites dos sistemas informáticos utilizados na avaliação psicológica.

2.1. Diferentes tecnologias na avaliação psicológica

A utilização de diferentes tipos de tecnologias na avaliação psicológica está directamente relacionada com a evolução dos equipamentos e a sua acessibilidade do ponto de vista

económico. Por outro lado, cada uma das diferentes fases do processo de avaliação psicológica solicita a utilização de diferentes tecnologias.

Na avaliação psicológica, os sistemas computadorizados utilizados na avaliação de diferentes dimensões psicológicas mais frequentemente referenciados são: (1) leitores ópticos; (2) leitores ópticos com ligação a microcomputadores e *scanners*; (3) testes em *software*; (4) *software* com *hardware* específico; (5) redes locais; e (6) realidade virtual.

Os leitores ópticos não trazem nada de novo relativamente à construção de testes que avaliem dimensões já avaliadas por testes de papel-e-lápis, uma vez que o seu valor acrescentado reside sobretudo na correcção das respostas. Têm como principal objectivo a ajuda na correcção, substituindo, assim, os profissionais encarregues desta fase do processo de avaliação, com mais-valia em termos de rapidez e precisão.

Quando ligados a computadores, os leitores ópticos aumentam substancialmente as potencialidades, uma vez que toda a informação relativa ao candidato e aos itens fica armazenada automaticamente no sistema informático, permitindo que esses dados sejam utilizados posteriormente para tratamento estatístico com vista à actualização das tabelas de normalização dos dados. Este tipo de sistema permite, ainda, um sumário dos resultados (tipo relatório), o qual poderá ser impresso.

Nos testes em *software*, as provas são administradas, cotadas e os resultados impressos através do computador. Hoje em dia, um número significativo de testes em papel-e-lápis têm disponível uma versão em computador. A criação de *software* gestor de testes para funcionar informaticamente tem sido bastante dinâmica nos últimos anos. A diversidade do *software* criado para avaliação psicológica tem acompanhado a evolução das capacidades dos computadores, tanto em termos de rapidez de processamento como das capacidades de gestão dos componentes multimédia. Por exemplo, começam a aparecer provas psicológicas em que

é apresentada ao sujeito uma situação por intermédio de um pequeno filme, sobre a qual terá de indicar o comportamento que teria. Provas como esta, destinam-se, geralmente a avaliar aspectos da personalidade e ultrapassam a avaliação clássica dos testes de papel-e-lápis. Substituem as provas situacionais tradicionais em que o sujeito é solicitado a imaginar uma situação para responder como se comportaria nela, reduzindo a subjectividade relativa à imaginação da situação. A informatização permite também a criação de provas em que os itens se movimentam, permite avaliar dimensões impossíveis de serem medidas com os convencionais testes de papel-e-lápis, como é o caso da já “baptizada” aptidão espacial dinâmica (Pellegrino & Hunt, 1989).

Uma outra grande vantagem associada aos testes em *software*, é a possibilidade de acesso ao tempo de latência das respostas, que corresponde ao tempo que separa a apresentação da situação ao sujeito e o momento de resposta. O conhecimento do tempo de latência do sujeito nos diferentes itens permite conhecer melhor o seu funcionamento cognitivo na execução de uma tarefa. Com os testes psicológicos administrados no formato papel-e-lápis, o avaliador apenas ficava com uma ideia do rendimento do sujeito, através da proporção de respostas correctas no tempo permitido.

Os testes administrados em computador, tanto podem derivar quase directamente das suas formas convencionais papel-e-lápis, como podem obedecer ao modelo dos testes adaptativos em que os itens a apresentar ao sujeito estão dependentes do seu próprio desempenho.

Existem também testes em *software* que exigem um *hardware* específico. Incluem todas as vantagens do modelo anterior, mais a possibilidade de maior contextualização da situação de avaliação. É o caso dos simuladores utilizados na selecção de pilotos ou de controladores aéreos. Nestas provas, tenta-se dar o maior realismo possível à situação de avaliação, em que o *hardware* é composto por manípulos, pedais e consolas de botões. Neste tipo de sistemas, é

geralmente esperada uma resposta manipulativa por parte do sujeito.

As redes locais correspondem a salas informatizadas onde existem várias estações de trabalho, permitindo avaliar vários sujeitos ao mesmo tempo, como na situação clássica de aplicação colectiva de testes de papel-e-lápis. Existem, contudo, diferenças significativas entre estas duas formas de aplicação de testes. As redes locais vão usufruir das vantagens enumeradas para os testes em *software*, mas agora com a possibilidade de as diferentes estações de trabalho estarem ligadas em rede, o que permite que os avaliados estejam fisicamente distanciados, bem como garante a centralização num servidor de toda a informação relativa às provas efectuadas. Também se torna possível desenvolver testes que exijam interacção. As redes locais tanto podem conter simples terminais como estações de trabalho ou computadores inteligentes e *hardware* específico associado.

Por último, temos a possibilidade de utilizar a tecnologia inerente à realidade virtual para a concepção de provas que avaliem novos atributos psicológicos. A este nível, o interesse reside essencialmente em desenvolver situações de avaliação que se possam aproximar das situações reais, e que, de alguma forma, permitam algum controlo sobre elas, de modo a padronizar a situação para os diferentes sujeitos. O sistema da realidade virtual, enquanto sistema de avaliação psicológica, pode permitir a avaliação de factores personalísticos de ordem relacional, algo que até agora só era permitido com as chamadas provas de grupo. Os sistemas de realidade virtual associados à possibilidade de ligação em rede permitirão ainda que seja possível efectuar provas de carácter relacional, mas com os sujeitos em zonas geograficamente diferentes.

2.2. Vantagens da informatização para a avaliação psicológica

Analisar os ganhos com a informatização para a avaliação psicológica será sempre uma tarefa

inacabada e terá de ser analisada segundo diferentes perspectivas. Consideramo-la uma tarefa inacabada porque assistimos a uma evolução tecnológica exponencial das máquinas (*hardware*), implicando, conseqüentemente, um aumento das possibilidades da programação (*software*), proporcionando situações de avaliação completamente novas e praticamente limitadas apenas pela criatividade humana.

Segundo Bartram (1989), torna-se necessário decompor o processo de avaliação em sete componentes para melhor identificarmos as vantagens da informatização para cada uma delas. A estas sete componentes - selecção de um instrumento de avaliação, administração do instrumento, correcção, análise das medidas obtidas, interpretação das medidas, dar *feedback* aos candidatos e tomada de decisão com base nos resultados - podemos acrescentar uma outra: construção de novos testes. Embora não relacionada com o procedimento de avaliação, pode contribuir, positivamente, para a qualificação da mesma.

Iremos, seguidamente, analisar em detalhe as vantagens e desvantagens da utilização dos sistemas informatizados em cada um dos momentos da avaliação: (1) construção e/ou escolha do teste; (2) aplicação; (3) correcção e padronização de resultados; (4) interpretação dos resultados e redacção de relatórios; e (5) tomada de decisão.

2.2.1. Construção e/ou escolha do teste

A utilização dos computadores na avaliação psicológica permite avaliar dimensões psicológicas impossíveis de serem avaliadas com os testes tradicionais tal como permite a aplicação das provas de diferente modo. Não pretendendo ser exaustivos, chamamos particular atenção para os testes adaptativos e para as aptidões espacial-dinâmica e *time-sharing*.

2.2.1.1. Testes adaptativos

A construção de testes adaptativos visa essencialmente a apresentação ao sujeito de um conjunto de itens adaptado ao seu desempenho, permitindo uma avaliação do atributo em causa de forma precisa e económica em termos de tempo. O computador utiliza um algoritmo que selecciona o item a apresentar ao sujeito com base nas respostas dadas anteriormente. Assim, por exemplo, se o sujeito acertou os itens 1, 2, 3, e 4, o computador, com base nos parâmetros dos diversos itens disponíveis, selecciona eventualmente o item 7 e não o 5. A construção de testes deste tipo tem várias implicações. Uma, é de que é necessário que o programa tenha acesso a um banco de itens suficientemente grande. Outra, é de que para criar um banco de itens operacional, é necessário que os parâmetros relativamente a cada um possuam suficiente estabilidade, o que obriga a que se tenha utilizado, para o seu cálculo, uma amostra bastante alargada. Neste tipo de testes, apenas os itens mais relevantes e mais informativos são apresentados ao sujeito (Bartram & Bayliss, 1984). Uma outra consequência de utilização de testes adaptativos é a de que se torna muito raro aplicar-se dois testes a dois sujeitos exactamente com os mesmos itens.

A inclusão de um item no banco de itens de um teste adaptativo depende, em grande parte, dos parâmetros calculados de acordo com a Teoria de Resposta ao Item (TRI), que se revela mais eficaz que a Teoria Clássica dos Testes (TCT) quando se utilizam grandes amostras. No entanto, a TCT possui várias estatísticas úteis para a análise dos itens quando as amostras são pequenas. Recordemos aqui, a título de exemplo, o grau de dificuldade do item e a correlação bisserial entre o item e a nota total obtida no teste. A primeira permite-nos conhecer a proporção das pessoas que respondeu correctamente ao item, enquanto a segunda indica o grau de contribuição do item para a obtenção da nota global do teste. O coeficiente de

correlação bisserial entre o item e o total do teste pode ser um melhor indicador que o parâmetro de discriminação do item, de acordo com a TRI, quando a amostra utilizada para o seu cálculo foi de centenas e não de milhares. Drasgow e colaboradores (1993) recomendam nunca incluir no banco de itens um item que possua um coeficiente de correlação bisserial inferior a 0,20.

Por outro lado, a TCT que suportava toda a prática da avaliação psicométrica é desafiada com o desenvolvimento da TRI, em que o investigador passa a ter acesso a outro tipo de informação sobre o comportamento de cada um dos itens da prova. A unidade conceptual mais importante da TRI corresponde à curva característica do item, que é uma função da relação entre a probabilidade de ocorrência de uma resposta correcta num item e a aptidão subjacente ao constructo do item. Embora possam ser considerados vários modelos matemáticos para expressar a relação de probabilidade de sucesso num item e a aptidão medida pelo teste, geralmente são considerados três com maior relevância, diferenciando-se em função do número de parâmetros que utilizam para descrever o item: os que avaliam somente a dificuldade; os que consideram a dificuldade e a discriminação; e os que avaliam a dificuldade, discriminação e a probabilidade de responder correctamente ao acaso (Paquali, 1996).

A TRI surge em alternativa à TCT e caracteriza-se fundamentalmente por, tal como o nome indica, centrar o seu estudo nos itens e não no teste como acontece na TCT. A aplicação dos modelos da TRI aos testes psicológicos possui vantagens relativamente à TCT, sobretudo quando pretendemos conhecer melhor o funcionamento diferencial de cada item. Primeiro, os parâmetros dos itens estimados pela TRI, tal como os da dificuldade e da discriminação, são menos influenciados pelas características da amostra quando comparados com os coeficientes obtidos pela TCT, nomeadamente a fidelidade. Segundo, as características intrínsecas a cada um dos itens são definidas de uma forma muito mais precisa. Por último, é possível

representar de uma forma gráfica mais elucidativa as propriedades estatísticas dos itens.

Em testes adaptativos, Skinner e Pakula (1986) estimam uma redução do tempo envolvido na avaliação na ordem dos 50 a 75%. Muito embora estas vantagens possam parecer prosaicas, elas revelam-se extremamente importantes quando se considera a selecção de pessoal em larga escala (Hunt & Pellegrino, 1985).

Outra vantagem dos testes adaptativos está associada à facilidade de introdução de alterações nos próprios testes. Alterações no formato de resposta ou ainda nos próprios itens (e.g., eliminar itens não discriminativos) são fáceis de efectuar sem que isso implique custos elevados adicionais. No formato papel-e-lápis, tais alterações implicam a elaboração e impressão de novos cadernos de teste e/ou formulários de resposta. Em termos de custos é necessário considerar não só os associados ao tempo necessário para fazer essas alterações como os das próprias alterações.

2.2.1.2. Testes novos, novas dimensões

Uma outra área onde os computadores poderão introduzir substanciais inovações na avaliação psicológica está relacionada com a identificação de novas aptidões humanas (Fleishman, 1988). O que hoje se sabe sobre a caracterização dos factores cognitivos deriva, quase exclusivamente, da análise dos resultados obtidos nos testes de papel-e-lápis. A possibilidade de construir testes sem equivalência a provas de papel-e-lápis, pode levar à descoberta de novas dimensões psicológicas. Salientam-se aqui os testes com movimento e os com apresentação de tarefas simultâneas (*time-sharing*). Os testes com tarefas simultâneas permitem avaliar a estratégia de partilha da atenção entre uma ou várias tarefas prioritárias e outra(s) de rotina, que correspondem a situações mais próximas do que se passa no nosso dia-a-dia.

A exploração das capacidades dinâmicas do computador para a criação de itens com movimento, pode levar à reformulação da conceptualização das capacidades humanas. Por exemplo, a avaliação do movimento na dimensão espacial coloca a hipótese da existência de uma nova aptidão cognitiva que é a aptidão espacial dinâmica (Pellegrino & Hunt, 1989), definida como a capacidade para raciocinar sobre algo que está em movimento, requerendo julgamentos relativos a velocidades e distâncias.

O desenvolvimento de itens com movimento coloca, para algumas funções, o teste mais próximo da simulação de tarefas. Por exemplo, a avaliação da aptidão espacial dinâmica em candidatos a pilotos e controladores de tráfego aéreo permite uma maior proximidade com as funções executadas, elevando a validade facial do processo avaliativo.

As capacidades actuais dos sistemas informáticos permitem, também, conceber tarefas para serem resolvidas em simultâneo, como por exemplo, tarefas em que a estimulação visual é apresentada em simultâneo e coordenada com a estimulação auditiva (sistemas multimédia). A literatura demonstra que as capacidades solicitadas ao sujeito para lidar com estas informações de forma isolada são distintas das capacidades para lidar com elas em simultâneo (Yee, Hunt & Pellegrino, 1991). Em contexto de selecção de pessoal, os sistemas multimédia podem permitir a construção de testes que avaliem os atributos psicológicos numa situação mais próxima da função a que o sujeito se está a candidatar, aproximando o teste a situações simuladas do posto de trabalho, elevando os coeficientes de validade relativa ao critério (McHenry & Schmitt, 1994). Nalguns postos de trabalho os sujeitos têm de dar igual atenção a mais que uma tarefa cognitiva em simultâneo. É o caso dos controladores aéreos num aeroporto com grande movimento em que os aviões, provenientes de vários espaços aéreos, chegam a aterrar na mesma pista com intervalos inferiores a um minuto. As provas avaliadoras das tarefas em simultâneo têm sido utilizadas em selecção de pessoal sobretudo ao nível de pessoal navegante, na predição da *performance* do sucesso no treino de pilotagem

(Carretta, 1987c).

Nos testes de tarefas simultâneas podemos também avaliar a resistência da pessoa à sobrecarga de trabalho mental, o que está directamente relacionado com o stress induzido. Este tipo de situação acaba por melhor traduzir as diversas realidades profissionais, onde a informação apresentada simultaneamente ao sujeito tem diferentes naturezas e fontes.

Projectos mais recentes sobre a utilização dos sistemas informáticos na avaliação psicológica, visam a criação de tarefas baseadas em sistemas de realidade virtual, em que o sujeito não está a dar respostas a um teste, mas a participar numa situação que evolui em função da sua interacção (realidade virtual interactiva).

A criação de novos testes está directamente relacionada com as potencialidades das novas tecnologias, mas também com a criatividade de quem concebe o *software*. Apesar de a construção de novos testes não implicar, de uma forma directa, a "descoberta" de novas dimensões psicológicas, é importante enquadrá-los no conhecimento que temos do funcionamento cognitivo. Também não deixa de ser importante, pelo menos para aqueles que utilizam a avaliação psicológica numa perspectiva de intervenção, conhecer-se a utilidade das novas dimensões para o processo de avaliação em que se está a trabalhar. Não basta que o teste seja aplicado informaticamente para ser considerado um bom teste e substituir os equivalentes de papel-e-lápis. Torna-se necessário estudar a sua validade.

A construção de testes informatizados, tal como a adaptação de testes de papel-e-lápis para suporte informático, deverá obedecer a regras bem delineadas, de forma a que as qualidades métricas a atingir sejam satisfatórias para a sua aplicação. No caso de adaptação de testes de formato papel-e-lápis para formato informatizado, a avaliação das qualidades psicométricas de ambos deverá ser desenvolvida como se de formas paralelas se tratasse (Kline, 1993).

2.2.2. Aplicação

Relativamente à administração, a informatização consegue assegurar que as instruções são exactamente as mesmas para todos os sujeitos. Atendendo à rigidez com que as instruções são apresentadas aos sujeitos, e uma vez que a maioria dos sistemas não possui versões alternativas, de modo a explicar o mesmo por outras palavras (como faria um técnico de aplicação de provas), é de extrema importância a concepção e validação das instruções. Esta fase da construção do teste tem de ser conduzida com um cuidado redobrado.

Mesmo quando se utilizam vários computadores em rede para uma aplicação colectiva, a maioria dos sistemas de testes informatizados permite que cada um dos avaliados possa gerir à sua maneira o tempo para a compreensão das instruções, assim como pode decidir o momento para iniciar a prova propriamente dita. O sujeito passa a ter maior autonomia no momento da aplicação. Mesmo em aplicações colectivas, com os computadores ligados em rede, não se torna necessário que todos comecem a prova ao mesmo tempo.

A evolução tecnológica tem proporcionado maior flexibilidade na programação, concebendo-se *software* que assemelhe o manuseamento das provas informatizadas aos testes de papel-e-lápis, como por exemplo, permitir ao sujeito voltar aos itens anteriores e alterar a resposta dada.

Apesar de se garantir, nesta fase da avaliação psicológica, maior autonomia ao sujeito, retira-se, ou diminui-se, o contacto com o elemento humano. Em salas com várias estações de trabalho, o sujeito praticamente não interage com o responsável de sala.

2.2.3. Correção e padronização

Os sistemas informatizados utilizados na correção das provas permitem, essencialmente,

aumentar a rapidez, eliminar os erros prováveis na correcção manual e alimentar uma base de dados de forma a permitir que os mesmos sejam utilizados posteriormente para o necessário tratamento psicométrico. Alguns dos testes informatizados possuem rotinas que efectuem uma actualização quase automática das tabelas utilizadas na normalização de resultados.

2.2.4. Interpretação dos resultados e redacção do relatório

O grande contributo que os sistemas informatizados podem dar nesta etapa da avaliação psicológica, consiste em aceder, de uma forma mais rápida, a um maior volume de informação. Mesmo que as provas sejam aplicadas em locais afastados, se os terminais de resposta estiverem ligados em rede, a informação relativa aos resultados pode ser centralizada num computador, e eventualmente visualizada por todos os envolvidos no processo de selecção que estejam legitimamente autorizados a aceder a essa informação.

Os resultados do sujeito em cada um dos itens, no teste e comparativamente com as médias do grupo, podem ajudar o avaliador a ter uma ideia mais realista das capacidades do avaliado. Por outro lado, alguns sistemas produzem relatórios não só baseados nos resultados de provas de aptidões, como também nos questionários auto-descritivos. Podemos considerar que são bons indicadores para a condução de posteriores entrevistas e/ou para a redacção dos relatórios finais, mas nunca para serem entregues directamente ao cliente, como se fosse o próprio relatório final. É importante que a informação constante dos relatórios informatizados seja validada, antes de ser utilizada como forma final. Alguns autores questionam, inclusive, a natureza ética de tais procedimentos.

A inclusão de informação do foro fisiológico e neuropsicológico no relatório informatizado poderá ser um novo passo na avaliação psicológica. O avaliador pode, deste modo, tomar em consideração, para a produção do relatório final, alguns dos indicadores da actividade cerebral

do sujeito (Matarazzo, 1992).

2.2.5. Tomada de decisão

Para quem utiliza sistemas informatizados na avaliação psicológica em contexto de selecção de pessoal, pode beneficiar da possibilidade de o computador submeter automaticamente os resultados a várias equações de ponderação, de forma a estimar o desempenho posterior do sujeito em várias áreas profissionais e de o hierarquizar relativamente aos restantes candidatos.

Também é possível efectuar a comparação de perfis. O sistema compara os resultados obtidos pelo sujeito nas dimensões consideradas críticas para o desempenho da função com os valores limite constantes do perfil exigido, indicando o grau de ajustamento a essa função, comparativamente com os outros candidatos. Esta opção será bastante útil para quem trabalhe em selecção/colocação de pessoal e esteja a avaliar um conjunto de candidatos que concorre em simultâneo para vários postos de trabalho com exigências diferentes.

Tal como na fase anterior, podem colocar-se constrangimentos de ordem ética quanto à utilização das novas tecnologias. As interpretações efectuadas pelos sistemas, assim como as tomadas de decisão sugeridas, devem ser abordadas com cuidado e cada dimensão deve ser objecto de análise fina por parte do psicólogo. Apesar de as novas tecnologias fornecerem um conjunto de informação que permite ao utilizador uma tomada de decisão cientificamente mais precisa, os psicólogos continuam a ser os peritos do comportamento humano, logo responsáveis pela emissão de qualquer parecer derivado da avaliação psicológica. Estas questões não podem ser negligenciadas, pois são determinantes no crédito que a avaliação psicológica terá perante a comunidade.

2.3. Testes de papel-e-lápis versus testes informatizados

A informatização da avaliação psicológica começou pela digitalização de itens que eram apresentados num ecrã ao sujeito. Não obstante de ser um avanço relativamente à aplicação de testes de papel-e-lápis, tinha reduzidas vantagens. Pretendia-se que os indivíduos não riscassem os cadernos e que a correcção dos testes fosse mais rápida e precisa. Actualmente, o cenário é diferente. As capacidades actuais dos computadores têm estimulado o trabalho de pesquisa e sistematização de informação conducente a uma avaliação psicológica mais dinâmica e com ganhos substanciais relativamente ao método de papel-e-lápis.

As tendências mais recentes ao nível da criação de testes informatizados passam pela possibilidade de criar provas que se adaptem ao rendimento progressivo do sujeito (testes adaptativos), pela possibilidade de apresentar itens com movimento e pela possibilidade da apresentação visual dos itens ser acompanhada da componente auditiva, entre outras (Drasgow *et al.*, 1993).

Em termos de *output*, o computador permite uma avaliação mais individualizada e mais completa. Nas versões tradicionais o desempenho dos sujeitos era avaliado por uma nota bruta, geralmente correspondente ao número de itens respondidos correctamente. O resultado final de um sujeito em determinado teste espelhava quase sempre esse grau de correcção ou, eventualmente, o tempo total gasto na prova quando se tratava de aparelhos psicómotores.

Com a informatização, o avaliador tem a possibilidade de conhecer a evolução do sujeito ao longo da prova, não só pelas respostas que deu aos itens mas também pela forma como geriu o tempo para responder aos mesmos. A informatização permite obter informação mais pormenorizada sobre o desempenho do sujeito e a consequente inclusão no processo de avaliação psicológica (Skinner & Pakula, 1986). O tempo gasto, não em termos globais mas para cada um dos itens, passa a ser uma nova variável a que o avaliador tem acesso e que os

estudos terão que demonstrar a sua validade/utilidade, a qual não deve ser simplesmente assumida porque a prova é agora aplicada numa plataforma mais sofisticada.

As objecções psicométricas à utilização de computadores na avaliação psicológica são raras. No entanto, é comum a referência a um aspecto polémico, relacionado com a equivalência entre formas paralelas de testes administrados por computador e por papel-e-lápis. Poderemos construir conjuntos de testes que possam ser considerados verdadeiramente equivalentes? Isto é, ao aplicarmos uma bateria de papel-e-lápis e outra informatizada, exactamente com os mesmos testes, estamos a avaliar as mesmas aptidões e iremos obter os mesmos resultados? Alguns estudos, nomeadamente os de Segall e Moreno (1999) foram desenvolvidos neste sentido, e mostraram que a bateria vocacional utilizada pelas Forças Armadas dos EUA, conhecida pela abreviatura ASVAB (*Armed Services Vocational Aptitude Battery*), que anualmente é aplicada a mais de dois milhões de alunos finalistas do ensino secundário, ao ser informatizada mostrou ser uma forma paralela da sua congénere no formato em papel-e-lápis. Contudo, num estudo efectuado dez anos antes, por Almeida e Sá (1987) onde foram aplicados dois testes da bateria de provas de raciocínio diferencial (Raciocínio Verbal e Abstracto a dois grupos emparelhados de alunos do 8º Ano de escolaridade no modo de papel-e-lápis e em computador, verificaram haver diferenças significativas nas médias quer no número de itens correctos, quer no número de itens atingidos. Os alunos que realizaram as provas através de computador apresentam resultados inferiores aos que as realizaram em papel-e-lápis. Os sujeitos tendem a obter resultados mais baixos nos testes de conteúdo verbal, quando estes são aplicados informaticamente. No entanto, estes resultados podem estar associados à menor resolução gráfica dos equipamentos utilizados, se considerarmos o ano em que foi realizado o estudo. (Almeida & Sá, 1987).

A maior parte dos investigadores defende que o processo de desenvolvimento de um teste informatizado deve ser idêntico ao das formas paralelas (Kline, 1993, 2000). Kline defende

que se devem efectuar estudos para analisar a equivalência das versões computadorizadas de testes de papel-e-lápis, devendo as duas versões apresentar correlações superiores a 0,90. Para Burke e Normand (1987) não só os coeficientes de correlação devem ser elevados, como a própria distribuição de frequências deverá ser semelhante em ambas as versões. Ainda segundo os mesmos autores, apesar da necessidade de se demonstrar a equivalência de versões, são poucos os estudos desenvolvidos com tais objectivos, sendo, por vezes, os respectivos resultados contraditórios.

A “conversão” de uma versão papel-e-lápis para a versão automatizada de um teste deve também cuidar dos aspectos directamente relacionados com a aplicação, nomeadamente assegurar a clareza das instruções; temporizar o treino antes do teste começar; poder voltar atrás para fazer alterações às respostas; poder “saltar” itens, etc. Por tudo isto, não basta colocar o teste em computador. A partir do momento em que o computador passa a ser o meio de aplicação, é necessário ter todos os cuidados de modo a se garantir um bom *interface* homem-máquina. Como se referiu anteriormente, esses aspectos poderão influenciar a medição do constructo.

De acordo com Bartram (1994), para que duas versões possam ser consideradas formas paralelas do mesmo teste, elas têm de: (1) ter idênticos coeficientes de consistência interna; (2) correlacionarem-se a um nível esperado pelas consistências internas encontradas; (3) apresentarem coeficientes de correlação idênticos com outras variáveis; e (4) mostrarem idênticas médias e desvios-padrão. Ainda segundo Bartram, para se poder avaliar a equivalência das formas, torna-se necessário utilizar um *design* de quatro grupos para se efectuar o teste-reteste em que dois grupos funcionam como controlo (condições computador/computador e papel-e-lápis/papel-e-lápis) e os outros dois com aplicações cruzadas (computador/papel-e-lápis e papel-e-lápis/computador).

A informatização directa de um teste de papel-e-lápis pode introduzir variáveis que levem a não podermos considerar as duas formas equivalentes, apesar de os itens conterem exactamente a mesma informação. Variáveis como, por exemplo, a eventual dificuldade de leitura dos itens no ecrã do computador ou o menor controlo sobre a situação, podem baixar o rendimento do sujeito. Neste último caso, estamos-nos a referir ao facto de o sujeito não poder "caminhar" no teste informatizado com a mesma liberdade que o faz com o caderno do teste. No entanto, cada vez mais, a programação dos testes tem-se desenvolvido no sentido de facilitar o trabalho dos testados, tornando a sua utilização mais amigável. Uma posição mais drástica é assumida por Green, Bock, Humphreys, Linn e Reckase (1984) ao sugerirem que os próprios constructos avaliados podem alterar-se quando o modo de administração do teste é modificado.

Mead e Drasgow (1993) defendem que a conversão de um teste de formato papel-e-lápis para formato automatizado pode deparar-se com duas dificuldades: (1) mudança do resultado final devido, por exemplo, a uma mudança na escala utilizada na nova versão; e (2) baixa correlação entre as duas versões. A baixa correlação entre a versão papel-e-lápis e a versão automatizada poderá ser indicativa da avaliação de diferentes constructos, ou seja, o modo como as provas foram administradas introduz variáveis específicas que faz com que a dimensão psicológica que está a ser avaliada não seja exactamente a mesma. Por exemplo, o facto de na versão papel-e-lápis o teste avaliar a aptidão espacial isso não significa que meça essa mesma dimensão na sua versão automatizada. É preciso efectuar estudos que demonstrem essa equivalência.

Num estudo de meta-análise, Mead e Drasgow (1993), ao avaliarem a equivalência de testes de aptidão cognitiva aplicados de forma informatizada e em papel-e-lápis, concluíram que a alteração do meio de administração dos testes produz efeitos relativamente reduzidos quando os constructos são avaliados por testes de rendimento. Por outro lado, os constructos

avaliados por testes de velocidade parecem apresentar diferenças relativamente à forma de aplicação. Não podemos, no entanto, a partir do estudo de Mead e Drasgow, generalizar, para todo e qualquer teste de rendimento ou de velocidade, os resultados encontrados.

Um dos aspectos importantes aquando da comparação entre testes aplicados em diferentes formatos é a avaliação das suas capacidades preditivas. São raros, no entanto, os estudos publicados sobre esta matéria. O próprio trabalho de Mead e Drasgow (1993) não inclui a análise dos coeficientes de validade preditiva.

No estudo de MacLennan, Jackson e Bellantino (1988) encontraram-se diferenças entre os resultados obtidos nos testes administrados no formato papel-e-lápis e via computador no que diz respeito às suas correlações com os tempos de latência. A relação entre os tempos de latência e o rendimento nos testes é inferior quando os testes são aplicados via papel-e-lápis. A razão para tais resultados prende-se com as características da própria aplicação em que o estudo foi desenvolvido. O sujeito não pode, por exemplo, voltar atrás para corrigir os itens, que assinalou incorrectamente. A gestão do tempo disponível também é feita de forma diferente. O facto do estudo ter sido desenvolvido em 1988 levanta alguns problemas de generalização dos resultados, já que actualmente é possível dar a possibilidade ao sujeito de poder recuar no teste alterando as respostas dadas.

MacLennan e colaboradores (1988) são da opinião que os indivíduos ao efectuarem um teste via computador obtêm um rendimento superior aos que efectuam o mesmo teste na versão de papel-e-lápis, por conseguirem uma maior concentração, já que só têm acesso a um item de cada vez. Uma das conclusões do estudo de MacLennan e colaboradores é a de que a latência das respostas a itens de testes verbais avaliados por computador apresentam uma relação inversa com o rendimento global no teste, confirmando estudos anteriores (Jensen, 1982a; Vernon, 1983).

2.4. Potencialidades e limites dos sistemas informatizados

Mas se a balança tende para as vantagens na utilização de testes computadorizados, temos de considerar alguns aspectos menos positivos na aplicação da informatização na avaliação psicológica. A primeira grande desvantagem relaciona-se com o investimento financeiro implicado na aquisição do sistema base e dos testes, caso a opção seja por testes informatizados. O investimento engloba não só o equipamento necessário, mas igualmente na qualificação do pessoal a ser envolvido na sua utilização. Com um sistema computadorizado de avaliação psicológica cada pessoa tem o seu terminal, o que limita o número de pessoas que podem ser testadas ao mesmo tempo. Burke e Normand (1987) defendem que num ano este investimento poderá ser recuperado, mas esta será talvez uma perspectiva demasiado optimista. Aliás, e contrariamente às provas de papel-e-lápis, os equipamentos a adquirir carecem de uma actualização constante e conseqüente garantia de manutenção do sistema. Por sua vez, a actualização, tanto do sistema como das provas nele incluídas, poderá ter implicações imediatas naquilo que é avaliado pelo teste. Ao colocar o mesmo teste numa máquina, tecnologicamente mais evoluída ou ao instalar uma versão mais actual do teste, é necessário garantir que tal “migração” não afecte a avaliação do constructo (Bartram & Bayliss, 1984).

Um segundo conjunto de desvantagens está associado às reacções perante os equipamentos (teclado, pedais, etc.) por parte das pessoas envolvidas no processo de avaliação. Um primeiro grupo de pessoas a considerar são os técnicos envolvidos na selecção. Alguns autores defendem que a automação pode ser vista como uma ameaça à “*expertise*” profissional (Bartram & Bayliss, 1984). No entanto, esta resistência à mudança por parte dos profissionais pode ser ultrapassada com formação (Burke & Normand, 1987) e, ainda, com o seu envolvimento nas diversas fases do processo de automação (Skinner & Pakula, 1986).

O segundo grupo de pessoas a considerar são os próprios testados. De um modo geral, a literatura aponta para reacções positivas em relação a este novo tipo de testes (Skinner & Pakula, 1986). Uma das explicações apontada para essa boa aceitação é o facto de cada vez haver um maior contacto com os computadores, quer a nível profissional quer a nível pessoal. Uma outra explicação apresentada para a boa aceitação é a semelhança que os sujeitos percebem existir entre os testes que estão a efectuar e as funções que no futuro pensam executar. Isto é, os sujeitos reconhecem que os instrumentos em si poderão ser úteis para a selecção, por estarem próximos do posto de trabalho (validade facial). Um dos exemplos diz respeito às provas utilizadas para a selecção de pilotos em que o grafismo apresentado aos sujeitos está muito ajustado ao cenário que irão encontrar no *cockpit* do avião. Considerando que existem profissões que cada vez mais recorrem ao uso da informática, como é o caso dos pilotos e controladores aéreos, mais fácil se torna fazer a aproximação dos cenários apresentados pelos testes em computadores às situações da prática profissional.

A literatura reconhece que esta boa aceitação dos computadores é fortemente influenciada pelas experiências prévias que os sujeitos têm com computadores (Burke, Normand & Raju, 1987; Skinner & Pakula, 1986). Mas esta maior familiarização com os computadores pode, ela própria, levantar questões específicas dos testes automatizados. Alguns autores defendem que a prática de jogos de computador pode influenciar os resultados nos testes informatizados (Eliot & Smith, 1983). No entanto, os estudos não indicam diferenças significativas dos resultados de provas informatizadas entre sujeitos com contacto anterior com computadores e os sem contacto anterior (Almeida & Sá, 1987).

Por fim, os testes em computador têm levantado algumas questões do foro ético. As possíveis interpretações que o computador pode oferecer (i.e., breves relatórios sumariando os resultados dos sujeitos) constituem um alvo privilegiado dessas críticas (Skinner & Pakula, 1986). A situação ideal consistirá em ter um técnico para analisar e explicar ao sujeito o

significado dos seus resultados, independentemente do tipo de testes utilizado (Kline, 1993).

O facto de a avaliação psicológica ser efectuada em computador, em nada deve diminuir a responsabilidade dos técnicos envolvidos no processo. Em última análise, e tal como acontece com qualquer outro tipo de testes, é a estes técnicos que cabe a responsabilidade de assegurar a privacidade, confidencialidade e segurança dos resultados e dos envolvidos no processo de avaliação psicológica. Tal como em qualquer área profissional, não é a tecnologia por si só que poderá colocar novas questões éticas, mas sim a forma como esta é utilizada (D'Oliveira & Bártole Ribeiro, 2000).

2.5. Comentários finais

A evolução das novas tecnologias e a expansão da utilização está a ter consequências a vários níveis, incluindo a avaliação psicológica. Salientemos algumas: (1) ganhos substanciais em termos de tempo e de precisão na correcção e padronização de resultados; (2) concepção de testes que nos permitem conhecer novas dimensões psicológicas, de forma a melhor compreender e prever o funcionamento psico-afectivo do sujeito; (3) concepção de novas formas de avaliação que permitam uma avaliação mais personalizada, como a aplicação da TRI à construção de provas; e (4) acesso a um novo conjunto de informação, tanto de ordem psicológica (tempos de latência), como de ordem fisiológica, permitindo fazer a ponte entre a Psicometria e a Psicologia Cognitiva. Quer a informatização nos leve, ou não, para a construção de novos itens ou testes, e nos permita, ou não, conhecer outras dimensões da esfera cognitiva, a certeza é de que ela nos vai permitir avanços importantes na avaliação das diferenças individuais (Drasgow *et al.*, 1993).

A avaliação psicológica com a informatização também se poderá desvirtuar afastando-se dos padrões éticos, se não forem tomadas em consideração um conjunto de medidas. O facto de

que quanto mais *user friendly* for o sistema e mais global for a sua actuação, maior será a tentação para ser usado por quem não possui os necessários conhecimentos ao nível do comportamento humano, que poderá levar a efectuar interpretações abusivas e sem o necessário enquadramento técnico-científico. Consideramos que nesta área as associações científicas e profissionais da Psicologia deverão estar atentas e activas para que possam ser o garante da boa prática da avaliação psicológica, segundo os padrões éticos e deontológicos consignados na directrizes estipuladas a nível nacional e internacional (Fernández-Ballesteros & Steyer, 2001; ITC, 2000).

3. Avaliação Psicológica em Contexto da Selecção de Pessoal

A selecção de pessoal foi uma das primeiras áreas de intervenção da psicologia nas organizações e um dos primeiros psicólogos a trabalhar neste campo foi Lahy que teve a seu cargo o desenvolvimento de baterias de testes com vista à selecção. Constatou que as causas da fadiga e os acidentes estavam relacionadas com a selecção de pessoal. Relativamente à fadiga, verificou que os indivíduos que se cansavam mais tinham menor rendimento na avaliação de determinadas capacidades/aptidões. Sabendo as causas, poder-se-ia rejeitar os sujeitos com pouca resistência à fadiga no processo de selecção. Para Lahy, a principal causa dos acidentes seria a má adaptação do indivíduo ao posto de trabalho, isto é, o acidente surgia por deficiente ajustamento do sujeito àquele posto de trabalho.

São várias as razões que levam as organizações a recorrer a processos de selecção sustentados em pressupostos psicológicos ao invés de escolher um trabalhador novo com base exclusivamente na informação e referências que se têm dele, ou na impressão que ele eventualmente criou num primeiro contacto. Uma das razões prende-se com a justiça do processo.

É um processo justo para todos os candidatos, já que são colocados em igualdade de condições, de forma a que os que possuem maior facilidade de adaptação às exigências da função têm maiores probabilidades de virem a ser admitidos. Os candidatos são escolhidos com base em critérios objectivos e definidos a partir das exigências da função. Podemos considerar que existem vantagens para a organização e para o indivíduo numa selecção deste

tipo. As pessoas não têm que ter sucesso em todos os tipos de postos de trabalho. Diferentes funções/postos de trabalho possuem características diferentes, que fazem com que os perfis de exigências sejam diferentes. Por outro lado, o insucesso organizacional é oneroso tanto para a organização como para o trabalhador. Para a organização, porque ela investe dinheiro no recrutamento, selecção e formação; mas também é oneroso para o trabalhador porque também ele investe psicologicamente num trabalho novo. O tempo que ele vai dedicar ao seu novo trabalho, para aprender as tarefas que lhe serão adstritas, poderia ter sido dedicado a outro projecto potencialmente mais produtivo. Na gestão de recursos humanos, os custos provocados pela perda de trabalhadores é difícil de quantificar, mas os peritos nesta área desenvolvem um esforço considerável para otimizar os benefícios da utilização dos sistemas de selecção de pessoal.

Com base numa análise do impacto financeiro do comportamento humano nas organizações, Cascio (1982) mostra vários exemplos da utilidade da selecção de pessoal. Por exemplo, numa academia americana regional de treino policial, uma média de 15 por cento dos 300 recrutas anualmente aceites são eliminados. Considerando um custo de USD\$10,000 por cada aluno rejeitado, esta taxa de atrição representa uma perda anual para a academia de USD\$450,000. A eventual introdução de um sistema de selecção com uma validade relativamente superior ao anterior, reduzindo a taxa de atrição para 12 por cento, pode provocar uma poupança anual de USD\$90,000.

A maioria das empresas portuguesas ainda não submetem os seus processos de selecção a análises em termos de custo-eficácia, não estimando, assim, os benefícios que se têm no desenvolvimento de uma selecção eficaz. As consequências em utilizar um sistema de selecção que não possua a eficácia necessária são extremamente onerosas se tomarmos em consideração as verbas investidas no recrutamento, selecção e treino, para já não se falar nas consequências organizacionais do seu mau desempenho. A título de exemplo, e tomando em

consideração as estatísticas americanas, de acordo com Hunter (1989), a Força Aérea dos Estados Unidos estima entre USD\$50,000 e USD\$80,000 o custo por cada candidato que não termine o treino de pilotagem. Estes dados baseiam-se numa taxa de abandono na ordem dos 30%, que é o normal nos cursos de pilotagem militar. Se considerarmos um fluxo de cerca de 2.000 novos pilotos por ano durante os anos oitenta, o custo anual para a Força Aérea dos Estados Unidos relativo aos que não completaram o curso foi entre 30 a 48 milhões de dólares. A projecção destes valores para outras áreas, ou ainda dentro do âmbito da pilotagem para outras nações, é fácil, bastando apenas a substituição dos valores nas variáveis consideradas.

No caso das organizações militares em geral, torna-se difícil a identificação realista das diferentes componentes da análise custo-eficácia aplicada à selecção de pessoal e à formação, uma vez que são várias as entidades envolvidas e muitas vezes operam de forma não integrada.

3.1. Os princípios da selecção

É inquestionável que as empresas recrutam pessoas para poderem sobreviver, assim como as pessoas procuram trabalho para poderem adquirir os bens que necessitam para satisfazer as suas necessidades básicas e garantirem o bem-estar e qualidade de vida inerente às suas ambições. O que pode ser questionado é a razão de se recorrer a modelos complexos, nem sempre totalmente objectivos, e relativamente onerosos, como a avaliação psicológica (ainda hoje, erradamente designada por psicotécnicos e que inclui a entrevista com um psicólogo). Porque não admitir uma pessoa recomendada por alguém com uma simples entrevista com o futuro chefe imediato? As empresas sabem que a admissão de pessoas simplesmente porque são conhecidas ou indicadas por alguém sem serem submetidas a um processo formal de

selecção constitui um risco, que os responsáveis pela gestão dos recursos humanos estão cada vez menos dispostos a assumir (Schuler & Jackson, 1997).

A selecção de pessoal assente em procedimentos de índole psicológica tem por objectivo a convergência dos interesses do indivíduo e da organização, de forma a otimizar a relação satisfação-rendimento. Esta convergência é conseguida através de dois princípios base. Um é o de que as pessoas são diferentes a vários níveis, nomeadamente em aspectos cognitivos, de personalidade, e de motivação, entre outros. Face à mesma situação, diferentes pessoas apresentam reacções diferentes em termos de rendimento, empenho e satisfação. A aplicação deste conceito ao posto de trabalho é directa. Existem postos de trabalho que poderão proporcionar maior satisfação e estimular o empenho para uma pessoa e não para outra. O conhecimento do posto de trabalho e das competências assumem, desta forma, uma importância vital no processo de selecção de pessoal.

A selecção óptima corresponderia à escolha da pessoa que possuísse as características do foro psicológico (motivação, aptidões e personalidade) exigidas pelo posto de trabalho no seu nível mais elevado. Desta análise, resultam três considerações: a primeira fundamenta o facto de serem os psicólogos a efectuarem a avaliação dos sujeitos por se estar a lidar com a sua área de conhecimento – o comportamento humano. A segunda consideração relaciona-se com os métodos de selecção. Os mais utilizados na avaliação do foro psicológico das pessoas não possuem as características máximas no âmbito de psicometria (sobretudo no que respeita à validade). Uma terceira consideração relaciona-se com a impossibilidade de se ter a certeza de que não existiria mais ninguém capaz de desempenhar melhor a mesma função, para além dos candidatos encontrados.

O segundo princípio de selecção de pessoal assenta na avaliação. As dimensões psicológicas geralmente avaliadas são mais ou menos estáveis e as pessoas tendem a manifestar comportamentos idênticos em situações análogas. Assim, com base nos resultados obtidos na

avaliação psicológica é possível, pelo menos de forma parcial, prever o comportamento futuro das pessoas. A partir deste segundo princípio também podemos fazer algumas considerações. Uma está relacionada com as limitações estatísticas dos métodos utilizados na avaliação, como veremos mais à frente. Outra diz respeito a que as situações de avaliação quanto mais próximas estiverem das situações funcionais reais futuras, menor será o erro na predição do comportamento. Por último, também podemos considerar que as condições envolventes das situações não permanecem inalteráveis ao longo do tempo, ou seja, a atitude com que o indivíduo se submeteu à avaliação na selecção não será a mesma com que está agora no posto de trabalho a desempenhar a função.

3.2. O paradigma da selecção de pessoal

A selecção ainda é vista no enquadramento do paradigma do "Homem certo para o lugar certo", ou seja, para uma função existe um único Homem que é necessário encontrar para o colocar nesse posto de trabalho. No entanto, este paradigma dá uma visão reducionista da selecção porque: (1) não tem em conta as alterações do posto de trabalho; nem (2) as modificações do indivíduo. Relativamente ao primeiro aspecto nada garante que o posto de trabalho se mantenha estruturado sempre da mesma forma. A introdução de novas tecnologias nos postos de trabalho provocou alterações substanciais nas exigências cognitivas. Um sujeito altamente eficiente a alimentar um arquivo de fichas de processos, pode não ter a mesma rentabilidade ao trabalhar com um computador. Por outro lado, não podemos esperar que o sujeito colocado naquele posto de trabalho mantenha sempre o mesmo nível de ambição e motivação. A pessoa tende a modificar o seu comportamento face à organização.

Se este paradigma do "Homem certo para o lugar certo" não satisfaz plenamente a selecção, coloca-se então a questão se se deve escolher as pessoas para um posto de trabalho ou para uma

carreira? A opção da selecção para uma carreira implica, por outro lado, que as empresas possuam um planeamento de carreiras minimamente estruturado. Sabemos, também, que existem diferenças entre as organizações no que concerne ao planeamento das carreiras. Nalgumas empresas os sujeitos ao serem seleccionados para determinados postos de chefia, têm necessariamente de passar por todos os escalões intermédios. Se existem empresas em que o trajecto do sujeito na organização está formalmente determinado, noutras isso não acontece, nem faz parte da orientação estratégica da organização.

Estes dois aspectos, projecto de carreira e a eventual evolução do posto de trabalho, são pertinentes para a selecção. Eles vão ter implicações na estruturação do processo, como, por exemplo, na escolha das dimensões a avaliar e na forma de as avaliar. Numa tentativa de integração destes dois aspectos, poderemos considerar que a selecção de pessoal visa sobretudo a escolha da pessoa que garanta uma adaptação rápida ao posto de trabalho que é necessário preencher e que possua potencial suficiente para progredir no projecto de carreira.

Os riscos decorrentes de uma má selecção, ou seja, a escolha incorrecta dos sujeitos poderá originar acidentes de trabalho, perda de produtividade ou levar a situações negativas, mesmo que menos evidentes como, por exemplo, a perda de clientes, danificação de equipamentos, constante tendência para a doença, etc. Aliás, uma má selecção de pessoal não diz respeito unicamente à escolha dos sujeitos que não se deveriam ter seleccionados, mas também à não admissão dos indivíduos que são realmente bons, deixando-os para a concorrência. Podemos afirmar que, se o princípio de Murphy se verificar, um mau processo de selecção de pessoal acarreta dois grandes erros: escolhemos os maus para nós; e deixamos os verdadeiramente bons para os nossos concorrentes.

O processo de selecção tem de ser desenvolvido de acordo com o enquadramento estratégico da organização. Os responsáveis pela selecção de pessoal poderão debater-se entre escolher os indivíduos que apresentem indicadores moderados de eficiência profissional, mas garantam uma

rápida acomodação às normas vigentes; e escolher os indivíduos que revelaram elevado potencial mas com dificuldades de acomodação às normas. A resposta a estas questões prende-se com algo mais vasto que é a dinâmica da própria instituição face à envolvente em que se insere. Organizações em envolventes demasiado turbulentas, com uma forte componente concorrencial, optarão pela flexibilidade e eficácia, enquanto que as empresas mais estáveis interessam-se sobretudo, por trabalhadores com necessidades de estabilidade e segurança, que não questionem as regras laborais instituídas.

Considerando que as características das empresas e as suas envolventes são determinantes para a escolha dos critérios de selecção, somos da opinião que o processo de selecção de pessoal deverá desenvolver-se enquadrado pela estratégia de gestão de recursos humanos, independente se a entidade responsável pela selecção é interna ou externa à organização.

3.3. A selecção de pessoal como processo

Quando se inicia um processo de escolha, existe sempre um conceito base que orienta toda a dinâmica: o critério. Qual o critério em que nos vamos basear para escolher aquele candidato e não outro?

O critério mais divulgado é o da produtividade, segundo Hunter e Schmidt (1982). No entanto, outra noção aparece habitualmente relacionada com a produtividade e que é a satisfação profissional. Afirma-se frequentemente que os índices de produtividade podem ser incrementados se controlarmos os níveis de satisfação. Nesta base são seleccionados os que, quer pelo seu desempenho anterior, quer pelas características avaliadas, apresentam maior probabilidade de se virem a sentir satisfeitos naquele posto de trabalho, ou para determinada carreira profissional. Desta forma, estima-se que alcancem níveis elevados de produtividade. No entanto, esta relação causal entre a satisfação e a produtividade carece de validação empírica

(Lima, Vala & Monteiro, 1988; Mitchell & Larson, 1987).

A magnitude da hipotética relação entre os resultados obtidos nas avaliações e o critério, eventualmente a produtividade posterior, é determinante na escolha do método de selecção. Esta capacidade do sistema (preditor) em prever, com determinado grau de confiança, o comportamento de um determinado critério é designada por validade preditiva. A capacidade preditiva dos sistemas de selecção difere consoante o critério utilizado. Podemos, então, ter um mesmo sistema de selecção com vários valores preditivos. Por exemplo, um teste de raciocínio verbal pode ser um preditor adequado do sucesso num curso de formação, mas não ser capaz de prever o desempenho futuro na função de um montador de componentes de electrónica.

Não existe um procedimento ou um conjunto de métodos de selecção que possam ser considerados universais para todas as funções, nem mesmo para uma mesma função em postos de trabalho com características diferentes, como seja o caso de diferentes empresas. Por exemplo, em instituições bancárias diferentes, a função de um caixa pode ter requisitos de desempenho diferentes. Contudo, existem conjuntos de procedimentos que são comuns a qualquer processo de selecção. O primeiro é inerente ao posto trabalho e a toda a informação que de lá se pode extrair. Descritos genericamente por análise de funções, permite-nos conhecer, não só o que o sujeito efectua, mas como o faz, que meios/equipamentos utiliza, etc. Ficamos também a conhecer o posicionamento do posto de trabalho no organograma da empresa e conseqüentemente, as relações formais que lhe são inerentes. Partindo da análise de funções, torna-se necessário conhecer as dimensões de carácter psicológico (aptidões e personalidade) que se constituem como fundamentais para o desempenho óptimo. Se, por um lado, existe um conjunto razoável de instrumentos e metodologias de apoio à análise de funções que permitem uma descrição objectiva da função, já a transposição do que se faz, para o que é necessário para que se faça bem, segue regras menos objectivas, geralmente baseadas no bom senso.

Um segundo conjunto de procedimentos diz respeito à avaliação dos atributos psicológicos exigidos a um eficiente desempenho da função a que concorre e que permitam uma adequada adaptação à organização. A avaliação das dimensões psicológicas pertinentes para a função e par uma normal adaptação, deve seguir as regras formais de utilização dos instrumentos psicológicos, assim como os instrumentos escolhidos devem possuir as necessárias características de ordem psicométrica.

Com o objectivo de conhecer de forma mais pormenorizada o processo de selecção, apresentam-se de seguida, as fases mais comuns, com base no paradigma de selecção de Smith e Robertson (1989), esquematizado na Figura 3.1.

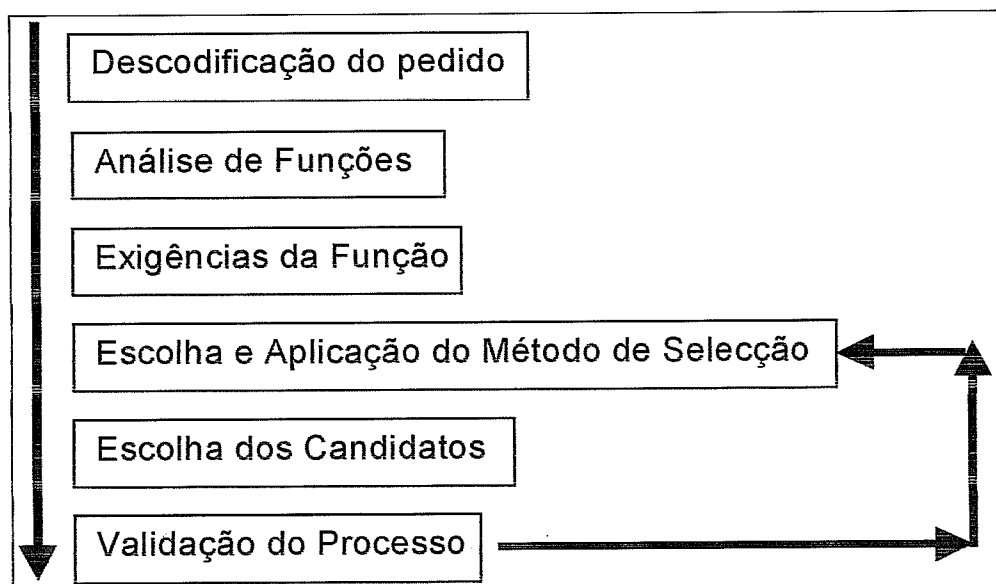


Figura 3.1. Paradigma de selecção baseado em Smith e Robertson (1989).

3.3.1. Análise de funções

A análise de funções corresponde a uma metodologia que tem por principal objectivo a

descrição, tão exacta e exaustiva quanto possível, das diversas componentes de uma função. A informação recolhida pode ser utilizada para vários fins que não somente o recrutamento e a selecção, como o levantamento de necessidades de formação, o redesenho dos postos de trabalho, entre outros.

Quando a análise de funções é desenvolvida com vista ao recrutamento e selecção, o tipo de informação a recolher vai-se consubstanciar, essencialmente nos aspectos a requerer ao sujeito para um correcto desempenho da função. É usual recolher-se informação com vista a um conhecimento preciso sobre aptidões e competências exigidas, características de ordem personalística, experiência profissional necessária, formação necessária, necessidade de deslocações (raio e frequência), conhecimento de línguas, carteiras profissionais necessárias, carta de condução, conhecimentos de informática entre outros.

A análise de funções pode ser desenvolvida segundo várias metodologias, sendo contudo, quase obrigatória a referência à técnica dos incidentes críticos de Flanagan (1954), segundo Algera e Greuter (1989). Qualquer que seja o método pelo qual o responsável do processo de selecção adquira a informação necessária, é importante que as chefias directas estejam envolvidas no processo e que concordem com a pertinência dessas especificações.

3.3.1.1. Especificações da função

Com base na análise de funções, identificam-se as dimensões pertinentes para um desempenho adequado, tornando-se necessário determinar o grau de importância a atribuir a cada um deles. Os métodos utilizados para a valorização destes atributos poderão assentar em bases puramente qualitativas, por intermédio de julgamentos subjectivos ou determinados de forma quantitativa, a partir de procedimentos estatísticos mais ou menos complexos.

Métodos qualitativos

Nos métodos qualitativos, são utilizados julgamentos. A importância de cada atributo é calculada a partir do valor que os próprios trabalhadores ou supervisores consideram ser mais determinante para um desempenho correcto da função. Por exemplo, um piloto de avião para poder executar uma manobra acrobática, em formação com outros aviões, ou para efectuar um voo de navegação com base num mapa, necessita de possuir aptidão espacial a três dimensões. No entanto, um piloto de fórmula 1, cujo desempenho profissional também é bastante exigente, poderá obter uma boa *performance* sem necessitar de uma elevada aptidão espacial tridimensional (factor essencial para o desempenho da função de piloto de avião).

O grau de importância a atribuir aos factores psicológicos exigidos ao desempenho profissional depende do tipo de função para que estamos a seleccionar, devendo a sua especificação ser realizada por alguém que conheça bastante bem a função. Um dos grandes inconvenientes na utilização de um sistema deste tipo, quando se utilizam os próprios sujeitos e/ou seus supervisores, consiste na tendência destes sobrevalorizarem os requisitos da função, distorcendo a realidade.

As especificações são geralmente pontuadas de modo a permitir uma análise comparativa entre elas, tendo como base a distribuição desses atributos na população correspondente.

Métodos quantitativos

Com os métodos quantitativos, pretende-se chegar à determinação exacta da importância de cada uma das especificações. Apesar de ser um método mais objectivo, não é frequente a divulgação de estudos a este respeito (Robertson & Smith, 1989).

A divisão da função em componentes de trabalho é uma das formas mais utilizadas. Consiste

em decompor a função em todas as suas componentes, pontuá-las e hierarquizá-las a partir da pontuação obtida. A pontuação de uma componente resulta da quantificação da frequência da sua utilização e da sua importância para uma eficiente execução da função. A etapa seguinte consiste em identificar para cada uma das componentes as bases psicológicas ou físicas envolvidas. Assim, quanto mais importante for uma componente para o desempenho profissional, maior peso deverão ter os atributos psicológicos subjacentes a essa componente.

Os pesos a atribuir às especificações podem ser determinados através do procedimento estatístico de regressão múltipla em que as especificações são consideradas como variáveis preditoras e o desempenho como variável critério.

3.3.2. Escolha dos métodos de selecção

A escolha dos métodos a serem utilizados nos processos de recrutamento e selecção deverá estar dependente essencialmente das suas características psicométricas, nomeadamente da validade relativa ao critério.

Muchinsky (1986) considera quatro critérios de avaliação dos métodos de selecção: (1) a amplitude de aplicação (possibilidade de poder ser aplicado a mais que um grupo profissional); (2) o custo; (3) o *fairness* (capacidade de o método poder ser aplicado a dois conjuntos ou mais de candidatos diferenciados em termos de sexo, raça, idade, etc., sem produzir enviesamentos); e, (4) a capacidade de predição ou validade. A estes quatro critérios, Robertson e Smith (1989) acrescentaram oito características (disponibilidade no mercado, qualificações necessárias ao utilizador, aceitabilidade pelos candidatos e impacto nos candidatos, entre outras).

3.3.3. A validade relativa ao critério

A validade relativa ao critério, e com maior pertinência a validade preditiva dos métodos de selecção é um dos aspectos mais pertinentes na tomada de decisão relativamente à escolha dos instrumentos a utilizar na selecção, podendo provocar um aumento significativo da qualidade da selecção efectuada.

A validade preditiva de um método pode ser operacionalizada pelo coeficiente estatístico de correlação (r), que oscila entre -1.0 e +1.0. Um coeficiente de correlação zero significa que não existe qualquer relação entre as variáveis. Quando a correlação entre duas variáveis é máxima, o coeficiente assume um valor quantitativo igual a -1.0 ou +1.0, o que significa que uma alteração numa das variáveis implica uma modificação directa, de igual proporção, na outra. Em termos de previsão, seria possível prever, sem margem de erro, o comportamento do sujeito na variável critério a partir do resultado na variável preditora. Em selecção de pessoal não existe, infelizmente, nenhum método com uma validade tão precisa. Os métodos de selecção considerados mais válidos possuem coeficientes de correlação que raramente são superiores a 0,50.

Graficamente, os coeficientes de validade podem ser representados por gráficos de dispersão. Quando mais elevada for a correlação maior será a tendência de os resultados se aglomerarem junto da recta de regressão. À medida que o coeficiente de validade vai diminuindo, aproximando-se de zero, a recta vai-se transformando numa elipse. Quanto menor a correlação mais imprecisa é a previsão do comportamento do sujeito na variável critério (cf. Figura 3.2).

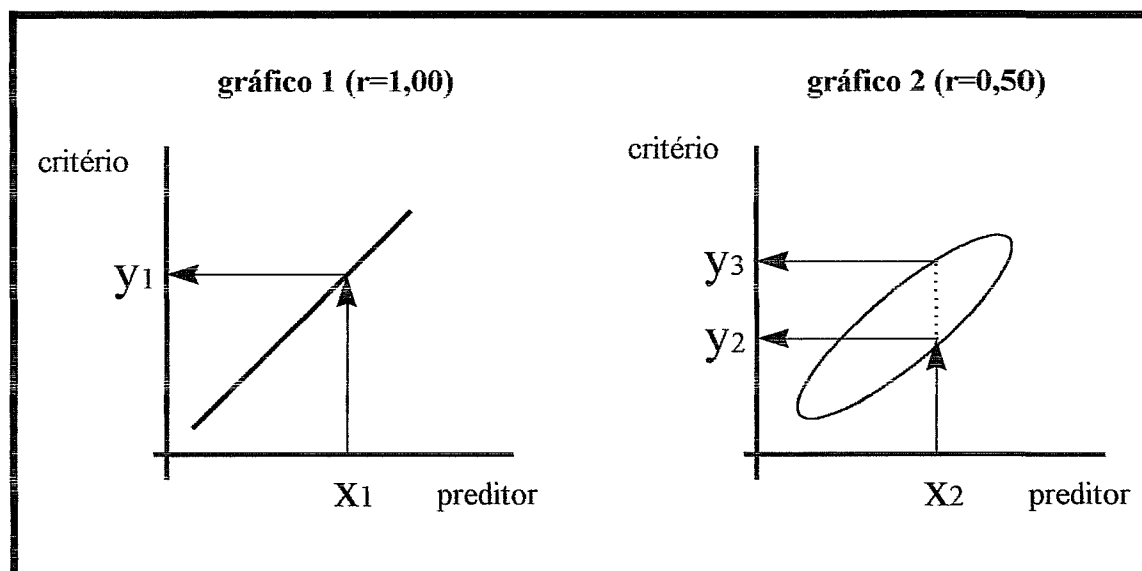


Figura 3.2. Representação gráfica de dois hipotéticos coeficientes de correlação.

Na Figura 3.2 é possível observar a representação gráfica da correlação entre preditor e critério de dois hipotéticos métodos de selecção, com validades diferentes e a sua repercussão ao nível da estimativa do critério. No gráfico 1 a predição é feita pela deslocação de um traço na vertical, que ao encontrar a recta de regressão projecta-se na horizontal, indicando o valor de critério correspondente. Assim, por exemplo, se um sujeito tiver uma nota x_1 no preditor, irá obter uma nota de y_1 na variável considerada para critério. Neste caso estamos na presença de uma associação máxima que leva a predições precisas, sem qualquer margem de erro. O preditor utilizado possui a validade máxima de $r=1,00$. No caso de considerarmos um método com validade menor que 1, que é o que acontece com todos os métodos de selecção, a recta de regressão é convertida numa elipse, que tenderá para uma circunferência à medida que o coeficiente se aproxima de zero. O gráfico 2 da Figura 3.2 mostra como a predição é processada quando estamos na presença de um método com uma validade hipotética de 0,50. Neste caso, a projecção já não se faz para um valor exacto, mas sim para um intervalo. Observando o gráfico 2 e considerando que o sujeito obteve o mesmo resultado que no

exemplo anterior ($x_2=x_1$), a nota prevista para o critério irá agora oscilar entre y_2 e y_3 . O intervalo da predição será tanto maior quanto menor for o coeficiente de validade, ou seja, quanto mais se aproximar de zero. Estando o conceito de validade preditiva directamente associado à capacidade de previsão, é fácil de entender que esta é uma das características mais importantes quando procedemos à escolha de um método para o incluirmos no processo de selecção.

3.3.4. Restrições aos coeficientes de validade

Os coeficientes de validade encontrados nos estudos de validade relativa ao critério não são geralmente elevados devido a um conjunto de factores (Cronbach, 1990). Podemos distinguir dois grandes tipos de restrições relativas aos coeficientes de validade. O primeiro diz respeito à falta de qualidade psicométrica dos critérios, ou seja, das medidas utilizadas para avaliar os comportamentos posteriores, como sejam as escalas de avaliação de desempenho. É frequente encontrar escalas de avaliação de desempenho com baixa fidelidade e fraco poder discriminativo. Mostram-se pouco objectivas provocando flutuações de resultados aquando da sua repetição ou quando são utilizadas por diferentes avaliadores. As escalas construídas com itens ancorados a comportamentos, permitem uma maior objectividade na medição dos factores que se pretende avaliar (Caetano, 1996). As escalas com comportamentos ancorados aos diferentes níveis de cada um dos itens, poderão originar melhores qualidades métricas, comparativamente às escalas mais tradicionais em que o avaliador tem espaços em branco para marcar a cruz relativamente a um determinado adjectivo ou característica do sujeito. As escalas ancoradas a comportamentos, permitem uma avaliação mais objectiva e mais específica do desempenho do sujeito, não sendo tão sensível às diferenças entre os avaliadores.

Outro tipo de restrições relaciona-se com a amostragem. Quando se fazem estudos de validade, tendemos a considerar um grupo restrito de pessoas. Se existem 400 candidatos para 20 vagas, o nosso estudo vai centrar-se nas 20 pessoas que entraram e não nas 400 que concorreram. Logicamente que as 20 pessoas que foram admitidas deverão ser as melhores do grupo de 400. Atendendo a que esses 20 candidatos resultam da selecção sobre uma amostra maior, os seus resultados tenderão a ser muito homogéneos originando baixos coeficientes de validade. Se os candidatos com resultados muito fracos fossem admitidos para a organização, muito provavelmente iriam ter um desempenho profissional negativo, o que a acontecer, provocaria maior variabilidade dos resultados, e consequentemente coeficientes mais elevados. Como será fácil de entender, as organizações não estão dispostas a admitir um número de candidatos superior ao de vagas, exclusivamente para fins de investigação e consequentemente suportar os encargos resultantes deste tipo de estudos.

3.3.5. Métodos de selecção e validade preditiva

A validade preditiva é um factor crucial quando se pretende avaliar o valor dos métodos de selecção, não obstante da importância de outros critérios. Tomando por amostra os estudos de validade publicados entre 1964 e 1982 nas revistas com maior divulgação deste tipo de trabalhos, nomeadamente o *Journal of Applied Psychology* e o *Personnel Psychology*, Schmitt, Gooding, Noe e Kirsch (1984) apresentam os métodos mais conhecidos e/ou mais utilizados em selecção de pessoal e os respectivos coeficientes de validade relativa ao critério que foram encontrados (a este respeito ver também Bártolo Ribeiro, 1995, 1996).

No Quadro 3.1 encontram-se diferentes métodos de selecção e os correspondentes coeficientes de validade encontrados nas pesquisas de Robertson e Smith (1989). A amplitude dos coeficientes encontrados reflecte os coeficientes mínimos e máximos desses preditores

quando correlacionados com diferentes critérios e em diferentes contextos em que os estudos foram efectuados.

Os coeficientes transcritos no Quadro 3.1 são tendencialmente baixos, o que numa primeira análise poderá provocar alguma desconfiança relativamente à utilização desses métodos na selecção de pessoal. Algumas das causas para os coeficientes serem tão baixos já foram referidas anteriormente.

Quadro 3.1. Amplitude dos coeficientes de validade para um conjunto de métodos de selecção segundo I. Robertson e M. Smith (1989).

Método de selecção	Amplitude dos coeficientes por critério
<i>Work sample</i>	0.38 - 0.54
Aptidões combinadas	0.53
<i>Assessment centres</i>	0.41 - 0.43
Avaliação por pares / supervisores	0.43
Aptidão mental geral	0.25 - 0.45
Dados biográficos	0.24 - 0.38
Referências	0.17 - 0.26
Entrevistas	0.14 - 0.23
Avaliação da personalidade	0.15
Avaliação dos interesses	0.10
Auto-avaliação	0.15
Grafologia	0.00

Pelo lado positivo, destacam-se métodos menos divulgados entre a gestão de recursos humanos nas empresas portuguesas, como sejam os *Work Sample* e os *Assessment Centre*, que possuem dois campos de aplicação perfeitamente diferenciados. O primeiro, consistindo numa amostra de trabalho, destina-se a uma população cujo trabalho seja essencialmente manipulativo e de pouca interacção pessoal. Os segundos são sobretudo escolhidos para a selecção de quadros superiores.

O Quadro 3.1 permite-nos, ainda, observar os baixos índices de validade que possuem alguns métodos bastante utilizados, como sejam a avaliação da personalidade, as Entrevistas e as Referências. Por exemplo, os métodos Entrevista e Referências aparecem como os mais utilizados na selecção de quadros da indústria britânica, entrando em 81% e 67% dos processos, respectivamente (Robertson & Smith, 1989). Os Testes Cognitivos, de Personalidade, os Dados Biográficos e os *Assessment Centres* são outros métodos utilizados, mas não ultrapassam 5% dos casos, ainda para a mesma população.

Dos métodos com baixos resultados em termos de validade, destacam-se as entrevistas, que, no entanto, são dos métodos de selecção mais utilizados. Estima-se que, no mercado português, a entrevista seja utilizada em mais de 80% dos processos de selecção (Bártolo Ribeiro, 1995). Sendo um método a que os técnicos atribuem grande importância, têm sido desenvolvidos esforços para contrariar esta tendência. Dos diferentes tipos de entrevista, o mais conhecido é o tipo psicológico. Tem por objectivo conhecer a motivação, a história familiar e profissional do sujeito, entre outros aspectos. Esta modalidade de entrevista possui, de facto, uma validade baixa. No entanto, as entrevistas mais estruturadas, designadas por situacionais e comportamentais, podem apresentar coeficientes mais elevados, na ordem de 0,30 e 0,40 (Wiesner & Cronshaw, 1988).

Os testes psicométricos, especificamente os que avaliam as aptidões mentais, possuem valores de validade que oscilam entre 0,25 e 0,53. Estudos recentes (Olea & Ree, 1994; Ree & Earles, 1992; Ree, Earles & Teachout, 1992,1994), indicam o *factor g* como o principal preditor do desempenho profissional.

Relativamente à avaliação da personalidade, estamos a atravessar uma fase de transição. Em simultâneo com os modelos factoriais clássicos, começam a surgir os testes de personalidade que assentam no modelo dos *Big Five Factors*, e que, eventualmente, vão permitir aumentar os tão magros coeficientes de validade atribuídos aos antigos testes. As cinco dimensões que

constituem este modelo, são: Extroversão (*Extraversion*), Amabilidade (*Agreeableness*), Consciencioso (*Conscientiousness*), Neuroticismo (*Neuroticism*), e Abertura à experiência (*Openness*), segundo Barrick e Mount (1991) e Digman (1990). Ainda dentro da avaliação da personalidade, estão em desenvolvimento/aperfeiçoamento os chamados testes de integridade que permitem avaliar comportamentos contra-produtivos, nomeadamente a tendência para o absentismo e para o roubo (Ones, Schmidt & Viswesvaran 1994; Ones, Viswesvaran & Schmidt, 1993). Curiosamente, estudos recentes mostram uma forte relação entre as dimensões avaliadas por estas provas e a adesão à cultura organizacional da empresa, assim como à dimensão Consciencioso do modelo dos *Big Five Factors*.

Os *Work Samples* são amostras de trabalho, em que os candidatos têm de realizar pequenos exercícios directamente relacionados com a função a que estão a concorrer, como, por exemplo, passar a computador um texto que inclua tabelas e gráficos para uma função administrativa. A magnitude dos coeficientes de correlação destas provas com o desempenho das tarefas no início de actividade é elevada mas tende a diminuir à medida que o sujeito vai evoluindo nesse posto de trabalho. As correlações permitem, assim, prever o grau de adaptação do sujeito às novas funções.

A Informação Biográfica diferencia-se do “tradicional” *Curriculum Vitae* e possui uma validade bastante satisfatória. Consiste em questionários, onde constam perguntas que envolvem a experiência profissional dos candidatos. São questões contextualizadas, isto é, questões que têm a haver não só com a função que ele desempenha mas também com a que se está a candidatar.

Já foi referido anteriormente que os coeficientes de validade tendem a reflectir a robustez psicométrica dos critérios utilizados, para além da própria capacidade dos preditores. Em psicologia organizacional, o critério por excelência tem sido a eficácia, geralmente avaliada pela *performance* no trabalho. A conceptualização deste constructo tem sido efectuada em

duas diferentes perspectivas. Para Hunter (1986) o rendimento no trabalho deve ser tido como uma variável unidimensional onde os sujeitos se situam. Drasgow e colaboradores (1993) não partilham desta opinião e são mais de acordo que o rendimento tem uma natureza multidimensional. Prieto (1993), por outro lado, considera que a perspectiva de olharmos o sujeito exclusivamente pelo rendimento individual é redutora da realidade. Existem funções em que o rendimento individual pouco interessa uma vez que o sujeito está integrado numa equipa e a produtividade ultrapassa as variáveis intrínsecas aos indivíduos.

Os resultados obtidos nos estudos de validade não podem ser generalizados de uma determinada situação organizacional para outras onde perspectivemos trabalhar. Os estudos de meta-análise de Ree e Earles (1991b) revelam que a variável que consegue maior estabilidade na predição do rendimento no trabalho é a inteligência geral. Em estudos de carácter mais aplicado a situações específicas, os resultados vão no mesmo sentido (Olea & Ree, 1994; Ree & Earles, 1991a; 1991b; Ree, Earles & Teachout, 1994). Assim sendo, parece que pouco mais haveria a trabalhar ao nível da predição do rendimento. Esta noção deixa, no entanto, de fazer sentido se olharmos para o rendimento numa perspectiva mais alargada. Por exemplo, Drasgow e colaboradores (1993) fazem referência a diferentes dimensões que estão presentes no comportamento do líder, pelo que, forçosamente, a eficácia será estimada por mais que um preditor. Se olharmos apenas para a componente essencialmente cognitiva e técnica que contribui para o rendimento do sujeito, talvez o *factor g* seja suficiente para o estimar. Contudo se a nossa perspectiva for mais alargada, então outros preditores devem ser considerados.

Ree e Earles (1994) no seu capítulo "*The ubiquitous predictiveness of g*" fazem uma síntese dos vários estudos que demonstram o poder do *factor g* na predição do sucesso, não só na formação específica para o desempenho de uma função, como também, no próprio desempenho das diferentes actividades desse posto de trabalho. Também abordam o impacto

do *factor g* para diferentes postos de trabalho. Se realmente o *factor g* tem esta capacidade de estar omnipresente na predição, tanto no sucesso da formação como do desempenho, coloca-se a questão da pertinência de utilização de medidas de avaliação de outras dimensões psicológicas. No entanto, os próprios autores não são radicais, pois acreditam que existam outros factores que permitam aumentar a percentagem de variância explicada das variáveis critério. Referimo-nos, por exemplo, a variáveis do plano personalístico individual ou relacional.

Num estudo para demonstrar o que seria possível incrementar, em termos de validade preditiva, ao que já é possível com o *factor g*, Ree e Earles (1989) conduziram um estudo com a *Armed Services Vocational Aptitude Battery (ASVAB)*, que avalia essencialmente *factor g*, composta por testes que medem múltiplas aptidões. Neste estudo, o *g* psicométrico foi responsável pelo sucesso nos cursos de formação para 89 funções técnicas que envolveram mais de 78.000 sujeitos. A correlação média do *g* com o critério foi de 0,76 após correcção da restrição da amplitude (*range restriction*). Quando as equações de predição tomaram em consideração também os factores designados por *não-g*, a correlação aumentou, em média, apenas 0,02.

Mesmo quando os estudos são delineados para averiguar a capacidade preditiva do *factor g* e dos *factores não-g* do sucesso no desempenho profissional, obtêm-se resultados muito idênticos aos encontrados com o critério sucesso no treino. Numa investigação desenvolvida por Morales (1991, citado por Ree & Earles, 1994) foram analisados cerca de 5.500 oficiais recém-formados para estudar em que medida os *factores g* e *não-g* prediziam o sucesso tanto no treino da navegação e pilotagem como no próprio desempenho. Os resultados encontrados foram muito semelhantes aos do estudo com funções técnicas. O *g* psicométrico continuava a ser o melhor preditor. Quando os coeficientes foram corrigidos, para serem representativos daquela amostra particular, as correlações atingiram um valor médio de 0,33. O contributo

devido a *factores não-g* não ultrapassou os 0,05, em termos da equação de predição.

Mesmo quando são utilizados diferentes tipos de critério para medir o desempenho na função, os resultados relatados vão no mesmo sentido (Ree, et al. 1992). O *factor g* continua a ser o melhor preditor, com um coeficiente de correlação corrigido médio de 0,44. Os *factores não-g* apenas acrescentaram 0,06 ao modelo de predição. Neste estudo foram utilizados três critérios diferentes e uma amostra aproximada de 1.500 elementos da aviação com cerca de 2 anos de experiência. Os três tipos de critério utilizados foram: (1) resultados da resolução de amostras de trabalho desenvolvido no próprio posto de trabalho; (2) avaliações de desempenho com base em entrevistas técnicas em que os sujeitos eram solicitados a explicar, passo por passo, como efectuavam determinadas tarefas; e, (3) avaliações dos seus supervisores.

Carretta (1990), num estudo desenvolvido ainda em contexto aeronáutico, mostrou que mesmo quando se utilizam testes psicomotores conjuntamente com testes de *factor g*, os primeiros pouco adiantam em termos de predição do critério relacionado com o desempenho da pilotagem. Os factores psicomotores avaliados pela bateria computadorizada americana apenas acrescentaram 0,02 para além do *factor g* ao estimar o resultado final no treino de pilotagem em termos de apto/inapto. Estes resultados não são, no entanto, confirmados por Bártole Ribeiro (1992), num estudo de validação de uma bateria de preditores do sucesso no estágio de adaptação ao voo em candidatos ao curso de pilotagem da Força Aérea Portuguesa, em que a variável preditora exactidão na prova de coordenação motora tem uma correlação de 0,45 (coeficiente não corrigido) com o nível de rendimento em voo. Existem, contudo, diferenças entre os dois estudos. No de Bártole Ribeiro, os factores psicomotores foram avaliados por intermédio de instrumentos clássicos e não informaticamente. A segunda diferença consiste no critério. Enquanto o critério de Carretta diz respeito ao treino na pilotagem, o de Bártole Ribeiro corresponde ao desempenho num número diminuto de horas (10 horas), servindo apenas para avaliar o grau de adaptação do sujeito ao voo.

O estudo de Ree e Carretta (1994) aprofunda este problema, relacionado com a não independência dos factores psicomotores relativamente ao *factor g*. Encontraram correlações na ordem de 0,73 entre os testes ditos de *factor g* (medidos por intermédio da bateria ASVAB) e testes psicomotores da bateria *Basic Attributes Tests* (BAT), validada para a selecção de candidatos ao treino de pilotagem da Força Aérea Americana. Os coeficientes de correlação, após correcção, mostraram que os dois tipos de medidas não são totalmente independentes. A razão apontada para esta não independência pode ser justificada pelo facto de a componente *g* das notas nos testes psicomotores não ser independente da componente *g* dos testes de papel-e-lápis.

3.4. Validade *versus* utilidade

Se por um lado os coeficientes de validade são fáceis de entender por quem possui conhecimentos do ponto de vista estatístico, torna-se mais difícil explicar as consequências práticas da aplicação de métodos pouco válidos em termos de selecção de pessoal a um gestor menos habituado a trabalhar com estatística.

Existe, contudo, a possibilidade de realizar estudos que, embora baseados no conceito de validade relativa ao critério, utilizam uma linguagem mais económica do que estatística. Estes estudos são designados por estudos de utilidade. Os estudos de validade distinguem-se dos de utilidade porque, enquanto os primeiros nos informam, com coeficientes estatísticos, do grau de previsão de um comportamento com base nas medições efectuadas com os preditores, os estudos de utilidade dão-nos a mesma informação, mas retratada de forma económica. Reforçando esta ideia, diríamos que os estudos de validade preditiva respondem à pergunta: qual o grau de probabilidade de este sujeito, que foi presente a este conjunto de provas, passar nesta formação, ou atingir um determinado nível de desempenho? Os estudos de utilidade, por

sua vez, são desenvolvidos para responderem a uma pergunta mais pragmática em termos de gestão: o que é que a organização ganha em admitir este candidato e não um outro que tenha ficado de fora?

A diferença entre os estudos de validade e os de utilidade não reside exclusivamente no tipo de unidade utilizada - coeficiente de correlação para os primeiros e custo financeiro para os segundos. Não é assim tão simples, até porque os estudos de utilidade são desenvolvidos com base nos estudos de validade. Os estudos de utilidade são mais recentes que os de validade e praticamente ainda não saíram dos livros para serem testados no campo. São raros os estudos de utilidade com dados reais e as fórmulas que aparecem nos corpos teóricos diferem de autor para autor (ver Cascio, 1982; e, Arnold, Cooper & Robertson, 1995). De qualquer modo, os estudos de utilidade possibilitam uma comunicação mais eficaz com os gestores. Eles podem influenciar as tomadas de decisão por se basearem em valores quantitativos, se bem que os critérios utilizados para a obtenção desses valores careçam de maior objectividade.

Se, por um lado, o conceito psicométrico de validade preditiva aplicado aos instrumentos e métodos de selecção é um dado extremamente importante do ponto de vista estatístico, torna-se pouco útil quando se tem de o transmitir a alguém que não domine a linguagem estatística. Neste sentido, o conceito de utilidade está mais adaptado a uma linguagem de gestão. Os estudos de utilidade têm por objectivo a quantificação económica dos ganhos em se utilizar um determinado método de selecção. Segundo Brogden (1949), a análise de utilidade dos métodos de selecção pode basear-se em 4 parâmetros: (1) variabilidade do rendimento no trabalho medido em unidades económicas; (2) rendimento no método pelos candidatos seleccionados; (3) custo do procedimento de selecção; e, (4) validade do método. Desta forma, mesmo preditores com baixa validade podem ser de grande utilidade se o valor do rendimento no trabalhador for elevado.

Cascio e Ramos (1986) propõem um sistema que permite avaliar a *performance* no trabalho

em termos económicos: *Cascio-Ramos Estimate of Performance in Dollars* (CREPID). A avaliação do rendimento do sujeito passaria a ser feita em termos do valor da sua actividade e não simplesmente das clássicas, e tão pouco objectivas, avaliações de desempenho. Arnold e colaboradores (1995), por outro lado, apresentam uma fórmula que permite determinar a relação custo-eficácia da utilização de um determinado método. A utilidade é fortemente influenciada por dois parâmetros: (1) a taxa de selecção; e, (2) o benefício económico provocado pelo aumento de rendimento no trabalho. A taxa de selecção traduz-se pelo quociente entre o número de vagas que a empresa tem para oferecer e o número de candidatos. Tendencialmente este número é inferior a um, uma vez que, na maioria dos casos, o número de candidatos é superior ao número de vagas. Nestas condições é fácil demonstrar que as empresas ganham quando optam por métodos de selecção válidos. Por exemplo, quando a taxa de selecção é 0,1, ou seja, são considerados 10 candidatos para uma vaga, será esperado que a organização tenha benefícios económicos, admitindo que o candidato seleccionado obteve melhores resultados que os restantes nove. Mesmo no caso de uma taxa de selecção ser superior a um, isto é, quando a empresa possui mais vagas que candidatos, deverá, em nossa opinião, optar-se por admitir os candidatos somente depois de terem passado por um processo de selecção, como forma de evitar a admissão de pessoas não adaptáveis àquele posto de trabalho e aos valores da organização.

O segundo parâmetro, anteriormente designado por benefício económico, é composto, de acordo com a fórmula de Arnold e colaboradores (1995), por quatro variáveis específicas ao processo de selecção e ao método utilizado e que são: (1) média padronizada obtida pelo grupo seleccionado; (2) coeficiente de validade do método de selecção escolhido; (3) número de anos que se estima que o(s) candidato(s) permaneça(m) na organização; e, (4) número de candidatos a seleccionar. A primeira variável, que corresponde à média padronizada obtida pelo grupo seleccionado, diz respeito à comparação da média obtida por este grupo relativamente à população trabalhadora que já se encontra na organização. Para cálculo da

média padronizada considera-se a fórmula das chamadas notas reduzidas e que neste caso se expressa da seguinte forma: $z(m_g) = (m_g - m_p) / \sigma_p$. Em que $z(m_g)$ corresponde à nota padronizada da média do grupo seleccionado; m_g a média do grupo seleccionado; m_p a média do grupo dos trabalhadores; e σ_p o desvio-padrão do grupo dos trabalhadores. Esta média padronizada representa o afastamento do grupo seleccionado relativamente à população trabalhadora em termos de desvio-padrão. Por intermédio das tabelas da curva normal é possível converter as notas padronizadas em proporção de sujeitos, uma vez que se parte do princípio que o atributo psicológico avaliado se distribui segundo a curva normal. Se a média padronizada obtida for negativa, isso significa que a média obtida naquele atributo pelo grupo seleccionado é inferior à média da população, o que logicamente significa que aquele lote de indivíduos escolhidos irá ter um desempenho na função inferior, em termos médios, aos que já lá trabalham. Esta dedução será tanto mais verdadeira quanto mais válido for o método de selecção em uso.

Apesar de não existir uma correspondência rigorosa entre a magnitude da diferença entre as médias dos dois grupos (o dos seleccionados e o dos trabalhadores), e o ganho em termos económicos, é frequente utilizar-se uma regra de algibeira que atribui um ganho de 40% sobre o salário por cada desvio-padrão de rendimento. Por outras palavras, significa que se a empresa admitir um conjunto de sujeitos com um rendimento médio equivalente a um desvio-padrão acima da média dos actuais trabalhadores, será esperada uma mais-valia, em termos de produtividade, que se traduz num aumento equivalente a 40% do salário.

A segunda variável diz respeito à validade preditiva e que já foi anteriormente desenvolvida. As restantes variáveis são fáceis de determinar e dependem quase em exclusivo da estratégia da empresa. Enquadrado na estratégia organizacional será determinado quantos sujeitos se pretendem admitir e quanto tempo se espera que eles permaneçam na instituição. Embora esta última variável corresponderá, logicamente, a uma estimativa a partir da experiência da organização.

A título de exemplo, aplicámos a fórmula desenvolvida por Arnold e colaboradores (1995) a duas situações hipotéticas em que mantivemos constantes os valores de todas as variáveis à excepção do valor correspondente à validade do método de selecção (ver Quadro 3.2). Os valores da fórmula considerados para os dois métodos foram os seguintes: (1) média padronizada pelo grupo seleccionado [0,8]; (2) coeficiente de validade do método de selecção [$r=0,10$ ao método A] e [$r=0,30$ para o método B]; (3) nº de anos que se prevê que o candidato vá ficar na empresa [3]; e (4) nº de candidatos que se vai admitir [2]. Considerámos um benefício económico de 5.600 €, correspondente a 40% de um salário anual de 14.000€.

Quadro 3.2. Comparação entre dois métodos com diferentes validades num estudo de utilidade.

Var.		Método A	Método B
(1)	Média padronizada pelo grupo seleccionado [0,8]	0,80	0,80
(2)	Validade	0,10	0,30
(3)	Nº de anos	3	3
(4)	Nº de candidatos a admitir	2	2
(5)	(1) * (2) * (3) * (4)	0,48	1,44
(6)	Benefício económico 40% * 14.000€ (salário anual)	5.600€	5.600€
	(5) * (6)	2.688€	8.064€

Mantendo constantes todas as variáveis, à excepção do coeficiente de validade (var. 2 do Quadro 3.2), verifica-se que o Método A apresenta um ganho de 2.688€ contra os 8.064€ conseguido pelo Método B (cf. Quadro 3.2). A diferença entre eles aponta para uma vantagem de cerca de 5.376€ em se utilizar o Método B para a selecção daqueles dois colaboradores por um tempo mínimo de três anos. O ganho conseguido pelo facto de se ter aplicado um método

mais válido pode compensar a organização, a curto-prazo, dos custos inerentes à utilização de um método, eventualmente mais dispendioso.

Apesar de não ser prática corrente a realização de estudos de utilidade, é importante, pelo menos, a simulação de situações para a estimativa dos ganhos económicos com a utilização de métodos de selecção com melhores capacidades preditivas.

Parte Empírica

Introdução aos Estudos Empíricos

Na parte prática deste trabalho efectuaram-se dois estudos que apesar de aparentemente diferentes possuem idênticas características base. Primeiro, foram aplicados num contexto real de selecção. Os participantes nos dois estudos efectuaram os respectivos testes em consciência de que os resultados obtidos teriam uma influência directa na decisão final relativamente à sua admissão para a função em concurso. Segundo, ambos avaliam dimensões cognitivas e são de aplicação informática. Por último, mas não menos importante, ambos os instrumentos utilizados nos dois estudos permitem o registo dos tempos de latência às respostas aos itens.

O grande objectivo dos estudos empíricos que a seguir se apresentam consistiu em conhecer melhor o comportamento dos sujeitos na realização de testes informatizados, a partir da dimensão velocidade mental, medida através dos tempos de latência.

Contudo, foram colocadas outras hipóteses específicas a cada um dos estudos. No primeiro, com base na aplicação de um teste de figuras embebidas inserido na bateria de selecção de candidatos ao curso de pilotagem militar, pretendeu-se conhecer a relação entre a exactidão das respostas dadas pelos sujeitos e os tempos de latência às mesmas. Face ao facto de o instrumento ser de resposta múltipla e de o sujeito saber o tempo restante para terminar o teste, colocou-se a hipótese de interferência de factores não cognitivos na estratégia de resolução de tarefas cognitivas. Deste modo, espera-se que os sujeitos tendam a ser mais lentos a responder quando existe maior incerteza na resposta a dar.

No segundo estudo trabalhou-se com um dos testes mais divulgados ao nível da avaliação da inteligência geral: As Matrizes Progressivas de Raven na sua forma *Standard* mas de aplicação em suporte informático. A versão informatizada possui diferenças relativamente à sua congénere de papel-e-lápis, pelo que foi nossa intenção conhecer melhor a evolução do sujeito ao longo da prova, em termos de estratégia de controlo do esforço e tempo. Também procurámos, complementarmente, testar duas hipóteses. A primeira analisa a relação entre dificuldade e complexidade da tarefa; estando a complexidade associada ao tempo de latência às respostas será de esperar uma relação forte entre os tempos gastos nas respostas dadas e a correcção das mesmas. A segunda, relacionada com a capacidade de armazenamento de informação na memória de trabalho, pressupõe que os sujeitos mais inteligentes se diferenciem dos menos inteligentes em termos de rapidez com que processam informação. Consecutivamente, será esperado que os sujeitos com diferentes capacidades intelectuais gerais obtêm idênticos resultados em termos de tempos de latência quando o grau de dificuldade é muito reduzido, ou seja, quando o item tem uma elevada probabilidade de ser respondido correctamente. Por outro lado, quanto maior o nível de complexidade dos itens maior a probabilidade de ocorrência da diferença entre os grupos com diferentes níveis de rendimento intelectual global. Espera-se que o grupo mais favorecido consiga tempos médios de resposta significativamente inferiores.

4. Estudo 1 - Estratégia nos Testes de Escolha Múltipla

A realização deste estudo teve por objectivo conhecer melhor um dos testes inseridos na bateria de selecção de candidatos ao curso de pilotagem militar, a partir dos tempos de latência às respostas aos itens. Foi utilizada a prova “Padrões” avaliadora da *flexibility of closure* e de resposta por escolha múltipla. Pretendeu-se conhecer a relação entre a exactidão da resposta dada pelo sujeito e os tempos de latência às mesmas. Também foi objecto de estudo a eventual interferência de factores não cognitivos na execução de itens de escolha múltipla. Como hipótese, espera-se que os sujeitos tendam a ser mais lentos a responder quando existe maior incerteza na resposta a dar.

Considerando que esta prova foi aplicada num contexto real de selecção, e incluída numa bateria com outras provas, procurámos caracterizar melhor a natureza das variáveis exactidão e velocidade de realização, por intermédio da correlação com dois testes psicométricos: um de inteligência geral e outro de rapidez perceptiva.

4.1. Instrumento

Foi escolhido o teste “Padrões” por ser uma prova de escolha múltipla e porque se pretendia conhecer melhor esta prova ao nível da selecção de pessoal navegante. É um teste informatizado do tipo das figuras embebidas, desenvolvido por Eugene Burke em 1995, sendo uma medida de *perceptual closure*. Solicita-se ao sujeito que identifique uma figura-alvo em

dois padrões desorganizados visualmente que representam um fundo não estruturado (People Technologies, 2000).

Se considerarmos a tarefa que é solicitada ao sujeito, podemos enquadrar o teste “Padrões” na gama dos testes de figuras ocultas, já que o sujeito tem de identificar se a figura-alvo que funciona como estímulo se encontra em alguma, nas duas, ou em nenhuma das figuras complexas desestruturadas que lhe são apresentadas (ver Figura 4.1).

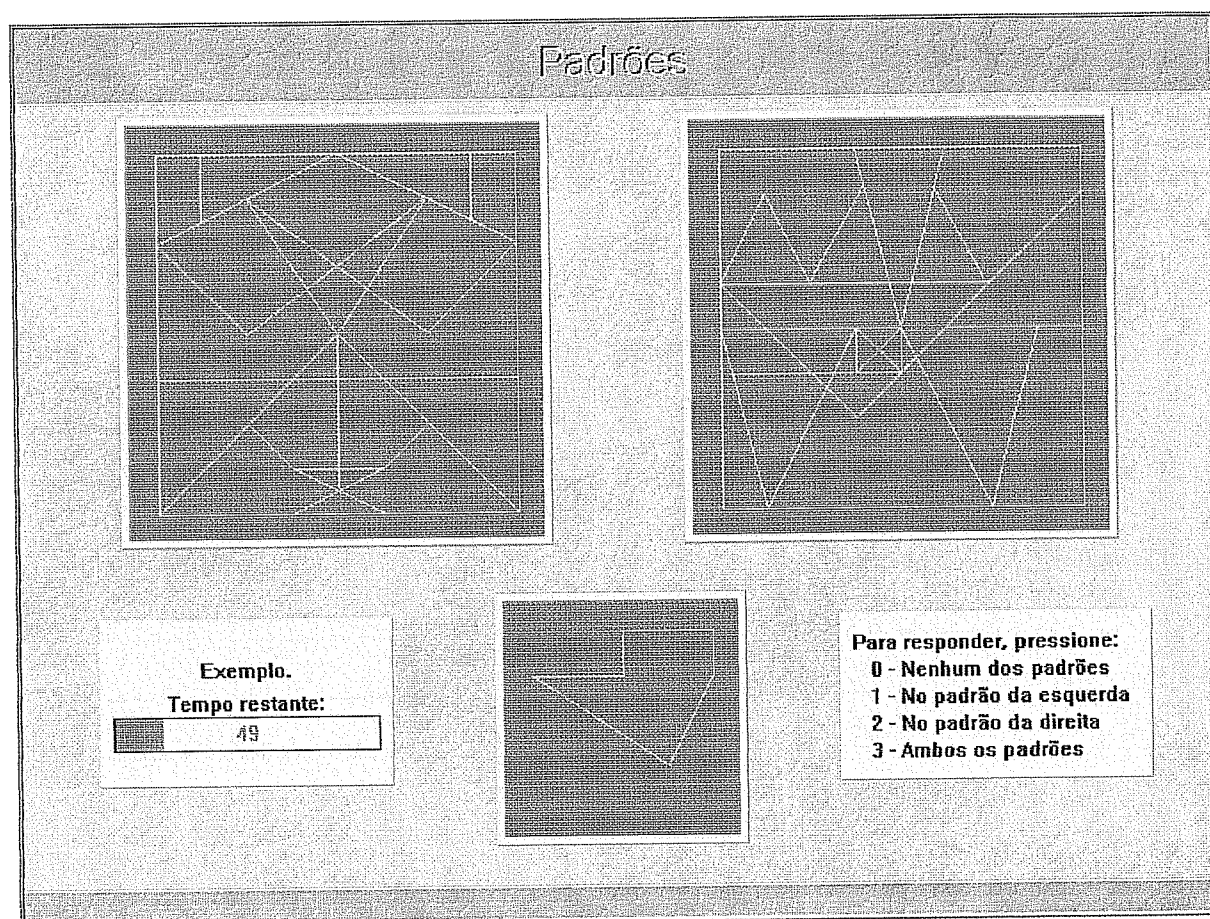


Figura 4.1. Exemplo de um item do teste “Padrões”

A aptidão avaliada por este teste insere-se no domínio da percepção visual, de acordo com a categorização de Carroll (1993), segundo a qual o factor avaliado é o *Flexibility of Closure* que deriva do factor designado por *Gestalt Flexibility* segundo French (1951, citado por

Carroll, 1993). A evidência e singularidade deste factor ainda está longe de uma total clarificação (Kline, 2000).

Para Carroll (1993), *flexibility of closure* corresponde a um factor directamente relacionado com a pesquisa, num campo perceptivo, de uma forma espacial retida anteriormente e em que o campo perceptivo funciona como distrator. Este factor distingue-se do *closure of speed*, na medida em que neste último a forma espacial procurada pelo sujeito não é apresentada antecipadamente. Ele apenas sabe que o campo perceptivo onde se encontra a efectuar a pesquisa contém uma forma espacial com significado. A complexidade da tarefa advém do facto de a figura-alvo se encontrar camuflada num contexto, sem uma clara estruturação lógica.

Os trabalhos de Witkin e Goodenough (1981) realçam a importância de provas deste tipo (figuras embebidas/embutidas – *embedded figures*) para a avaliação da estrutura cognitiva e do desempenho intelectual, mas também de aspectos da personalidade, estando fortemente relacionada com o “estilo cognitivo” do sujeito. Consideraram estas provas avaliadoras da dependência/independência de campo. Kline (2000) mostra-se bastante crítico relativamente a esta designação de factor para este tipo de provas, já que considera que se enquadram dentro dos testes de *flexibility of closure*.

Comparativamente com a generalidade dos testes de *flexibility of closure* e *closure of speed*, o “Padrões” tem a singularidade de permitir que o sujeito revisualize a figura-alvo enquanto está a tentar resolver o item (ver Figura 4.1). O facto de a figura-alvo se manter presente durante a realização do item, contrariamente aos testes de *flexibility of closure*, em que o sujeito primeiro retém a figura-alvo e só depois é que a vai pesquisar no campo perceptivo, pode eventualmente, exigir ao sujeito menor utilização da memória que será compensada por factores de comparação perceptiva entre a figura-alvo e as várias configurações perceptivas que compõem o campo distrator que podem, ou não, ser idênticas à figura-alvo.

A prova “Padrões” é constituída por doze itens, que são apresentados ao sujeito numa sequência pré-determinada. Os candidatos têm de responder dentro do tempo restrito de um minuto a cada um dos doze itens. Esta tarefa também inclui pressão do tempo, já que é apresentado ao sujeito, numa janela própria, em termos gráficos e digitais, o tempo restante em contagem decrescente (ver Figura 4.1). A barra que indica de forma digital o tempo em falta, muda para encarnado nos últimos 10 segundos. O reduzido tempo associado ao facto de o sujeito poder, em qualquer momento, verificar o tempo em falta contribuem para um aumento da ansiedade na realização da tarefa. A pressão do tempo vivida pelo sujeito pode constituir um factor não cognitivo significativo e limitador do sucesso da tarefa. Quando se utiliza esta prova na selecção de pilotos, ela pode constituir um preditor válido, já que no desempenho real da função são executadas várias tarefas extremamente exigentes do ponto de vista perceptivo, em que muitas das decisões têm de ser tomadas com constrangimentos de tempo.

O candidato deverá assinalar a sua resposta num teclado numérico, premindo a tecla zero (0) para indicar se a figura-alvo não se encontra em nenhuma das figuras complexas; a tecla um (1), se a figura-alvo se encontra na figura complexa da esquerda; a tecla dois (2), se a figura-alvo se encontra na figura complexa da direita; e a tecla três (3) se a figura-alvo se encontra em ambas as figuras complexas. Como para cada item existem quatro hipóteses de resposta, o sujeito tem uma probabilidade de 0,25 de acertar, se responder ao acaso.

O sistema calcula duas notas globais para cada candidato: o número de itens respondidos correctamente, que corresponde ao resultado final (**P_Corr**); e a média dos tempos de resposta a todos os itens respondidos dentro do tempo estipulado, independentemente de terem sido respondidos correctamente ou não (**P_Tempo**). O tempo é registado em centésimos de segundo.

Para além das notas globais, o sistema também armazena informação relativamente a cada um

dos itens (i), nomeadamente: (1) a resposta dada nesse item (**P_Respi**); (2) se a resposta dada está correcta ou não (**P_Corri**); e (3) o tempo de latência na resposta ao item (**P_Timei**), ou seja o tempo, em centésimos de segundo, desde a apresentação do item até o sujeito pressionar a tecla correspondente à resposta escolhida. Para melhor compreensão das variáveis consideradas no estudo deve ser consultado o Quadro 4.1.

Quadro 4.1. Identificação das variáveis do teste de Padrões

Código	Designação	Tipo de Métrica	Nota Mínima	Nota Máxima
P_Corr	Número total de itens respondidos correctamente	Intervalar	0	12
P_Tempo	Média dos tempos de resposta a todos os itens respondidos dentro do tempo normal	Intervalar (centésimos de segundo)	1	5999
P_Respi	Resposta dada pelo sujeito no item especificado em i	Nominal (0, 1, 2 ou 3)	—	—
P_Corri	Resultado obtido pelo sujeito no item especificado em i	Ordinal 0 – Errou 1 – Acertou	—	—
P_Timei	Tempo gasto pelo sujeito na resolução do item especificado em i	Intervalar (centésimos de segundo)	1	5999

4.2. Amostra

O teste foi aplicado a 731 candidatos ao curso de Pilotagem Aeronáutica da Academia da Força Aérea Portuguesa nos anos de 1998 e 1999. As idades dos candidatos oscilam entre os 17 e os 24 anos. Em termos de género, 90,7% da amostra é do sexo masculino, o que significa que, aproximadamente, menos do que um candidato em cada 10 é do sexo feminino. As condições de ingresso no curso de Pilotagem Aeronáutica, em termos de habilitações literárias, são idênticas às exigidas para o ingresso no ensino superior naqueles anos, ou seja,

necessitavam de ter, pelo menos, o 12º Ano de escolaridade.

4.3. Apresentação e discussão dos resultados

Os resultados vão ser apresentados em dois grupos. Um primeiro grupo de estatísticas descritivas directamente relacionadas com as variáveis principais do teste e suas qualidades métricas, nomeadamente, os coeficientes de consistência interna. Um segundo grupo de resultados prende-se com a análise dos testes estatísticos efectuados aos diferentes tipos de respostas dadas nos itens. Para o estudo das qualidades métricas analisou-se a distribuição dos resultados relativamente às duas variáveis globais, análise da consistência interna às variáveis tempo e correcção dos itens.

4.3.1. Qualidades Métricas

Para análise das qualidades métricas foram calculadas as consistências internas com recurso ao coeficiente alpha de Cronbach para as duas variáveis globais e foi efectuado um teste de aderência à normalidade por intermédio da fórmula de Kolmogorov-Smirnov (K-S).

O alpha (α) encontrado para a variável **P_Corr** (resultado obtido pelo sujeito na totalidade dos itens) foi de 0,45, o que é um coeficiente significativamente baixo se o compararmos com o de 0,92 da variável **P_Tempo** (tempo gasto pelo sujeito na resolução da totalidade dos itens). A estatística Kolmogorov-Smirnov, para testar a aderência à distribuição normal, veio de algum modo corroborar estes resultados tendo sido obtidos os resultados de $K-S = 0,12$ para **P_Corr** e $K-S = 0,03$ para o **P_Tempo**, fazendo com que para a variável **P_Corr** se rejeitasse a hipótese de normalidade da distribuição ($P \geq 0,05$).

4.3.2. Análise das médias e desvios-padrão dos tempos de latência aos itens

Para se analisar os tempos de latência aos itens foi construída a Figura 4.2, onde, graficamente, estão representadas as médias dos tempos de resposta em centésimos de segundo, para cada item e em função do resultado obtido: acerto ou erro do item.

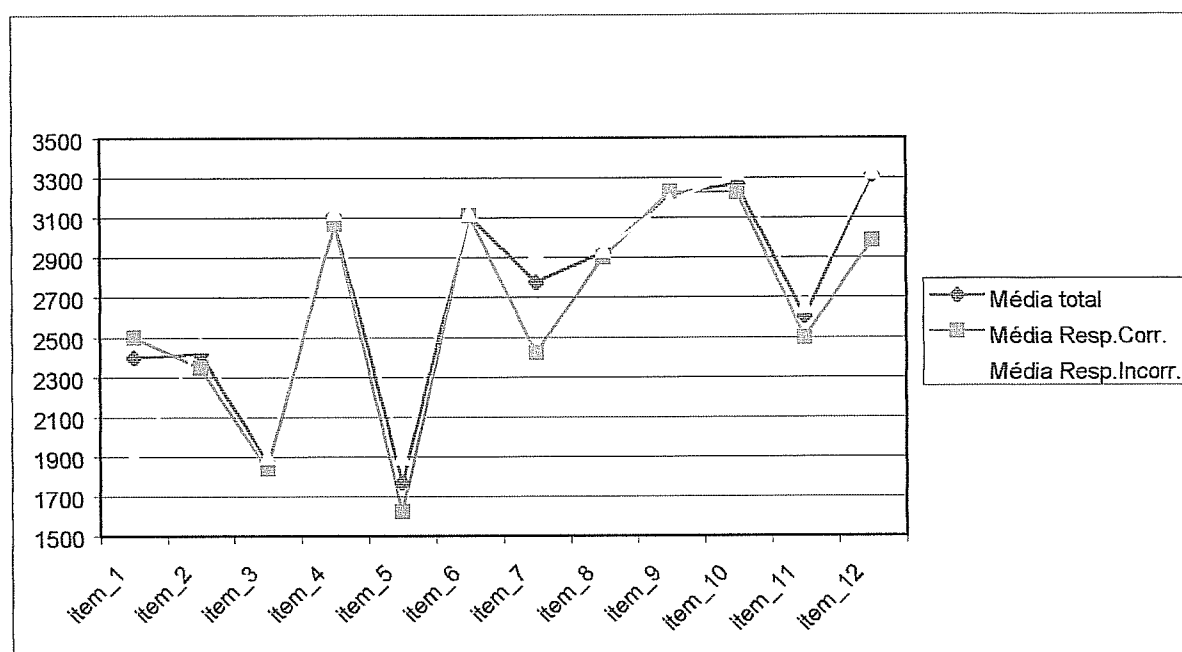


Figura 4.2. Médias dos tempos de resposta aos itens

Ao analisarmos, de forma comparativa, as médias dos tempos de resposta aos itens respondidos correctamente com os respondidos incorrectamente, e numa perspectiva progressiva de evolução da apresentação dos mesmos aos sujeitos, verificamos que a maioria dos itens respondidos correctamente apresentam tempos médios de resposta mais baixos que os respondidos incorrectamente. Esta tendência é idêntica para os desvios-padrão (ver Figura 4.3), ou seja, o desvio-padrão dos tempos de resposta nos itens respondidos correctamente é inferior aos desvios-padrão nos itens errados.

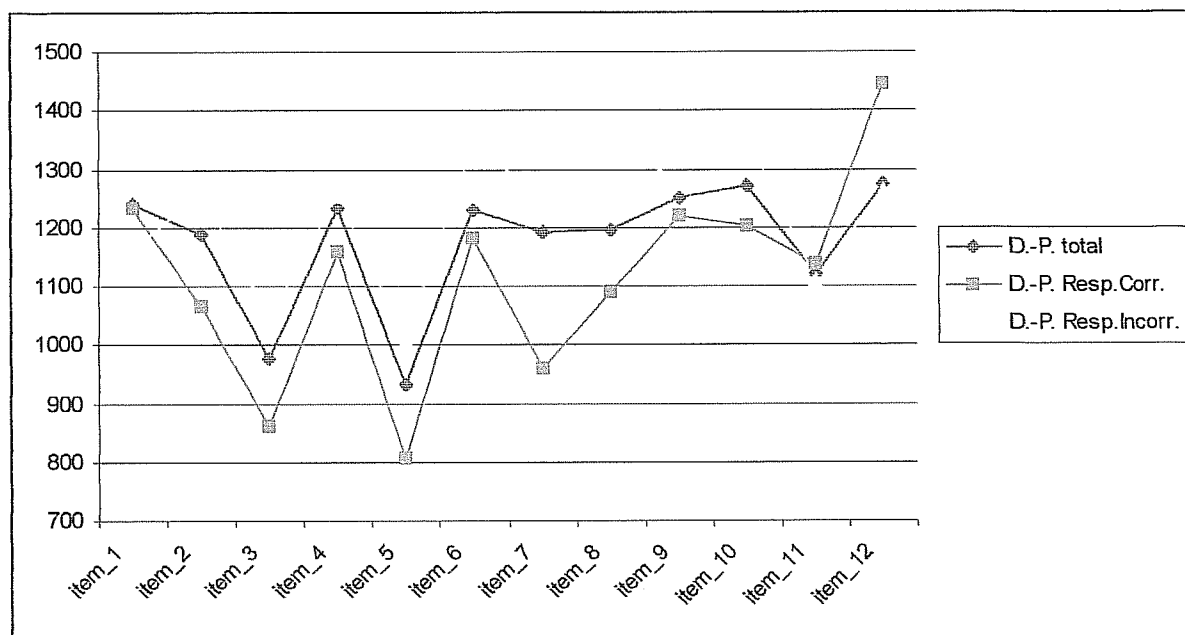


Figura 4.3. Desvios-padrão nos tempos de resposta aos itens

Para se poder efectuar uma análise mais rigorosa foi construído o Quadro 4.2 em que são apresentados os tempos médios e os desvios-padrão para cada um dos itens e em função do resultado obtido. No mesmo Quadro, também é apresentado o número de sujeitos que acertaram e erraram o item, assim como a análise de variância para testar a diferença de médias entre os tempos gastos nos itens respondidos correctamente e nos respondidos incorrectamente.

Uma análise mais pormenorizada das Figuras 4.2 e 4.3 e do Quadro 4.2, mostra que o primeiro e o nono item apresentam um tempo de resposta superior quando é respondido correctamente, contrariamente ao que acontece nos restantes itens. O comportamento discrepante no primeiro item pode encontrar justificação no facto de os primeiros itens de um teste de aptidão tenderem a absorver uma grande parte da variância devida à insegurança dos sujeitos. Por outro lado, o item 1 e o 9 são dos que possuem maior número de respostas correctas conjuntamente com o item 6 (Quadro 4.2).

Quadro 4.2. Médias e desvios-padrão dos tempos de resposta aos itens respondidos correcta e incorrectamente e correspondente análise da significância da diferença de médias

	Itens respondidos correctamente			Itens respondidos incorrectamente			F	Sig
	N	Tempo Médio	Desvio-Padrão	N	Tempo Médio	Desvio-Padrão		
item_01	603	2498,0	1233,1	124	1901,9	1161,3	24,50	0,000
item_02	292	2344,7	1067,5	435	2462,6	1260,6	1,72	0,190
item_03	292	1836,5	860,9	436	1898,4	1048,1	0,70	0,402
item_04	163	3065,1	1158,6	563	3123,5	1253,3	0,28	,0595
item_05	333	1620,4	805,8	393	1895,2	1012,2	15,97	0,000
item_06	509	3109,0	1181,8	218	3116,9	1339,2	0,01	0,937
item_07	225	2420,0	961,5	502	2934,2	1253,2	29,96	0,000
item_08	267	2900,3	1091,1	460	2934,5	1255,1	0,14	0,711
item_09	491	3228,0	1218,9	234	3177,5	1317,1	0,26	0,611
item_10	407	3225,1	1201,0	316	3323,1	1356,9	1,06	0,304
item_11	220	2495,1	1137,2	505	2649,9	1109,0	2,94	0,087
item_12	36	2982,2	1445,5	690	3327,2	1264,5	2,51	0,114

Relativamente aos desvios-padrão também encontramos um padrão constante na comparação entre respostas correctas e incorrectas, ou seja, nos itens respondidos correctamente verifica-se um desvio-padrão menor do que nos respondidos incorrectamente, à excepção do primeiro e dos dois últimos itens. Uma hipótese explicativa para o comportamento dos sujeitos no primeiro item já foi avançada anteriormente. No último item a discrepância pode ser devida ao facto de estarmos na presença de uma questão de elevada dificuldade (apenas se verificaram 5% de respostas correctas). Considerando que o tempo médio gasto neste item para dar uma resposta correcta, é inferior ao tempo médio para quando a resposta é dada incorrectamente, e, que o desvio-padrão do tempo para quando a resposta é respondida

correctamente é superior, associado ao facto de ser um item de elevada dificuldade, podemos hipotetizar que algumas respostas possam ter sido dadas correctamente, devido ao acaso, e não de forma pensada. O facto de ser um item em que apenas 36 dos 726 sujeitos responderam correctamente e de ser o último item reforça a ideia de que as respostas dadas possam ter sido menos reflectidas e, conseqüentemente, ter provocado uma diminuição do tempo médio e uma maior oscilação do desvio-padrão. Da observação da Figura 4.3 (e também do Quadro 4.2) constata-se que é no último item que se verifica o maior desvio-padrão, dois segundos a mais que o item 1 que contém o segundo desvio-padrão mais elevado, para quando é respondido correctamente.

O número de respostas correctas entre o item 1 e o item 12 (primeiro e último itens do teste), está praticamente invertido. Enquanto o item 1 é o que tem maior número de respostas correctas (item mais fácil), o item 12 é o que possui menor número de respostas correctas (item mais difícil). Enquanto no item 1 o tempo médio de resposta é mais elevado quando o item é respondido correctamente do que quando é respondido incorrectamente, o oposto acontece no item 12 em que os sujeitos são mais rápidos em média quando acertam este item. Este dado pode ser uma consequência do grau de dificuldade do item. Enquanto o primeiro item tem um grau reduzido de dificuldade (82,9% das respostas foram dadas correctamente), o último item tem um elevado grau de dificuldade (apenas 5% dos sujeitos acertaram). A generalização destes resultados, só com base no item 1 e 12, levar-nos-ia a colocar a hipótese de que o grau de dificuldade varia numa razão inversa do tempo médio de resposta, tomando em consideração apenas quando são respondidos correctamente. Na Figura 4.4 tentou-se representar graficamente esta hipotética relação.

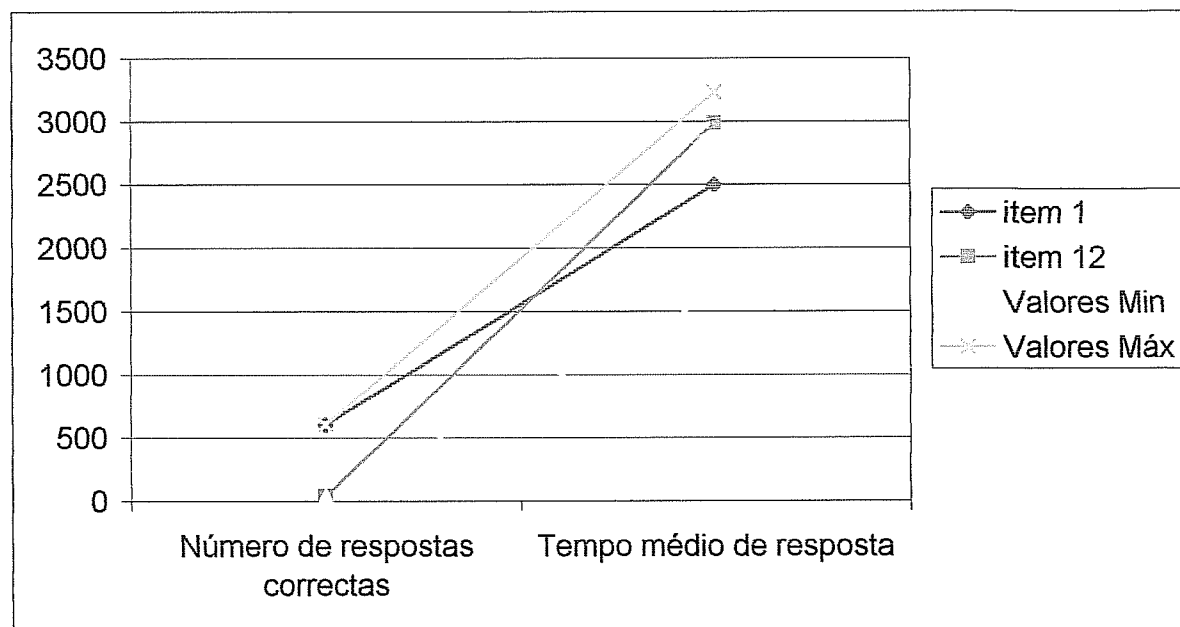


Figura 4.4. Relação entre o tempo médio de resposta e o número de respostas correctas nos itens 1 e 12

Para analisarmos se os tempos médios de resposta nos itens respondidos de forma correcta diferem dos respondidos incorrectamente, efectuou-se uma análise de variância entre as médias nos diferentes itens. Apenas se verificaram diferenças significativas de tempos entre os itens respondidos correctamente e os respondidos incorrectamente, os itens 1, 5 e 7 (ver Quadro 4.2).

4.3.3. Factores não cognitivos

Como já foi referido anteriormente, este estudo teve, também, por objectivo explorar a hipótese de interferência de factores não cognitivos na execução de provas com itens de escolha múltipla. Na tentativa de abordar o problema, levantou-se a hipótese de os tempos de resposta serem diferentes face ao tipo de resposta dada. Recorde-se que nesta prova existem três alternativas de resposta possível: (0) para indicar se a figura-alvo não se encontra em

nenhuma das figuras complexas; (1) se a figura-alvo se encontra na figura complexa da esquerda; (2) se a figura-alvo se encontra na figura complexa da direita; e (3) se a figura-alvo se encontra em ambas as figuras complexas. A hipótese nula é de que os tempos médios de resposta são idênticos indiferentemente das alternativas escolhidas como resposta correcta, ou seja, os sujeitos deverão demorar tempos não significativamente diferentes independentemente da resposta correcta ser 0, 1, 2, ou 3.

A nossa hipótese alternativa à hipótese nula é a de que, quanto maior for a incerteza do sujeito na resposta a dar, maior é o tempo gasto no processamento de informação com vista à resolução do problema. Será confirmada se os sujeitos demorarem mais tempo a tomar uma decisão nos itens em que a figura-alvo não se encontra em nenhuma das figuras (alternativa de resposta 0). Por oposição, nos itens em que a figura-alvo se encontra nas duas figuras (alternativa de resposta 3) os tempos serão menores. Os tempos de latência na resolução de um item irão depender do grau de certeza da resposta a dar.

Para testar esta hipótese construíram-se o Quadro 4.3 e a Figura 4.5. No Quadro 4.3 colocaram-se os itens por resposta pretendida e não pela ordem normal de apresentação aos sujeitos. Assim, são apresentados, em primeiro lugar, os itens cuja resposta pretendida é «0», ou seja, a figura-alvo não se encontra em nenhuma das figuras complexas e que compreende os itens 1, 6 e 9. Depois os itens em que a resposta pretendida é de «1» (itens 8, 10 e 11), seguidos dos itens cuja resposta correcta corresponde à alternativa «2» (itens 2 e 4) e, finalmente, os de resposta «3» (itens 3, 5, 7 e 12). A elaboração do gráfico constante da Figura 4.5 obedeceu ao mesmo princípio de apresentação dos itens.

Quadro 4.3. Médias e desvios-padrão dos tempos de resposta por item e as respostas pretendidas

	Resposta pretendida	N	Número de respostas correctas	Número de respostas incorrectas	Tempo Mínimo	Tempo Máximo	Média dos tempos	Desvio padrão
P_time1	0	727	603	124	162	5840	2396,29	1240,83
P_time6	0	727	509	218	415	5849	3111,40	1230,18
P_time9	0	725	491	234	280	5939	3211,70	1250,72
P_time8	1	727	267	460	222	5872	2921,91	1196,86
P_time10	1	723	407	316	296	5962	3267,96	1271,55
P_time11	1	725	220	505	356	5764	2602,94	1119,11
P_time2	2	727	292	435	323	5919	2415,27	1187,47
P_time4	2	726	163	563	330	5820	3110,38	1232,11
P_time3	3	728	292	436	320	5608	1873,59	977,19
P_time5	3	726	333	393	264	5368	1769,15	932,74
P_time7	3	727	225	502	254	5726	2775,06	1193,96
P_time12	3	726	36	690	266	5901	3310,09	1275,13

Na análise do quadro 4.3 e do gráfico da Figura 4.5, verificamos que os itens em que a resposta correcta é de «0», ou seja, quando a figura-alvo não se encontra em nenhuma das figuras complexas (itens 1, 6 e 9) os tempos são significativamente maiores do que quando a resposta correcta é «3» (itens 3, 5, 7 e 12) isto é, quando a figura-alvo se encontra em ambas as figuras complexas. A nossa hipótese ainda ganha maior robustez se eliminarmos os itens 1 e 12 que se encontram nos dois grupos virtualmente criados, pelas razões que já mereceram a nossa consideração.

De acordo com a nossa hipótese, uma possível explicação para estes resultados relaciona-se com a interferência de factores não cognitivos na realização da tarefa. O sujeito quando confirma a existência da figura-alvo numa das figuras complexas, passa automaticamente para a figura complexa seguinte. Se também nesta visualiza a figura-alvo dá a resposta correspondente (tecla 3) e passa ao item seguinte.

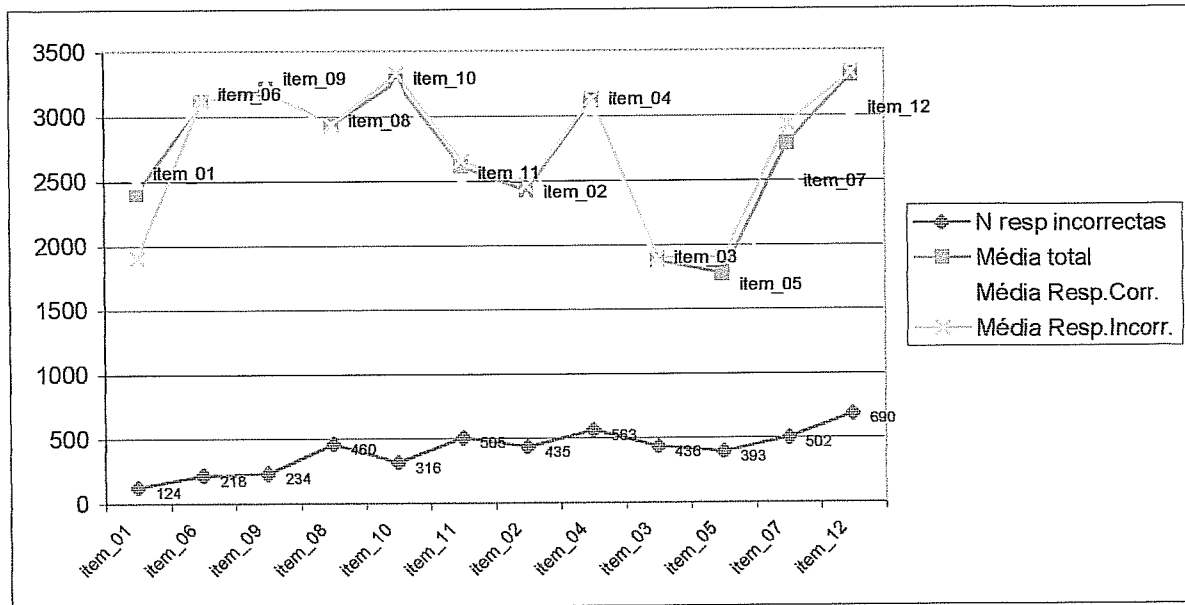


Figura 4.5. Médias dos tempos de latência e o número de respostas incorrectas para cada um dos itens

Por oposição, no caso de a figura-alvo não se encontrar em nenhuma das figuras complexas, o sujeito ao não localizar a figura-alvo em nenhuma das figuras complexas poderá tender a efectuar uma segunda observação de confirmação. Esta estratégia de dupla verificação pode ser uma consequência da interferência de factores não cognitivos, nomeadamente no plano da autoconfiança. O tempo de resposta do sujeito pode não significar, assim, em exclusivo, um índice da sua capacidade para processar rapidamente informação visual, mas também um índice de confiança na tomada de decisão.

Se analisarmos o Quadro 4.4, onde se comparam as médias dos tempos de latência entre os dois grupos de itens, quando as respostas são dadas correctamente, verificamos que, à excepção das diferenças de médias em que o item 12 intervém, todas as outras são significativas a $p < .001$, pelo que confirmamos parcialmente a nossa hipótese alternativa, de que quando a figura-alvo se encontra nas duas figuras complexas, os sujeitos são mais rápidos a dar a resposta. A confirmação desta hipótese pode significar que nos itens que proporcionam

menor incerteza na resposta a dar, os sujeitos têm um tempo de latência menor, o que pode não corresponder forçosamente a uma maior rapidez no processamento da informação.

Quadro 4.4. Teste de diferença de médias entre os itens em que a figura-alvo não aparece nas figuras complexas e os itens em que a figura-alvo aparece nas duas figuras complexas.

		Itens em que a figura-alvo se encontra nas duas figuras complexas			
		Item_3 M=1836,5	Item_5 M=1620,4	Item_7 M=2420,0	Item_12 M=2982,2
Itens em que a figura-alvo não se encontra em nenhuma das duas figuras complexas	Item_1 M=2498,0	12,136 gl=253 ,000	16,106 gl=283 ,000	3,512 gl=193 ,001	-3,476 gl=25 ,002
	Item_6 M=3109,0	17,924 gl=190 ,000	21,353 gl=224 ,000	9,868 gl=154 ,000	,322 gl=19 ,751
	Item_9 M=3228,0	18,604 gl=220 ,000	23,306 gl=233 ,000	10,699 gl=155 ,000	,567 gl=20 ,577

Uma hipótese *post-hoc* que é possível colocar é a de encontrarmos correlações diferentes entre os diferentes itens da prova e determinados traços de personalidade. Por exemplo, os itens em que a resposta pretendida é zero (a figura-alvo não se encontra em nenhuma das figuras complexas) serão mais diferenciadores dos sujeitos em termos da sua capacidade de tomada de decisão. Espera-se, desta forma, que os sujeitos que acertem os itens de resposta zero e que sejam mais rápidos a dar a resposta, possuam valores mais elevados em factores de personalidade ao nível da segurança na tomada de decisão, do que os sujeitos que também acertem estes itens mas sejam mais lentos na resposta.

O presente estudo não nos permitiu testar esta hipótese *post-hoc*, já que o desenho do teste não está concebido para o efeito, possuindo poucos itens por cada uma das condições. Também não foi possível controlar com rigor outras variáveis, como sejam a complexidade da tarefa e a identificação e medição de factores não cognitivos.

4.3.4. Identificação dos factores avaliados

Sendo este teste avaliador de um factor cognitivo será esperado encontrar correlações entre a eficiência nos diferentes itens e o resultado no teste psicométrico de inteligência. Por outro lado, este teste enquadra-se dentro dos testes que avaliam o factor perceptivo à semelhança de outros, nomeadamente o teste psicométrico de rapidez perceptiva.

Para confirmar esta hipótese, e considerando que esta prova foi aplicada em conjunto com outras dentro de um processo real de selecção, efectuaram-se as correlações momento-produto de Pearson com dois testes psicométricos: um de inteligência geral baseado no teste dos Dominós (**Int. Geral**) e um de rapidez perceptiva baseado nas formas idênticas (**Rap. Perc.**). No Quadro 4.5 encontram-se as correlações entre os dois scores globais da prova “Padrões” em termos de respostas certas e tempos médios de resposta (**P_Corr** e **P_Tempo**, respectivamente) e os dois testes psicométricos (**Int. Geral** e **Rap. Perc.**).

Quadro 4.5. Correlações momento-produto, níveis de significância e número de sujeitos entre dois testes psicométricos e as notas globais do “Padrões”

	Int. Geral	Rap. Perc.
P_Corr	,299	,161
	,000	,000
	659	602
P_Tempo	,001	-,144
	,989	,000
	659	602

A análise do Quadro 4.5 leva-nos a duas constatações. Uma é a confirmação da hipótese de que o teste “Padrões” avalia um factor dentro da esfera cognitiva e a outra é a de que está fortemente saturado de uma componente perceptiva. Se considerarmos o número de respostas correctas dadas por cada sujeito, será de esperar que, quanto mais inteligente for o sujeito, maior será o número de respostas correctas dadas no teste “Padrões”.

De acordo com o tipo de dimensão avaliada pelo teste, seria de esperar que a correlação maior fosse com o teste de rapidez perceptiva e não com a inteligência geral, já que pertencem à mesma família de factores segundo a orientação de Carroll (1993). Pode-se, desta forma, hipotetizar que o teste de rapidez perceptiva possui uma componente de velocidade mental essencial ao processamento de informação cognitiva.

A segunda interpretação está relacionada com a variável tempo médio de resposta aos itens. Como esperado, correlaciona-se de forma significativa com o teste de rapidez perceptiva que é um dos testes considerados dentro do factor velocidade cognitiva (Carroll, 1993). Apesar de as provas de rapidez perceptiva se enquadrarem dentro do factor visual perceptivo, contribuem também para a identificação do factor velocidade cognitiva. O facto de o coeficiente de correlação ser negativo, significa que quanto maior for o número de figuras correctamente comparadas num espaço de tempo reduzido, menos tempo o sujeito demora a dar as suas respostas aos itens do teste “Padrões”.

A análise do quadro de correlações leva-nos à identificação de dois factores: um mais relacionado com a exactidão e outro mais com a velocidade. Para confirmar esta afirmação foi efectuada a análise factorial pelo método das componentes principais com rotação *varimax* e cujos resultados se encontram no Quadro 4.6.

Quadro 4.6. Matriz factorial após rotação ortogonal

	Componente 1	Componente 2
P_Corr	,752	
P_Tempo		,864
Int. Geral	,708	
Rap. Perc.	,527	-,574

Método de extracção: Análise das Componentes Principais

Método de Rotação: Varimax com normalização de Kaiser

Pesos inferior a 0,30 foram retirados da tabela

Os resultados da análise factorial permitiu-nos confirmar a separação entre o factor exactidão e o factor velocidade na resposta aos itens do teste “Padrões”. Apesar de termos encontrado uma correlação significativa entre o teste psicométrico de papel-e-lápis de inteligência geral e o de rapidez perceptiva ($r=.177$, $p<0,000$), eles vão orientar-se para factores diferentes. O teste de rapidez perceptiva ao aparecer com um peso factorial relativamente forte nas duas componentes extraídas, pode constituir um bom indicador da velocidade de processamento de informação do ponto de vista perceptivo, ou seja, a velocidade mental.

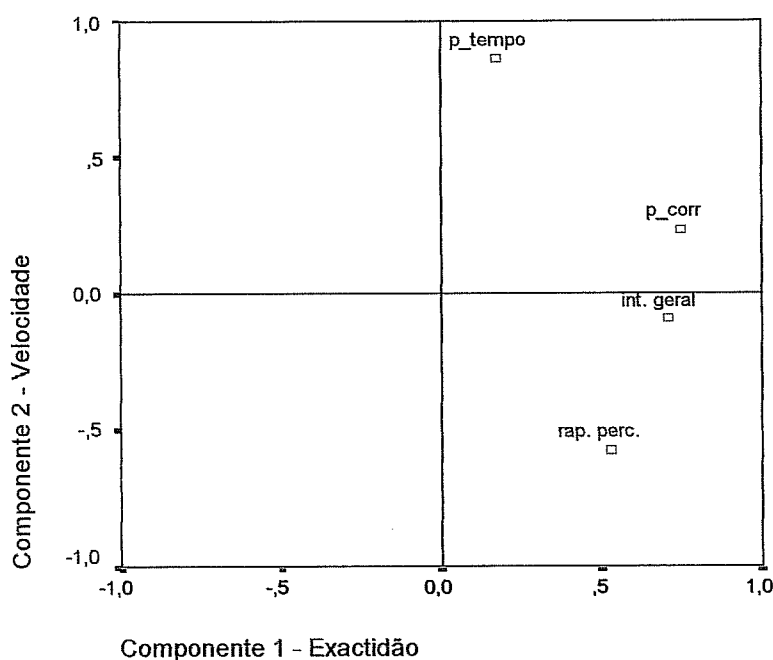


Figura 4.6. Representação espacial das variáveis tempo e erros do teste de padrões e dos testes de inteligência geral e rapidez perceptiva.

A Figura 4.6 é elucidativa da distribuição espacial das quatro variáveis consideradas no espaço dimensional confinado pelas duas componentes extraídas. Apesar de o teste de rapidez perceptiva se encontrar no mesmo factor que a variável relativa aos tempos médios de resposta, ele aparece na representação gráfica no lado oposto à variável **P_Tempo**. Este facto é justificado, como já referido, pelo tipo de scores respectivos. Enquanto a variável **P_Tempo** diz respeito ao tempo médio gasto pelo sujeito na resposta aos itens em centésimos de segundo, o score da variável rapidez perceptiva corresponde ao número de respostas correctas dadas num curto espaço de tempo.

4.4. Sumário dos resultados

Das análises estatísticas efectuadas a este teste podemos sintetizar as seguintes considerações finais:

Os tempos médios de resposta aos itens são diferentes quando os sujeitos acertam ou erram os itens. Os tempos médios de resposta são inferiores quando as respostas são cotadas correctamente, comparativamente com os tempos correspondentes às respostas cotadas incorrectamente.

A proficiência dos sujeitos nesta prova avaliadora de *flexibility of closure*, medida através do número de respostas correctas, diferencia-se da rapidez como processam a informação avaliada pelo tempo médio das respostas dadas. Uma medida mais válida da dimensão *flexibility of closure* deveria englobar tanto a correcção de respostas como a velocidade.

A rapidez com que o sujeito processa informação para dar as respostas não está relacionada com a capacidade mental geral ($r = 0,001$) mas correlaciona-se significativamente ($r = -0,144$, $p < 0,00$) com o teste psicométrico de rapidez perceptiva que poderá significar que é uma prova

válida para avaliar a velocidade cognitiva no processamento de informação perceptiva.

O grau de incerteza da resposta é um factor que se mostrou significativo para a análise dos resultados nos itens. Quanto maior é o grau de incerteza, mais tempo o sujeito tende a demorar a dar a resposta. Estes resultados parecem fortalecer a posição de Witkin e Goodenough (1981), em que defendem que este tipo de provas é permeável a aspectos de personalidade, transformando-se mais num teste avaliador do estilo cognitivo do sujeito, do que propriamente de uma aptidão específica.

Os primeiros e últimos itens comportam-se de modo contrário à maioria dos restantes itens. Podemos especular que os primeiros tendem a absorver forte variância devido à adaptação do sujeito à prova.

Apesar de não se ter objectivado a complexidade da tarefa para cada um dos itens, podemos considerar que o facto de o sujeito não saber que pelo menos uma das figuras complexas contém a figura-alvo é um elemento que contribui para a complexidade do item. Este aspecto é agravado pelo conhecimento do tempo em falta, que pode induzir um factor de pressão adicional.

5. Estudo 2 – Desempenho nas Matrizes Progressivas de Raven

Com este segundo estudo procurou-se conhecer melhor a evolução do comportamento de resposta dos sujeitos ao longo de um teste informatizado de inteligência geral numa situação real de avaliação. Escolhemos a prova das Matrizes Progressivas de Raven na sua forma *Standard* mas de aplicação em suporte informático.

O principal objectivo consistiu na avaliação da eficiência dos sujeitos na prova em termos de correcção das respostas dadas, bem como o modo como geriram o tempo para responder a cada um dos itens. Foi também efectuada uma análise exploratória da relação entre os tempos gastos nas respostas dadas e a correcção das mesmas, em dois grupos com diferentes níveis de rendimento intelectual global.

Com este estudo procurámos testar duas hipóteses. A primeira diz respeito à análise da relação da variável por nós definida por complexidade da tarefa com a dificuldade da mesma. Com a segunda hipótese pretendeu-se estudar a relação entre a capacidade de processamento de informação dos sujeitos em termos de velocidade mental e o seu nível de capacidade mental geral.

5.1. Amostra

Foram utilizados os resultados de 300 candidatos que se submeteram a provas psicológicas inseridas num processo real de selecção para condutores de veículos pesados de passageiros. A

média de idades é de 26,97 anos ($d.-p. = 6,63$), sendo 58% do sexo masculino e 42% do sexo feminino. A escolaridade foi codificada em cinco grupos, de acordo com as normas europeias. Encontrámos 36,4% do grupo 2, 40,4% do grupo 3, 13,9% do grupo 4 e 9,3% do grupo 5. A amostra foi de conveniência por terem sido considerados para estudo os resultados dos primeiros trezentos participantes que completaram o teste.

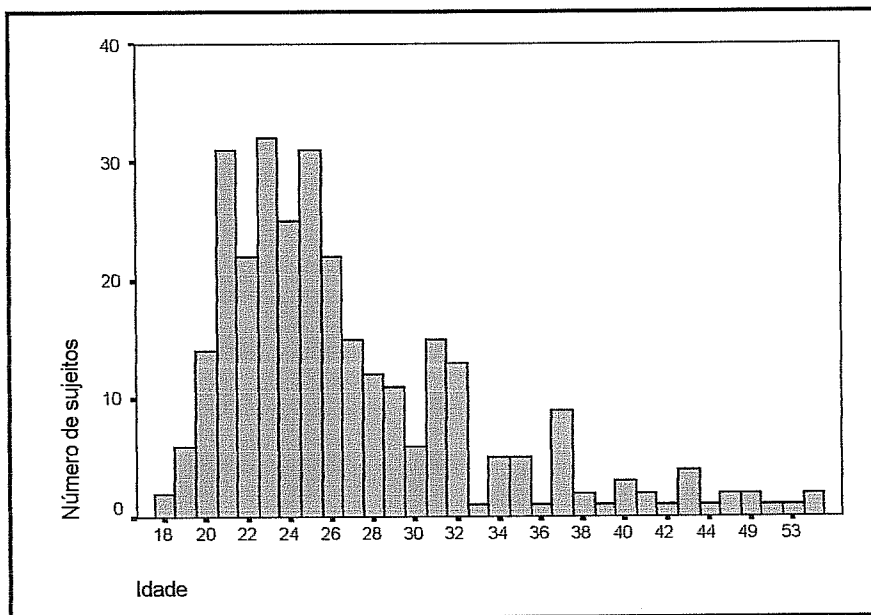


Figura 5.1. Distribuição dos sujeitos por idades

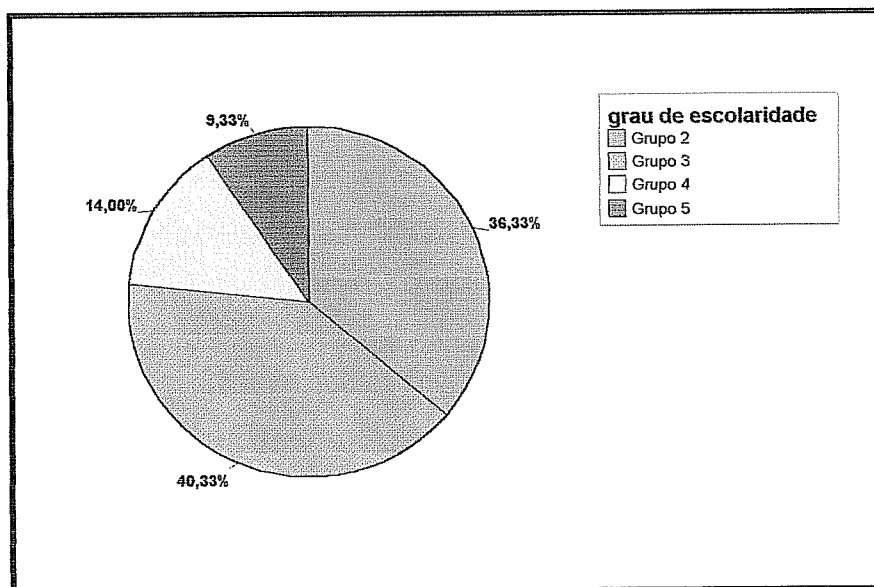


Figura 5.2. Distribuição dos sujeitos por grupos de habilitações literárias

5.2. Instrumento

Utilizou-se o teste das Matrizes Progressivas de Raven na sua versão *Standard* (MPR-S) para medir a inteligência, sendo este teste aplicado na versão computadorizada numa plataforma gerida pelo sistema de testes de Viena. As MPR-S foram aplicadas com um tempo limite de 30 minutos e estavam inseridas na bateria de selecção para a função a que os candidatos estavam a concorrer: condutores de veículos pesados de passageiros.

As MPR-S são compostas por 5 séries de doze itens. O grau de complexidade é progressivo em cada série, fazendo com que em termos globais a dificuldade seja crescente mas não constante. O sujeito é, assim, será confrontado, de tempos a tempos, com itens menos complexos que os realizados anteriormente e que correspondem aos primeiros de cada série. Enquanto que na versão de papel-e-lápis o sujeito pode facilmente intuir que existe mudança de complexidade, na versão informatizada o sujeito não tem essa percepção. Na versão clássica, a folha de resposta possui cinco colunas de quadrados que correspondem a cada uma das séries. Na versão informatizada, os itens são apresentados sequencialmente do 1 ao 60 e o sistema não lhe dá qualquer indicação de que mudou de série e que irá responder a itens menos complexos.

5.3. Análise e discussão de resultados

Iniciamos a apresentação dos resultados com a estatística descritiva relativamente às duas variáveis principais do estudo: número de respostas correctas (RC) e o tempo gasto em segundos na execução dos itens cuja resposta foi cotada correctamente num primeiro ensaio (TEMP). Considerando o interesse deste estudo em conhecer o modo como o sujeito geriu o tempo para dar a resposta, não foram tomados em consideração os tempos e as respostas que foram dadas em modo de alteração, ou seja, que foram alteradas por parte do sujeito. Apesar de o sistema permitir que o sujeito volte atrás para alterar respostas, não guarda as primeiras

respostas dadas, nem se foram dadas correctamente, assim como também não guarda o tempo acumulado para dar as respostas respectivamente a cada item. Apenas é gravado o estado final do item, isto é, se no final foi respondido correctamente, ou não, e o tempo que o sujeito demorou na última vez que acedeu à questão (cf. Quadro 5.1). Por exemplo, se um indivíduo corrigiu o item x, o seu resultado não será considerado, independentemente de ter acertado ou não o item. Os scores de respostas correctas e do tempo total não incluirá o resultado nesse item.

Quadro 5.1. Estatística descritiva para as variáveis respostas correctas (RC) e tempo gasto em seg (TEMP) para a totalidade do teste

	RC	TEMP
N	300	300
Média	43,93	884,80
Mediana	45	894,50
Moda	48	906,00
Desvio-Padrão	6,44	270,55
Variância	41,47	73199,53
<i>Skewness</i>	-0,904	-0,019
Erro-Padrão da <i>Skewness</i>	0,141	0,141
Curtose	0,950	-0,561
Erro-Padrão da Curtose	0,281	0,281
Valor Mínimo encontrado	18	248,00
Valor Máximo encontrado	55	1546,00

5.3.1. Distribuição dos resultados

Da análise do Quadro 5.1 verifica-se que a distribuição de resultados para a variável RC aponta para uma assimetria negativa (*skewness*/E.-P. = -6,738), conforme se pode verificar no gráfico de barras da Figura 5.3 (encontra-se desviada para a esquerda com uma média de 43,93 e uma mediana de 45,00).

A distribuição dos resultados aponta para menor sensibilidade do teste, já que a média encontrada (43,93) fica apenas a 16 pontos do resultado máximo da prova e que é de 60. Esta

tendência para a maioria dos resultados aparecerem à direita da média, provoca um enviezamento dos resultados à direita da distribuição.

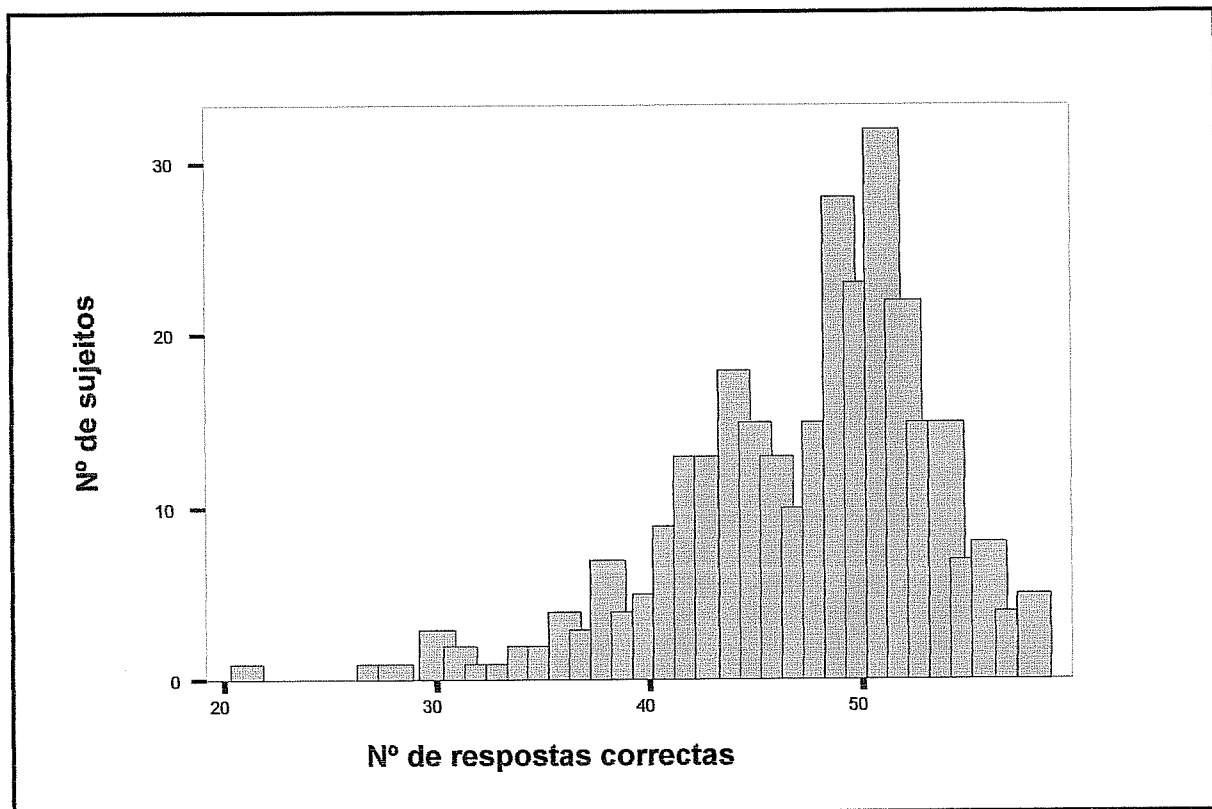


Figura 5.3. Gráfico de barras com as frequências dos resultados finais em termos de respostas correctas

Relativamente à outra variável considerada, o tempo gasto na execução dos itens cotados correctamente (TEMP), a distribuição dos resultados está mais próxima da distribuição normal como pode ser constatado no histograma apresentado na Figura 5.4. No teste Kolmogorov-Smirnov, efectuado para testar a normalidade da distribuição, obteve-se um resultado de 0,038 ($p = 0,200$), pelo que podemos concluir que se encontra dentro da normalidade esperada.

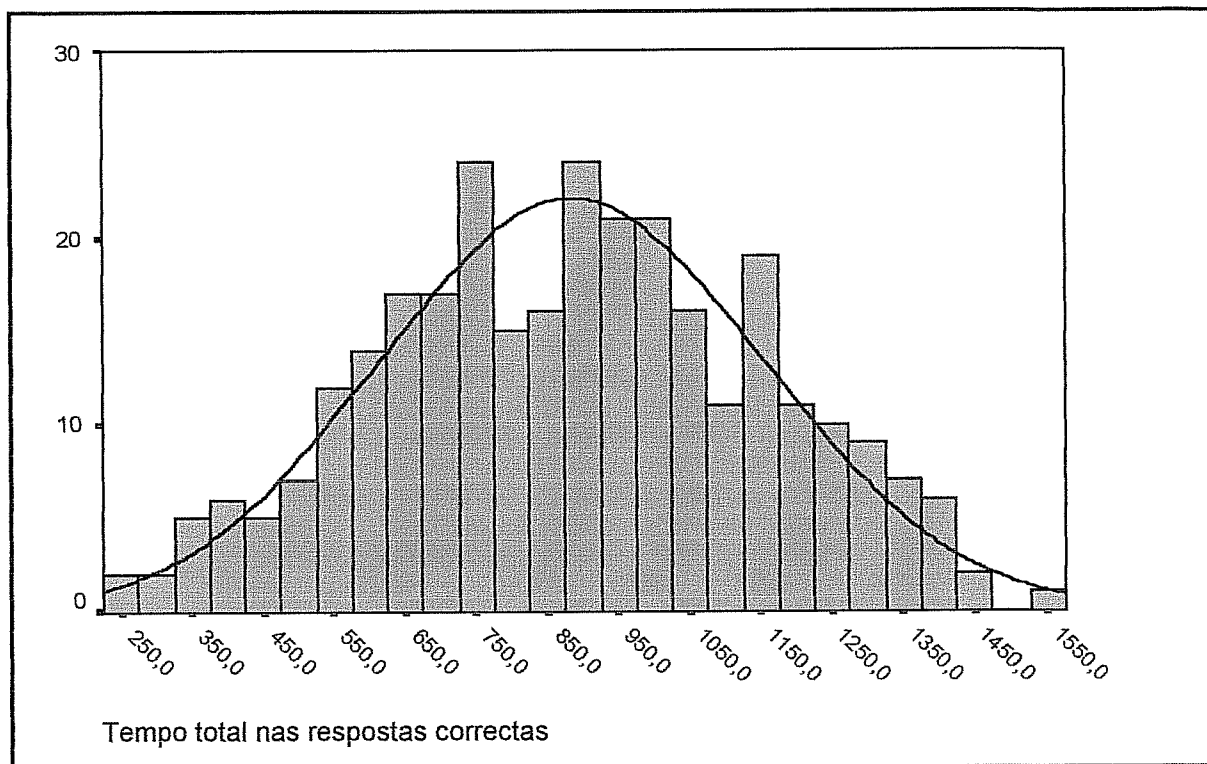


Figura 5.4. Histograma com os tempos gastos nos itens respondidos correctamente

5.3.2. Consistência interna

Foi calculada a consistência interna, tendo sido obtido um α de 0,867 para a variável RC que nos dá uma garantia elevada nos resultados encontrados. No Quadro 5.2 podemos constatar que são poucos os itens que prejudicam a escala, ou seja, que permitem um incremento do valor α da escala quando são retirados/reformulados. Esses itens (47 e 48) permitem aumentar o α da escala, embora não possa ser considerado significativo, já que no máximo, passaria de 0,867 para 0,870.

Quadro 5.2. Hipotéticos alphas para as Matrizes Progressivas de Raven na forma *Standard* se os itens respectivos forem eliminados

IT01	0,867	IT13	0,867	IT25	0,867	IT37	0,867	IT49	0,864
IT02	0,867	IT14	0,867	IT26	0,865	IT38	0,863	IT50	0,863
IT03	0,867	IT15	0,867	IT27	0,865	IT39	0,865	IT51	0,858
IT04	0,867	IT16	0,867	IT28	0,865	IT40	0,865	IT52	0,859
IT05	0,867	IT17	0,864	IT29	0,868	IT41	0,867	IT53	0,858
IT06	0,867	IT18	0,867	IT30	0,863	IT42	0,864	IT54	0,860
IT07	0,862	IT19	0,865	IT31	0,862	IT43	0,863	IT55	0,864
IT08	0,869	IT20	0,864	IT32	0,861	IT44	0,863	IT56	0,859
IT09	0,865	IT21	0,863	IT33	0,865	IT45	0,864	IT57	0,866
IT10	0,867	IT22	0,864	IT34	0,867	IT46	0,863	IT58	0,863
IT11	0,862	IT23	0,860	IT35	0,861	IT47	0,870	IT59	0,865
IT12	0,866	IT24	0,863	IT36	0,862	IT48	0,870	IT60	0,868

Alpha das matrizes para a totalidade dos 60 itens = **0,867**

Foi efectuada seguidamente uma análise ao grau de dificuldade dos 60 itens em termos do número de respostas correctas, tendo sido criado um gráfico de dispersão dos itens e que constitui a Figura 5.5.

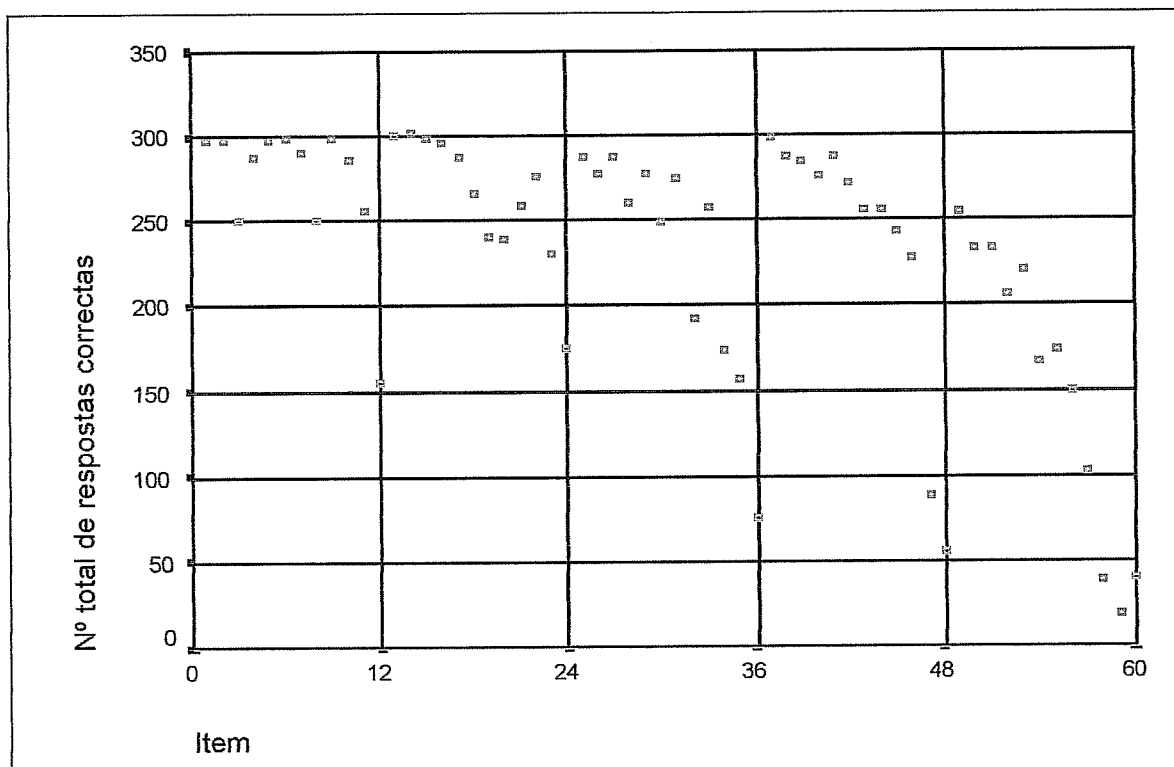


Figura 5.5. Gráfico de dispersão dos itens pelo número total de respostas correctas

Como se pode verificar no gráfico de dispersão da Figura 5.5 existe uma clara diferenciação entre cada uma das séries, embora se constate um padrão idêntico dentro de cada uma delas. Os primeiros itens de cada um das séries apresentam um número de respostas correctas muito superior aos últimos da série anterior, e, à medida que se avança na série, o número de respostas correctas vai diminuindo. Considerando o facto de o teste ser apresentado em formato informático e com tempo limite, pode-se questionar se a melhor forma de apresentar os itens é por séries, mantendo a estrutura original do teste, ou por grau de dificuldade crescente em termos globais.

De acordo com a primeira hipótese colocada, relativamente à relação entre as variáveis dificuldade e complexidade da tarefa, será de esperar uma relação forte entre ambas. Os conceitos de dificuldade e complexidade foram operacionalizados do seguinte modo: a dificuldade dos itens corresponderá à probabilidade de serem respondidos correctamente. Um item com elevada percentagem de respostas correctas é considerado um item fácil. Dito de outra forma, quanto mais difícil for o item, menor a probabilidade de ser respondido correctamente.

A complexidade da tarefa torna-se mais difícil de operacionalizar. A maioria dos estudos nesta área aborda a complexidade da tarefa pelo somatório de operações cognitivas fundamentais a efectuar, podendo-se afirmar, por exemplo, que uma tarefa com 24 peças de Lego é mais complexa que outra construção com 12 peças. Isso não significa, porém, que uma seja mais difícil que a outra, já que ambas podem ter idêntica probabilidade de serem resolvidas correctamente, dentro do tempo que lhes foi atribuído. Se tomarmos por critério de complexidade da tarefa, a quantidade de componentes cognitivas elementares envolvidas, torna-se difícil a operacionalização desta variável para cada um dos itens das MPR-S já que não existe uma descrição exaustiva das componentes cognitivas que são solicitadas para a realização de cada um deles.

A variável complexidade foi, assim, operacionalizada de um modo indirecto, ou seja, através do tempo gasto na resolução dos itens. O tempo que o sujeito gasta para resolver um item (tempo de latência), executando todas as componentes cognitivas nele envolvidas é definido, por alguns autores (por e.g., Schweizer, 1996; Sternberg, 1977), como um indicador da velocidade mental estando directamente relacionada com a complexidade do item. Apesar de desconhecermos as componentes cognitivas envolvidas em cada um dos itens da MPR-S, será de supor que quanto maior o tempo gasto na resolução do item, maior será o número de componentes cognitivas elementares envolvidas (Sternberg, 1977).

Os tempos gastos na execução dos itens tenderão a aumentar à medida que aumenta o grau de complexidade dos mesmos. O nível de complexidade dos itens influencia não só o tempo gasto na resolução do problema, como também, a probabilidade de ser respondido correctamente. O aumento de complexidade está directamente relacionado com as limitações do sistema de processamento de informação. Quanto maior for a complexidade, ou seja, quanto maior o número de componentes cognitivas a considerar na resolução do item, mais tempo o sujeito demora a dar a resposta. Esta limitação estará eventualmente relacionada com as limitações da memória de trabalho (Schweiser, 1996). Os sujeitos diferenciar-se-ão consoante as suas limitações ao nível da memória de trabalho, quer em termos de processamento de informação, quer em termos da capacidade de armazenamento.

Dado que o teste das MPR-S é considerado um teste de inteligência e muitas vezes utilizado como exemplo de uma prova de inteligência fluída, com fraco apelo a conhecimentos adquiridos e uma vez que praticamente todo o suporte informacional necessário para a resolução de cada questão se encontra no próprio estímulo (item), os tempos gastos pelos sujeitos na realização dos itens irão dar-nos uma ideia de como se diferenciam relativamente à velocidade de processamento de informação com diferentes graus de complexidade. Para testar a hipótese de que os tempos gastos no processamento de informação inerente a cada item variam na razão inversa do número de respostas correctas, efectuámos a correlação entre estas

variáveis. A correlação obtida ($r=-0,923$, $p<0,000$) é reveladora da forte interligação entre estas variáveis. Esta relação está graficamente representada na Figura 5.6 que conjuga as variáveis número total de respostas correctas e a média dos tempos de latência.

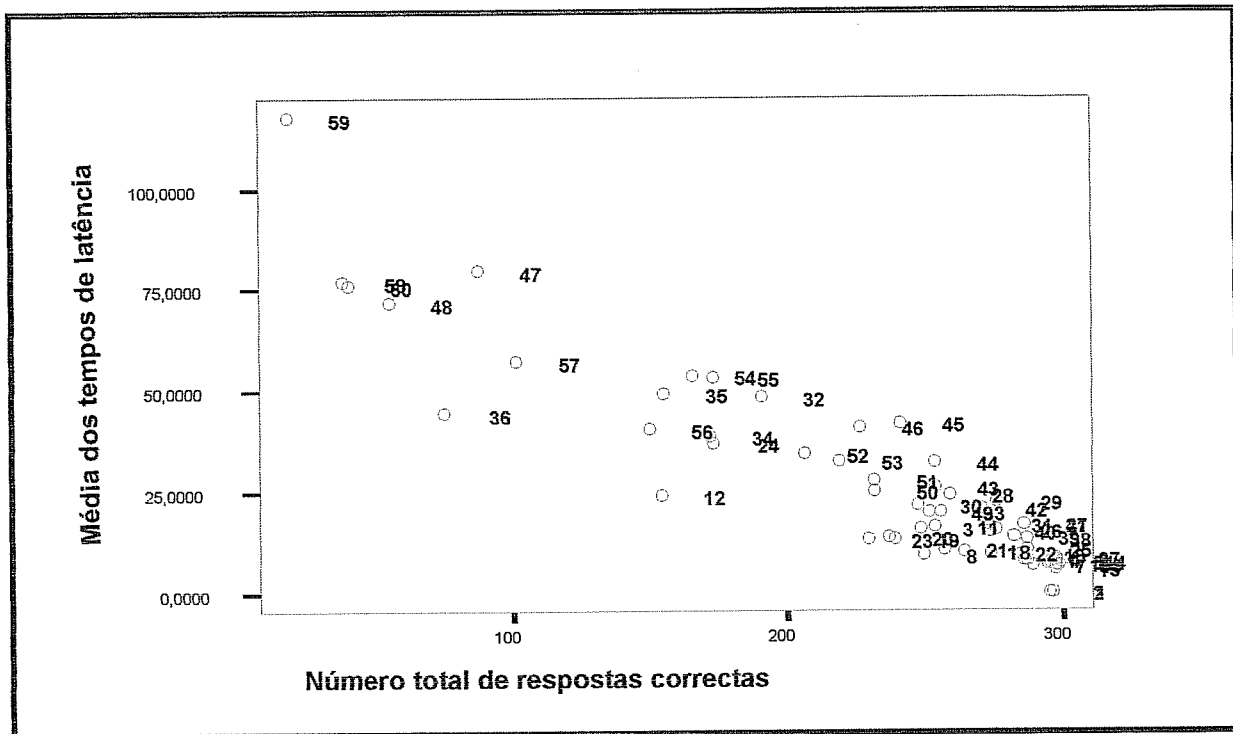


Figura 5.6. Gráfico de dispersão dos itens pelo número total de respostas correctas e pelas médias dos tempos de latência

A Figura 5.6 mostra também que mais de 2/3 da amostra dá respostas correctas em cerca de 45 dos 60 itens da prova, ou seja, pela inversa, apenas 15 itens são respondidos por menos de 200 sujeitos, o que, de algum modo, pode condicionar a evidência dos resultados.

5.4. Velocidade mental e inteligência psicométrica

Para estudar a relação entre a capacidade de processamento de informação dos sujeitos em termos de velocidade mental e a dita inteligência psicométrica, dividiu-se a amostra em três

grupos, de acordo com o resultado global obtido nas próprias MPR-S, ou seja, do número total de respostas correctas. No Quadro 5.3 podemos verificar a distribuição dos sujeitos pelos três níveis de capacidade mental geral.

Quadro 5.3. Distribuição dos sujeitos pelos três níveis de capacidade mental geral

Grupo	N	%
Baixo nível de capacidade mental geral	72	24.0%
Médio nível de capacidade mental geral	153	51.0%
Elevado nível de capacidade mental geral	75	25.0%
	300	100%

Consideraram-se no grupo de baixo nível de capacidade mental geral todos aqueles que deram um total de respostas correctas correspondente a um percentil igual ou inferior a P25. No grupo de elevado nível de capacidade mental geral foram consideradas as pessoas cujo score total na mesma prova correspondesse a um percentil igual ou superior a P75. No grupo médio ficaram os restantes sujeitos.

Considerando que a amostra foi dividida com base nos seus próprios resultados, tendo como critério os quartis, a percentagem de indivíduos em cada um dos grupos está muito equilibrada e de acordo com o esperado que seria 25% nos dois grupos extremos e 50% no grupo central.

Como se pretendeu verificar se havia diferenças nos tempos de latência na resolução de itens com diferentes graus de dificuldade entre os sujeitos dos dois grupos extremos, relativamente ao nível da capacidade mental geral, construiu-se o Quadro 5.4 onde se apresentam as médias e desvios-padrão dos tempos de latência e o número de respostas correctas para cada um dos itens por nível de capacidade mental geral, assim como as diferenças de médias entre os dois grupos extremos e as respectivas análises de variância.

Quadro 5.4. Médias e desvios-padrão dos tempos de latência, e número de respostas correctas para cada um dos itens por nível de capacidade mental geral e respectivas análises de variância.

Item	Baixo Nível de Capacidade Mental Geral			Médio Nível de Capacidade Mental Geral			Elevado Nível de Capacidade Mental Geral			Total			Dif Bx-AI	F.	Sig.
	Méd.	D.P.	N.R .C.	Méd.	D.P.	N.R .C.	Méd.	D.P.	N.R .C.	Méd.	D.P.	N.R .C.			
1ª Série de questões															
1	0,00	0,00	70	0,00	0,00	151	0,00	0,00	75	0,00	0,00	296			
2	0,00	0,00	69	0,00	0,00	151	0,00	0,00	75	0,00	0,00	295			
3	18,92	13,26	64	14,50	10,24	121	14,46	12,75	63	15,63	11,85	248	4,46	3,389	0,035
4	8,64	7,97	66	7,73	4,99	146	6,92	3,50	74	7,73	5,54	286	1,72	1,686	0,187
5	7,65	4,45	69	7,60	5,70	153	6,36	3,20	73	7,31	4,92	295	1,30	1,818	0,164
6	7,31	7,36	71	6,24	4,42	152	5,74	4,55	74	6,37	5,31	297	1,57	1,684	0,187
7	8,13	4,75	63	6,58	3,34	150	5,75	2,28	75	6,70	3,56	288	2,38	8,209	0,000
8	9,84	4,08	55	9,44	5,04	124	8,50	5,31	70	9,26	4,93	249	1,34	1,288	0,278
9	8,86	4,30	70	7,68	4,56	152	6,85	2,23	75	7,75	4,08	297	2,00	4,518	0,012
10	11,27	8,07	64	8,93	6,59	146	7,51	3,18	74	9,09	6,41	284	3,75	6,184	0,002
11	19,70	13,39	47	16,42	16,50	136	12,93	7,45	71	16,05	14,09	254	6,77	3,434	0,034
12	26,73	26,78	26	27,18	37,33	78	17,98	10,47	49	24,16	29,60	153	8,75	1,584	0,209
2ª Série de questões															
13	6,71	3,23	72	6,93	7,06	152	5,84	3,70	74	6,60	5,60	298	0,87	0,958	0,385
14	9,03	7,34	71	7,19	3,90	153	6,87	3,08	75	7,55	4,84	299	2,16	4,587	0,011
15	6,46	3,30	70	5,28	2,45	152	5,01	1,68	75	5,49	2,57	297	1,44	7,025	0,001
16	8,75	4,53	68	6,64	3,39	151	5,77	2,44	75	6,90	3,64	294	2,98	13,932	0,000
17	10,95	7,63	60	7,36	2,93	151	7,01	3,13	74	8,02	4,63	285	3,94	17,015	0,000
18	12,60	8,84	55	9,88	5,18	136	8,38	2,84	73	10,03	5,84	264	4,22	8,753	0,000
19	19,28	16,83	46	12,87	10,84	128	9,60	5,06	65	13,21	11,57	239	9,68	10,302	0,000
20	19,90	11,06	50	12,85	8,81	119	10,47	4,35	68	13,65	9,01	237	9,43	19,356	0,000
21	13,50	10,27	46	10,90	8,95	138	7,86	4,32	73	10,50	8,38	257	5,64	7,027	0,001
22	12,31	6,97	55	9,31	5,44	144	7,79	3,61	75	9,49	5,57	274	4,52	11,442	0,000
23	19,74	15,66	34	12,38	8,82	128	10,63	6,17	67	12,96	9,92	229	9,11	10,860	0,000
24	49,15	45,08	27	35,97	25,57	86	33,23	33,49	60	37,08	32,31	173	15,91	2,399	0,094
3ª Série de questões															
25	11,29	5,37	66	10,91	5,14	149	9,63	3,13	71	10,68	4,81	286	1,65	2,407	0,092
26	16,90	14,52	59	15,20	7,61	146	14,17	7,35	71	15,30	9,47	276	2,73	1,359	0,259
27	20,48	15,54	62	16,68	10,03	149	13,36	6,01	74	16,65	10,93	285	7,12	7,479	0,001
28	32,88	28,38	52	23,48	13,75	137	18,97	9,18	70	24,15	17,44	259	13,91	10,412	0,000
29	26,65	23,53	54	21,66	12,92	148	20,81	9,52	74	22,41	14,98	276	5,84	2,802	0,062
30	28,06	25,20	47	20,52	12,77	128	19,35	11,91	72	21,61	15,94	247	8,72	5,040	0,007
31	18,82	11,47	57	17,18	11,84	141	13,88	9,01	75	16,62	11,16	273	4,94	3,621	0,028
32	57,22	41,66	27	47,87	28,41	98	45,60	30,03	65	48,42	31,18	190	11,62	1,362	0,259
33	24,46	18,54	50	19,79	11,23	135	16,79	6,49	71	19,87	12,28	256	7,67	5,956	0,003
34	53,42	54,20	26	36,22	20,24	86	36,10	19,52	59	38,80	28,38	171	17,32	4,228	0,016
35	64,80	51,95	15	49,14	25,41	80	44,90	24,74	60	49,01	29,02	155	19,90	2,893	0,058
36	51,33	33,32	3	48,24	33,65	33	40,59	22,61	39	44,39	28,25	75	10,74	0,745	0,478

Quadro 5.4. (cont.). Médias e desvios-padrão dos tempos de latência, e número de respostas correctas para cada um dos itens por nível de capacidade mental geral e respectivas análises de variância.

Item	Baixo Nível de Capacidade Mental Geral			Médio Nível de Capacidade Mental Geral			Alto Nível de Capacidade Mental Geral			Total			Dif Bx-AI	F.	Sig.
	Méd.	D.P.	N.R. .C.	Méd.	D.P.	N.R. .C.	Méd.	D.P.	N.R. .C.	Méd.	D.P.	N.R. .C.			
4ª Série de questões															
37	10,46	6,74	69	7,76	3,57	153	7,31	2,40	75	8,28	4,46	297	3,16	11,873	0,000
38	13,94	7,05	64	13,22	8,64	148	12,30	7,27	74	13,14	7,96	286	1,64	0,741	0,478
39	13,60	9,48	62	13,52	10,57	148	13,39	7,85	72	13,50	9,67	282	0,21	0,008	0,992
40	20,85	10,50	55	14,18	7,00	147	11,67	3,59	72	14,86	7,85	274	9,19	26,826	0,000
41	23,06	14,85	62	15,57	8,19	150	12,78	5,27	73	16,48	10,15	285	10,28	21,118	0,000
42	24,30	14,22	54	20,45	10,65	143	17,75	7,42	73	20,49	10,93	270	6,54	5,765	0,004
43	30,52	26,13	48	25,76	17,06	136	22,27	10,18	70	25,70	17,83	254	8,25	3,099	0,047
44	38,57	29,60	49	30,90	16,20	135	29,66	15,15	70	32,04	19,45	254	8,91	3,594	0,029
45	57,24	43,07	42	39,63	23,15	130	36,59	16,58	69	41,83	27,11	241	20,64	9,070	0,000
46	43,97	23,06	38	42,27	25,73	116	36,61	18,26	72	40,75	23,23	226	7,36	1,769	0,173
47	71,00	47,36	17	76,50	34,78	42	90,07	51,37	28	79,79	43,33	87	-19,07	1,267	0,287
48	37,50	26,16	2	72,04	34,46	27	74,44	25,46	25	71,87	30,61	54	-36,94	1,368	0,264
5ª Série de questões															
49	22,41	13,42	44	20,58	12,09	136	16,83	8,87	72	19,83	11,66	252	5,58	3,823	0,023
50	28,89	16,39	36	26,10	14,38	125	20,62	8,87	71	24,85	13,58	232	8,27	5,794	0,004
51	36,50	18,33	32	29,72	18,95	127	20,27	8,81	73	27,68	17,19	232	16,23	13,120	0,000
52	56,87	25,86	15	34,84	27,43	116	28,52	21,78	75	34,15	26,24	206	28,35	7,886	0,001
53	38,42	23,89	26	33,51	26,81	119	28,31	25,73	74	32,34	26,21	219	10,11	1,705	0,184
54	61,64	42,17	14	54,96	29,85	85	50,00	30,08	66	53,55	31,11	165	11,64	0,991	0,373
55	47,00	21,17	18	51,26	30,50	88	57,72	38,43	67	53,32	33,09	173	-10,72	1,091	0,338
56	40,67	18,20	9	44,27	26,90	71	36,10	23,97	69	40,27	25,30	149	4,57	1,845	0,162
57	39,60	22,81	5	55,69	28,50	49	60,96	40,22	47	57,35	34,32	101	-21,36	0,985	0,377
58	33,00		1	60,09	29,84	11	85,54	45,37	26	76,79	42,66	38	-52,54	2,022	0,148
59			0	118,67	76,85	3	117,20	74,75	15	117,44	72,78	18			
60	50,40	23,61	5	78,50	53,05	16	81,05	40,78	19	76,20	44,79	40	-30,65	0,960	0,392

Verifica-se a existência de um padrão idêntico nas séries dois, três e quatro. O tempo de latência aumenta à medida que diminui o número de respostas correctas. Esta tendência já tinha sido referida anteriormente para o teste global, mas aparece agora de uma forma mais evidente ao analisar-se série por série.

No Quadro 5.4 podem, também, ser analisadas as diferenças entre as médias dos tempos de latência aos itens para o grupo de baixo nível de capacidade mental geral e as do grupo de

elevado nível de capacidade mental geral que se encontram na antepenúltima coluna. Nas duas últimas colunas encontra-se o coeficiente “F” de Snedecor relativo à análise de variância para testar a diferença das notas médias observadas nos três grupos com diferentes níveis de capacidade mental geral e o correspondente nível de significância.

Dentro de cada uma das séries, a diferença entre as médias dos tempos de latência para a resolução dos itens nos dois grupos extremos, em termos de capacidade mental geral, tende a aumentar à medida que aumenta o grau de dificuldade dos itens, ou seja, à medida que baixa a probabilidade de ocorrência de respostas correctas. Esta observação leva-nos a pensar que os sujeitos com maior capacidade mental geral diferenciam-se ainda mais dos do grupo inferior, quanto mais exigentes forem as tarefas. Poderíamos afirmar que a relação ou *trade-off* entre a correcção das respostas e o tempo de latência aumenta à medida que aumenta a complexidade da tarefa, quando se comparam grupos com diferentes níveis de capacidade mental geral. Os resultados encontrados vão para além dos encontrados no estudo de Schweizer (1996) em que, de acordo com as teorias da limitação da capacidade do sistema de processamento de informação, nomeadamente em relação à memória de trabalho, a transição entre velocidade e correcção torna-se mais óbvia à medida que aumenta o grau de exigência das tarefas.

De uma forma geral, podemos concluir que os sujeitos com melhores índices de rendimento global do ponto de vista intelectual são mais rápidos a responder aos itens comparativamente aos sujeitos de rendimento mais fraco. No entanto, a diferença entre as médias dos tempos de latência dos dois grupos extremos (alto e baixo nível de capacidade mental geral) é progressiva, mas não rigorosamente contínua ao longo de todo o teste. Nas duas últimas séries, e só relativamente aos itens com menor número de respostas correctas, encontramos uma diferença de médias negativa, ou seja, os sujeitos do grupo de baixo nível de capacidade mental geral obtiveram uma média do tempo de latência inferior aos do grupo de elevado nível de capacidade mental geral. Esta inversão acontece apenas em dois itens da 4ª série e quatro da 5ª série. Nestes itens verificou-se um menor número de respostas correctas em termos globais.

Esta diferença entre as médias dos tempos de latência não é, no entanto, comprovada estatisticamente, ou seja, nos itens onde se verifica a inversão da diferença de médias (nos itens onde os sujeitos do grupo de baixo nível de capacidade mental geral são mais rápidos do que os do grupo elevado), a análise de variância aos tempos para os três grupos não é estatisticamente significativa. Uma hipótese de explicação para estes resultados pode consistir no facto de as MPR-S terem sido aplicadas com limitação de tempo. Considerando que os indivíduos do grupo de baixo nível de capacidade mental geral são mais lentos a dar as respostas que os do grupo elevado, e sendo a prova aplicada com tempo limite, estes chegam às últimas perguntas do teste com muito pouco tempo disponível, arriscando respostas aleatórias às questões mais difíceis e/ou às últimas do teste, na esperança de poder pontuar. Recorde-se que os candidatos nesta prova tem acesso a um cronógrafo digital que está sempre presente no ecrã e lhes dá informação do tempo ainda disponível. Este facto pode explicar porque é que os sujeitos do grupo com menor capacidade apresentam nos itens finais tempos de latência mais rápidos mas com um número de respostas correctas significativamente inferior. Esta aparente maior eficácia de velocidade mental não é validada empiricamente, já que como se disse anteriormente a diferença de resultados não é significativa do ponto de vista estatístico.

Os indivíduos com menor capacidade cognitiva poderão tender a utilizar uma estratégia de “fuga para a frente” quando se vêem confrontados com pouco tempo para resolver itens muito difíceis. Recorde-se que nos itens em que os sujeitos do grupo de baixo nível de capacidade mental geral são mais rápidos de que os do grupo elevado (itens 47, 48, 55, 57, 58 e 60), eles são respondidos correctamente por um número reduzido de sujeitos do grupo baixo, respectivamente 17, 2, 18, 5, 1 e 5.

É, assim, reforçada a ideia de que as diferenças individuais em termos cognitivos não são as únicas causadoras dos resultados obtidos em termos do tempo de latência. Factores não cognitivos podem funcionar como limitadores da expressão da velocidade mental no

processamento de informação do sujeito. No entanto, o presente trabalho não permite a confirmação desta hipótese *post-hoc*, pelo que se reconhece a necessidade de delinear outros estudos que permitam analisar, com maior profundidade, o funcionamento cognitivo do sujeito numa perspectiva integradora das diferentes componentes que influenciam a forma como os sujeitos processam a informação.

5.5. Comentários finais

Os resultados apontam para a confirmação da hipótese de que os itens que são respondidos correctamente por menos pessoas, tender a consumir mais tempo para serem resolvidos. Em estudo semelhante desenvolvido por Schweizer (1996) constatou-se efeito idêntico, ou seja, a transformação de velocidade em precisão, numa relação directa com o aumento da complexidade da tarefa. De acordo com o autor, o sistema de memória de trabalho é apresentado como principal responsável dos resultados obtidos. No entanto, o desenho do nosso estudo não permite concluir que o principal motivo porque em determinados itens os sujeitos dão menos respostas correctas, e demoram mais tempo a responder, esteja relacionado unicamente com as limitações ao nível do sistema da memória de trabalho.

O trabalho parece evidenciar diferenças individuais na execução de tarefas cognitivas. Os sujeitos com maior índice de capacidade mental geral, tendem a ser mais rápidos a dar as respostas correctas que os sujeitos com menor índice de capacidade. Apesar de as diferenças dos tempos de latência entre estes dois grupos tenderem a aumentar à medida que aumenta o grau de dificuldade dos itens, este padrão é significativamente influenciado pela limitação temporal da aplicação do teste. Em provas aplicadas sem limite de tempo, colocamos a hipótese de encontrar resultados diferentes por parte dos testados.

O estudo chama ainda a atenção para dois aspectos específicos desta prova. O primeiro

prende-se com o facto de a prova não estar construída com os itens dispostos numa sequência contínua do grau de complexidade/dificuldade. Esta situação permite que factores de ordem não-cognitiva se possam expressar e desta forma mascarar o resultado obtido pelos sujeitos. Os tempos de latência gastos pelos sujeitos nos itens iniciais, intermédios e finais de cada série são diferentes entre si. Estas diferenças dos tempos de latência dos itens relativamente às suas posições na série, têm, inclusive, repercussões quando comparamos os sujeitos em termos do seu nível de capacidade mental geral. Se analisarmos o Quadro 5.4, verifica-se que as diferenças entre as médias dos três grupos com diferentes capacidades mentais são apenas estatisticamente significativas quando os itens se encontram no meio da série, e não nos iniciais e finais, com excepção para a primeira série e última séries. Numa análise, não suficientemente fundamentada, podemos lançar a hipótese de que se os itens iniciais e finais de cada série são de algum modo, desestruturantes para os sujeitos, atendendo ao grau de complexidade percebido comparativamente com os intermédios. Apesar de o sujeito estar avisado de que o grau de dificuldade é crescente mas não constante, quando depara com um item significativamente mais fácil que os itens resolvidos até aí (mudar de série), pode acontecer que o seu rendimento seja contaminado por factores não cognitivos. O confronto com um item substancialmente menos exigente do que os anteriores, pode levar a que o sujeito duvide da rapidez com que chegou à solução. O rendimento do sujeito é, assim, afectado, não só pelas suas limitações em termos de processamento de informação, como também por aspectos de ordem não-cognitiva. Provavelmente, indivíduos com menor autoconfiança poderão gastar mais tempo que outros com maior autoconfiança ao tentarem confirmar que a resposta a que chegaram é a que julgam ser correcta. Assim sendo, os primeiros itens de cada série são mais sensíveis à ocorrência deste fenómeno que os restantes. Esta sensibilidade poderá justificar a ausência de diferenças significativas entre os grupos de diferentes níveis de capacidade mental geral nos itens iniciais. Se estivessem em causa, exclusivamente as capacidades de processamento de informação, a diferença entre os dois grupos extremos tenderia a manter o mesmo diferencial em termos de tempos de latência (ver Quadro 5.4). Pelo atrás exposto,

colocamos a hipótese que a média dos tempos de latência observada para os primeiros itens de cada uma das séries não reflecte apenas as limitações cognitivas dos sujeitos, mas também o grau de afectação da sua capacidade de processamento da informação por factores de ordem não cognitiva.

O outro aspecto específico desta prova prende-se com o facto de ter sido utilizado um teste cognitivo com limitação do tempo de aplicação. A análise comparativa entre grupos extremos em termos de eficiência cognitiva permitiu evidenciar uma tendência para se encontrar tempos mais reduzidos na realização dos itens para o grupo dos mais favorecidos, reforçando a existência de um *trade-off* entre a rapidez e a exactidão das respostas já identificado anteriormente (Schweiser, 1996). Esta relação apesar de ser significativa em termos gerais, não se revelou universal para os 60 itens do teste. Para uma análise mais fina seria necessário ter acesso a uma outra variável que permitisse identificar o grau de certeza com que o sujeito deu a resposta, de modo a evitar a inclusão no estudo dos tempos de latência que não correspondessem a decisões convictas do sujeito. Essa nova variável iria reduzir a probabilidade de que algumas respostas fossem resultantes de “fugas para a frente”, ou seja, face ao pouco tempo disponível, o sujeito decide arriscar, respondendo de forma aleatória.

Considerações Finais

Se bem que a grande preocupação na avaliação cognitiva tenha consistido, essencialmente, no aperfeiçoamento da medição da capacidade do sujeito num atributo específico, através do número de respostas correctas a um conjunto de itens com dificuldade variada, o presente trabalho chama a atenção para uma outra componente da medição, que é o tempo gasto na execução da tarefa. De acordo com a perspectiva de Simon (1978, citado por Simões, 1994), a resolução dos itens de um teste depende de três componentes: (a) as capacidades cognitivas do sujeito; (b) a maneira como representa o problema; e (c) a estrutura da própria tarefa. A combinação destas componentes pode dar-nos a informação necessária para a compreensão do processamento cognitivo individual e identificação da estratégia predominante do sujeito. Entre vários sujeitos apresentando a mesma capacidade, pode haver diferenças de estilo ligada à distribuição do tempo e do esforço (Simões, 1994). Nos estudos empíricos deste trabalho procurou-se abordar a capacidade cognitiva dos sujeitos, não somente pela produtividade dos indivíduos, mas sobretudo pela velocidade do processamento cognitivo para a realização de dois testes. Os tempos de latência foram, assim, o eixo transversal do enquadramento teórico e do trabalho empírico realizados, como forma de se conhecer melhor as estratégias de resolução de itens pelos sujeitos.

Começámos por clarificar os conceitos associados à velocidade mental no processamento cognitivo de informação. Nesta perspectiva, diferenciámos os conceitos associados ao funcionamento cognitivo, nomeadamente os termos ingleses *capacity*, *aptitude* e *ability*. A

distinção entre eles não é clara e perfeitamente estanque. Encontraram-se posturas diferentes e por vezes contraditórias (Cooper, 1999; Cronbach, 1984; Howe, 1996; Kline, 2000; Marañón, 1997; Sternberg, 1982). *Capacity* corresponde a uma noção de carácter inato e está relacionada com uma área específica, influenciando a realização nos factores que lhe estão subjacentes. O conceito *ability* deriva directamente do de *capacity*, correspondendo a atributos mais específicos, cujo valor resulta da nota obtida numa prova específica para a avaliar. O conceito *aptitude* está relacionado com o sucesso futuro numa área profissional específica. Neste estudo, considerámos aptidão como tradução de *ability*, correspondendo à capacidade de um atributo específico capaz de se expressar na realização de uma tarefa particular. Capacidade corresponderia, por seu lado, a todo o volume do atributo em causa que o sujeito possui, e que se identificaria com a respectiva aptidão, se a tarefa na qual o sujeito está a ser avaliado fosse 100% representativa desse atributo e realizada nas condições teoricamente ideais.

A relação entre o tempo de reacção e a inteligência geral mereceu uma atenção particular no nosso estudo, assim como a memória de trabalho cuja capacidade de armazenamento/processamento se pressupõe ser limitadora da expressão da inteligência geral (Süb *et al.*, 2002). Também desenvolvemos o conceito de complexidade cognitiva e apresentámos as suas implicações na execução de tarefas cognitivas.

Um estudo de Ree e colaboradores (1994) demonstra, através de análise de regressão múltipla, com correcção para a restrição de amplitude, que o *factor g* é o melhor preditor do rendimento num conjunto de postos de trabalho, nomeadamente em contexto aeronáutico. A importância atribuída ao *factor g* na predição de comportamentos relacionados com o treino e o desempenho aumentou o interesse pelo conhecimento do funcionamento cognitivo. É neste propósito que a avaliação psicológica informatizada pode ajudar a compreender melhor como o sujeito processa a informação, por intermédio da análise da velocidade mental. Os tempos

de latência aos itens de testes de aptidão, podem fornecer informação pertinente para clarificar a forma como o sujeito processa a informação neste tipo de situações. Em contexto organizacional, a avaliação psicológica informatizada pode, também, incrementar a capacidade de predição dos modelos utilizados na selecção de pessoal, que por intermédio da utilização da variável velocidade mental, quer pela avaliação de novos constructos.

Efectuámos dois estudos empíricos, com base na aplicação de dois testes informatizados em contexto de selecção, com o objectivo principal de conhecermos melhor o rendimento dos sujeitos nos itens a partir da variável relativa aos tempos de latência. No primeiro estudo, pretendeu-se conhecer a relação entre a exactidão das respostas dadas pelos sujeitos e os tempos de latência às mesmas, explorando a possibilidade de interferência de factores não cognitivos na execução de itens de escolha múltipla. Foi aplicado o teste “Padrões” a 731 candidatos ao curso de Pilotagem Aeronáutica da Academia da Força Aérea Portuguesa nos anos de 1998 e 1999. Foi aplicado informaticamente e é um teste do tipo das figuras embebidas, desenvolvido por Eugene Burke em 1995. É uma medida de *perceptual closure*, onde se solicita ao sujeito que identifique uma figura-alvo em dois padrões desorganizados visualmente, que representam um fundo não estruturado (People Technologies, 2000).

Dos resultados obtidos, verificou-se que os tempos médios de resposta aos itens, são diferentes quando os sujeitos acertam ou erram os itens. Os sujeitos que acertam os itens tendem a ser mais rápidos e mais homogéneos no tempo gasto para dar as suas respostas do que os que erram. A proficiência dos sujeitos nesta prova de *flexibility of closure*, medida através do número de respostas correctas, diferencia-se da rapidez com que processam a informação avaliada pelo tempo médio das respostas dadas. A rapidez com que o sujeito processa a informação, parece pouco relacionada com uma capacidade mental geral (correlação da inteligência geral com o tempo médio de resposta é de 0,001), mas sobretudo com as componentes perceptiva e velocidade cognitiva. Verificou-se, também, que os

primeiros e os últimos itens comportam-se de modo contrário à maioria dos restantes. Apesar de, não se ter objectivado a complexidade cognitiva de cada um dos itens, podemos considerar, que o facto de o sujeito não saber que pelo menos uma das figuras complexas contém a figura-alvo, é um elemento contribuidor para a complexidade do item. Este aspecto é agravado pelo conhecimento que o sujeito tem do tempo que falta para terminar, que pode constituir um indicador da resistência a situações de pressão. Se considerarmos que o grau de incerteza na resposta pretendida contribui para a complexidade do item, então os resultados encontrados vão ao encontro à posição de Schweizer (1996), em que o número de respostas correctas tende a decrescer e o tempo de reacção a aumentar, à medida que aumenta o grau de complexidade; do mesmo modo, as respostas correctas nas tarefas mais complexas tende a apresentar correlações mais elevadas com as aptidões cognitivas do que nas tarefas menos complexas. De acordo com a perspectiva de Wilhem e Schulze (2002), podemos considerar a velocidade mental como um indicador da inteligência geral, segundo duas condições: quando as tarefas complexas de velocidade mental se opõem às simples e quando as tarefas de raciocínio em velocidade, ou com tempo limite, se opõem às de rendimento ou sem limitações de tempo.

No segundo estudo, procurou-se conhecer em detalhe o modo como os sujeitos evoluem nas respostas a dar na realização de um teste informatizado de inteligência geral numa situação real de avaliação. Utilizámos as Matrizes Progressivas de Raven na sua forma *Standard*, mas de aplicação em suporte informático. A prova foi aplicada a 300 sujeitos candidatos a condutores de veículos pesados de passageiros. Os resultados apontam para a confirmação da hipótese de que à medida que os itens vão apresentando maior índice de complexidade, maior é o tempo gasto na sua resolução. A hipótese realçada no estudo de Schweizer (1996), é de novo confirmada neste segundo estudo: a transformação da velocidade em precisão, apresenta uma relação directa com o aumento da complexidade da tarefa. Schweizer coloca como principal hipótese explicativa, a limitação do sistema de processamento de informação devido

à capacidade da memória de trabalho. Contudo, no nosso estudo não nos foi possível testá-la, já que não foi utilizada qualquer prova avaliadora da memória de trabalho.

Quando se dividiu a amostra em dois grupos com diferentes níveis de rendimento intelectual global, verificou-se que os sujeitos com maior índice de capacidade mental geral tendem a dar as respostas correctas mais rapidamente que os sujeitos com menor índice de capacidade mental geral. Apesar de as diferenças dos tempos de latência entre estes dois grupos tenderem a aumentar à medida que aumenta o grau de complexidade dos itens, este padrão é significativamente influenciado pelo tempo limite para a realização do teste. Estes resultados corroboram os estudos efectuados por Kyllonen e Christal (1990) e por Vernon e colaboradores (1985b), em que a velocidade de processamento traz duas vantagens para o sujeito mais inteligente ao realizar um teste de aptidão mental com complexidade crescente. Nos itens iniciais (menos complexos) o indivíduo acede mais rapidamente à informação necessária e responde mais rapidamente às perguntas. Nas questões com menor probabilidade de se obterem respostas correctas, os sujeitos mais inteligentes, ao serem mais rápidos a processarem informação, têm a possibilidade de executá-las sem que a capacidade (limitada) da memória de trabalho entre em sobrecarga (Süb *et al.*, 2002).

Este segundo estudo, à semelhança do primeiro, veio evidenciar a influência de factores de ordem não-cognitiva na resolução de tarefas cognitivas. Os tempos de latência gastos pelos sujeitos nos itens iniciais, intermédios e finais de cada série são diferentes entre si. Estas diferenças dos tempos de latência de resposta aos itens, relativamente à posição que ocupam na série, têm repercussões quando comparamos os sujeitos em termos do seu nível de capacidade mental geral. As diferenças entre os grupos de diferentes capacidades mentais, em termos gerais, são apenas estatisticamente significativas para os itens intermédios de cada série. Não se encontraram diferenças significativas nos itens iniciais e finais, salvo para a primeira e última séries. A hipótese que se poderá formular, embora sem sustentação

definitiva, é a de que os itens iniciais e finais de cada série do teste das Matrizes Progressivas de Raven na sua forma *Standard* são, de algum modo, desestruturantes para os sujeitos, atendendo à complexidade percebida. Provavelmente, estará relacionado com o tipo de componentes cognitivas fundamentais envolvidas comparativamente com os itens intermédios. Esta hipótese *pos-thoc* encontra consistência nos trabalhos de Stankov (2000), que associa o aumento de complexidade cognitiva das tarefas à maior probabilidade de emergência de factores não cognitivos.

Não obstante de os sujeitos saberem que o grau de dificuldade é crescente em cada série de itens do teste das Matrizes Progressivas de Raven na sua forma *Standard*, o confronto com os primeiros itens de cada série, substancialmente menos complexos que os últimos da série imediatamente anterior, leva o sujeito a duvidar da rapidez com que chegou à solução e consequentemente da resposta dada. O rendimento do sujeito é assim afectado, não só pelas suas limitações em termos de processamento de informação, como também por aspectos de ordem não-cognitiva. Provavelmente, indivíduos com menor autoconfiança poderão gastar mais tempo que outros com maior autoconfiança ao tentarem confirmar que a resposta a que chegaram é a que julgam ser correcta.

Por outro lado, o facto de ter sido utilizado um teste cognitivo com tempo limite de aplicação pode explicar os tempos de latência encontrados nas respostas aos últimos itens, quando se comparam os resultados dos dois grupos extremos em termos de capacidade mental. Os sujeitos com nível de capacidade mental geral mais baixo foram mais rápidos a responder aos últimos itens, do que os sujeitos com melhor nível de capacidade mental geral, apesar de apresentarem um número reduzido de respostas correctas. Para uma análise mais fina e correcta seria necessário ter acesso a uma outra variável que permitisse identificar o grau de certeza com que o sujeito deu a resposta, evitando que sejam considerados na análise estatística os tempos de latência não correspondentes a decisões convictas do sujeito. Essa

nova variável iria reduzir a probabilidade de inclusão como boas respostas, as que foram dadas como “fugas para a frente”, ou seja, face ao pouco tempo disponível, o sujeito decide arriscar, podendo acertar o item de forma não pensada.

Dos dois estudos realizados, parece ser claro que o acesso ao tempo de latência das respostas, conseguido através da informatização dos testes, permite-nos conhecer melhor, não só as diferenças individuais ao nível do funcionamento cognitivo do atributo que se está a avaliar, como também a própria prova.

Foi evidenciado na parte prática deste trabalho a possibilidade de interferência de factores não-cognitivos no desempenho final aquando da realização de provas vincadamente mentais e com tempo limite de aplicação. Os tempos de latência conseguidos por intermédio da utilização de provas informatizadas poderão ajudar a compreender melhor a relação entre a capacidade cognitiva do sujeito e a aptidão demonstrada. O estudo da possibilidade de interferência de factores não-cognitivos poderá dar contributos válidos para quem trabalha em selecção de pessoal e tem por objectivo escolher os mais capazes nas aptidões consideradas pertinentes, como também que sejam bons gestores da componente não-cognitiva. Defendemos que devem ser desenvolvidos estudos que cruzem a avaliação das capacidades intelectuais do sujeito com características não-cognitivas (autoconfiança, resistência a situações de pressão e de frustração, entre outras) de modo a permitir um conhecimento mais integrado do sujeito. Desenvolver um processo de selecção de pessoal com informação que permita conhecer na globalidade o sujeito, torna a escolha mais válida e justa perante os vários candidatos ao posto de trabalho. Quando a validade preditiva assume um particular interesse, como é o caso da selecção de pessoal, a utilização de provas de inteligência geral de realização em velocidade (ou com tempo limite de aplicação), poderão ser melhor opção do que provas de constructo idêntico mas de rendimento (ou sem tempo limite), sobretudo se o

rendimento tomado para critério conter semelhante pressão de tempo (Wilhem & Schulze, 2002).

Quando falamos da interferência e controlo dos factores não-cognitivos na realização de tarefas do foro intelectual estamos a estabelecer pontos de ligação entre a capacidade pretensamente avaliada e todos os aspectos que do ponto de vista personalístico possam limitar a emergência da aptidão do sujeito. Um pouco à semelhança do que é comumente designado por Quociente Intelectual (domínio das competências funcionais específicas aquele trabalho) e Quociente Emocional (capacidade para gerir emoções) para diferenciar as diferentes competências que contribuem para a obtenção de sucesso em determinados postos de trabalho, consideramos que os resultados alcançados na realização de tarefas intelectuais não dependem em exclusivo das respectivas capacidades cognitivas que pressupostamente se estão a avaliar. Características não-cognitivas poderão interferir na estratégia de controlo do esforço e tempo.

Os nossos dois estudo têm, no entanto, limitações de ordem confirmatória, já que não permite objectivar de modo sustentado a interferência de variáveis não-cognitivas no processamento mental do sujeito, como por exemplo, características da personalidade, motivação individual e questões de ordem fisiológica. No entanto, ao nível da motivação, verificou-se um controlo mínimo, já que todos os sujeitos se encontravam na mesma situação, ou seja, num processo real de selecção de pessoal.

Realça-se, mais uma vez, o facto de que os estudos ao terem sido desenvolvidos em contexto real de selecção, permite-nos evidenciar a utilidade desta pesquisa, já que os resultados encontrados podem ter consequências directas nos procedimentos de selecção actualmente em uso.

Referências Bibliográficas

Academia das Ciências de Lisboa & Editora Verbo. (2001). *Dicionário de Língua Portuguesa Contemporânea da Academia das Ciências de Lisboa*. Lisboa: Verbo.

Algera, J. A., & Greuter, M. A. M. (1989). Job analysis for personnel selection. In M. Smith & I. T. Robertson (Eds.), *Advances in selection and assessment* (pp. 7-30). Chichester: Wiley.

Almeida, L. S. (1988). *Teorias da inteligência* (2ª Ed.). Porto: Ed. Jornal de Psicologia.

Almeida, L. S. (1994). *Inteligência: Definição e medida*. Porto: CIDInE.

Almeida, L. S., & Sá, J. M. S. (1987). Uso dos computadores na avaliação psicológica: Estudos em testes de aptidão intelectual. *Jornal de Psicologia*, 395-404.

Arnold, J., Cooper, C. L., & Robertson, I. T. (1995). *Work psychology*. London: Pitman Publishing.

Baddeley, A. (1986). *Working Memory*. Oxford: Clarendon Press.

Baddeley, A., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8, pp. 47-90). New York: Academic.

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44, 1-26.

Bártolo Ribeiro, R. (1992). Predição da performance em psicologia aeronáutica: Validação de uma bateria de selecção. *Análise Psicológica, Série X*, 3, 353-365.

Bártolo Ribeiro, R. (1995). Métodos de selecção: Soluções teóricas vs realidade portuguesa. *Comportamento Organizacional e Gestão*, 1, 49-65.

Bártolo Ribeiro, R. (1996). Selecção de pessoal: Métodos e aplicações. In C. A. Marques & M. P. Cunha (Orgs.), *Comportamento organizacional e gestão de empresas* (pp. 93-115). Lisboa: Publicações D. Quixote.

- Bartram D., & Bayliss R. (1984). Automated testing: Past, present and future. *Journal of Occupational Psychology*, 57(3), 221-237.
- Bartram, D., (1989). Computer-based assessment. In P. Herriot (Ed.), *Handbook of assessment in organisations*. London: Wiley.
- Bartram, D., (1994). Computer-based assessment. In C. L. Cooper & I. T. Robertson (Eds.), *International Review of Industrial and Organizational Psychology*, 9, 31-69.
- Brogden, H. E. (1949). When testing pays off. *Personnel Psychology*, 2, 171-183.
- Burke M. J., Normand J., & Raju N. S. (1987). Examinee attitudes toward computer-administered ability testing. *Computers in Human Behavior*, 3, 95-107.
- Burke, M. J., & Normand, J. (1987). Computerized psychological testing: Overview and critique. *Professional Psychology: Research and Practice*, 18, 42-51.
- Caetano, A. (1996). *Avaliação de desempenho: Metáforas, conceitos e práticas*. Lisboa: Editora RH.
- Carretta, T. R. (1987a). Basic attributes tests (BAT) system: Development of an automated test battery for pilot selection. (AFHRL-TR-87-9). Brooks Air Force Base, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.
- Carretta, T. R. (1987b). The basic attributes tests: An experimental selection and classification instrument for USAF pilot candidates. In R. S. Jensen (Ed.), *Proceedings of the fourth International Symposium on Aviation Psychology*. Columbus, OH: Ohio State University.
- Carretta, T. R. (1987c). Time-sharing ability as a predictor of flight training performance. (AFHRL-TP-86-69). Brooks Air Force Base, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.
- Carretta, T. R. (1990). Cross-validation of experimental USAF pilot training performance models. *Military Psychology*, 2, 257-264.
- Carroll, J. B. (1976). Psychometric tests as cognitive tasks: A new "structure of intellect". In L. B. Resnick (Ed.), *The nature of intelligence*. Hillsdale, NJ: Erlbaum.
- Carroll, J. B. (1980a). Individual difference relations in psychometric and experimental cognitive tasks. Chapel Hill, NC: The L. L. Thurstone Psychometric Laboratory, University of North Carolina, Report N° 163.
- Carroll, J. B. (1980b). Remarks on Sternberg's "Factor theories of intelligence are all right almost". *Educational Researcher*. 9 (8), 14-18.

- Carroll, J. B. (1992). Cognitive abilities: The state of the art. *Psychological Science*, 3, 266-270.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. N. Y. : Cambridge University Press.
- Cascio, W. F. (1982). *Costing human resources: The financial impact of behavior in organisations*. Kent, Boston, MA.
- Cascio, W. F., & Ramos, R. A. (1986). Development and application of a new method for assessing job performance in behavioral/economic terms. *Journal of Applied Psychology*, 71, 20-28.
- Cohn, S. J., Carlson, S. J., & Jensen, A. R. (1985). Speed of information processing in academically gifted youths. *Personality and Individual Differences*, 6, 621-629.
- Conway, A. R. A., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, 30, 163-184.
- Cooper, C. (1999). *Intelligence and abilities*. London: Routledge.
- Cronbach, L. J. (1984). *Essentials of psychological testing*. (4th Ed.). New York: Harper & Row.
- Cronbach, L. J. (1990). *Essentials of psychological testing*. (5th Ed.). New York: Harper & Row.
- D'Oliveira, T., & Bártolo Ribeiro, R. (2000, Outubro). Novas tecnologias e suas repercussões na selecção de pessoal. Comunicação apresentada no IV Encontro sobre Comportamento Organizacional e Gestão. Lisboa: Instituto Superior de Psicologia Aplicada.
- Dennis, I., & Evans, J. (1990). The consequences of speed-accuracy trade-offs for latency based testing (HAL Tech. Rep. 5-1990). Plymouth, England: Polytechnic South West, Department of Psychology, Human Assessment Laboratory.
- Dennis, I., & Evans, J. (1991). Speed-accuracy tradeoff in testing: An empirical study (HAL Tech. Rep. 1-1990). Plymouth, England: Polytechnic South West, Department of Psychology, Human Assessment Laboratory.
- Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, 41, 417-440.
- Doron, R., & Parot, F. (2001). Dicionário de Psicologia. Lisboa: CLIMEPSI. (Tradução de *Dictionnaire de Psychologie*. PUF, 1991)

- Drasgow, F., Olson, J. B., Keenan, P. A., Moberg, P., & Mead, A. D. (1993). Computerized assessment. *Research in Personnel and Human Resources Management, 11*, 163-206.
- Eliot J., & Smith I. M. (1983). *International Directory of Spatial Tests*. Windsor: NFER-Nelson.
- Estes, W. K. (1982). Learning, memory and intelligence. In R. J. Sternberg (Ed.), *Handbook of Human Intelligence* (pp. 170-224). Cambridge: Cambridge University Press.
- Eysenck, H. J. (1986). Inspection time and intelligence: A historical introduction. *Personality and Individual Differences, 7*, 603-607.
- Fernández-Ballesteros, R. (1995, October). Crucial issues in the field of psychological assessment and evaluation during the next decades. *International Association of Applied Psychology, Newsletter, 7(2)*, 13-20.
- Fernández-Ballesteros, R., & Steyer, R. (2001). Special section on psychological assessment standards and guidelines. *European Journal of Psychological Assessment, 17(3)*, 155-156.
- Flanagan, J. L. (1954). The critical incidents technique. *Psychological Bulletin, 51*, 327-358.
- Fleishman, E. A. (1988). Some new frontiers in personnel selection research. *Personnel Psychology, 41(4)*, 679-701.
- French, J. W. (1951). The description of aptitude and achievement tests in terms of rotated factors. *Psychometrics Monographs, 5*
- Green, B. F., Bock, R. D., Humphreys, L. G., Linn, R. L., & Reckase, M. D. (1984). Technical guidelines for assessing computerized adaptive tests. *Journal of Educational Measurement, 21*, 347-360.
- Guilford, J. P. (1982). Cognitive psychology's ambiguities: Some suggested remedies. *Psychological Review, 89*, 48-59.
- Gustafsson, J-E. (1984). A unifying model for the structure of intellectual abilities. *Intelligence, 8*, 179-203.
- Gustafsson, J-E. (1989). Broad and narrow abilities in research on learning and instruction. In R. Kanfer, P. L. Ackerman & R. Cudeck (Eds.), *Abilities, motivation and methodology: The Minnesota symposium on learning and individual differences* (pp. 203-237). Hillsdale, NJ: Erlbaum.
- Haier, R. J. (1990). Editorial: The end of intelligence research. *Intelligence, 14*, 371-374.
- Haier, R. J., Siegel, B.V. Jr., Nuechterlein, K. H., Hazlett, E., Wu, J. C., Paek, J., Browning, H.L., & Buchsbaum, M. S. (1988). Cortical glucose metabolic rate correlates of abstract reasoning and attention studied

with positron emission tomography. *Intelligence*, 12, 199-217.

Howe, M. J. A. (1996). Concepts of ability. In I. Dennis & P. Tapsfield (Eds.), *Human abilities: Their nature and measurement*. NJ: Lawrence Erlbaum Associates, Inc.

Hunt E., & Pellegrino J. W. (1985). Using interactive computing to expand intelligence testing: A critique and prospectus. *Intelligence*, 9, 207-236.

Hunt, E. (1980). Intelligence as an information processing concept. *British Journal of Psychology*, 71, 449-474.

Hunter, D. R. (1989). Aviator selection. In Wiskoff & Rampton (Eds.), *Military personnel measurement: Testing, assignment, evaluation*. Nova York: Praeger.

Hunter, J. E. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior*, 29, 340-362.

Hunter, J. E., & Schmidt F. L. (1982). Fitting people to jobs: The impact of personnel selection on national productivity. In M. D. Dunnette & E. A. Fleishman (Eds.), *Human performance and productivity* (Vol. 1, pp. 233-284). Hillsdale, NJ: Erlbaum.

International Test Commission (2000). *International guidelines for test use*. Stockholm: Autor.

Jackson, D. N. (1983). *Multidimensional aptitude battery*. Port Huron, MI: Research Psychologists Press, Inc.

Jensen, A. R. (1982a). Reaction time and psychometric g. In H. J. Eysenck (Ed.), *A model for intelligence* (pp. 93-132). Nova York: Springer-Verlag.

Jensen, A. R. (1982b). The chronometry of intelligence. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 1, pp. 255-310). Hillsdale, NJ: Erlbaum.

Kaplan, R. M., & Saccuzzo, D. P. (1997). *Psychological testing: Principles, applications, and issues* (4th Ed.). Pacific Grove: Brooks/Cole Publishing Company.

Keating, D. P., & Bobbitt, B. L. (1978). Individual and developmental differences in cognitive-processing components of mental ability. *Child Development*, 49, 155-167.

Kline, P. (1986). *A handbook of test construction: Introduction to psychometric design*. London: Methuen & Co. Ltd.

Kline, P. (1993). *Handbook of psychological testing*. London: Routledge.

- Kline, P. (2000). *Handbook of psychological testing* (2nd Ed.). London: Routledge.
- Kyllonen, P. C. (1994). Cognitive abilities testing: An agenda for the 1990s. In M. G. Rumsey, C. B. Walker & J. H. Harris (Eds), *Personnel selection and classification*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working memory capacity?!. *Intelligence, 14*, 389-433.
- Larson, G. E., Merritt, C. R., & Williams, S. E. (1988). Information processing and intelligence: Some implications of task complexity. *Intelligence, 12*, 131-147.
- Lima, M. L., Vala, J., & Monteiro, M. B. (1988). Os determinantes da satisfação organizacional: Confronto de modelos. *Análise Psicológica, VI* (3/4), 441-457.
- Lohman, D. F. (1989). Estimating individual differences in information-processing using speed-accuracy models. In R. Kanfer, P. L. Ackerman & R. Cudeck (Eds.), *Abilities, motivation and methodology: The Minnesota symposium on learning and individual differences* (pp. 119-164). Hillsdale, NJ: Erlbaum.
- Lohman, D. F. (1994a). Component scores as residual variation (or why the intercept correlates best). *Intelligence, 19*, 1-11.
- Lohman, D. F. (1994b). Implications of cognitive psychology for ability testing: Three critical assumptions. In M. G. Rumsey, C. B. Walker & J. H. Harris (Eds.), *Personnel selection and classification* (pp. 145-172). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- MacLennan, R. N., Jackson, D. N., & Bellantino, N. (1988). Response latencies and the computerized assessment of intelligence. *Personality and Individual Differences, 9*, 811-816.
- Marañón, B. R. C. (1997). *Capacidades humanas*. Madrid: Ediciones Pirámide.
- Marshalek, B., Lohman, D. F., & Snow, R. E. (1983). The complexity continuum in the radex and hierarchical models of intelligence. *Intelligence, 7*, 107-127.
- Matarazzo, J. D. (1992) Psychological testing and assessment in the 21st century. *American Psychologist, 47*, 1007-1018.
- Maxwell, A. E., Fenwick, P. B. C., Fenton, G. W., & Dollimore, J. (1974). Reading ability and brain function: A simple statistical model. *Psychological Medicine, 4*, 274-280.
- McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems in cascade. *Psychological Review, 86*, 287-330.

- McHenry, J. J., & Schmitt, N. (1994). Multimedia testing. In M. G. Rumsey, C. B. Walker & J. H. Harris (Eds.), *Personnel selection and classification* (pp. 193-232). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- McHenry, J. J., Hough, L. M., Toquam, J. L., Hanson, M. A., & Ashworth, S. (1990). Project a validity results: The relationship between predictor and criterion domains. *Personnel Psychology, 43*, 335-353.
- Mead, A. D., & Drasgow, F. (1993). Equivalence of computerized and paper-and-pencil cognitive ability tests: A meta-analysis. *Psychological Bulletin, 114*, 449-458.
- Miller, L. T., & Vernon, P. A. (1992). The general factor in short-term memory, intelligence and reaction time. *Intelligence, 16*, 5-29.
- Mitchell, T. R., & Larson, J. R. (1987). *People in organizations: An introduction to organizational behavior*. Singapore: McGraw-Hill.
- Morales, M. (1991). *The function of general and specific abilities in the validity of the Air Force Officer Qualifying Test*. Unpublished master's thesis, St. Mary's University, San Antonio, Texas.
- Muchinsky, P. M. (1986). Personnel selection methods. In C. L. Cooper and I. T. Robertson (Eds.), *International Review of Industrial and Organizational Psychology*. Chichester and New York: John Wiley.
- Necka, E. (1991a). Reaction time and intelligence. *European Journal for High Ability, 1*, 211-221.
- Necka, E. (1991b, July). States, traits, and dealing with novelty. Paper presented at the Vth Meeting of the International Society for the Study of Individual Differences. Oxford.
- Necka, E. (1992). Cognitive analysis of intelligence: The significance of working memory processes. *Personality and Individual Differences, 13*, 1031-1046.
- Nettlebeck, T., & Lally, M. (1976). Inspection time and measured intelligence. *British Journal of Psychology, 67*, 17-22.
- Olea, M. M., & Ree, M. J. (1994). Predicting pilot and navigator criteria: Not much more than g. *Journal of Applied Psychology, 79*, 845-851.
- Ones, D. S., Schmidt, F. L., & Viswesvaran, C. (1994, July). Correlates of pre-employment Integrity Tests. Artigo apresentado no 23º Congresso Internacional de Psicologia Aplicada, Madrid.

- Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (1993). Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology, 78*, 679-703.
- Pasquali, L. (1996). Teoria da resposta ao item – IRT: Uma introdução. In L. Pasquali (Org.), *Teoria e métodos de medida em ciências do comportamento* (pp. 173-195). Brasília: Laboratório de Pesquisa em Avaliação e Medida / Instituto de Psicologia / UnB: INEP.
- Pellegrino, J. W., & Glaser, R. (1979). Cognitive correlates and component in the analysis of individual differences. *Intelligence, 3*, 187-214.
- Pellegrino, J. W., & Hunt, E. B. (1989). Computer-controlled assessment of static and dynamic spatial reasoning. In R. F. Dillon & J. W. Pellegrino (Eds.), *Testing: Theoretical and applied perspectives* (pp. 174-198). Westport, CT: Praeger.
- People Technologies (2000). *PILAPT Handbook*. Surrey: People Technologies, Inc.
- Petrill, S. A., Luo, D., Thompson, L. A., & Detterman, D. K. (2001). Inspection time and the relationship among elementary cognitive tasks, general intelligence, and specific cognitive abilities. *Intelligence, 29*, 487-496.
- Prieto, J. M. (1993). The team perspective in selection and assessment. In H. Schuler, J. L. Farr & M. Smith (Eds.), *Personnel selection and assessment: Individual and organizational perspectives* (pp. 221-234). Hillsdale, Nova Jersey: Lawrence Erlbaum Associates.
- Ree, M. J., & Carretta, T. R. (1994). The correlation of general cognitive ability and psychomotor tracking tests. *International Journal of Selection and Assessment, 2*, 209-216.
- Ree, M. J., & Earles, J. A. (1989). *The differential validity of a differential aptitude test*. (AFHRL-TR-89-59). Brooks Air Force Base, TX: Manpower and Personnel Division.
- Ree, M. J., & Earles, J. A. (1991a). Predicting training success: Not much more than g. *Personnel Psychology, 44*, 321-332.
- Ree, M. J., & Earles, J. A. (1991b). The stability of g across different methods of estimation. *Intelligence, 15*, 271-278.
- Ree, M. J., & Earles, J. A. (1992). Intelligence is the best predictor of job performance. *Current Directions in Psychological Science, 1*, 86-89.

- Ree, M. J., & Earles, J. A. (1994). The ubiquitous predictiveness of g. In C. Walker, M. Rumsey, & J. Harris (Eds.), *Personnel selection and classification* (pp. 127-135). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Ree, M. J., Earles, J. A., & Teachout, M. S. (1992). *General cognitive ability predicts job performance*. (AL-TR-1991-0100). Brooks Air Force Base, TX: Human Resources Directorate.
- Ree, M. J., Earles, J. A., & Teachout, M. S. (1994). Predicting job performance: Not much more than g. *Journal of Applied Psychology*, 79, 518-524.
- Robertson, I. T., & Smith, M. (1989). Personnel selection methods. In M. Smith, & I. T. Robertson (Eds.), *Advances in selection and assessment* (pp. 89-112). Chichester: Wiley.
- Sanders, A. F. (1983). Towards a model of stress and human performance. *Acta Psychologica*, 53, 61-97.
- Schmitt, N., Gooding, R. Z., Noe, R. A., & Kirsch, M. (1984). Meta-analysis of validity studies published between 1964 and 1982 and the investigation of study characteristics. *Personnel Psychology*, 37, 407-422.
- Schuler, R., & Jackson, S. (1997). Gestão de recursos humanos: Tomando posição para o século XXI. *Comportamento Organizacional e Gestão*, 3(2), 255-274.
- Schweizer, K. (1996). The speed-accuracy transition due to task complexity. *Intelligence*, 22, 115-128.
- Segall, D. O., & Moreno, K. E. (1999). Development of the computerized adaptive testing version of the Armed Services Vocational Aptitude Battery. In F. Drasgow & J. B. Olson-Buchanan (Eds.), *Innovations in Computerized Assessment* (pp. 35-65). New Jersey: Lawrence Erlbaum Associates, Publishers.
- Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, 171, 701-703.
- Simões, M. M. R. (1994). *Investigações no âmbito da aferição nacional do teste das Matrizes progressivas Coloridas de Raven* (Tese de Doutoramento). Coimbra: Universidade de Coimbra.
- Simon, H. A. (1978). Information-processing theory of human problem solving. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes: Human information processing* (vol.5). Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Skinner H. A., & Pakula A. (1986). Challenge of computers in psychological assessment. *Professional Psychology: Research and Practice*, 17(1), 44-50.
- Smith, M., & Robertson, I. T. (1989). Perspectives. In M. Smith, & I. T. Robertson (Eds.), *Advances in selection and assessment* (pp. 1-6). Chichester: Wiley.

- Smith, M., Farr, J. L., & Schuler, H. (1993). Individual and organizational perspectives on personnel procedures: Conclusions and horizons for future research. In H. Schuler, J. L. Farr & M. Smith (Eds.), *Personnel selection and assessment: Individual and organizational perspectives* (pp. 333-351). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Stankov, L. (2000). Complexity, metacognition, and fluid intelligence. *Intelligence*, 28(2), 121-143.
- Stankov, L., & Roberts, R. D. (1997). Mental speed is not the “basic” process of intelligence. *Personality and Individual Differences*, 22, 69-84.
- Sternberg, R. J. (1977). *Intelligence, information processing and analogical reasoning: The componential analysis of human abilities*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Sternberg, R. J. (1980). Sketch of a componential subtheory of human intelligence. *The Behavioral and Brain Sciences*, 3, 573-614.
- Sternberg, R. J. (Ed.). (1982). *Handbook of human intelligence*. Cambridge: Cambridge University Press.
- Sternberg, R. J., & Gardner, M. K. (1983). Unities in inductive reasoning. *Journal of Experimental Psychology: General*, 112, 80-116.
- Sternberg, R. J., & Salter, W. (1982). Conceptions of intelligence. In R. J. Sternberg (Ed.) *Handbook of human intelligence* (pp. 3-28). Cambridge: Cambridge University
- Sternberg, S. (1969). The discovery of processing stages: Extensions of Donder’s method. In W. G. Koster (Ed.), *Acta Psychologica*, 30: *Attention and Performance II*, 276-315.
- Süb, H-M., Oberauer, K., Wittmann, W. W., Wilhelm, O., & Schulze, R. (2002). Working-memory capacity explains reasoning ability – and a little bit more. *Intelligence*, 30, 261-288.
- Undheim, J. O., & Gustafsson, J-E. (1987). The hierarchical organization of cognitive abilities: Restoring general intelligence through the use of linear structural relations (LISREL). *Multivariate Behavioral Research*, 22, 149-171.
- Vernon, P. A. (1983). Speed of information processing and general intelligence. *Intelligence*, 7, 53-70.
- Vernon, P. A. (1986). The g-loading of intelligence tests and their relationship with reaction times: A comment on Ruchella et al. *Intelligence*, 10, 93-100.
- Vernon, P. A. (1989). The generality of g. *Personality and Individual Differences*, 10, 803-804.

- Vernon, P. A., & Kantor, L. (1986). Reaction time correlations with intelligence tests scores obtained under either timed or untimed conditions. *Intelligence, 10*, 315-330.
- Vernon, P. A., Nador, S., & Kantor, L. (1985a). Group differences in intelligence and speed of information-processing. *Intelligence, 9*, 137-148.
- Vernon, P. A., Nador, S., & Kantor, L. (1985b). Reaction-times and speed-of-processing: Their relationship to timed and untimed measures of intelligence. *Intelligence, 9*, 357-374.
- Weiss, V. (1986). From memory span and mental speed toward the quantum mechanics of intelligence. *Personality and Individual Differences, 7*, 737-749.
- Weiss, V. (1989). From short-term memory capacity toward the EEG resonance code. *Personality and Individual Differences, 10*, 501-508.
- Whimbley, A., Fischof, V., & Silikowitz, R. (1969). Memory span: A forgotten capacity. *Journal of Educational Psychology, 60*, 56-58.
- Wickelgren, W. A. (1977). Speed-accuracy trade-off and information-processing dynamics. *Acta Psychologica, 41*, 67-85.
- Wiesner, W. H., & Cronshaw, S. F. (1988). A meta-analytic investigation of the impact of interview format and degree of structure on the validity of the employment interview. *Journal of Occupational Psychology, 61*, 275-290.
- Wilhelm, O., & Schulze, R. (2002). The relation of speeded and unspeeded reasoning with mental speed. *Intelligence, 30(6)*, 537-554.
- Witkin, H. A., & Goodenough, D. R. (1981). *Cognitive styles: Essence and origins – Field dependence and independence*. New York: International Universities Press.
- Yee, P. L., Hunt, E., & Pellegrino, J. W. (1991). Coordinating cognitive information: Task effects and individual differences in integration information from several sources. *Cognitive Psychology, 23*, 615-680.